

Pushing frontiers - imaging for photon science

Edited by

Jiaguo Zhang, Iain Sedgwick and Cornelia B. Wunderer

Published in

Frontiers in Physics



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-5812-6
DOI 10.3389/978-2-8325-5812-6

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Pushing frontiers - imaging for photon science

Topic editors

Jianguo Zhang — Paul Scherrer Institut (PSI), Switzerland

Iain Sedgwick — Rutherford Appleton Laboratory, United Kingdom

Cornelia B. Wunderer — German Electron Synchrotron, Helmholtz Association of German Research Centres (HZ), Germany

Citation

Zhang, J., Sedgwick, I., Wunderer, C. B., eds. (2024). *Pushing frontiers - imaging for photon science*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-5812-6

Table of contents

- 05 **Editorial: Pushing frontiers—imaging for photon science**
Iain Sedgwick, Cornelia B. Wunderer and Jianguo Zhang
- 09 **VeryFastCCD: a high frame rate soft X-ray detector**
Azriel Goldschmidt, Carl Grace, John Joseph, Amanda Krieger, Craig Tindall and Peter Denes
- 18 **A low noise CMOS camera system for 2D resonant inelastic soft X-ray scattering**
Nord Andresen, Christos Bakalis, Peter Denes, Azriel Goldschmidt, Ian Johnson, John M. Joseph, Armin Karcher, Amanda Krieger and Craig Tindall
- 29 **Considerations about future hard x-ray area detectors**
Sol M. Gruner, Gabriella Carini and Antonino Miceli
- 34 **First operation of the JUNGFR AU detector in 16-memory cell mode at European XFEL**
Marcin Sikorski, Marco Ramilli, Raphael de Wijn, Viktoria Hinger, Aldo Mozzanica, Bernd Schmitt, Huijong Han, Richard Bean, Johan Bielecki, Gábor Bortel, Thomas Dietze, Gyula Faigel, Konstantin Kharitonov, Chan Kim, Jayanath C. P. Koliyadu, Faisal H. M. Koua, Romain Letrun, Luis M. Lopez, Nadja Reimers, Adam Round, Abhisakh Sarma, Tokushi Sato, Miklós Tegze and Monica Turcato
- 48 **X-ray detector requirements for laser–plasma accelerators**
Chris D. Armstrong, G. G. Scott, S. Richards, J. K. Patel, K. Fedorov, R. J. Gray, K. Welsby and P. P. Rajeev
- 57 **Detector developments for photon science at DESY**
Heinz Graafsma, Jonathan Correa, Sergei Fridman, Helmut Hirsemann, S. M. A. Hosseini-Saber, Alexandr Ignatenko, Alexander Klujev, Sabine Lange, Torsten Laurus, Alessandro Marras, David Pennicard, Seungyu Rah, Sandra Schneider, Ofir Shefer Shalev, Thorsten Stoye, Ulrich Trunk and Cornelia B. Wunderer
- 68 **X-ray beam test of fabricated photo-diodes for Pohang-Accelerator-Laboratory X-ray free-electron laser**
S. C. Lee, H. J. Hyun, J. M. Baek, S. M. Hwang, H. Jang, J. Y. Kim, S. H. Kim and H. Park
- 81 **Turning European XFEL raw data into user data**
Philipp Schmidt, Karim Ahmed, Cyril Danilevski, David Hammer, Robert Rosca, Thomas Kluyver, Thomas Michelat, Egor Sobolev, Luca Gelisio, Luis Maia, Maurizio Manetti, Janusz Malka, Krzysztof Wrona, Jolanta Sztuk-Dambietz, Vratko Rovensky, Marco Ramilli, Nuno Duarte, David Lomidze, Ibrahim Dourki, Hazem Yousef, Björn Senfftleben, Olivier Meyer, Monica Turcato, Steffen Hauf and Steve Aplin
- 93 **Novel sensor developments for photon science at the MPG semiconductor laboratory**
J. Ninkovic, A. Bähr, R. H. Richter and J. Treis
- 113 **Single-photon counting detectors for diffraction-limited light sources**
Erik Fröjd, Anna Bergamaschi and Bernd Schmitt

- 124 **Data reduction and processing for photon science detectors**
David Pennicard, Vahid Rahmani and Heinz Graafsma
- 140 **Probing the potential of CdZnTe for high-energy high-flux 2D X-ray detection using the XIDer incremental digital integrating readout**
Marin Collonge, Oriane Baussens, Paolo Busca, Pablo Fajardo, Peter Fischer, Thierry Martin, Michael Ritzert, Marie Ruat, David Schimansky and Morag Williams
- 152 **Neural network methods for radiation detectors and imaging**
S. Lin, S. Ning, H. Zhu, T. Zhou, C. L. Morris, S. Clayton, M. J. Cherukara, R. T. Chen and Zhehui Wang
- 179 **Operational experience with Adaptive Gain Integrating Pixel Detectors at European XFEL**
Jolanta Sztuk-Dambietz, Vratko Rovensky, Alexander Klujev, Torsten Laurus, Ulrich Trunk, Karim Ahmed, Olivier Meyer, Johannes Möller, Andrea Parenti, Natascha Raab, Roman Shayduk, Marcin Sikorski, Gabriele Ansaldi, Ulrike Bösenberg, Lopez M. Luis, Astrid Muenich, Thomas R. Preston, Philipp Schmidt, Stephan Stern, Richard Bean, Anders Madsen, Luca Gelisio, Steffen Hauf, Patrick Gessler, Krzysztof Wrona, Heinz Graafsma and Monica Turcato
- 201 **Data reduction activities at European XFEL: early results**
Egor Sobolev, Philipp Schmidt, Janusz Malka, David Hammer, Djelloul Boukhelef, Johannes Möller, Karim Ahmed, Richard Bean, Ivette Jazmin Bermúdez Macías, Johan Bielecki, Ulrike Bösenberg, Camille Carinan, Fabio Dall'Antonia, Sergey Esenov, Hans Fangohr, Danilo Enoque Ferreira de Lima, Luís Gonçalo Ferreira Maia, Hadi Firoozi, Gero Flucke, Patrick Gessler, Gabriele Giovanetti, Jayanath Koliyadu, Anders Madsen, Thomas Michelat, Michael Schuh, Marcin Sikorski, Alessandro Silenzi, Jolanta Sztuk-Dambietz, Monica Turcato, Oleksii Turkot, James Wrigley, Steve Aplin, Steffen Hauf, Krzysztof Wrona and Luca Gelisio
- 213 **Resolving soft X-ray photons with a high-rate hybrid pixel detector**
Viktoria Hinger, Rebecca Barten, Filippo Baruffaldi, Anna Bergamaschi, Giacomo Borghi, Maurizio Boscardin, Martin Brückner, Maria Carulla, Matteo Centis Vignali, Roberto Dinapoli, Simon Ebner, Francesco Ficorella, Erik Fröjd, Dominic Greiffenberg, Omar Hammad Ali, Shqipe Hasanaj, Julian Heymes, Thomas King, Pawel Kozłowski, Carlos Lopez-Cuenca, Davide Mezza, Aldo Mozzanica, Konstantinos Moustakas, Giovanni Paternoster, Kirsty A. Paton, Sabina Ronchin, Christian Ruder, Bernd Schmitt, Patrick Sieberer, Dhanya Thattil, Xiangyu Xie and Jiaguo Zhang
- 223 **Low gain avalanche diodes for photon science applications**
Matteo Centis Vignali and Giovanni Paternoster
- 244 **Perspective of perovskite-based X-ray hybrid pixel array detectors**
Michael Fiederle and Tilo Baumbach



OPEN ACCESS

EDITED AND REVIEWED BY

Cinzia Da Via,
The University of Manchester,
United Kingdom

*CORRESPONDENCE

Iain Sedgwick,
✉ iain.sedgwick@stfc.ac.uk
Cornelia B. Wunderer,
✉ cornelia.wunderer@desy.de
Jiaguo Zhang,
✉ jiaguo.zhang@psi.ch

RECEIVED 06 November 2024

ACCEPTED 07 November 2024

PUBLISHED 09 December 2024

CITATION

Sedgwick I, Wunderer CB and Zhang J (2024)
Editorial: Pushing frontiers—imaging for
photon science.
Front. Phys. 12:1523545.
doi: 10.3389/fphy.2024.1523545

COPYRIGHT

© 2024 Sedgwick, Wunderer and Zhang. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Editorial: Pushing frontiers—imaging for photon science

Iain Sedgwick^{1*}, Cornelia B. Wunderer^{2,3*} and Jiaguo Zhang^{4*}

¹Technology Department, UKRI-STFC Rutherford Appleton Laboratory, Didcot, United Kingdom,

²Photon Science, Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ³Center for
Free-Electron Laser Science CFEL, Hamburg, Germany, ⁴Center for Photon Science, Paul Scherrer
Institute, Villigen, Switzerland

KEYWORDS

X-ray imaging detectors, sensors, readout ASICs, data reduction, machine learning,
detector operation, synchrotrons, free-electron lasers

Editorial on the Research Topic

Pushing frontiers—imaging for photon science

1 Introduction

The dramatic improvement in photon sources such as Free Electron Lasers (FELs) and Diffraction-Limited Storage Rings (DSLRS) over the last two decades has significantly expanded the range of science that is possible at these facilities. In order to take full advantage, detectors with similarly advanced capabilities are needed. Developing such detectors, however, is extremely challenging; they typically take a decade to deploy and often require several iterations, necessitating considerable resources. Their integration in experiments is also not trivial. As a result, many experiments are still detector limited, as described by [Gruner et al.](#)

Therefore, we have solicited papers on progress in this field. This editorial includes an overview of key challenges reported by the authors and new technologies they described that help overcome them. Of course, many other developments are underway; here, we largely focus on those submitted by the authors.

2 Challenges for the future

The development of new detectors for photon science presents several challenges. The first is to meet the well-documented [1, 2] performance increase of new FELs and DSLRS. Second, photon science detectors must accommodate a wide range of experimental operating modes ([Gruner et al.](#), [Andresen et al.](#), [Armstrong et al.](#)). Even within a single facility, detectors supporting a variety of applications are required ([Graafsma et al.](#)). They are also frequently adapted for experiments for which they were not originally optimized, and are increasingly fitted with multiple sensor types to address the need for wider X-ray energy ranges. Designing with all these possible cases in mind is difficult and time-consuming—and can lead to compromise solutions not optimised for any one experiment.

The range of requirements is not entirely open-ended, as some specifications have practical limits, and advancements in radiation sources can even lead to detector

consolidation. For instance, at higher photon energies, there is a sensible limit to pixel size related to absorption length of secondary particles (Frojd et al.), and the increasing brilliance of DLSRs requires event or frame rates comparable to the CW-pulse repetition rates of future FELs, suggesting that similar detectors may be suitable for both (Graafsma et al.). Despite this convergence in frame rate requirements, however, specific needs do persist for much higher frame rates (GHz) for burst imaging (Gruner et al.) and much slower, low-noise imaging (sub-Hz) for RIXS (Andresen et al.). This, and the fact that high repetition-rate CW operation at some facilities remains a long-term project, suggests requirements will remain divergent at least in the near future.

The increase in performance of modern detectors also poses challenges for downstream systems. Multi-megapixel detectors with MHz frame rates generate vast quantities of data. This needs not just to be captured but calibrated and stored for many years. The calibration itself is a major task, since for some detectors the number of parameters can exceed 10^9 (Sztuk-Dambietz et al.). The associated difficulty can be strongly impacted by decisions taken at the detector development stage many years earlier (Pennicard et al.). The need for reproducibility of calibrations years later further adds to the complexity, since it must also be possible to apply more advanced calibrations as the understanding of an installed detector improves whilst still re-producing older results (Schmidt et al.).

A “gold standard” would be an integrating sensor with single-photon resolution which could convert to photon counts and compress to the Poisson limit for ultimate data reduction with zero science loss (Frojd et al., Pennicard et al.).

When addressing these challenges, it is crucial to identify the primary bottleneck in the system, which could be anywhere from sensor to data transfer. If this bottleneck cannot be mitigated, optimizing other parts of the system for higher performance may be inefficient or unnecessary.

3 New technologies

The need for detectors to span an energy regime from 10^1 to 10^5 eV pushes both hard and soft X-ray sensor developments. For hard X-rays, in addition to GaAs, Ge, and CdTe, research into the manufacture and use of CdZnTe has led to improved leakage current and stability under high flux conditions (Collonge et al.), making its use in detectors more viable, but much work remains to be done. Other high-Z options, such as Perovskites, are also being investigated, but are at an earlier stage (Fiederle et al.).

Similarly, in the soft X-ray regime, several useful technologies exist. For monolithic systems such as the pnCCD (Ninkovic et al.), backside-illuminated CCDs (Goldschmidt et al.) and CMOS imagers (Andresen et al.), entrance window processing technologies have been developed that make these devices sensitive down to the double digit eV range with good efficiency and reasonable signal-to-noise ratio (SNR) thanks to their relatively low noise. High-quality entrance windows are key for any soft X-ray detector (Lee et al.). For hybrid detector systems, typically with higher noise due to the bump-bonding process, segmented LGAD sensors (Vignali et al., Sikorski et al.) and DEPFETs (Ninkovic et al.) provide good sensor options for the improvement of SNR in a different manner.

Technical advances in the commercial semiconductor market can also help improve performance. For example, CMOS technology nodes of 180 nm and below are routinely used in photon science ASICs. Their high transistor density allows much functionality to be implemented on-chip. This has enabled several developments, particularly in the high flux area. XIDER, CORDIA, and Matterhorn all use different methods to overcome challenges associated with the combined need for high frame or count rates and high dynamic range (Collonge et al., Graafsma et al., Frojd et al.).

Cutting-edge commercial designs and even other scientific fields use much smaller nodes [3] than the 65 nm and 110 nm used here. However, commercial effort focuses on reducing the cost per transistor whereas for large area detector applications, the cost per area is most important, and this tends to increase as the node shrinks [4]. This may eventually limit what node is used for large-area applications. In addition, while smaller nodes are superior for digital circuitry, for analog circuits larger nodes have advantages as well. Older nodes may continue to be employed, or the use of chiplets to best match cost and performance may become more common.

New CMOS functionalities beyond node size also allow improved performance. However, these are sometimes not available for small-batch developments. A prominent example is 3-D integration, which has been commercially common for many years but has only been sporadically employed for photon science. Whether such technologies will permit higher-performing detectors in the future will likely be a question of access.

Advances can also be made when commercial detector systems in other fields turn out to be suitable for photon science use. In some cases, in particular in terms of cost and time, these constitute a viable or even better alternative to custom-developed systems.

Handling the vast amount of data produced by modern detectors is a particularly critical area, discussed in greater detail in the next section.

4 Data reduction and processing

Data reduction and processing is a vast field which, even 10–15 years ago, was—at least in photon science—firmly linked to “data analysis” which occurred long after data was first recorded. Since then, source and detector advancements have resulted in a paradigm change. Today, data reduction during or shortly after detection is unavoidable to keep recorded data volumes manageable (Sobolev et al., Pennicard et al.). Reducing stored data volumes while maintaining science content may be ‘the’ key to future advances in photon science experiments. This is not merely a technical problem, but also has legal and social ramifications (Sobolev et al.).

Many in the photon science user community are reluctant to reduce raw data before detailed inspection, and data reduction is complicated by the vast range of experiments (Sobolev et al.). This contrasts with fields such as particle physics, where in-detector data reduction has been standard for decades (Pennicard et al.). For photon science, technique-specific data reduction is needed, and it is important to keep both reproducibility and improved data processing in mind—i.e., it must remain possible to recreate results from old processing tools even as improved tools allow better results or systems are updated (Schmidt et al.).

Data reduction can be carried out in the frontend ASIC itself (see ‘on-chip reduction’ examples in Pennicard et al.). However, most of the schemes submitted to this Research Topic take place in the processing FPGA or further downstream. Promising examples today often involve machine learning (ML) methods (Lin et al.). However, these are sometimes not transparent—making them difficult to understand and trust (Pennicard et al.). Partly as a result, the majority of processing is still performed without machine learning (Sztuk-Dambietz et al.), but it is clear that ML will become increasingly common.

From a detector developers’ viewpoint, the key point to realize is that the complexity of data processing and calibration depends largely on the detector design (Pennicard et al.). ASIC design decisions in particular, often among the first taken in the system design, can have a significant impact on the complexity of later data reduction processes. With ever-increasing raw data volumes, a system that delivers the most science content per recorded Gigabyte in a variety of scientific contexts is likely to become the most sought-after. Furthermore, simplifying system integration is also critical, and this is treated in the next section.

5 Operational complexities

Running full-scale imaging systems at photon science facilities constitutes a challenge in itself. Partly, this is inherent in the diverse user needs and facility parameters, but it also originates in the imaging systems’ design. Prioritizing ease of operation, maintenance, calibration, and data processing during the design phase will significantly enhance user interest in the final system.

Anticipating and simplifying both assembly and disassembly of the full-scale system is crucial, as the associated risk and time in turn impacts decisions on replacement, refurbishment, or upgrades (Sztuk-Dambietz et al.). A clear and fast route to (re-)calibration is also critical in simplifying deployment and increasing adoption.

Even when designing individual components, one should keep in mind the envisioned system scale and strive for simplicity. Some detector systems have more than 10^9 calibration parameters, resulting in obvious complications in terms of calibration, parameter storage, and data correction. Multi-gain systems are a very good way to address the need for high dynamic range. However, the gain transition regions add significant complexity to calibration (Sikorski et al., Sztuk-Dambietz et al.). A goal of next-generation detectors should be to dramatically reduce these complexities to enable simplified operation.

To a facility, complexity is not only related to operating one imaging detector, but also to the range of systems in use. The more common components, the easier to operate the entirety of systems at a facility. Ideally, this means largely identical systems with, e.g., different geometric arrangements (AGIPD 1M vs. 4M (Graafsma et al.)) or sensor type (hybrids mated to high-Z, Si, or LGADs for instance (Graafsma et al., Hinger et al., Vignali et al., Collonge et al.)). Even “just” shared control, DAQ, or cooling systems already reduce operational complexity for the facility.

It is also important to note that the facility will choose the pragmatic route to a functional user experiment—this might mean running a well-integrated and stable detector outside its usual envelope (Sikorski et al.), or using a stable or already-installed

imaging system over a fledgling, ultra-fragile one, trading stability for maximized performance. For detector development, this means that the simpler to use and optimize the system, the more likely it will actually get used at its full potential.

The bottom line to keep in mind as an imaging system developer is: “Data quality is the paramount measure of detector performance” (Sztuk-Dambietz et al.)—and too-complex calibration or module exchange can compromise this just as much as a noisy frontend.

6 Outlook

The development and optimization of imaging detectors for photon science is a wide and vibrant field, and progress is being made on many fronts—including many outside the scope of this Research Topic. Exciting challenges remain, and new ones develop as experiments as well as sources advance. The community as a whole can look forward to the future, and the many exciting developments yet to come.

Author contributions

IS: Writing—original draft, Writing—review and editing.
CW: Writing—original draft, Writing—review and editing. JZ: Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to thank all the authors and reviewers involved in this Research Topic for their contributions and hard work. They would also like to acknowledge the help and advice of all their colleagues, including several fruitful discussions on specific topics. CW acknowledges support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Huang Z. Brightness and coherence of synchrotron radiation and fels moycb101. *Proc IPAC (Shanghai, China)* (2013).
2. Pellegrini C. The history of x-ray free-electron lasers. *Eur Phys J H* (2012) 37:659–708. doi:10.1140/epjh/e2012-20064-5
3. Traversi G, Gaioni L, Ratti L, Re V, Riceputi E. Characterization of a 28 nm cmos technology for analog applications in high energy physics. *IEEE Trans Nucl Sci* (2024) 71:932–40. doi:10.1109/TNS.2024.3382348
4. Flamm K. *Measuring moore's law: evidence from price, cost, and quality indexes*. University of Chicago Press (2019). p. 403–70. doi:10.7208/chicago/9780226728209.001.0001



OPEN ACCESS

EDITED BY

Jianguo Zhang,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY

Gabriele Giacomini,
Brookhaven National Laboratory (DOE),
United States
Konstantin Stefanov,
The Open University, United Kingdom
Robert Andritschke,
Max Planck Institute for Extraterrestrial
Physics, Germany

*CORRESPONDENCE

Peter Denes,
✉ pdenes@lbl.gov

RECEIVED 29 August 2023

ACCEPTED 16 October 2023

PUBLISHED 26 October 2023

CITATION

Goldschmidt A, Grace C, Joseph J,
Krieger A, Tindall C and Denes P (2023),
VeryFastCCD: a high frame rate soft X-
ray detector.
Front. Phys. 11:1285350.
doi: 10.3389/fphy.2023.1285350

COPYRIGHT

© 2023 Goldschmidt, Grace, Joseph,
Krieger, Tindall and Denes. This is an
open-access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

VeryFastCCD: a high frame rate soft X-ray detector

Azriel Goldschmidt¹, Carl Grace¹, John Joseph¹, Amanda Krieger¹,
Craig Tindall¹ and Peter Denes^{2*}

¹Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, ²Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

Direct X-ray detection in silicon has been transformative for scattering experiments in biology and materials science. While bump-bonded hybrid pixels have been attractive for hard X-ray detection, the challenge for single photon soft X-ray detection is sufficiently low noise. CCD structures on thick, high-resistivity silicon have been successfully used as sensors over the range of soft to hard X-rays at storage rings and FELs. The VeryFastCCD is a high frame rate, column-parallel CCD sensor with 48 μm pixels. Combined with 256-channel custom readout ASICs, frame rates of 5–10 kHz have been achieved with readout noise as low as 20 e and full-well capacity $>4 \times 10^5$ e/pixel. Thin (10 nm and 100 nm) entrance window contact processes have also been developed which provide $>85\%$ quantum efficiency for 285 eV X-rays. Systems are currently being developed for several beamlines at the upgraded Advanced Light Source.

KEYWORDS

soft X-ray detector, direct-detection CCD, readout ASIC, thin contacts, sensors

1 Introduction

The charge-coupled device [1] (CCD) has been a backbone of scientific imaging for 50 years. The CCD, together with a fiber-coupled phosphor, forms a detector that has been used for decades in X-ray and electron microscopies: incident radiation creates optical photons by ionization, which are transported by a fiber optic light guide onto a CCD. Removing the phosphor and directly detecting X-rays in CCDs has proven over the last ~15 years to be attractive for storage ring and free electron laser X-ray light sources, since compared to indirect detection in a phosphor, direct detection provides much better spatial resolution (scintillation photons are emitted in all angles) along with higher detection efficiency and signal/noise (since it takes 3.6 eV to create an e/h pair in silicon vs. 10s of eV to create a scintillation photon).

Key considerations for direct X-ray detection in silicon are the thickness of the sensitive volume and the amount of dead material in front of the sensitive volume. Since the X-ray absorption length λ , in silicon is a strong function of energy— $\lambda = 40$ nm at 100 eV, $\lambda = 130$ μm at 10 keV and $\lambda = 23$ mm at 100 keV—high efficiency X-ray detection is practical up to ~10 keV (100s of μm thick sensitive volume). To maintain high spatial resolution and charge collection efficiency, a fully-depleted sensitive volume is desired (so that the ionization charge is collected by drift to a collection diode rather than by diffusion into all angles). This requires that the sensitive volume is of sufficiently high resistivity that it can be depleted by the application of a substrate voltage below the breakdown voltage [2]. To apply the substrate voltage, an entrance contact is required, and for operation to the lowest energies, the

entrance contact must be thin enough that the incident X-rays are not all absorbed in the contact.

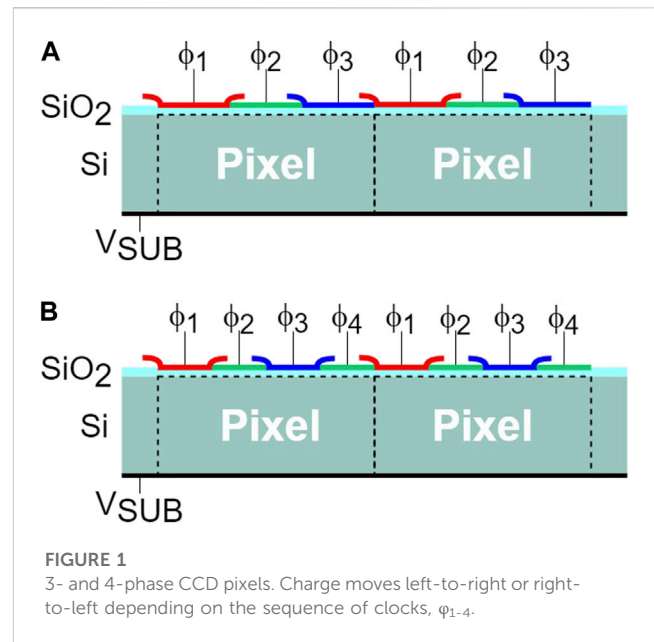
Due to their lower noise compared to hybrid pixel detectors, direct detection CCDs are particularly attractive for soft X-rays ($E < 2$ keV). CCDs are comparatively slow, however, since the CCD structure is based on shifting charge from pixel-to-pixel—rather than reading each pixel directly. In the simplest case, all pixels are individually shifted and digitized through a single readout port. It thus takes time $t_{FRAME} = \frac{N}{f}$ to read out N pixels at a readout frequency f . As source brightness increases, less time is required to collect the same number of photons. Similarly, t_{FRAME} sets the temporal resolution. For these reasons, reducing t_{FRAME} is required as sources improve, and it can be reduced by adding more readout ports, so that $t_{FRAME} = \frac{N}{fm}$ for m ports. Direct X-ray detection CCDs, on thick high-resistivity substrates, with large numbers of readout ports have been developed: such as the pnCCD [3], the FastCCD [4], and the MPCCD [5].

Our prior work is the FastCCD which is a >120 Hz frame rate, megapixel CCD optimized for soft X-rays that has one readout port for every 10 columns. Below we describe the VeryFastCCD, which is the natural follow-on to the FastCCD as a general-purpose soft X-ray scattering detector—increasing speed by having one readout port for every column. Compared to the FastCCD, the VeryFastCCD is much faster, with a 5 kHz frame rate for a 512×512 pixel sensor. As described below, through the use of a multi-gain gated integrator, it is possible to tune the full-scale signal vs. noise both by gain selection and variation of integrator timing.

2 CCD design

The metal-oxide-semiconductor (MOS) CCD structure consists of polysilicon gates on a gate oxide above a silicon substrate. The most common CCD pixel, as illustrated in Figure 1A, has three separate gates (ϕ_1, ϕ_2, ϕ_3). By manipulating the voltages on the gates, charge can be stored (integrated) or shifted from pixel-to-pixel. The simplest CCD detector, Figure 2A has a 2D array of imaging pixels and a 1D array of pixels used as an output shift register. In operation (1) an image is exposed, (2) the first row is shifted into the output shift register via the Array Transfer Gate (ATG), (3) the output shift register pixels individually shifted out and captured, (4) and the next row is then transferred to the output shift register. In most CCDs, an MOS source-follower is used to drive the signal off the sensor. An Output Transfer Gate (TG), Figure 2B, collects the charge from the output shift register, and pushes it over a barrier potential (OG) onto the gate of the source-follower. The voltage on the gate is then $V = \frac{Q}{C_G}$, where C_G is the total capacitance of the source-follower gate. A RESET transistor is used to remove charge from the previous sample before the next sample is presented. The digitized pixel value is thus the *voltage* difference between the RESET and SIGNAL levels.

To increase the frame rate of the CCD readout, additional parallelism is used. The simplest is reading out from both sides of the CCD: The 4-port CCD (Figure 2C) is a trivial modification of the single port CCD (Figure 2A). Incorporating more output ports becomes more challenging, since the output shift registers must be modified in order to provide the space required for the output stage. Numerous CCDs with more than 4 readout ports (Figure 2D) have been produced: for example, the FastCCD has one port for every



10 columns. The VeryFastCCD is a fully-column-parallel CCD (Figure 2E)—which means that every CCD column has a readout port. Because there is not enough space for an output stage (Figure 2B) for each column, the VeryFastCCD directly outputs a *charge* rather than a *voltage*.

The FastCCD and VeryFastCCD are fabricated in the Dalsa 2.5 μm , triple poly, triple metal process¹ on very high resistivity n-type silicon with buried p channels, using a method developed at LBNL(2). For CCDs operating at high speed, the long polysilicon gates must be “metal strapped”—that is covered with metal lines, which are $\sim 1/1000$ the resistivity of the polysilicon, and connected as frequently as possible to the polysilicon. For the FastCCD, we chose to use a single metal layer, and designed a 30 μm , 3-phase (Figure 1A) pixel. For the VeryFastCCD, we chose to use 3 metal layers. Since here a charge, rather than voltage, output, is used, parasitic charge injection (through substrate or other capacitive coupling to the output) can present large transients to any downstream amplifier, and potentially saturate the amplifier output. For this reason, we designed a 4-phase (Figure 1B) pixel in order to minimize parasitic charge injection from the CCD clocks since 4-phase clocking provides first order cancellation at each clock transition ($\phi_3 = -\phi_1$ and $\phi_4 = -\phi_2$). The larger the pixel, the larger the working distance between the sample and detector can be. The power needed to clock the CCD, though, is proportional to area. To balance the two the pixel size was set at 48 μm .

Since the CCDs are generally thick (100s of μm) to have good X-ray absorption, they are “self-shielding” against radiation damage when back-illuminated. Figure 3 shows the cumulative photon flux per pixel to accumulate 1 MRad in the gate oxide for a 200 μm thick sensor (the thickness of the VeryFastCCD). Measurements at DORIS

1 <https://www.teledynedalsa.com/en/products/foundry/ccd/>

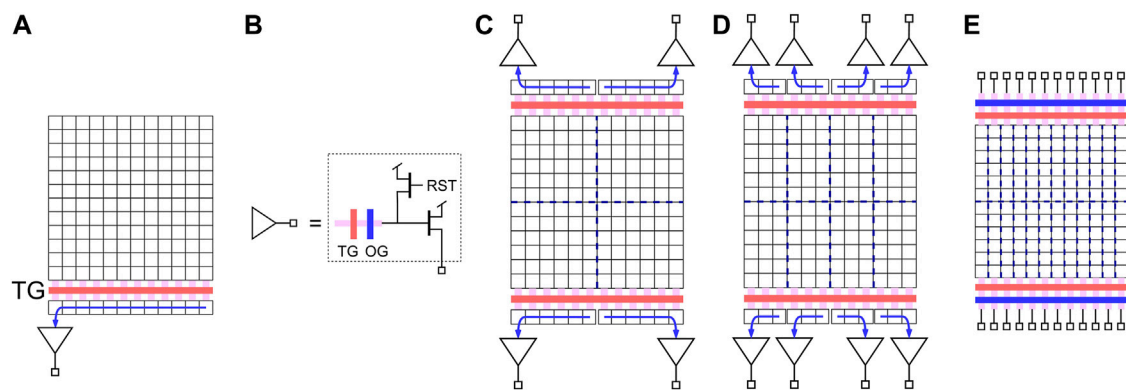


FIGURE 2
CCD structures with increasing parallelism.

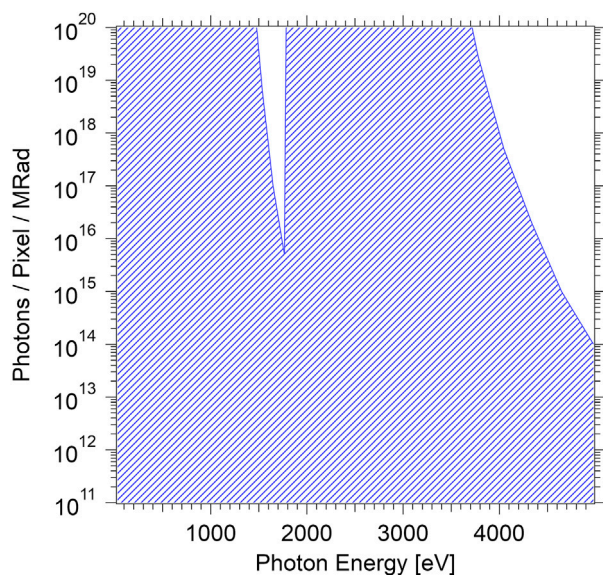


FIGURE 3
Calculated number of photons per VeryFastCCD pixel required to deposit 1 MRad in the CCD gate oxide as a function of energy.

with hard X-rays on the FastCCD indicate that 1 MRad will induce sufficient leakage current to limit performance with 10 ms exposure times. This suggests no radiation damage for soft X-ray storage ring use, nor at FELs (under normal circumstances).

The VeryFastCCD uses all of the same sensor fabrication and processing methods as the FastCCD. A fabricated wafer is shown in Figure 4. Both 256 × 256 pixel and 512 × 512 devices are present. The 512 devices are arranged in such a way that four of them can be diced to create a 1024 × 1024 sensor with 90% fill factor.

2.1 Thin entrance windows

High efficiency detection of soft X-rays requires a back side illuminated CCD with minimal dead material between the incident

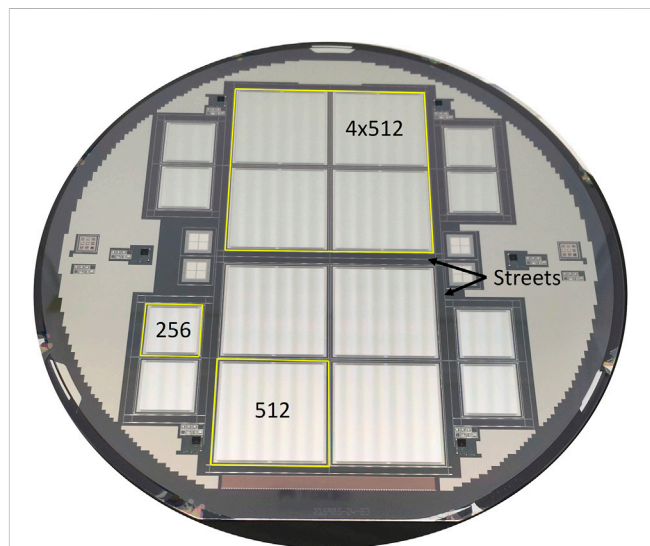
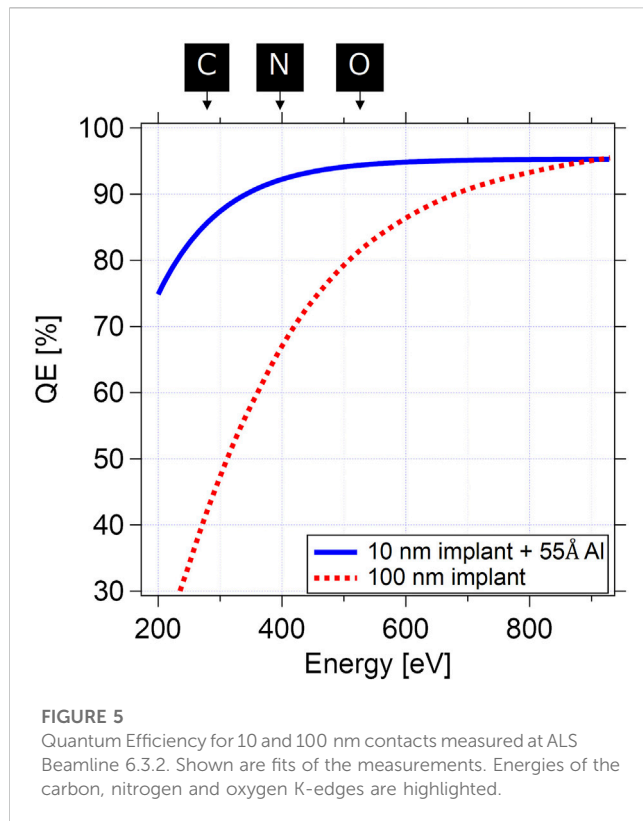


FIGURE 4
A 150 mm VeryFastCCD production wafer, containing 8 256 × 256 and 8 512 × 512 CCDs. Four 512 × 512 CCDs can be diced out as a single 1024 × 1024 CCD.

X-ray and the active silicon. To deplete the thick active volume, a voltage must be applied to the back (X-ray entrance) side of the CCD (V_{SUB} in Figure 1). A very thin contact layer is thus required. Soft X-ray detection also requires very low noise, hence the contact must ensure minimal leakage current. While Schottky barrier contacts exhibit a very thin dead layer, they result in leakage currents that are high for soft X-ray applications. An ohmic contact is thus preferred and used here.

The technique initially developed at LBNL for fabricating thick, high resistivity CCDs is to use doped polysilicon to form the thin X-ray entrance contact [6]. However, polysilicon deposition requires temperatures that are too high to be compatible with aluminum metallization. When this technique is used, wafers are (a) implanted and gates formed at the factory (b) thinned and then the doped polysilicon contact is deposited at LBNL (c) final metalization is performed at LBNL. As the LBNL metal process has only one layer, and larger-sized vias than the triple-metal Dalsa process, having a

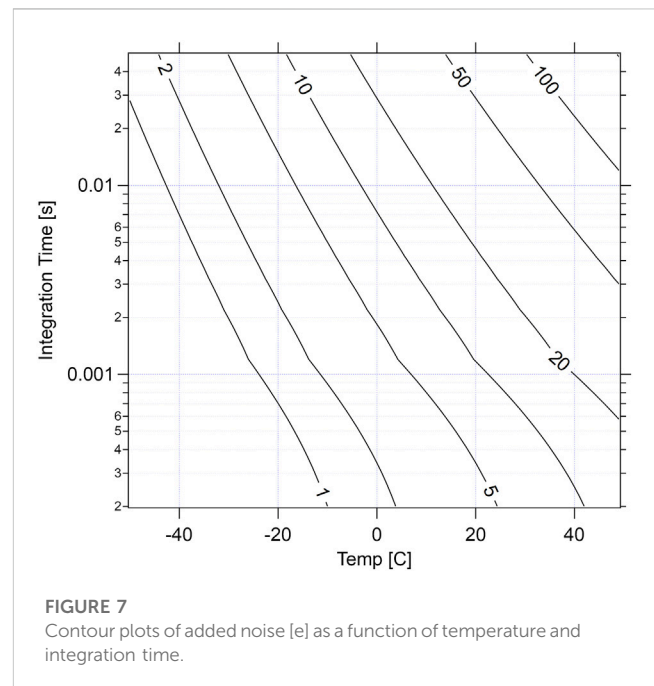
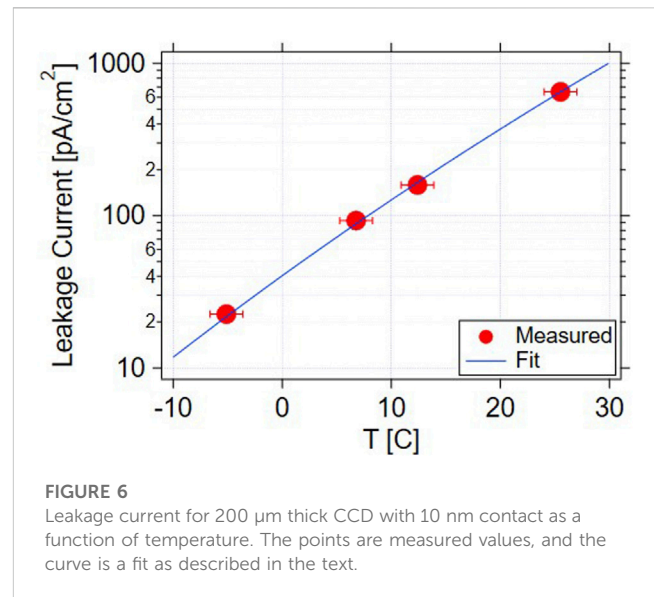


low temperature process that can be performed on fully metallized wafers (a) simplifies the CCD fabrication process since all of the steps except the thin contact fabrication can first be performed at the foundry (b) enables the use of more metal layers, which improves the high-speed CCD clock distribution.

The thin X-ray entrance contact is then added to the device at LBNL as the last processing step. To this end we have developed two low temperature thin contact processes: (1) implantation followed by low temperature annealing (~ 100 nm thick), and (2) molecular beam epitaxy (MBE) (~ 10 nm thick). The MBE contact has the added benefit of being even thinner than the doped polysilicon contact (~ 30 nm thick).

The ion implanted contact generally finds use for applications in which the energies of the incident photons are 500 eV or more. This contact has the advantage of being relatively simple to fabricate. The MBE contact is used to extend the reach to lower energies.

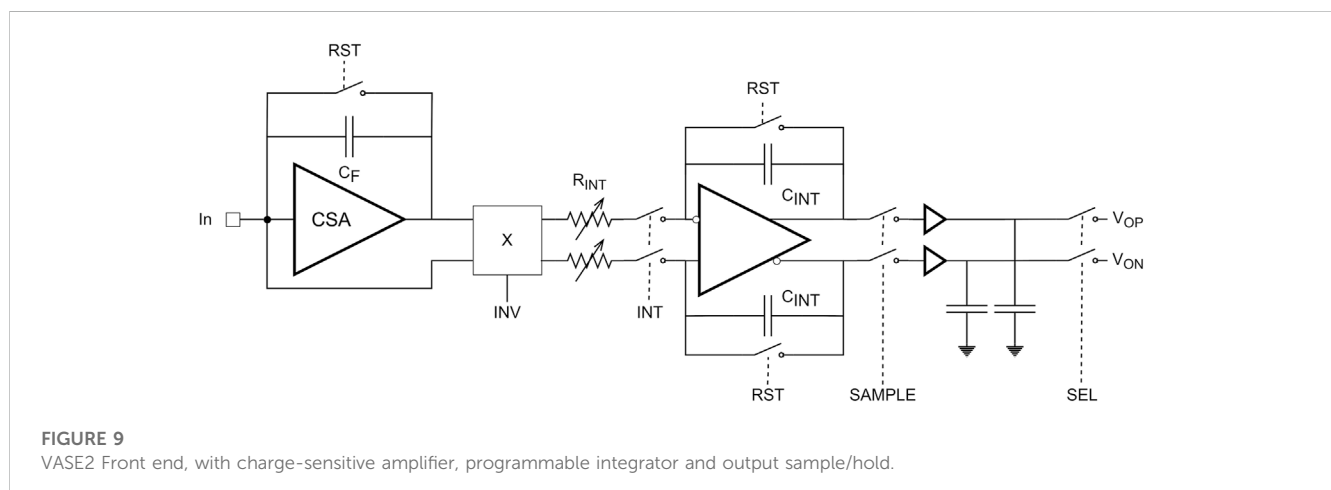
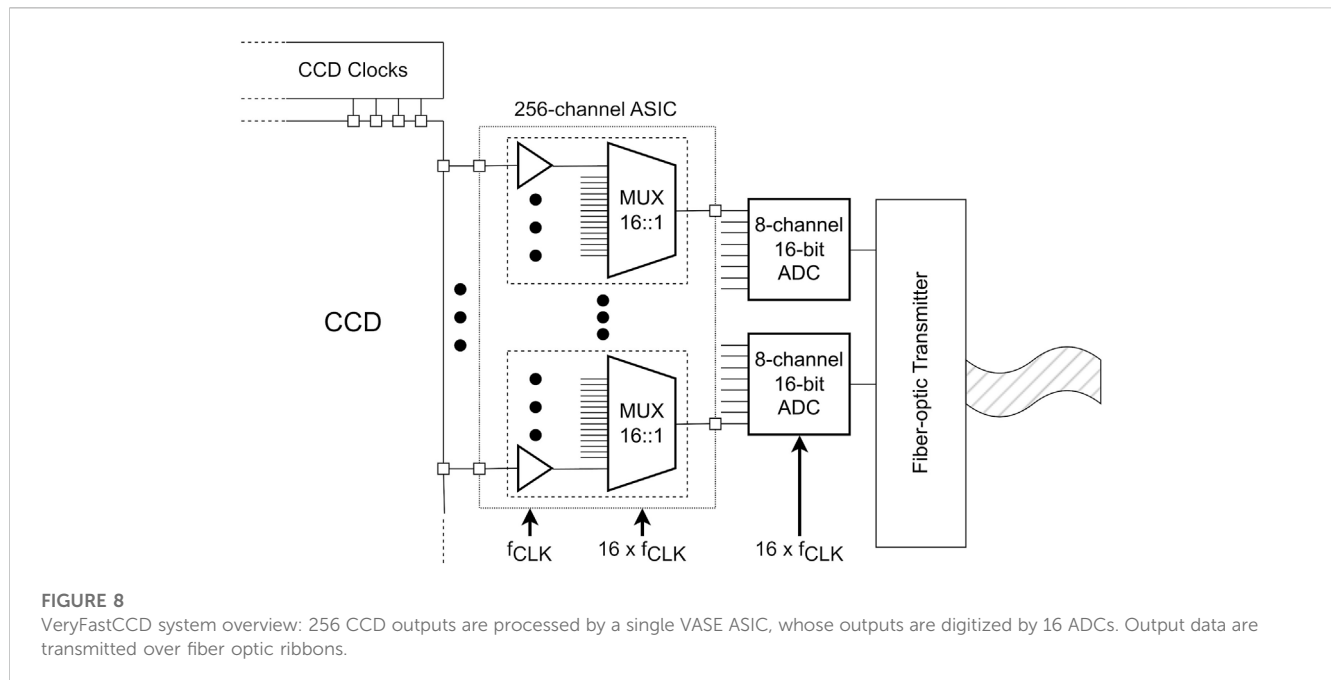
We have evaluated the quantum efficiency (QE) of both of these types of contacts on numerous samples (CCDs and diodes) using the metrology beamline² (6.3.2) at the Advanced Light Source [7]. Figure 5 shows the averages of QE measurements for the two types of contacts. Figure 6 shows the leakage current for the 10 nm contact on a 256×256 VFCCD as a function of temperature, T . The temperature dependence of the leakage current is modeled [8, 9] as $I_L \propto T^n e^{-E/2kT}$, usually with $n = 2$, and effective energy $E = E_g + 2\Delta$ ($E_g = E_0 - \alpha T$ is the temperature dependent bandgap of Silicon, and Δ arises from the presence of a charge-generating trap). The curve in Figure 6 is a fit to the model above, with $\Delta = 0.11$ eV. The low value



of Δ , and of the overall leakage current, indicates that the MBE contact does not add significant leakage current.

While the VeryFastCCD has a high frame rate, different experimental conditions necessitate different exposure times. A key system design parameter is the camera operating temperature, and this is determined by the noise added by leakage current: that added noise is proportional to $\sqrt{I_L T_{EXP}}$, where I_L is the leakage current, and T_{EXP} is the total exposure time (which for a given row of pixels is the total time that the row can collect leakage current: the image exposure time plus the time needed to shift out the charge from that row). Figure 7 shows noise contours of added noise due to the leakage current shown in Figure 6 (in electrons, where 3.6 eV of absorbed energy is required to create one electron/hole pair in silicon). For FEL

² <https://cxro.lbl.gov/als632/>



applications, operation near room temperature is possible. For storage ring applications, the operating temperature will depend on the longest exposure time required, and the desired total noise, but operation at -20°C should cover most use cases.

3 Data acquisition

Data acquisition is illustrated in Figure 8. The CCD signal is read out on two sides (Figure 2E), and the other two sides are used to provide clocks. A 256-channel integrated circuit VASE2 has been developed to read out the charge from the VeryFastCCD. (A prior version VASE1 is described in [10] and was subsequently simplified for VASE2 by removing a $\Sigma\Delta$ gain stage.). One VASE2 handles 256 CCD channels. Each CCD output is wire-bonded to a VASE2 input. Internally, VASE2 consists of 16 modules: each module has 16 front-end circuits and a 16:1 analog multiplexer. Each 8 of the

VASE2 outputs are input to a commercial 8-channel, 16-bit ADC (TI ADS52J65). High-speed outputs from the ADCs are converted to optical signals, driven off the camera via fiber-optic ribbons.

For a 256×256 pixel camera, 2 VASE2 and 4 ADCs are used; for a 512×512 camera, 4 VASE2 and 8 ADCs are used, and this doubles again for a 1024×1024 camera.

A single VASE2 front end is shown in Figure 9. Similar to the FastCCD [11], the CCD signal is voltage-integrated, with a programmable gain and analog inversion to allow correlated double sampling (CDS). Unlike conventional CCDs, the VeryFastCCD signal is charge: VASE2 thus incorporates a charge-sensitive amplifier (CSA) with reset. As shown in Figure 10A, the RST switches are closed, removing charge on capacitors C_F and C_{INT} . The RST switches are then opened and switch INT is closed for time T_{INT} , and the reset level (V_{RST}) is integrated on capacitors C_{INT} . Switch INT is then opened, and CCD charge is presented to the CSA. The CSA output is $V_{SIG} = \frac{Q_{CCD}}{C_F} + V_{RST}$. The INV signal inverts the CSA output, so that

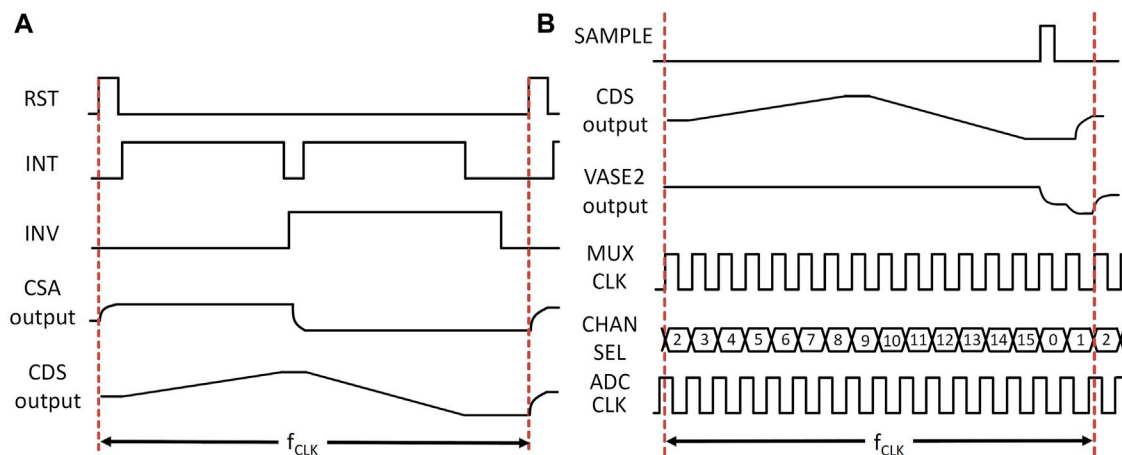


FIGURE 10
(A) VASE2 front-end timing (B) back-end timing.

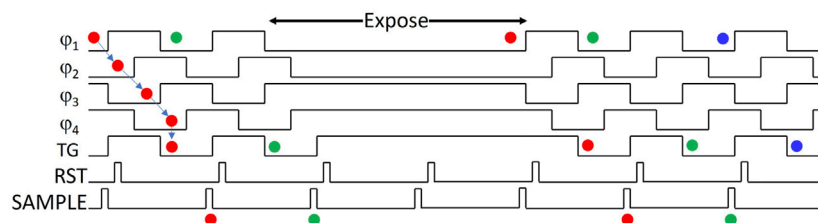


FIGURE 11
CCD and VASE front end timing: VASE clocks run continuously whereas CCD clocks are paused during signal integration. The colored circles represent pixel signals—as they are shifted by the CCD clocks and finally captured by the VASE SAMPLE signal.

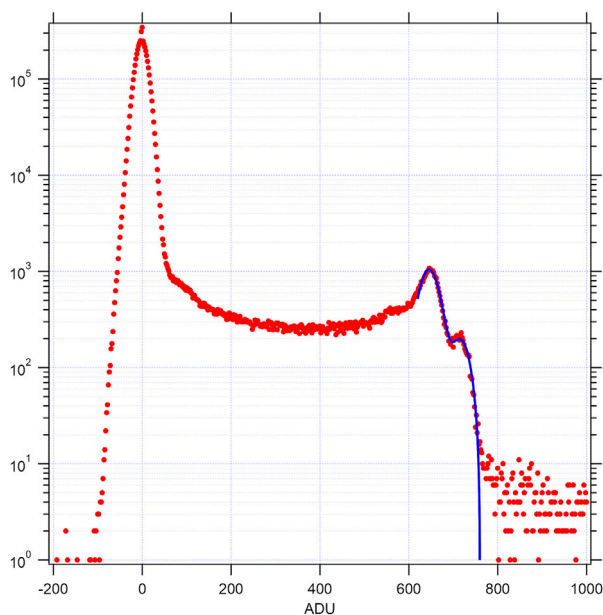


FIGURE 12
Spectrum (red) from all pixels illuminated by a ^{55}Fe source. The blue curve shows a fit to the combined K_α/K_β lines.

after the second integration the integrator output is $V_{INT} = \frac{1}{R_{INT}C_{INT}} \int_0^{T_{INT}} dt \left(\frac{Q_{CCD}}{C_F} + V_{RST} - V_{RST} \right)$. CDS thus removes the kTC noise associated with the CSA reset. Each front end contains 6 different values of R_{INT} , allowing a factor of approximately 5 in gain adjustment. Two versions of VASE2 were fabricated: a high gain (HG) version with $C_F = 14.5$ fF (including parasitic capacitance) and a low gain (LG) version with $C_F = 44$ fF. Further, as discussed below, additional gain adjustment is possible by changing T_{INT} .

The CCD row clocks and the VASE2 front end clocks (shown in Figure 10A) both operate at frequency f_{CLK} . The VASE2 back end and ADC operates at $16 \times f_{CLK}$. As illustrated in Figure 10B, each front end stores the value of V_{INT} for CCD row N on a sample-and-hold circuit when the SAMPLE command is issued. While the front end proceeds to acquire the signal from CCD row $N+1$, the 16 stored values for row N are multiplexed and digitized.

All of the measurements described below were performed with 256×256 CCD cameras. All of the CCDs were $200 \mu\text{m}$ thick, and operated at reverse bias of up to 80V. Several versions, equipped with HG and LG VASE2 chips, along with both 10 nm and 100 nm contacts have been tested. Cameras were also tested at the COSMIC beamline³ at ALS [12], and the SXR beamline at LCLS I. The cameras

³ <https://als.lbl.gov/beamlines/7-0-1-2/>

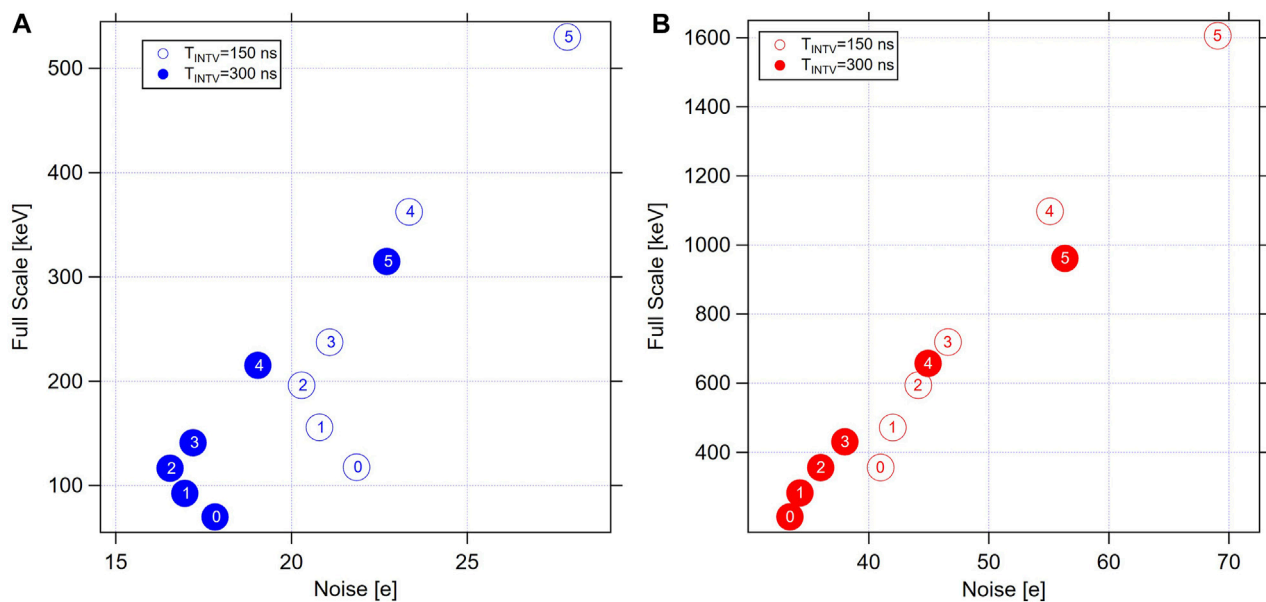


FIGURE 13

Maximum photon energy per pixel that can be recorded vs. Equivalent Noise Charge (ENC) for (A) VASE2 HG (B) VASE2 LG. The numbers in the circles are the gain setting, with 0 being the highest gain and 5 being the lowest. For both (A,B) values are shown for the nominal $T_{INT} = 300$ ns VASE integration time (solid circles) and for $T_{INT} = 140$ ns (open circles).

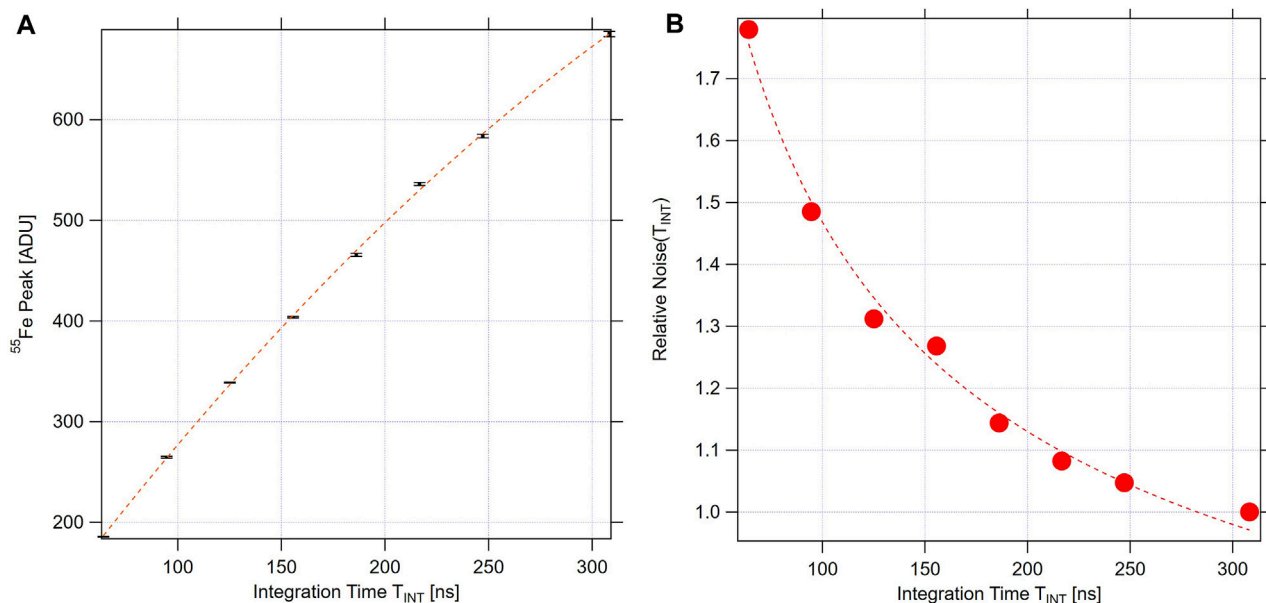


FIGURE 14

(A) ^{55}Fe peak position as a function of integration time. The curve is a quadratic fit. (B) Noise relative to nominal integration time of 305 ns. The curve is a fit to $1/\sqrt{T_{INT}}$.

were operated at $f_{CLK} = 1.28$ MHz, corresponding to up to 10,000 frames/s (equivalently 5,000 frames/s for a 512×512 device).

A timing diagram is shown in Figure 11. Clocks to the VASE2s and ADCs are continuous. CCD clocks also operate at $f_{CLK} = 1.28$ MHz. For storage ring applications, with signal integration times $\gg 1 \mu\text{s}$, the CCD clocks are paused for an integral number of

1.28 MHz cycles. An FPGA-based data acquisition constructs images based on those ADC samples that are flagged as corresponding to charge readout. The FPGA operates at $256 \times f_{CLK}$, and programs the position and duration of each clock (CCD clocks, VASE clocks) using 8-bit registers to store values for when they should turn on and off.

4 Results

The digitized signal value is $G_{ADC}V_{INT} = G_{ADC}G_{INT}\frac{Q_{CCD}}{C_F}$ where G_{ADC} is the number of ADU per volt for the ADC, $G_{INT} = \frac{T_{INT}}{R_{INT}C_{INT}}$ is the integrator gain, giving a nominal conversion gain $G_{ADC}G_{INT}\frac{1}{C_F}$ (ADU/e). The actual conversion gain for the HG and LG versions of VASE2 at different gain settings was determined by illuminating the CCD with a ^{55}Fe source and constructing spectra of single pixels. Figure 12 shows such a spectrum: the conversion gain is determined from a combined Gaussian fit to the $K_{\alpha}K_{\beta}$ lines.

Figure 13A shows the maximum photon energy deposited in a pixel that can be recorded (before the ADC saturates) vs. the ENC for different gain settings of the high gain version of VASE2. Figure 13B shows the values for the low gain version of VASE2. On a given gain range, the full-scale divided by the noise is 10–12 bits (larger as the gain is reduced).

The digitized noise is $\hat{\sigma}_n[ADU] = G_{ADC}(G_{INT}\sigma_{FE} \oplus \sigma_{BE})$ where $\sigma_{FE}[V]$ is the front-end noise (from the CSA and any added noise due to leakage current), and $\sigma_{BE}[V]$ is the back-end noise (from the everything after the CSA and the ADC). The ADC noise is 4.7 ADU^4 . Keeping RST continuously asserted allows us to measure $\hat{\sigma}_{BE} = G_{ADC}\sigma_{BE} = 5.23 \pm 0.01\text{ ADU}$ (independent of gain setting). The VASE back end thus contributes the quadrature difference of 2.3 ADU. The equivalent noise charge (ENC) is $\hat{\sigma}_n$ divided by the conversion gain. Operating the high gain VASE2 without a CCD gives an ENC of $14.0 \pm 0.5\text{ e}$ on the highest gain settings, and $17.1 \pm 0.5\text{ e}$ with a CCD, so that the addition of the CCD increases the ENC by 9.8 e in quadrature. Note that for an ideal integrator, the noise would be independent of G_{INT} . In VASE2, though, there are two competing effects: as the gain decreases, the noise contribution from R_{INT} increases, while due to slight high frequency peaking in the integrator frequency response, there is a noise increase as the gain increases. This can be seen in the slight noise increase for the highest gains in Figure 13A.

As described above, the VASE signal is proportional to integration time. By reducing the integration time, a larger full scale signal can be accommodated on a given gain range. Figure 14A shows the roughly linear reduction in signal with integration time. Of course, decreasing the integration time increases the noise bandwidth by $1/\sqrt{T_{INT}}$ as seen in Figure 14B. This can be seen in Figures 13A, B which shows noise and full scale for two different integration times. At 50% of the nominal integration time, the maximum signal on the low gain VASE2 corresponds to $4 \times 10^5\text{ e}$. Adjusting the integration time provides additional flexibility in optimizing full scale signal with noise, such as making the same kind of measurement at different energies, or with large sample-to-sample variations.

5 Conclusion

The VeryFastCCD is the fully-column-parallel successor to the FastCCD. Fabricated in the same CCD process, it has a 25 times higher frame rate. With 10 nm and 100 nm entrance window contacts, it has high quantum efficiency for soft and tender X-rays. The custom readout circuit has programmable gains and

can also use integration time as a way to trade full-scale signal for noise. Prototype versions have been tested at ALS and LCLS, and systems for ALS are currently being prepared.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

AG: Formal Analysis, Writing–review and editing. CG: Conceptualization, Methodology, Writing–review and editing. JJ: Methodology, Writing–review and editing. AK: Methodology, Writing–review and editing. CT: Methodology, Writing–review and editing. PD: Conceptualization, Methodology, Project administration, Writing–original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the U.S. Department of Energy, Office of Basic Energy Sciences, Scientific User Facilities Division, through grants from the Accelerator and Detector Research Program, and by the Linac Coherent Light Source and the Advanced Light Source. Work was performed at the Lawrence Berkeley National Laboratory, which is supported by the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

Acknowledgments

The authors would like to acknowledge members of the LBNL Engineering Division: Ian Johnson, Thorsten Stezelberger and Vamsi Vytla for design and development of the data acquisition system, and Armin Karcher for electronics design. We also acknowledge design concepts and measurement support from Rebecca Armenta, Gabriella Carini, Philip Hart, Mark McKelvey, and Kaz Nakahara of the Linac Coherent Light Source and Rich Celestre and David Shapiro of the Advanced Light Source.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

4 <https://www.ti.com> > dataconverters > ads52j65.

References

1. Boyle WS, Smith GE. Charge coupled semiconductor devices. *Bell Syst Tech J* (1970) 49(4):587–93. doi:10.1002/j.1538-7305.1970.tb01790.x
2. Stover RJ, Wei M, Lee Y, Gilmore DK, Holland SE, Yi X, et al. High-performance CCD on high-resistivity silicon. In: *Imaging System Technology for Remote Sensing: Proceedings of SPIE* (1998). p. 13–8.
3. Struder L, Brauning H, Briel U, Hartmann R, Hartner G, Hauff D, et al. A 36 cm² large monolithic pn-charge coupled device x-ray detector for the European XMM satellite mission. *Rev Scientific Instr* (1997) 68(11):4271–4. doi:10.1063/1.1148341
4. Denes P, Doering D, Padmore HA, Walder JP, Weizeorick J. A fast, direct x-ray detection charge-coupled device. *Rev Scientific Instr* (2009) 80(8):083302. doi:10.1063/1.3187222
5. Kameshima T, Ono S, Kudo T, Ozaki K, Kiriara Y, Kobayashi K, et al. Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Rev Scientific Instr* (2014) 85(3):033110. doi:10.1063/1.4867668
6. Holland SE, Wang NW, Moses WW. Development of low noise, back-side illuminated silicon photodiode arrays. *IEEE Trans Nucl Sci* (1997) 44(3):443–7. doi:10.1109/23.603687
7. Underwood JH, Gullikson EM, Koike M, Batson PJ, Denham PE, Franck KD, et al. Calibration and standards beamline 6.3.2 at the advanced light source. *Rev Scientific Instr* (1996) 67(9):3372. doi:10.1063/1.1147338
8. Sze SM, Ng K. p-n Junctions. In: *Physics of semiconductor devices*. New Jersey, United States: Wiley (2006). p. 77–133.
9. Chilingarov A. Temperature dependence of the current generated in Si bulk. *J Instrumentation* (2013) 8(10):P10003. doi:10.1088/1748-0221/8/10/p10003
10. Grace CR, Denes P, Fong E, Goldschmidt A, Papadopoulou A. A 4-MHz, 256-channel readout ASIC for column-parallel CCDs with 78.7-dB dynamic range. *IEEE Trans Nucl Sci* (2020) 67(5):823–31. doi:10.1109/tns.2020.2980769
11. Walder JP, Denes P, Grace C, Hvd L, Zheng B. A fast low noise CMOS charge sensitive preamplifier for column parallel CCD readout. In: 2011 IEEE Nuclear Science Symposium Conference Record; 30 October - 6 November 2010; Knoxville, Tennessee, USA (2011). p. 721–4.
12. Shapiro DA, Celestre R, Enders B, Joseph J, Krishnan H, Marcus MA, et al. The COSMIC imaging beamline at the advanced light source: a new facility for spectro-microscopy of nano-materials. *Microsc Microanalysis* (2018) 24(S2):8–11. doi:10.1017/s1431927618012485



OPEN ACCESS

EDITED BY

Jiaguo Zhang,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY

Andrea Castoldi,
Polytechnic University of Milan, Italy
Herman Larsen,
United Kingdom Research and
Innovation, United Kingdom
Aldo Mozzanica,
Paul Scherrer Institut (PSI), Switzerland

*CORRESPONDENCE

Azriel Goldschmidt,
✉ agoldschmidt@lbl.gov

RECEIVED 29 August 2023

ACCEPTED 26 October 2023

PUBLISHED 20 November 2023

CITATION

Andresen N, Bakalis C, Denes P,
Goldschmidt A, Johnson I, Joseph JM,
Karcher A, Krieger A and Tindall C (2023),
A low noise CMOS camera system for 2D
resonant inelastic soft X-ray scattering.
Front. Phys. 11:1285379.
doi: 10.3389/fphy.2023.1285379

COPYRIGHT

© 2023 Andresen, Bakalis, Denes,
Goldschmidt, Johnson, Joseph, Karcher,
Krieger and Tindall. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

A low noise CMOS camera system for 2D resonant inelastic soft X-ray scattering

Nord Andresen, Christos Bakalis, Peter Denes,
Azriel Goldschmidt*, Ian Johnson, John M. Joseph,
Armin Karcher, Amanda Krieger and Craig Tindall

Lawrence Berkeley National Laboratory, Engineering Division, Berkeley, CA, United States

Resonant Inelastic X-ray Scattering (RIXS) is a powerful spectroscopic technique to study quantum properties of materials in the bulk. A novel variant of RIXS, called 2D RIXS, enables concurrent measurement of the scattered X-ray spectrum for a wide range of input energies, improving on the typically low throughput of 1D RIXS. In the soft X-ray domain, 2D RIXS demands an X-ray camera system with small pixels, large area, high quantum efficiency and low noise to limit the false detection rate in long duration exposures. We designed and implemented a 7.5 Megapixel back-illuminated CMOS detector with 5 μm pixels and high quantum efficiency in the 200–1,000 eV X-ray energy range for the QERLIN 2D RIXS spectrometer at the Advanced Light Source. The QERLIN beamline and detector are currently in commissioning. The camera noise from *in-situ* 3 s long dark exposures is $7e^-$ or less and the leakage current is $6.5 \times 10^{-3} e^-/(\text{pixel} \cdot \text{s})$. For individual 500 eV X-rays, the expected efficiency is greater than 75% and the false detection rate is $\sim 1 \times 10^{-5}$ per pixel.

KEYWORDS

RIXS, x-ray, CMOS sensor, spectrometer, back-illumination, low noise, synchrotron, light source

1 Introduction

Resonant Inelastic X-ray Scattering (RIXS) is a technique useful to study quantum properties of materials in the bulk [1]. In its simplest and most common implementation, a focused monochromatic X-ray beam impinges on the material under study with an energy very near a chosen element's atomic electronic transition. When an atom of the specific element in the sample absorbs a beam photon and de-excites through the same atomic transition (resonant condition), the emitted outgoing photon can have the same energy as the beam photons (elastic scattering) or slightly smaller energy (inelastic scattering). In the inelastic case, the outgoing X-ray photon spectrum, which is intrinsically sharp because the final state of the emitting system is the ground state of the atom, carries information about the intrinsic excitations of the molecule/material in which the atom is embedded. In a typical RIXS experiment a spectrometer collects a fraction of the outgoing photons and disperses them by energy, with energy resolving gratings, and a camera captures the image at the end of a long free-flight spectrometer arm to measure the photon-out spectrum.

The cross section for resonant inelastic scattering is small and so is the typical angular acceptance of the spectrometers' optics due to the mirrors' dimensions. The dispersing gratings have less than 5% efficiency. These three factors make RIXS, comparatively, a photon-starved technique. Bright light sources, such as synchrotrons and free electron lasers,

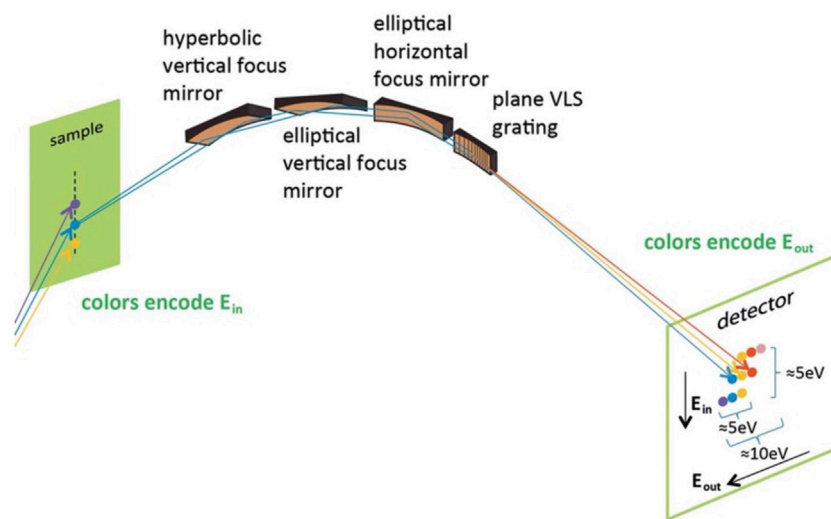


FIGURE 1

2D RIXS spectrometer scheme. Reproduced with permission from the Journal of Synchrotron Radiation. The monochromator that disperses the incoming photon energies before impinging on the sample is not shown. An (x,y) point on the detector image encodes the impinging X-ray energy (y-coordinate) and the scattered X-ray energy (x-coordinate), thus providing a full RIXS map in one exposure [3].

along with optimized optics designs and efficient X-ray detectors help mitigate this limitation. Still, typical RIXS experiments require exposure times of tens of minutes to obtain a spectrum for a single beam energy value and many such spectra, captured at varying beam energies, are required to produce a full RIXS map.

The novel 2D RIXS concept [2] further addresses the throughput of RIXS. In this variant of RIXS the input X-ray beam is broad spectrum in order to measure the full range of input energies of interest simultaneously. An optical element disperses the input X-ray beam by energy in one dimension (e.g., y). This results in an illuminated segment of a line on the sample such that adjacent points have slightly different X-ray energy photons impinging on them. Like in normal 1D RIXS, optical elements, in this case elliptic and hyperbolic mirrors, focus a fraction of the outgoing scattered X-rays and a grating disperses them in the orthogonal dimension (e.g., x) by their outgoing energy. The mirrors and gratings are arranged such that the 2D pixelated camera at the focal plane measures the full RIXS map. The x-dimension of the image encodes the scattered, outgoing, photon energy while the y-dimension encodes the incoming photon energy. Figure 1 shows the 2D RIXS beam/spectrometer/detection scheme.

The concept of 2D RIXS led to the design of the new QERLIN [3] beamline at the Advanced Light Source (ALS) with a RIXS spectrometer. The space-constrained 4.5 m photon flight path length, the targeted spectrometer resolving power of $\lambda/\Delta\lambda = 30,000$, and the 5–10 eV range of in/out photon energies around the atomic excitation level, constrain the requirements for key dimensions of the camera sensor. In particular, the sensor pixel needs to be 5 μm or less. Furthermore, the requirements of a high quantum efficiency in the 200–1000 eV X-ray energies of interest and of the ability to detect individual X-rays with a false rate (false positives) negligible with respect to the RIXS signal level constrain the sensor technology, sensor post-processing, temperature of operation and electronic readout noise.

Because of the lack of commercially available cameras that could fulfill all the requirements at the time when the QERLIN system was developed, the design of the custom QERLIN sensor and camera was co-designed with the QERLIN beamline and the spectrometer. Very recently, cameras with similar specification have become commercially available [4].

In this paper, we describe the QERLIN camera system, including the custom built sensor, readout electronics, camera/sensor cooling, mechanical and vacuum components, data acquisition, and image post-processing. We show the camera performance in bench-top testing, including dark images and 5.9 keV soft X-rays response and sensitivity results from a dedicated measurement at the ALS metrology beamline with 500 and 900 eV X-rays.

At the time of this writing, the QERLIN spectrometer is being commissioned and the camera has not yet seen first spectrometer light. We show, however, the camera performance in dark images as installed at the end of the spectrometer and measure the expected false rate as a function of X-ray energy based on the real dark images and on the expected and partly characterized X-ray signal response.

2 CMOS sensor

2.1 Design

The QERLIN sensor consists of a 2,048 rows by 3,840 columns array of 5 $\mu\text{m} \times 5 \mu\text{m}$ pixels that satisfies the spectrometer resolving power requirements and in and out energy ranges given the 4.5 m long QERLIN spectrometer arm and its optics. It is a thinned back-illuminated CMOS Active Pixel Sensor with a 4T architecture [5]. As such, there are 4 transistors per pixel. Charge is collected in a pinned photodiode structure in the pixel. A transfer gate moves the charge to the output node (also referred to as the floating diffusion) which is previously cleared by means of a reset transistor. A third transistor, in a source-follower

**FIGURE 2**

Picture of back side of the thinned QERLIN sensor. A full thickness frame is left for rigidity and ease of handling.

configuration, outputs a voltage proportional to the collected charge. The fourth transistor, in a switch configuration, selects the pixel for readout.

The QERLIN sensor is segmented into 16 regions of 2048 (rows) by 240 (columns) pixels, where each region has an independent analog output. The fourth in-pixel transistor selects the row, while a multiplexer at the bottom of the columns selects the column to be presented at the analog output for each channel. Externally provided signals control in-chip logic that can a) turn on the transfer gate transistors for all the pixels or only for the selected row b) reset the floating diffusions for all the pixels or only for the selected row c) select the next row d) select the next column and e) reset the digital logic to select the first row and first column of each channel. The sensor logic supports both global shutter and rolling shutter operation, the latter by resetting the previous row's pixels automatically.

Besides digital power, the sensor requires externally supplied DC voltages, 1.8 V and 3.3 V, and bias currents for its operation: a reset voltage, a pixel source follower voltage, rails for the transfer gate and for the reset transistor gate, along with a voltage and two currents for the bottom-of-column circuitry and a voltage for the output stage.

The QERLIN sensor has 16 analog outputs along the bottom edge of the chip. Voltage supplies to power the bottom-of-column (BOC) and for the output stage are also along the bottom side along with bias currents for the BOC analog circuitry. On the top side there are pads to supply the voltage for the sensor columns and the voltage for pixel reset. On the left side are pads for LVDS clocks and for other LVDS digital signals that control sensor timing and reset/exposure/readout sequencing.

2.2 Fabrication and post-processing

Fabrication of the sensor was done using a UMC 180 nm CMOS image sensor process. The starting material is a p-type silicon substrate with a moderate resistivity (with type p-) epitaxial (EPI) layer. The EPI layer, which functions as the sensitive volume of the sensor, is 4 μm thick.

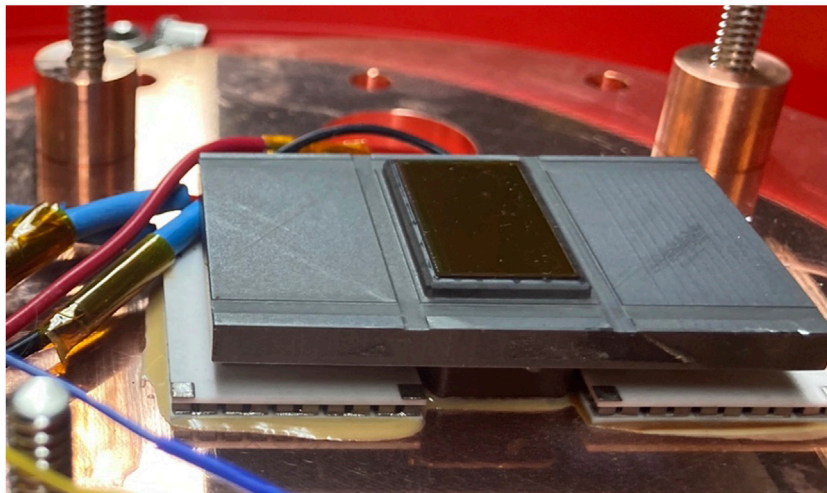
Soft X-rays in the 200–1,000 eV have attenuation lengths in silicon from 63 nm to 2.7 μm . The passivation, metal and oxide layers on the transistor-implanted/patterned side of the sensor are several micrometers thick. Thus, illumination from the non-patterned side, often referred to as back-illumination, is required for efficient detection of X-rays in this energy range. Furthermore, the silicon substrate material, where charge carriers readily recombine, needs to be removed. This process, referred to as thinning, exposes the sensitive EPI layer to the X-ray illumination.

In order to maintain the sensor's mechanical rigidity, only the imaging area of the QERLIN sensor was thinned to the EPI (by an outside vendor) after dicing, leaving a full thickness frame around it. The picture in Figure 2 shows the back side of the fully post-processed QERLIN sensor.

The etched surface of the thinned sensor has a relatively high number of microscopic defects left over from the etching process. These defects produce an unacceptable level of thermally generated leakage current. Therefore a high quality, p-type layer on the entrance side of the device is needed to isolate these defects from the electric field of the device and prevent them from injecting current into the active volume of the detector. No explicit contact to ground of the back surface is made but the conductive implanted layer is effectively grounded on the edges of the sensor.

Since thinned devices cannot be further processed at the foundry, it is necessary to have a contact fabrication process that can be performed on fully metallized chips. This means that the maximum processing temperature that can be used has to be low enough to avoid damaging the aluminum metallization and the interlevel dielectric layers (ILD). Temperatures that are too high can alloy the metal with the underlying silicon or crack the ILD, either of which will destroy the chip. The exact temperature at which this happens is highly dependent on the particular foundry process used to make the sensor.

For the QERLIN sensor, we chose to use a process that we have developed to fabricate contacts at low temperatures. This process uses ion implantation followed by annealing to electrically activate

**FIGURE 3**

The internal thermal/mechanical assembly: The 1.2 cm x 2.0 cm QERLIN sensor is in the center (soft X-rays entrance is from the bottom of the image -the thinned back side of the sensor-), glued to a SiC which in turn is glued to the cold-end of 2 TECs.

the dopant. This process produces high quality contacts that are ~100 nm thick. Contacts fabricated using this method effectively suppress the leakage current to a level that is acceptable for this application.

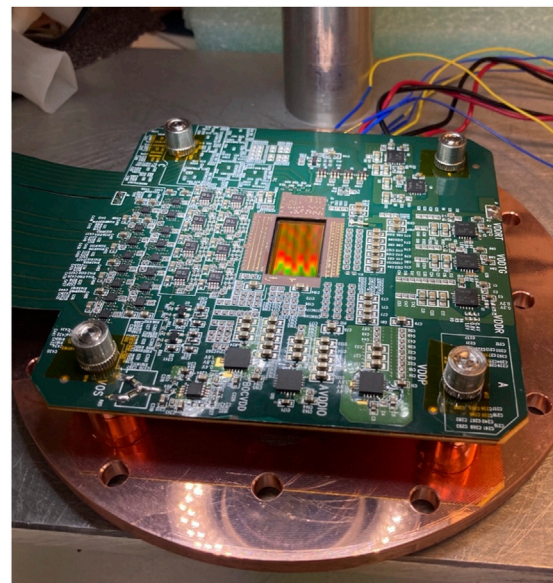
We have developed an additional contact fabrication process based on molecular beam epitaxy (MBE) that produces high quality contacts of 10 nm thickness or less. With such thin contacts the efficiency at the low end of the soft X-ray energy range can be increased four-fold. For instance, for 200 eV X-rays the attenuation length in silicon is 63 nm and the 10 nm dead layer absorbs a very small fraction of the photons. We plan to replace the sensor in the QERLIN camera with an MBE contact device in the near future.

3 Camera system

3.1 Mechanical design and temperature control

Even in a thinned detector, leakage current is significant at room temperatures. Therefore, cooling the sensor is important. Cooling, in turn, enables the long exposure times useful in a soft X-ray spectrometer due to its low photon flux. In addition, long exposure times (without saturation from X-ray signal or from dark current) are desirable in order to increase the signal level with a fixed readout noise contribution.

The thermal design has two main requirements: to cool the sensor to a stable temperature between -20°C and -50°C and to prevent the overheating ($<60^{\circ}\text{C}$) of the in-vacuum electronics (pre-amps, etc.). The sensor is glued to a silicon carbide (SiC) thermal-bus with a thin film of thermally conducting epoxy around the non-thinned frame of the sensor. Besides being a good thermal conductor, SiC has a similar coefficient of thermal expansion to the silicon sensor. The cold sides of the two three-stage solid-state thermoelectric coolers (TECs) are glued to the SiC thermal-bus with thermally conductive epoxy. The hot surface of the TECs is likewise

**FIGURE 4**

In-vacuum electronics. While the sensor is operated at -50°C , the in-vacuum electronics is thermally isolated from the cold side of the TECs and is kept near room temperature.

glued to a thick copper plate. When operating at full power, the TECs dissipate about 45 W each and are the main thermal load on the system. The picture in Figure 3 shows the sensor/SiC/TECs/thick-copper-plate assembly. The picture shows the side of the sensor where the wire bonds are made, while the X-rays impinge from the bottom thinned-side. Two PT-100 thermistors monitor the SiC/sensor cold temperature and the thick-copper-plate (near room) temperature. The wires to power the TECs and to measure the PT-100 resistances come out of the vacuum enclosure through a dedicated 9-pin feedthrough.

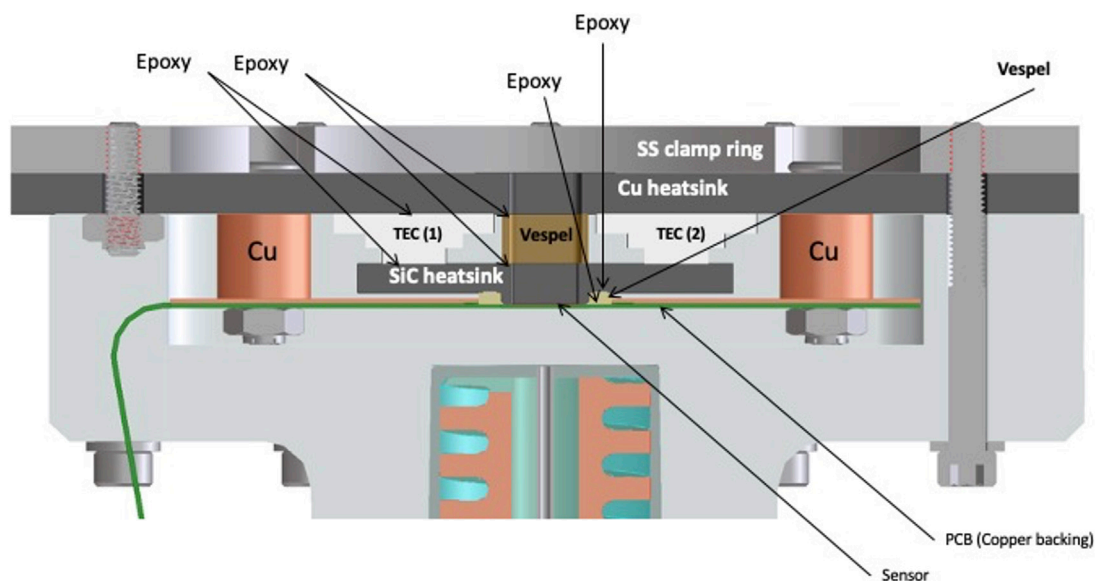


FIGURE 5

Details of the thermo-mechanical design of the QERLIN camera. X-rays come from the top side and the layers before the sensor have cut-outs.

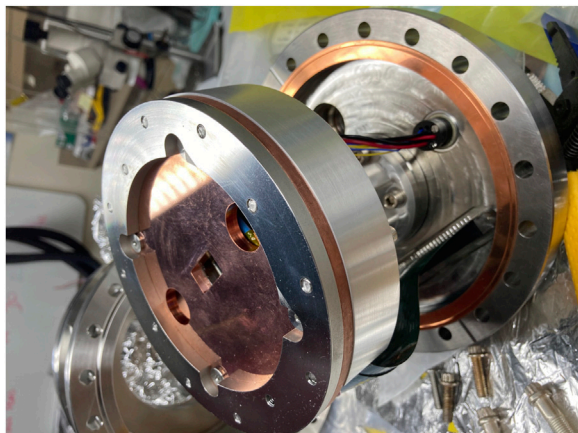


FIGURE 6

Flange-mounted in-vacuum components of the QERLIN camera. The rectangular opening in the copper plate is the entrance port for the X-rays.

The in-vacuum electronics board is a flex-circuit board with discrete components for pre-amplifiers, current sources, and voltage sources for the sensor operation. The flex-circuit board is glued to a backing of thin copper sheet. This sheet is mechanically attached and thus thermally coupled to the externally cooled thick-copper-plate. A thin Vespal frame around the sensor thermally isolates the flex-circuit board from the SiC to minimize the heat load on the TEC's cold-side. Figure 4 shows the in-vacuum assembly right before wire-bonding of the sensor.

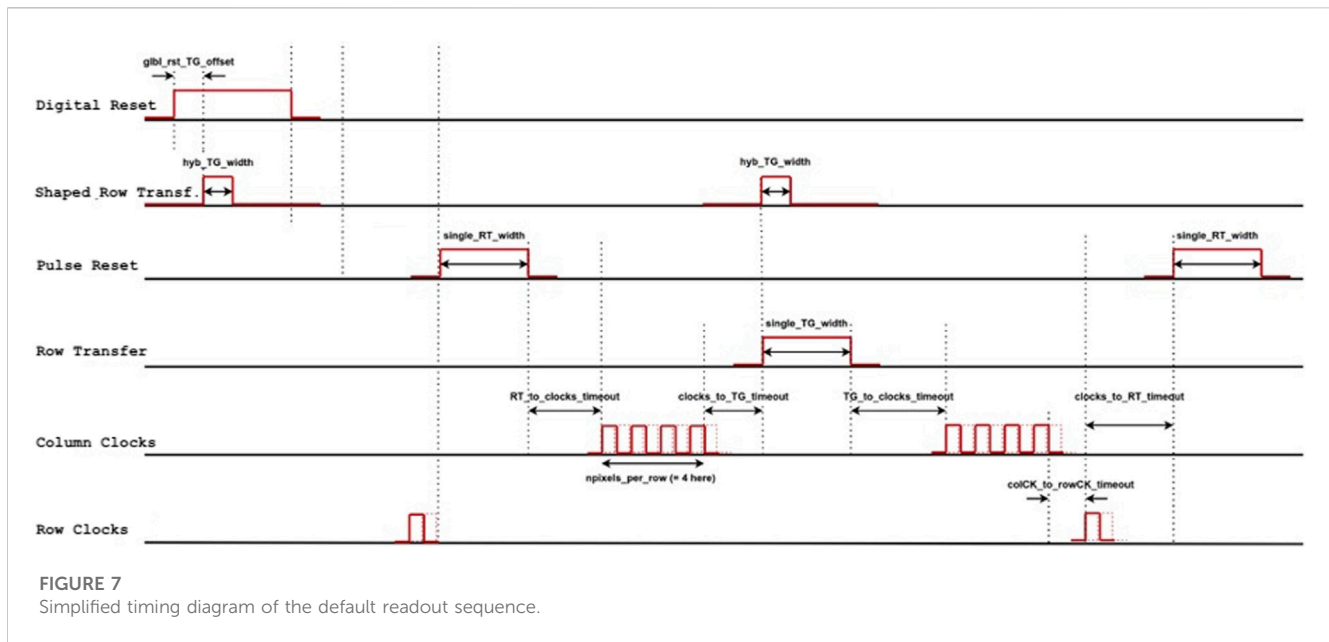
The thick-copper-plate with the sensor assembly is sandwiched between a clamping ring and the aluminum-body which, in turn, is mated to a 2.5" Conflat flange at the center of the main 8" camera vacuum flange. The aluminum-body accepts, from the air-side, a

tightly fitting stainless steel cold-probe. Coolant flows through the internal manifold of this cold-probe to facilitate the room-temperature cooling circuit. The nearly 100 W of power dissipated in-vacuum is thus removed by the coolant loop of the cold-probe and maintains the temperature of the thick-copper-plate and electronics to less than 40°C and the QERLIN sensor (via the TECs) near -50°C. Figure 5 shows the design of the thermal components around the sensor and the in-vacuum electronics and Figure 6 shows the fully assembled QERLIN camera mounted on the main 8" flange. The rectangular cutout in the thick-copper-plate is the entrance window for the X-rays from the spectrometer to the back-side of the sensor. The sensor is recessed with respect to the copper plate but this has no impact on the camera X-ray acceptance because the incident rays are nearly perpendicular to the sensor. The two circular openings, adjacent to the entrance-window, contain reed-valves (movable flaps of polyamide film). These reed-valves prevent a large pressure differential (during system pump-down, venting and vacuum failures) between the backside and the front-side of the highly fragile thinned sensor. The sensor and reed-valves act as a vacuum barrier to isolate the non-VHV (very-high-vacuum), internal camera components operating at high-vacuum from the main system operating at VHV.

3.2 Readout system

The camera readout hardware is organized as follows: 1) the in-vacuum flex board that connects to the main flange vacuum feedthroughs with two 51-pin connectors, 2) an in-air electronics box that connects to the air-side of those feedthroughs and delivers digital data through an optical fiber and 3) a Linux server that receives the digital data via a 10 GigE optical fiber network.

The in-vacuum flex board has single-ended preamplifiers and single-ended-to-differential amplifiers for the 16 analog output



channels of the QERLIN sensor. The 16 differential analog outputs come out through one of the 51-pin feedthroughs. A single adjustable voltage controls the offset level of all the sixteen channels. Linear regulators on the board provide the DC voltages required for the sensor operation. Three current mirror circuits provide the DC currents needed and are powered with regulated voltages. All digital control signals and clocks come through one of the feedthrough interfaces as LVDS pairs. On the board, some of the signals (e.g., two non-overlapping clocks) are just passed through to the sensor, while others are converted to single-ended signals (like the global reset) as required by the sensor design.

The in-air electronics box has a simple pass-through motherboard, a power board, a transfer gate pulser board and the main DAQ board (ADAQ). This ADAQ board has 16 channels of differential amplifiers, 4xQuad ADS5263 16-bit ADC chips capable of 100 mega-sample per second per channel, an Enclustra KX1-325 FPGA module and two 10GigE network interfaces. The custom FPGA code orchestrates the sensor signaling, through a set of 16 LVDS signals, and the output digitization. It organizes the digital output as full physical image frames. It presents the organized digitized data through the 10GigE network interface as UDP packages. Control of the FPGA DAQ cycle is through a set of about 30 control registers that are accessible from the Linux server through a dedicated network port opened by the FPGA on the 10GigE connection.

3.3 Readout modes

The QERLIN sensor supports both global shutter and rolling shutter readout modes. In addition, the device provides a large degree of operational flexibility because most of its control signals are externally supplied. As discussed above, the 2D RIXS application benefits from as low a noise figure as possible and the required and desirable frame rate is from slow O(Hz) to very slow O (mHz). In what follows we describe the default readout cycle for the QERLIN

camera with rolling shutter with pause, sample/reset averaging and correlated double sampling. This readout cycle was used to take the characterization data presented in the next sections. Figure 7 shows a simplified timing diagram of the readout sequence.

A Digital Reset signal prepares the device for selection of the first physical row of pixels. Then a Row Clock signal selects the next (first) row. A Pulse Reset signal then resets the voltage on the output node (floating diffusion) of all the pixels in that row. A Column Clock signal then selects the next (first) column in the channel (or super-column). Eight consecutive ADC samples (for the same reset pixel) are acquired to be averaged in the FPGA to reduce the readout thermal noise component. A new Column Clock signal then selects the following column in the super-channel followed by the corresponding 8-sample read. This is repeated until the averaged reset samples from all 240 columns of the channel have been acquired. Next, the charges stored in the pixels' photodiodes of the selected row are transferred to the output nodes. This is achieved by issuing a Row Transfer signal. After the charge transfer is completed the pixels are read out analogously to the reset read sequence (*i.e.* 8 ADC samples per pixel that are then averaged). After the readout of the entire row, a new Row Clock signal selects the next row for corresponding reset/signal readout. While the next row is being read out, the previous row is (automatically) fully reset by having its transfer gate on while the reset is issued (thus resetting both the output nodes and the photodiodes). After the last row of the sensor is read out and reset, an arbitrary duration pause of all signals extends the pixels' integration time to the desired total frame exposure time. At the pause's end the cycle restarts with a new Digital Reset signal. In this fashion, all the pixels in the sensor have equal (although not fully contemporaneous) exposure time.

The entire cycle takes about 3.2 s when the pause between frames is removed. With the ADCs clocked at 12.5 MHz the ADC sampling (8 + 8 samples per pixel) takes only 0.6 s (80 ns per sample \times 240 columns per channel \times 2048 rows \times 16 samples per pixel). The bulk of the additional time, 2 out of the remaining 2.4 s, is used for the charge transfer operation between the photodiodes and

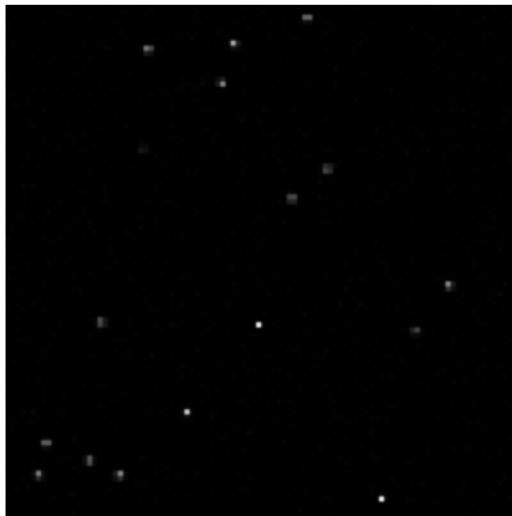


FIGURE 9

A random 100×100 pixel patch with several X-rays clusters. The image contrast is set such that the brightest X-ray pixels are white and it is not zero-suppressed.

(non-dark) frames. For data runs where the external illumination is very sparse (such as the ^{55}Fe 5.9 keV X-ray calibration data sets described in the following section) the CDS dark image can be obtained from the illuminated data set by computing the per-pixel median over a set of CDS images. The per-pixel median is a very good approximation to the most-probable-value in these very sparsely illuminated images and is used for convenience.

4 Camera performance

4.1 Characterization with 5.9 keV X-rays from ^{55}Fe source

Measurements of the camera response to 5.9 keV X-rays from a ^{55}Fe source provide information on the conversion factor from ADU (arbitrary digital units) to ionization electrons, on the sensor response uniformity and on the effective point spread function of the camera.

The camera was first pumped to a 1×10^{-5} torr pressure and then the TECs were turned on to maximum power until a stable temperature of about -50°C was reached. A 12 mCi ^{55}Fe source was placed outside the vacuum enclosure in front of a thin aluminized mylar window a few inches away from the back side of the QERLIN sensor.

In evenly illuminated dark-subtracted CDS frames almost point-like clusters of pixels from individual 5.9 keV X-rays are distributed uniformly over the entire sensor area. Figure 9 shows a random zoomed-in 100×100 pixel patch of an image. Typical 5.9 keV X-ray depositions have 1–4 pixels with a significant fraction of the X-ray induced charge depending on the exact location of the X-ray absorption within the pixel. Based on the roughly approximated experimental geometry, the source activity and the estimated X-ray absorption between the source and the sensor surface, the estimated

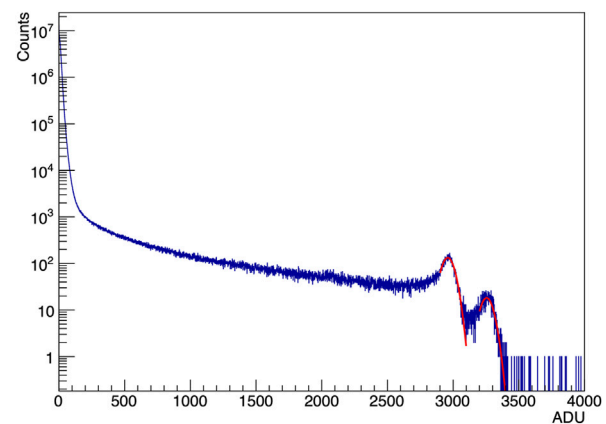


FIGURE 10

Single-pixel-value distribution for dark subtracted CDS images with ^{55}Fe X-rays. The two peaks near 3,000 are due to the two X-ray lines at 5.9 and 6.4 keV from the source. The fitted FWHM of the peaks is 61 e, using the ADU to e^- conversion derived from the peaks' locations.

impinging X-ray flux is ~ 0.006 X-rays per pixel per second. At this X-ray energy only about 4.5% of those convert in the 2.5 micron thick sensitive volume, therefore, for 3.6 s long exposures we expect 9.5 X-rays detected in a 100×100 pixel patch, in reasonable agreement with our observations.

Figure 10 shows the single pixel value distribution (after CDS and dark subtraction) from many frames. The Gaussian fitted peaks near 3000 ADU are due to “single-pixel” energy depositions from the K_α and K_β lines of the ^{55}Fe source. Their relative position matches to better than 1% the nominal 5.88 keV 6.49 keV source lines. These single-pixel depositions occur when the X-ray conversion happens in a fully depleted part of the pixel volume under the photodiode implant. The region below the peak in Figure 10 but above the residual noise near zero is due to pixel charge sharing from multi-pixel clusters. A calibration factor of $1.8 \text{ ADU}/e^-$ was deduced from the measured position of the single-pixel K_α peak and the average deposited energy to produce an e-h pair in silicon $W_{\text{Si}} = 3.6 \text{ eV}$.

Since a large fraction of the X-ray hits deposit their charge over multiple pixels we perform a simple cluster analysis. Seed pixels are identified as local maxima with values greater than $3\sigma_n$, where σ_n is the standard deviation of the pixel values in dark images. For each seed pixel the sum of the pixel values in the 3×3 region around the seed is computed. Figure 11 shows the distribution of the 3×3 cluster charge (now in a linear plot). The peaks near 3,000 and 3,300 are from 3×3 clusters seeded on “single-pixel” depositions. The broader and much larger peak near 2,500 is from clusters with shared charge. For these the charge collection is incomplete, leaving about 15% of the charge unaccounted for. Larger cluster regions (5×5 , 7×7) do not recover the missing charge. On the other hand, the spectrum of the 3×3 integrated signal is, as expected, much cleaner with far fewer entries between the noise peak and the iron peaks than the “single-pixel” distribution of Figure 10.

In order to study the camera response uniformity (excluding quantum efficiency effects), we select all clusters with a 3×3 integrated signal in a broad window around the peak (between

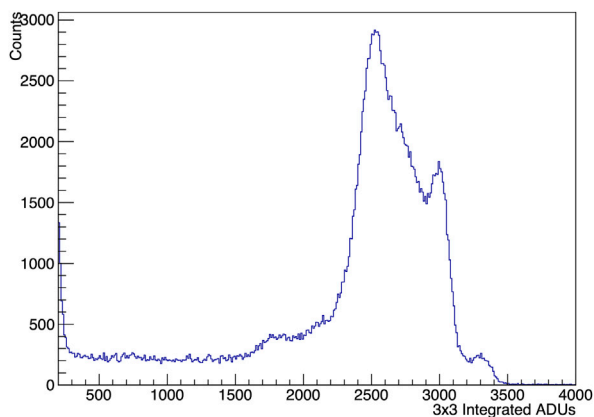


FIGURE 11
Distribution of the signal sum over 3×3 pixel clusters.

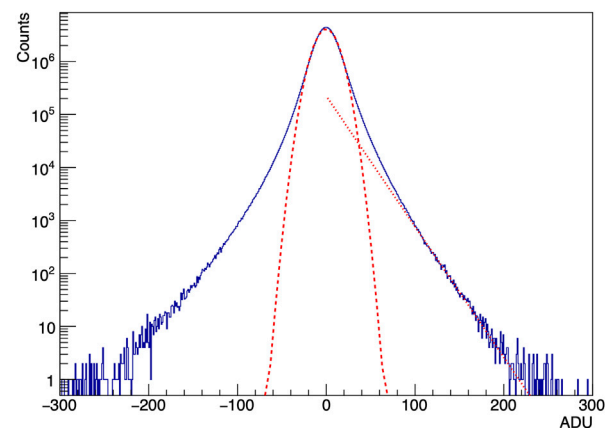


FIGURE 13
Noise residuals from a dark run (-50°C and 3.2 s exposure time).

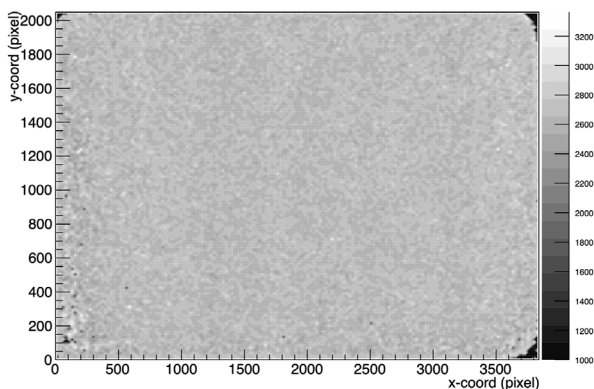


FIGURE 12
QERLIN camera response uniformity. The mean response is 2,650 ADU per X-ray and the RMS is 108 ADU. This corresponds to a 4% response uniformity (without any corrections to the data).

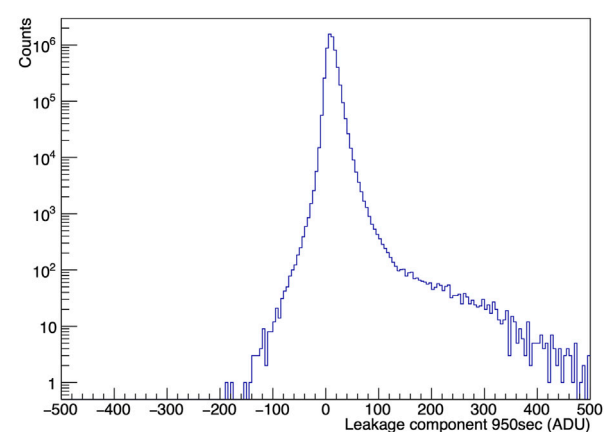


FIGURE 14
Raw dark current signal contribution from 950 s supplementary exposure. It corresponds to a very low $6.5 \times 10^{-3} \text{ e}^-/(\text{pixel} \cdot \text{s})$ leakage.

2,000 and 3,500 ADU). We re-bin the sensor area into patches of 16×24 pixels (to ensure enough statistics in each bin) and calculate the average value of the cluster charges. Figure 12 shows the sensor response to 5.9 keV X-rays. Within the statistical limit of the measurement the response is uniform. The dark corners are the result of the absence of clusters in those regions because the thinning process reached a slightly undersized region and the 5.9 keV X-rays are fully absorbed in the thick substrate of the un-thinned frame.

4.2 Noise

4.2.1 Noise characterization

We took dark data using the standard detector configuration (e.g., -50°C and 3.2 s/frame) with the QERLIN detector installed at the end of the QERLIN spectrometer arm at the ALS. Figure 13 shows the residuals from the median subtracted CDS dark images. The central component of the residuals distribution is well fitted to a

Gaussian (in the $[-30, 30]$ range) with 11.6 ADU sigma. Using the conversion factor from the X-ray data this corresponds to 6.5 e^- noise. There is, however, a clear non-Gaussian tail to the noise residuals. We fit these noise tails in the $[100, 300]$ range to a single exponential. Noise tails can be due to hot pixels with excess leakage current or other electronics effects. A correlation study between the dark current images (see next section) and noise images showed only a 0.13 correlation (where 0 is no correlation and 1 is full correlation) between those. This indicates that the bulk of the noise tail is not due to high leakage hot pixels.

4.2.2 Dark current

Two dark data sets with per-frame exposure times of 50 s and 1,000 s were utilized to determine the dark current in standard operating conditions (-50°C). For each data set we compute the (per-pixel) median of the set of CDS frames. We then subtract pixel-by-pixel the median image of the 50 s exposure from the median image of the 1,000 s exposure. Figure 14 shows the pixel value

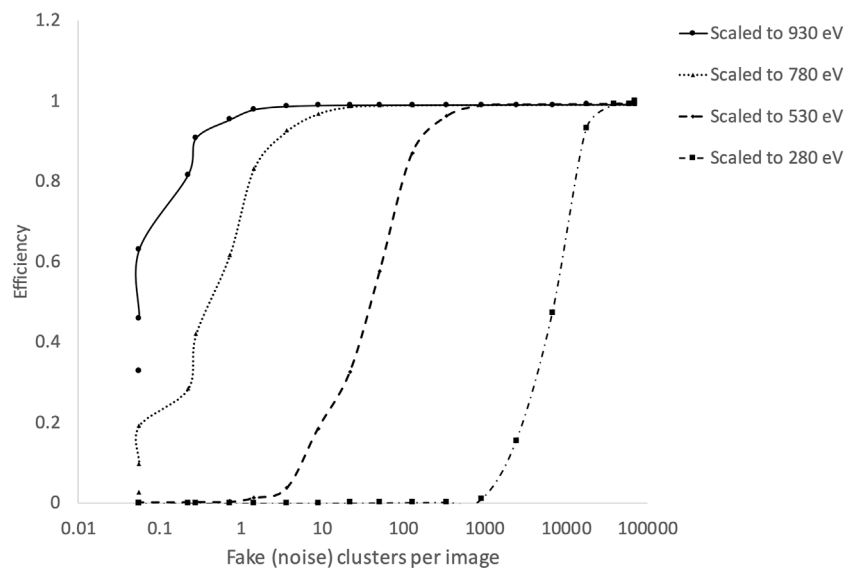


FIGURE 15

Calculated X-ray cluster detection efficiency vs. measured fake hit rate per image for various typical soft X-ray energies used in RIXS experiments. The tradeoff is controlled with a threshold on q_4 (see text). The efficiency calculation assumes a fully efficient (like from an MBE) back contact.

distribution of the difference image. The distribution has a mean of 11.2 ADU. Accounting for the additional exposure time this corresponds to 6.5×10^{-3} e/(pixel · s). This very low leakage component enables long up to 1,000 s exposure times without significantly increasing the overall noise. The distribution has a tail on the positive side due to $\sim 0.1\%$ of all pixels that have higher dark current.

4.2.3 Fake hits

As described in the introduction, in the 2D RIXS application a point in the image corresponds to an energy-in energy-out pair. Some regions of 2D RIXS images may have very low photon flux but still carry important information. For this reason, it is important to have a fake hit rate negligible compared to the RIXS signal. A fake hit is the measurement of localized pixel values consistent with the signal expected from an individual X-ray detection in a region where no X-ray hit.

We estimate the expected fake hit rate using the dark data runs and the expected soft X-ray signal. To estimate the expected signal from 200–1,000 eV soft X-rays we use the measured detector response to individual 5.9 keV ^{55}Fe X-rays and linearly scale the charge depositions by the X-ray energy. To find fake hits in the dark data set we use a two-threshold method. The first threshold requires that a local pixel maximum is greater than 3σ (like before, σ is obtained from the fit to the $[-30, 30]$ region of the noise residuals). For each local maximum, the sum of the four largest pixel values within a 3×3 pixel region around the maximum is calculated (q_4). The second threshold is on the value of q_4 . The q_4 threshold is scanned to compute, for each q_4 threshold value, the fake hit rate from the noise images and the soft X-ray detection efficiency from the scaled ^{55}Fe response. Figure 15 shows the calculated single X-ray detection efficiency *versus* the corresponding measured fake hit rate (the number of fake hits per image) as the q_4 is varied and for multiple soft X-ray energies of relevance in RIXS. For example, for individual 500 eV X-rays, one can

choose a threshold for which the expected efficiency is greater than 75% and the false detection rate is $\sim 1 \times 10^{-5}$ per pixel.

This calculation assumes no significant inefficiency from the absorption of X-rays in the thin dead layer on the illuminated back-side of the sensor. For 530 eV X-rays on a dead silicon layer of 120 nm this reduces the efficiency by 22%, while the reduction is 65% at 280 eV and just 5% at 930 eV. A future version of the sensor with an MBE contact dead layer of less than 10 nm should show efficiencies similar to those plotted.

To understand the effect of the expected fake hit rate, it is useful to look at an example. Take the hardest case of 280 eV X-rays (a sensor with an MBE contact is assumed): for a suitably chosen threshold we expect $>80\%$ quantum efficiency and an estimated 10,000 fake hits per image (see Figure 15), randomly and uniformly distributed. On the other hand, the “true” signal component over the entire image would have millions to 100s of millions of X-ray hits. Thus the fake hits represent a small contribution, between 0.01% and 1%. Additionally, in 2D RIXS maps produced from series of 1D RIXS spectra, the typical features span contiguous regions of many pixels, typically 100s. As a consequence, the spatially uncorrelated fake hits have the effect of marginally reducing the overall signal-to-noise of those multi-pixel features by adding a dim and spatially flat component.

4.3 Soft X-ray sensitivity and linearity

While most of the detector characterization was obtained from dark images and ^{55}Fe X-rays, the QERLIN camera was also installed and briefly tested at the ALS 6.3.2 metrology beamline. Unfortunately, the geometry of the setup was constrained and the beam illuminated an area at the sensor’s edge, with partial coverage. Given those circumstances, no reliable quantum efficiency

measurement could be obtained from that data. However, selecting the part of the images with a much dimmer beam halo we could identify and measure individual well-isolated X-rays at the two beam energies we used, 500 and 900 eV. From those, we verified the camera response linearity to $\sim 10\%$ in the 500 eV–5.9 keV range.

5 Conclusion

At the time of this writing, the 2D RIXS QERLIN beam line and spectrometer at ALS are being commissioned. The QERLIN camera described in this article is installed and operational at the spectrometer. The camera data acquisition is fully integrated with the beam line and spectrometer controls. We demonstrated the camera performance in dark images as installed and measured the expected fake rate as a function of X-ray energy based on the real dark images and the expected and partly characterized X-ray signal response.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

NA: Conceptualization, Investigation, Writing–review and editing. CB: Methodology, Software, Writing–review and editing. PD: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing–review and editing. AG: Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing–original draft. IJ: Methodology, Software, Writing–review and editing. JJ: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing–review and editing. AKa: Investigation, Methodology, Writing–review and editing. AKr: Conceptualization, Investigation, Methodology, Writing–review and editing. CT: Conceptualization, Methodology, Writing–review and editing.

References

1. Ament LJP, van Veenendaal M, Devereaux TP, Hill JP, van den Brink J. Resonant inelastic X-ray scattering studies of elementary excitations. *Rev Mod Phys* (2011) 83(2): 705–67. doi:10.1103/RevModPhys.83.705
2. Strocov VN. Concept of a spectrometer for resonant inelastic X-ray scattering with parallel detection in incoming and outgoing photon energies. *J Synchrotron Radiat* (2010) 17(1):103–6. doi:10.1107/S0909049509051097
3. Warwick T, Chuang Y, Voronov DL, Padmore HA. A multiplexed high-resolution imaging spectrometer for resonant inelastic soft X-ray scattering spectroscopy. *J Synchrotron Radiat* (2014) 21(4):736–43. doi:10.1107/S1600577514009692
4. Marana-X sCMOS. Examples of these are the Marana-X sCMOS cameras, with slightly bigger 11 or 6.5 μm pixels and $>90\%$ quantum efficiency in the soft X-ray regime along with $<2\text{ e}^-$ noise and fast readout (2023).
5. Lee P, Gee R, Guidash M, Lee T-H, Fossum ER. An active pixel sensor fabricated using CMOS/CCD process technology, Proceedings of the presented at the IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, Dana Point, California, April 1995. 1995.
6. Teranishi N, Mutoh N. Partition noise in CCD signal detection. *IEEE Trans Electron Devices* (1986) 33(11):1696–701. doi:10.1109/T-ED.1986.22730
7. Goiffon V, Estribeau M, Cervantes P, Molina R, Gaillardin M, Magnan P. Influence of transfer gate design and bias on the radiation hardness of pinned photodiode CMOS image sensors. *IEEE Trans Nucl Sci* (2014) 61(6):3290–301. doi:10.1109/TNS.2014.2360773

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Advanced Light Source and by the U.S. Department of Energy, Office of Basic Energy Sciences, Scientific User Facilities Division, through grants from the Accelerator and Detector Research Program. Work was performed at the Lawrence Berkeley National Laboratory, which is supported by the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

Acknowledgments

We thank Yi-De Chuang, Harold Barnard and Alastair McDowell for their help throughout the camera development, installation and commissioning at the ALS QERLIN spectrometer. We thank Martin Keller for his work on the integration of the camera with the beam and spectrometer control and monitoring. Braeden Benedict designed and built the camera temperature controller. Michael Holmes performed finite element analysis of the in-vacuum thermal design. We thank Terry McAfee and Eric Gullikson for their help with the camera measurements at the ALS 6.3.2 metrology beam line.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Cornelia B. Wunderer,
Helmholtz Association of German
Research Centres (HZ), Germany

REVIEWED BY

Takaki Hatsui,
RIKEN SPring-8 Center, Japan
Bernd Schmitt,
Paul Scherrer Institut (PSI), Switzerland

*CORRESPONDENCE

Sol M. Gruner,
✉ smg26@cornell.edu

RECEIVED 30 August 2023

ACCEPTED 13 November 2023

PUBLISHED 23 November 2023

CITATION

Gruner SM, Carini G and Miceli A (2023),
Considerations about future hard x-ray
area detectors.
Front. Phys. 11:1285821.
doi: 10.3389/fphy.2023.1285821

COPYRIGHT

© 2023 Gruner, Carini and Miceli. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Considerations about future hard x-ray area detectors

Sol M. Gruner^{1*}, Gabriella Carini² and Antonino Miceli³

¹Physics Department, Cornell University, Ithaca, NY, United States, ²Instrumentation Division, Brookhaven National Laboratory, Upton, NY, United States, ³X-ray Science Division, Argonne National Laboratory, Lemont, IL, United States

X-ray sources continue to advance in both intensity and temporal domains, thereby opening new ways to analyze the structure and properties of matter, provided that the resultant x-ray images can be efficiently and quantitatively recorded. In this perspective we focus on specific limitations of pixel area x-ray detectors. Although pixel area x-ray detectors have also advanced in recent years, many experiments are still detector limited. Specifically, there is need for detectors that can acquire successive images at GHz rates; detectors that can accurately measure both single photon and millions of photons per pixel in the same image at frame rates of hundreds of kHz; and detectors that efficiently capture images of very hard x-rays (20 keV to several hundred keV). The data volumes and data rates of state-of-the-art detection exceeds most practical data storage options and readout bandwidths, thereby necessitating on-line processing of data prior to, or *in lieu* of full frame readouts.

KEYWORDS

wide dynamic range, online detector processing, synchrotron radiation, x-ray burst rate detector, x-ray pixel array detectors

1 Introduction

X-ray analysis of matter has advanced greatly with the advent of brighter and more intense synchrotron radiation (SR) sources. This has enabled the development and application of techniques that were practically unfeasible only a decade or two ago. Examples include the real-time monitoring of microstructural details during materials synthesis (e.g., 3-D printing, thin film processing), nanometer-level ptychography of complex non-periodic objects (e.g., bone, integrated circuits, alloys), rapid optimization of the composition of multi-component thin-film catalysts, etc. In all these cases, improved x-ray sources enable experiments by providing the required numbers of x-rays arriving at the specimen with requisite time-structure, divergence, photon energies, and focal spot size.

However, getting x-rays to the sample is only part of the challenge: One must also efficiently detect the x-rays emanating from the specimen. X-ray detection technology has historically lagged source development and continues to constrain practical performance of many experiments.

The purpose of this perspective is to draw attention to several detector limitations presently constraining specimen analysis. It is impossible in a short Perspective to cover all types of experiments done at SR sources, or on all types of detectors. The focus here will be on “direct” (see below) detectors for very hard (>20 keV) x-ray diffraction experiments at both storage ring and XFEL applications. The focus will also be limited to integrating pixel array area detectors because photon-counting detectors cannot handle many x-rays/pixel/x-ray pulse often encountered at modern storage rings and XFELs [1].

2 Detector limitations: where are the needs?

2.1 Burst rate imaging

Much of the Universe consists of “warm, dense matter” in stars and planets where the densities are greater than Earth surface solids and the constituent atoms or ions have eV-scale thermal energies. Pulsed laser heating can produce transient warm, dense matter in the lab but it quickly explodes; hence, a need to capture x-ray images very quickly. SR sources can deliver sufficiently intense successive x-ray pulses at very fast rates, thereby enabling study of warm, dense matter dynamics. This requires “burst-rate detectors” [1] to record successive diffraction patterns, or “frames” within the time envelope of the event, typically in the ps to μ s range. Other experiments requiring burst-rate detection include analysis of shock waves and study of crack propagation and materials failure under sudden stress.

Burst-rate detectors may utilize either “direct” or “indirect” detection of the results of x-ray absorption in a “sensor screen”. In the former, x-rays absorbed in a sensor screen, such as a sheet of appropriately biased semiconductor, produce electron-hole pairs. This yields an electrical current that is directly processed by in-pixel electronics. Indirect detectors use a sensor screen that produce other types of quanta, such as visible light from a scintillator that is then recorded by a camera. Attention in this short Perspective is confined to direct detection.

State-of-the-art burst-rate detectors frame at nearly 1 GHz [2–5], a rate that is too slow for processes occurring on ps time scales but adequate for slower processes. A primary limitation is the detector readout rate: Even at 10 MHz framing, a 1B/pixel full-frame readout of, say, 10^5 pixels implies an off-detector data rate of 10^{12} B/s; this is beyond current capabilities. In practice, a small number of frames (~ 10) often suffices; hence, state-of-the-art burst-rate imagers store images in-pixel for later, much slower readout after the experiment is over. Going beyond 10 MHz framing will require new technology, e.g., faster in-pixel electronics and more intelligent readout schemes. In most high-frame rate situations the images consist of mostly null pixels. This is because even x-ray bursts delivered to a sample from an XFEL tend to top out at $\sim 10^{12}$ x-rays, only a small fraction of which are scattered over the many pixels of the detector. In-pixel electronics that reject null events from readout (“sparsification”) can greatly reduce the required detector readout rate, an approach typical of many high energy physics detectors.

2.2 Continuously framing, wide dynamic range imaging with single x-ray sensitivity

Another class of experiments requires very many frames in which the x-ray dose/pixel/frame may vary over 5 or more decades. Examples include high spatial resolution ptychography of extended objects (e.g., bone, integrated circuits, composites), dynamic SAXS, and simultaneous collection of Bragg and diffuse scatter from labile crystals. Often the weak parts of the image are photon starved, thereby necessitating single x-ray detection sensitivity, even as the low-Q diffraction receives many millions of x-rays/pixel/frame. Note that most such experiments rarely require measurement of

the x-ray dose per pixel to better than a few tenths per cent accuracy. This is fortunate because detector calibration errors and x-ray Poisson noise would otherwise impose limitations. (Users are often surprised to discover that practically all area detectors on SR beamlines are not calibrated to better than 0.5% accuracy.)

Most fast, continuously framing cameras use a charge-to-voltage converting amplifier to integrate the charges resulting from x-ray absorption in the sensor screen. This amplifier operates between fixed voltage limits, V_{SAT} , of typically less than a few volts. All amplifiers have some front-end noise, σ_{IN} . Robust single x-ray detection requires that the signal, S , from a single x-ray be such that $S \approx 5\sigma_{IN}$. For a linear amplifier, the dynamic range, D_R is then $< V_{SAT}/S$ x-rays; this is typically in the range of 10^2 – 10^4 x-rays. Larger values of D_R could be obtained with an amplifier with a nonlinear (e.g., logarithmic) response at the cost of complex calibration. A detector with a high-gain linear amplifier can have an extended dynamic range *via* implementation of electronics that remove fixed quantities of charge from the feedback integration capacitor as the signal accumulates. This dynamic integrator charge removal scheme is the basis for the Mixed-Mode Pixel Array Detector (MM-PAD) family [6–8]. Some examples of other detector efforts exploring high sensitivity, wide dynamic range realm include JUNGFRU [9], AGIPD [10], CITIUS [11, 12], and CoRDIA [13].

2.3 High atomic weight (hi-Z) sensor screens

Sensors providing high quantum efficiencies for very hard X-rays (>20 keV) are critically important to extend effective photon science beyond the reach of existing silicon sensors. The development of silicon x-ray sensors leveraged processes originally developed for the microelectronics industry with modifications and customizations required for the full silicon-thickness usage. However, high-quality, hi-Z sensors have not had this luxury. The technical challenges of hi-Z weight sensors are in several elements of sensor production, e.g., synthesis of crystals of sufficient size and quality, fabrication of pixelated sensors with the required pitch, and the *ad hoc* processes for integration and bonding to multipixel ASICs. Hi-Z sensors must have not only good energy resolution, temporal stability, homogeneity, carrier mobility (μ) and lifetimes (τ), low lag and dark current but also equally important must be readily available. Recently, the medical imaging industry has turned its attention towards “direct” detection computed tomography (CT) systems (e.g., photon counting CT and SPECT-CT imaging), resulting in the availability of new hi-Z weight sensors.

One of the first hi-Z weight sensor material to be readily available commercially for photon science applications is Cadmium Telluride (CdTe) [14]. Currently, CdTe offers a compromise between performance and availability while simultaneously providing some degree of radiation protection to underlying electronics. CdTe is arguably the predominate hard x-ray sensor used for commercial high energy x-ray pixel array detectors. However, CdTe sensors are prone to the buildup of excessive space-charge at relatively modest flux available at storage rings and XFELs. Cadmium Zinc Telluride (CZT) has been extensively used for low-flux spectroscopic applications. Recently, so called ‘high-flux-capable’ grade CZT from Redlen [15] appears to hold great promise, but the availability for photon science applications has been challenging. This variant of CZT has been designed with more nearly equal $\mu\tau$ -product for both holes and

electrons compared to its low flux equivalent [16]. The resulting material has shown stable performance at very high flux, even under conditions observed at XFELs [17]. Gallium arsenide (GaAs) sensors have also garnered attention for moderate hard x-ray energies. Unlike CdTe or CZT, GaAs lacks a troublesome absorption edge above 20 keV. A number of groups have examined the considerable promise of chromium-compensated gallium arsenide (GaAs:Cr), though obtaining material of sufficient and consistent quality, area and thickness has been an issue [18–20]. There are currently two sources providing GaAs:Cr for x-ray science detector applications [21, 22]. Besides availability, prominent problems with GaAs:Cr include lengthy charge collection times and nonlinearities [18, 23]. Finally, new classes of hi-Z materials such as perovskites (e.g., CsPbBr₃) are being studied for radiation detection [24, 25], but significant R&D will still be required.

The sensor materials discussed above are generally single crystal boules that are processed into thick (>0.5 mm) wafers. There is another class of sensor materials that are grown using thin film techniques (e.g., physical vapor deposition or molecular-beam epitaxy (MBE)) for moderate x-ray energies. Amorphous selenium (a-Se) sensor screens have been used by the medical imaging industry for static applications, such as mammography. Time-resolved, high-dynamic range imaging with a-Se sensors have been hindered by flux-dependent afterglow issues [26, 27], but there are applications in photon sciences requiring simultaneous high-energy and high-spatial resolution at low-flux levels [28]. CdTe [29] and GaAs [30] can also be grown using MBE deposition techniques; these materials show promise for ultra-fast applications, e.g., [31]. Solution-processed perovskite thin films are being studied [32].

Finally, the collection rate of quanta resulting from the absorption of x-rays is an important consideration for burst-mode imaging approaching GHz frame rates. In direct detection the thickness and carrier mobilities of the x-ray absorbing layer are key parameters for both detector quantum efficiency and temporal response. This is especially true for the hi-Z sensor screens required for very hard x-rays. Si sensor screens thick enough to efficiently absorb >20 keV x-rays require ~10s of ns to collect charge into the input node of the collection electronics [33, 34]. Hi-Z semiconductors with higher carrier mobilities than Si are known [35] but few of these are available in the quality or size needed for direct x-ray detection. This is yet another reason to continue to develop new hi-Z sensors.

2.4 Continuous on-line image analysis

The combination of increases in detector frame rates and pixel number, together with the evolution of multimodal and concurrent techniques, challenge existing capabilities in data acquisition, online fast feedback, and computing capacity. Currently, detectors and data reduction methods are not tightly integrated. In addition, the continued demand for faster detectors presents severe challenges of limited data bandwidth off the detector front-end and of the deluge of data that such new detectors will generate. The flood of data also limits the ability of scientists to extract actionable insights to steer experiments. It is estimated that the U.S. light sources will generate exabytes (EB) of data over the next decade, requiring tens to 1,000 PFLOPS of peak on-demand computing resources, and utilization of billions of core hours

per year [36]. Data loads of TB/s at LCLS-II, 1.3 EB per year at SPring-8 and similar loads at the European XFEL [36, 37] can only be managed by the implementation of strategic data reductions as close as possible to the signal generation point, i.e., edge processing. Light sources around the world are developing strategies for data reduction at different stages of the data flow. Several tradeoffs should be considered when choosing how early data reduction can be implemented effectively. Reconfigurable and flexible implementations can occur directly after the first front-end readout functions or further downstream of the detector [38, 39]. However, they require significant bandwidth and massive parallelism for data streaming. On the other hand, specialized hardware solutions offer the opportunity to send out only selected data with optimal information content. Dedicated readout architectures can be designed to optimize data quality and implement pre-processing, e.g., reconstruction of partial signals shared between pixels to improve either position or energy resolution while reducing the number of pixels to be recorded [40]. They can be specific to a typology of scientific data and require advanced technologies to achieve suitable processing density while maintaining reasonable power consumption.

A concern when implementing any type of reduction technique is if important information is lost that might alter the result of an experiment. Lossless data compression is an efficient and popular option that can be achieved by zero suppression at the detector readout level or at successive intermediate steps and offline (e.g., frame summing, Hcompress). The former requires dedicated readout architectures that work only for some categories of experiments. In some cases, selections of relevant regions of interest are also possible *via* relatively simple hardware and software options; this allows a focus on key information at the timescale important for the phenomena being investigated. Suppressing redundant information in the data stream may also be achieved with the acceptance of some losses. Lossy compression is very much experiment-specific and requires user involvement to evaluate quality and value of the different implementations.

The use of artificial intelligence and/or machine learning inspired data handling for on-line images analysis is becoming very popular. Pattern identification and recognition can be a powerful method to improve performance and efficiency. It allows identification of the class of problems and associated solutions, as has been studied for light source applications [41]. For example, online autonomous Bragg peak finders and analysis are much sought after and new tools utilizing deep neural networks are now becoming available [42, 43].

As mentioned earlier, on-detector electronics to reduce data flow is very promising and is being explored in various fields [44]. At the other end of the implementation space, predictive approaches are also becoming very popular and can help optimize data flows (e.g., compression ratio and speed), but typically require careful validation [45, 46].

3 Conclusion

Advances in pixel area x-ray detection will be needed to exploit capabilities provided by the rapid advance of x-ray sources. New technologies, materials, and concepts provide great opportunities for advancement in the next decade. In this Perspective we noted some of the challenges (i.e., frame rate, dynamic rate, high-energy sensors

and on-line data analysis) and point to the current state-of-the-art. There is much yet to be done.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SG: Conceptualization, Writing—original draft, Writing—review and editing. GC: Writing—original draft, Writing—review and editing. AM: Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Detector research by the authors is supported by the U.S. Department of

Energy, Office of Basic Energy Sciences. Detector research at Cornell is supported by U.S. Dept. of Energy Grant DE-SC0021026 and Argonne National Lab contract 2F-60279.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Graafsma H, Becker J, Gruner SM. Integrating hybrid area detectors for storage ring and free-electron laser applications. In: Jaeschke E, Khan S, Schneider JR, Hastings JB, editors. *Synchrotron light sources and free-electron lasers: accelerator physics, instrumentation and science applications*. Cham: Springer International Publishing (2018). p. 1–31.
- Looker Q, Colombo AP, Kimmel M, Porter JL. X-ray characterization of the Icarus ultrafast x-ray imager. *Rev Scientific Instr* (2020) 91(4):043502. doi:10.1063/1.50004711
- Hodge DS, Leong AFT, Pandolfi S, Kurzer-Ogul K, Montgomery DS, Aluie H, et al. Multi-frame, ultrafast, x-ray microscope for imaging shockwave dynamics. *Opt Express* (2022) 30(21):38405–22. doi:10.1364/oe.472275
- Mendez J. *Radiographic detector development for DARHT and ASD scorpius (U)*. Los Alamos national laboratory (LANL). Los Alamos, NM (United States): OSTI (2019). Report No.: LA-UR-19-23511. doi:10.2172/1508537
- Garafalo AM, Carpenter A, Troselle C, Funsten BT, Dean J, Benedetti L, et al. Characterization suite of a 1 ns, multi-frame hybridized CMOS imager for the ultra-fast x-ray imager program. In: *Hard X-ray, gamma-ray, and neutron detector physics XXV*. San Diego, CA: Optics and Photonics (2023).
- Tate MW, Chamberlain D, Green KS, Philipp HT, Purohit P, Strohmman C, et al. A medium-format, mixed-mode pixel array detector for kilohertz x-ray imaging. *J Physics: Conf Ser* (2013) 425:062004. doi:10.1088/1742-6596/425/6/062004
- Weiss JT, Shanks KS, Philipp HT, Becker J, Chamberlain D, Purohit P, et al. High dynamic range X-ray detector pixel architectures utilizing charge removal. *IEEE Trans Nucl Sci* (2017) 64(4):1101–7. doi:10.1109/tns.2017.2679540
- Gadkari D, Shanks K, Hu H, Philipp H, Tate M, Thom-Levy J, et al. Characterization of 128 × 128 MM-PAD-2.1 ASIC: a fast framing hard x-ray detector with high dynamic range. *JINST* (2022) 17(03):P03003. doi:10.1088/1748-0221/17/03/p03003
- Leonarski F, Mozzanica A, Brückner M, Lopez-Cuenca C, Redford S, Sala L, et al. JUNGFRAU detector for brighter x-ray sources: solutions for IT and data science challenges in macromolecular crystallography. *Struct Dyn* (2020) 7(1):014305. doi:10.1063/1.5143480
- Klačković I, Sztuk-Dambietz J, Graafsma H, Hosseini-Saber S, Klyuev A, Laurus T, et al. Five years operation experience with the AGIPD detectors at the European XFEL. *Proc SPIE, X-Ray Free-Electron Lasers. Adv Source Dev Instrumentation VI* (2023) 12581:142–50. doi:10.1117/12.2666402
- Grimes M, Pauwels K, Schulli TU, Martin T, Fajardo P, Douissard P-A, et al. Bragg coherent diffraction imaging with the CITIUS charge-integrating detector. *J Appl Crystallogr* (2023) 56(4):1032–7. doi:10.1107/s1600576723004314
- Takahashi Y, Abe M, Uematsu H, Takazawa S, Sasaki Y, Ishiguro N, et al. High-resolution and high-sensitivity X-ray ptychographic coherent diffraction imaging using the CITIUS detector. *J Synchrotron Radiat* (2023) 30(5):989–94. doi:10.1107/S1600577523004897
- Marras A, Kluev A, Lange S, Laurus T, Pennicard D, Trunk U, et al. Development of CoRDIA: an imaging detector for next-generation photon science X-ray sources. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1047:167814. doi:10.1016/j.nima.2022.167814
- Shiraki H, Funaki M, Ando Y, Kominami S, Amemiya K, Ohno R. Improvement of the productivity in the THM growth of CdTe single crystal as nuclear radiation detector. *IEEE Trans Nucl Sci* (2010) 57(1):395–9. doi:10.1109/tns.2009.2035316
- Iniewski K. CZT sensors for Computed Tomography: from crystal growth to image quality. *J Instrumentation* (2016) 11(12):C12034. doi:10.1088/1748-0221/11/12/c12034
- Thomas B, Veale M, Wilson M, Seller P, Schneider A, Iniewski K. Characterisation of redlen high-flux CdZnTe. *J Instrumentation* (2017) 12(12):C12045. doi:10.1088/1748-0221/12/12/c12045
- Veale MC, Angelsen C, Booker P, Coughlan J, French M, Hardie A, et al. Cadmium zinc telluride pixel detectors for high-intensity x-ray imaging at free electron lasers. *J Physics D: Appl Phys* (2018) 52(8):085106. doi:10.1088/1361-6463/aaf556
- Becker J, Tate MW, Shanks KS, Philipp HT, Weiss JT, Purohit P, et al. Characterization of chromium compensated GaAs as an x-ray sensor material for charge-integrating pixel array detectors. *JINST* (2018) 13:P01007. doi:10.1088/1748-0221/13/01/p01007
- Greiffenberg D, Andrä M, Barten R, Bergamaschi A, Brückner M, Busca P, et al. Characterization of chromium compensated GaAs sensors with the charge-integrating jungfrau readout chip by means of a highly collimated pencil beam. *Sensors* (2021) 21(4):1550. doi:10.3390/s21041550
- Fiederle M, Procz S, Hamann E, Fauler A, Fröjd C. Overview of GaAs und CdTe pixel detectors using Medipix electronics. *Cryst Res Tech* (2020) 55(9):2000021. doi:10.1002/crat.202000021
- Tyazhev A, Budnitsky D, Koretskay O, Novikov V, Okaevich L, Potapov A, et al. GaAs radiation imaging detectors with an active layer thickness up to 1 mm. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 509(1–3):34–9. doi:10.1016/s0168-9002(03)01545-6
- Kalliopuska J. *Chromium compensated gallium arsenide sensors evaluation using Timepix1, Timepix3, Medipix3 and Timepix2 readout electronics. 24th international workshop on radiation imaging detectors*. Norway: Oslo Science Park (2023). <https://indico.cern.ch/event/1247911/contributions/5394505/2023>.
- Greiffenberg D. *High-Z sensors for synchrotron sources and FELs ultrafast imaging and tracking instrumentation, methods and applications conference (ULITIMA 2023)*. Menlo Park, CA USA: SLAC (2023). <https://indico.slac.stanford.edu/event/7076/book-of-abstracts.pdf2023>.

24. Tsai H, Liu F, Shrestha S, Fernando K, Tretiak S, Scott B, et al. A sensitive and robust thin-film x-ray detector using 2D layered perovskite diodes. *Sci Adv* (2020) 6(15): eaay0815. doi:10.1126/sciadv.aay0815
25. Pan L, Pandey IR, Miceli A, Klepov VV, Chung DY, Kanatzidis MG. Perovskite CsPbBr₃ single-crystal detector operating at 1010 photons s⁻¹ mm⁻² for ultra-high flux X-ray detection. *Adv Opt Mater* (2023) 11(7):2202946. doi:10.1002/adom.202202946
26. Scott CC, Farrier M, Li Y, Laxer S, Ravi P, Kenesei P, et al. High-energy micrometre-scale pixel direct conversion X-ray detector. *J Synchrotron Radiat* (2021) 28(4):1081–9. doi:10.1107/s1600577521004835
27. Han Z, Mukherjee A, Albert A, Rumaiz AK, Harding I, Tate MW, et al. High spatial resolution direct conversion amorphous selenium X-ray detectors with monolithically integrated CMOS readout. *J Instrumentation* (2023) 18(04):P04021. doi:10.1088/1748-0221/18/04/p04021
28. Kiesel E, Poudyal I, Kenesei P, Engbretson M, Last A, Basak R, et al. Full-field nanoscale X-ray diffraction-contrast imaging using direct detection (2022). Available at: <https://arxiv.org/abs/2212.07303>.
29. Niraula M, Yasuda K, Watanabe A, Kai Y, Ichihashi H, Yamada W, et al. MOVPE growth of CdTe on Si substrates for gamma ray detector fabrication. *IEEE Trans Nucl Sci* (2009) 56(3):836–40. doi:10.1109/tns.2008.2010256
30. Looker Q, Wood MG, Lake PW, Kim JK, Serkland DK. GaAs x-ray detectors with sub-nanosecond temporal response. *Rev Scientific Instr* (2019) 90(11):113505. doi:10.1063/1.5127294
31. Looker Q, Wood MG, Miceli A, Niraula M, Yasuda K, Porter JL. Synchrotron characterization of high-Z, current-mode x-ray detectors. *Rev Scientific Instr* (2020) 91(2):023509. doi:10.1063/1.5139403
32. Tsai H, Shrestha S, Pan L, Huang HH, Strzalka J, Williams D, et al. Quasi-2D perovskite crystalline layers for printable direct conversion X-ray imaging. *Adv Mater* (2022) 34(13):2106498. doi:10.1002/adma.202106498
33. Looker Q, Colombo AP, Porter JL. Detector thickness effects on nanosecond-gated imager response. *Rev Scientific Instr* (2021) 92(5):053504. doi:10.1063/5.0048519
34. Parker S, Kok A, Kenney C, Jarron P, Hasi J, Despeisse M, et al. Increased speed: 3D silicon sensors; fast current amplifiers. *IEEE Trans Nucl Sci* (2011) 58(2):404–17. doi:10.1109/tns.2011.2105889
35. Owens A, Peacock A. Compound semiconductor radiation detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2004) 531(1–2):18–37. doi:10.1016/j.nima.2004.05.071
36. Schwarz N, Campbell S, Hexemer A, Mehta A, Thayer J. Enabling scientific discovery at next-generation light sources with advanced AI and HPC. In: *Driving scientific and engineering discoveries through the convergence of HPC*. In: *Big data and AI: 17th smoky mountains computational sciences and engineering conference, SMC 2020*. Oak Ridge, TN, USA: Springer (2020).
37. Ishikawa T Spring-8-ii conceptual design report 2014 (2014). Available at: <http://rsc.riken.jp/eng/pdf/SPRing-8-II.pdf>.
38. Kandel S, Zhou T, Babu AV, Di Z, Li X, Ma X, et al. Demonstration of an AI-driven workflow for autonomous high-resolution scanning microscopy (2023). Available at: <https://arxiv.org/abs/2301.05286>. doi:10.1038/s41467-023-40339-1
39. Babu AV, Zhou T, Kandel S, Bicer T, Liu Z, Judge W, et al. Deep learning at the edge enables real-time streaming ptychographic imaging (2022). Available at: <https://arxiv.org/abs/2209.09408>. doi:10.1038/s41467-023-41496-z
40. Otfinowski P, Deptuch GW, Maj P. Asynchronous approximation of a center of gravity for pixel detectors' readout circuits. *IEEE J Solid-State Circuits* (2018) 53(5):1550–8. doi:10.1109/jssc.2018.2793530
41. Vescovi R, Chard R, Saint ND, Blaiszik B, Pruyne J, Bicer T, et al. Linking scientific instruments and computation: patterns, technologies, and experiences. *Patterns* (2022) 3(10):100606. doi:10.1016/j.patter.2022.100606
42. Wang C, Li P-N, Thayer J, Yoon CH. PeakNet: Bragg peak finding in X-ray crystallography experiments with U-Net (2023). Available at: <https://arxiv.org/abs/230315301>.
43. Liu Z, Sharma H, Park J-S, Kenesei P, Miceli A, Almer J, et al. BraggNN: fast X-ray Bragg peak analysis using deep learning. *IUCrJ* (2022) 9(1):104–13. doi:10.1107/s2052252521011258
44. Carini G, Deptuch G, Dickinson J, Doering D, Dragone A, Fahim F, et al. Smart sensors using artificial intelligence for on-detector electronics and ASICs (2022). Available at: <https://arxiv.org/abs/2204.13223>.
45. Lotter W, Kreiman G, Cox D. Deep predictive coding networks for video prediction and unsupervised learning (2016). Available at: <https://arxiv.org/abs/1605.08104>.
46. Roy R, Sato K, Bhattacharya S, Fang X, Joti Y, Hatsui T, et al. Compression of time evolutionary image data through predictive deep neural networks. In: *Proceedings of the 2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*; May 2021; Melbourne, Australia (2021). p. 41–50.



OPEN ACCESS

EDITED BY

Iain Sedgwick,
Rutherford Appleton Laboratory,
United Kingdom

REVIEWED BY

Carl Grace,
Berkeley Lab (DOE), United States
Vandana Sharma,
Indian Institute of Technology
Hyderabad, India

*CORRESPONDENCE

Marcin Sikorski,
✉ marcin.sikorski@xfel.eu
Marco Ramilli,
✉ marco.ramilli@xfel.eu

[†]These authors have contributed equally
to this work and share first authorship

RECEIVED 27 September 2023

ACCEPTED 13 November 2023

PUBLISHED 29 November 2023

CITATION

Sikorski M, Ramilli M, de Wijn R, Hinger V,
Mozzanica A, Schmitt B, Han H, Bean R,
Bielecki J, Bortel G, Dietze T, Faigel G,
Kharitonov K, Kim C, Koliyadu JCP,
Koua FHM, Letrun R, Lopez LM,
Reimers N, Round A, Sarma A, Sato T,
Tegze M and Turcato M (2023), First
operation of the JUNGFRAU detector in
16-memory cell mode at European XFEL.
Front. Phys. 11:1303247.
doi: 10.3389/fphy.2023.1303247

COPYRIGHT

© 2023 Sikorski, Ramilli, de Wijn, Hinger,
Mozzanica, Schmitt, Han, Bean, Bielecki,
Bortel, Dietze, Faigel, Kharitonov, Kim,
Koliyadu, Koua, Letrun, Lopez, Reimers,
Round, Sarma, Sato, Tegze and Turcato.
This is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

First operation of the JUNGFRAU detector in 16-memory cell mode at European XFEL

Marcin Sikorski^{1*†}, Marco Ramilli^{1*†}, Raphael de Wijn¹,
Viktoria Hinger², Aldo Mozzanica², Bernd Schmitt², Huijong Han¹,
Richard Bean¹, Johan Bielecki¹, Gábor Bortel³, Thomas Dietze¹,
Gyula Faigel³, Konstantin Kharitonov¹, Chan Kim¹,
Jayanath C. P. Koliyadu¹, Faisal H. M. Koua¹, Romain Letrun¹,
Luis M. Lopez¹, Nadja Reimers¹, Adam Round¹, Abhisakh Sarma¹,
Tokushi Sato¹, Miklós Tegze³ and Monica Turcato¹

¹European XFEL, Schenefeld, Germany, ²Paul Scherrer Institut, Villigen, Switzerland, ³Wigner Research Centre for Physics, Institute for Solid State Physics and Optics, Budapest, Hungary

The JUNGFRAU detector is a well-established hybrid pixel detector developed at the Paul Scherrer Institut (PSI) designed for free-electron laser (FEL) applications. JUNGFRAU features a charge-integrating dynamic gain switching architecture, with three different gain stages and 75 μm pixel pitch. It is widely used at the European X-ray Free-Electron Laser (EuXFEL), a facility which produces high brilliance X-ray pulses at MHz repetition rate in the form of bursts repeating at 10 Hz. In nominal configuration, the detector utilizes only a single memory cell and supports data acquisition up to 2 kHz. This constrains the operation of the detector to a 10 Hz frame rate when combined with the pulsed train structure of the EuXFEL. When configured in so-called *burst* mode, the JUNGFRAU detector can acquire a series of images into sixteen memory cells at a maximum rate of around 150 kHz. This acquisition scheme is better suited for the time structure of the X-rays as well as the pump laser pulses at the EuXFEL. To ensure confidence in the use of the *burst* mode at EuXFEL, a wide range of measurements have been performed to characterize the detector, especially to validate the detector alibration procedures. In particular, by analyzing the detector response to varying photon intensity (so called 'intensity scan'), special attention was given to the characterization of the transitions between gain stages. The detector was operated in both dynamic gain switching and fixed gain modes. Results of these measurements indicate difficulties in the characterization of the detector dynamic gain switching response while operated in *burst* mode, while no major issues have been found with fixed gain operation. Based on this outcome, fixed gain operation mode with all the memory cells was used during two experiments at EuXFEL, namely in serial femtosecond protein crystallography and Kossel lines measurements. The positive outcome of these two experiments validates the good results previously obtained, and opens the possibility for a wider usage of the detector in *burst* operation mode, although compromises are needed on the dynamic range.

KEYWORDS

hybrid detector, EuXFEL, adaptive gain, characterization, calibration

1 Introduction

The European X-ray Free-Electron Laser (EuXFEL) [4] generates high brilliance X-ray pulses at megahertz repetition rate. The pulses are grouped into bursts, also known as “trains,” occurring at a rate of 10 Hz. Within each train, pulses are generated at an intra-burst frequency reaching up to 4.5 MHz. The duration of a burst is approximately 0.6 ms, followed by a 99.4 ms gap period between subsequent pulse trains. This unique time structure opens novel scientific opportunities but also poses technological and engineering challenges, particularly in the domain of X-ray detection. Three types of large, mega-pixel imaging detectors, namely, AGIPD [5], LPD [6], and DSSC [7] were developed explicitly for the EuXFEL to enable pulse-resolved measurements at megahertz repetition rate. They are at present successfully used in various types of scientific experiments and are critical components of the scientific instruments at the facility. The detectors’ capability to capture images at a repetition rate of 4.5 MHz also entails a series of constraints, such as a relatively large pixel size (ranging from 200 to 500 μm), elaborate calibration procedures [6, 8, 9], large physical dimensions, and extensive support infrastructure. Therefore, the MHz detectors are frequently complemented by smaller pixelated detectors, which offer smaller pixel size, lower noise, compact/modular design, and well established calibration procedures. Among those, the JUNGFRÄU detector [1–3] is currently the most widely deployed at EuXFEL. The pixel size of 75 μm combined with a relatively low noise (~ 80 ENC) and a dynamic range of 10^4 12 keV photons [3] make the JUNGFRÄU well suited for a wide range of scientific applications, from protein crystallography [10] to spectroscopy [11].

As the JUNGFRÄU was developed primarily for SwissFEL [12, 13] and synchrotron applications, its compatibility with EuXFEL’s bunch structure is not optimal. The JUNGFRÄU detector has so far been mostly operated at 10 Hz at EuXFEL across a number of instruments. In many cases, the signal is integrated over multiple pulses per train. Therefore, the benefit of the sub-microsecond temporal resolution offered by the MHz intra-bunch repetition rate is lost. However, due to the high signal-to-noise ratio of the detector at hard X-ray photon energies (ranging between 6 keV and 20 keV), longer integration times do not significantly compromise the performance.

Due to the detector’s architecture (see Section 2.1.2), the 15 additional memory cells open the possibility of a more effective exploitation of the pulse-train structure. However, the relatively large RC time characteristic of the pixel readout architecture limits the frame rate in *burst* mode to a maximum of ~ 150 kHz; nevertheless its implementation remains an attractive option. For example, operating the detector in *burst* mode in pulse-resolved measurements such as protein crystallography will increase data throughput 16-fold. In addition, for experiments using low-viscosity buffer media it will offer an attractive compromise between detector performance and sample consumption. For many other types of experiments, it will improve time resolution while maintaining all the advantages of the JUNGFRÄU detector.

The 16-memory-cell operation of the detector, however, requires complete characterisation and calibration procedures to be established and validated before it can be provided for user operation.

In this work, the performance of the JUNGFRÄU detector operating in *burst* (16-cell) mode was studied for various detector configurations and illumination levels. Special attention was given to the

transition regions between the high-, medium-, and low-gain stages. Serial femtosecond crystallography data were collected using lysozyme as a model system to enable statistical comparisons, operating the detector in fixed medium gain. Additionally, single shot Kossel line measurements were used to study the fast dynamics of single crystal materials. The quality of the data collected in *burst* mode was validated against literature data.

2 Materials and methods

2.1 The JUNGFRÄU detector

JUNGFRÄU is a hybrid pixel detector designed and produced at Paul Scherrer Institut (PSI), in Villigen (CH), consisting of pixelated Application-Specific Integrated Circuits (ASICs) bump-bonded to a semiconductor sensor (current versions use Si of 320 μm or 450 μm thickness; high-Z materials such as GaAs and CdZnTe are under investigation).

Each ASIC is segmented into pixels of 75 μm pitch, each featuring a charge-integrating Dynamic Gain Switching (DGS) architecture, with a dynamic range on the order of 110 dB. Each pixel has an array of 16 memory cells that can store the collected analog signal before the readout. A matrix of 256×256 pixels comprises a single ASIC. An array of 4×2 ASICs is bump-bonded to a single monolithic silicon sensor, constituting a JUNGFRÄU Front-End Module (FEM) of 1,024 columns and 512 rows, for a total of about 0.5 megapixels.

Each ASIC is divided into supercolumns of 64 columns \times 256 rows, whose pixel output signal is multiplexed to an individual off-chip Analog-to-Digital Converter (ADC) for digitization; the whole module is therefore read out by a total of 32 different ADCs.

The FEM is operated via a dedicated electronic component, the JUNGFRÄU Master Control Board (MCB), equipped with a 40 MHz Field Programmable Gate Array (FPGA) actively controlling the ASICs, and a small processor with embedded Linux Operating System (OS) running a server allowing slow control of the module.

A third electronic component is necessary to deliver the external Transistor-Transistor Logic (TTL) signal trigger to the JUNGFRÄU MCB.

Each JUNGFRÄU module comprises a FEM coupled with a JUNGFRÄU MCB and operates individually, running on the clock of its own FPGA; however, more modules can be simultaneously controlled and configured, and their acquisition can be synchronized with a common external trigger, allowing the operation of multi-module configurations.

2.1.1 Dynamic gain switching (DGS)

In order to comply with the dynamic range requirements of a FEL, the pixel architecture of the JUNGFRÄU detector is designed with a DGS mechanism: a pixel-wise threshold comparator switches additional capacitors into the pre-amplifier feedback loop if the signal rises above a certain value (set module-wise), thus increasing the feedback capacitance and reducing the gain. In the JUNGFRÄU pixels, the Correlated Double Sampling (CDS) stage (see Figure 1) is bypassed after gain switching, hence causing a sign inversion of the signal curve slope.

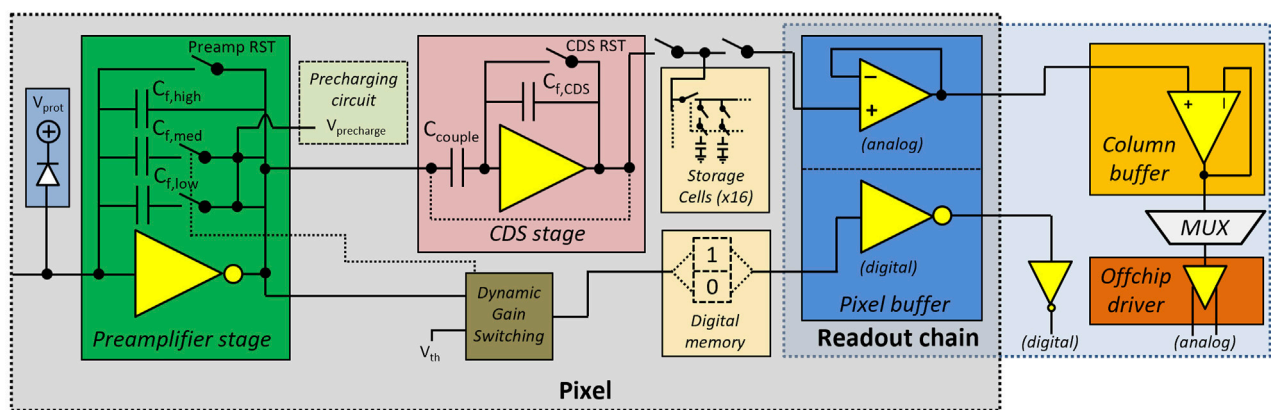


FIGURE 1
Schematics of the JUNGFR AU pixel architecture.

There are a total of three gains (respectively G_0 , G_1 , and G_2) which can be independently triggered in each pixel, with relative amplification ratios:

$$R_1 = \frac{G_0}{G_1} \sim 30 \quad (1)$$

$$R_2 = \frac{G_1}{G_2} \sim 20. \quad (2)$$

During acquisition the integrated charge is temporarily stored in dedicated capacitors (memory cells) and, during the readout phase, the signal stored in each memory cell is multiplexed image-wise to one of the 32 14-bit ADCs for digitization. The state of the DGS mechanism is digitally recorded for each memory cell by changing the value of two so-called *gain bits* (00, 01, and 11 for G_0 , G_1 , and G_2 , respectively): this value is then appended to the ADC output.

Additionally, there is the possibility of turning the DGS mechanism off, by fixing the value of the pre-amplifier feedback capacitance to the one corresponding to G_1 or G_2 . This operation mode is referred to as *fixed gain mode*. When operating in fixed gain mode, the CDS stage is not bypassed, therefore the gain conversion factor for $G_{1,fix}$ and $G_{2,fix}$ are higher with respect to their DGS counterparts.

Since the detector was primarily designed for SwissFEL and synchrotron applications [1, 13], the pixel architecture has not been optimized for MHz-class frame-rate acquisition. Its most widely employed and well-characterized operation mode is *single cell* operation, where only one of the 16 available memory cells is utilized: in this operation mode, a continuous frame rate of ~ 2 kHz can be achieved. This is sufficient to allow the detector to acquire at least one image per EuXFEL train and to produce useful scientific data [14–18]; however, the possibility of better exploiting the EuXFEL pulse train structure is attractive for many experiments.

2.1.2 Sixteen memory cell operation

Due to the 15 additional memory cells available in each pixel, it is possible to acquire up to 16 images in rapid succession and to read them out at the end of each measurement cycle. This is the so-called *burst* mode of operation, where pixels are operated differently from the *single cell* mode:

- In single cell mode, one memory cell is continuously connected to the memory cell bus. However, in *burst* mode the capacitors are sequentially connected during the acquisition phase and again during the readout;
- The reset of the pre-amplifier and the CDS is performed in the short time between acquisitions instead of during the long readout between trains.

The default memory cell used in single cell mode (cell 15) is the first cell in the *burst*, which is then followed by the fifteen additional cells. Both the exposure time and the dead time between two consecutive exposures can be adjusted to experimental needs. Frequently, in a similar way to single cell operation, the signal from multiple X-ray pulses is detected within a single memory cell. To allow sufficient time for the signal to settle after the reset perturbation, photons should arrive around $\sim 1 \mu\text{s}$ after the opening of the integration gate, while, to avoid late gain switching effects, they should not arrive later than $\sim 2 \mu\text{s}$ before the end of the integration gate. The shortest dead time between two exposures is $2.1 \mu\text{s}$, dictated by the time needed to reset the pre-amplifier and CDS stage. Therefore, the maximum operational acquisition rate of the JUNGFR AU in *burst* mode is around 150 kHz.

2.1.2.1 Calibration of the raw data

In principle, determination of the calibration constants in *burst* mode should follow the same procedure as in single cell mode [19]. Below, we shortly outline the calibration strategy applied to data collected in *burst* mode.

2.1.2.1.1 Pedestal evaluation. The detector is equipped with two special operation modes, called *forceswitchg1* and *forceswitchg2*, which are used to estimate the pedestal value for the G_1 and G_2 gain stages, respectively, by forcing the pre-amplifier to switch gain. Although this procedure is used routinely for single cell operation, validation is required for *burst* mode.

Preliminary tests of the forced gain switch of the detector operated in *burst* mode, showed the presence of artefacts in the resulting pedestal values due to the very high peak current

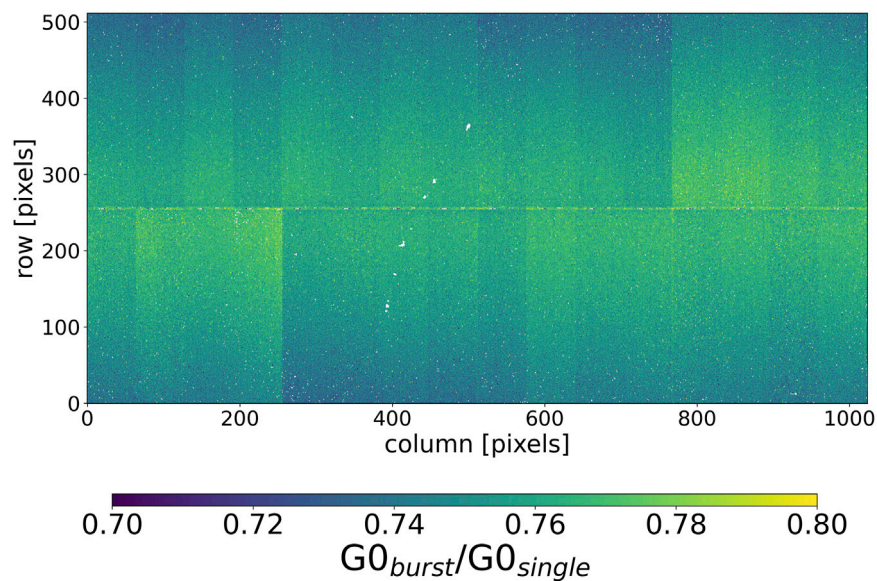


FIGURE 2
Ratio of G0 gains measured in *burst* and single memory cell configuration.

consumption during high repetition forced gain switching. Since the first memory cell connected to the bus was not yet affected, a pedestal measurement procedure has been outlined:

1. G0 pedestal values are measured by acquiring dark images with all the memory cells connected to the bus, and then calculating the average for each cell;
2. For G1 and G2 the forced gain switch is used, but only data from the *first* memory cell connected to the bus is used to calculate the pedestal value; hence, to acquire valid data for all cells, the forced switch gain measurement is repeated 16 times, each time connecting a different memory cell as first.

One of the main goals of the present study is indeed to validate this procedure.

2.1.2.1.2 Gain evaluation. Gain conversion maps for *burst mode* must be re-evaluated with respect to the single-cell mode for two main reasons:

1. The memory cells with index from 0 to 14 have, by design, a smaller capacitance than memory cell 15 (the default cell), due to space limitations in the pixel layout;
2. Memory cells are connected to the memory cell bus in turn instead of having a constant connection. This implies that the signal is smaller in *burst* mode than in single-cell mode. In fact, in the latter case the voltage written on the $C_{\text{memorycell}} + C_{\text{bus}}$ capacitance is directly presented to the pixel output voltage buffer, while in the former case the charge on the $C_{\text{memorycell}}$ is redistributed on the $C_{\text{memorycell}} + C_{\text{bus}}$, thus lowering the voltage at the pixel buffer input. This effect is illustrated in [Figure 2](#).

In principle, the full calibration procedure outlined in [19] should not need to be repeated for all memory cells in all gain

stages, as the gain ratios R_1 and R_2 do not depend on the individual memory cell, but only on the pre-amplifier (see [Figure 1](#)). Therefore, after having measured the gain G0 in *burst mode* for each memory cell, the G1 and G2 values are calculated using the gain ratios previously measured [19].

2.2 Dynamic range scan with EuXFEL beam

To validate the calibration strategies outlined in [Sections 2.1.2.1.1](#) and [2.1.2.1.2](#), a JUNGFRÄU detector was illuminated with X-ray pulses of varying intensity (so called ‘intensity scan’) to probe the detector response across at least the first two gain stages, a measurement similar to those performed in [20]. Ideally, the detector output (corrected for pedestal and gain for each memory cell) should have a linear response as a function of the impinging illumination, in particular:

- There should be no artifacts around the ‘gain switching region’ (GSR), i.e., the incoming photon flux range at which enough signal is generated in order to cause the pre-amplifier to switch gain;
- If gain calibration is correct, the slope of the corrected response curve should be the same for each gain stage.

2.2.1 Experimental setup

The measurements described in this document were carried out using the 4 megapixel JUNGFRÄU detector (JF4M) installed at the downstream interaction region (IRD) of the SPB/SFX scientific instrument [21] at atmospheric pressure. The IRD is mainly used for serial femtosecond crystallography (SFX) experiments, therefore the detector and infrastructure design are optimised for these demands. The JF4M detector is shown in [Figure 3A](#). It consists of eight 0.5 megapixel JUNGFRÄU units arranged in two columns.

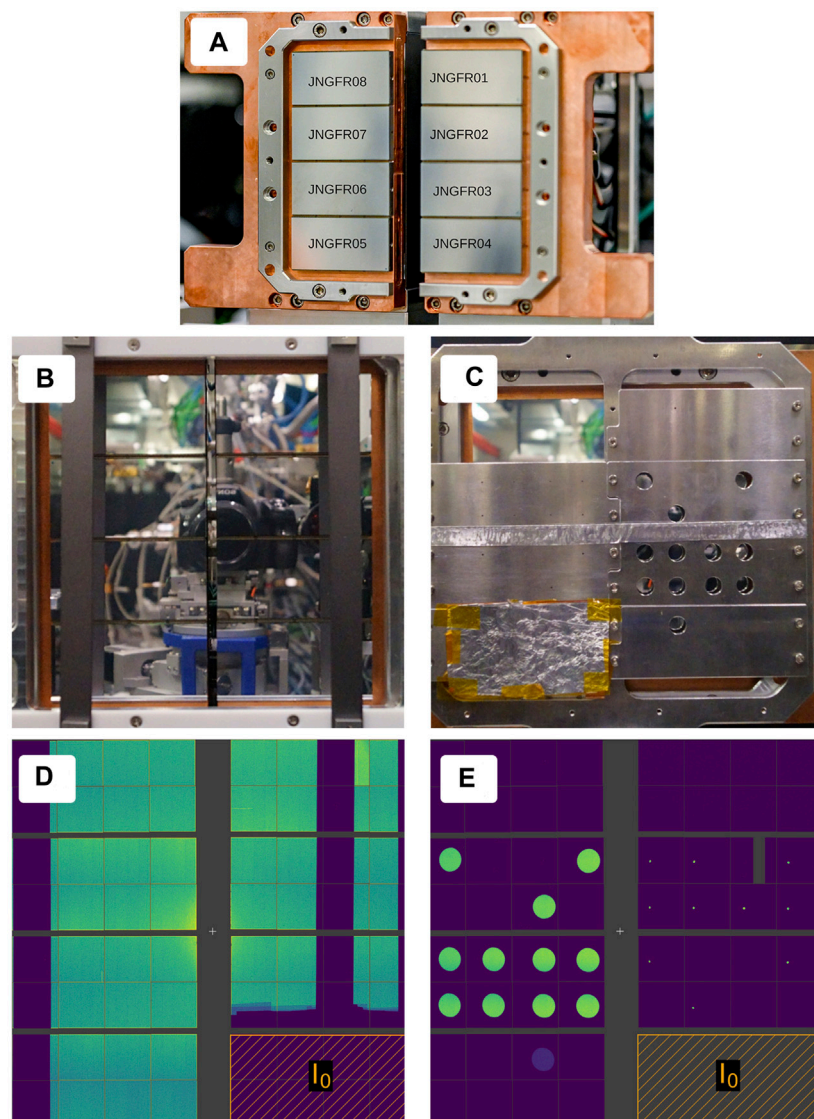


FIGURE 3

(A) The four megapixel JUNGFR01 to JUNGFR08 detector modules at the SPB/SFX instrument at EuXFEL. For reference, the individual modules were labelled as JUNGFR01 to JUNGFR08. (B,C) Two types of masks used to characterize the *burst* mode. Position of the vertical tantalum bars and configuration of the aluminium panels were adjusted between various measurement. (D,E): Example images acquired with both masks. Modules designated as the incident intensity monitors are marked and labelled as “ I_0 .”

For reference, individual modules are labelled JUNGFR01 to JUNGFR08 (Figure 3A). At the beginning of each pulse train, the eight modules receive the same trigger signal to synchronize the start of their image acquisition to the EuXFEL light delivery. Typically, the beam not diffracted by the sample passes through the gap between the columns and interacts with the beam diagnostics downstream of the detector. Both size and position of the gap with respect to the primary X-ray beam are optimized by adjusting the horizontal position of each detector half independently.

During a typical SFX measurement, crystals are suspended in a dedicated delivery buffer and injected into the interaction point in the form of a jet. The interaction with the focused beam results in the explosion of the jet. To protect the detector from debris created

during such explosions, kapton shields are installed in front of the sensor. The position of the shield can be adjusted in both horizontal and vertical directions. During the measurements reported here, the shield was replaced by various masks to be able to select the size of the exposed areas, referred to in the following sections as the *occupancy*, as well as the level of illumination of individual modules. For the measurements requiring flat-field-like illumination from copper fluorescence, the mask shown in Figure 3B was used. It consists of two vertical tantalum bars, each 19 mm wide and 2 mm thick, which shadowed a large fraction of each module. A typical image collected in such a configuration is presented in Figure 3D. Pixels behind the bars were used, mostly, to determine the baseline-shift corrections. The JUNGFR04 served as an incident intensity monitor. A stack of

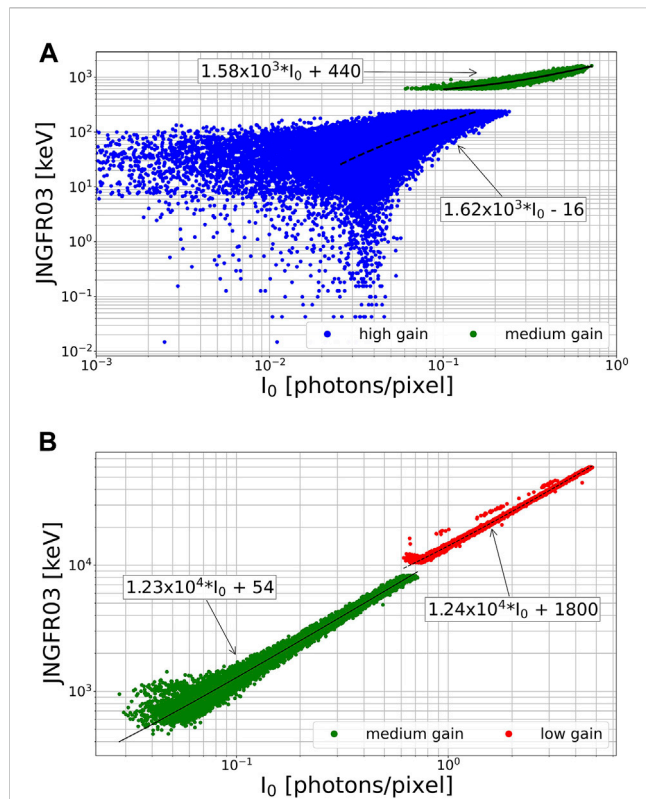


FIGURE 4

Examples of the intensity scans across (A) high to medium and (B) medium to low transition regions measured in the first image of the burst. The response of the pixel in different gain stages have each been fitted with linear functions, and fit results are shown in the plots. In order to access both transition regions, the scans were performed in two different experimental arrangements. Therefore values of the normalized I_0 in the two plots should not be compared directly.

aluminum foils (not shown in the figure) was placed in front of it to adjust the average intensity of photons per pixel to the desired levels. The mask shown in Figure 3C was used to study the correlation between the occupancy levels and performance of the JUNGFR03

modules, especially across the transition regions. It consists of two sets of panels with 1 mm and 8 mm holes. The position of the holes with respect to ASICs was optimized by adjusting the position of each detector half as well as the position of the mask itself. An example image acquired with this mask is shown in Figure 3E.

In order to measure, in relative terms, the intensity of the fluorescence signal on the detector, one of the eight modules was used as an intensity (I_0) monitor for each measurement. The module was shielded with an aluminum foil so that the intensity of the transmitted fluorescence photons reaching the module itself would always be at a sufficiently low intensity to allow the module not to switch gain but to remain in the independently calibrated G0 region; the I_0 value is then presented as the average number of photons per pixel behind the foil. Consequently, the thickness of the aluminum absorber has been changed to keep the overall maximum signal on the I_0 module at the same level, for each individual measurement. So it is important to keep in mind that the I_0 values from measurements obtained in severely different illumination conditions cannot be compared directly.

The data presented in this report was collected during multiple experiments. The experimental setup during each session was adjusted to accommodate the goals of the particular measurements as well as the current configuration of the SPB/SFX instrument but the core set of parameters was preserved across all the activities. All data were acquired in the transmission geometry using the fluorescence from a 5 μm -thick copper foil. The size of the 9.3 keV beam at the interaction point was set to approximately $400 \times 400 \mu\text{m}^2$. The incident flux was tuned with the help of diamond filters, providing up to 64 different attenuation levels with the smallest step of 7%. Contamination from the 9.3 keV air scattering at low angles was reduced by placing a 15 μm -thick nickel foil downstream of the interaction point. The sample-to-detector distance was adjusted for each type of measurement between 120 and 500 mm to find the best compromise between the flatness of the signal across the entire detector and the maximum number of photons per pixel. For larger distances, an additional He-filled pipe was installed to further suppress the air scattering.

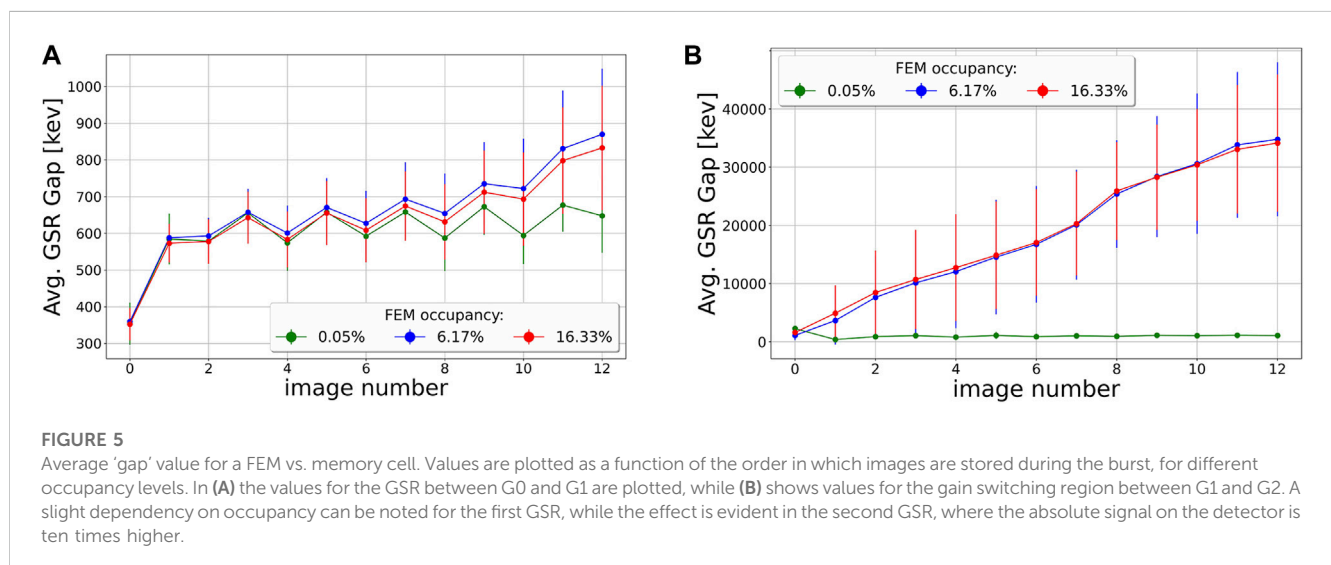


FIGURE 5

Average 'gap' value for a FEM vs. memory cell. Values are plotted as a function of the order in which images are stored during the burst, for different occupancy levels. In (A) the values for the GSR between G0 and G1 are plotted, while (B) shows values for the gain switching region between G1 and G2. A slight dependency on occupancy can be noted for the first GSR, while the effect is evident in the second GSR, where the absolute signal on the detector is ten times higher.

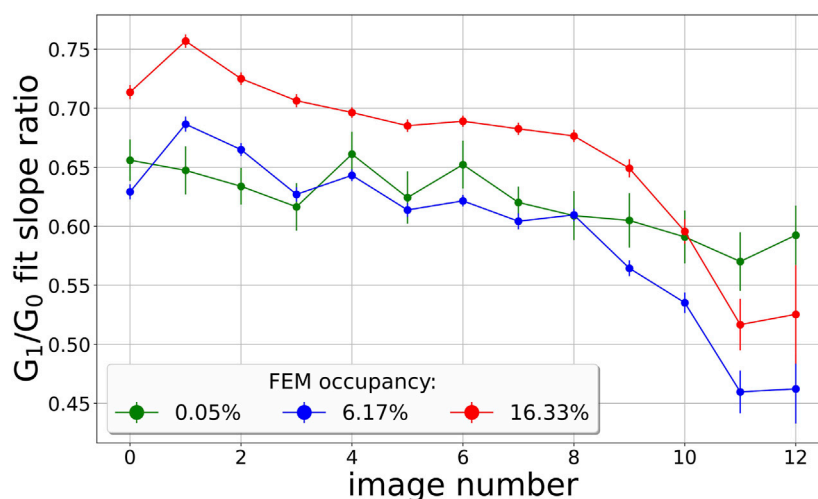


FIGURE 6

The average value for a FEM for the ratio between the slope of the G1 and the G0 part of the intensity scan is displayed. The error bars indicate the RMS of the distribution of values across the pixels.

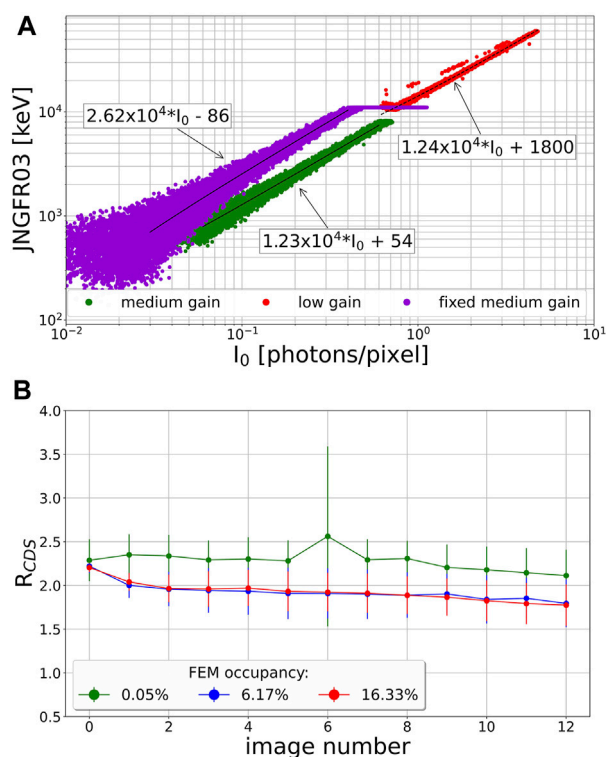


FIGURE 7

In (A) the output of an intensity scan in fixed medium gain (fixG1) is compared with a scan taken in the dynamic gain switching (DGS). Both scans were carried out under the same experimental conditions. The difference in slope of the lines is due to the intercalibration factor R_{CDS} . Impact of the occupancy level on the value of R_{CDS} is shown in (B).

Prior to the intensity scans, the detector G0 stage was calibrated using Cu K_{α} fluorescence photons at an intensity low enough so that individual photons could be resolved; the gain conversion factors for

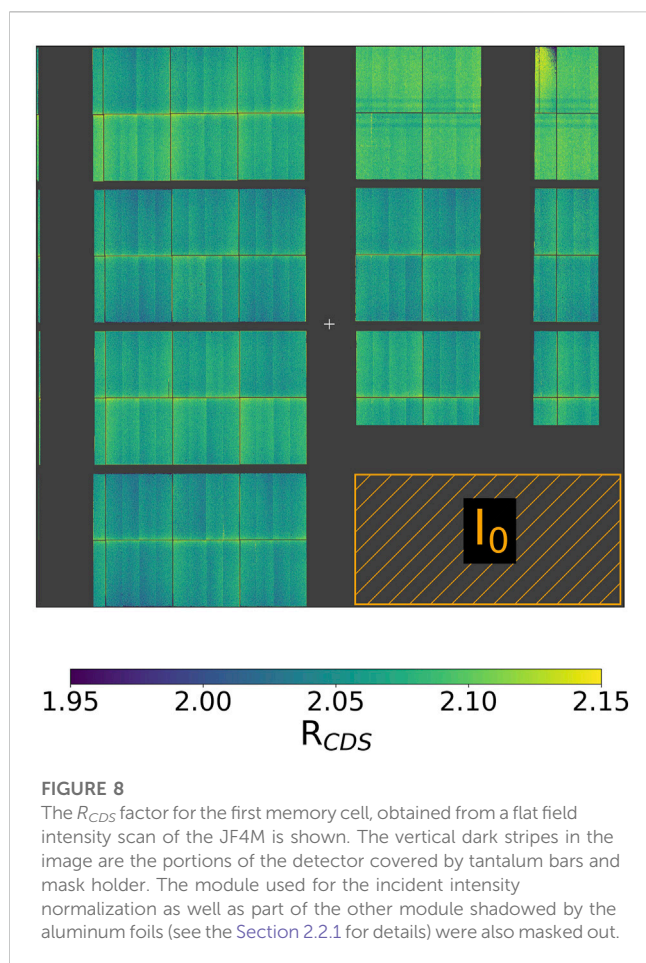
the G1 and G2 stages with respect to G0 were calculated for each pixel and memory cell, with a procedure similar to what is described in [19].

2.3 Experimental tests in *burst* mode

Based on the promising results obtained from the *burst* mode characterization, the performance of the detector was validated against data collected in single cell mode. Two experiments were conducted, each utilizing the detector at very different photon intensity levels and signal distribution across the sensors. For reasons outlined in Section 3, the operation of the detector in *burst* mode was restricted to the fixed G1 stage. The lack of absolute gain calibration for this detector setting did not pose a major problem for these measurements.

2.3.1 Serial femtosecond crystallography

In this section, we describe the SFX test experiment using the JF4M detector in *burst* mode with the gain set to fixed medium gain. One of the calibration samples used routinely at the SPB/SFX instrument was chosen; crystals of the hen egg-white lysozyme (HEWL) with a diameter of 2–3 μm . Crystals were grown using the well-established protocol [22] and injected into the interaction point beam as a liquid jet using a Gas Dynamic Virtual Nozzle (GDVN) type B, as described in [23], with a sample flow rate of 70 $\mu\text{L}/\text{min}$. To minimize the air scattering background, the sample chamber enclosing the interaction point was continuously flushed with helium. The crystals were exposed to a 12.55 keV X-ray beam focused to a 3 μm spot. The detector was positioned 118.3 mm downstream from the interaction point. This configuration corresponded to the maximum resolution of 1.6/1.5 \AA at the edges/corner of the detector. Both photon energy and sample-to-detector distance were refined from the SFX data based on expected unit cell parameters of the HEWL crystals.



2.3.2 Kossel line measurement

When fluorescent radiation emitted by the atoms in a single crystal is scattered by the crystal itself, intensity modulations, so-called Kossel lines, can be observed in the diffraction intensity [24]. As the phase of the structure factor is encoded in the profiles of those lines, and a single image contains multiple lines, complex insight into the structure of the illuminated crystal can be gained from the analysis of a single diffraction pattern. Measurement of the Kossel lines can be a way to overcome constraints faced by techniques relying on the accumulation of thousands of images such as serial femtosecond crystallography [22] or single particle imaging [25]. Access to such rich information about the structural changes within a sample from a single pulse is critical for pump-probe-type experiments, especially when the given pumped state of the sample cannot be easily repeated thousands of times. The MHz repetition rate of the EuXFEL enables the tracing of changes in sample structure with sub- μ s resolution. To support this type of measurement, the detector needs to meet stringent conditions. First, it has to allow measurements of the profiles of the Kossel lines with sufficient spatial resolution on top of the fluorescent background. Second, it has to be fast enough to take advantage of the MHz intra-bunch repetition rate. The JUNGFR AU detector, operated in *burst* mode, has suitable specifications for this use case, at least within the 150 kHz limit highlighted before.

Proof-of-principle single-pulse Kossel line measurements were recently conducted at the SPB/SFX instrument using the JF4M detector. Here, the feasibility of the *burst* mode to measure

Kossel lines for one demonstration sample, a 100 μ m thick Ge crystal, is reported. The single crystal Ge wafer was illuminated by a series of 16 pulse bursts of 11.5 keV beam focused down to 25 μ m \times 25 μ m. For each burst, a fresh spot of the wafer was exposed. The X-ray beam energy at the sample position was adjusted using diamond foils to several μ J, which was sufficient to observe the Kossel lines in several images before the radiation damage-induced signal dominated in the recorded patterns. The alignment of the wafer and the quality of its surface were monitored with the help of a microscope, inline with the X-ray illumination. Data were acquired in transmission geometry with the detector positioned 120 mm from the interaction point. To reduce air scattering contamination, the sample was encapsulated in a He-flushed chamber.

Based on measurements taken in the single memory cell configuration, for which the absolute calibration was available, the number of 9.25 keV fluorescent photons per pixel per pulse ranged from 20/120 to 200/800 at the edges/center of the detector. The exposure time was set to 20.12 μ s to provide the desired time resolution for this measurement.

3 Results

3.1 Intensity scan results

Various intensity scans were conducted in *burst* mode in both DGS as well as in fixed gain. Results are summarized in the following sections.

3.1.1 Gain switching region gap

It became immediately evident that the pedestal evaluation via forced switching of the feedback capacitor of the pre-amplifier shows substantial limitations in accuracy. This manifests itself as a positive ‘gap’ in the GSR. It indicates that this method underestimates the actual pedestal value after gain switching occurs. An example of the intensity scan for a pixel of JNGFR03 module is shown in Figure 4.

The value of this ‘gap’ can be estimated by calculating the difference between the lowest value in G1 and the highest in G0: the average ‘gap’ value per memory cell and per module shows a clear incremental trend with the memory cell filling order. However, when looking for dependencies of this value on the total occupancy of the FEM, a minimal effect has been noted for the first GSR (see Figure 5A), while the effect is evident when inspecting the GSR between G1 and G2 (Figure 5B), where a strong dependency is evident, on both occupancy and filling order of the memory cell.

3.1.2 Gain calibration

The results concerning the validation of the gain calibration are presented in the subsequent paragraphs, for the DGS and the fixed gain operation separately.

3.1.2.1 Lower gains in dynamic gain switching

The corrected output of the intensity scan as a function of I_0 has been fitted with linear functions for each pixel, cell and gain stage. An example can be seen in Figure 4. As mentioned in Section 2.1.2.1.2, the ratio of the slopes of the G1 and the G0 parts of the intensity scan, $G1_{DGS}$ and $G0_{DGS}$, should be equal to 1. Instead, we

TABLE 1 SFX data and refinement statistics.

Parameter	Value
Photon energy (eV)	12,550
X-ray focus, FWHM (μm)	3
Electron bunch length (fs)	24
Frame count	97,128
Indexed crystals	18,420
Indexing rate (%)	18.96
Space group	P4 ₃ 2 ₁ 2
Point group	4/mm
Cell dimensions (Å)	
a	79.75
b	79.75
c	38.60
Resolution	17.91–1.55 (1.6–1.55)
R _{split}	10.72 (60.21)
CC _{1/2}	98.14 (63.32)
SNR	6.84 (1.26)
Completeness	100 (100)
Multiplicity	241.7 (113.3)
No. reflections (R free set)	18,039 (902)
R _{work} /R _{free}	0.1808/0.1976
Bond length (Å)	0.005
Bond angle (°)	0.883
Ramachandran	
favored/allowed/outlier	99.21/0.79/0.0
Average B-factor	22.53

Data were acquired in the fixed medium gain and using a single X-ray pulse per exposure. Due to restrictions on the time window, only the first eight memory cells were illuminated.

found that, on average, it is substantially different from unity. The values of the slope ratios of G1 and G0 on a module level show that the gain conversion factor for medium gain in DGS mode is overestimated by approximately 30% when we assume that gain ratios calculated in single cell mode will hold true for all storage cells. In Figure 6, the module average per memory cell is shown; the results are plotted according to the order in which the memory cells are filled during a burst. It can clearly be noted that not only is the ratio substantially lower than unity, but there appears to be a decreasing trend with filling order. No clear dependence on the FEM occupancy appears from the data.

3.1.2.2 Fixed medium gain

Operation in *burst* mode in fixed medium gain (fixG1) presents fewer calibration challenges. On top of the obvious absence of a GSR with the resulting pedestal evaluation difficulties highlighted in the

previous paragraph, an absolute gain calibration is not necessary to maintain a linear response across its entire dynamic range (Figure 7A). Thus fixed gain setting appears to be a more straightforward candidate for *burst* operation mode.

As mentioned in Section 2.1.1, the CDS stage is not bypassed in fixed gain, hence the gain conversion factor calculated for G1 in DGS does not provide a full correction. To test the possibility to cross-calibrate the two factors (and hence measure the amplification provided by the CDS stage), the intensity scans in fixG1 have been corrected with the gain factor calculated for DGS. Data has been fitted with linear functions and the intercalibration factor R_{CDS} was estimated using the formula:

$$R_{CDS} = G1_{fix}/G1_{DGS}.$$

(3)

The average values of R_{CDS} as a function of the memory cell number for various occupancy levels are plotted in Figure 7B. Results indicate a weak dependence of the R_{CDS} factor on the memory cell, while the total amount of signal in the FEM has to be taken into account when estimating the values of R_{CDS} . At this point, it has to be stressed that each occupancy level data set was collected for a different module. To verify to what extent the observed trends can be explained by the unique characteristics of the individual modules, the intensity scans were repeated in the flat-field illumination geometry using the mask shown in Figure 3B. As presented in Figure 8, the values of the R_{CDS} factor are similar across all the modules and consistent with the lower occupancy level data (Figure 7B).

3.2 First experiments

Due to the issues encountered in the characterization of the operation of the JUNGFRAU detector in *burst* mode with DGS, and the promising results during *fixed gain* operation, the experimental tests of SFX and Kossel line measurement have been performed without the DGS mechanism. These types of measurements represent different regimes of detector operation. In the case of SFX, diffracted photons are confined, mostly, to clusters of pixels. The accumulated charge varies among the clusters by few orders of magnitude. Resolving Kossel lines requires the detection of small signals superimposed by a large, uniform background. In the following sections, we show that *burst* mode operation in fixed gain can be used in both scenarios.

3.2.1 Serial femtosecond crystallography

For this particular experiment, an exposure time of 12.7 μs and a single X-ray pulse per frame were chosen. The accelerator was configured to produce 24 fs electron bunches. Although all 16 memory cells were used only the first 8 memory cells were illuminated by X-rays, due to the limited RF-window available for this particular experiment. Diffraction patterns were processed with the CrystFEL software [26] with PEAKFINDER8 and XGANDALF used for peakfinding and indexing, respectively. The PHENIX package [27] was then used to perform molecular replacement (with PDB 6FTR as search model) and structure refinement. Data from 97,128 bursts was acquired in two 10 min data sets. In total, 18,420 crystals were identified with an indexing rate of 18.96%,

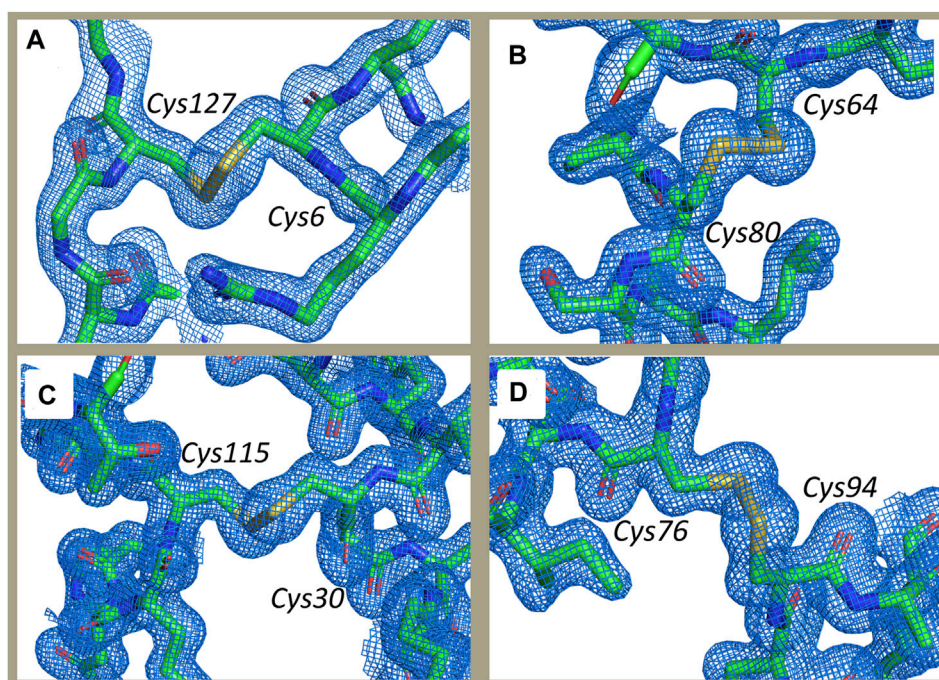


FIGURE 9

The electron density maps around the four disulfide bonds area in the hen egg-white lysozyme. Regions around cysteines 6 and 127 (A), 64 and 80 (B), 30 and 115 (C), 76 and 94 (D) display continuous densities. The 2Fo-Fc map is shown in blue at 1σ using PyMOL version 2.3.0 (ref: The PyMOL Molecular Graphics System, Version 2.3 Schrodinger, LLC).

only counting the illuminated frames. The figures of merit showing very good statistics and outcome of the structure refinement are summarized in Table 1. The obtained electron density map at 1.55 Å resolution appears free from local radiation damage as illustrated by the well-defined disulfide bond areas and presented in Figure 9. Distribution of the indexed frames across the illuminated memory cells is shown in Figure 10. As desired, the indexed patterns are uniformly distributed among the first eight cells, while no crystals were identified in any of the “dark” cells, which indicates a lack of cross-talk between the memory cells.

3.2.2 Kossel line measurement

Example data from a single 16-pulse burst showing the evolution of the Kossel lines as function of the radiation dose are presented in Figure 11. Despite the large intensity gradient in the collected images, the Kossel lines can clearly be resolved on the single-pulse level. Visibility of the lines increased as a function of the pulse number in the burst, reflecting the change of the mesoscopic structure of the sample. With the 120 mm sample-to-detector distance, the sensitive area of the detector covered a sufficient fraction of the Kossel line pattern to solve the 3D crystalline structure of the sample but the resolution of the fine structure of the lines was limited. With increasing accumulated dose towards the end of the burst, radiation damage of the wafer becomes evident, which is reflected by the sharp Bragg peaks from polycrystalline Ge. For future experiments, the quality of the data could be greatly improved by thinning the sample, switching to the backscattering geometry, and extending the sample-to-detector distance. However, the outcome of this experiment proved the feasibility of pulse-resolved Kossel line measurements at the European XFEL. Most importantly, it showed

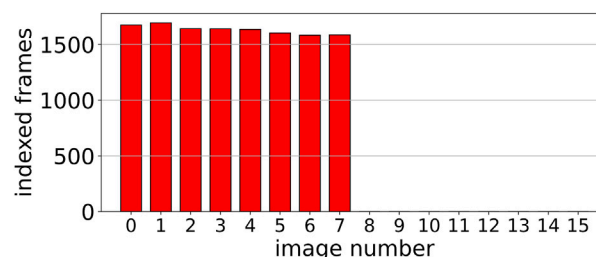


FIGURE 10

Number of the indexed crystals per memory cell for the serial femtosecond crystallography data set from the hen egg-white lysozyme collected in the *burst* mode and fixed medium gain. Only the first eight frames were illuminated and the rest was kept dark.

that the JUNGFRU detector operated in *burst* mode is a well-suited tool to support such demanding measurements at kHz time-scales.

4 Discussion

These characterization results indicate that the operation of the JUNGFRU detector in *burst* mode with DGS requires a new calibration procedure in order to properly correct the raw output after gain switching:

- The current pedestal evaluation for the G1 and G2 gain stages through forced gain switching does not produce predictive results;

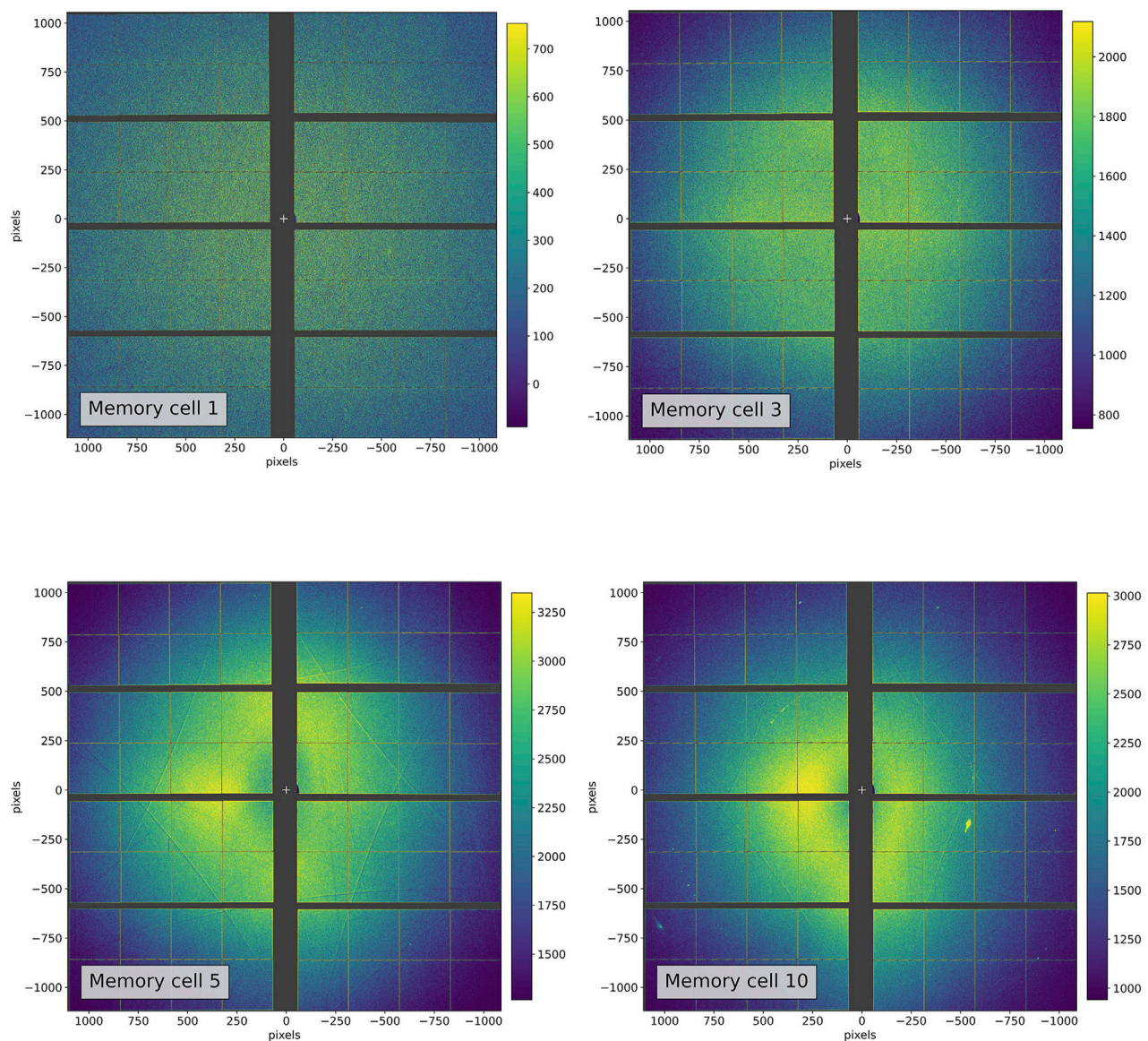


FIGURE 11

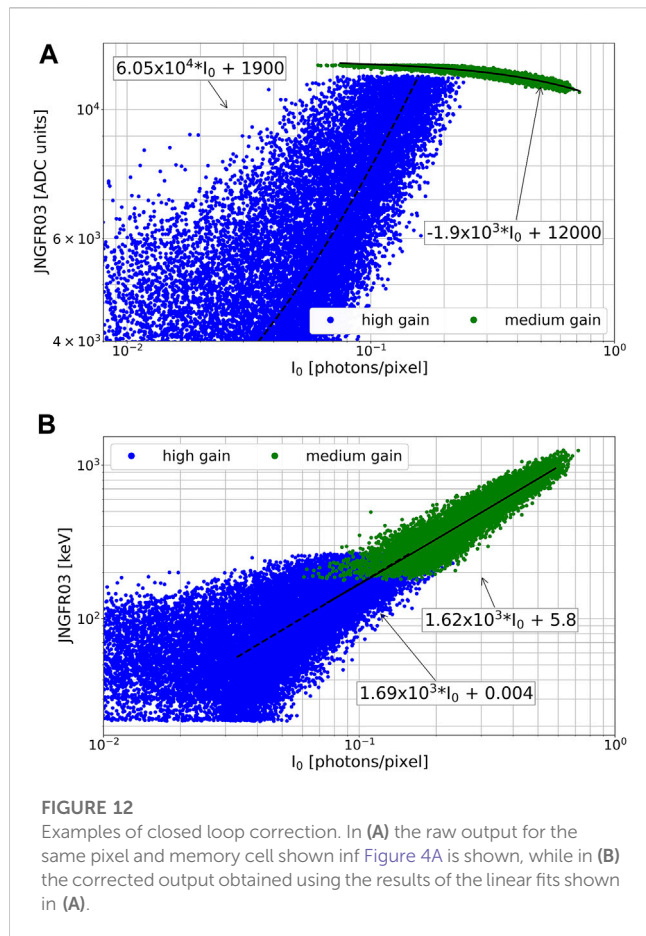
The Kossel lines measured for 100 μm thick Ge single crystal using the JUNGFRÄU detector in fixed medium gain configuration. Photon energy of the incident radiation was set to 11.5 keV. The sample was illuminated by a single X-ray pulse per each 20.12 μs exposure. Visibility of the lines improves with increasing accumulated dose. Eventually, the polycrystalline Ge becomes well-pronounced.

- The gain ratios R_1 and R_2 estimated for *single cell* operation do not accurately predict the gain conversion factors for lower gain stages. These are overestimated by $\sim 30\%$, with a dependency on the order in which the memory cells are filled. This indicates that we could experience one (or more) of the following effects:
 1. As well as the $G_{0_{burst}}$ gain being lower than expected due to the memory cell read/write operation, the feedback capacitance of the pre-amplifier in $G_{0_{burst}}$ is lower than predicted. This could be due to additional parasitic capacitances coming from the different way in which the pixel is operated in burst mode.
 2. It is not possible to completely preset the pre-amplifier charge in between images during the burst, which will introduce inter-cell mixing effectively resulting in a lower gain.

Dedicated measurements are planned to understand the issue in detail.

Operation at photon fluxes so low that gain switching is not triggered does not seem to present particular calibration issues.

Concerning *fixed gain* operation, no issue with the pedestal evaluation has been noticed, while the issue of calculating the intercalibration factors R_{CDs} for each pixel and memory cell remains unresolved. It is however worth stressing that in the



experimental tests conducted in *fixed gain*, the absence of an absolute gain calibration did not impact the successful results, as presented in Sections 3.2.1 and 3.2.2.

4.1 Prospects for calibration strategies

4.1.1 Pedestal calibration

At the moment two possible calibration strategies for pedestal evaluation are being envisaged:

1. Gap correction: continue using the forced gain switch for pedestal estimate and apply an additional correction for the 'gap';
2. Linear fit estimate: fit with a linear function the raw output of the intensity scan, and use the linear function offset fit parameter as an estimate of the pedestal, in a similar way to the closed loop correction described in [20].

An example of the results obtained using *linear fit estimates* of offset is shown in Figure 12. There, the same data are shown as in Figure 4A, however they have been corrected with the results of the fit of the raw output, and as a result the 'gap' is closed and the slopes align.

Since the re-calibration of the pedestal via an intensity scan is not practical in many experimental situations, the calibration constants produced via both the methods proposed above need to be tested for their portability, their dependence on temperature and exposure time. Results will be presented in subsequent communications.

4.1.2 Gain calibration

It is evident that *burst* operation mode in DGS requires a re-calibration of the gain conversion factors R_1 and R_2 , in order to maintain a linear response through the whole dynamic range. For this purpose, the method described in (19), must be applied to the detector when operating in *burst* mode. In addition, the dynamic range scans necessary to calculate R_1 and R_2 may also be employed to extract information about the pedestal position for lower gain stages.

Concerning *fixed gain* operation, instead, the calculation of R_{CDS} is not fundamental to maintain a linear response, and it is useful only if absolute gain calibration is needed. If this would be the case, the intercalibration factor can be calculated using dynamic range scans with photons, or external charge injection. Dependence of the R_{CDS} as function of occupancy levels or memory cell did not impact the quality of the presented scientific data. Additional calibration of the gains, may be required when the total signal within the images would fluctuate by several orders of magnitude.

4.2 Experimental tests in fixed gain

The results obtained with the detector operated in *burst* mode with *fixed gain* settings prove that the detector can produce scientifically relevant results. At present this operation mode is limited to cases where the absolute gain calibration is not critical for the experimental outcome, but a seemingly straightforward gain intercalibration presents a good prospect to further widen the range of accessible experiments.

4.3 Conclusions

The calibration strategy proposed in Section 2.1.2.1 for the JUNGFRAU DGS mechanism in *burst* operation mode has been put to test extensively. Results indicate that forced gain switching produces a largely underestimated evaluation of the pedestal for the lower gains (G1 and G2), which results in strong artificial discontinuities in the corrected data. Concerning the gain calibration, the assumption that the ratio between feedback capacitors estimated with the *single cell* calibration procedure should remain valid leads to an overestimation of the actual gain conversion factor.

Both these results indicate the need to re-think the calibration procedure required for *burst* operation mode with DGS mechanism: the possibility of using dynamic range scans as outlined in Section 4.1 seems promising, but the method needs to be validated and tested most of all against the stability of its results. The outcome of this study will be presented in future communications.

On the other hand, the operation in *fixed gain* seems not to present major obstacles to its usage in experiments. Two separate tests have been conducted, namely, protein crystallography and Kossel line measurements. The results validated the data produced by the JUNGFRAU detector in *burst* mode and showed benefits of this mode in making better use of the EuXFEL pulse train structure, although compromises are needed on the dynamic range.

Besides benefiting SFX, implementations of this fixed-gain operation mode of JUNGFRAU will be relevant for many experimental techniques utilized at the European XFEL. For

example, ultrafast X-ray diffraction experiments, where the intensity and position of a Bragg peak are monitored as a function of laser excitation, will allow pulse-resolved measurements to be performed at up to 16 pulses per train, providing better signal-to-noise ratio and statistics on compatible samples. As the Bragg diffraction intensity does not vary over several orders of magnitude between pulses, these measurements do not necessarily require the DGS capabilities for data collection. A similar method will also work for time-resolved X-ray emission spectroscopy (XES) measurements, where the per-pulse X-ray signals are often weak, allowing a fixed gain mode to be defined for data collection. This approach will allow XES measurements to take advantage of the excellent low noise characteristics of the JUNGFRU detector, while benefiting from the pulse-resolved measurement capabilities. This latter ability will allow interleaved laser excitation, where every second X-ray pulse in the train is laser-excited, resulting in significant improvements in the signal-to-noise ratio of laser ON - laser OFF differences due to the better intra-train stability of the X-ray parameters compared to the inter-train stability. Many pump-probe experiments require the highest possible laser powers, which are often achieved at the cost of repetition rates reduced to a few hundreds of kHz. In those cases, the *burst* mode of JUNGFRU becomes an attractive option.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

MR: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Validation, Writing-original draft, Writing-review and editing. MS: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Validation, Writing-original draft, Writing-review and editing. RW: Formal Analysis, Methodology, Software, Validation, Writing-review and editing. VH: Conceptualization, Methodology, Supervision, Writing-review and editing. AM: Conceptualization, Methodology, Supervision, Writing-review and editing. BS: Conceptualization, Investigation, Supervision, Writing-review and editing. HH: Resources, Writing-review and editing. RB: Writing-review and editing, Project administration, Resources. JB: Resources, Writing-review and editing. GB: Conceptualization, Formal Analysis, Methodology, Resources, Writing-review and editing. TD: Resources, Writing-review and editing. GF: Conceptualization, Formal Analysis, Methodology, Resources, Writing-review and editing. KK: Resources,

Writing-review and editing. CK: Resources, Writing-review and editing. JK: Resources, Writing-review and editing. FK: Resources, Writing-review and editing. RL: Resources, Writing-review and editing. LL: Resources, Writing-review and editing. NR: Resources, Writing-review and editing. AR: Writing-review and editing. AS: Resources, Writing-review and editing. TS: Resources, Writing-review and editing. MiT: Conceptualization, Formal Analysis, Methodology, Resources, Writing-review and editing. MoT: Supervision, Writing-review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by Wellcome Trust grant. Bundesministerium für Bildung und Forschung grants 05K13GU7; and 05E13GU1. Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky grant. Svenska Forskningsrådet Formas grant 822-2013-2014. Knut och Alice Wallenbergs Stiftelse grant. Röntgen-Ångström Cluster grant. Stiftelsen för Strategisk Forskning grant. Australian Research Council Center of Excellence in Advanced Molecular Imaging grant CE140100011. Australian Nuclear Science and Technology Organisation grant. Max-Planck-Gesellschaft grant. Helmholtz Association grant. MSCA PSI-FELLOW-III-3i (EU Grant Agreement No. 884104).

Acknowledgments

The authors would like to acknowledge the contribution of the European XFEL and in particular of the scientific instruments SPB/SFX for the beam time allocations, preparation and support during operation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Mozzanica A, Andrä M, Barten R, Bergamaschi A, Chirioti S, Brückner M, et al. The JUNGFRU detector for applications at synchrotron light sources and XFELs. *Synchrotron Radiat News* (2018) 31:16–20. doi:10.1080/08940886.2018.1528429
2. Mozzanica A, Bergamaschi A, Cartier S, Dinapoli R, Greiffenberg D, Johnson I, et al. Prototype characterization of the JUNGFRU pixel detector for SwissFEL. *J Instrumentation* (2014) 9:C05010. doi:10.1088/1748-0221/9/05/C05010

- 3 . Mozzanica A, Bergamaschi A, Brueckner M, Cartier S, Dinapoli R, Greiffenberg D, et al. Characterization results of the JUNGFRÄU full scale readout ASIC. *J Instrumentation* (2016) 11:C02047. doi:10.1088/1748-0221/11/02/C02047
- 4 . Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, et al. A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nat Photon* (2020) 14:391–7. doi:10.1038/s41566-020-0607-z
- 5 . Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Radiat* (2019) 26:74–82. doi:10.1107/S1600577518016077
- 6 . Wheeler R, Hart M, Veale M, Wilson M, Doblas-Jiménez D, Turcato M, et al. Development of data correction for the 1M large pixel detector at the EuXFEL. *J Instrumentation* (2022) 17:P04013. doi:10.1088/1748-0221/17/04/P04013
- 7 . Porro M, Andricek L, Aschauer S, Castoldi A, Donato M, Engelke J, et al. The MiniSDD-based 1-mpixel camera of the DSSC Project for the European XFEL. *IEEE Trans Nucl Sci* (2021) 68:1334–50. doi:10.1109/TNS.2021.3076602
- 8 . Mezza D, Allahgholi A, Delfs A, Dinapoli R, Goettlicher P, Graafsma H, et al. New calibration circuitry and concept for AGIPD. *J Instrumentation* (2016) 11:C11019. doi:10.1088/1748-0221/11/11/C11019
- 9 . Mezza D, Becker J, Carraresi L, Castoldi A, Dinapoli R, Goettlicher P, et al. Calibration methods for charge integrating detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1024:166078. doi:10.1016/j.nima.2021.166078
- 10 . Tolstikova A, Levantino M, Yefanov O, Hennicke V, Fischer P, Meyer J, et al. 1 kHz fixed-target serial crystallography using a multilayer monochromator and an integrating pixel detector. *IUCrJ* (2019) 6:927–37. doi:10.1107/S205225251900914X
- 11 . Preston T, Göde S, Schwinkendorf JP, Appel K, Brambrink E, Cerantola V, et al. Design and performance characterisation of the HAPG von Hámos spectrometer at the high energy density instrument of the European XFEL. *J Instrumentation* (2020) 15:P11033. doi:10.1088/1748-0221/15/11/P11033
- 12 . Milne CJ, Schietinger T, Aiba M, Alarcon A, Alex J, Anghel A, et al. SwissFEL: the Swiss X-ray free electron laser. *Appl Sci* (2017) 7:720. doi:10.3390/app7070720
- 13 . Redford S, Bergamaschi A, Brückner M, Cartier S, Dinapoli R, Ekinci Y, et al. Calibration status and plans for the charge integrating JUNGFRÄU pixel detector for SwissFEL. *J Instrumentation* (2016) 11:C11013. doi:10.1088/1748-0221/11/11/C11013
- 14 . Khakhulin D, Otte F, Biednov M, Bömer C, Choi TK, Diez M, et al. Ultrafast X-ray photochemistry at European XFEL: capabilities of the femtosecond X-ray experiments (FXE) instrument. *Appl Sci* (2020) 10:995. doi:10.3390/app10030995
- 15 . Bacellar C, Kinschel D, Mancini GF, Ingle RA, Rouxel J, Cannelli O, et al. Spin cascade and doming in ferric hemes: femtosecond X-ray absorption and X-ray emission studies. *Proc Natl Acad Sci* (2020) 117:21914–20. doi:10.1073/pnas.2009490117
- 16 . Šmíd M, Baehtz C, Pelka A, Laso García A, Göde S, Grenzer J, et al. Mirror to measure small angle x-ray scattering signal in high energy density experiments. *Rev Scientific Instr* (2020) 91:123501. doi:10.1063/5.0021691
- 17 . Voigt K, Zhang M, Ramakrishna K, Amouretti A, Appel K, Brambrink E, et al. Demonstration of an x-ray Raman spectroscopy setup to study warm dense carbon at the high energy density instrument of European XFEL. *Phys Plasmas* (2021) 28:082701. doi:10.1063/5.0048150
- 18 . Kaa JM, Konôpková Z, Preston TR, Cerantola V, Sahle CJ, Förster M, et al. A von Hámos spectrometer for diamond anvil cell experiments at the high energy density instrument of the European X-ray free-electron laser. *J Synchrotron Radiat* (2023) 30:822–30. doi:10.1107/S1600577523003041
- 19 . Redford S, Andrä M, Barten R, Bergamaschi A, Brückner M, Dinapoli R, et al. First full dynamic range calibration of the JUNGFRÄU photon detector. *J Instrumentation* (2018) 13:C01027. doi:10.1088/1748-0221/13/01/c01027
- 20 . Redford S, Andrä M, Barten R, Bergamaschi A, Brückner M, Chirioti S, et al. First full dynamic range scan of the JUNGFRÄU detector performed at an XFEL with an accurate intensity reference. *J Instrumentation* (2020) 15:C02025. doi:10.1088/1748-0221/15/02/c02025
- 21 . Mancuso AP, Aquila A, Batchelor L, Bean RJ, Bielecki J, Borchers G, et al. The single particles, clusters and biomolecules and serial femtosecond crystallography instrument of the European XFEL: initial installation. *J Synchrotron Radiat* (2019) 26:660–76. doi:10.1107/S1600577519003308
- 22 . Boutet S, Lomb L, Williams GJ, Barends TR, Aquila A, Doak RB, et al. High-resolution protein structure determination by serial femtosecond crystallography. *Serial Femtosecond Crystallogr* (2012) 337:362–4. doi:10.1126/science.1217737
- 23 . Vakili M, Vasireddi R, Gwozdz PV, Monteiro DCF, Heymann M, Blick RH, et al. Microfluidic polyimide gas dynamic virtual nozzles for serial crystallography. *Rev Scientific Instr* (2020) 91:085108. doi:10.1063/5.0012806
- 24 . Faigel G, Bortel G, Tegze M. Experimental phase determination of the structure factor from kossel line profile. *Scientific Rep* (2016) 6:22904. doi:10.1038/srep22904
- 25 . Bielecki J, Maia FRNC, Mancuso AP. Perspectives on single particle imaging with x rays at the advent of high repetition rate x-ray free electron laser sources. *Struct Dyn* (2020) 7:040901. doi:10.1063/4.0000024
- 26 . White TA, Kirian RA, Martin AV, Aquila A, Nass K, Barty A, et al. CrystFEL: a software suite for snapshot serial crystallography. *J Appl Crystallogr* (2012) 45:335–41. doi:10.1107/S0021889812002312
- 27 . Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Section D* (2010) 66:213–21. doi:10.1107/S0907444909052925



OPEN ACCESS

EDITED BY

Cornelia B. Wunderer,
Helmholtz Association of German
Research Centres (HZ), Germany

REVIEWED BY

Christopher John Hall,
Australian Nuclear Science and
Technology Organisation, Australia
David Pennicard,
Helmholtz Association of German
Research Centres (HZ), Germany
Uddhab Chaulagain,
ELI Beamlines, Czechia

*CORRESPONDENCE

Chris D. Armstrong,
✉ chris.armstrong@stfc.ac.uk

[†]PRESENT ADDRESS

K. Welsby, UK Research and Innovation,
Swindon, United Kingdom

RECEIVED 31 August 2023

ACCEPTED 16 November 2023

PUBLISHED 21 December 2023

CITATION

Armstrong CD, Scott GG, Richards S,
Patel JK, Fedorov K, Gray RJ, Welsby K
and Rajeev PP (2023), X-ray detector
requirements for
laser–plasma accelerators.
Front. Phys. 11:1286442.
doi: 10.3389/fphy.2023.1286442

COPYRIGHT

© 2023 Armstrong, Scott, Richards, Patel,
Fedorov, Gray, Welsby and Rajeev. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

X-ray detector requirements for laser–plasma accelerators

Chris D. Armstrong^{1*}, G. G. Scott¹, S. Richards¹, J. K. Patel¹,
K. Fedorov¹, R. J. Gray², K. Welsby^{1†} and P. P. Rajeev¹

¹Central Laser Facility, Rutherford Appleton Laboratory, Harwell, United Kingdom, ²Department of Physics
SUPA, University of Strathclyde, Glasgow, United Kingdom

Laser–plasma interactions (LPIs) are an emerging source of a range of energetic radiation. LPI experiments drive ultra-short (< ps) and brilliant sources of X-rays from keV to MeV energies. Designing detectors to maximise the sensitivity and resolution achievable with these sources is paramount to optimising laser-driven accelerators. In this article, we explore the key parameters associated with laser-driven X-ray sources and the detector systems required to characterise them. We present a concise approach to modelling the sensitivity and resolution for indirect detector systems factoring in both the optical collection and the X-ray attenuation within the scintillator.

KEYWORDS

laser–plasma interactions, bremsstrahlung, betatron, inverse Compton scattering, laser wakefield, high-intensity laser, X-ray generation

1 Introduction

A unique aspect of high-intensity laser–plasma interactions (LPIs) is the ability to generate broad ranges of energetic radiation by tuning the plasma and laser conditions. The different radiation types share a common set of characteristics: they are of short pulse (< ps) [1–3], emanate from a small source (100 nm–1 mm) [4–8], and can be tuned by subtle variations in the laser parameters or target [9–13]. In general, when high-intensity lasers are focussed on a target (either gas, liquid, or solid), atoms are ionised, and the freed electrons are accelerated to high energies [14, 15]. These electrons continue to radiate via one of the mechanisms described below or seed secondary reactions such as collisional ionisation, prompting further emission. A schematic diagram of these mechanisms is shown in [Figure 1](#), with the expected photon emission characteristics for a PW (30 J, 30 fs) laser system to establish a baseline—a full description of each calculation is given in [Supplementary Material](#). These energies and conditions are considered to mirror the upcoming high-repetition laser facility EPAC at the Central Laser Facility [16].

In laser wakefield acceleration, the laser is focussed onto a gas target at low ($\sim 10^{18-19} n_e/\text{cm}^3$) [9] density, which ionizes the gas and forms a plasma channel. The ponderomotive potential [14] of the laser drives electrons out of the focus and forms a plasma “wake” behind the laser pulse. This displacement of electrons sets up a strong electric field gradient, and electrons, injected into the wakefield, are accelerated to high energies. As the electrons are accelerated, they oscillate within the channel due to the background positive ion population and emit a synchrotron-like emission, betatron radiation [[Figure 1A](#)] [17–21], which is near-collimated and emanates from a small source area. For narrow-band X-ray emission, we can utilise inverse Compton scattering (ICS) ([Figure 1B](#)), where a secondary laser irradiates the relativistic electron beam and laser photons interacting with the electrons are upshifted by $\sim 4\gamma^2$, where γ is the electron Lorentz factor [22–25]. By

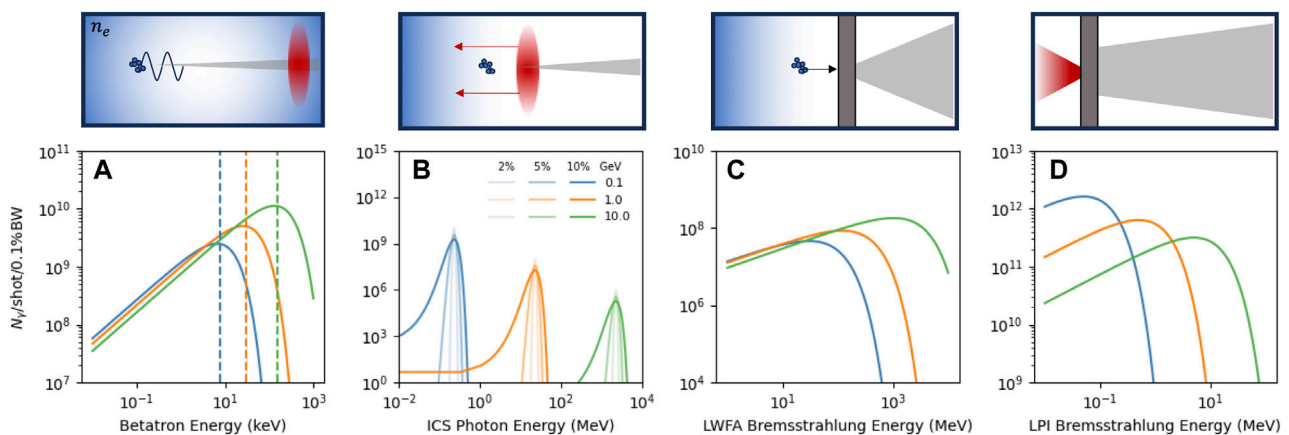


FIGURE 1

Schematic representation and example emission spectra for each X-ray generation mechanism by laser–plasma acceleration. (A) Betatron radiation, (B) inverse Compton scattering, (C) LWFA bremsstrahlung, and (D) direct target bremsstrahlung. Details of the calculations for each spectrum are given in the [Supplementary Material](#). The three curves for each correspond to different laser–plasma conditions; in (A–C), these correspond to intensities and plasma densities that drive 100 MeV, 1 GeV, and 10 GeV electron bunches, respectively; in (D), this corresponds only to the intensity with the effective bremsstrahlung temperature scaling from 100 keV to 10 MeV. See [Supplementary Material](#) for further details.

controlling the electron energies and focus of the scattering laser, the resultant X-ray emission can have a narrow spectral distribution at MeV energies and small divergence [23, 25–27]. Alternatively, the electron population can be accelerated into a solid converter foil to produce laser-wakefield bremsstrahlung (Figure 1C). This emission can be tuned by varying the thickness and material of the converter target [12, 28–30].

In laser–solid (overdense) interactions, the laser is again focussed on the target where electrons are ionised from the surface and accelerated up to MeV energies within the laser field before being driven into the target. Electrons accelerated by this mechanism generally have a broad Maxwell- or Boltzmann-like distribution of energies, and when interacting with the solid target, atoms produce a similarly broad distribution of laser–solid bremsstrahlung (Figure 1D) extending up to the peak energies of the electrons [4, 31–33]. The emission can be optimised by tuning the target material and thickness [34], lateral dimensions [4], and incident laser parameters [10, 33, 35]. The X-rays are typically more divergent than the laser-wakefield mechanisms but can still emanate from < 100- μ m source sizes [4, 6, 10, 33].

This article outlines the primary approaches to spatially resolving the emission from laser-driven sources, a brief introduction to two distinct methods we can use to detect X-rays, and then a discussion on how to apply these techniques to each of the different sources discussed above. Here, we distinguish between detectors (i.e., the sensor/substrate/scintillator that records the incident X-ray radiation) and diagnostics which translate the measured signal into a useful characteristic of the beam. As there are numerous regimes laser-driven sources can produce, we consider that no single solution/detector will suffice and, instead, describe the necessary parameters across three case studies.

- High-resolution imaging with betatron radiation—3.1
- High-energy imaging with ICS and bremsstrahlung radiation—3.2

- High instantaneous flux with direct or proximity-focussing detectors—3.3

The expected photon parameters, where necessary, are taken from the scaling shown in Figure 1.

2 X-ray imaging approaches for laser-driven sources

With the adoption of higher-repetition high-power laser facilities, there has been considerable effort by the community to produce high-stability interactions. There are several recent proposals using LPI as the initial accelerator for FELs [36], ion beamlines [37], and compact ICS sources [26]. Maier *et al.* demonstrated significant improvements in the stability of laser-driven sources [38], and new facilities are being built and designed with mechanical and machine stability in mind, including active feedback to ensure long-term stability of the driving laser. However, especially during commissioning of new facilities, detectors and diagnostics need to be able to fully characterise the emission on a single-shot basis and not integrate or scan, assuming that the emission is constant. The requirement for single-shot acquisitions limits what diagnostic techniques are possible; for example, super-resolution methods [39] and single-event processing techniques [40] are not possible. This requirement is, therefore, an important factor in detector design for laser-driven sources.

2.1 Direct detection

Indirect and direct detection schemes are both useful in laser–plasma interactions. For direct-detection schemes, electrons within the substrate material are freed by incident radiation and swept towards an electronic readout. The total charge is then

digitised by an ADC. High-grade Si, CdTe, and CZT semiconductors are often used for direct detection. Detectors based on this mechanism have demonstrated sub-% energy resolution [41], large area [42–44], and X-ray energies of up to MeV [45]. Typically, these devices are operated in an integrating mode where the counts after the ADC correspond to an amount of charge deposited in the pixel. This can be interpreted as either the number of photons at a given energy or the energy of an individual photon. With the latter mode, it is, therefore, possible to retrieve spectroscopic information as well. Recent efforts with direct-detector systems aimed to increase the full frame rate of the camera beyond MHz repetition rates [46–48] for use in high average brilliance facilities. As the pulse duration for high-intensity laser sources is typically on the order of femto- to pico-seconds, increases in the frame rate of the detector have a limited impact on reducing the number of photons per frame. Instead, the area of the pixels must be considered to control the chance of interactions per pulse, the thickness of the substrate to tune the stopping power, and the full-well capacity of the readout to maximise the total energy that can be measured.

2.2 Indirect detection

In contrast to direct detection, indirect detection uses scintillators to convert the X-ray radiation into optical light that is then imaged with a standard optical camera and lens system. By tuning the scintillator and the optical relay, we can optimise detection as needed. The optical collector is either a lens or a fibre-optic plate. For imaging, the typical lenses considered are either microscope objectives or camera/machine vision lenses. Both options are compound lenses with many elements to correct aberrations or minimise distortion; however, the microscope objectives are generally of high numerical aperture with a magnification greater than unity, whereas machine vision lenses are typically of lower numerical aperture and demagnify the image onto the sensor. Critically, with either option, we must factor in the distance to the scintillator, (d_o), and opening aperture, D , of the optic to determine the effective numerical aperture ($NA = D/2d_o$) of the system with the magnification (M_o) to determine what the collection efficiency and resolution will be. There have been several studies on the relationship between these parameters [32, 49, 50]; in general, resolution and collection efficiency are inversely proportional. The resolution limit in the optically limited regime can be approximated as follows:

$$R[\mu\text{m}] = \sqrt{\underbrace{\left(\frac{p}{NA}\right)^2}_{\text{Diffraction}} + \underbrace{(q\ell NA)^2}_{\text{Defect of focus}} + \underbrace{\left(\frac{2dx}{M_o}\right)^2}_{\text{Sampling}}}, \quad (1)$$

where p and q are fitting parameters determined by Koch *et al.* [49] and ℓ is the thickness of the scintillator. Throughout this article, we define the resolution as 90% of the integrated line-spread-function to mirror the work by Koch *et al.* [49]. This is notably larger than the full-width half-maximum, and so, features below the limits given here could still be visible—albeit at a lower contrast. The first term in Eq. 1 relates to the cumulative effects of diffraction and spherical aberrations stemming from the scintillator–air interface,

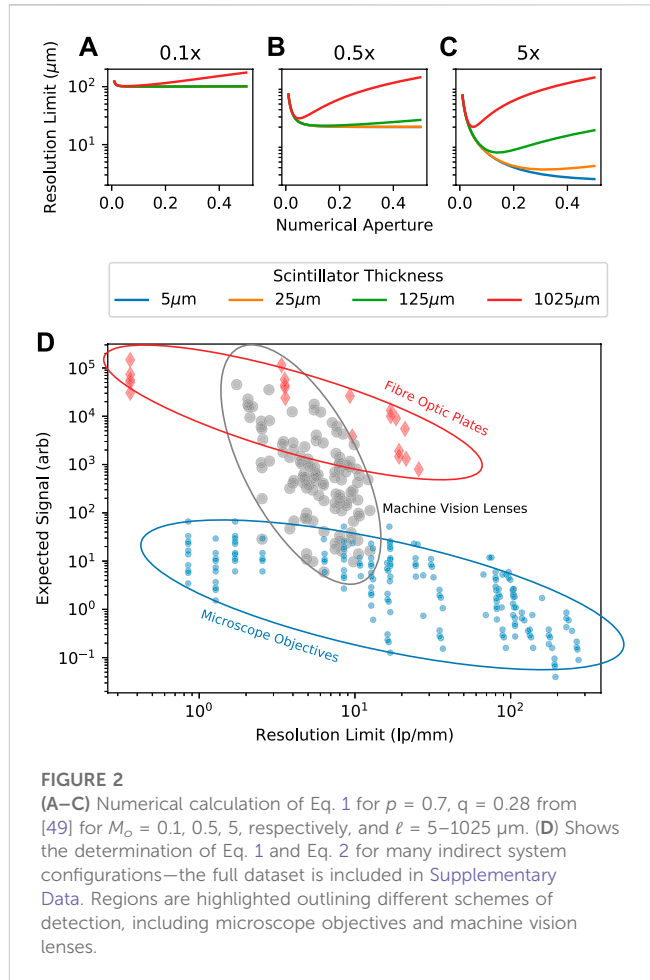


FIGURE 2

(A–C) Numerical calculation of Eq. 1 for $p = 0.7$, $q = 0.28$ from [49] for $M_o = 0.1, 0.5, 5$, respectively, and $\ell = 5\text{--}1025\ \mu\text{m}$. (D) Shows the determination of Eq. 1 and Eq. 2 for many indirect system configurations—the full dataset is included in [Supplementary Data](#). Regions are highlighted outlining different schemes of detection, including microscope objectives and machine vision lenses.

and the second term is due to planes beyond the point of best focus contributing. We include the third term factoring in the pixel size dx and the optical magnification M_o to account for the limit in the low-magnification domain. In addition, this relationship holds assuming a pencil-like deposition throughout the scintillator. For the purposes of this discussion, we omit effects due to non-uniform attenuation throughout the scintillator [49] and large electron cascades with high-energy incident X-rays [51]. Assuming that we have a relatively thin ($\ell < \lambda_{\text{mfp}}$) scintillator and a limited electron cloud radius, Eq. 1 sets the resolution at the detector plane, and so, by controlling the geometric magnification, M_g , the system resolution, can exceed this for samples or images. The signal detected in each pixel, however, is dependent on the attenuation of the scintillator, the conversion efficiency into optical photons κ , the collective power of the lens system, and the magnified area of pixels:

$$\text{Signal} \left[\frac{\text{Counts}}{\text{px}} \right] = \underbrace{\frac{E_{\text{dep}}}{\text{mm}^2}}_{\text{Photons Produced}} \cdot \underbrace{\kappa}_{\text{Collection}} \cdot \underbrace{QE(\lambda) \frac{NA^2}{4n^2}}_{\text{Collection}} \cdot \underbrace{\frac{A_{\text{px}}}{M_o^2} G}_{\text{Sampling}}, \quad (2)$$

where $QE(\lambda)$ is the quantum efficiency of the chosen sensor at the wavelength of emission, κ is the yield in photons per deposited energy for the scintillator material, n is the refractive index of the scintillator material (see [Supplementary Material](#) for collection efficiency derivation), and G a linear gain term to convert the

number of photons to counts recorded by the sensor. This equation is similar in form to that determined by Cardarelli et al. [32]; however, the F-stop terms they used are replaced by terms for the effective numerical aperture and optical magnification directly. Where the deposited energy per mm^2 is dependent on both the incident spectral shape, $f(E)$, and the attenuation of the scintillator, to a first order, this can be expressed using the Beer–Lambert model for transmission as

$$\frac{E_{\text{dep}}}{\text{mm}^2} = \int_0^{\infty} f(E) (1 - \exp^{-\rho \ell \sigma(E)}) dE, \quad (3)$$

where ρ is the density of the scintillator and σ is the attenuation cross section. Eq. 1 and Eq. 2 are both dependent on the length of the scintillator, ℓ , the numerical aperture of the lens system, and the optical magnification, resulting in a complex trade off of parameters between lens choices.

We demonstrate the calculation of both the signal and the resolution at the plane from parameters for typical systems, using both microscope objectives and camera lenses to image scintillators of different thicknesses, as shown in Figures 2A–C; we set the pixel size to $5 \mu\text{m}$ for each magnification and vary the numerical aperture of the lens to demonstrate the scaling. It is clear from evaluating Eq. 1 that the magnification and pixel size act as a hard limit on the resolution. When operating with high magnification, there becomes a significant dependence on the thickness of the scintillator and numerical aperture of the system. However, if we consider only the numerical aperture of the system, a clearer picture emerges—with a low numerical aperture; therefore, with a small angle being observed, there is little dependence on scintillator thickness, whereas with a high numerical aperture, the scintillator thickness becomes paramount to the achieved resolution. Figure 2D summarises how different commercially available optics vary in terms of achievable resolution and expected signal. The full dataset used is detailed in Supplementary Material. The data points assume a Gaussian X-ray distribution with a 100-keV centre and a full-width half-maximum of 40 keV.

3 Imaging case studies

In the following sections, we outline three distinct imaging challenges with laser-driven sources where advancements in detector characteristics would enhance the potential applications of laser-driven sources.

3.1 Single-shot high-resolution imaging with laser-driven betatron radiation

Laser-wakefield betatron sources have sub-micron source sizes [8]; however, typically, the smallest resolvable element achieved in experiments is much higher than this and is, instead, limited by the detector [52]. To achieve a high resolution at the detector plane, the best route is to use high-magnification objectives, as shown in Figure 2D; however, this limits the expected signal due to spreading over many pixels. Microscope objectives exhibit a general scaling between the numerical aperture and optical

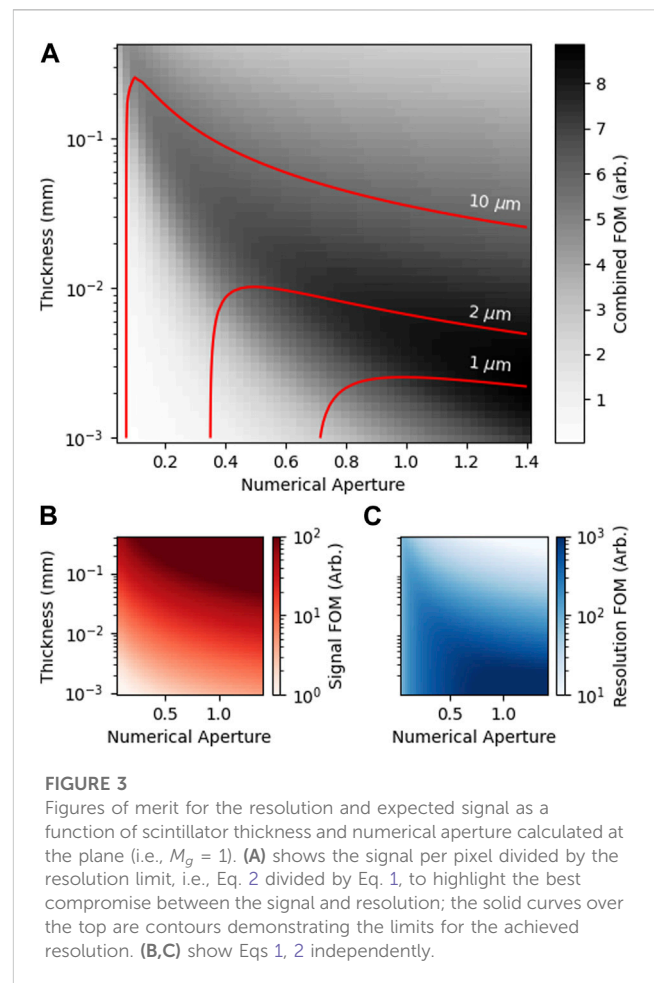


FIGURE 3

Figures of merit for the resolution and expected signal as a function of scintillator thickness and numerical aperture calculated at the plane (i.e., $M_o = 1$). (A) shows the signal per pixel divided by the resolution limit, i.e., Eq. 2 divided by Eq. 1, to highlight the best compromise between the signal and resolution; the solid curves over the top are contours demonstrating the limits for the achieved resolution. (B,C) show Eqs 1, 2 independently.

magnification of the form $M_o \propto \sqrt{\text{NA}}$. Then, for a given spectral distribution and scintillator material, we can express Eq. 1 and Eq. 2 as a function of the numerical aperture and scintillator thickness, respectively. The result of these two functions is shown in Figures 3B,C, and a combination figure of merit, calculated as the signal per pixel divided by the resolution limit, i.e., Eq. 2 divided by Eq. 1, is shown in Figure 3A. The contours on the first panel demonstrate the fundamental limits in resolving power for different numerical apertures. Achieving a sub-micron resolution at the plane requires both a high ($\text{NA} > 0.7$) numerical aperture and an ultra-thin ($\ell < 5 \mu\text{m}$) scintillator.

Betatron emission is divergent, albeit narrowly, ($\theta \approx \text{mRad}$), and so, the geometric magnification can increase the system resolution. This is also true of conventional X-ray tube sources; however, as their source area is typically large [53, 54] (excluding micro- and nano-focus systems [55]), the resolution is fundamentally limited by the extent of the source. The divergence of betatron emission combined with a small source opens up additional pathways to achieve sub-micron imaging, and we can look at the trade-offs between different optical systems by considering how the resolution and sensitivity vary as a function of each variable in 1–2.

Figure 4 shows the variation in the resolution and effective signal due to several parameters—numerical aperture (green), optical magnification (orange), geometric magnification (blue), and scintillator length (pink)—for a selection of microscope

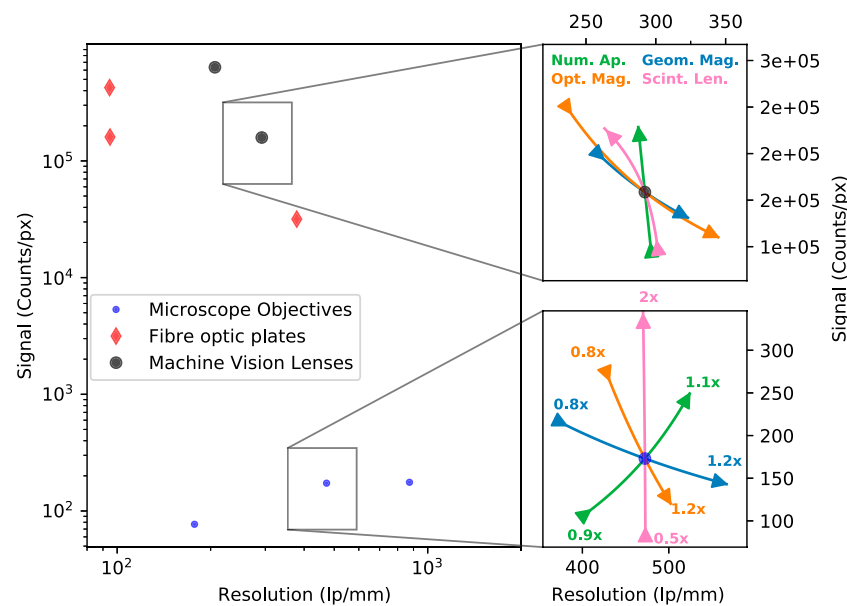


FIGURE 4

Expected signal (counts/px) and resolution (lp/mm) for microscope objectives, fibre-optic plates, and machine vision lenses, calculated by Eqs 1, 2.

The subplots demonstrate the variation in performance due to numerical aperture (green), optical magnification (orange), geometric magnification (blue), and scintillator length (pink).

objectives, fibre-optic plates (FOPs), and high numerical aperture machine vision lenses. These three approaches represent distinct options for optical imaging. Generally speaking, microscope objectives have a numerical aperture between 0.1 and 0.5 and high magnification, fibre-optic plates of a high numerical aperture >0.7 and $\sim \times 1$ magnification, and machine vision lenses have lower numerical aperture ~ 0.1 and typically demagnify the image ($M_o < 0.5$). However, what is interesting between these different approaches is how subtle changes in their respective parameters can change the resultant system performance. It is clear that the complex interplay between parameters makes it challenging to identify the *ideal* candidate. In the inset plots of Figure 4, we vary each parameter independently to demonstrate its effect. The arrows indicate the direction of increasing value for each parameter.

Interestingly, due to the difference in their initial values, increasing the numerical aperture for the highlighted machine vision lens results in a decrease in the expected resolution, leading to a significant increase in the highlight microscope objective. Conversely, if we consider increasing the thickness of the scintillator, we would observe an increase in the signal with no loss in resolution for the microscope objective, and yet, for the machine vision lens example, we observe a decrease in the resolution with minimal gains to the expected signal. Whilst these contradictions make it difficult to identify a general pattern, we note that the system resolution for the microscope objectives can be matched using the machine vision lenses but with a much greater signal. Since geometric magnification depends on the *sample* position rather than just the *detector* position, we can design systems to use lower optical magnification and increase the system resolution by reducing the sample position. The signal values are calculated assuming 7×10^{11} photons with a

divergence of 5 mRad and a critical energy of 50 keV—the parameters given for the example distribution in Figure 1A. For a single shot, it is clear that high-magnification objectives will result in a low signal-to-noise ratio, and so, many shots would likely need to be combined to achieve sufficient imaging quality.

In summary, to achieve a maximum signal with high-resolution imaging with betatron radiation, the best approach is to use relatively low optical magnification combined with high geometric magnification and keep the detector plane at the smallest distance from the source. In laser-wakefield acceleration (LWFA) experiments, the highly relativistic electron bunch must be deflected prior to interacting with the sample to minimise background contributions or damage to the sample. In practice, this sets a lower limit on the sample position ($Z \approx 1$ m), and therefore, a compromise will need to be sought for different experimental configurations. In addition, objectives that exceed the $M_o \propto \sqrt{\text{NA}}$ scaling offer potential routes to offset the losses in the signal due to magnification.

3.2 MeV imaging with ICS and laser-driven bremsstrahlung

It is pertinent to consider at this point the extreme case, where instead of a few hundred microns of the scintillator, we require thousands to ensure sufficient stopping power. Laser-driven sources can readily generate X-rays greater than 1 MeV via either inverse Compton sources or bremsstrahlung processes with energies extending up to 100 s of MeV for LWFA mechanisms. At these energies, the dominant X-ray attenuation mechanisms are no longer photo-electric absorption or scattering but nuclear and electronic-pair production [56]. This significantly alters the effects on resolution. No longer can we consider only the optical spreading

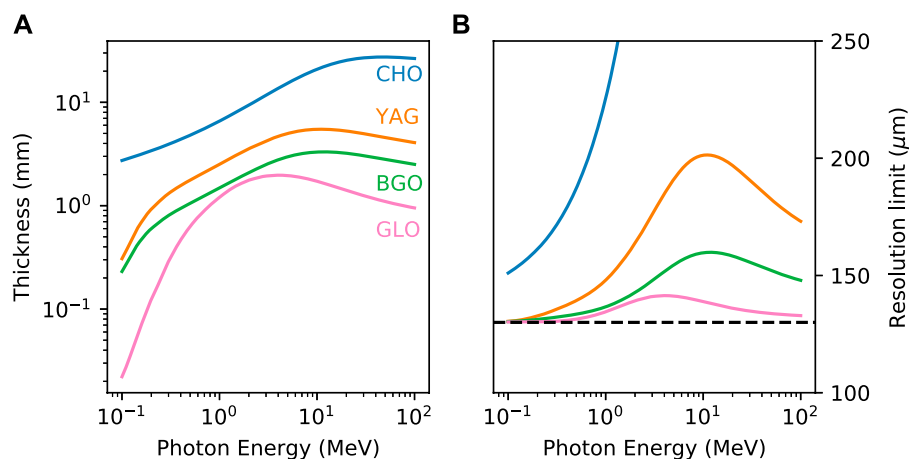


FIGURE 5

(A) Attenuation length for 10% absorption in each scintillator and (B) equivalent resolution limit for a scintillator of that length. Calculated via the NIST data tables [56] as a function of incident X-ray energy.

determined by Koch et al. [49] (Eq. 1); we must now also factor in the lateral diffusion that occurs as the X-ray deposits energy through a scintillator. Solving these additional factors requires in-depth Monte Carlo simulations to be conducted. However, we can determine some useful bounds from the aforementioned equations. First, we consider that for a given energy, we can calculate the required thickness, ℓ , as a function of the desired absorption, i.e., $\ell = -\log(T)/\sigma\rho$. Second, we wish to calculate what the intrinsic (i.e., PSF limited) resolution would be for that thickness to inform how we select scintillator materials. The calculation of this is shown in Figure 5 for a plastic scintillator (CHO, e.g., EJ260, BC422q), YAG:Ce, BGO, and the high-Z and high-density GLO scintillator [57]. For each material, the effective resolution limit determined in Figure 5 could then be further refined by Monte Carlo simulations adding a fourth term to Eq. 1 to account for the radial spreading as the X-rays deposit their energy. The energy spreading in materials is also dependent on the density and effective- z , and so, whilst the scattering at higher energies will reduce the resolution, this effect will be reduced with higher-density materials [58, 59].

This scaling with density and atomic number underlines why research on the GLO scintillator is beneficial to high-energy X-ray imaging; the higher density is intrinsically linked to the minimum thickness and relative attenuation to high-energy X-rays. Consequently, continued research and development into exotic scintillators and their manufacture will greatly benefit radiography with high-energy X-rays [60].

3.3 High instantaneous X-ray flux imaging

In the previous cases, we considered an optical system to image the scintillation onto the sensor. However, the lens system introduces significant losses in the collection efficiency, and by switching to either direct detection, the sensor in contact with the scintillator, or the proximity focussing of FOPs, we can consider a far more efficient detection method for imaging the X-rays. With the losses from the optical system minimised, the

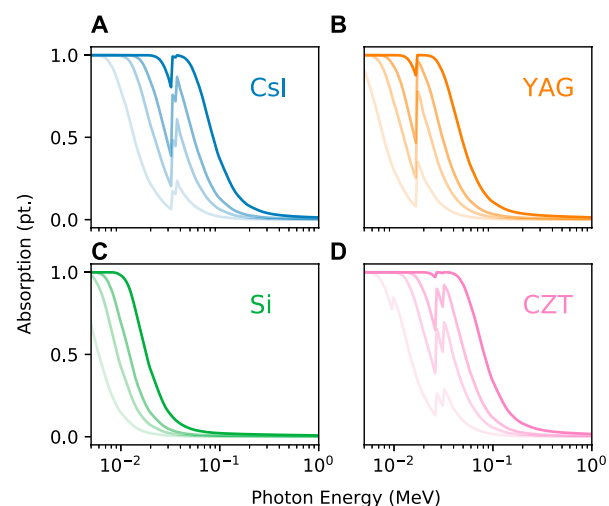


FIGURE 6

Absorption fraction for scintillators and direct-detector substrates, and the transparency of the line corresponds to the thickness of the layer—the order of increasing opacity corresponds to 20, 70, 150, and 500 μm thickness, respectively. Calculated via the NIST data tables [56] for (A) CsI, (B) YAG:Ce, (C) Si, and (D) CZT.

detection efficiency can be calculated directly from NIST-XCOM data tables [56] and the Beer–Lambert law [61]. Figure 6 shows that the efficiency of stopping in the material can be up to 100% for betatron-like energies (~50 keV) of X-rays for both scintillators (CsI and YAG) and thick direct-detection sensors such as CZT. The resolution for the direct detection is dependent on the pixel size and on PSF of the incident photon energy, whereas the FOP method is limited by the spreading in the scintillator, similar to Eq. 1.

One desired application for such a detector is to measure the unattenuated beam in a single shot. The number of X-ray photons arriving per pixel is dependent on the emitted number, the divergence of the emission, and the solid angle subtended by each pixel. With

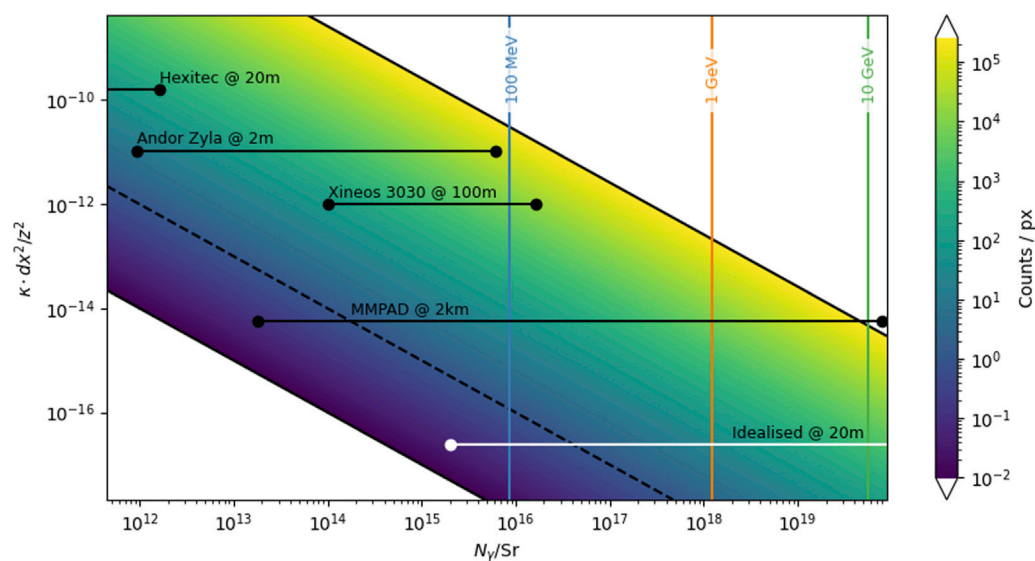


FIGURE 7

Calculated number of counts per pixel as a function of incident photons per Sr, and the relative pixel area (dx^2/z^2) and effective efficiency/gain (K) for a given detector system. The vertical lines correspond to the emission characteristics given in Figure 1A. The solid horizontal lines correspond to the effective dynamic range of the labelled detectors at representative distances, taking the background level and saturation point from the respective datasheets [42, 62–64].

these criteria and the photon distributions shown in Figure 1, we can generate the plot shown in Figure 7, showing the effective number of counts as a function of the incident photon intensity and the effective signal (Eq. 2) of a given detector. Here, we express the y-axis as Kdx^2/z^2 , where K is the overall efficiency term from incident photons per pixel to counts. Expressing Eq. 2 like this provides a clear route to improving detector performance. Figure 7 shows the approximate dynamic ranges for various detectors [42, 62–64] available that can operate in the desired range; additionally, we include an “idealised” detector that operates with a high dynamic range similar to the MM-Pad diagnostics [64] but a pixel size of $dx = 10 \mu\text{m}$ at a distance of $z = 20 \text{ m}$ and $K = 10^{-4} [\text{Counts}/\gamma_x]$ to compensate. Additionally, we include a dashed line at 1 count/px to indicate the relative threshold for single-photon spectroscopy techniques. Current detector systems, at reasonable distances ($< 100 \text{ m}$), cannot operate in the single-photon mode in the direct beam, leading to secondary scattering techniques to utilise such detectors [65]. Reducing the pixel size can, in principle, work; however, this approach is fundamentally limited since the charge cloud, the energy deposition volume, and lateral diffusion through the substrate will start to dominate as the pixel size is reduced. Extending the propagation further might be possible in certain facilities; however, it is impractical to consider kilometre-scale propagation distances in the near future, and since the divergence of betatron radiation is expected to drop with increasing electron energy [17], this approach will rapidly become impractical. In practice, to access the higher incident intensities, relatively inefficient systems, where $K \ll 1$, become the only realistic route to ensure sufficient sampling of the directly emitted beam. This is not to say thin scintillators and lossy lens systems are the only route to achieve it as this would result in significant disparity between low- and high-energy X-ray sensitivity. Instead, a high-attenuation substrate paired with a variable gain register [66] would provide a more appropriate solution for highly instantaneous flux environments.

4 Summary

Herein, we demonstrated, using analytical equations for resolution (Eq. 1) and signal (Eq. 2), approaches to X-ray imaging with laser-driven sources. The variety of X-ray energies and fluences poses a challenge to detectors but provides an opportunity for a wide variety of applications. Taking advantage of the diverging X-ray emission and the proximity to the source location, we can exploit geometric magnification to maximise the signal collected for a given resolution. With improvements to scintillator materials, we can design efficient scintillators for high-energy ($> \text{MeV}$) X-ray imaging while maintaining a $\sim 100\text{-}\mu\text{m}$ resolution as laser-driven source intensity increases further, i.e., by increasing the total flux per pulse and reducing the emission divergence.

Data Availability Statement

The original contributions presented in the study are publicly available. This data can be found here: <https://edata.stfc.ac.uk/handle/edata/945>.

Author contributions

CDA: Conceptualization, Investigation, Methodology, Writing—original draft. GS: Conceptualization, and Writing—review and editing. SR: Conceptualization, Investigation, and Writing—review and editing. JP: Writing—review and editing. KF: Writing—review and editing. RG: Conceptualization, Investigation, Methodology, and Writing—review and editing. KW: Supervision, and Writing—review and editing. PR: Funding acquisition, Investigation, Supervision, and Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Strategic Priorities Fund and EPSRC grant EP/V049232/1.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Hamster H, Sullivan A, Gordon S, White W, Falcone R. Subpicosecond, electromagnetic pulses from intense laser-plasma interaction. *Phys Rev Lett* (1993) 71:2725–8. doi:10.1103/physrevlett.71.2725
- Jaroszynski D, Bingham R, Brunetti E, Ersfeld B, Gallacher J, van Der Geer B, et al. Radiation sources based on laser–plasma interactions. *Phil Trans R Soc A: Math Phys Eng Sci* (2006) 364:689–710. doi:10.1098/rsta.2005.1732
- Khachatryan A, Van Goor F, Boller KJ, Reitsma A, Jaroszynski D. Extremely short relativistic-electron-bunch generation in the laser wakefield via novel bunch injection scheme. *Phys Rev Spec Topics-Accelerators Beams* (2004) 7:121301. doi:10.1103/physrevstab.7.121301
- Armstrong C, Brenner C, Jones C, Rusby D, Davidson Z, Zhang Y, et al. Bremsstrahlung emission from high power laser interactions with constrained targets for industrial radiography. *High Power Laser Sci Eng* (2019) 7:e24. doi:10.1017/hpl.2019.8
- Pirozhkov A, Esirkepov TZ, Pikuz T, Faenov AY, Ogura K, Hayashi Y, et al. Burst intensification by singularity emitting radiation in multi-stream flows. *Scientific Rep* (2017) 7:17968. doi:10.1038/s41598-017-17498-5
- Park HS, Maddox BR, Giraldez E, Hatchett SP, Hudson LT, Izumi N, et al. High-resolution 17–75 keV backlighters for high energy density experiments. *Phys Plasmas* (2008) 15. doi:10.1063/1.2957918
- Gruse JN, Streeter M, Thornton C, Armstrong C, Baird C, Bourgeois N, et al. Application of compact laser-driven accelerator x-ray sources for industrial imaging. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2020) 983:164369. doi:10.1016/j.nima.2020.164369
- Köhler A, Couperus J, Zarini O, Jochmann A, Irman A, Schramm U. Single-shot betatron source size measurement from a laser-wakefield accelerator. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 829:265–9. doi:10.1016/j.nima.2016.02.031
- Albert F, Thomas AG. Applications of laser wakefield accelerator-based light sources. *Plasma Phys Controlled Fusion* (2016) 58:103001. doi:10.1088/0741-3335/58/10/103001
- Armstrong C, Brenner C, Zemaityte E, Scott G, Rusby D, Liao G, et al. Bremsstrahlung emission profile from intense laser-solid interactions as a function of laser focal spot size. *Plasma Phys Controlled Fusion* (2019) 61:034001. doi:10.1088/1361-6587/aaf596
- Wood J. *Betatron radiation from laser wakefield accelerators and its applications*. London: Imperial College London (2017). Ph.D. thesis.
- Underwood C, Baird C, Murphy C, Armstrong C, Thornton C, Finlay O, et al. Development of control mechanisms for a laser wakefield accelerator-driven bremsstrahlung x-ray source for advanced radiographic imaging. *Plasma Phys Controlled Fusion* (2020) 62:124002. doi:10.1088/1361-6587/abbebe
- Finlay O, Gruse J, Thornton C, Armstrong C, Baird C, Bourgeois N, et al. Characterisation of a laser plasma betatron source for high resolution x-ray imaging. *Plasma Phys Controlled Fusion* (2021) 63:084010. doi:10.1088/1361-6587/ac0fcf
- Gibbon P. *Short pulse laser interactions with matter: an introduction*. Singapore: World Scientific (2005).
- McKenna P, Neely D, Bingham R, Jaroszynski D. *Laser-plasma interactions and applications*. Cham: Springer (2013).
- EPAC. Extreme photonics applications centre (2022). Available at: <https://www.clf.stfc.ac.uk/Pages/EPAC.aspx> (Accessed August 28, 2023).
- Albert F, Lemos N, Shaw J, King P, Pollock B, Goyon C, et al. Betatron x-ray radiation in the self-modulated laser wakefield acceleration regime: prospects for a novel

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphy.2023.1286442/full#supplementary-material>

- probe at large scale laser facilities. *Nucl Fusion* (2018) 59:032003. doi:10.1088/1741-4326/aad058
- Lu W, Huang C, Zhou M, Mori W, Katsouleas T. Limits of linear plasma wakefield theory for electron or positron beams. *Phys Plasmas* (2005) 12. doi:10.1063/1.1905587
- Lu W, Tzoufras M, Joshi C, Tsung F, Mori W, Vieira J, et al. Generating multi-gev electron bunches using single stage laser wakefield acceleration in a 3d nonlinear regime. *Phys Rev Spec Topics-Accelerators Beams* (2007) 10:061301. doi:10.1103/physrevstab.10.061301
- Fourmaux S, Hallin E, Chaulagain U, Weber S, Kieffer J. Laser-based synchrotron x-ray radiation experimental scaling. *Opt Express* (2020) 28:3147–58. doi:10.1364/oe.383818
- Kozlova M, Andriyash I, Gautier J, Sebban S, Smartsev S, Jourdain N, et al. Hard x rays from laser-wakefield accelerators in density tailored plasmas. *Phys Rev X* (2020) 10:011061. doi:10.1103/physrevx.10.011061
- Chen S, Powers N, Ghebregziabher I, Maharjan C, Liu C, Golovin G, et al. MeV-energy x rays from inverse Compton scattering with laser-wakefield accelerated electrons. *Phys Rev Lett* (2013) 110:155003. doi:10.1103/physrevlett.110.155003
- Tsai HE, Wang X, Shaw JM, Li Z, Arefiev AV, Zhang X, et al. Compact tunable Compton x-ray source from laser-plasma accelerator and plasma mirror. *Phys Plasmas* (2015) 22. doi:10.1063/1.4907655
- Graves W, Bessuille J, Brown P, Carbajo S, Dolgashev V, Hong KH, et al. Compact x-ray source based on burst-mode inverse Compton scattering at 100 kHz. *Phys Rev Spec Topics-Accelerators Beams* (2014) 17:120701. doi:10.1103/physrevstab.17.120701
- Sanri G, Corvan D, Schumaker W, Cole J, Di Piazza A, Ahmed H, et al. Ultrahigh brilliance multi-MeV γ -ray beams from nonlinear relativistic Thomson scattering. *Phys Rev Lett* (2014) 113:224801. doi:10.1103/physrevlett.113.224801
- Eggl E, Dierolf M, Achterhold K, Jud C, Günther B, Braig E, et al. The Munich compact light source: initial performance measures. *J synchrotron Radiat* (2016) 23:1137–42. doi:10.1107/s160057751600967x
- Döpp A, Guillaume E, Thaury C, Gautier J, Andriyash I, Lifschitz A, et al. An all-optical Compton source for single-exposure x-ray imaging. *Plasma Phys Controlled Fusion* (2016) 58:034005. doi:10.1088/0741-3335/58/3/034005
- Cipiccia S, Wiggins S, Shanks R, Islam M, Vieux G, Issac R, et al. A tuneable ultra-compact high-power, ultra-short pulsed, bright gamma-ray source based on bremsstrahlung radiation from laser-plasma accelerated electrons. *J Appl Phys* (2012) 111. doi:10.1063/1.3693537
- Dong K, Zhang T, Yu M, Wu Y, Zhu B, Tan F, et al. Micro-spot gamma-ray generation based on laser wakefield acceleration. *J Appl Phys* (2018) 123. doi:10.1063/1.4997142
- Li S, Shen B, Xu J, Xu T, Yu Y, Li J, et al. Ultrafast multi-MeV gamma-ray beam produced by laser-accelerated electrons. *Phys Plasmas* (2017) 24:093104. doi:10.1063/1.4996020
- Giulietti D, Gizzi LA. X-ray emission from laser-produced plasmas. *La Rivista Del Nuovo Cimento* (1978-1999) (1998) 21:1–93. doi:10.1007/bf02874624
- Cardarelli P, Paterno G, Di Domenico G, Consoli E, Marziani M, Andreotti M, et al. A gamma beam profile imager for ELI-NP gamma beam system. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2018) 893:109–16. doi:10.1016/j.nima.2018.03.023
- Brenner C, Mirfayzi S, Rusby D, Armstrong C, Alejo A, Wilson L, et al. Laser-driven x-ray and neutron source development for industrial applications of plasma accelerators. *Plasma Phys Controlled Fusion* (2015) 58:014039. doi:10.1088/0741-3335/58/1/014039

34. Fiorini F, Neely D, Clarke RJ, Green S. Characterization of laser-driven electron and photon beams using the Monte Carlo code FLUKA. *Laser Part Beams* (2014) 32: 233–41. doi:10.1017/S0263034614000044
35. Borm B, Khaghani D, Neumayer P. Properties of laser-driven hard x-ray sources over a wide range of laser intensities. *Phys Plasmas* (2019) 26. doi:10.1063/1.5081800
36. Assmann R, Weikum M, Akhter T, Alesini D, Alexandrova A, Anania M, et al. Eupraxia conceptual design report. *Eur Phys J Spec Top* (2020) 229:3675–4284. doi:10.1140/epjst/e2020-000127-8
37. Aymar G, Becker T, Boogert S, Borghesi M, Bingham R, Brenner C, et al. Lhara: the laser-hybrid accelerator for radiobiological applications. *Front Phys* (2020) 8: 567738. doi:10.3389/fphy.2020.567738
38. Maier AR, Delbos NM, Eichner T, Hübner L, Jalas S, Jeppe L, et al. Decoding sources of energy variability in a laser-plasma accelerator. *Phys Rev X* (2020) 10:031039. doi:10.1103/physrevx.10.031039
39. Tian J, Ma KK. A survey on super-resolution imaging. *Signal Image Video Process.* (2011) 5:329–42. doi:10.1007/s11760-010-0204-6
40. Lu G, Fei B. Medical hyperspectral imaging: a review. *J Biomed Opt* (2014) 19: 010901. doi:10.1117/1.jbo.19.1.010901
41. Jones L, Seller P, Wilson M, Hardie A. Hexitec asic—a pixellated readout chip for czt detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2009) 604:34–7. doi:10.1016/j.nima.2009.01.046
42. Wilson M, Dummott L, Duarte DD, Green F, Pani S, Schneider A, et al. A 10 cm × 10 cm CdTe spectroscopic imaging detector based on the HEXITEC ASIC. *J Instrumentation* (2015) 10:P10011. doi:10.1088/1748-0221/10/10/p10011
43. Veale M, Seller P, Wilson M, Liotti E. Hexitec: a high-energy x-ray spectroscopic imaging detector for synchrotron applications. *Synchrotron Radiat News* (2018) 31: 28–32. doi:10.1080/08940886.2018.1528431
44. Ballabriga R, Campbell M, Llopart X. Asic developments for radiation imaging applications: the medipix and timepix family. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2018) 878:10–23. doi:10.1016/j.nima.2017.07.029
45. Meuris A, Limousin O, Gevin O, Lugiez F, Le Mer I, Pinsard F, et al. Caliste hd: a new fine pitch cd (zn) te imaging spectrometer from 2 kev up to 1 mev. In: 2011 IEEE Nuclear Science Symposium Conference Record (IEEE); 30 October - 6 November 2010; Knoxville, Tennessee (2011). p. 4485–8.
46. Jones L, Bell S, Cline B, Gardiner T, Hart M, Prydderch M, et al. Spectroscopic x-ray imaging at mhz frame rates—the hexitecmhz asic. *J Instrumentation* (2022) 17: C10012. doi:10.1088/1748-0221/17/10/c10012
47. Zhang J, Andrä M, Barten R, Bergamaschi A, Brückner M, Chirioti-Alvarez S, et al. Design and first tests of the gotthard-ii readout asic for the european x-ray free-electron laser. *J Instrumentation* (2021) 16:P04015. doi:10.1088/1748-0221/16/04/p04015
48. Allahgholi A, Becker J, Bianco L, Delfs A, Dinapoli R, Goettlicher P, et al. Agipd, a high dynamic range fast detector for the european xfel. *J Instrumentation* (2015) 10: C01023. doi:10.1088/1748-0221/10/01/c01023
49. Koch A, Raven C, Spanne P, Snigirev A. X-ray imaging with submicrometer resolution employing transparent luminescent screens. *JOSA A* (1998) 15:1940–51. doi:10.1364/josaa.15.001940
50. Martin T, Koch A. Recent developments in x-ray imaging with micrometer spatial resolution. *J synchrotron Radiat* (2006) 13:180–94. doi:10.1107/s0909049506000550
51. Iniewski K, Chen H, Bindley G, Kuvvetli I, Budtz-Jorgensen C. Modeling charge-sharing effects in pixellated czt detectors. In: 2007 IEEE Nuclear Science Symposium Conference Record (IEEE)vol. 6; 26 October 2007 - 03 November 2007; Hilton, Hawaiian (2007). p. 4608–11.
52. Cole J, Wood J, Lopes N, Poder K, Abel R, Alatabi S, et al. Tomography of human trabecular bone with a laser-wakefield driven x-ray source. *Plasma Phys Controlled Fusion* (2015) 58:014008. doi:10.1088/0741-3335/58/1/014008
53. Mattsson A. Some characteristics of a 600 kv flash x-ray tube. *Physica Scripta* (1972) 5:99–102. doi:10.1088/0031-8949/5/1-2/017
54. Hemberg O, Otendal M, Hertz HM. Liquid-metal-jet anode x-ray tube. *Opt Eng* (2004) 43:1682–8. doi:10.1117/1.1737787
55. Brunke O, Brockdorf K, Drews S, Müller B, Donath T, Herzen J, et al. Comparison between x-ray tube-based and synchrotron radiation-based μ ct. *Dev X-ray tomography VI (Spie)* (2008) 7078:260–71. doi:10.1117/12.794789
56. Berger MJ, Hubbell JH, Seltzer SM, Chang J, Coursey JS, Sukumar R, et al. XCOM: photon cross sections database. *NIST Stand Reference Database* (1998) 8: 87–3597. doi:10.18434/T48G6X
57. Cherepy N, Sealey Z, Payne S, Swanberg E, Beck P, Schneberk D, et al. Transparent ceramic scintillators for gamma spectroscopy and mev imaging. *Hard X-ray, Gamma-ray, neutron detector Phys XVII (Spie)* (2015) 9593:74–80. doi:10.1117/12.862503
58. Ashley J, Ritchie R, Brandt W. Z 1 3 effect in the stopping power of matter for charged particles. *Phys Rev B* (1972) 5:2393–7. doi:10.1103/physrevb.5.2393
59. Wilson G, Dennison JR. Approximation of range in materials as a function of incident electron energy. *IEEE Trans Plasma Sci* (2012) 40:291–7. doi:10.1109/tps.2011.2176515
60. Derenzo S, Weber M, Bourret-Courchesne E, Klintonberg M. The quest for the ideal inorganic scintillator. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 505:111–7. doi:10.1016/s0168-9002(03)01031-3
61. Swinehart DF. The beer-lambert law. *J Chem Educ* (1962) 39:333. doi:10.1021/ed039p333
62. Andor zyla. Andor zyla datasheet (2023). Available at: <https://andor.oxinst.com/assets/uploads/products/andor-zl41-wave-specifications.pdf> (Accessed August 28, 2023).
63. Xineos. Xineos-3030hr datasheet (2022). Available at: <https://www.teledynedalsa.com/download/3b835500-052b-469a-bfaf-5aac905fb2ec/> (Accessed August 28, 2023).
64. Philipp HT, Tate MW, Shanks KS, Purohit P, Gruner SM. High dynamic range cdte mixed-mode pixel array detector (mm-pad) for kilohertz imaging of hard x-rays. *J Instrumentation* (2020) 15:P06025. doi:10.1088/1748-0221/15/06/p06025
65. Rakowski R, Golovin G, O'Neal J, Zhang J, Zhang P, Zhao B, et al. Single-shot structural analysis by high-energy x-ray diffraction using an ultrashort all-optical source. *Scientific Rep* (2017) 7:16603. doi:10.1038/s41598-017-16477-0
66. Wunderer C, Marras A, Bayer M, Glaser L, Göttlicher P, Lange S, et al. The percival soft x-ray imager. *J Instrumentation* (2014) 9:C03056. doi:10.1088/1748-0221/9/03/c03056



OPEN ACCESS

EDITED BY

Gabriella Carini,
Brookhaven National Laboratory (DOE),
United States

REVIEWED BY

Gabriele Giacomini,
Brookhaven National Laboratory (DOE),
United States
Christopher John Hall,
Australian Nuclear Science and
Technology Organisation, Australia

*CORRESPONDENCE

Heinz Graafsma,
✉ Heinz.graafsma@desy.de

[†]These authors have contributed equally
to this work

RECEIVED 14 October 2023

ACCEPTED 01 December 2023

PUBLISHED 03 January 2024

CITATION

Graafsma H, Correa J, Fridman S,
Hirseman H, Hosseini-Saber SMA,
Ignatenko A, Klugev A, Lange S, Laurus T,
Marras A, Pennicard D, Rah S, Schneider S,
Shalev OS, Stoye T, Trunk U and
Wunderer CB (2024), Detector
developments for photon science
at DESY.

Front. Phys. 11:1321541.

doi: 10.3389/fphy.2023.1321541

COPYRIGHT

© 2024 Graafsma, Correa, Fridman,
Hirseman, Hosseini-Saber, Ignatenko,
Klugev, Lange, Laurus, Marras, Pennicard,
Rah, Schneider, Shalev, Stoye, Trunk and
Wunderer. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Detector developments for photon science at DESY

Heinz Graafsma^{1,2,3*†}, Jonathan Correa^{1,2†}, Sergei Fridman^{1,2†},
Helmut Hirsemann^{1,2†}, S. M. A. Hosseini-Saber^{1,2†},
Alexandr Ignatenko^{4†}, Alexander Klugev^{1,2†}, Sabine Lange^{1,2†},
Torsten Laurus^{1,2†}, Alessandro Marras^{1,2†}, David Pennicard^{1,2†},
Seungyu Rah^{1,2†}, Sandra Schneider^{1†}, Ofir Shefer Shalev^{1†},
Thorsten Stoye^{1†}, Ulrich Trunk^{1,2†} and Cornelia B. Wunderer^{1,2†}

¹Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ²Center for Free-Electron Laser Science CFEL, Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ³Mid-Sweden University, Sundsvall, Sweden, ⁴Friedrich Schiller University Jena, Jena, Germany

The past, current and planned future developments of X-ray imagers in the Photon-Science Detector Group at DESY-Hamburg is presented. The X-ray imagers are custom developed and tailored to the different X-ray sources in Hamburg, including the storage ring PETRA III/IV; the VUV-soft X-ray free electron laser FLASH, and the European Free-Electron Laser. Each source puts different requirements on the X-ray detectors, which is described in detail, together with the technical solutions implemented.

KEYWORDS

synchrotron storage rings, free-electron lasers, X-ray imagers, integrating detectors, time-stamping detectors

Introduction

The continuous exponential improvement in X-ray sources at synchrotron storage rings [1, 2] and Free-Electron Lasers [3] means that significant improvements in X-ray imagers are needed in order to maximize the scientific output, both in quality and quantity. Keeping up with the increase in source brilliance by an order of magnitude every 3 years, is, however, a major challenge. To put this into perspective, to develop a new detector from idea to installation at the experimental station takes roughly 10 years when the concept is new. This means that the source brilliance increases by 3 orders of magnitude during the development time of the detector. In addition, the large variety of experimental techniques, especially at storage rings [2], results in very diverse, and often conflicting detector requirements, so that different detector systems need to be developed.

In Hamburg the situation is even more challenging due to very different photon sources. PETRA III is a 6 GeV storage ring catering mainly to the medium to high photon energy synchrotron community. This is very different to the FLASH facility, which is a 1.35 GeV Free-Electron Laser using superconducting accelerator technology. It produces photons in the VUV to soft X-ray regime in a burst mode operation. It produces pulse trains up to 800 micro-seconds long, with a train repetition frequency of 10 Hz and up to 500 pulses in each train. The European X-ray Free-Electron Laser uses the same superconducting accelerator technology as FLASH, but runs at 17 GeV and produces photons in the soft, medium and hard X-ray regime. The pulse trains are 600 micro-seconds long, are repeated with a 10 Hz frequency, and can contain up to 2,700 pulses.

There are many fundamental differences between a storage ring and a Free-Electron Laser, both from an experimental and a detector point of view. Storage rings can often be regarded as continuous sources and data is mostly collected with many consecutive pulses. Free-Electron Lasers on the other hand are mostly used as pulsed sources, where an image is recorded for every single pulse. This is both useful, since single FEL pulses are often strong enough to produce a statistically significant image, and also required, since a single pulse often destroys the sample. This has many consequences for the detectors at Free-Electron Lasers. First, the photon-counting technique, widely applied at storage rings, cannot be used, since all photons arrive at the same time, and charge integrating front-ends are required instead. At storage rings, the desired dynamic range can be obtained by either extended exposure times, or multiple exposures. At Free-Electron Lasers, a single shot image has to be able to cover the entire dynamic range. For Free-Electron Lasers based on superconducting accelerators, frame rates of up to 4.5 MHz, corresponding to the pulse rate of the source, are additionally required. Even though the time structure of the European XFEL and FLASH are comparable, their photon energies are so different that conceptually different detectors are required.

In the following section we will present the detector systems that we have developed specifically for the European XFEL and FLASH, including ongoing improvements and additional systems under construction. Based on these developments and other previous projects, we are developing two new systems, CoRDIA and TEMPUS, which are tailored for the upgraded source PETRA-IV. The developments and current status will be presented in the subsequent section. In the final section we will discuss the future directions of our developments.

AGIPD detectors for the European XFEL

With the construction of the European X-ray Free-Electron Laser, conceptually new X-ray imagers were required [4]. A few of the most challenging requirements were that in a single shot the entire dynamic range from single photons to more than 10^4 photons per pixel had to be covered, with a frame rate of 4.5 MHz. Additionally, as many frames as possible needed to be recorded during the 600 micro-second long pulse trains. Following a call for proposals three projects were selected; The Large Pixel Detector (LPD) [5], the DePMOS Sensor with Signal Compression (DSSC) [6], and the Adaptive Gain Integrating Pixel Detector, AGIPD [7]. Each system tackled the challenges with a different concept and targeted different applications [4]. Here only the AGIPD system, developed under the leadership of the DESY photon science detector group, will be described. The LPD system was developed by the Rutherford-Appleton Laboratory in the United Kingdom. The DSSC system was developed by a larger consortium, initially under the leadership of the Max-Planck Semiconductor Laboratory in Munich, and later under the leadership of the European XFEL itself.

The AGIPD system was developed by a consortium of DESY, PSI, University of Hamburg and University of Bonn. The AGIPD ASIC employs a charge integrating frontend with adaptive gain, where every pixel automatically and autonomously adapts its gain to

the number of incoming photons [7]. Every pixel starts in high gain mode, with a small feedback capacitor, allowing for single photons to be distinguished from the noise. As soon as the voltage over the feedback capacitor reaches a preset threshold, a second, larger, feedback capacitor is added reducing the gain and extending the dynamic range. For very strong signals a third, even larger, feedback capacitor is added, resulting in a low gain and corresponding dynamic range of 10^4 photons of 12 keV. At the end of the integration time the final voltage over the feedback capacitors is stored in an analog memory, together with the information of the final gain setting, i.e., which feedback capacitors were used. In this manner the required dynamic range is achieved. In pixel frame storage allows for frame rates up to 4.5 MHz, compatible with the highest pulse rate of the European XFEL. During the 99.4 millisecond interval between pulse trains, the analogue memory is read out and digitized by external ADCs and the digital information is transmitted over 10G links to the DAQ system. The pixel concept is given in Figure 1. The main parameters of the AGIPD readout ASIC are given in Table 1.

An additional complication for building the system was the requirement of maximum flexibility. Two 1-million-pixel systems operate in a moderate vacuum of $\approx 10^{-6}$ mbar and consist of four individually movable quadrants that protrude into the sample chamber. Figure 2 shows a picture of the systems in operation at the MID and SPB/SFX instruments of the European XFEL.

The 1M-AGIPD systems have been in user operation at the SPB/SFX and MID stations since 2017 and 2019 respectively, and together produced the vast majority of the scientific user publications of the European XFEL.

Since commissioning, various improvements have been made to the system [8].

One improvement concerned the so-called late gain-switching. Wrong signal values occur when a pixel changes its gain too close to the end of the integration window, so that the output signal of the pre-amp is not yet settled when it is stored into the storage cell matrix. The proper solution to the problem is to block gain switching within the last 40 ns of the integration window, but this would require a significant redesign of the ASIC. Instead, two mitigation strategies were implemented. For experiments that do not require single photon sensitivity, automatic gain switching is disabled and the system is operated in fixed medium gain. For experiments that do not require the highest repetition rate of 4.5 MHz, extended integration times can be used. Studies have shown that extending the integration time reduces the occurrence of late gain switching in the transition region by at least three orders of magnitude.

Both MID and SPB/SFX systems will undergo an upgrade, where all the front-end modules will be replaced with new ones equipped with AGIPD1.2 ASICs. This new version fixes an issue of poor gain bit separation in the AGIPD1.1 ASICs, which resulted in an ambiguity whether medium or low gain was used.

We are currently constructing two more AGIPD cameras for the European XFEL. The first one is a 4 megapixel version for the Serial Femto-second Crystallography (SFX) user consortium, which will also be installed at the SPB/SFX instrument. This in-vacuum camera consists of two-halves, which can move independently, by 400 mm along the beam axis and ± 15 mm perpendicular to it. The frontend modules will be equipped with AGIPD1.2 ASICs and silicon sensors. A CAD rendering of the 4M-AGIPD is given in Figure 3.

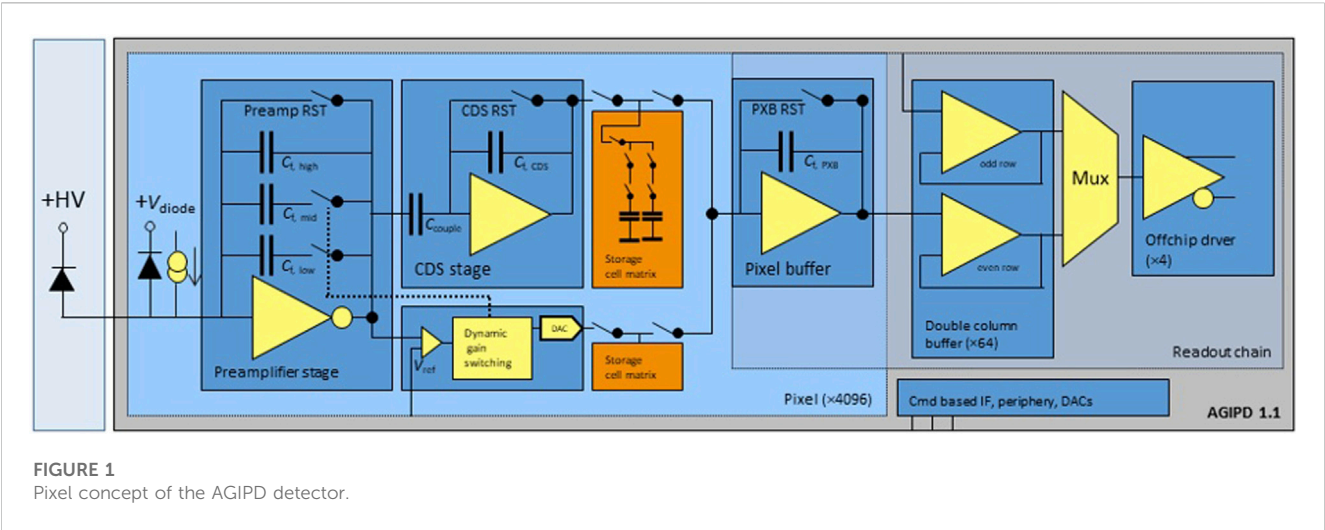
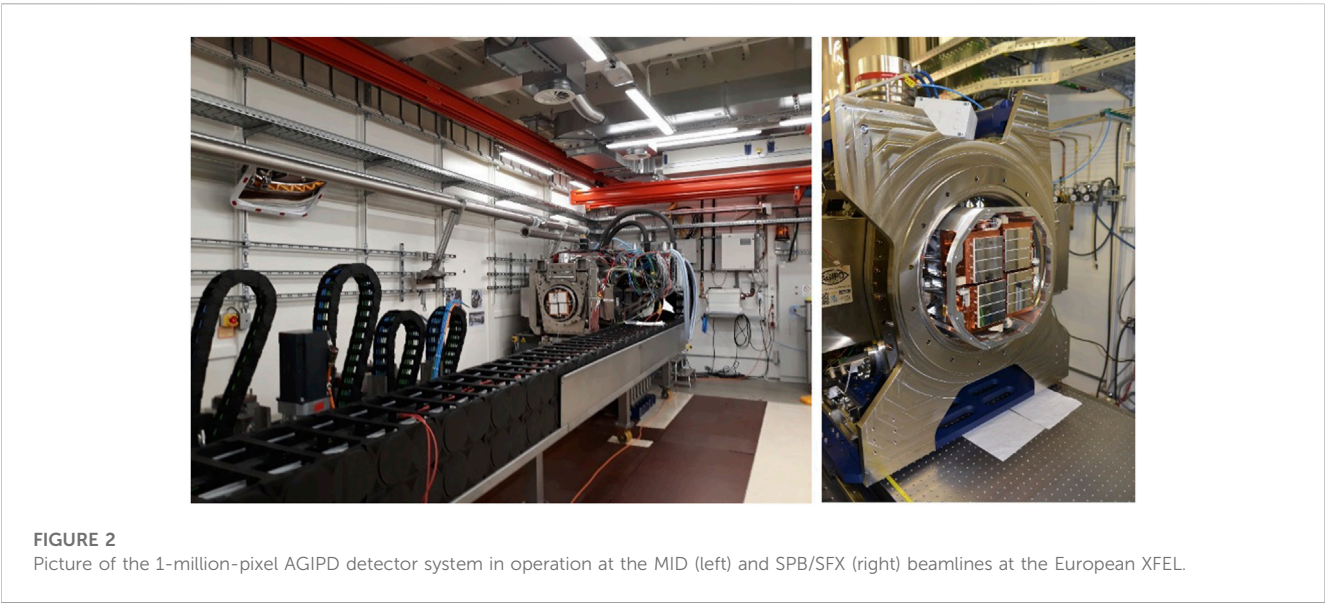
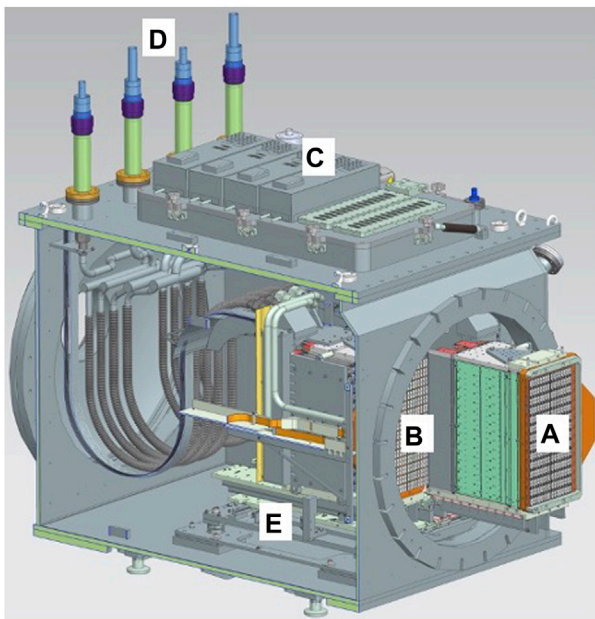


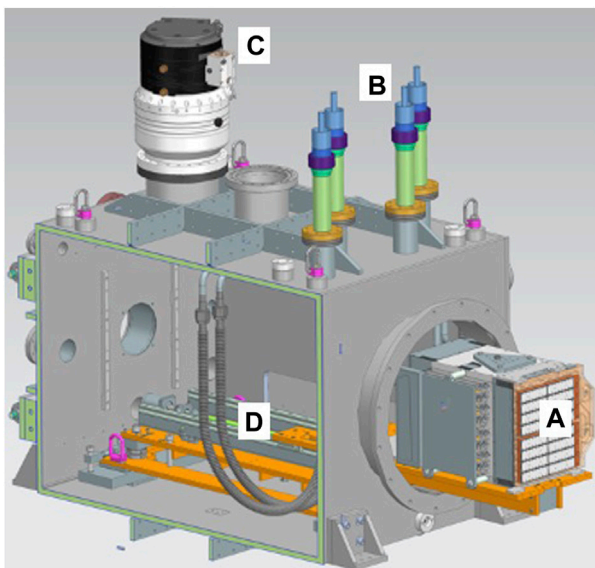
TABLE 1 Main parameters of the AGIPD readout ASIC.

Parameter	Value	Comment
Working principle	Charge integrating	Required for use at FELs
Adaptive gain	3 stages	Gain adaptively selected per pixel by incoming signal
Frame rate	4.5 MHz	XFEL bunch frequency
Number of frames	352 during burst	3520 per second
Pixel size	200 μm x 200 μm	Dictated by number of frames to be stored during burst
Photon energy	>6 keV	6 keV is 5 σ above noise
Noise	300 e^-	1.2 keV in silicon
Dynamic Range	$3 \times 10^7 e^-$	10k photons at 12 keV in silicon



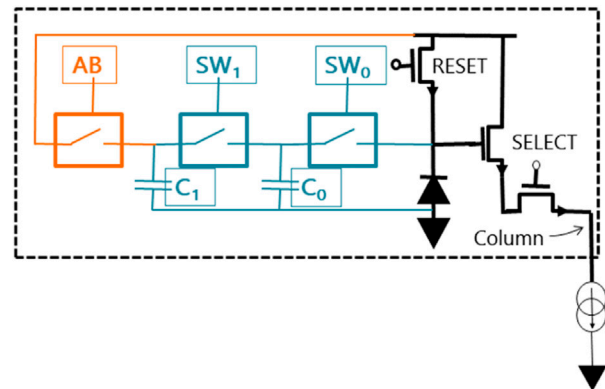
**FIGURE 3**

CAD rendering of the AGIPD-4M system in vacuum for the SFX user consortium at the European XFEL. (A) right half of the detector in forward position; (B) left half of detector in retracted position; (C) electrical feed-through connections for signal, power and control; (D) cooling inlets and outlets for frontend modules and readout boards; (E) translation stages.

**FIGURE 4**

CAD rendering of the AGIPD-1M system in vacuum for the HIBEF user consortium at the European XFEL. (A) frontend modules in forward position; (B) cooling inlets and outlets for frontend and readout boards; (C) turbo pump; (D) translation stage.

The second system is a 1 megapixel version with high-Z sensors for the Helmholtz International Beamline for Extreme Fields (HIBEF), which will be installed at the HED instrument. High-Z

**FIGURE 5**

Pixel concept of the PERCIVAL sensor. The standard CMOS 3T pixel is enhanced by three switches and two capacitors. All three switches are biased at ~ 0.7 V, enabling lateral overflow as the diode, then C0, then C1 reach low voltages. While SW0 and SW1 enable adding in more capacitance as needed, the third "AntiBleeding" switch provides controlled overflow for even larger charges, a preventive measure to avoid 'blooming' effects.

sensor material is required for high photon-energies of up to 25 keV. Since the original AGIPD is hole collecting an electron collecting version of the AGIPD ASIC had to be developed [9]. The camera will initially be equipped with 500 μm silicon sensors and hole-collecting AGIPD1.2 ASICs and will later be upgraded to front-end modules with the electron collecting AGIPD1.3 ASICs and high-Z sensors. The camera is in vacuum and can be moved along the beam over 550 mm. Special care was taken for electromagnetic compatibility (EMC), required by the very high pulsed electro-magnetic fields in the experiments. A CAD rendering of the system is given in Figure 4.

To enable the long in-vacuum travel ranges along the beam axis of the two new cameras, the entire readout electronics were revised. The core idea was to combine all the functionality previously distributed over many different boards on a single readout board, which is located directly behind the frontend module in vacuum. This drastically increases modularity and improves reliability, which was compromised by the large number of connectors in the first generation systems.

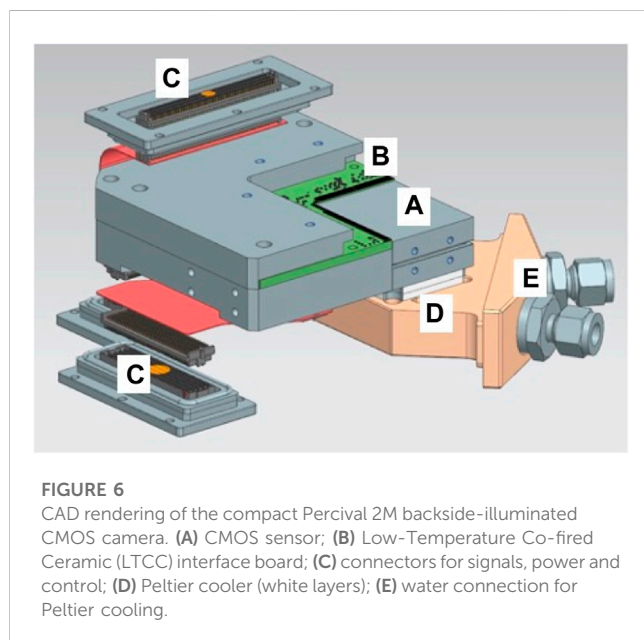
A 0.5 megapixel prototype camera system based on the new readout electronics, often referred to as mini-half, is in operation at HED instrument as part of the dDAC (dynamic Diamond Anvil Cell) compression platform [10]. Numerous user experiments have already been successfully completed with the prototype. Installation and commissioning of both new AGIPD cameras, the AGIPD-4M for the SFX user consortium and the AGIPD-1M for the HIBEF user consortium, are planned for 2024.

PERCIVAL soft X-ray CMOS imagers for FLASH

The Free-Electron Laser in Hamburg (FLASH) produces photons in the XUV and soft X-ray range. Developing imagers for soft X-rays is particularly hard, since in order to get the low-energy photons into the sensitive volume of the sensor, entrance

TABLE 2 Main parameters of the Percival imager.

Parameter	Value	Comment
Frame rate	300 Hz	83 Hz in current version
Readout mode	Rolling shutter	
Pixel size	27 μm x 27 μm	
Sensor size	1484 x 1408 pixels	4 cm x 4 cm, stitched monolithic
Main photon energy	250 eV–1 keV	Photons below 100 eV can be detected, 250 eV is 4 σ above noise, higher energies can be detected with lower quantum efficiency
Noise	14 e ⁻	In optimized very high gain mode
Dynamic range (“full well”)	5.7 ke ⁻ /3.6 Me ⁻ per pixel per frame	In very high gain/low gain mode. Corresponds to 1–50k photons @ 250 eV
Dead area in focal plane	0	Continuous sensitive area



windows and passivation layer thicknesses have to be kept to a minimum. As for photons below 250 eV, attenuation lengths fall significantly below 100 nm, the dead layer at the entrance should be on the order of 10 nm or less. As each photon does not create many electron-hole pairs (e.g., 69e⁻ for 250 eV in Si), in addition care must be taken to have favorable electric fields at the surface, and the readout electronics must have low noise, preferably on the order of 10–15 electrons or less to avoid false positives in a megapixel-scale system. In order to meet the requirements for imaging experiments at FLASH, a backside-illuminated CMOS sensor, PERCIVAL, was custom designed by Rutherford Appleton Laboratory. Like AGIPD, every pixel can adapt its gain according to the incoming signal strength, using an in-pixel lateral overflow that auto-selects the capacitance used depending on the charge deposited in the pixel [11]. During pixel readout, thresholding is used to determine which capacitor(s) were needed and thus which gain was used, the corresponding voltage level is digitized on-chip, and ADC output

and gain are transmitted via LVDS. The pixel concept is given in Figure 5. The main parameters of the system are given in Table 2.

It is the combination of these features that makes Percival an especially well-suited detector for soft X-ray scattering experiments. Detailed description of the system and its performance can be found in [11–18].

Even though the current first version of the sensor has a number of shortcomings [16] it has been used in a number of proof-of-principle user experiments, including Holography and XPCS of magnetic skyrmions at Petra III's soft X-ray beamline P04, and single-shot Ptychography at FLASH. These exploratory experiments showed the great potential of the large dynamic range and high frame rates of Percival [18]. The shortcomings of the first version sensor—crosstalk between ADC and digital control lines, and insufficient ground connections for the pixel array that resulted in uneven biasing—are understood and are corrected in a second version of the sensor which was recently submitted to the foundry. In addition, a more compact version of the camera head has been designed; a CAD rendering is given in Figure 6.

New developments

Thanks to the technological developments in accelerator technology, synchrotron storage rings world-wide are currently being upgraded [19] giving 2 or more orders of magnitude improvement in the horizontal emittance and overall source brilliance. At DESY the plan is to upgrade the accelerator PETRA III to PETRA IV, which will give an improvement of the source brilliance by a factor 500 at 10 keV and a factor of 1,000 at 60 keV photon energies [2]. This creates many new scientific possibilities, but also requires an upgrade of the current X-ray imagers. Even though an increased source brilliance does not always translate to the same increase in the number of photons to be detected, it is clear that the X-ray imagers need to have both higher flux handling capabilities (count rates), and higher frame rates. As an example, there will be enough incident photons to obtain statistically significant ptychographic images by raster scanning entire microelectronic chips with more than 100 kHz frame rates. This makes it possible, for example, to investigate a

TABLE 3 Target specifications for CoRDIA.

Parameter	Value	Comment
Frame rate	>135 kHz continuous	Revolution frequency of PETRA IV
Readout dead time	0 sec	Continuous RW
Reset blind time	0.6 msec	10% of integration time at highest frame rate
Pixel size	110 mm x 110 mm	Compatible with 55 mm pitch of Medipix systems
Photon energy	0.5–150 keV	Using different sensors
Noise	1 keV	Giving single photon sensitivity
Dynamic range	5×10^3 photons/pixel/image	12 keV in silicon
Dead area in focal plane	Minimal	Using TSV technology

silicon chip for either defects or unwanted structures within a reasonable period of days. Another example is 3D tomography of batteries on millisecond timescales during charging and discharging, in order to follow ongoing processes at the microscopic level. All these and many more examples require X-ray imagers with a large dynamic range and frame rates of more than 100 kHz.

As described above, the European XFEL operates in a burst mode, where all X-ray pulses arrive in less than 1% of the time, the remaining 99% of the time are needed for the superconducting accelerator to cool down again. For many reasons this mode of operation is not ideal for the experiments and more evenly distributed pulses would be beneficial. Studies and developments are underway for the accelerator to stretch the pulse trains, and ultimately to distribute the pulses evenly over time, which is called continuous wave or CW operation. From a detector point of view, PETRA IV and a CW-XFEL would become more similar than PETRA III and European XFEL are at the moment. However, a 100 kHz frame rate is not sufficient for a CW-XFEL operation, where pulse frequencies of 1 MHz are expected. Since PETRA IV will almost certainly be realized before an upgrade of the European XFEL towards more CW operation, we have started the development of an imager driven by the specifications of PETRA IV. This development, CoRDIA, will be described in the next section. PETRA IV will have a large variety of experiments and techniques and not all needs will be covered by the capabilities of CoRDIA. Specifically, there will be needs for photon counting based systems, particularly in experiments with moderate X-ray flux but extremely high time resolution requirements. For these applications a system called TEMPUS is being developed, which is based on the TimePix-4 readout chip and can operate in either photon counting or time stamping modes. This system will be described in the subsequent section.

The continuous readout digitizing imaging array: CoRDIA

For the diffraction limited storage ring PETRA IV, a dedicated development was started at DESY building on the successful development of the AGIPD system. Some of the target specifications are given in Table 3.

The CoRDIA (Continuous Readout Digitising Imager Array) project is a collaboration between DESY and University of Bonn. Like AGIPD, CoRDIA is a hybrid pixel detector in which read-out ASICs will be bump bonded to the semiconductor sensor. CoRDIA utilizes the same architecture for the analogue front-end in the ASIC as AGIPD, a charge-integrating pre-amplifier with adaptive gain. This allows for a dynamic range from single up to thousands of photons. To provide continuous operation without any dead time beyond the reset phase of the front-end, the readout architecture had to be developed from scratch. Compared to the systems developed for the European XFEL, frame recording is slower by a factor of ≈ 30 , while readout frame rate needs to be faster by 2 orders of magnitude. The latter rules out analogue signal transmission due to the poor performance expected, while the input rate of ≈ 150 kHz conveniently allows for the digitization of multiple pixels with a single ADC. In such a scenario, memory is only needed on the level of a 2-stage S&H (sample and hold) per pixel, in order to allow simultaneous frame recording and digitization of the previous frame. As a benefit, pixel size can be shrunk considerably.

An analog stage front-end (FE) is used to collect charge generated by photons in a bump-bonded sensor. The FE circuit is compatible with electron-collecting sensors, so that a p-doped silicon sensor could be used for the main energy range (around 12 keV), while high-Z sensors could be used for harder X-ray imaging. The circuit uses the same adaptive-gain approach as AGIPD to modulate the signal amplification to the incoming photon flux. The response of the analog circuit to a charge injection, emulating photon fluxes of variable intensity (with modulation of the gain), has been tested on a test chip, at a frequency of 150k frame/s, i.e., exceeding our frame-rate requirements [20].

While a new image is stored as an analog signal in one memory circuits, the previous image stored in the other memory circuit is retrieved, processed by a correlated-double-sampling (CDS) circuit, and digitized. As soon as digital bits are produced by the ADC, they are packaged, encoded, and transmitted off-chip by a high-speed driver.

To reach a suitable trade-off between frame rate, layout size and power dissipation, a successive approximation register (SAR) ADC

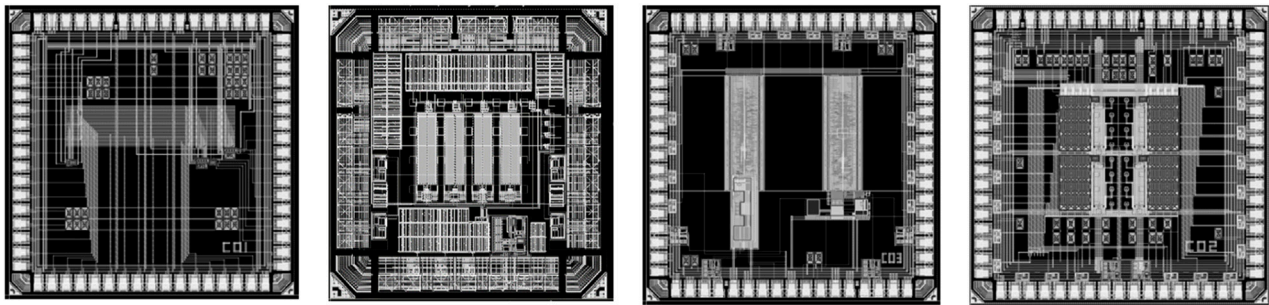


FIGURE 7

MPW prototypes including standalone circuits for validation. From left to right: CoRDIA01 - front-end and CDS, HIS-ADC01 - SAR ADC, CoRDIA03 - multi-gigabit transmitter, CoRDIA02 "superpixel" layout including front-end, CDS and ADC.

was designed, capable of digitizing 11 bits at 2.5 MS/s while reaching 10 ENOBs (effective number of bits net resolution). While its layout is reasonably compact, it is too large to fit within our target pixel size ($110\ \mu\text{m} \times 110\ \mu\text{m}$); on the other hand, the ADC sampling speed is high enough that the signal from 16 FE circuits can sequentially be processed by the same ADC without impairing our frame rate requirements. We defined our basic structure as a $440\ \mu\text{m} \times 440\ \mu\text{m}$ pixel block ("super pixel"), each including 16 analog FE circuits, a CDS and ADC circuit serving sequentially the 16 pixels, and 1/128 of the digital readout circuit.

Space has been reserved in the pixel block layout for through silicon via (TSV) landing pads that could be used for vertical integration as an alternative to traditional periphery wirebonding, so that blind silicon areas could be minimized. The inputs of the FE circuits are distributed evenly to pads on the "super pixel" surface, allowing the ASIC to be bump-bonded to a $110\ \mu\text{m}$ -pixel-pitch sensor.

The readout circuit has been developed by NIKHEF, and implements the physical layer defined in the IEEE 802.3ae standard. The circuit consists of the Physical Coding Sublayer (digital) part, preparing the data for high-speed throughput (scrambling to ensure DC balance, and 64-to-66-bit encoding), and fast driver (Gigabit Wire Transmitter) able to stream out data at several GHz. The circuit has already been used in the Timepix4 ASIC (discussed later) and has been successfully tested by the Medipix4 collaboration [21] at 5.12 GHz. In our architecture, each block of 2k pixels is streamed out by one such circuits: since each pixel output is digitized with less than 16 bits, the 5.12 GHz frequency is enough to cover our frame rate requirements.

Several MPW prototypes have been designed and manufactured in TSMC 65 nm technology, see Figure 7, to validate the circuits described above as standalone blocks, up to now confirming expected performances. We acknowledge CERN and the RD53 collaboration for developing several IO blocks that were used in the design of the MPWs. A further ASIC prototype, embedding some of the blocks in a "superpixel" layout, is currently under test. We aim at developing the first version of the imager so that it will be ready for the first experiments at the PETRA-IV upgraded ring, expected in 2029.

The TimePix-4 based Edgeless Multi-Purpose Sensor: TEMPUS

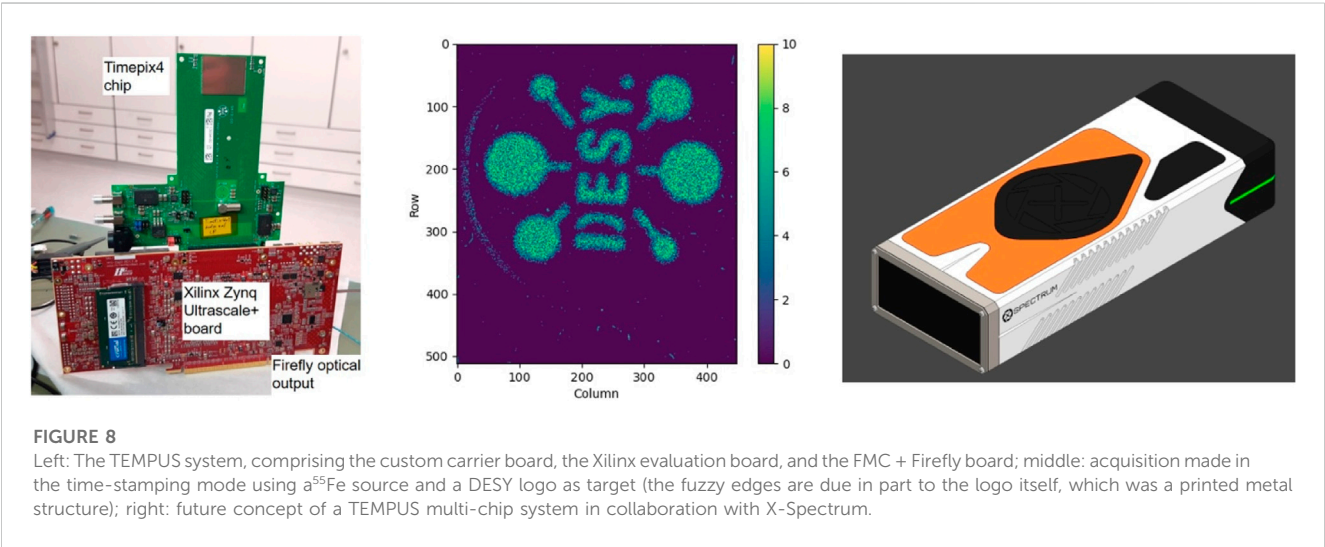
A large number of different experiments at PETRAIII have benefited from the use of one of the detectors developed at DESY, the LAMBDA detector [22]. Based on the Medipix3 photon counting readout chip [23], it has become a workhorse for many different experimental stations. The combination of a small pixel size, of $55\ \mu\text{m}$, with relatively high readout rates have made it attractive for X-ray imaging at synchrotrons. Moreover, the use of high-Z sensors (mainly CdTe and GaAs) and the option of tiling up several ASICs covering a large area, has pushed the project to a full commercialization of the detector by the DESY spin-off company X-Spectrum GmbH [24].

In that context, a new readout chip has been recently developed by the Medipix4 collaboration using 65 nm technology, Timepix4 [21]. This chip, almost 4 times as large as the previous generation while keeping the same $55\ \mu\text{m}$ pixel pitch, offers an array of 448×512 pixels. One crucial difference with this new device is the combination of two distinct operation modes, whose specifications can be found in Table 4. Firstly, there is a Photon Counting mode, similar to the Medipix3, but it improves frame rate up to 40 kfps, and maximum counting rates by an order of magnitude compared to its predecessor. Secondly, there is a time-stamping mode, where for each photon detected a packet of data is sent out containing the pixel coordinates, a timestamp and an energy measurement; this can provide extremely high time resolution in experiments with moderate fluxes. Compared to the previous generation, the Timepix3 readout chip [25], the hit rate has been increased by an order of magnitude and the time resolution can achieve a value of 192 ps with suitable sensors.

To reach these high frame and event rates, the chip has 16 high speed readout links developed by Nikhef; the design of these links is being re-used in the CoRDIA detector. In the chip's original specifications, these were targeted at 5.12 Gbps, but during development a 10.24 Gbit/s mode was also developed. Since this speed has not yet been fully demonstrated, Table 4 conservatively lists the frame and event rates with 5.12 Gbps readout, but the readout system is designed to be compatible with 10.24 Gbps readout.

TABLE 4 Timepix4 chip readout specifications.

Parameter		Value
Sensitive area		6.94 cm ²
Pixel size		55 x 55 μm ²
Pixel arrangement		512 x 448
Photon Counting Mode	Max count rate	2 x 10 ⁶ hits/pixel/s
	Max frame rate (@ 5.12 Gbps/link)	40 kfps @ 8-bits depth
Time-stamping Mode	Chip readout rate (@ 5.12 Gbps/link)	1.24 x 10 ⁹ events/s
	Corresponding average pixel rate	5.4 kHz/pixel
	Time resolution	192 ps
	Energy resolution	<1 keV
Readout bandwidth @ 10.24 Gbps/link		~163 Gbps



At DESY, the development of a readout system based on this new ASIC has started: TEMPUS, the Timepix4-based Edgeless Multi-Purpose Sensor. The initial system developed is a single-chip system, consisting of a custom chip carrier board connected to an off-the-shelf Xilinx evaluation board—see Figure 8 left. The board has a Zynq system-on-chip, incorporating a CPU for high-level control functions, and an FPGA fabric for low-level interfacing with Timepix4. In particular, this provides high-speed transceivers for receiving the high amount of data from the chip and sending serialized data out over 100 Gigabit Ethernet links. To send the data from the Xilinx evaluation board to the outside world, an FMC + FireFly daughter board with up to 6 Samtec FireFly links is used to reach the 160 Gbps data bandwidth. To make effective use of Timepix4, TEMPUS uses high-data-rate board designs and firmware.

The chip can be configured for sending the data out using the control link, as a debugging tool. This, although limited to a

much lower hit rate of ~5,000 hits/s, has already allowed us to obtain a number of images of background radiation and measurements of X-ray sources—see Figure 8 middle. Operating in timestamping mode, Timepix4 offers nanosecond timing while detecting X-rays with a silicon sensor.

There are a range of anticipated applications in photon science. Photon counting detectors nowadays are widely-used in X-ray diffraction experiments, and Timepix4’s high frame rate in photon counting mode will enable higher-speed measurements, for example, when rapidly raster-scanning large samples. Time-stamping mode will enable much higher time resolution in X-ray diffraction experiments with moderate flux; for example, experiments such as XPCS and XCCA studying dynamics of proteins in solution. Furthermore, Timepix4 can potentially replace detectors such as delay lines or APDs in experiments measuring emission of photoelectrons or nuclear fluorescence from a sample excited by an X-ray pulse.

Finally, the Timepix4 chip has been designed to be fully TSV compatible. As described in the next section, this technology will allow powering, control and data read out from the back of the chip, eliminating the use of wire-bonds. This 4-side buttable chip will make it possible to build multi-chip systems with reduced gaps between the ASICs, therefore minimizing the dead area in the imaging plane. DESY, together with X-Spectrum, has already started the development of a multi-chip module—see [Figure 8](#) right—which in turn can be tiled to create multi-megapixel systems.

Near future developments

In addition to the above-described developments of new systems, there are also a number of technological and component developments in the DESY photon science detector group. One of them is focused on high-Z sensors for higher photon energies. This is a very important energy range for PETRA and the European XFEL, both being high energy machines. Over the past decades DESY together with partners have successfully developed large area GaAs sensors with acceptable quality [26]. We will continue these activities, further improving the performance and contributing to establish multiple suppliers. GaAs shows a number of advantages over other high-Z sensor materials like Cd(Zn)Te, one being the much shorter range of fluorescent photons generated by the sensor itself. The shorter range leads to less blurring and makes charge summing techniques more efficient [27]. However, for photon energies above 50 keV the efficiency of GaAs sensors drops to unacceptable levels, and higher Z sensors are required. High-flux CdZnTe is the most promising material also for experiments with high instantaneous intensities, e.g., at the European XFEL [28]. We at DESY are participating in the world-wide efforts to further improve the quality, availability and size of this material.

At the other end of the energy spectrum, we are involved with the development of inverted-Low Gain Avalanche Diodes (i-LGADs) with thin entrance windows. The high energy physics community has initiated the developments of LGADs mainly for their timing performance. The photon science community is now pushing the development of inverted LGADs in order to boost the weak signals produced by low energy photons [29]. The enormous advantage of such sensors becoming available is, that the entire detection chain developed over many decades for the intermediate photon energies can be used also for the low photon energies. Even though imaging at DESY is mainly done in the medium X-ray energy range, soft X-ray imaging, especially at FLASH, remains a scientifically important area.

Another technological development that will bring significant improvements to the X-ray imaging experiments is the use of through silicon vias (TSVs). TSVs with redistribution layers will almost completely eliminate the dead areas required for wire bond connections in the imaging plane of large multi-module imagers. This avoids the loss of important information, in particular for non-repeatable scattering experiments. One example is high pressure experiments on small unit cell compounds using dynamic diamond anvil cells, where only a

few reflections are produced and some fall in the gap between modules. At DESY we successfully developed TSV based modules with the Medipix3 readout chip [30]. However, manufacturing time was too long and the yield too low for it to be a viable option to produce multi-module systems. One of the reasons was the non-ideal layout of the metal stack in Medipix3; the TSV landing pad consisted of a single layer with a mesh structure, and over-etching could result in making contact with structures in the next layer. In Timepix4 and other chips, multiple metal layers are used for landing pads, to avoid this problem. Another reason was the developmental nature of the chip processing we were using. TSV technology has been around for a long time, but has not been readily available to the scientific community. An important reason is that the required volumes are often too small for foundries to be interested. It is hoped that a concerted effort by the larger photon science detector community might change this situation.

Conclusion

Over last decades, a number of X-ray imagers have successfully been custom developed by DESY for the different photon sources operating in the Hamburg area. Due to the continual improvement in source performance these systems are continuously upgraded, and new systems are developed. For the future source PETRA IV, DESY is focusing its developments on a high frame rate integrating imager, CORDIA, and a high time resolution photon-counting imager TEMPUS.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HG: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing—original draft, Writing—review and editing. JC: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Supervision, Writing—original draft, Writing—review and editing. SF: Conceptualization, Data curation, Software, Validation, Writing—review and editing. HH: Conceptualization, Methodology, Writing—review and editing. SH-S: Data curation, Formal Analysis, Investigation, Software, Validation, Writing—review and editing. AI: Data curation, Formal Analysis, Investigation, Software, Writing—review and editing. AK: Conceptualization, Data curation, Formal Analysis, Investigation, Validation, Writing—review and editing. SL: Conceptualization, Methodology, Writing—review and editing. TL: Conceptualization, Data curation, Formal Analysis, Investigation, Supervision, Validation, Writing—original draft, Writing—review and editing. AM: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation,

Writing—original draft, Writing—review and editing. DP: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Writing—review and editing. SR: Conceptualization, Methodology, Writing—review and editing. SS: Methodology, Validation, Writing—review and editing. OS: Conceptualization, Methodology, Software, Validation, Writing—review and editing. TS: Conceptualization, Methodology, Writing—review and editing. UT: Conceptualization, Formal Analysis, Investigation, Methodology, Writing—review and editing. CW: Conceptualization, Data curation, Formal Analysis, Investigation, Supervision, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Acknowledges funding from QuCoLiMa (Quantum Cooperativity of Light and Matter) CRC-TR 306 Project A03.

References

- Raimondi P, Benabderrahmane C, Berkvens P, Biasci JC, Borowiec P, Bouteille J-F, et al. The extremely brilliant source storage ring of the European synchrotron radiation facility. *Commun Phys* (2023) 6:82. doi:10.1038/s42005-023-01195-z
- Schroer CG, Wille HC, Seck OH, Bagschik K, Schulte-Schrepping H, Tischer M, et al. The synchrotron radiation source PETRA III and its future ultra-low-emittance upgrade PETRA IV. *Eur Phys J Plus* (2022) 137:1312. doi:10.1140/epjp/s13360-022-03517-6
- Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, et al. A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nat Photon* (2020) 14:391–7. doi:10.1038/s41566-020-0607-z
- Graafsma H. Requirements for and development of 2 dimensional X-ray detectors for the European X-ray free electron laser in Hamburg. *J Instrum* (2009) 4:P12011. doi:10.1088/1748-0221/4/12/P12011
- Hart M, Angelsen C, Burge S, Coughlan L, Halsall R, Koch A, et al. Development of the LPD, a high dynamic range pixel detector for the European XFEL. In: 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC); 27 October 2012 – 03 November 2012; Anaheim, CA, USA. IEEE (2012). p. 534–7. doi:10.1109/NSSMIC.2012.6551165
- Porro M, Andricke L, Aschauer S, Bayer M, Becker J, Bombelli L, et al. Development of the DEPFET sensor with signal compression: a large format X-ray imager with megapixel readout capability for the European XFEL. *IEEE Trans Nucl Sci* (2012) 59(6):3339–51. doi:10.1109/TNS.2012.2217755
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Rad* (2019) 26:74–82. doi:10.1107/s1600577518016077
- Klačkova I, Sztuk-Dambietz J, Graafsma H, Hosseini-Saber SMA, Klyuev A, Laurus T, et al. Five years operation experience with the AGIPD detectors at the European XFEL. *Proc SPIE* (2023) 23 12581. doi:10.1117/12.2666402
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Graafsma H, et al. Megapixels @ Megahertz – the AGIPD high-speed cameras for the European XFEL. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 942:162324. doi:10.1016/j.nima.2019.06.065
- Husband RJ, Strohm C, Appel K, Ball OB, Briggs R, Buchen J, et al. A MHz X-ray diffraction set-up for dynamic compression experiments in the diamond anvil cell. *J Synchrotron Rad* (2023) 30:671–85. doi:10.1107/s1600577523003910
- Marsh B, Das D, Sedgwick I, Turchetta R, Bayer M, Correa J, et al. PERCIVAL: the design and characterisation of a CMOS image sensor for direct detection of low-energy X-rays. In: 2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); 08–15 November 2014; Seattle, WA, USA. IEEE (2014). p. 1–4. doi:10.1109/NSSMIC.2014.7431113
- Wunderer CB, Marras A, Bayer M, Correa J, Lange S, Shevyakov I, et al. Percival: an international collaboration to develop a MAPS-based soft X-ray imager. *Synchrotron Radiat News* (2014) 27(27):30–4. doi:10.1080/08940886.2014.930808
- Wunderer CB, Marras A, Bayer M, Glaser L, Göttlicher P, Lange S, et al. The PERCIVAL soft X-ray imager. *J Instrum* (2014) 9:C03056. doi:10.1088/1748-0221/9/03/c03056
- Correa J, Bayer M, Göttlicher P, Lange S, Marras A, Niemann M, et al. Characterisation of a PERCIVAL monolithic active pixel prototype using synchrotron radiation. *J Instrum* (2016a) 11:C02090. doi:10.1088/1748-0221/11/02/c02090
- Correa J, Marras A, Wunderer CB, Göttlicher P, Lange S, Reza S, et al. On the charge collection efficiency of the PERCIVAL detector. *J Instrum* (2016b) 11:C12032. doi:10.1088/1748-0221/11/12/c12032
- Wunderer CB, Correa J, Marras A, Aplin S, Boitrelle B, Goettlicher P, et al. The Percival 2-Megapixel monolithic active pixel imager. *J Instrum* (2019) 14:C01006. doi:10.1088/1748-0221/14/01/c01006
- Marras A, Correa J, Lange S, Vardanyan V, Gerhardt T, Kuhn M, et al. Characterization of the Percival detector with soft X-rays. *J Synchrotron Rad* (2021) 28:131–45. doi:10.1107/s1600577520013958
- Correa Magdalena J, Mehrjoo M, Battistelli R, Lehmkuhler F, Marras A, Wunderer C, et al. The PERCIVAL detector: first user experiments. *J Synchrotron Rad* (2023) 1(1):242–50. doi:10.1107/S1600577522010347
- Einfeld D, Plesko M, Schaper J. First multi-bend achromat lattice consideration. *J Synchrotron Rad* (2014) 21:856–61. doi:10.1107/S160057751401193X
- Marras A, Klugev A, Lange S, Laurus T, Pennicard D, Trunk U, et al. (2023) Development of CoRDIA: an imaging detector for next-generation photon science X-ray sources. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* 1047:167814. doi:10.1016/j.nima.2022.167814
- Llopart X, Alozy J, Ballabriga R, Campbell M, Casanova R, Gromov V, et al. Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning. *J Instrum* (2022) 17, C01044. doi:10.1088/1748-0221/17/01/C01044
- Pennicard D, Lange S, Smoljanin S, Becker J, Hirsemann H, Eppele M, et al. Development of LAMBDA: large area medipix-based detector array. *J Instrum* (2011) 6: C11009. doi:10.1088/1748-0221/6/11/C11009
- Ballabriga R, Campbell M, Heijne EHM, Llopart X, Tlustos L. The Medipix3 prototype, a pixel readout chip working in single photon counting mode with improved spectrometric performance. *IEEE Trans Nucl Sci*. (2007) 54:1824–9. doi:10.1109/TNS.2007.906163
- X-Spectrum GmbH. X-ray and electron detectors. Available at: <https://x-spectrum.de/> (Accessed 01-11-2023).

Acknowledgments

Acknowledge support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

25. Poikela T, Plosila J, Westerlund T, Campbell M, De Gaspari M, Llopart X, et al. Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout. *J Instrum* (2014) 9:C05013. doi:10.1088/1748-0221/9/05/C05013
26. Pennicard D, Smoljanin S, Pithan F, Sarajlic M, Rothkirch A, Yu Y, et al. LAMBDA 2M GaAs—a multi-megapixel hard X-ray detector for synchrotrons. *J Instrum* (2018) 13:C01026. doi:10.1088/1748-0221/13/01/C01026
27. Pennicard D, Graafsma H. Simulated performance of high-Z detectors with Medipix3 readout. *J Instrum* (2011) 6:P06007. doi:10.1088/1748-0221/6/06/P06007
28. Veale MC, Angelsen C, Booker P, Coughlan J, French MJ, Hardie A, et al. Cadmium zinc telluride pixel detectors for high-intensity x-ray imaging at free electron lasers. *J Phys D: Appl Phys* (2019) 52:085106. doi:10.1088/1361-6463/aaf556
29. Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of LGAD sensors with a thin entrance window for soft X-ray detection. *J Instrum* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/C11011
30. Sarajlić M, Pennicard D, Smoljanin S, Fritzsche T, Zoschke K, Graafsma H. Progress on TSV technology for Medipix3RX chip. *J Instrum* (2017) 12:C12042. doi:10.1088/1748-0221/12/12/C12042



OPEN ACCESS

EDITED BY

Jiaguo Zhang,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY

Massimo Camarda,
SenSic, Sweden
Xudong Ju,
ShanghaiTech University, China

*CORRESPONDENCE

H. J. Hyun,
✉ hjhyun@postech.ac.kr

†PRESENT ADDRESSES

J. M. Baek,
Department of Physics, University of
Illinois at Chicago, Chicago, IL,
United States
J. Y. Kim,
Capacitor Institute of Technology,
Sungho electronics, Gyeonggi, South
Korea

RECEIVED 27 October 2023

ACCEPTED 08 December 2023

PUBLISHED 04 January 2024

CITATION

Lee SC, Hyun HJ, Baek JM, Hwang SM,
Jang H, Kim JY, Kim SH and Park H (2024),
X-ray beam test of fabricated photo-
diodes for Pohang-Accelerator-
Laboratory X-ray free-electron laser.
Front. Phys. 11:1328639.
doi: 10.3389/fphy.2023.1328639

COPYRIGHT

© 2024 Lee, Hyun, Baek, Hwang, Jang,
Kim, Kim and Park. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

X-ray beam test of fabricated photo-diodes for Pohang-Accelerator-Laboratory X-ray free-electron laser

S. C. Lee¹, H. J. Hyun^{1*}, J. M. Baek^{2†}, S. M. Hwang¹, H. Jang¹,
J. Y. Kim^{1†}, S. H. Kim¹ and H. Park²

¹XFEL Beamline Division, Pohang Accelerator Laboratory, Pohang University of Science and Technology, Pohang, Republic of Korea, ²Department of Physics, Kyungpook National University, Daegu, Republic of Korea

The Pohang-Accelerator-Laboratory X-ray free-electron laser (PAL-XFEL) emits intense, ultra-short X-ray pulses, enabling studies on ultra-small and -fast dynamics in various fields. We designed and fabricated silicon p-intrinsic-n photo-diodes (PDs) to detect strong X-ray pulses. Four types of PDs were fabricated with different metal shapes on the P- and N-sides. The fabrication process was split into two parts, based on the thickness of the Al metal and anti-reflective coating (ARC) layers, to evaluate the fabrication process feasibility and the detection performance. A beam test was performed with X-ray pulses at the soft X-ray beamline of the PAL-XFEL using the proposed PDs and a commercial PD. The test was designed to evaluate the effects of the metal and ARC layers and compare the detection efficiency depending on the X-ray entrance side and signal readout side. The waveforms were saved during the test and integrated to obtain the charges. Thereafter, they were analyzed and normalized, and compared. This paper presents the preparation, setup, and procedure of the beam test, and the test results are described.

KEYWORDS

PIN photo-diode, fabrication, metal thickness, anti-reflective coating, signal readout, X-ray beam test, Pohang-Accelerator-Laboratory X-ray free-electron laser (PAL-XFEL)

1 Introduction

The Pohang-Accelerator-Laboratory X-ray free-electron laser (PAL-XFEL) [1, 2] is a scientific research facility that provides intense, ultra-short, and coherent X-ray pulses, typically generated through self-amplified spontaneous emission (SASE) process [3, 4]. The PAL-XFEL has recently begun offering hard X-ray self-seeded FELs [5] for specific experiments that require a narrow bandwidth and high spectral brightness. The PAL-XFEL comprises a hard X-ray (HX) beamline [6] and a soft X-ray (SX) beamline [7, 8]. The HX and SX beamlines operate with photon energies in the ranges of 2–15 keV and 0.25–1.25 keV, respectively. The unique properties of XFEL beams enable the study of ultra-small and -fast dynamics in various fields through pump-probe experiments [9]. Optical laser pump with wavelengths of 266, 400, or 800 nm and X-ray probe experiments are commonly conducted at the PAL-XFEL.

The XFEL beam unpredictably fluctuates between pulses because of the SASE process; thus, online photon diagnostics is essential. As representative online diagnostics devices, the quadrant beam position monitor (QBPM) and gas monitor detector (GMD) are used at the

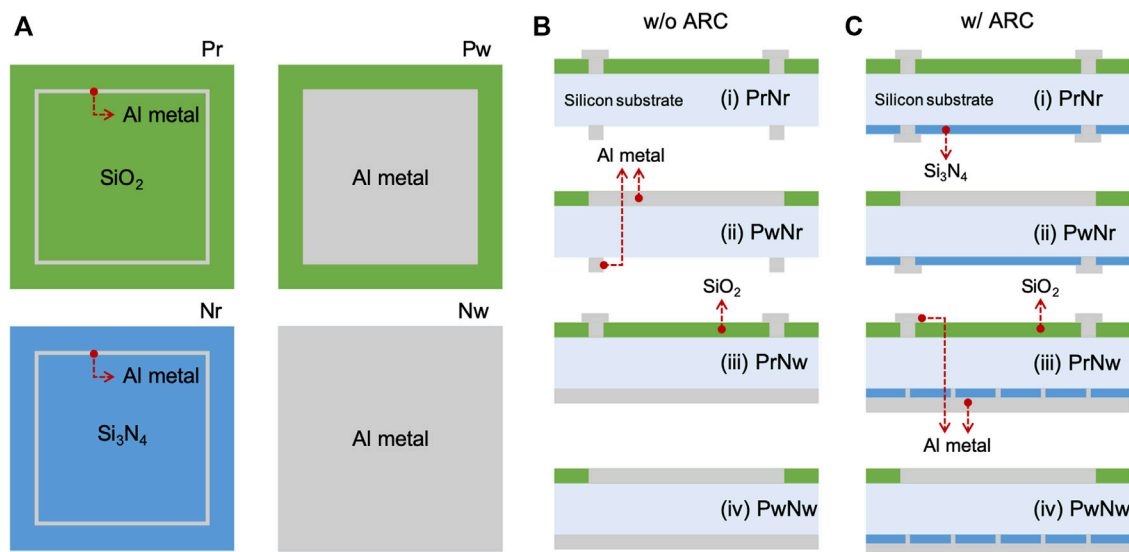


FIGURE 1

(A) Simple schematics of four metal types; P- and N-side rings and whole metals and (B, C) cross-sectional views of their combinations; (i) PrNr, (ii) PwNr, (iii) PrNw, and (iv) PwNw depending on the existence of an ARC layer. In their schematics, sky blue represents the silicon substrate, green represents the passivation layer of SiO_2 , gray represents the Al metal layer, and blue represents the ARC layer of Si_3N_4 . Pr had the SiO_2 layer, and Nr may or may not have the Si_3N_4 layer depending on the ARC process. The implant region of Pw and Nw is covered by the Al metal layer.

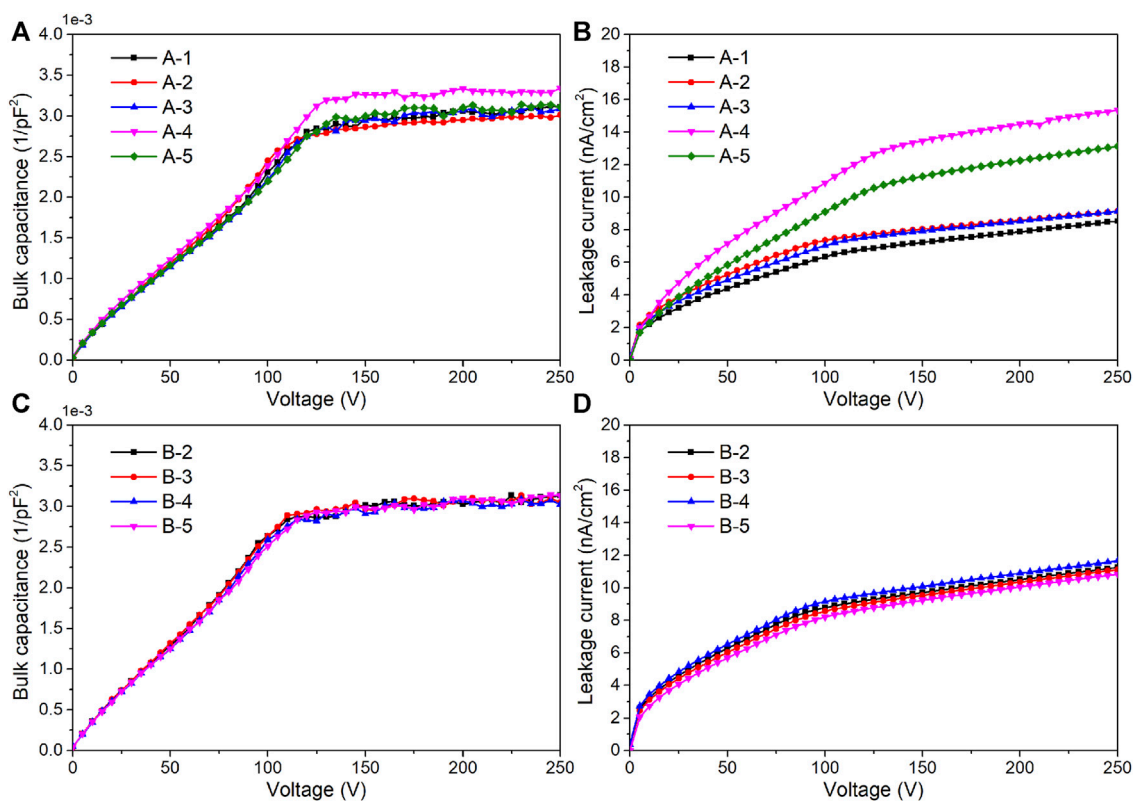


FIGURE 2

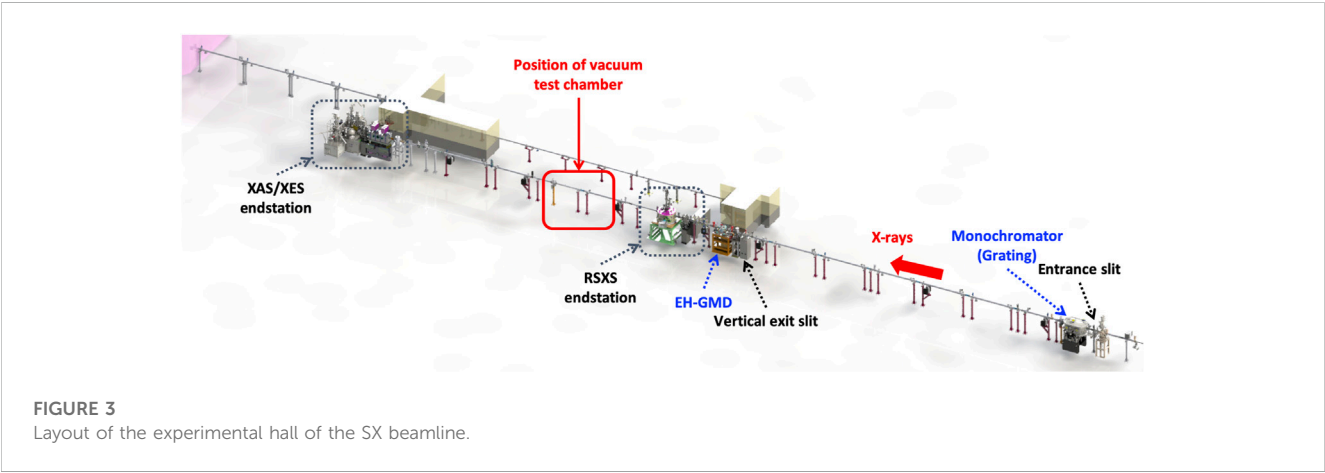
Electrical characteristics of fabricated PDs. (A, C) Capacitances and (B, D) leakage currents of Sets A and B.

TABLE 1 Evaluating the effect of Al and ARC layers on the N-side. Set A comprised PwNr and PwNw structures, with cross-sectional views depicted in [Figure 1B\(ii\)](#), [Figure 1C\(ii\)](#), [Figure 1B\(iv\)](#), and [Figure 1C\(iv\)](#), respectively. The fabrication parameters for the Al metal and ARC layers in Set A, as well as the target and measured values are presented. Set A has light entrance on the N-side and signal readout on the P-side.

PD id	PD type	Cross-section view	TiW metal (nm)	Al metal (nm)	ARC (nm)
A-1	PwNr	Figure 1B(ii)	-	-	-
A-2	PwNw	Figure 1B(iv)	150	100 (92.4 ± 0.6)	-
A-3	PwNw	Figure 1B(iv)	150	200 (189.7 ± 1.3)	-
A-4	PwNr	Figure 1C(ii)	-	-	48.8 ± 2.0 (48.3 ± 0.5)
A-5	PwNw	Figure 1C(iv)	150	400 (373.6 ± 3.0)	48.8 ± 2.0 (48.3 ± 0.5)

TABLE 2 Investigating the effect of light entrance and signal readout sides. Set B consisted of the HPK-PD and four fabricated PrNr type PDs, with cross-sectional views depicted in [Figure 1B\(i\)](#). Definitions of light entrance and signal readout sides of Set B are presented.

PD id	PD type	Cross-section view	Light-in	Signal-out	Conditions of both sides
B-1	HPK-PD	N/A	P-side	P-side	N/A
B-2	PrNr	Figure 1B(i)	P-side	P-side	No ARC layer on N-side and thick passivation layer on P-side
B-3	PrNr	Figure 1B(i)	P-side	N-side	
B-4	PrNr	Figure 1B(i)	N-side	P-side	
B-5	PrNr	Figure 1B(i)	N-side	N-side	



HX and SX beamlines, respectively. The QBPM comprises four photo-diodes (PDs) and a thin film in the beam path [10, 11]. The operation of the GMD is based on the photo-ionization of noble gases [12]. In addition, PDs are used in various ways at both beamlines as destructive or non-destructive methods, including beam alignment, beam size measurement, timing jitter measurement, and scattering signal measurement (a method for normalizing experimental data).

The PDs used at the PAL-XFEL are commercial, and it can be challenging to optimize their specifications for various purposes and obtain them on time owing to external factors, such as international circumstances. Thus, we designed and fabricated PDs for strong X-ray and optical laser detection at the foundries of institutes in South Korea using an N-type silicon substrate with a thickness of 500 μm , diameter of 6 inches, and high resistivity (more than

5 $\text{k}\Omega\text{-cm}$). We designed four types of p-intrinsic-n (PIN) PDs depending on the metal structure on the P-side (junction side, signal readout) and N-side (ohmic side, light entrance window). The metal on each side was designed as a ring or whole shape and labeled as Pr (P-side ring), Pw (P-side whole), Nr (N-side ring), or Nw (N-side whole). Anti-reflective coating (ARC) was applied to the N-side to reduce the reflection of visible light. A thick SiO_2 layer was deposited on the P-side to protect it from environmental degradation. [Figure 1](#) shows the different metal types used on the P- and N-sides. The fabricated PDs have been previously characterized [13] in terms of design, fabrication, electrical characteristics, quantum efficiency, signal-to-noise ratio, and energy resolution. The PDs were evaluated for the effects of metal thickness and detection efficiency using an XFEL beam. This paper presents the electrical characteristics of selected PDs



FIGURE 4

Vacuum test chamber installed downstream of the RSXS endstation with PDs from Sets A and B assembled inside.

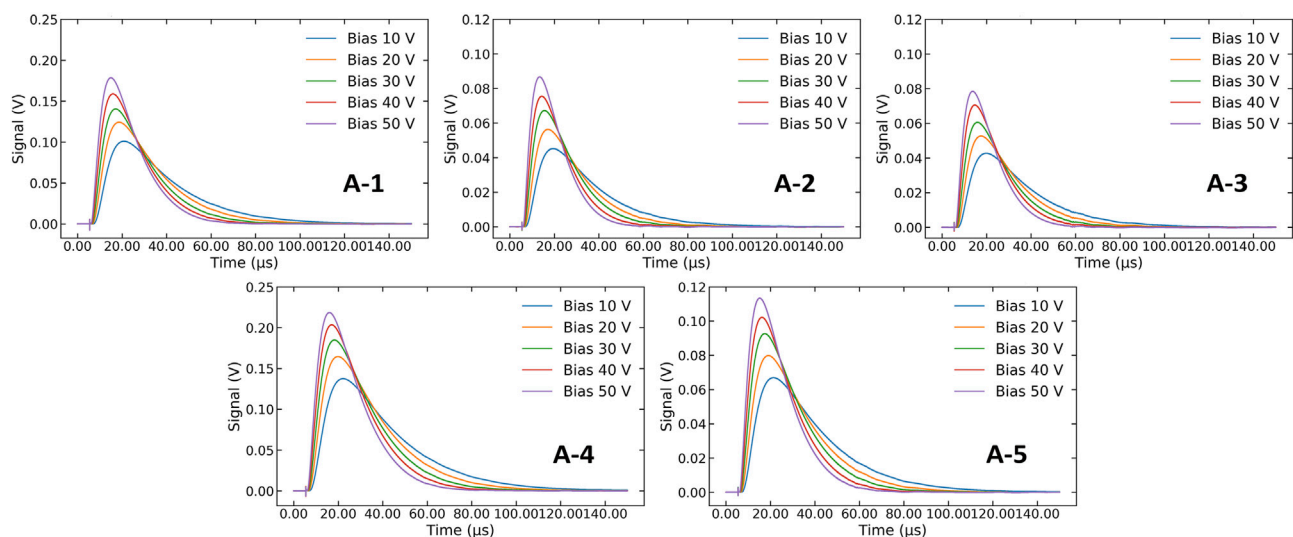


FIGURE 5

Signal waveforms of fabricated PDs in Set A under applied biases from 10 to 50 V for a photon energy of 900 eV.

for an X-ray beam test and describes the test results depending on the metal and readout types of the PDs, including the existence of the ARC layer. The fabricated PDs were compared with a commercial PD (S3590-09, Hamamatsu (HPK-PD) [14]).

2 Beam test with X-ray pulses

2.1 Purpose and preparation

The main purposes of the X-ray beam test with the fabricated PDs were to evaluate the effect of the metal thickness and ARC layer and compare the detection efficiencies of the fabricated PDs with that of the HPK-PD, depending on the light entrance and signal readout sides. Nine fabricated PDs were selected based on their

electrical test results and categorized into Sets A (A-1 to A-5) and B (B-2 to B-5), to address the test objectives. The HPK-PD was designated B-1. The capacitance and leakage current of the PDs with good and similar electrical test results were measured using an LCZ meter (4277A, HP) and a picoammeter (6487, Keithley) (Figure 2). The measured capacitances at the full depletion voltage of 125.6 ± 4.9 V correlated with the calculated bulk capacitance of 20.7 ± 0.9 pF/cm². The leakage currents of the PDs except A-4 and A-5 were below 10 nA/cm² at the full depletion voltage. A printed circuit board (PCB) for detecting high intensity beams in the photoconductive mode was used to mount the PDs. The pads of the PCB were wire-bonded to the electrodes of the PDs for signal readout and biasing.

Set A was used to evaluate the effects of the metal thickness and ARC layer on the N-side. Set A PDs were the Pw type and shared a

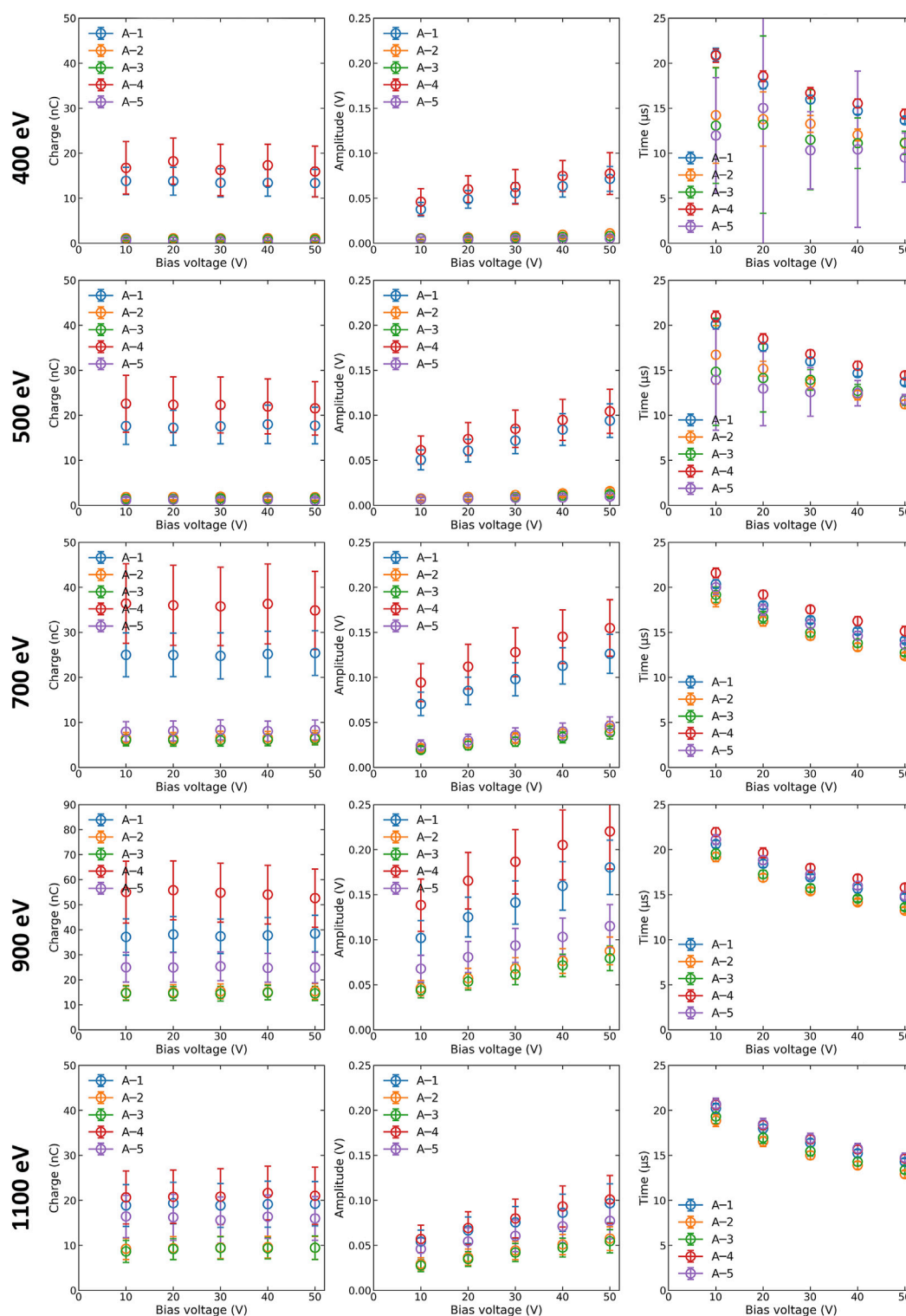


FIGURE 6
Signal charge, amplitude, and time of Set A at photon energies of 400, 500, 700, 900, and 1,100 eV.

common light entrance and signal readout configuration (X-ray entrance to the N-side and signal readout from the P-side). Table 1 lists the fabrication parameters of the metal and ARC layers of Set A. A-1 was of the Nr type, with no metal or ARC layers on its N-side (Figure 1B(ii)). A-2 and A-3 both had 150 nm-thick TiW layers and

Al layers of 100 and 200 nm, respectively, on their N-sides (Figure 1B(iv)). The average thickness of the ARC layer on A-4 (Figure 1C(ii)) and A-5 (Figure 1C(iv)) was 48.3 ± 0.5 nm, as measured at five points on the wafer (top, bottom, left, right, and center). A-5 was of the Nw type with a metal stack of 150 nm-thick TiW and 400 nm-

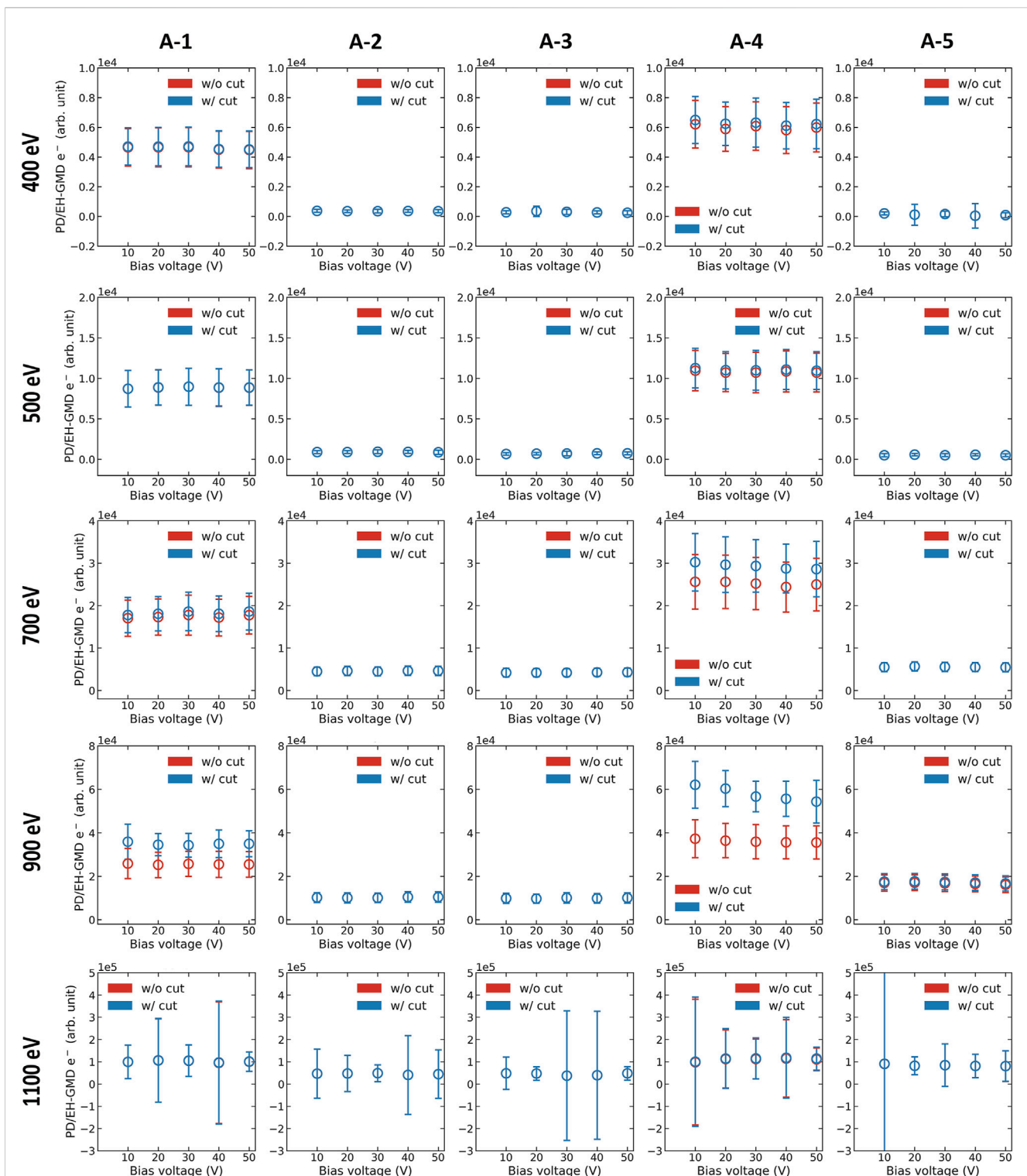


FIGURE 7
Normalized charge distributions as a function of applied bias depending on cuts applied, for A-1–A-5 (left to right) and 400–1,100 eV (top to bottom).

thick Al. The thicknesses of the Al layers were determined by measuring the sheet resistance of the Al at five points on the wafer, using the average sheet resistance and resistivity of Al ($2.735 \times 10^{-8} \Omega \cdot m$). Set B comprised the HPK-PD (B-1) and four fabricated PrNr type PDs without metal or ARC layers (Figure

1B(i)), to examine the effects of the light entrance and signal readout sides on charge collection and rising time. Table 2 defines the light entrance and signal readout sides of each PD. B-1 and B-2 had the light entrance and signal readout on the P-side, whereas B-3 had the light entrance and signal readout on the P- and N-sides, respectively.

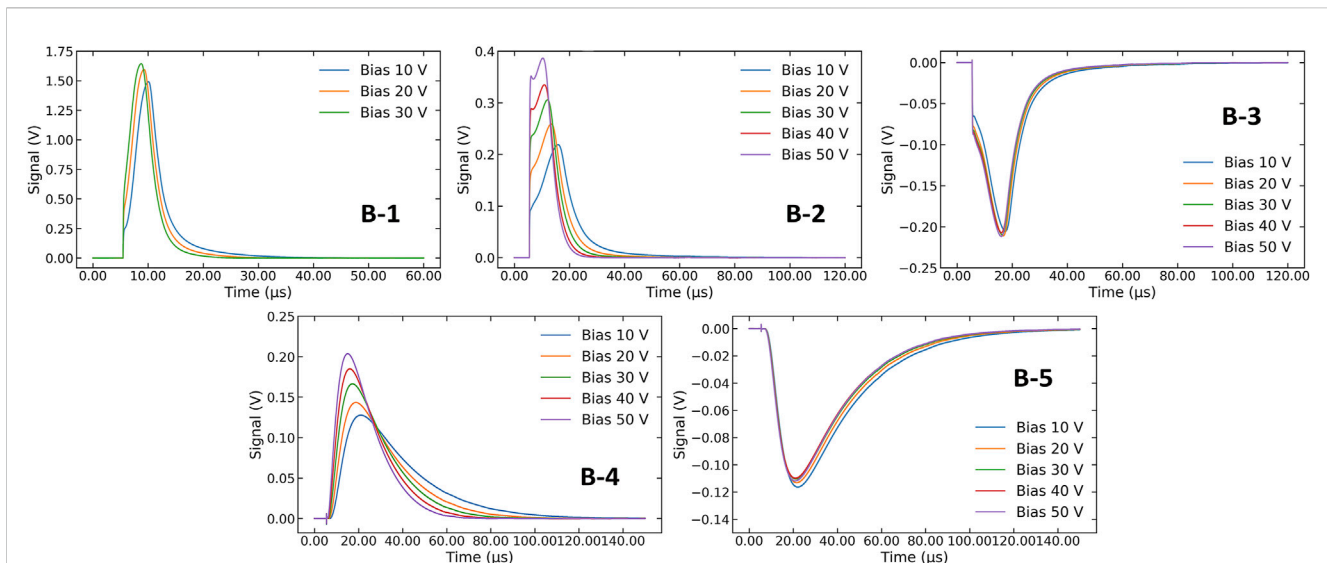


FIGURE 8

Signal waveforms of the HPK-PD (B-1) and fabricated PDs (B-2–B-5) with applied biases from 10 to 30 V or 50 V at 900 eV.

B-4 was the opposite of B-3, and B-5 had the light entrance and signal readout on the N-side.

2.2 Setup and procedure

The X-ray beam test was performed at the SX beamline of the PAL-XFEL. Figure 3 shows the layout of the experimental hall of the SX beamline. A vacuum test chamber was installed between a resonant soft X-ray scattering (RSXS) endstation and an X-ray absorption and emission spectroscopy (XAS/XES) endstation. An X-Y manipulator was assembled on the test chamber, and Sets A and B were assembled in a holder and successively tested. Charge signals from the PDs were processed using a current amplifier (DHPCA-100, FEMTO) and a digitizer (PXIe-5160, National Instruments) synchronized with the PAL-XFEL event timing system. Figure 4 shows the vacuum test chamber installed on the SX beamline, with the PDs. Prior to the data acquisition, the PD holder was aligned, and the center position of each PD was recorded using the X-Y manipulator.

Monochromatic photons with energies of 400, 500, 700, 900, and 1,100 eV were used. The size of the vertical exit slit was 0.1 mm, except for 400 eV photons, which used a larger slit with a size of 0.2 or 0.25 mm. As a result of previous X-ray beam test, we decided to apply the under-bias voltage which showed similar signal charge compared to the full depletion or over-bias voltages [15]. Therefore, at each energy, the reverse bias voltages of the fabricated PDs (HPK-PD) were varied from 10 to 50 V (30 V), with steps of 10 V, applied to the opposite side of the signal readout for Sets A and B. The waveforms of all the PDs, as well as the waveforms of photo-ions and photo-electrons from the GMD at the experimental hall (EH-GMD), were saved for each setting. A total of 1,200 pulses were collected for each setting at a repetition rate of 60 Hz. No attenuators were used, and the X-ray photon flux was adjusted by changing the sizes of the entrance and vertical exit slits for Set B at 900 eV.

The charges of the PDs and EH-GMD were obtained by integrating the saved waveforms. The PD charges correlated better with the EH-GMD photo-electron charges than the photo-ion charges. Thus, the PD charges were normalized using the EH-GMD photo-electron charges according to the following equation:

$$Q_{norm} = \frac{1}{n} \sum_{i=1}^n \frac{q_i(PD)}{q_i(EH-GMD)}, \quad (1)$$

where $q_i(PD)$ is the i th charge of the PD, $q_i(EH-GMD)$ is the i th charge of the EH-GMD photo-electrons, and n is the total number of X-ray pulses.

2.3 Test results

2.3.1 Effects of metal and ARC layers

Figure 5 shows the averaged signal waveforms of the PDs in Set A depending on applied bias voltages for 900 eV monochromatic X-rays. The PDs in Set A exhibited similar waveform shapes. With an increase in the bias voltages, the waveforms exhibited increasing heights and rapidly rising times. However, the peak heights were dependent on the thickness of the metal layer and the presence of an ARC layer. Figure 6 shows the averaged charges, amplitudes, and times as a function of bias voltages for Set A at each photon energy. The charges were calculated by integrating the waveforms. The amplitudes were the maximum heights of the waveforms, and the times were those at which the waveforms reached their maximum heights. The data from 1,200 pulses were averaged for each setting, and the averages and standard deviations were plotted. The facts that the charges were not affected by the bias voltage, the amplitudes increased, and the times decreased as the bias voltage increased were clearly shown over all photon energies. However, the charges varied from A-1 to A-5. Contrary to expectations that the ARC layer of Si_3N_4 would function as an attenuator, A-4 and A-5 with ARC layers

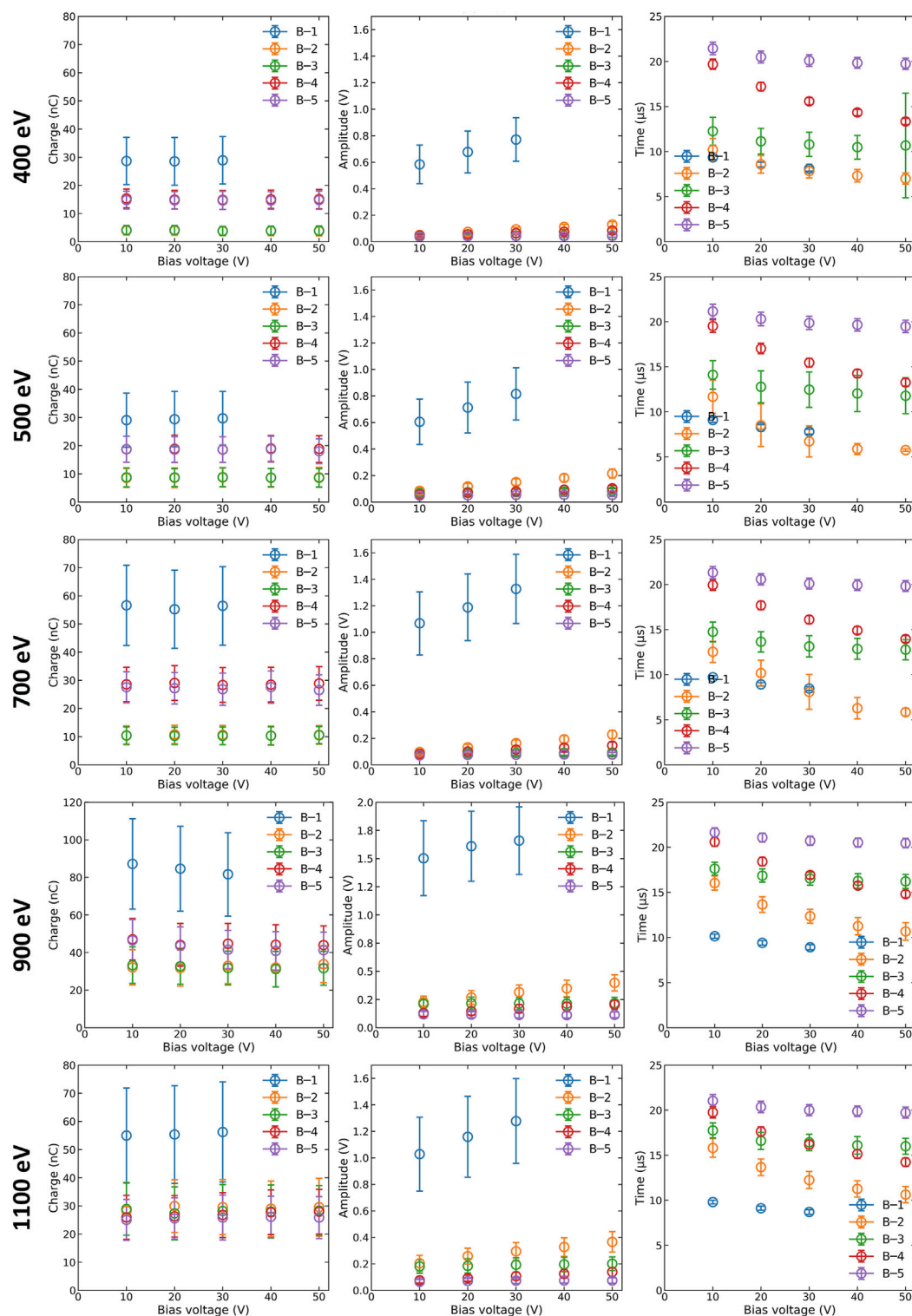


FIGURE 9

Signal charge, amplitude, and time of Set B at photon energies of 400, 500, 700, 900, and 1,100 eV.

offered better responses than A-1 and A-2-3, respectively. This suggested that the ARC layer, which had a medium refractive index, played an anti-reflection role on the silicon and metal surfaces even for X-rays. Thus, A-4 exhibited the largest charges and amplitudes at all photon energies because of the absence of the metal layer and the

effect of the ARC layer. A-5, with the thickest metal layer, exhibited slightly larger charges than A-2 and A-3, particularly at higher photon energies, probably because metal attenuation decreased at higher energies. The signal charges of A-1 and A-4 were sufficiently large such that their rising times were unaffected by photon energy.

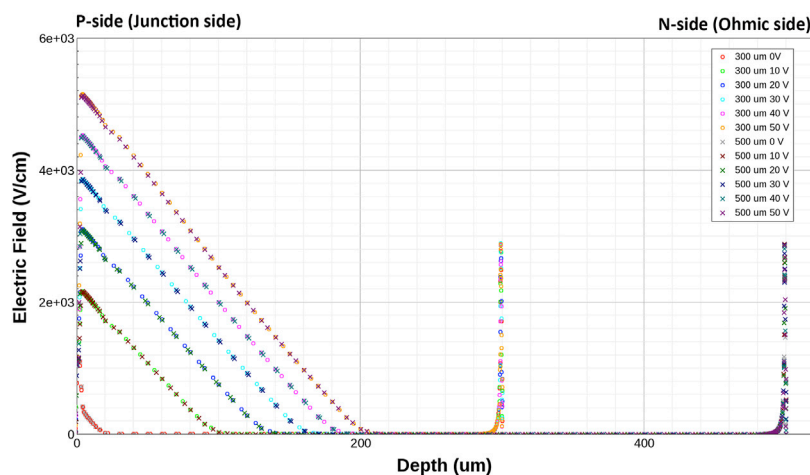


FIGURE 10

TCAD simulation results of the electric field distribution of PDs with thicknesses of 300 and 500 μm .

Contrarily, the signal charges of A-2, A-3, and A-5 were small at 400 and 500 eV because of their metal layers, resulting in short rising times with large errors (Figure 6).

The PD charges were normalized to the charges of the EH-GMD photo-electrons using Eq. (1) for each experimental parameter. Considering that the PDs were always saturated at high incident flux in the correlation plot between the PDs and EH-GMD photo-electron charges, the normalized charges of the PDs were obtained after the cuts were applied. The cuts were determined to be 2.0×10^{-8} C and 3.0×10^{-8} C for 400 eV and the other energies, respectively. Figure 7 shows the normalized charge distributions as a function of bias voltages for each PD and photon energy, with and without cuts applied. For A-1 and A-4, particularly at a photon energy of 900 eV, which had the highest intensity, the signal charges of the PDs were sufficiently large to be saturated. Thus, the normalized charges significantly increased after the cuts were applied. The detection efficiency of A-4 improved by 55.4% (36.6%) over that of A-1 at 900 eV (on average for all the photon energies). The large errors of 1,100 eV shown in Figure 7 were due to the poor signal-to-noise ratio of the EH-GMD signal. To address this challenge, the gain of the EH-GMD was increased at 1,100 eV for Set B.

To examine the thickness of the Al metal layer, the ratios of the normalized charges, $\bar{Q}_{\text{norm_A-2}}/\bar{Q}_{\text{norm_A-1}}$, $\bar{Q}_{\text{norm_A-3}}/\bar{Q}_{\text{norm_A-1}}$, and $\bar{Q}_{\text{norm_A-5}}/\bar{Q}_{\text{norm_A-4}}$, were calculated after applying cuts and compared with theoretical values considering X-ray attenuation in matter [16]. The obtained ratios were smaller than the theoretical values at all energies. The ratios of the theoretical values to the measured values ranged from 1.004 to 3.895 and increased with a decrease in the photon energy. The observed differences could have been due to a thicker TiW layer, unexpected layers stacked during other fabrication processes, such as SiO₂ or reflections at the metal/silicon interfaces. The thicknesses of TiW and SiO₂ were calculated to be about 300 and 500 nm, respectively, considering the effect of single layer only. X-ray reflection at the interface was observed for A-4.

2.3.2 Detection effects according to light incident and signal readout surfaces

Figure 8 shows the averaged signal waveforms of Set B. B-1, B-2 and B-3, which used their P-side for light entrance, had similar sharp rising shape of waveforms. Because B-3 used its N-side for reading the signals, the waveforms were inverted. For the same reason, B-4 and B-5, which used their N-side for light entrance, showed the similar gradual rising shape of waveforms with an opposite sign. B-1, B-2, and B-4 read signals from the P-side, similar to Set A, and B-4 had the same light incident direction as Set A. Thus, their behavior as a function of bias voltages was similar to Set A; increased heights and shorter rising times were observed with an increase in the bias voltage (Figure 9). B-3 and B-5 used their N-side to read the signals; therefore, their waveforms were inverted. Additionally, the waveforms of B-3 and B-5 appeared unaffected by changes in the bias voltage, resulting in minimal change in amplitude and rising time across all the photon energies (Figure 9). Technology CAD (TCAD) simulation [17] was performed to understand the difference in the signal readouts of the P- and N-sides. The simulation revealed that the electric field on the P-side increased and the depletion depth expanded as the bias voltage increased while the electric field on the N-side remained constant (Figure 10). This explains why the N-side signal readout exhibited different waveform characteristics, in terms of the amplitudes, and rising time, compared with the P-side signal readout.

Among Set B PDs, B-1 had the highest charges and amplitudes at all photon energies. The rising times of B-2 were close to or shorter than those of B-1 at photon energies below 700 eV, which were results of the same cause for the shorter rising times of A-2, A-3, and A-5 within the same energy range. A thick SiO₂ passivation layer on the P-side significantly attenuated the incident X-ray flux at a lower energy. The charges of B-2 and B-3, which used the P-side for X-ray entrance, were slightly lower than those of B-4 and B-5, which used the N-side for light entrance, except at 1,100 eV. This was consistent with the normalized charge distributions as a function of bias voltage shown in Figure 11. The effect of the SiO₂ was investigated by

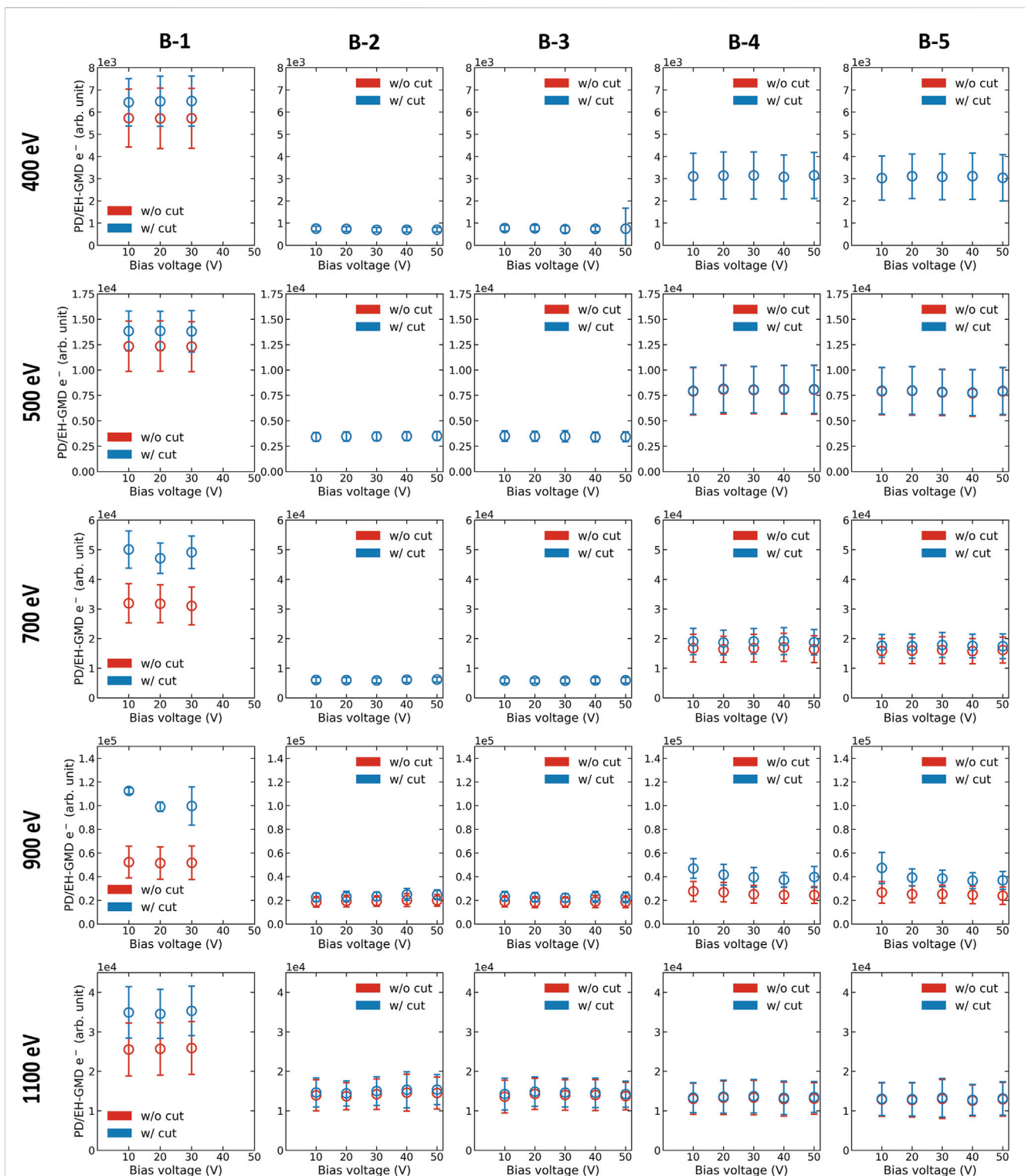


FIGURE 11

Normalized charge distributions as a function of applied bias depending on whether cuts were applied or not, for B-1–B-5 (left to right) and 400–1,100 eV (top to bottom).

comparing the ratios of normalized charges $\bar{Q}_{norm_B-2(3)}/\bar{Q}_{norm_B-4(5)}$ to the theoretical values of 1,350 nm-thick SiO_2 [16] after the cuts applied. The cut value for Set B was determined to be 3.0×10^{-8} C for photon energies below 1,100 eV and 4.0×10^{-8} C for 1,100 eV.

Figure 12A shows the theoretical values and measurement results ($\bar{Q}_{norm_B-2(3)}/\bar{Q}_{norm_B-4(5)}$), as a function of photon energy, with cuts. Figure 12B shows the ratios of the measured values to the theoretical values. Errors were calculated by error propagation using

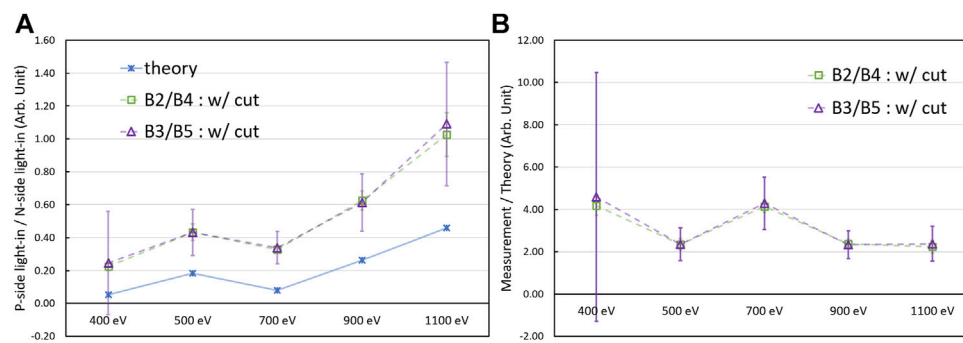


FIGURE 12 (A) Ratios of normalized charge of B-2 (B-3) to that of B-4 (B-5) after cuts were applied and the theoretical values as a function of photon energy. (B) Ratios of the measured to the theoretical values as a function of photon energy with cut.

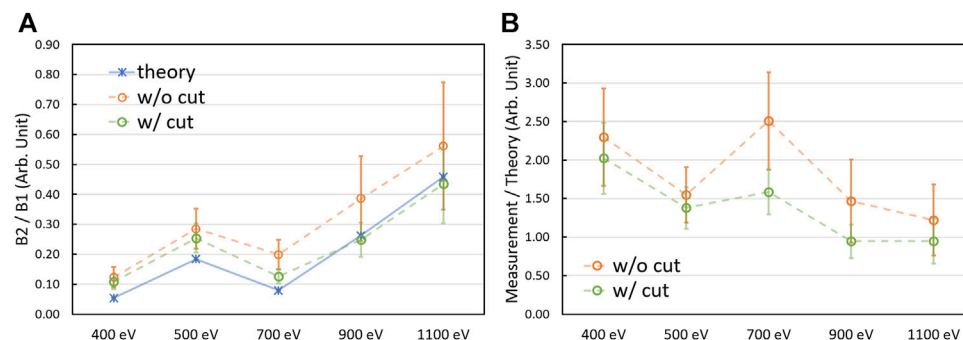


FIGURE 13 (A) Ratios of normalized charge of B-2 to that of B-1 and the theoretical values as a function of photon energy. (B) Ratios of the measured to the theoretical values as a function of photon energy, depending on whether cuts were applied or not.

the standard deviations of each normalized charge. The measured values exceeded the theoretical values by a factor of more than 2, indicating the anti-reflection effect of the SiO₂ layer.

To compare the fabricated PD with commercial PD, we calculated the ratios of normalized charges $\bar{Q}_{norm_B-2}/\bar{Q}_{norm_B-1}$ and compared with theoretical values of a 1,350 nm-thick SiO₂ depending on the cuts applied. Figure 13A shows the theoretical values and measurement results ($\bar{Q}_{norm_B-2}/\bar{Q}_{norm_B-1}$) as a function of photon energy, with and without applied cuts. Figure 13B shows the ratios of the measured values to the theoretical values. The $\bar{Q}_{norm_B-2}/\bar{Q}_{norm_B-1}$ ratios at 900 and 1,100 eV correlated with the theoretical values obtained after cuts were applied. Thus, the ratio of the measured to theoretical values were close to one. A difference of 2.2σ in the maximum value was observed at 700 eV and below, suggesting a potential improvement in anti-reflection effect compared to the commercial PD.

B-2 and B-4 had a higher amplitude and shorter rising time than B-3 and B-5, respectively. This implied that reading signals from the P-side was appropriate when fast reading was required in case of partial depletion. At 900 eV, the rising times of B-1 to B-5 at maximum bias voltage were 8.9 ± 0.3 , 10.7 ± 1.0 , 16.2 ± 0.8 , 14.8 ± 0.4 , and 20.5 ± 0.5 μ s, respectively. B-1 and B-2, which used the P-side for light-in and signal-out, had the best rising times, and B-2 was comparable to B-1.

The incident flux was adjusted to three by changing the sizes of the entrance slit and vertical exit slit at a photon energy of 900 eV. At the standard slit setting, all the Set B PDs exhibited different normalized charge distributions depending on whether or not cuts were applied (Figure 11), although there were differences between the PDs. The difference between the normalized charge distributions depending on the existence of cuts for each PD decreased as the incident flux decreased with smaller slit sizes. It was within 1σ for B-1 and disappeared for B-2 to B-5 at the smallest incident flux.

3 Conclusion

We designed and fabricated PDs for use as diagnostic devices at the PAL-XFEL for detecting X-rays or optical lasers. The PDs were manufactured with different Al metal thicknesses and ARC layers to evaluate the feasibility of fabricating thin metal PDs and assess their functionalities. The electrical characteristics, leakage current and capacitance, of the fabricated PDs were measured, and nine PDs with good electrical characteristics were selected for X-ray beam testing. The X-ray beam test was performed at the soft X-ray beamline of the PAL-XFEL with photon energies ranging from

400 to 1,100 eV. The under-bias voltages were applied to the PDs, and the PD charges obtained integrating waveforms were normalized using EH-GMD charges. Due to saturation observed in the PDs when exposed to unattenuated strong X-ray pulses, the normalized charges were compared after applying cuts to the data. The five PDs with different Al metal thickness and the presence or absence of ARC layers on their N-side were initially tested to evaluate the effect of the metal and ARC layers. The temporal and amplitude characteristics of signal waveforms were modulated by the bias voltage while maintaining a constant normalized charge. The PDs with ARC layers exhibited the better detection efficiency, and particularly the PD with only ARC layer, indicating that the ARC reduced reflection at the silicon surface, even for X-rays. When the ARC and metal layers were present, the anti-reflection effect was evident at photon energies above 700 eV and minimal at the low energy of 500 eV. The estimated transmissions by the metal layers were 1–4 times lower than the theoretical predictions, and this discrepancy widened as the photon energy decreased. These could have been due to the thickness of the TiW layer exceeding the target thickness, unexpected additional layers stacked during other fabrication processes, or reflections at the interfaces of the metal and silicon layers. To investigate the effect of light entrance and signal readout side configurations and compare with a commercial PD, we assembled PDs onto the PCB in four distinct arrangements. The signal waveforms exhibited varying behavior depending on the signal readout side. For the P-side signal-out configurations, the signal waveforms were affected by the bias voltage, as the increasing bias voltage resulted in an enhanced electric field and expanded depletion depth from the P-side. The charge discrepancies between the N-side light-in and the P-side light-in configurations were examined, taking into account the 1,350 nm-thick SiO₂ passivation layer on the P-side. Interestingly, the SiO₂ layer appeared to serve not only as an attenuating layer but also as an anti-reflection layer, similar to the Si₃N₄ ARC layer. The P-side light-in and signal-out PD exhibited rapid signal rise times, comparable to the commercial PD which was achieved the highest charge and amplitude values across all photon energies. Accounting for the thick passivation layer of the PD, the normalized charge ratios of the commercial PD and the PD were consistent with the theoretical values at 900 and 1,100 eV. Moreover, the use of an optimized SiO₂ layer could potentially enhance the performance of the PD below 700 eV, surpassing that of the commercial PD. Based on the beam test results, we will optimize PDs with different metal and ARC layer parameters, and readout types for various detection purposes.

References

1. Ko IS, Kang HS, Heo H, Kim C, Kim G, Min CK, et al. Construction and commissioning of PAL-XFEL facility. *Appl Sci* (2017) 7(5):479. doi:10.3390/app7050479
2. Kang H, Min CK, Heo H, Kim C, Yang H, Kim G, et al. Hard X-ray free-electron laser with femtosecond-scale timing jitter. *Nat Photon* (2017) 11:708–13. doi:10.1038/s41566-017-0029-8
3. Kondratenko AM, Saldin EL. Generating of coherent radiation by A relativistic electron beam in an undulator. *Part Accel* (1980) 10:207.
4. Bonifacio R, Pellegrini C, Narducci LM. Collective instabilities and high-gain regime in a free electron laser. *Opt Commun* (1984) 50:373–8. doi:10.1016/0030-4018(84)90105-6
5. Nam I, Min CK, Oh B, Kim G, Na D, Suh YJ, et al. High-brightness self-seeded X-ray free-electron laser covering the 3.5 keV to 14.6 keV range. *Nat Photon* (2021) 15:435–41. doi:10.1038/s41566-021-00777-z
6. Park J, Eom I, Kang TH, Rah S, Nam KH, Park J, et al. Design of a hard X-ray beamline and end-station for pump and probe experiments at Pohang Accelerator

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

SL: Formal Analysis, Methodology, Writing—original draft, Writing—review and editing. HH: Formal Analysis, Methodology, Project administration, Writing—original draft, Writing—review and editing. JB: Methodology, Writing—review and editing. SH: Methodology, Software, Writing—review and editing. HJ: Methodology, Resources, Writing—review and editing. JK: Methodology, Writing—review and editing. SK: Methodology, Resources, Writing—review and editing. HP: Supervision, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the Korean Government MSIT (Grant No. RS-2022-00164805, Accelerator Application Support Project), the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant No. 2022R1I1A1A01068225), and the Ministry of Science and ICT (Grant No. 2018R1A6B4023605, RS-2023-00237194). The EDA tool was supported by the IC Design Education Center (IDEC), Korea.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Laboratory X-ray Free Electron Laser facility. *Nucl Instrum Methods Phys Res A* (2016) 810:74–9. doi:10.1016/j.nima.2015.11.147

7. Park SH, Kim M, Min CK, Eom I, Nam I, Lee HS, et al. PAL-XFEL soft X-ray scientific instruments and X-ray optics: first commissioning results. *Rev Sci Instrum* (2018) 89(5):055105. doi:10.1063/1.5023557

8. Jang H, Kim HD, Kim M, Park SH, Kwon S, Lee JY, et al. Time-resolved resonant elastic soft x-ray scattering at Pohang Accelerator Laboratory x-ray free electron laser. *Rev Sci Instrum* (2020) 91(8):083904. doi:10.1063/5.0016414

9. Dong PT, Cheng JX. Pump-probe microscopy: theory, instrumentation, and applications. *Spectroscopy* (2017) 32:24.

10. FMB OXFORD. *Quadrant beam position monitor* (2023). Available from: <https://fmb-oxford.com/products/detectors-diagnostics/beam-position-monitors/qbpms/> (Accessed October, 2023).

11. Tono K, Kudo T, Yabashi M, Tachibana T, Feng Y, Fritz D, et al. Single-shot beam-position monitor for x-ray free electron laser. *Rev Sci Instrum* (2011) 82:023108. doi:10.1063/1.3549133

12. Tiedtke K, Sorokin AA, Jastrow U, Juranić P, Kreis S, Gerken N, et al. Absolute pulse energy measurements of soft x-rays at the Linac Coherent Light Source. *Opt Express* (2014) 22(18):21214. doi:10.1364/oe.22.021214

13. Lee SC, Baek JM, Hwang SM, Hyun HJ, Kim JY, Kim SH, et al. Development of photo-diodes for Pohang-Accelerator-Laboratory X-ray free-electron laser. *Nucl Instrum Methods Phys Res A* (2023) 1045:167598. doi:10.1016/j.nima.2022.167598

14. Hamamatsu. *Hamamatsu Si PIN photo-diode* (2023). Available from: <https://www.hamamatsu.com/jp/en/product/optical-sensors/photodiodes/si-photodiodes/S3590-09.html> (Accessed October, 2023).

15. Baek JM, Hwang SM, Hyun HJ, Jang H, Kim JY, Kim SH, et al. Performance measurements of photodiodes for X-ray detection. *IEEE Trans Nucl Sci* (2022) 69:1953–9. doi:10.1109/tns.2022.3186481

16. Center for X-Ray Optics. *X-Ray interaction with matter of the center for X-ray optics* (2023). Available from: https://henke.lbl.gov/optical_constants/ (Accessed October, 2023).

17. Silvaco. *TCAD simulation software* (2023). Available from: <https://silvaco.com/ko/tcad/> (Accessed October, 2023).



OPEN ACCESS

EDITED BY

Iain Sedgwick,
Rutherford Appleton Laboratory,
United Kingdom

REVIEWED BY

Naoto Yagi,
Japan Synchrotron Radiation Research
Institute, Japan
Suren Chilingaryan,
Karlsruhe Institute of Technology (KIT),
Germany

*CORRESPONDENCE

Philipp Schmidt,
✉ philipp.schmidt@xfel.eu

RECEIVED 14 October 2023

ACCEPTED 14 December 2023

PUBLISHED 08 January 2024

CITATION

Schmidt P, Ahmed K, Danilevski C,
Hammer D, Rosca R, Kluyver T,
Michelat T, Sobolev E, Gelisio L, Maia L,
Manetti M, Malka J, Wrona K,
Sztuk-Dambietz J, Rovensky V, Ramilli M,
Duarte N, Lomidze D, Dourki I, Yousef H,
Senfftleben B, Meyer O, Turcato M, Hauf S
and Aplin S (2024), Turning European
XFEL raw data into user data.
Front. Phys. 11:1321524.
doi: 10.3389/fphy.2023.1321524

COPYRIGHT

© 2024 Schmidt, Ahmed, Danilevski,
Hammer, Rosca, Kluyver, Michelat,
Sobolev, Gelisio, Maia, Manetti, Malka,
Wrona, Sztuk-Dambietz, Rovensky,
Ramilli, Duarte, Lomidze, Dourki, Yousef,
Senfftleben, Meyer, Turcato, Hauf and
Aplin. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Turning European XFEL raw data into user data

Philipp Schmidt*, Karim Ahmed, Cyril Danilevski, David Hammer, Robert Rosca, Thomas Kluyver, Thomas Michelat, Egor Sobolev, Luca Gelisio, Luis Maia, Maurizio Manetti, Janusz Malka, Krzysztof Wrona, Jolanta Sztuk-Dambietz, Vratko Rovensky, Marco Ramilli, Nuno Duarte, David Lomidze, Ibrahim Dourki, Hazem Yousef, Björn Senfftleben, Olivier Meyer, Monica Turcato, Steffen Hauf and Steve Aplin

European XFEL, Schenefeld, Germany

The European X-ray Free Electron Laser is a research facility located close to Hamburg, offering X-ray pulses with ultra-high brilliance and femtosecond duration at megahertz repetition rates. The detection systems necessary to unlock the full scientific potential made possible by this machine poses considerable challenges both in terms of data volume and rate, as well as the interpretation of their recorded signal. To provide optimal data quality, expert and detector-specific knowledge not easily accessible to external facility users is essential, and its implementation must cope with the generated volumes. We therefore aim to perform these preparatory processing steps and offer users a dataset suitable for further analysis as the primary data product. This work describes the machinery and workflows providing this data to users in an automatic, configurable and reproducible manner, both online during the experiment, and offline for scientific analysis afterward on the way to publication.

KEYWORDS

photon science, free electron laser, big data, data analysis, data quality, detector correction, HPC

1 Introduction

The advent of X-ray free electron laser sources and in particular their recent advance into data rates in the kHz regime continues to push the boundaries of data analysis techniques. The *European X-ray Free Electron Laser (European XFEL)* [1, 2], in operation since 2017, is such a facility located in the area of Hamburg, Germany. Its superconducting linear accelerator produces electron bunches with an energy of up to 17.5 GeV in a unique burst mode time structure as shown in [Figure 1](#). The resulting X-ray pulses are arranged in trains of up to 2,700 pulses, with trains arriving at a rate of 10 Hz. Within each train, the pulses are separated by as little as 222 ns, which is equivalent to an intra-train repetition rate of up to 4.5 MHz. They are currently delivered to three beamlines in parallel covering the soft X-ray to hard X-ray photon energy regime. At each beamline, up to three instrument endstations are installed spanning a large range of different experiment techniques.

This unique train-pulse time structure offers the benefit of high pulse energies and small wavelengths at comparably high repetition rates, but incurs additional challenges in terms of detector technologies able to keep up with this intra-pulse distance and duty cycle. These challenges led to the development of multiple custom X-ray 2D imaging cameras—*AGIPD*

[3], *LPD* [4], and *DSSC* [5]—capable of capturing up to 8,000 frames per second at the pulse repetition rate of 4.5 MHz, by using large memory cell arrays, while also being able to cover large dynamic ranges of photon intensities. Achieving optimal data quality with these detectors requires intimate technical knowledge and the sheer data volume they produce means that processing must be highly scalable. For users, this complexity can impose a high barrier of entry to make use of their data and to achieve scientific results for their proposals.

We are aware of the impact of this complexity on data analysis for users, and aim to offer the data taken during a proposal in a form useful in the scientific context of its experiment. This form of data we call *user data*, and it is provided in the same data format alongside the original *raw data* as it was acquired by detectors. What constitutes user data can be highly variable from experiment to experiment and depends on the technique, experimental conditions, and of course detectors used. It may range from image corrections per pixel, over clustering or integration of neighbouring intensities, to event reconstruction across correlated signal sources. To this end, established and essential data preparation steps are offered as a service running on the facility infrastructure, where they can be efficiently and reproducibly applied at scale. Data processing is provided both for real time applications during an experiment—delivering data streams at latencies of a few seconds or less—as well as for exhaustive processing of data recorded to disk with a focus on completeness, precision, and reproducibility, scaling to up the petabyte regime for single experiments. Rather than replacing the raw data product, however, these systems are designed to maintain configurability and integrate into user workflows with custom adaptations for each scientific application. This article reports on the general infrastructure and systems developed for this purpose as well as the specific detectors and methods it was applied to over the past 6 years of facility operation. Their impact on user experiments and the facility is discussed, leading up to a comparison with the originally envisioned concept and an outlook into upcoming developments.

2 Methods

Data processing at European XFEL is generally separated into the two paradigms of *online* and *offline*. This separation is reflected in the facility-side machinery and tools that provide user data.

Online processing happens during the experiment on the direct data streams from the detectors and other acquisition devices, it is near real-time and provides immediate feedback and monitoring to steer the experiment. Given these requirements, and the key role online analysis plays in the success of an experiment, its primary focus is low latency to provide analysis results within a few seconds or less with high reliability for the operator. To this end, it may only operate on a relevant subset of data and employ less sophisticated algorithms to guarantee a result at the highest possible throughput, potentially at the cost of accuracy.

Offline processing on the other hand operates on data stored in persistent files for deeper data exploration and analysis. This may take place minutes to hours after acquisition to guide experimental decisions and extend for months after the experiment is concluded until a clear scientific picture emerges. Such analyses aim at accuracy, completeness, and reproducibility and, as such, they are generally more efficient at scales which enable the use of computationally expensive methods.

Distinct solutions have been developed to optimally serve both of these requirements, built around streams and files respectively, with a common ecosystem for tracking metadata. Both operate on the same input of raw data, but are strictly split on the volatility of their results. Those obtained from online processing are generally not stored to disk to prevent any compromise in data quality or reproducibility, which may result from their *a priori* configuration or performance requirements. Instead, any permanent results are produced through the offline processing system to allow for continuous tuning of parameters and behaviour for optimal and traceable results.

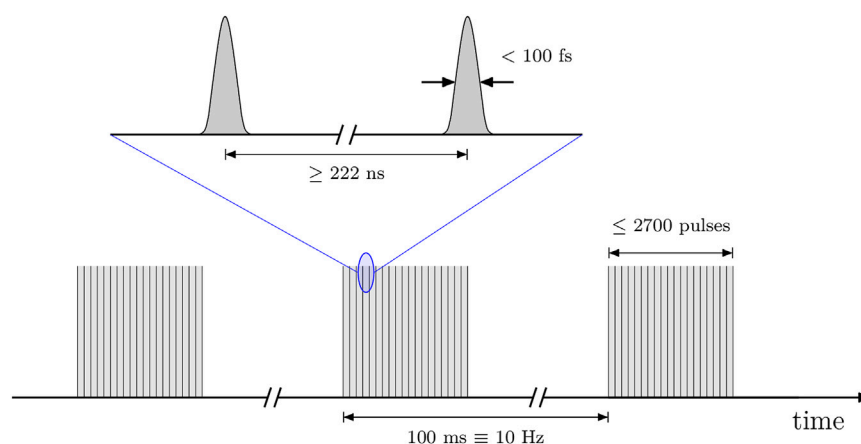


FIGURE 1

The time structure of European XFEL consisting of pulse trains with up to 2,700 individual pulses at a train repetition rate of 10 Hz. Within one train, the spacing between bunches is in the order of several hundred nanoseconds, while each bunch by itself has a length of typically less than 100 fs [1].

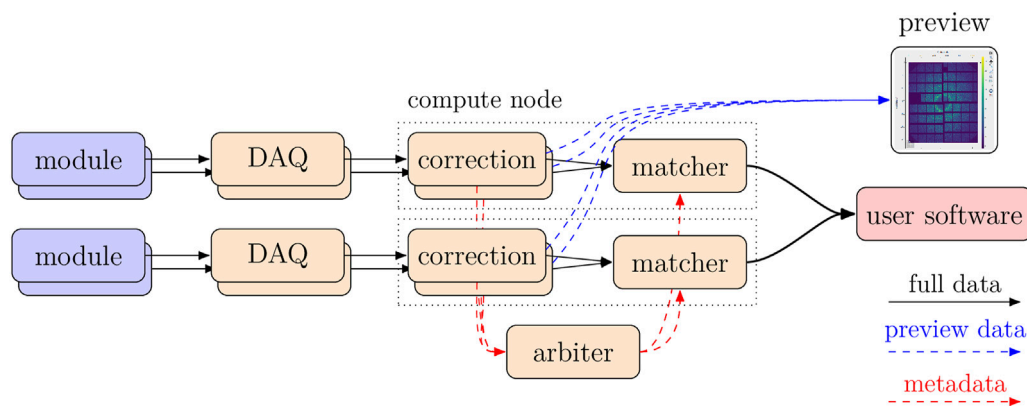


FIGURE 2

Data flows in the online processing pipeline for multi-module 2D pixel detectors with up to 16 modules. Detector hardware sends raw data through the data acquisition system (DAQ) to correction software devices. Each compute node typically hosts up to four correction devices called a group, each processing the data of a single module. From the correction devices, the low-latency preview stream is limited to a single frame and provided for immediate operator display. The full data stream for further downstream analysis can include additional metadata introduced by calibration add-ons running in the correction device. This metadata across all modules may be used for data reduction decisions in an arbiter device. In each group, a matcher aggregates the individual data streams taking this data reduction feedback into account.

2.1 Online data processing

Providing an interpretable result in near real-time during an experiment is often essential to performing successful user beamtimes in photon science, in particular, to make the most efficient use of the short time allocated to each experiment. In addition, it serves a monitoring role for the experimental hardware and environment, helping to ensure safe and effective operation.

The facility-provided online processing system is integrated into *Karabo* [6], a distributed control system developed at European XFEL to address three main challenges: i) acquisition and processing of the large data volumes generated at the facility, ii) provision of global time synchronisation across most control variables, and iii) the flexibility required to efficiently control both static and highly dynamic setups in prototype and facility scale. Physical hardware is represented by corresponding *Karabo* devices written in C++ or Python, which may communicate with each other via a central message broker and direct point-to-point connections. Additional functionality can be provided in the form of pure software devices, as is the case with the online data processing system. The processing is performed in the online computing cluster (ONC), which consists of dedicated nodes for each of the three beamlines operating in parallel and located physically close to their endstations. This is equipped with datacenter-grade graphics processing units (GPUs), which can accelerate the most demanding processing steps.

The topology for online image correction for the large-area multi-module 2D imaging detectors is illustrated in Figure 2, as this presents the most demanding application due to the high data rates involved. A single *Karabo* endpoint device per physical detector module feeds data to a single correction device which performs processing per module. Multiple correction devices are grouped together on a single computing node to balance bandwidth and computing capabilities. From the correction devices, two separate types of output streams are provided: preview output and full data output.

The preview path is limited to at most a single frame per train via configurable reduction methods and provides quick feedback with minimal latency. It is geared for direct monitoring on screen and includes assembly steps of the individual detector modules into their physical geometry. The latency of the assembled preview—including corrections as well as assembly—is typically a few hundred milliseconds.

The full data path, on the other hand, delivers the complete data stream and can be tuned to best match the experimental analysis requirements to the available network performance. It may carry only a few specific detector modules of interest or assemble entire detector frames in a single stream, with modules grouped and processed together on the same machine as required to make optimal use of network bandwidth. The highest data rate among the currently used detectors is 9 Gbits per second for a single module and a total bandwidth of 140 Gbits per second for the entire assembly consisting of 16 modules. This output path is generally used for specialized real time analysis suites provided by facility users, which are tailored for each instrument via interfaces into the *Karabo* control system.

In addition to the built-in correction methods, custom processing code may be injected into the high-performance paths via so-called *correction add-ons*. In particular for implementations running on GPUs, this may take advantage of data already being present in device memory to perform further analysis after corrections. It should be noted that to preserve the monitoring aspect, the preview result always remains unchanged. An important application of the correction add-on mechanism is for the purpose of online data reduction. Any metadata generated by correction add-ons across all correction devices of a detector can be transported to a central arbiter device ahead of the actual detector data, where further custom code is executed in so-called reduction kernels. Here, the final decision can be made as to which data to include in the data stream available further downstream and make optimal use of the available bandwidth by minimizing the amount of data to be transferred and processed.

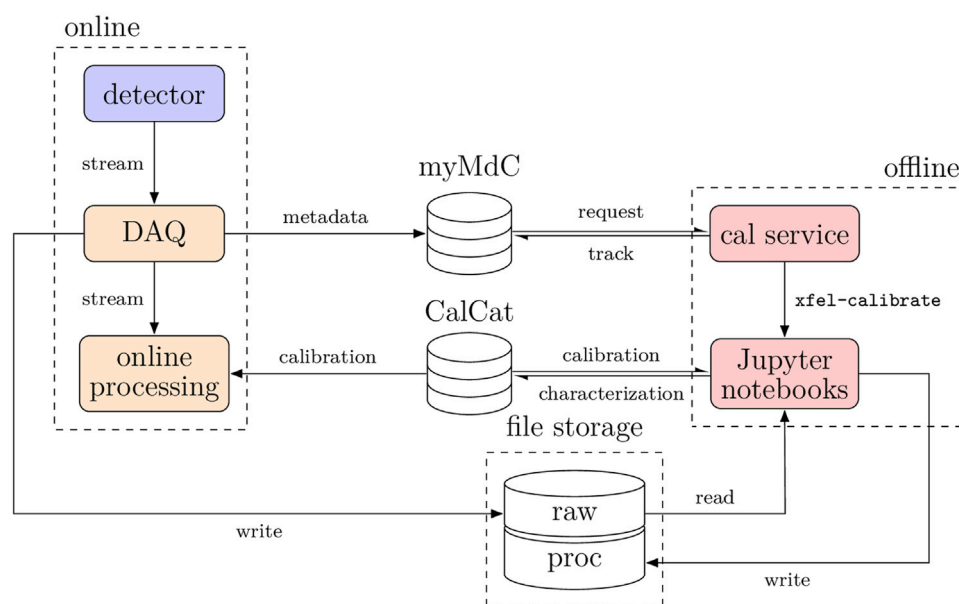


FIGURE 3

Overview of the services involved in the facility processing system at European XFEL. Detector data sent to the data acquisition system (DAQ) is sent as a stream to online processing systems and written to disk as raw data. For the files tracked by the myMdc application, requests can be sent to be calibration service to perform offline processing. This configures and runs Jupyter notebooks on the Maxwell compute cluster, interacting with the CalCat database to retrieve necessary calibration data or add new data after characterization. The resulting processed data is then stored and accessible alongside the raw data.

2.2 Offline data processing

Data saved to files constitutes the primary data product of beamtimes with more than 100 PB generated since European XFEL began user operation in 2017. It serves a critical role from early data exploration for decision-making during the beamtime to forming the basis of scientific publication. To address the FAIR [7] data principles and commitments of the common framework for scientific data management at photon and neutron facilities laid out by the PaN-data Europe Strategic Working Group [8], raw data is generally acquired via the facility data acquisition system (DAQ) and stored in the HDF5 format [9].

The raw data is complemented by data processed by the facility and saved in additional files alongside. This data is generated by a dedicated system either automatically upon the end of acquisition, or by explicit requests made through the data management portal for European XFEL users (myMdc). The portal is implemented as a web application and tracks all scientific data saved to disk with its physical storage location and metadata such as samples, techniques, and experiment types. Additionally, it includes administration of the experimental team, the electronic logbook, as well as management of digital object identifiers (DOIs) to the scientific data.

An overview of the service interactions and flows to generate this facility-processed data is provided in Figure 3. The requests triggered through myMdc are tracked and managed by the calibration service and run on the offline computing cluster Maxwell [10], which updates their status through the myMdc interface. The actual processing code at the heart of this system is implemented in Jupyter notebooks [11]. This allows the same code to scale from processing entire beamtimes automatically to manual,

interactive execution on a selected subset of data for exploration or development. Each notebook is identified by an *action* it performs on data of a particular *detector* and is written in such a way to receive input values using `nbparameterise` [12]. These specify the input data and the intended output location, as well as any other parameters in terms of format or scientific context. The `xfel-calibrate` runtime is then used to divide the workload and run several copies of the processing at once via the SLURM workload scheduler [13], each copy running on a subset of the data, spread across multiple compute nodes to maximise the efficient use of time and resources. At the end, the executed notebooks are compiled into a report documenting the processing, including plots intended for diagnostics and to monitor data quality. Additional single notebooks may be run before and after this central processing step to prepare the environment or reduce the results further.

Next to the automatic processing steps performed on acquired data, this system is also used to implement the characterization and generate the necessary calibration data for some of the aforementioned processing steps, e.g., image corrections. This processing differs in the degree of automation and interactivity depending on how often it has to be repeated and its robustness. The calculation of gain factors to calibrate intensity in absolute units is generally done by manual invocation of `xfel-calibrate` with suitable parameters, often running the underlying processing notebook manually first to exploring the parameter space. An example of fully automated characterization is the determination of baseline offsets from *dark* data, i.e., data in the absence of an external stimulus on the detector. It is performed at least daily during operation and triggered through the myMdc web application.

In both cases, the location of the generated calibration data is centrally indexed in a database called the Calibration Catalogue (CalCat). It is queryable by the detector identifier, the point in time the characterization took place and the conditions the data is applicable for, such as the sensor temperature, bias voltage or integration time. This enables the retrieval of the most suitable available calibration data anytime a given detector was used. Queries to CalCat happen both during the experiment using the current detector conditions, as well as for any data taken prior, with the respective conditions at the time of measurement. The conditions are described by key-value pairs, which are generally scalar numbers and have been assigned an allowed deviation at characterization time. Detectors are uniquely labelled throughout the facility by the physical detector unit (PDU) identifier, which is independent of their physical location at a particular experiment. This allows calibration data to seamlessly follow a detector to wherever it is used at a particular time, as long as calibration data does not depend strongly on the environment. In the case of multi-module detectors, each of these modules represents a single PDU to facilitate maintenance or exchange of individual modules. The calibration data stored in this database can be readily compared to past values for the same conditions to allow for regular monitoring by experts.

An important aspect of facility-processed data is reproducibility, which in this context denotes the ability to recreate the same output given the same input at a later point in time. Reproducibility aims to ensure a level of confidence in the scientific results derived from such data. It also alleviates the need to archive processed data in the longer term, as it can be recreated from archived raw data if needed. Here, it is important to acknowledge that in general, running the same arbitrary code irrespective of the software environment will not result in an identical result. Both changes to configuration and external services, e.g., calibration data received from CalCat, as well as differences in the lower lying soft- and hardware can lead to a numerically different result. Furthermore, as the processing code is developed further, its application to previous data may yield a different output than an earlier version.

For the offline processing system, reproducing an earlier result is therefore considered a distinct action from reprocessing it. It is implemented at the `xfel-calibrate` level, where for every invocation a special *metadata folder* contains all necessary parameters about the computation itself, the executed notebook with concretized parameters, the software environment it ran in as well as the captured responses from external services. A second command `xfel-calibrate-repeat` then uses this metadata directory to re-run the same code as before, with the same parameters, in a similar Python software environment, with the same external service responses. Some lower-level factors are not tracked in this implementation, such as the type of CPUs running the code or the compiler used for dependencies.

Essential for the data quality aspect of facility-provided processing is the continuous verification of its results. On a purely software engineering level, this is achieved by a wide coverage of unit tests [14] to test components individually. These tests are triggered automatically on every code change as part of a continuous integration workflow. In addition, an end-to-end approach from a scientific perspective is used, which processes data taken during regular user operation and compares the output against the expected result. For every supported

processing task, a collection of such reference data is curated alongside the intended and verified product. As part of ongoing improvements, this reference result is regularly replaced after manual examination. The list of configurations is also extended to cover significant or incompatible changes, e.g., a different data structure on the detector side, to ensure that newer and enhanced processing code also works on older data.

3 Results

The described machinery for facility-provided processing has been used for a wide range of actions, chief among them characterization and image corrections of the custom large-area 2D imaging detectors. Recently, this has been extended to special operating modes for these as well as entirely different but essential pre-processing steps for types of experiments not involving pixel-based detectors (see below).

3.1 Supported detectors and actions

A primary data driver of several instruments at European XFEL are the AGIPD, DSSC, and LPD detectors, which are developed specifically to exploit the unique burst mode time structure. As such, their uniqueness necessitated establishing new characterization and correction methods, and we consider it critical to offer an implementation ourselves. Common to all these systems, and a particular challenge for any applied processing method, is their very high data rate on the order of 100 Gbit/s for a Mpixel detector.

The *Adaptive Gain Integrating Pixel Detector* (AGIPD) [3] is a fast, integrating detector in the hard X-ray regime with adaptive gain. It offers single photon sensitivity at 12 keV and a dynamic range of up to 10^4 12 keV photons while being able to take up to 352 consecutive images at the facility's pulse repetition of 4.5 MHz. This image burst is then read out at 10 Hz between pulse trains (compare Figure 1), resulting in a total frame rate of up to 3,520 Hz. There are currently two 1 Mpixel installations consisting of 16 modules each in use at the SPB/SFX [15] and MID [16] instruments, as well as another 0.5 Mpixel prototype system with 8 modules and an upgraded version of the readout ASIC (application-specific integrated circuit) at the HED instrument [17]. A rich set of image corrections is implemented for this detector. First, the gain stage each pixel was recorded in is chosen through a threshold procedure followed by offset subtraction. Both the required threshold and offset values are inferred from dark image characterization automatically performed during operation in regular intervals. In certain scenarios, baseline shifts and common mode effects, both spatially per ASIC and temporally across multiple trains, can be accounted for. Finally, gain calibration converts pixel amplitudes to intensity in units of absolute energy. Several methods have been established to obtain the necessary slope characterization data and implemented as part of the facility-processing package, and are generally invoked manually. A detailed description of this detector and its calibration can be found in [18].

The *Large Pixel Detector* (LPD) [4] is another fast detector system acquiring up to 512 images at 4.5 MHz in three parallel gain stages. From these stages, an auto-gain mode can choose the optimal signal to resolve up to 10^5 12 keV photons. A Mpixel installation

with 16 modules is in use at the FXE hard X-ray instrument [19] alongside several smaller single-module detectors called *LPD Mini*. The image corrections consist of the basic steps of offset subtraction based on automatically characterized dark images and subsequent gain calibration by pre-determined slopes obtained by manual analysis [20].

At the soft X-ray instruments SCS [21] and SQS [22], the *DEPFET Sensor with Signal Compression* (DSSC) [5] is available. This camera operates at a peak frame rate of 4.5 MHz and features on-chip digitization of 1-Mpixel images for up to 800 images. There are two versions of this camera, each employing different sensor technologies. The first version, which uses *MiniSSD* technology, has been in user operation since 2019. It offers several gain configurations to accommodate a broad range of soft X-ray photon energies. The camera exhibits a linear intensity response and thus, only offset subtractions are necessary before proceeding with further analysis. The second version of the camera utilizes *DEPFET* sensors and is currently in the commissioning phase. This technology offers superior noise performance, with an equivalent noise charge averaging 16 e⁻ rms [23]. It enables single-photon imaging capabilities down to a photon energy of 0.25 keV, all while maintaining a dynamic range of up to 10⁴. However, its nonlinear response necessitates additional correction steps to convert ADC counts into photon energy. These are currently in development.

In addition to these large-area burst mode detectors, several other X-ray pixel detectors are used across the instruments with corresponding support for characterization and image corrections provided by the facility. The *JUNGFRAU* [24], *ePix100* [25] and *pnCCD* [26] are 2D frame-based detectors known from other facilities with robust and mature processing methods available in literature and upon which dark image pedestal subtraction, common mode, and gain calibrations are based. In the burst mode operation at European XFEL, these detectors generally are unable to record the intra-train pulses, hence they are only operated at the train repetition rate of 10 Hz. The *Gotthard-II* [27] is a 1D strip detector developed at the Paul Scherrer Institute for use at European XFEL capable of matching the pulse repetition rate. It is particularly suited to spectroscopic measurements. Here, the corrections also include an essential linearization of the raw output of the analog-to-digital converter (ADC) before subtracting offset and calibrating intensity to absolute units. Common to these detectors are considerably lower bandwidth requirements either due to their operation at only the 10 Hz train repetition rate or due to the smaller data volume of a single frame.

A different set of processing actions is available for a detector built on *Timepix3* [28], a time-resolved and event-driven pixel read-out chip. One such device is in use at the SQS instrument and is primarily used for electron and ion spectroscopy. Rather than full frames, it acquires individual time-over-threshold events for each of its pixels. A time walk correction is offered alongside centroiding to group neighbouring pixels illuminated at the same time into single particle impacts, if applicable. The calibration data required for the former correction process is currently acquired and prepared manually in a similar fashion to gain calibrations for frame-based detectors, but planned to be further automated in the future like dark image characterization is.

The SQS instrument also employs time and position sensitive delay line detectors [29] for charged particle and photon spectroscopy and imaging techniques such as REMI (reaction microscope) [30, 31]. Here the reconstruction process to assemble concrete particle impacts on the detector is entirely implemented as part of the facility-based processing systems, starting from digitized traces in the acquired raw data. After common mode correction of the analog data and discrimination to pulse arrival times, these digital signals on each channel are sorted into tuples corresponding to the same detector hit. For optimal resolution and reconstruction quality, further time sum and position correction on the digital signals and sophisticated event sorting based on components of the vendor-provided *CoboldPC* package can be included. In this application, the output format involving time and position events differs entirely from the initial input of analog voltage signals.

3.2 Special operating modes

In general, the core functionality of the implemented processing actions aims to be generic and experiment-agnostic apart from tuning parameters. Over the course of facility operation, however, more and more toggle-able operation modes have been added to aid users in data preparation procedures particular to their beamtime or technique.

In the case of offline analysis, this allows us to automatically enjoy the same benefits of reproducibility and scalability for these steps. While these special operating modes are in most cases not specific to a detector, their exclusive use at a particular instrument typically ties them to one detector and is thus implemented as part of its processing notebook.

One example of this is the generation of *virtual CXI* files for serial femtosecond crystallography (SFX) experiments [32] with the *LPD* detector after image corrections. The native data format of the Coherent X-ray Imaging Data Bank [33] specifies a particular layout of HDF5 files for SFX experiments, and analysis software developed in this scientific community can often use these files directly. By generating these files using HDF5 *virtual datasets* to refer to the corrected result in European XFEL's data format, this is possible without the need for an additional full copy on disk while being immediately available for users after acquisition and processing.

Another example influencing the actual data result is *photonization* available for the *AGIPD* detector [34]. Under certain illumination conditions commonly present at the MID instrument, pixel intensity after gain calibration can be interpreted as singular photon events of a particular photon energy and represented by an integer count. Performing this operation during image corrections can be implemented particularly efficiently for immediate analysis based off it. Furthermore, it serves as a data reduction technique, as the resulting integer representation is significantly more compressible, resulting in space savings of up to 97% at a negligible runtime cost [35].

Some operating modes are exclusively for data reduction purposes before processing takes place, e.g., limiting the trains or frames within a train to be included in the result. Even in the case of the *AGIPD* detector, different methods are employed depending on the experimental environment. At the SPB and MID instruments, the so called *LitFrameFinder* software automatically aligns the X-ray pulse pattern with the detector frame pattern to discard any frames in the processed data output not directly illuminated by X-rays. At

the HED instrument, however, an optical chopper device is used to pick out entire pulse trains irrespective of the actual pulse filling pattern, and the processing result is thus reduced by exploiting this pulse picker information.

In the context of online processing, the main goal of offering special purpose analysis directly into the facility-provided processing system is exploiting data locality, especially in the case of high-bandwidth data accessed on GPUs. Currently, this is primarily done for real-time image corrections of large-area detectors via the mechanism described in 2.1, with the first implementations covering the computation of integrated intensity, counting lit pixels, and performing peak-finding in the context of SFX. In most other cases, special requirements in an online setting are often not mature or standard enough to warrant an application within these systems, and are left to more flexible online analysis solutions developed in-house [36] or those from the corresponding scientific communities.

3.3 Processing performance

For all detector implementations listed in 3.1, the online image corrections performed in real-time are able to cope with incoming data rates delivered as data streams. In the case of the AGIPD, DSSC, and LPD large-area detectors however, certain rate limitations still remain when actually moving the full data through the network.

These occur in particular when all modules are desired on a single online cluster machine, entirely exhausting its network links. A single group of correction devices each processing a single module on the other hand can generally be transported at full rate. For a given experiment, the ideal distribution can therefore be chosen, e.g., concentrating modules critical for analysis in the same group. Further optimizations are possible by reducing the precision, e.g., to half-precision float16 per pixel as well as applying frame selections. Statically frames can be selected either by skipping entire trains or choosing fewer frames in each train, while calibration add-ons allow dynamic frame selection based on user-implemented criteria. Typically, up to 500 frames per second are currently achievable when assembling all modules to full frames for a single destination.

The achievable performance for basic image corrections is generally comparable on CPUs and GPUs when the total turnaround including memory transfer costs is considered. Still, these devices can perform these tasks more energy efficient and furthermore offer a surplus of computing resources that can be exploited for analysis or transformation steps. These can be integrated much easier with the data already local in device memory at that time. Currently it includes computational tasks like peak finding and azimuthal integration described in 3.2 as well as memory-intensive axis stacking and reordering when requested by downstream user applications using this data. At the same time, the CPUs remain available to perform processing tasks not well suited or not yet implemented on GPU architectures.

To feed this to analysis suites running outside of Karabo, the lightweight Karabo-bridge protocol is available (in the illustrations, each “matcher” can have one or more bridge outputs), with Python [37] and C++ [38] bindings for convenient integration with existing software. This has been used, for example, to feed data to Hummingbird [39] or OnDA [40], two packages in the field of X-ray imaging experiments developed by the scientific community and widely used across facilities.

The offline processing system aims to operate as efficiently as possible on data already fully present in files. Generally, these files are grouped into datasets called *runs*, which contain all the data collected during manually operated triggers. The data volume is automatically split into independent parts to parallelize the work in multiple jobs across the Maxwell computing cluster. The automatic creation of PDF reports and capturing other metadata relevant to reproducibility typically takes less than 1 minute in a trailing job after all processing jobs have returned. The distribution of resources on the Maxwell compute cluster prioritizes currently ongoing experiments to guarantee near-immediate allocation of nodes without any delay. At this time, swift access to processed results for interpretation is particularly valuable to make optimal use of beamtime. After the experiment has concluded, processing requests are queued with the same priority as regular users and may incur waiting times. On average, this results in about 650 jobs per day but also peaked at 2,200 jobs over a single day in the current year of operations. The per-week job statistics for the current year of operations are visualized in Figure 4.

For image corrections and other automatic data pre-processing tasks, each job is assigned a set of trains along file boundaries. The image corrections for the large-area pixel detectors with burst mode generally achieve on the order of 1,500–5,000 frames per second of a single module in each job, compared to data acquisition rates of up to 3,520 frames per second in the case AGIPD, for example. The pixel detectors not operated in burst mode and hence acquired at 10 Hz typically reach up to 140 frames per second. For the non-frame based detectors performance depends significantly on the experimental event rates, but train centroiding for *Timepix* is generally processed at least with 60 Hz, while *REMI* reconstruction varies between 40 and 80 Hz. By parallelization of larger datasets across multiple SLURM jobs, this generally matches or exceeds real-time across all processing options factoring in disk I/O and the constant costs of set-up and tear-down cluster jobs and processes.

In the case of the automatic dark characterization of pixel detectors, however, jobs are generally assigned all data belonging to a single detector module, if applicable. For burst mode detectors, these generally have runtimes of 10 min or less depending on the number of memory cells and trains considered for statistics. The runtimes of characterizing dark data of other detectors are negligible and complete within a minute or less. The non-automatic characterization tasks for gain calibration are not written with performance but completeness and traceability in mind with runtimes on the order of hours, as they are only repeated a few times per year.

For special circumstances where exceptionally fast file-based feedback is required, the offline processing machinery is capable of running on the online cluster usually reserved for streaming applications. This can take advantage of extremely fast disk I/O and avoid delays until data is accessible on the Maxwell compute cluster, but is limited to a too small number of machines to warrant the same level of parallelization. It is therefore typically restricted to cases with a high level of data reduction, e.g., pulse on demand techniques, that can take advantage of fast file-based data exploration. To this purpose, many of the features geared towards reproducibility and data tracking can be turned off to minimize run time to processing only. For example, corrected data written to file for a single AGIPD train with 352 frames can be made available within 30 s of acquisition.

The reproducibility of offline processing was confirmed for four different detector types (AGIPD, LPD, JUNGFRU and ePix100)

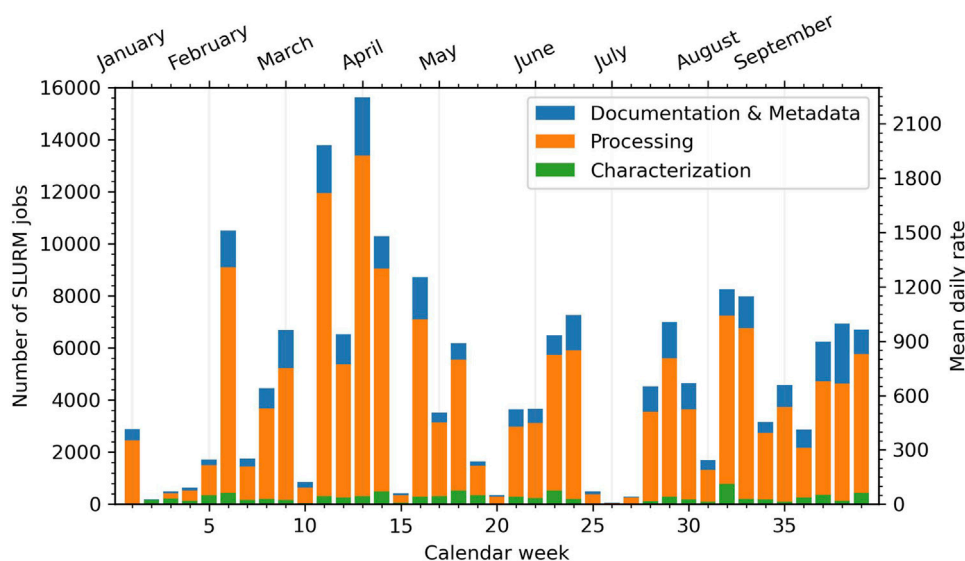


FIGURE 4

Number of processing jobs executed per week on the Maxwell SLURM cluster as part of the offline processing systems during the operations in 2023. Characterization jobs evaluate detector performance and generate calibration data, such as noise and pixel offsets. Processing jobs then transform experimental datasets, for example, by applying image corrections to 2D X-ray detector data. Documentation and metadata jobs run alongside all these actions to compile PDF reports with their results and capture metadata important for reproducibility.

using data from just after the reproducibility work was completed for each detector. This data was 18 months old for AGIPD at the time of testing, and 12 months for other detectors. In each case, the code used in the previous processing ran successfully and produced output data identical to the original results.

4 Discussion

4.1 Evolution of concept

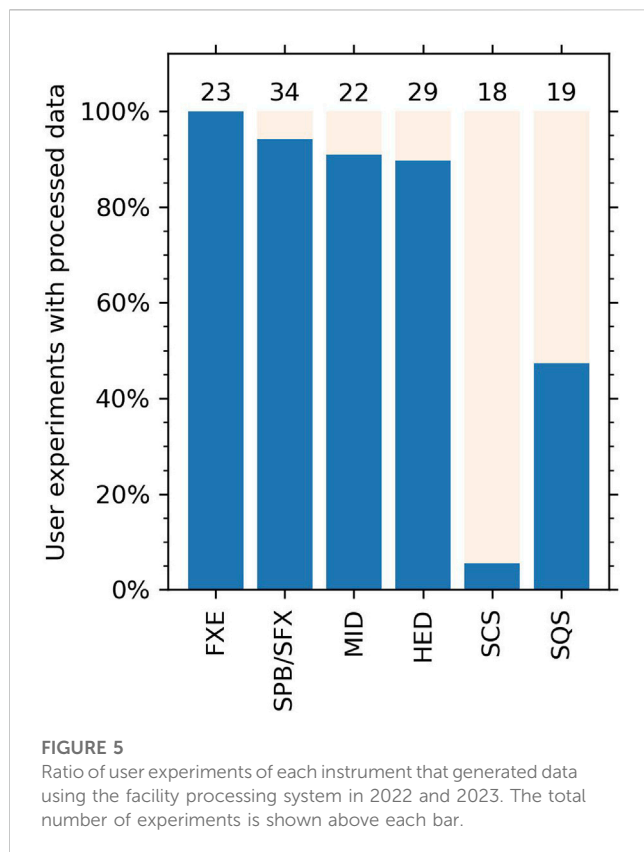
Prior to facility operation and facing the challenges ahead, the concept for such a system was already planned and described in [41], with a particular focus on the AGIPD, DSSC and LPD detectors. On a conceptual level, significant differences can be found and discussed between the current implementation and the original expectations in three areas: i) online analysis, ii) access to raw data, and iii) configurability. This stems from experience accumulated during operation and an increasing diversity of scientific applications.

The online monitoring of experimental data was not expected to deliver the full input rate, but rather a continuous and non-guaranteed stream targeted at visual monitoring by operators. While this is critical for monitoring from a technical and detector operation perspective, the diversity of metrics significant for online analysis from a user perspective has proven to be much wider. Beyond immediate and in particular visual evaluation of immediate detector signals, the near real-time evaluation of technique-specific quantities based on as many detector frames as possible can significantly enhance the efficiency of the running experiment. For example, single particle imaging (SPI) [42] typically suffers from very low hit rates for interaction of the

nanoscale targets with X-rays. An estimation of this rate during the experiment is more robust with access to at least some data of every acquired frame, rather than the entire data of a fraction of frames. This is primarily addressed by a flexible topology of the data stream to tailor any necessary compromise to each use case.

With the presence of a facility-provided processing system, no direct access to raw data affected by these systems was foreseen to be necessary by users. After calibrated files are produced by these systems, it was to only serve as the main archival data product, to be used when those files are no longer present in temporary storage. However, the operational experience so far has shown that a single truth for correction methods—particularly for this custom hardware, but also generally for the plethora of different experiments—has not been found. Established methods, like the aforementioned SFX, can almost exclusively rely on the processed datasets already, but this is not universal for other, often still developing, or novel techniques. Access to raw data remains essential for users with different requirements, either to expand on the processing methods already offered by us or replace them entirely with custom implementations. To this purpose, the utilized calibration data is available to users through the CalCat database as well. The collaboration on user's data treatment methods in this way allows to continuously adapt any improvements to the facility systems for the profit of all scientific users. Here, the distinction between reproducing an earlier result and reprocessing the same raw data was also underestimated, as it is possible that the current implementation may be newer and improved, giving a potentially better, yet different result. In the context of the scientific method, however, the capability to obtain an equal result is as essential as well.

To maximize the applicability to as many experiments as possible, and enable *user data* to be the data product sufficient for scientific analysis, processing cannot be limited to standardized blocks. The



emerging need for configurability covers immediate and often empiric parameters for computations, the implementation of special operating modes as discussed in 3.2, and tolerance to a constantly changing environment without stable interfaces. Even in the originally envisioned case of 2D imaging detectors, including such configurability in fundamental steps, like image correction and gain calibration, allows for greater adaptation to moving requirements. In contrast, limiting to standard methods risks focusing on the needs of established communities to the detriment of emerging fields. Instead, sensible defaults can aim to making common use cases as efficient as possible while leaving the option to expand for uncommon use cases as well. Parameters may range from flags for which algorithms are applied to tuneable numerical thresholds, but also includes technical configuration, like pipeline topology, for efficient network usage as described for 2.1. Whenever applicable, they are inferred from available metadata and hardware conditions and configured automatically. What remains may require manual tuning not just by operators and experts on the facility-side, but users as well. In line with the previous area of raw data access, this necessitates transparency and accessibility in this process. Finally, reliable operation of all these steps relies on countless implementation details of hard- and software upstream, which are themselves a moving target as the facility develops.

4.2 User operation

The facility processing system as presented here is in a mature and stable state for user operation at European XFEL. Since inception, it has seen widespread use during experiments

performed at the facility illustrated in Figure 5 for 2022 and 2023. More than 90% of experiments across the hard X-ray instruments—namely, FXE, SPB/SFX, MID, and HED—are generating processed data as part of their data product over this time period. In a sharp contrast, approximately half of all experiments at SQS and only a single experiment at SCS took advantage of this system. These soft X-ray instruments differ significantly from the hard X-ray instruments in experimental techniques and thus detection methods. As discussed previously in 4.1, the standardized blocks focused originally around image corrections for 2D pixel detectors were not flexible enough for, or did not at all cover, the requirements and use cases of the soft X-ray instruments. In total however, 75% of the user experiments performed at European XFEL in 2022 or 2023 were assisted by the presented system in its current state.

A significant impact of processed data being generated seamlessly and automatically by facility systems has been observed on the storage infrastructure. The data volume of the large area detectors, like AGIPD, can effectively be doubled after acquisition by having both copies of raw and processed data on disk at the same time, culminating in single experiments exceeding multiple petabytes in a single week. This has been primarily mitigated by the introduction of reproducibility, which allows to only keep processing results on storage when they are actively in use for analysis, as they can be safely recreated after deletion. In addition, it stimulated the development of data reduction techniques that are applied during these processing steps. As such reduced processed data is the result of the same still unmodified raw data, these represented ideal opportunities for research of such methods and their validation.

These systems are provided for and continuously monitored by experts from the Data Analysis and Detector groups at European XFEL with expertise in software development, data analysis and detector characterization. These experts also provide 24/7 on-call support for user experiments. The established testing infrastructure verifies that expected data quality is preserved, and further quality improvements are confirmed manually in dedicated commissioning campaigns.

4.3 Outlook

Driven by the evolution of the calibration and processing concept offered by European XFEL as discussed in 4.1, further improvements are in development. These focus on increasing the flexibility and performance of both online and offline processing with data reduction in mind throughout the entire process.

While the online processing pipelines are able to process the data volume of all burst mode detectors at their full input rate, moving this data across the network still imposes limitations. On the technical side, Remote Direct Memory Access (RDMA) technology will increase the achievable bandwidth in the near future between DAQ and the processing infrastructure. This is combined with expanding on the correction add-on mechanism to exploit data being present in the memory of high-performance GPUs to benefit custom user analysis as well. In fact, these devices are increasingly used to accelerate time-consuming analysis tasks in the analysis of X-ray experiments and enable their real-time application, in particular for techniques based on machine

learning [43–46]. These capabilities are also the ideal place to apply data reduction. As a result of the deep integration into the facility systems at this point, these decisions can not only be used to alleviate the bandwidth limitations in the online pipeline, but also reduce data before it hits the file storage. This removes the need for additional data reduction steps after writes have been performed.

For offline processing, the main focus is to use the existing scalable machinery for a wider catalogue of generic analysis steps beyond the facility-provided actions described in 3.1. Building on the first experiences here in the form of the special operating modes, this should expand to steps commonly re-implemented for many experiments as these are often agnostic of the underlying detector, such as azimuthal integration of detector frames for X-ray scattering experiments. Apart from the potential for highly optimized implementations, these can enjoy the same reproducibility guarantees as the existing actions. In the second step of this process, this machinery can be opened up to users to run their own, fully custom, analyses. The aim is for this to serve as a generic and accessible runtime for automatic, configurable, dataset-based offline analysis. This provides users instantly with a broad infrastructure related to managed code execution, monitoring, and parallelization. Both of these goals are contingent on further developments on the configurability and in particular interfaces first, like in the form of the myMdc web application. An important lesson already learned here for such an automated system is to clearly document and communicate the situations for which a particular analysis can be applied, for example, the photonization of absolute energy scales currently offered for the *AGIPD* detector. When the assumptions for such a method are not met, perhaps unknowingly, their application can result in a diminished data quality contrary to its intended purpose. Finally, this extends to more support for interoperability with existing solutions in the scientific community, for example, through the *NeXus* data format [47].

4.4 Summary and conclusion

A comprehensive and scalable system for processing scientific data has been developed at European XFEL. It aids users in essential preparatory processing steps, which are increasingly challenging due to high data rates and the use of custom detector technologies. Experimental data is delivered to the majority of user groups in a form suitable for further analysis at a constantly verified data quality. Future developments are focused on more support for a wider range of experimental techniques and integration of data reduction techniques.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://git.xfel.eu/calibration>.

Author contributions

PS: Conceptualization, Data curation, Investigation, Methodology, Project administration, Software, Validation,

Visualization, Writing—original draft, Writing—review and editing. KA: Data curation, Methodology, Software, Validation, Visualization, Writing—review and editing. CD: Data curation, Software, Writing—review and editing. DH: Data curation, Methodology, Software, Validation, Writing—review and editing. RR: Methodology, Software, Validation, Visualization, Writing—review and editing. TK: Methodology, Software, Validation, Visualization, Writing—review and editing. TM: Methodology, Software, Writing—review and editing. ES: Methodology, Software, Writing—review and editing. LG: Conceptualization, Funding acquisition, Resources, Supervision, Writing—review and editing. LM: Methodology, Software, Writing—review and editing. MM: Methodology, Software, Writing—review and editing. JM: Methodology, Resources, Software, Writing—review and editing. KW: Resources, Supervision, Writing—review and editing. JS-D: Data curation, Investigation, Methodology, Validation, Writing—review and editing. VR: Data curation, Investigation, Methodology, Software, Validation, Writing—review and editing. MR: Methodology, Validation, Writing—review and editing, Data curation, Investigation. ND: Data curation, Investigation, Methodology, Software, Validation, Writing—review and editing. DL: Data curation, Investigation, Methodology, Validation, Writing—review and editing. ID: Investigation, Methodology, Writing—review and editing. HY: Data curation, Investigation, Methodology, Software, Validation, Writing—review and editing. BS: Data curation, Investigation, Methodology, Software, Validation, Writing—review and editing. OM: Project administration, Resources, Writing—review and editing. MT: Funding acquisition, Resources, Supervision, Writing—review and editing. SH: Conceptualization, Investigation, Methodology, Software, Writing—review and editing. SA: Funding acquisition, Resources, Supervision, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The effort for the method, systems and infrastructure described in this work is carried by the entire staff of European XFEL. In particular all parts of the IT infrastructure maintained by the European XFEL ITDM and Deutsches Elektronen-Synchrotron (DESY) IT groups are critical for its operations. It relies on the reliable operation of detectors attended by instrument detector experts and other instrument staff. In addition, the authors acknowledge countless facility users that contributed feedback and ideas.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

References

- Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, et al. A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nat Photon* (2020) 14:391–7. doi:10.1038/s41566-020-0607-z
- Tschentscher T. Investigating ultrafast structural dynamics using high repetition rate x-ray FEL radiation at European XFEL. *The Eur Phys J Plus* (2023) 138:274. doi:10.1140/epjp/s13360-023-03809-5
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Radiat* (2019) 26:74–82. doi:10.1107/S1600577518016077
- Veale M, Adkin P, Booker P, Coughlan J, French M, Hart M, et al. Characterisation of the high dynamic range Large Pixel Detector (LPD) and its use at X-ray free electron laser sources. *J Instrumentation* (2017) 12:P12003. doi:10.1088/1748-0221/12/12/P12003
- Porro M, Andricek L, Aschauer S, Castoldi A, Donato M, Engelke J, et al. The MiniSDD-based 1-mpixel camera of the DSSC Project for the European XFEL. *IEEE Trans Nucl Sci* (2021) 68:1334–50. doi:10.1109/TNS.2021.3076602
- Hauf S, Heisen B, Aplin S, Beg M, Bergemann M, Bondar V, et al. The Karabo distributed control system. *J Synchrotron Radiat* (2019) 26:1448–61. doi:10.1107/S1600577519006696
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* (2016) 3:160018. doi:10.1038/sdata.2016.18
- H2020-EU14 – EXCELLENT SCIENCE – Research Infrastructures. *Photon and neutron open science cloud* (2018). doi:10.3030/823852
- Koranne S. *Hierarchical data format 5: HDF5*. Handbook of open source tools. Berlin, Germany: Springer US (2011). p. 191–200. doi:10.1007/978-1-4419-7719-9_10
- Deutsches Elektronen-Synchrotron. Maxwell-cluster (2023). Available at: <https://confluence.desy.de/display/MXW/Documentation>.
- Granger BE, Pérez F. Jupyter: thinking and storytelling with code and data. *Comput Sci Eng* (2021) 23:7–14. doi:10.1109/MCSE.2021.3059263
- Kluyver T. nbparameterise (2014). Available at: <https://github.com/takluyver/nbparameterise>.
- Jette MA, Wickberg T. *Job scheduling strategies for parallel processing*. Berlin, Germany: Springer (2023). p. 3–23. doi:10.1007/978-3-031-43943-8_1
- Hulzinga D, Kolawa A. *Automated defect prevention: best practices in software management*. Hoboken, New Jersey: Wiley-IEEE Computer Society Press (2007).
- Mancuso AP, Aquila A, Batchelor L, Bean RJ, Bielecki J, Borchers G, et al. The single particles, clusters and biomolecules and serial femtosecond crystallography instrument of the European XFEL: initial installation. *J Synchrotron Radiat* (2019) 26:660–76. doi:10.1107/S1600577519003308
- Madsen A, Hallmann J, Ansaldi G, Roth T, Lu W, Kim C, et al. Materials imaging and dynamics (MID) instrument at the European X-ray free-electron laser facility. *J Synchrotron Radiat* (2021) 28:637–49. doi:10.1107/S1600577521001302
- Zastrau U, Appel K, Baetz C, Baehr O, Batchelor L, Berghäuser A, et al. The high energy density scientific instrument at the European XFEL. *J Synchrotron Radiat* (2021) 28:1393–416. doi:10.1107/S1600577521007335
- Sztuk-Dambietz J, Klackova I, Klyuev A, Laurus T, Trunk U, Ahmed K, et al. Operational experience with adaptive gain integrating pixel detectors at European XFEL. *Front Phys* (2023).
- Galler A, Gawelda W, Biednov M, Bommer C, Britz A, Brockhauser S, et al. Scientific instrument Femtosecond X-ray Experiments (FXE): instrumentation and baseline experimental capabilities. *J Synchrotron Radiat* (2019) 26:1432–47. doi:10.1107/S1600577519006647
- Wheater R, Hart M, Veale M, Wilson M, Doblas-Jiménez D, Turcato M, et al. Development of data correction for the 1M large pixel detector at the EuXFEL. *J Instrumentation* (2022) 17:P04013. doi:10.1088/1748-0221/17/04/P04013
- Tschentscher T, Bressler C, Grünert J, Madsen A, Mancuso AP, Meyer M, et al. Photon beam transport and scientific instruments at the European XFEL. *Appl Sci* (2017) 7:592. doi:10.3390/app7060592
- Mazza T, Baumann TM, Boll R, De Fanis A, Grychtol P, Ilchen M, et al. The beam transport system for the Small Quantum Systems instrument at the European XFEL: optical layout and first commissioning results. *J Synchrotron Radiat* (2023) 30:457–67. doi:10.1107/S1600577522012085
- Castoldi A, Ghisetti M, Guazzoni C, Aschauer S, Strüder L, Hansen K, et al. Qualification of the X-ray spectral performance of the DEPFET pixels of the DSSC imager. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168686. doi:10.1016/j.nima.2023.168686
- Mozzanica A, Andrä M, Barten R, Bergamaschi A, Chirioti S, Brückner M, et al. The JUNGFRU detector for applications at synchrotron light sources and XFELs. *Synchrotron Radiat News* (2018) 31:16–20. doi:10.1080/08940886.2018.1528429
- Blaj G, Caragiulo P, Dragone A, Haller G, Hasi J, Kenney CJ, et al. X-ray imaging with ePix100a: a high-speed, high-resolution, low-noise camera. In: James RB, Fiederle M, Burger A, Franks L, editors. *Hard X-ray, gamma-ray, and neutron detector Physics XVIII*. Bellingham, Washington USA: International Society for Optics and Photonics (2016). doi:10.1117/12.2238136
- Meidinger N, Andritschke R, Hartmann R, Herrmann S, Holl P, Lutz G, et al. pnCCD for photon detection from near-infrared to X-rays. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2006) 565:251–7. doi:10.1016/j.nima.2006.05.006
- Zhang J, Andrä M, Barten R, Bergamaschi A, Brückner M, Chirioti-Alvarez S, et al. Design and first tests of the gotthard-II readout ASIC for the European X-ray free-electron laser. *J Instrumentation* (2021) 16:P04015. doi:10.1088/1748-0221/16/04/P04015
- Poikela T, Plosila J, Westerlund T, Campbell M, Gaspari MD, Llopart X, et al. Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout. *J Instrumentation* (2014) 9:C05013. doi:10.1088/1748-0221/9/05/C05013
- Jagutzki O, Mergel V, Ullmann-Pfleger K, Spielberger L, Spillmann U, Dörner R, et al. A broad-application microchannel-plate detector system for advanced particle or photon detection tasks: large area imaging, precise multi-hit timing information and high detection rate. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2002) 477:244–9. doi:10.1016/S0168-9002(01)01839-3
- Ullrich J, Moshhammer R, Dorn A, Dörner R, Schmidt LPH, Schmidt-Böcking H. Recoil-ion and electron momentum spectroscopy: reaction-microscopes. *Rep Prog Phys* (2003) 66:1463–545. doi:10.1088/0034-4885/66/9/203
- Boll R, Schäfer JM, Richard B, Fehre K, Kastirke G, Jurek Z, et al. X-ray multiphoton-induced Coulomb explosion images complex single molecules. *Nat Phys* (2022) 18:423–8. doi:10.1038/s41567-022-01507-0
- Barends TRM, Stauch B, Cherezov V, Schlichting I. Serial femtosecond crystallography. *Nat Rev Methods Primers* (2022) 2:59. doi:10.1038/s43586-022-00141-7
- Maia FRNC. The coherent X-ray imaging Data Bank. *Nat Methods* (2012) 9: 854–5. doi:10.1038/nmeth.2110
- Dallari F, Reiser M, Lokteva I, Jain A, Möller J, Scholz M, et al. Analysis strategies for MHz XPCS at the European XFEL. *Appl Sci* (2021) 11:8037. doi:10.3390/app11178037
- Sobolev E, Schmidt P, Malka J, Hammer D, Boukhelef D, Möller J, et al. Data reduction activities at European XFEL: early results. *Front Phys* (2023).
- Fangohr H, Aplin S, Barty A, Beg M, Bondar V, Boukhelef D, et al. Data analysis support in karabo at European XFEL. Proceedings of the 16th International Conference on Accelerator and Large Experimental Control Systems (2018) Barcelona, Spain, October 2018.
- European-XFEL. karabo-bridge-py (2018). Available at: <https://github.com/European-XFEL/karabo-bridge-py>.
- European-XFEL. karabo-bridge-cpp (2018). Available at: <https://github.com/European-XFEL/karabo-bridge-cpp>.
- Daurer BJ, Hantke MF, Nettelblad C, Maia FRNC. Hummingbird: monitoring and analyzing flash X-ray imaging experiments in real time. *J Appl Crystallogr* (2016) 49: 1042–7. doi:10.1107/S1600576716005926
- Mariani V, Morgan A, Yoon CH, Lane TJ, White TA, O'Grady C, et al. OnDA: online data analysis and feedback for serial X-ray imaging. *J Appl Crystallogr* (2016) 49: 1073–80. doi:10.1107/S1600576716007469

41. Kuster M, Boukhelef D, Donato M, Dambietz JS, Hauf S, Maia L, et al. Detectors and calibration concept for the European XFEL. *Synchrotron Radiat News* (2014) 27: 35–8. doi:10.1080/08940886.2014.930809
42. Chapman HN. X-ray free-electron lasers for the structure and dynamics of macromolecules. *Annu Rev Biochem* (2019) 88:35–58. doi:10.1146/annurev-biochem-013118-110744
43. Wang C, Florin E, Chang HY, Thayer J, Yoon CH. SpeckleNN: a unified embedding for real-time speckle pattern classification in X-ray single-particle imaging with limited labeled examples. *IUCrJ* (2023) 10:568–78. doi:10.1107/S2052252523006115
44. Ekeberg T, Engblom S, Liu J. Machine learning for ultrafast X-ray diffraction patterns on large-scale GPU clusters. *Int J High Perform Comput Appl* (2015) 29:233–43. doi:10.1177/1094342015572030
45. Ignatenko A, Assalauova D, Bobkov SA, Gelisio L, Teslyuk AB, Ilyin VA, et al. Classification of diffraction patterns in single particle imaging experiments performed at x-ray free-electron lasers using a convolutional neural network. *Machine Learn Sci Tech* (2021) 2:025014. doi:10.1088/2632-2153/abd916
46. Zhang Y, Yao Z, Ritschel T, Villanueva-Perez P. ONIX: an X-ray deep-learning tool for 3D reconstructions from sparse views. *Appl Res* (2023) 2:e202300016. doi:10.1002/appl.202300016
47. Könnecke M, Akeroyd FA, Bernstein HJ, Brewster AS, Campbell SI, Clausen B, et al. The NeXus data format. *J Appl Crystallogr* (2015) 48:301–5. doi:10.1107/S1600576714027575



OPEN ACCESS

EDITED BY

Jiaguo Zhang,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY

Giacomo Borghi,
Polytechnic University of Milan, Italy
Jonathan Correa,
Helmholtz Association of German Research
Centres (HZ), Germany

*CORRESPONDENCE

J. Ninkovic,
✉ Jelena.ninkovic@hll.mpg.de

RECEIVED 13 October 2023

ACCEPTED 09 January 2024

PUBLISHED 23 January 2024

CITATION

Ninkovic J, Bähr A, Richter RH and Treis J
(2024), Novel sensor developments for photon
science at the MPG semiconductor laboratory.
Front. Phys. 12:1321164.
doi: 10.3389/fphy.2024.1321164

COPYRIGHT

© 2024 Ninkovic, Bähr, Richter and Treis. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Novel sensor developments for photon science at the MPG semiconductor laboratory

J. Ninkovic*, A. Bähr, R. H. Richter and J. Treis

Semiconductor Laboratory of the Max-Planck-Society, Munich, Germany

The world of photon science experiences significant advancements since the advent of synchrotron light sources with unprecedented brilliance, intensity and pulse repetition rates, with large implications on the detectors used for instrumentation. Here, an overview about the work on this field carried out at the semiconductor laboratory of the Max-Planck-Society (MPG HLL) is given. Main challenges are high dynamic range to resolve faint features at the fringes of scatter images as well as structures in bright peaks, and high bandwidth to fully exploit the fast timing capability of the source. A newly developed device to improve the signal-to-noise-ratio (SNR) at high bandwidths is the so-called MARTHA (Monolithic Array of Reach-Through Avalanche Photodiodes) structure, which integrates an array of APDs on a monolithic substrate. The reach-through architecture assures near 100% fill factor and allows implementing a thin entrance window with optimized quantum efficiency for low energy X-rays. The structures operate in proportional mode with adjustable gain, and can serve as a drop-in replacement for PAD detectors in hybrid pixel systems. A more sophisticated solution for low to medium frame rate applications with high contrast requirement are pnCCDs with high dynamic range in the pixel area featuring DEPFET based readout nodes with non-linear amplification (NLA). The high dynamic range mode has been demonstrated for pnCCD devices with a pixel size down to $75\ \mu\text{m}^2$. Framerates of up to 1 kHz are possible for a 1 Megapixel detector. Small size prototypes of these structures have recently been manufactured. Modified DEPFET structures with build-in non-linear amplification are also used to implement active pixel detectors optimized for high dynamic range. Successfully prototyped for the DSSC sensors (DEPFET Sensor with Signal Compression) at the XFEL, these structures are increasingly being used in applications requiring high contrast and intensity, e.g., TEM imaging. Charge handling capability and output characteristics can be tailored to the requirements, as well as pixel geometry and size. The large intrinsic gain of the DEPFET provides excellent SNR even at fast timing. Pixels can be read with a speed of 100 ns, the resulting frame rate depends on the degree of readout parallelization.

KEYWORDS

photon science, silicon sensors, APD, DEPFET, pnCCD, pixel detectors, X-rays

1 Introduction

Many applications in photon science involve x-ray scattering experiments using high intensity x-ray pulses. The images thus generated are characterized by a distribution of high intensity, small size spots or speckles, surrounded by “halo” regions with an intensity rapidly declining with the distance from the intensity peaks. In this context, intensity is usually measured as number of photons with a fixed energy. Relevant information about the scattering process is not only contained in the distribution and size of the peaks, but also in the details and shape of the intensity distribution within the halos down to their fringe. Measurements at lower intensity, however, are often limited by the suppression of low signal amplitudes, driven by the necessity of noise discrimination. An ideal detector system therefore needs to acquire the complete scattering image, from the highest peak intensity to the faintest halo regions, ideally down to single photons, and still maintain robust noise suppression efficiency.

Key performance parameters of detector systems for photon science instrumentation are the charge handling capacity (CHC) of the detector resolution elements and the dynamic range of the system. In this context, CHC is the maximum amount of signal charge the system is able to handle per resolution element (pixel) without loss or deterioration of information. High CHC is required to suppress blooming or spillover effects compromising the imaging performance up to the requested maximum signal charge, so the scattering patterns can be imaged correctly.

As imaging of both low and high intensity regions of the image are equally important, an optimization of the dynamic range needs to target both, the high and the low dynamic range limit. Concerning the high dynamic range limit, the maximum preamplifier output needs to fit the maximum required signal charge. When optimizing the low end of the dynamic range, the gain characteristics of the detector/preamplifier structure have to be modified in order to push the low dynamic range limit to the minimum value achievable for the given noise conditions. To optimize both, the maximum dynamic range and its low range detection limit, a system with multilinear or non-linear behavior of the detector-preamplifier combo is superior compared to conventional single gain approaches.

Another requirement arising from the high rate capability of contemporary X-ray sources is the requested high readout speed. The frame rates required for megapixel sized detector arrays range from several kHz up to the MHz range. This brings about fast detector readout, high bandwidth of the amplifiers, and fast processing of the acquired data in the data acquisition backend.

Any optimization of a detector system for photon science therefore needs to target the detector structure itself as well as the respective readout electronics. Focus hereby lies on the optimization of the dynamic range, the signal-to-noise ratio and the increase of readout speed.

In the following, an overview about the work on this field carried out at the semiconductor laboratory of the Max-Planck-Society (MPG HLL) is given.

2 Gain characteristics of photon science detectors

Photon science experiments, e.g., in X-ray scattering, use photons of fixed energy to illuminate a scattering target. In the following, the interrelation between signal, gain, noise, accuracy and dynamic range of such a detector is outlined by means of a model for an ideal detector. For the sake of simplicity, the model neglects the effects of charge sharing between pixels, although in reality they greatly influence the sensitivity for lower amplitudes.

2.1 Signal and detector model

Using the simplification mentioned above, the stimulus S recorded by one image pixel consists of the individual, statistically independent signal depositions of N photons:

$$S = \sum_{i=1}^N S_i$$

The experimental requirement states that each pixel is able to record the signal from a maximum number of N_{Max} photons. For a monoenergetic source, it can be stated that

$$S = \sum_{i=1}^N S_i = N \cdot S_0$$

where S_0 is the average energy created by one photon. As the individual photons are statistically independent, the uncertainty of this energy deposition is $\Delta S = \sqrt{N} \cdot \Delta S_0$, where ΔS_0 is given by the energy dependent Fano noise.

A system matched to the requirements has a maximum pixel stimulus of $S_{Max} = N_{Max} \cdot S_0$. Sometimes, the values of S_{Max} or N_{Max} are used synonymously to the term “dynamic range”. In the context of photon science, however, also $S_{Min} = N_{Min} \cdot S_0$, is relevant for the application, which is the minimum stimulus distinguishable from noise. In this context, the quotient

$$R_{Dym} = \frac{S_{Max}}{S_{Min}} = \frac{N_{Max}}{N_{Min}}$$

is a useful definition for the dynamic range as relevant figure of merit. The value of N_{max} drives the instrument design in terms of e.g., CHC and defines the high limit of the dynamic range, and N_{min} is driven by the noise of the system and, indirectly, also by the value of N_{max} .

To understand the connection, an idealized model detector system is considered, which converts the stimulus N to a signal amplitude A from an amplitude range $A \in [0, A_{Max}]$ by using a gain function $A(N)$. From this, the uncertainty of the output amplitude $\Delta A_{tot}(N)$ of the detector can be calculated, taking into account the uncertainty of the stimulus ΔS , the statistical fluctuation of the number of quanta as defined by the Poisson statistics $\Delta N = \sqrt{N}$, and the RMS noise of the output amplitude of the detector, ΔA , which, for this model, is assumed to be constant. With $\Delta A_{tot}(N)$, the value for N_{min} can be determined using $\Delta A_{tot}(0)$ and an application-specific noise discrimination factor n by requiring that $A(N_{Min}) \geq n \cdot \Delta A_{tot}(0)$, where values for n are typically between 3 and 5.

TABLE 1 Comparison of model output for relevant performance parameters of the linear and nonlinear gain function. All parameters can be expressed as algebraic expressions.

Linear case	Nonlinear case
Gain function $A(N) =$	
$g \cdot N \cdot S_0 + A_0$	$\sqrt{q} \cdot N \cdot S_0 + A_0$
Gain coefficient	
$g = \frac{A_{Max} - \bar{A}_0}{N_{Max} \cdot S_0}$	$q = \frac{(A_{Max} - \bar{A}_0)^2}{N_{Max} \cdot S_0}$
Calibration function $\tilde{N}(A) =$	
$\frac{1}{g \cdot S_0} \cdot (A(N) - \bar{A}_0)$	$\frac{1}{q \cdot S_0} \cdot (A(N) - \bar{A}_0)^2$
Accuracy function $\Delta\tilde{N}(N) =$	
$N \cdot \sqrt{\frac{1}{N} \cdot (1 + (\frac{\Delta S_0}{S_0})^2) + (\frac{N_{Max}}{N})^2 \cdot (\frac{1}{A_{DNR}})^2}$	$N \cdot \sqrt{\frac{1}{N} \cdot (1 + (\frac{\Delta S_0}{S_0})^2) + \frac{4 \cdot N_{Max}}{N} \cdot (\frac{1}{A_{DNR}})^2}$
Normalized signal to noise ratio $SNR_N(N) =$	
$((\frac{\Delta S_0}{S_0})^2 + 1 + \frac{N_{Max}^2}{N} \cdot (\frac{1}{A_{DNR}})^2)^{-1/2}$	$((\frac{\Delta S_0}{S_0})^2 + 1 + 4 \cdot N_{Max} \cdot (\frac{1}{A_{DNR}})^2)^{-1/2}$
Low energy limit $N_{Min} =$	
$N_{Max} \cdot \frac{n}{A_{DNR}}$	$N_{Max} \cdot (\frac{n}{A_{DNR}})^2$
Dynamic range $R_{DYN} =$	
$\frac{A_{DNR}}{n}$	$(\frac{A_{DNR}}{n})^2$
Single quantum sensitivity limit $SNR_{Max} >$	
$n \cdot N_{Max}$	$n \cdot \sqrt{N_{Max}}$

In the application, a calibration function is used to retrieve a reconstructed value \tilde{N} from the measured value $A(N)$ using certain calibration coefficients, and with $\Delta A_{tot}(N)$ one can calculate $\Delta\tilde{N}$, an estimate for the uncertainty of \tilde{N} . In the following, the function $\Delta\tilde{N}(N)$, giving the uncertainty on the measurement for each value of N , is referred to as accuracy function. $\Delta\tilde{N}(N)$ describes, in the approximation outlined above, the uncertainty on the retrieved number $\tilde{N}(N)$ generated by the detector from a stimulus of N photons. The accuracy function of an ideal, noiseless detector would reproduce the Poisson limit $\Delta\tilde{N}(N) = \sqrt{N}$ assuming completely deterministic energy deposition.

To define the actual gain function, the most straightforward approach is that of a linear gain function of the form

$$A(N) = g \cdot N \cdot S_0 + A_0, \text{ with}$$

$$g = \frac{(A_{Max} - \bar{A}_0)}{N_{Max} \cdot S_0}$$

being the optimal gain matched to the required dynamic range. In this approach, A_0 represents the offset, which is the actual zero signal output of the amplifier at the time of measurement, and which is variable with an RMS of ΔA . \bar{A}_0 is the mean value of the offset distribution. The corresponding calibration function is

$$\tilde{N}(A) = \frac{1}{g} \cdot (A(N) - \bar{A}_0)$$

with \tilde{g} being the measured gain coefficient. As expected, a nonzero offset limits the dynamic range as it reduces the available signal output swing.

The alternative approach investigated here is the square-root shaped gain function of the form.

$$A(N) = \sqrt{q} \cdot N \cdot S_0 + A_0 \text{ with a quadratic gain coefficient}$$

$$q = \frac{(A_{Max} - \bar{A}_0)^2}{N_{Max} \cdot S_0}$$

which, in the following, is referred to as nonlinear gain function. This approach is chosen as an educated guess, because higher gain for lower amplitudes lifts the output signal above the noise threshold for yet smaller input stimuli. Accordingly, this yields a second degree polynomial without linear term for the calibration function:

$$\tilde{N}(A) = \frac{1}{q} \cdot (A(N) - \bar{A}_0)^2$$

with \tilde{q} being the gain coefficient retrieved by calibration.

Using these gain functions, expressions for the accuracy functions, low energy limit, dynamic range and single photon sensitivity limit were derived. Here, an ideal calibration (i.e., $\tilde{q} = q$ and $\tilde{g} = g$) was assumed for evaluation, although this approach can also be used to determine the effect of calibration errors.

2.2 Comparison

For the gain functions and conditions sketched above, the model gives a set of relations describing the relevant performance parameters of the model detector. Table 1 compares some of the findings for the performance characteristics for the nonlinear and linear gain function. For a more compact notation, the definitions in the table use the term

$$A_{DNR} = \frac{A_{Max} - \bar{A}_0}{\Delta A}$$

which corresponds to the ratio of the output dynamic range to the noise of the preamplifier.

In comparison, it is obvious that, for common values of A_{DNR} and n , the accuracy function for the nonlinear approach is superior for small signals. For equal A_{DNR} and n , the low signal limit for the nonlinear gain scales with the quotient n/A_{DNR} compared to the linear approach, and the dynamic range with the inverse ratio A_{DNR}/n .

For $A_{DNR} = 40$, an upper end of the dynamic range of $N_{Max} = 100$ and a noise discrimination threshold of $n = 5$, for instance, a value of $N_{Min} = 12.5$ is achieved for the linear case, while the nonlinear case yields $N_{Min} = 1.6$, and the dynamic range is 8 for the linear case compared to 64 for the nonlinear case. The single quantum sensitivity limit, giving a requirement on the minimum A_{DNR} to maintain $\Delta\tilde{N}(N) < 1$, scales with the square root of N_{Max} . For the figures given above, the A_{DNR} for the linear case required for single quantum sensitivity is 500, compared to only 50 for the nonlinear case.

The better low signal performance is countered by an accuracy worse compared to the linear gain function at higher stimuli.

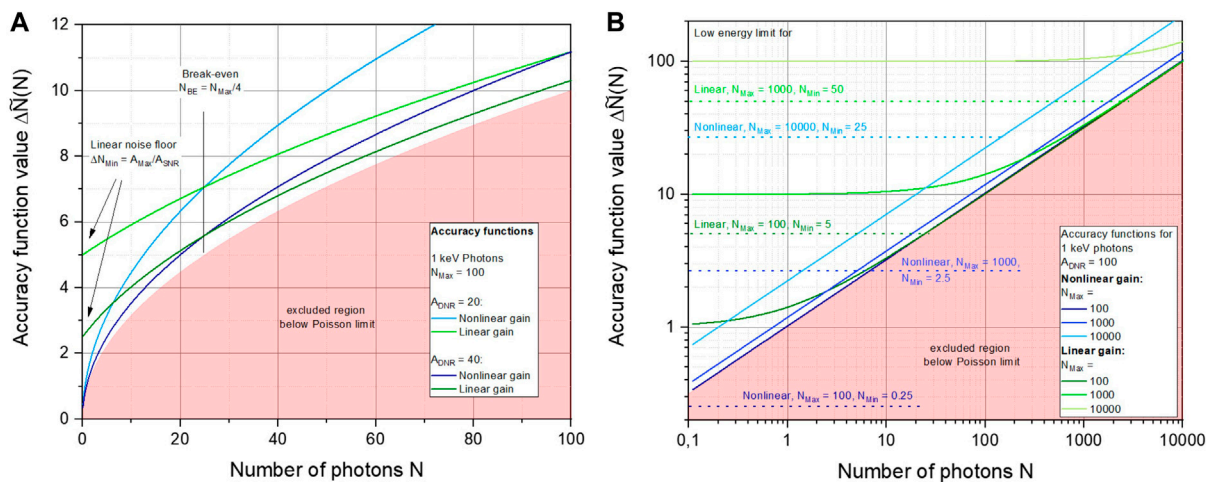


FIGURE 1

Accuracy functions for linear and nonlinear gain functions, for A_{DNR} values of 20 and 40 (A). The stimulus consists of up to $N_{Max} = 100$ photons with an energy of 1 keV. For the energy deposition of the photons, the Fano noise is taken into account. Below the break-even stimulus of $N_{BE} = N_{Max}/4$, the nonlinear gain produces much better results, as the linear gain function produces a constant noise floor. (B) shows the scaling behavior of the accuracy functions with N_{Max} , using 1 keV photons as stimulus, for an $A_{DNR} = 100$. In addition, N_{Min} is indicated by the dashed lines. For $N_{Max} = 100$, the single quantum sensitivity criterion is fulfilled.

Figure 1A shows the comparison of the accuracy functions for a maximum number $N_{Max} = 100$ photons with an energy of 1 keV, for an A_{DNR} of 20 and 40, with linear and nonlinear gain. At low stimuli, the term depending on A_{DNR} causes a constant noise floor value of $\Delta\tilde{N}_{Min} = N_{Max}/A_{DNR}$ at $N = 0$ for the linear gain function, while the corresponding term for the nonlinear gain converges to zero. Above a “break-even” stimulus of $N_{BE} = N_{Max}/4$, however, the linear gain function becomes more accurate and converges towards the Poisson limit.

The behavior scales accordingly for higher values of N_{Max} , as shown in the right panel of Figure 1. In this log/log plot, one can see that, while the accuracy for the linear gain converges to the Poisson limit, the discrepancy of the accuracy for the nonlinear gain function is always higher than the ideal Poisson limit accuracy by the same factor. This behavior can also be found in the expressions for the normalized SNR_N (see Table 1), being the actual SNR divided by the theoretically optimal SNR of the Poisson limit:

$$SNR_N = \frac{1}{\sqrt{N}} \cdot SNR$$

Here, it turns out that the SNR_N for the nonlinear gain is constant wrt. N , while the SNR_N for the linear gain shows a growth towards its optimum value at N_{Max} , showing an approximate \sqrt{N} characteristics.

2.3 Discussion

Using the model system as an example, it has been motivated that using a detector with a nonlinear gain function as described above has advantages over a linear approach when used within the context of photon counting applications. For low stimuli, the uncertainty on the retrieved photon number converges to zero for the nonlinear gain while the linear gain reaches a finite noise

floor level, which is due to the different scaling of the amplifier noise contributions. Main advantage of the nonlinear gain is a substantially increased dynamic range and improved low signal detection limit, as can be taken from the left plot in Figure 1.

The advantage of the linear gain approach, in contrast, is the higher accuracy for the largest part of the input stimulus range. In addition, it should not be ignored, that a detector system implementing a nonlinear gain also faces some practical challenges. The idealized square root behavior is in practice difficult to implement. Practical implementations therefore use approximations of the ideal behavior. In addition, the optimized behavior is valid only for a single target energy and intensity. When these values change, deviations from the ideal behavior are to be expected. Finally, the calibration of such a detector can be a cumbersome task, given that imaging detectors for photon science target a large sensitive region with many channels, and needs to cover the complete dynamic range.

But if a large dynamic range is requested and both low and high intensity features are to be acquired simultaneously, and the higher inaccuracy for larger signals can be tolerated, a nonlinear gain is the better solution. It is an extremely useful tool for the exploration of large dynamic ranges.

3 Detectors for photon science applications

The ideal gain function for applications limited by the counting error follows a square-root behavior. To implement or mimic this behavior, several approaches are commonly applied in state-of-the-art photon science instrumentation.

At the time of writing, standard pad or pixel detectors based on PIN diodes for conversion of the X-ray photons to signal charge are widely used. This type of detector is technologically simple and easy

to manufacture with high yield. When applied in large scale systems, integration is done using a hybrid pixel approach, in which front-end ICs providing one preamplifier channel per pixel are bump-bonded to the sensors, connecting preamplifier channel to its sensor pixel with solder bumps.

In this detector architecture, the task of applying the nonlinear gain is transferred to the front-end electronics, especially the preamplifier. Different approaches exist for implementing or approximating the requested shape of the gain function. The multilinear approach approximates the optimum nonlinear gain curve by a combination of multiple linear gains. One possibility to implement multilinearity is to having several (typically three) preamplifier channels per pixel operating in parallel and each having a different gain—one for low, one for medium and one for high intensities [1]. Depending on the actual stimulus, the information of the channel with the gain fitting the stimulus requirement in the best way is selected for data analysis.

A similar approximation can also be made using one single preamplifier per channel only, if the preamplifier can adapt automatically to the signal. Integrated, pixel-individual control electronics adjusts the preamplifier gain dynamically as a function of the signal level [2] by selecting one of three different gain values. The internal control logic monitors the output signal and switches the gain to the next higher value as soon as the amplitude exceed a certain threshold. In these and similar approaches, the design of the preamplifier determines the gain(s), and the shape of the effective gain function. The lowest gain is selected to meet the requirement on the upper limit of the dynamic range, while the highest gain defines the lower dynamic range limit.

The problem with Pad detectors is that their pixel capacitance is directly coupled to the pixel area, which, in combination with the high preamplifier bandwidth required for fast readout, results in relatively high noise figures, effectively limiting their low energy performance. This problem can be mitigated if the detector itself provides for a certain level of amplification, like for instance by using avalanche multiplication. Classical APDs however, suffer from homogeneity issues. The Monolithic Array of Reach-THrough APD (MARTHA) structure, however, uses avalanche multiplication for pushing the low energy dynamic range limit while avoiding the weak points of the classical approach.

3.1 MARTHA devices

MARTHA devices are based on the concept of Avalanche PhotoDiodes (APDs). APDs provide inherent localized amplification of the signal charge within the sensor, and the amplification is achieved by avalanche multiplication of the primary signal charge itself.

At the time of writing, APDs are applied in a large variety of contexts, but many more potential applications could benefit from their use. The main issue is here, that APD arrays show bad position resolution and homogeneity of sensitivity over larger areas, especially if benchmarked against what is possible for simple photo diode arrays. Consequently, APDs play a comparably minor role. Only small formats of 16 up to 64 pixels are commercially available, and the smallest dead space gaps between

pixels are in the range of 40 μm [3]. Institutional R&D is ongoing for several years to close the gap to the photo diode arrays.

For the fabrication of larger APD pixel or strip arrays one has to overcome two critical obstacles. The first problem concerns the homogeneity of the avalanche gain. The ionization rates in the avalanche process depend on the electric field strength, which by itself is determined by the actual doping profile in the implants defining the potential in the multiplication region. Consequently, any fluctuation in the doping process directly affects the magnitude of the electric field. Fluctuations in implant energies results in depth variations of the implants, affecting size and position of the multiplication region, and fluctuations in the implanted dose will directly modify the overall space charge. Both effects have an exponential impact on the ionization rate, and thus the gain. These gain fluctuations need to be minimized by careful technological processing using highly specialized, well controlled and monitored processes for all relevant implantations. The effect of the yet remaining gain variations can be mitigated by operating the APD arrays at a comparably low gain. In general, gain values of 20 and below are recommended to keep the effect of gain fluctuations in a manageable range. This is a common approach, for which the R&D activities are summarized under the name Low Gain Avalanche Diodes (LGAD) development—see, for example, references [4–6].

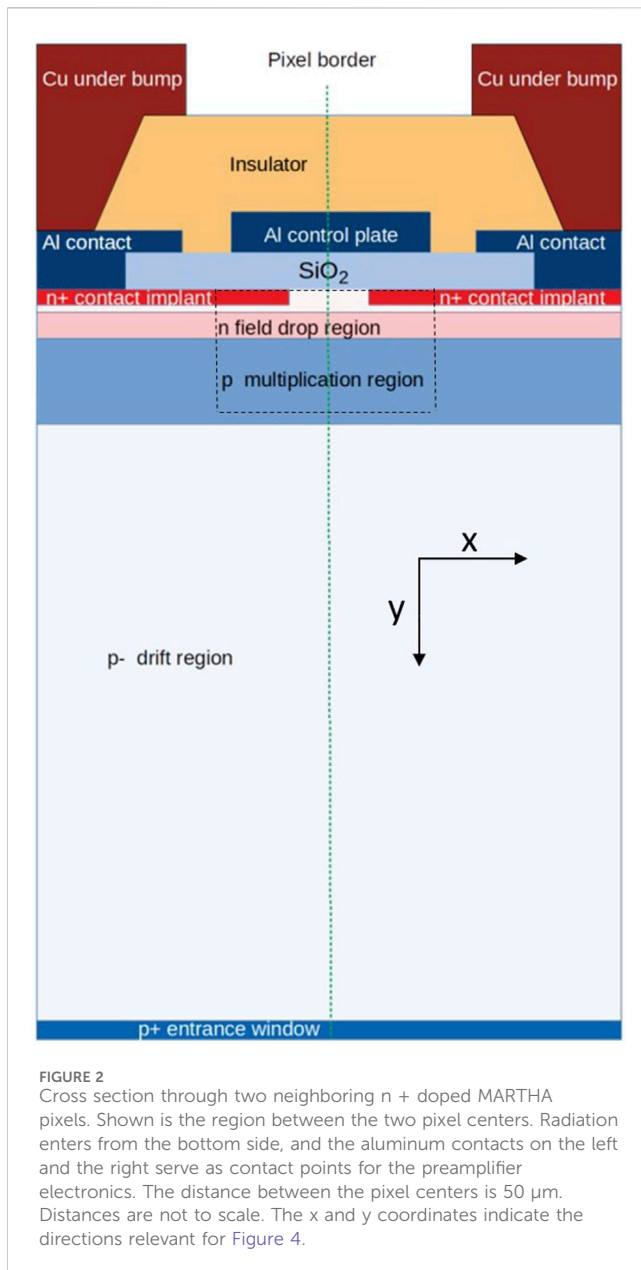
The second problem is the early breakdown at the edges of the segmented n + diodes defining the pixel electrodes. This is caused by the elevated field strength at convexly curved doping shapes as can be found at the edges of the anode implants. To avoid the breakdown, implanted guard ring structures are commonly provided, which rigorously suppress avalanche processes at the edges of the implants. The downside is that these arrays are blind in the inter-pixel gap regions.

Concepts to suppress edge breakdown while maintaining high detection efficiency in the gap region are currently in the focus of research and development [7–10]. The Inverse LGAD (iLGAD) [7, 8] is the most fundamental approach to achieve a homogeneous lateral electric field distribution. It uses a segmented p + cathode on a p type substrate as pixel/strip read out structure. One global non-structured n + anode and p type multiplication region on the opposite side of the substrate covers the entire sensor array. The pn-junction formed by the anode and the multiplication region, however, must not touch the cutting edge, requiring more expensive double sided processing, which is even more challenging in case thinned substrates are used.

If operated as X-ray detector, photons have to enter on the non-structured n-side. If a photon is absorbed in the multiplication region, the photon generated charge experiences different amplification, depending on the depth at which they are absorbed. Although there are successful attempts to mitigate this effect [8] it still compromises the detection efficiency especially for low energy X-rays.

Another approach is to replace the implanted standard inter-pixel guard structure by a trench isolation [9]. The blind inter-pixel region can be drastically but not completely reduced to a level sufficient for particle detection. For precise energy measurements or photon counting, a homogeneous response is required.

In the Deep LGAD concept [10], the p-avalanche multiplication structure is implanted before the growth of a thick, high resistive



n-type epitaxial layer, which separates avalanche multiplication from n + pixel implants located on top of the epitaxial layer. Reported simulation results show a very good edge breakdown suppression and a fairly homogeneous response behavior. Problematic with this approach is that the pn-junction between epitaxial layer and p-type bulk conceptually extends to the cutting edge of the sensor, leading to significant leakage current generation.

The MARTHA concept discussed here offers an alternative way to completely suppress edge breakdown while providing a fairly uniform signal response also in the inter-pixel gaps.

3.1.1 MARTHA concept

Figure 2 shows a cross section through two neighboring pixels of the proposed MARTHA structure. For any photon detector, optimum quantum efficiency can only be achieved without any

obstacles for the incident photons. For this purpose, a Reach Through APD [11] is an ideal solution. It operates at full depletion and provides a homogeneous photon entrance window with 100% fill factor on the backside. The bulk acts as drift region. If the bulk is p-doped, the structures can be manufactured using single sided processing only, with guard structures only on the n-side.

The pixel electrodes are formed by conventional n + anode implants. An aluminum control grid over the complete array provides flexibility in controlling the SiO₂/Si interface potential in the inter pixel gap regions. This additional electrode can later be replaced by a proper surface implantation in a future implementation. Key features of the MARTHA structure are the global multiplication region (MR) and the field drop region (FDR), which is located above the MR. The MR is defined by a p-type high energy (HE) implant in a depth of several microns from the surface, the FDR by a relatively deep-n implant below the anode implants. Both implants are global and unstructured. In this way, MR and pixel implants are vertically separated in a similar way as in the Deep LGAD concept mentioned above, but the MARTHA approach has the advantage that the required doping profiles can be confined to the pixel region.

If the structure is fully depleted, the MR extends between the MR HE implant and the FDR. The FDR is important for the mitigation of the edge breakdown. Its main effect is to suppress the aforementioned peaks in the field strength. Its effect can be deduced from equivalence between the depleted implant dose and the field strength, which can be concluded from the Poisson equation. Accordingly, the absolute maximum electric field at a pn-junction is

$$|E_{max}| = -q\epsilon D_D = q\epsilon D_A$$

with D_D and D_A being the corresponding depleted doping doses. In the APD array, D_A is composed of the dose of p-implantation creating the high electric field within the MR, the dose equivalent of the bulk concentration and the depleted fraction of the cathode implant at the entrance window. D_D , in contrast, is the sum of the implantation dose for the FDR, D_{FDR} , and the depleted fraction of the anode implant for the pixel contacts, with the latter contribution being omitted in the gap region. Assuming separated vertical concentration profiles for the n + anodes and the FDR, the electric field at the height of their transition point can be calculated to be the maximum field strength, scaled with the ratio of the depleted dose for the field drop layer to the overall depleted acceptor dose:

$$|E_{tp}| = |E_{max}| \cdot \frac{D_{FDR}}{D_A}$$

where D_A is dominated by the implant in the MR. In effect, the field drop layer does not avoid field peaks, but scales their amplitude down to uncritical values, i.e., below the field strength threshold for impact ionization. The ratio between FDR and MR implants were optimized using 2D TCAD simulations [12]. The results show that a FDR/MR dose ratio of about 2/3 is a good choice to keep the remaining electric field peak at the n + pixel edge below the critical value. Figure 3 shows TCAD 2D simulation results for the electric field distribution between two adjacent pixels as shown in Figure 2, with and without FDR. As the MR and FDR implants are

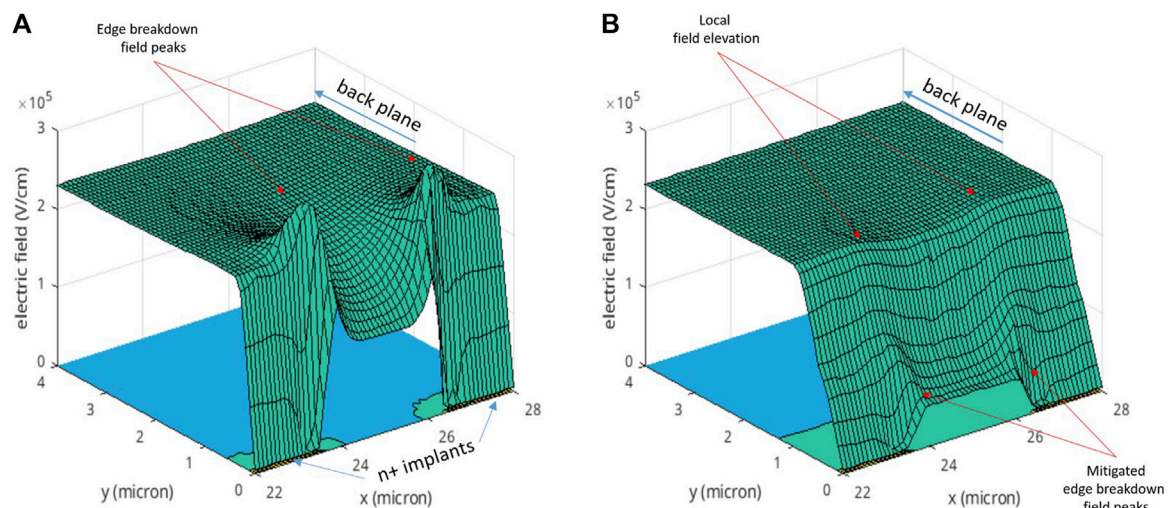


FIGURE 3

Zoom into simulation of inter pixel electric field distributions (absolute values), without (A) and with (B) field drop layer (FDR). The x and y coordinates correspond to the lateral surface coordinate and the depth in the wafer as indicated in Figure 2. The inter-pixel gap is 3 μm wide. Without FDR, the electric field peaks are prominent. A structure with FDR, in contrast, shows substantially smaller peaks, which are still visible, but suppressed to a large extent and well below the critical value for avalanche generation.

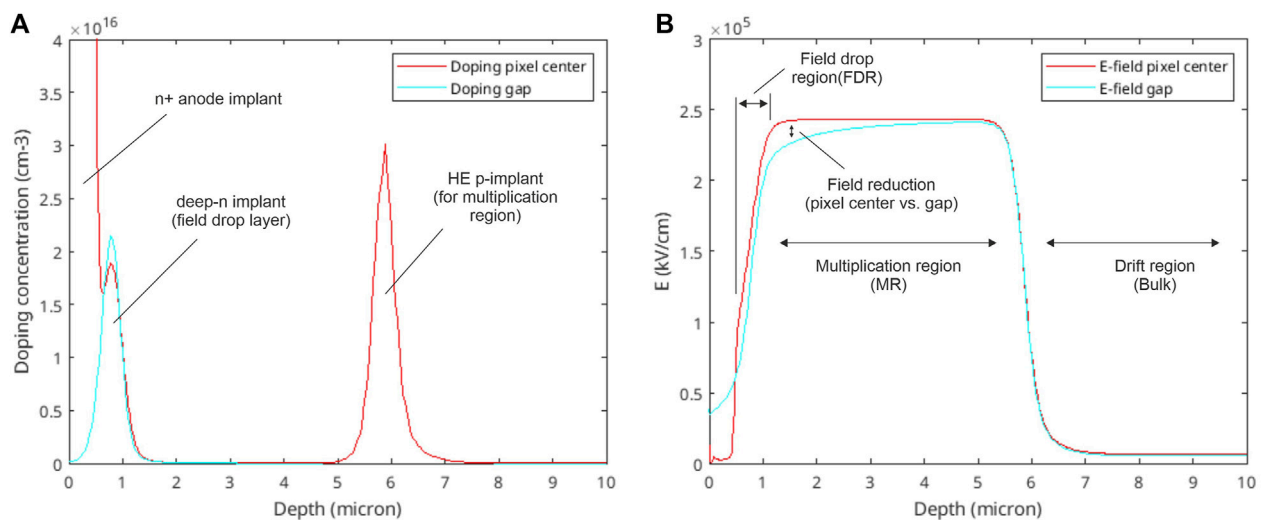


FIGURE 4

Vertical doping profiles (A) and resulting absolute electric fields (B) in the various vertical zones. Two different cuts, one through the pixel center and one through the inter-pixel gap are shown. Clearly visible are the near-homogeneous field in the multiplication region especially in the pixel center. The field reduction in the inter-pixel gap is caused by the more negative surface potential, but its influence is mitigated by compensating effects.

unstructured within the APD array, a homogeneous field distribution in horizontal direction is created, with slight inhomogeneity at the transition between gap and pixel regions. The statistical nature of the avalanche process effects an additional noise contribution for the APDs. The Excess Noise Factor (ENF) quantifies this contribution with respect to the shot noise of a classical photo diode:

$$ENF = k \cdot M + \left(2 - \frac{1}{M}\right) \cdot (1 - k)$$

where M is the gain and k is the ratio between the hole and electron ionization rates, which, in silicon, decreases with decreasing electric field [13]. The lower limit for the ENF of close to 2 can be obtained by a very small k value, at the cost of a low gain value M , as operation at low electric field also reduces M .

This loss can, however, be compensated if the depth of the multiplication region is increased, to that the gross charge amplification remains the same. For this purpose, the aforementioned HE implant is used to create a wide, vertically extended MR below the newly introduced deep n-doped FDR.

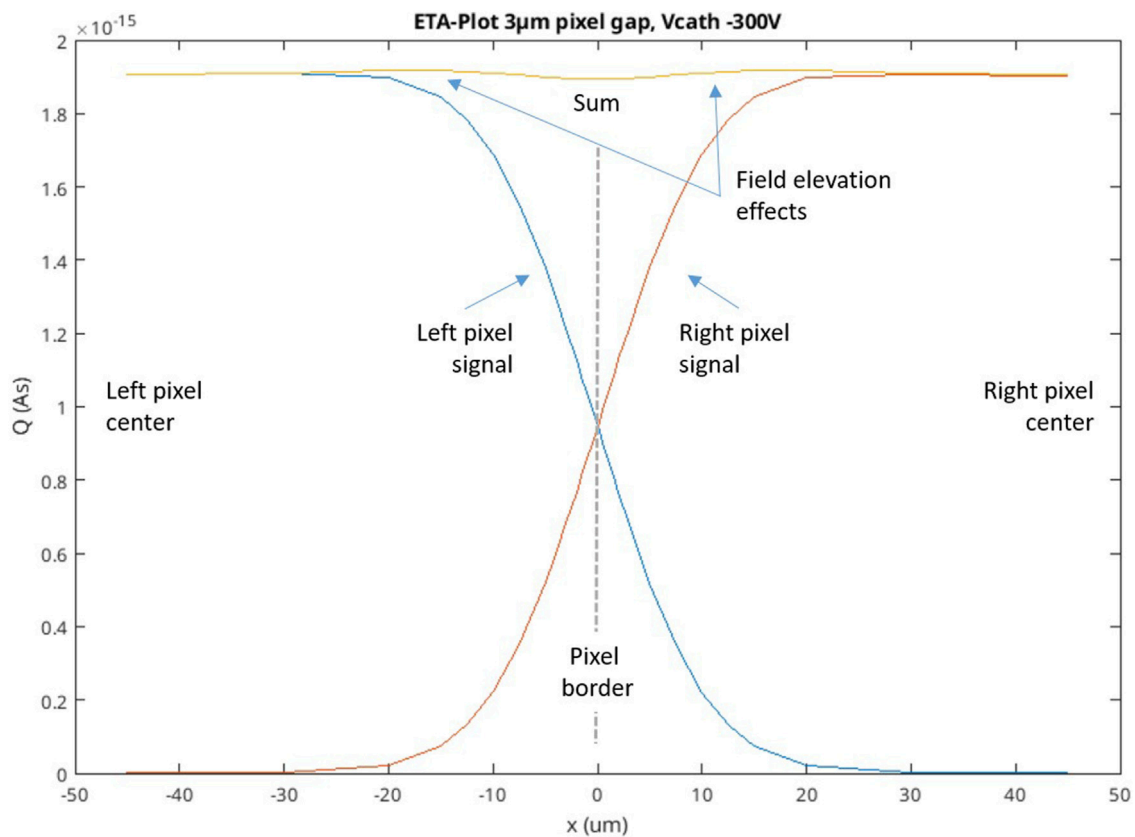


FIGURE 5

Simulation of signal charge response, to test charges generated at different horizontal positions x in a depth of $448\text{ }\mu\text{m}$. The signal level detected by the left anode is shown in blue, and the right anode signal is shown red. The sum of the signals is indicated by the yellow line. The simulation indicates that the structure does not have insensitive regions in the inter-pixel region. Overall simulated signal injection was about 800 electrons, and the structure operates at a gain of about 14.

3.1.2 Simulation results

The simulation refers to two neighboring half strips including the gap in between. Each half strip extends over $50\text{ }\mu\text{m}$. Taking into account the applied symmetric left/right Neumann boundary conditions the simulation domain corresponds to a strip detector array with $100\text{ }\mu\text{m}$ pitch. For this simulation, the gap between the pixels was $3\text{ }\mu\text{m}$ wide. When investigating the region below the pixel implants, the MR, between the MR implant and the FDR, shows a parallel-plate capacitor like electric field distribution. Reason for this is the very high doping concentration between FDR and the high-energy MR implant, where almost no space charge is integrated.

Figure 4 shows vertical cross sections of doping profiles through pixel and gap regions (left image) and the resulting electric fields (right image) of the upper part of a $450\text{ }\mu\text{m}$ thick sensor, determining the various regions of a MARTHA structure. The doping profile of the HE implant used in the TCAD simulations was measured by Secondary Ion Mass Spectroscopy (SIMS) on structures from a wafer from the prototyping production.

In the inter-pixel region, however, a reduction of the electric field in the MR can be observed, in the order of around some 10%, which increases with decreasing depth. This field reduction is a result of the potential at the SiO_2/Si interface, which must be more negative than the anode potentials in order to provide sufficient inter-strip/pixel isolation. But this electric field does not result in a

reduced avalanche generation in the gap region. Figure 5 shows the result of a charge generation scan simulation over the horizontal coordinate, which was done to evaluate the sensor response in the gap region. The simulation demonstrates that almost no charge loss in the gap region occurs. About 800 electrons were injected in a depth of $448\text{ }\mu\text{m}$ below the anode contact. In this simulation, the structure works with a gain of about 14 in the preset bias conditions.

Two mechanisms help to compensate the effect of the reduced electric field. At first, the negative gap potential creates an additional drift field, diverting charge generated below the gap region towards the n^+ doped anodes, where they are exposed to the nominal field strength and experience the specified amplification. In addition, the simulations suggest that a 2-D effect, already faintly visible in Figure 3B, also helps to compensate the effect of the field drop. A local field elevation appears on each side of the gap, which is laterally displaced from the n^+ pixel edge by a few μm towards the pixel center. Although the field elevation is of the order of about only few percent, its magnitude is sufficient to affect the amplification. Accordingly, the simulated response curves shown in Figure 5 indicate a region with slightly elevated response compared to the anode centers, at the order of up to 1%, in the regions left and right of the pixel borders, showing a maximum at about $\pm 15\text{ }\mu\text{m}$.

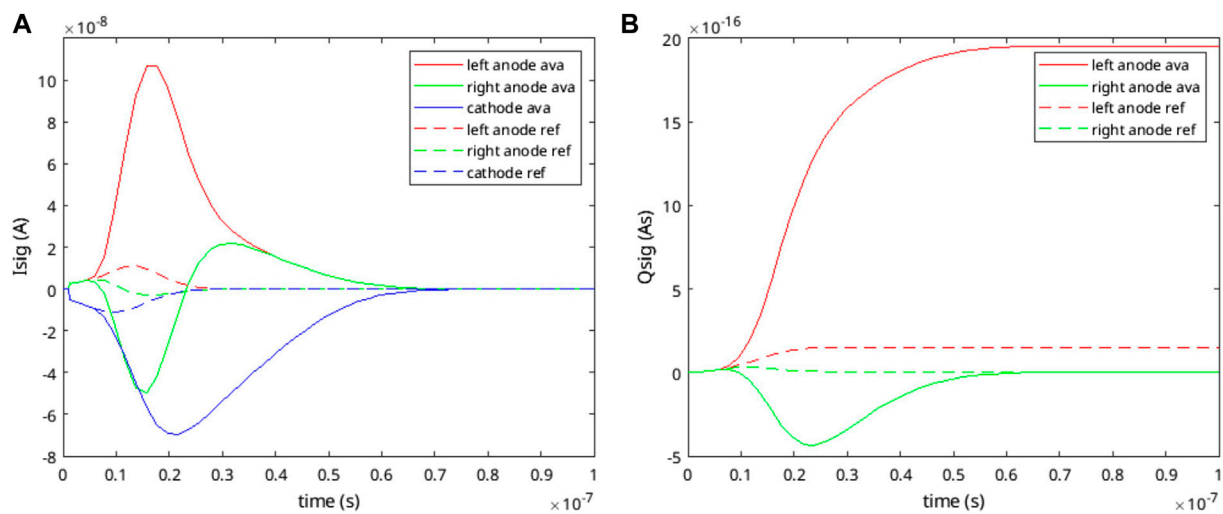


FIGURE 6

Evolution of signal currents (A) and signal charge (B) in a MARTHA structure after generation of a test charge of 0.134 fC at the backside of the sensor beneath the left anode. Generation depth is 448 μm . Solid lines represent the amplified currents and charges. For comparison, the same simulation was performed on a standard PIN diode structure without avalanche generation (dashed lines).

Similar simulations were made to study timing behavior and signal formation inside a MARTHA structure. A test charge of 0.134 fC was generated close to the backside in a depth of 448 μm beneath the left anode. Figure 6 shows the time dependence of signal currents and their integrals, the amplified signal charge. The plot compares the MARTHA pixel structure with standard pixel without avalanche multiplication. The simulation reveals that the detection of the full signal lasts longer in the MARTHA scenario, as the holes generated in the multiplication layer have to drift a significant distance towards the cathode until the full signal is established. This is a feature common for all types of APDs.

Accordingly, the induction current on the right anode is more pronounced compared to the situation at the standard diode. This is because in the standard diode the current is induced by drifting electrons according to the weighting field only, while for the MARTHA structure the right anode additionally ‘sees’ the induction current of generated holes drifting in the opposite direction. It gradually disappears the more the holes approach the cathode. Eventually the right anode net signal is zero in both cases. In summary, the signal rise time is comparably fast, but the simulations indicate that the detection of the full signal lasts longer compared to a diode.

Conceptually, the MARTHA structures are suitable for both X-ray and particle detection, but some optimization with respect to the application is required. Common to both applications is the need for high detection efficiency. MARTHA structures fulfill this requirement, as they are backside illuminated structures with high fill factor, and the FDR eliminates the need for insensitive guard ring structures. Applying suitable antireflective coating (ARC) on the backside entrance window to match the QE to the application requirements allows for customization of the MARTHA structures also for use with optical wavelengths.

Using MARTHA structures for the X-ray wavelength range requires low noise operation for good energy resolution and precise photon counting, especially when low photon energies are

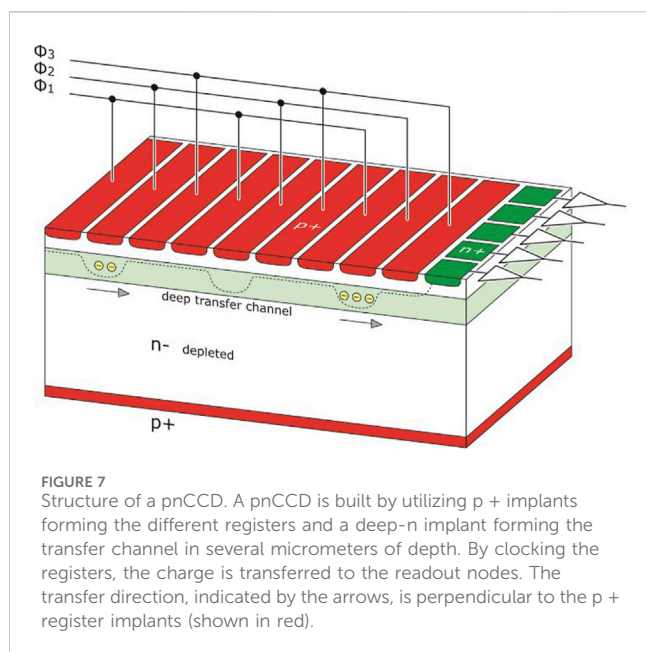
used. In addition, an adaption of the substrate thickness is required. To detect higher X-ray energies with acceptable QE, substrates with a thickness at the order of 450 μm or higher are required. These substrates are compatible with the HLL standard entrance window technology, which provides for excellent quantum efficiency up to an X-ray energy of 12 keV [14].

For applications in particle tracking, in contrast, thinned SOI substrates are a suitable option as an effective way to reduce the radiation length. Here, substrates with a thickness of 50 μm or less are feasible.

The devices in the ongoing prototyping production were optimized for X-ray detection. The structures are integrated on highly resistive p-type float-zone wafers with standard thickness of 450 μm . Reach through diodes need to be operated in full depletion. The high-energy implant defining the MR is applied to the area inside the pixel/strip array only, and the implant is lithographically masked at the sensor edges. This simplifies the design by avoiding interference with the guard ring structure on the n-side.

Silicon wafers with this bulk thickness are suitable for X-rays in a medium energy range up to 12 keV. Using a wide MR and a thick wafer leads to a relatively high operation voltage. For our chosen parameters, depletion of the high field region and the thick wafer needs about -140 V and -120 V, respectively. Adding a moderate over-depletion voltage results in an operation voltage of about -300 V. On the other hand, the electric field depends only weakly on the backside voltage. Therefore, if the sensor is operating with low gain only modest gain changes are to be expected, in case the bias voltage fluctuates or the bulk doping shows variations.

For soft x-rays below 1 keV, thin entrance window implants with minimum dead layer are required. These thin entrance windows suppress the recombination of signal charge in the undepleted part of the heavily p-doped cathode, so quantum efficiency and spectral performance improve. The respective technology steps are compatible to the MARTHA process.



To summarize, MARTHA structures have the potential to overcome the traditional limitations of APD arrays. In photon science applications, their gain properties help to push the low energy limit of the dynamic range. Concerning the high-energy limit of the dynamic range, gain-reducing effects observed at higher intensities [15] are subject of ongoing investigations. These could also help to accommodate the requirements from the application.

In case minor modifications concerning biasing are adopted, the devices can work together with most classical front-end solutions forming a hybrid pixel assembly. If combined with an advanced multilinear frontend, they can service the complete dynamic range as requested by state-of-the-art photon science experiments and provide for an attractive alternative to traditional pad and pixel detectors especially for detecting low energy X-rays.

Based on the abovementioned TCAD simulations, a generic MARTHA technology plan was conceived and optimized. A first prototyping production was finished recently at the MPG HLL; the prototype detector structures await qualification measurements.

3.2 pnCCDs for high dynamic range

Pad or Pixel detectors in a hybrid pixel detector are an appropriate solution if for instance extremely high framerates are required, as they provide for the option of full parallel readout of the complete detector. Disadvantages are high complexity of integration, high power consumption, comparably large volumes and the need for advanced mechanical and thermal interfaces, in combination with a low energy threshold, which is comparably high. In addition, the hybrid pixel approach demands similar pixel pitch on sensor and front-end electronics. This brings about limitations in terms of scalability of pixels sizes. In case resources are less abundant and the experiment requires a more compact sensor system, a lower system noise or smaller pixels are requested, and lower framerates can be tolerated, pnCCDs can be an attractive alternative.

3.2.1 pnCCD structure

The pnCCDs have originally been developed as sensors for imaging X-ray spectroscopy for the X-ray satellite mission XMM-Newton [16, 17] and since then found application in a wide variety of experiments in X-ray astrophysics as well as high luminosity X-ray sources [18–21].

Figure 7 illustrates the structure of a pnCCD. The pnCCD concept is based on sideways depletion, and the devices are manufactured in a double sided process. The backside provides a homogeneous, unobstructed, ultra-thin entrance window for radiation. Customization of the entrance window, e.g., by deposition of anti-reflective coatings or light blocking filters is possible if required by the application. On the frontside, a high-energy (HE) n implant several micrometers below the surface defines the potential minimum for electrons which forms the transfer channel. In transfer direction, p + implants connected to form a three-phase register structure define the CCD pixels. The pixel size can be scaled between 36 μm up to 150 μm to match the experimental requirements.

Perpendicular to the transfer direction, additional deep implants, the n-type channel guide and the p type channel stop, define the pixel structure. These implants create potential barriers between the pixel columns, confining the signal charge to the pixel. Utilizing a three phase clocking scheme, the charge is transferred from pixel to pixel, and finally handed over to the n + readout anode. The readout node is connected to the gate of the FirstFET, a low noise n-channel JFET with an equivalent gate capacitance of 50–100 fF. It is typically read out in a source follower configuration. With readout rates of 4 μs per line, an input capacitance this low provides for noise figures of a few electrons ENC.

A pixel of the pnCCD stores the signal charge below one of its registers inside a potential pocket in the n-type HE implant, which also serves as transfer channel. The register structure confines the signal charge transversally, and lateral confinement is given by the deep-n channel guide and deep-p channel stop implants. The depth of this potential well defines the pixels CHC. A state-of-the-art pnCCD, with a pixel size of 75 μm^2 , can store a few 100 ke-per pixel in standard operating conditions [22].

Although this is sufficient for most spectroscopic applications, the situation is different for ultra-luminous sources, as used for scattering experiments. Here, certain regions on the sensor are illuminated with high intensity, several thousand up to several hundred thousand photons and more can be collected within a single pixel. The amount of signal charge thus generated exceeds the charge-handling capacity of a pnCCD in standard operating conditions, leading to an overflow of the charge into neighboring pixels, an effect observable as blooming or smearing in the recorded images. Increasing the pixel CHC of a pnCCD is, however, possible by optimizing the operating conditions.

The standard operating conditions create the potential well defining the pixel in a depth of several micrometer below the surface, with the depth mainly being defined by the deep-n transfer channel implant. A more negative backside voltage, however, lowers the positive potential formed by this implant. Below a certain voltage level, the vertical potential barrier is overcome and the charge in the transfer channel is pushed closer to the surface, into a secondary potential well formed by the channel

guide implants. In effect, this increases the capacitance of the pixel and allows storing more charge within it. A pnCCD with $75 \mu\text{m}^2$ pixels operated in this high dynamic range mode exhibits a pixel CHC to up to 2×10^6 e⁻ and more, without exhibiting bleeding effects, which is sufficient for the use in photon science experiments [22].

3.2.2 pnCCD readout

Conceptually, pnCCDs are charge integrating devices, which are read out on demand. The image is generated by integrating signal charge for a defined integration time. Afterwards, the charge is read out by transferring it to the readout nodes.

A fundamental limitation of CCDs is that the charge needs to be transferred serially to the readout anode for every pixel, which puts up an intrinsic limit for the degree of readout parallelization. Unlike many conventional CCDs, pnCCDs provide one readout node per sensor column, and the front-end electronics processes the information of all sensor columns in parallel, so that no lateral transfer is required. This is referred to as full column-parallel readout. By suitable segmentation and matching interconnection of the register contacts, a split-frame readout can be realized. For such a device, readout anodes and first FETs are present on both edges of the CCD, to which the signal charge from the adjacent half of the sensor pixels is transferred. In this way, a two-fold column-parallel readout can be implemented, effectively reducing the readout time by a factor of two. A higher degree of parallelization cannot be realized on a monolithic pnCCD device. Further topological modifications are possible for pnCCDs, which, however, do not influence the readout speed. A frame store area [23], for instance, is primarily introduced to suppress so-called out-of-time events (OOTs), which are events registered by the CCD during the readout phase of the image. Due to the frame store area, OOTs can be suppressed to a large extent at the cost of substantially larger device area and additional transfer time, which slows down the readout. Therefore, a frame store is mainly useful for improving the image quality under continuous beam conditions with low to medium beam intensity.

Once the charge is transferred to the readout anode, it causes a voltage change at the gate of the FirstFET. In the classical JFET configuration, the FirstFET of a pnCCD can only be operated in source follower configuration for topological reasons. Here, constant current bias is provided to each FirstFET, usually by a current source integrated in the associated front-end channel. In first order approximation, the additional charge on the FirstFET gate increases the gate voltage by about:

$$\partial V_G \approx \frac{\partial q_{sig}}{C_G}, \text{ and the source follows this voltage step: } \partial V_S \approx \frac{\partial q_{sig}}{C_G}.$$

This source voltage step is acquired by comparing the source voltage level of the FirstFET. The baseline level before the transfer, and the signal level after the transfer are acquired and subtracted from each other. This technique is commonly referred to as correlated double sampling (CDS), and requires a time-variant preamplifier/shaper circuit. CDS provides superior 1/f noise suppression.

Several readout ASICs are available that are capable to process the signal of the FirstFETs via CDS, using either multi-correlated double sampling or a trapezoidal weighting function for filtering [24, 25]. Readout rates down to 2 μs per column can be realized with

excellent noise properties and fano-limited energy resolution for X-ray spectroscopy.

This excellent performance is also a consequence of the nature of the pnCCD being a sideways depleted device. The conceptually small capacitance of the FirstFET Gate/readout anode increase the signal voltage step, directly affecting the signal to noise ratio.

The FirstFET gate/readout anode contacts can be reset to a defined voltage level using the so-called ResetFET. It can be biased statically, causing slow discharge from the readout anode to the drain of the reset FET, keeping the readout anode voltage in a dynamic equilibrium depending on the reset FET bias and the amount of incoming charge. This mode works best for low to medium occupancy, e.g., in spectroscopic applications. Alternatively, the reset FET can be operated in pulsed mode, resetting the readout anode contacts after a fixed number of readout cycles. This mode is useful to restore defined biasing conditions in operation modes with large quantities of signal charge. In the most extreme cases, the readout anodes have to be reset after each readout cycle.

Considering a CCD with n_r rows, framerate ν_f for these systems is given by

$\frac{1}{\nu_f} = n_r \cdot (t_{rd} + t_{rst})$ for a standard mode CCD and $\frac{1}{\nu_f} = n_r \cdot (t_{rd} + t_t + t_{rst})$ for a framestore device, with t_{rd} being the readout time for one row, t_{rst} the time required for resetting the readout anode together with the gate of the FirstFET, and t_t being the time required for transferring the charge from one row to the next. At the time of writing, near Fano-limited resolution can be achieved, using dedicated preamplifiers implementing CDS, with readout times between 2 and 4 μs . Transfer from pixel to pixel takes of the order of 100 ns, which is also the time required for resetting the readout anode.

In this way, a state-of-the-art column-parallel pnCCD can be operated at framerates of between 480 Hz and 930 Hz for a standard mode device or 460–890 Hz for a framestore device, assuming a split-frame CCD topology with a $1 \text{ k} \times 1 \text{ k}$ pixel large imaging area. Accordingly, the readout of smaller sensors scales linearly with their size.

Conventional large area pnCCD systems have successfully been deployed to experiments at free electron lasers, namely, the CFEL-ASG Multi-Purpose instrument (CAMP) [19] and the Laser Applications in Material Processing (LAMP) [20] experiments at the Free-electron LASer in Hamburg (FLASH), and the Linac Coherent Light Source (LCLS), respectively. The concept for the LAMP instrument is shown in Figure 8B, a photo of one pnCCD module is shown in Figure 8A. The X-ray laser illuminates the specimen, and the various CCD detectors record the speckle. The image data is used to reconstruct the structure of the specimen [26]. The main challenge in this kind of experiment is the simultaneous observation of large and small scattering angles that typically have orders of magnitude difference in intensity. To resolve single scattered photons, a sufficiently low noise is required, while high numbers of photons demand a large charge handling capacity.

3.2.3 pnCCD gain

While the register structure of a pnCCD is capable of handling several 10^6 e⁻ within a single pixel, processing the charge puts up a challenge for the following readout chain. At first, the input stage of the preamplifier ASICs will face signal voltages of several volts, due

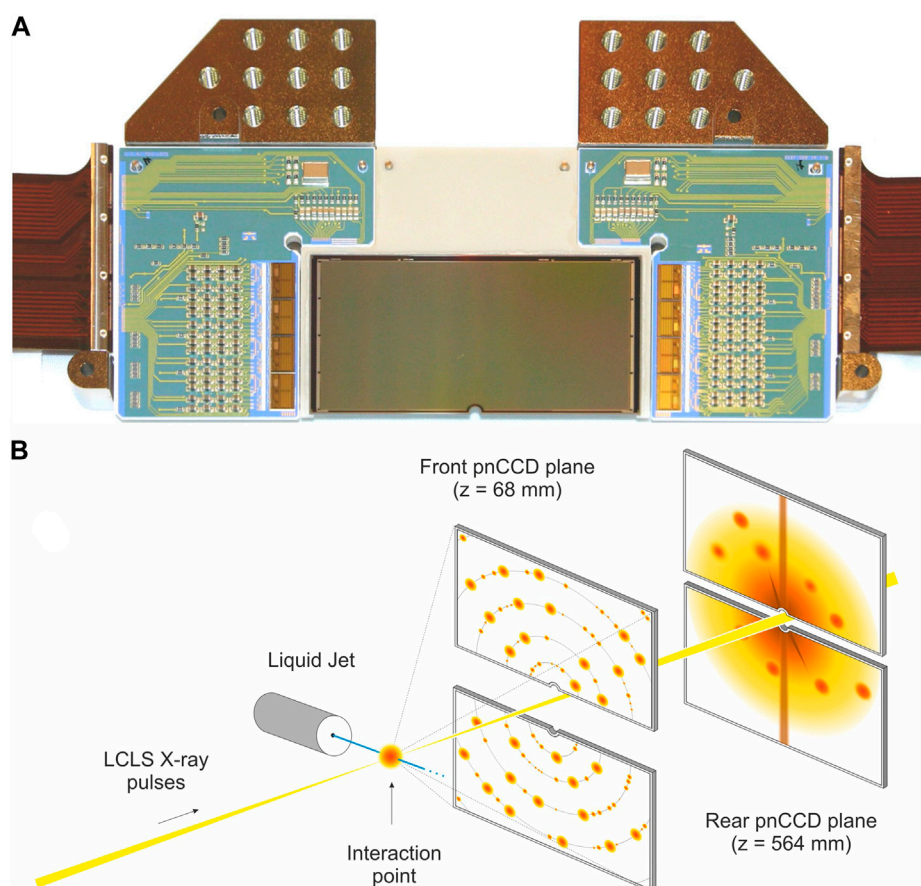


FIGURE 8

Photo of a pnCCD module as used in the CAMP and LAMP experiments (A). The structure is a standard mode, split frame pnCCD. To cover a larger solid angle, two modules can be placed close to each other with minimum dead area. The hole in the central region allows the un-scattered primary beam to pass. Setup of diffraction experiments at CAMP and LAMP using the pnCCD modules (B), using two planes of two pnCCD modules each. The setup corresponds to the one reported in [26].

to the small readout anode capacitance, so special precautions have to be taken not only in order to be able to process the signal levels correctly, but also to prevent damage to the preamplifiers. The respective ASICs need to be furnished with a HV input stage, with configurable attenuator circuitry if necessary.

What is more, the requirement to resolve single photons as well as the extremely high signals in the peaks simultaneously cannot be fulfilled by the conventional CCD readout systems applying preamplifiers with linear gain. As discussed before, multi-linear or non-linear amplification can optimize the system response for photon counting applications. The JFET used for the conventional pnCCD readout, however, is operated in a standard source follower configuration, has a linear characteristic and does not support a customized gain nonlinearity.

The nonlinear or multilinear gain could be integrated into the readout amplifier, but at the time of writing, a pnCCD compatible preamplifier solution implementing these features does not exist. Current pnCCD preamplifiers are conceived for spectroscopy and offer a number of different gain settings, but they can only be operated with one specific gain setting at a time. Depending on the gain setting, information from either high or low intensity features is therefore lost. As a workaround, scattering images from the same

specimen are taken with several gain settings, and the data is combined later on in the analysis. This pragmatic workaround is, however, only suitable for specific types of application scenarios.

A customized preamplifier with the option of multilinear amplification, which would be a new development, would be useful for standard CCDs with a simple anode-based readout, but would not necessarily be the optimum solution for the pnCCD. This is because of the pnCCDs FirstFET, which is the first element in the signal amplification chain. The gain nonlinearity is most effective if it is applied as early in the signal processing as possible.

As shown in section 2, a linear gain function introduces a constant offset to the accuracy function, while the accuracy function for the nonlinear gain converges to zero for small signals. In case of a linear amplifier (e.g., the FirstFET) as primary and a nonlinear amplifier as secondary stage, the FirstFET introduces an offset to the accuracy function, and the secondary stage amplifies this offset with a gain corresponding to its level. For a nonlinear first stage followed by a linear preamplifier, however, the offset to the accuracy function is zero, and, although the linear secondary stage introduces an offset on its own, this offset is not amplified further, so that this configuration has better accuracy at low signal levels.

Implementing the gain nonlinearity into the FirstFET will therefore be superior compared to solutions based on nonlinear or multilinear preamplifiers. An approach currently being developed at the MPG HLL is the replacement of the conventional FirstFET based on a JFET with DePFET based end-of-column amplifiers for the readout of pnCCDs. Critical issue here is the transfer of the signal charge from the transfer channel of the column defined by the HE implant to the internal gate of the DePFET serving as EOC amplifier. In this process, all charge needs to be transferred, even for very high signal levels, without compromising the image information. Once this process is complete, the device benefits from the unique capabilities of the DePFET.

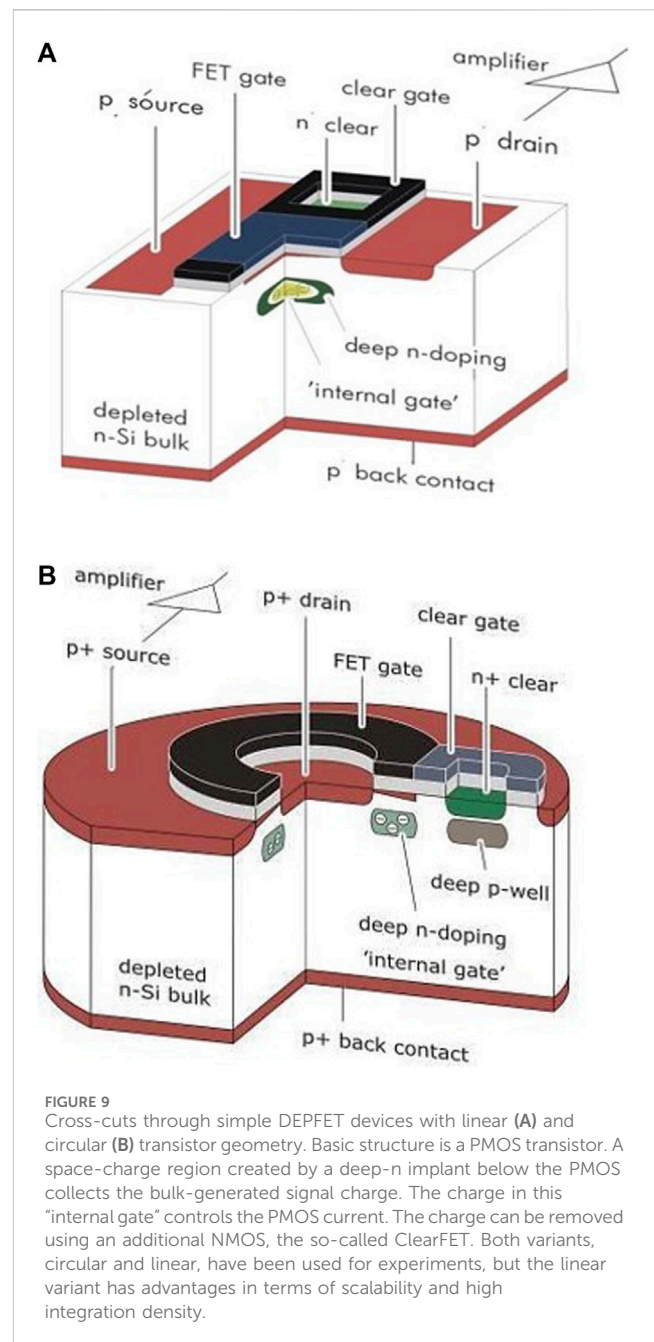
As described below, DePFET structures allow the implementation of a customized non-linear amplification already on sensor die level. A combination of a pnCCD pixel array with an optimized pixel CHC for photon counting with high dynamic range, and a DePFET based EOC amplifier with customized nonlinear amplification is a useful detector device for x-ray scattering imaging with high dynamic range. In addition, it could also help to speed up pnCCD readout.

Although a source follower configuration for readout has some benefits, e.g., its invariance against bias voltage settings and its robustness against radiation-induced threshold voltage shifts, it also puts a limitation on the readout speed. In the source follower configuration, the load capacitance connected to the source node needs to be recharged each time the source voltage changes. This recharge process is driven by the g_m of the FirstFET. Constraints on geometry and capacitance of the readout anode entail certain limitations concerning size and gate capacitance, and thus on the g_m of the FirstFET. The resulting settling time constants are usually not significant when operating with spectroscopic shaping times of several μ s. The picture changes, however, when the readout is to be sped up to the sub- μ s range. Here, the g_m dominated settling time constants could bring about a severe limitation to the readout speed. As explained below, a DePFET based EOC amplifier, in contrast, could be operated in a drain readout configuration, where the potentials on all terminals are kept constant by the front end circuit, eliminating the dependence of the settling time on the g_m of the EOC amplifier. As a side effect, front-end circuitry designed for drain readout eliminates the requirement for a HV input stage.

For photon science a large format pnCCD with a DePFET based readout node, incorporating a non-linear signal response would provide kHz readout rate and large sensitive area with maximized dynamic range. A first small format prototype combining a pnCCD with a DePFET based first amplifier has recently been manufactured. The test system is under construction.

3.3 DePFETs for photon science application

As mentioned above, DePFET [27] devices provide for a different approach to create required nonlinear characteristics, not in the preamplifiers, but in the sensor itself. This is the principle of in-sensor signal compression. Although a more sophisticated sensor is required, the design of the front-end



amplifiers can be simpler and less customization is required. DePFET devices can be used in multiple ways and in a variety of contexts. As mentioned, they can be combined with a pnCCD pixel structure replacing the traditional JFET based FirstFET, adding gain customization and higher readout speed.

They can, however, also be used as building block of an active pixel sensor (APS). The advantage here is a significantly higher framerate, as a much higher degree of readout parallelization is achieved in this way. One example is the so-called EDET sensor using a 4-fold multiparallel readout to achieve a framerate of 80 kHz for a 1 MPixel array. The extreme case is the full parallel readout using the DePFET APS as basic cell of a hybrid pixel detector, which has been realized for the so-called DSSC (DePFET Sensor with Signal Compression [34]).

3.3.1 The DEPFET detector/amplifier structure

The combined detector amplifier structure DEPFET [27] (Depleted P-channel FET) has found application in a variety of projects, ranging from high-energy physics [28, 29], extraterrestrial physics [30, 31], planetology [32, 33], and photon science [34]. A DEPFET structure essentially consists of a PMOS transistor integrated on the surface of a fully depleted, n-type silicon bulk. By means of sideways depletion, the potential minimum for electrons created in this way is shifted towards the surface on which the PMOS structure is located (see Figure 9). Using an additional n implant, the potential minimum is further enforced and confined to the region just below the PMOS gate. Bulk-generated electrons will follow the drift field created by the sideways depletion, and will be collected in the potential minimum beneath the PMOS gate. Here, their presence modulates the PMOS current in the very same way as charge on the external gate. Therefore, the potential minimum is referred to as internal gate. The figure of merit for the internal gate is the so-called charge transconductance, or $g_q = \partial I_{DS} / \partial q_{int}$, describing the change of DEPFET drain source current per change of charge in the internal gate. Typical values here range between 300 pA/e[−] and 1 nA/e[−] [35]. The PMOS current therefore is a measure for the charge present in the internal gate.

The DEPFETs nature as sideways depleted device brings about additional benefits. DEPFETs have an extremely low internal gate capacitance of only few fF, and therefore exhibit intrinsically high signal to noise ratios. They have in common with pnCCDs that in most cases they are illuminated from the backside, which allows for 100% fill factor and provides a certain degree of self-shielding, especially against low-energy X-rays. In addition, the sensors can be furnished with an entrance window configuration optimized for the respective application, e.g., antireflective coating for QE optimization for use in the optical wavelength, or thin entrance window implants for good spectral performance even for the lowest X-ray energies.

3.3.2 DEPFET clear

Conceptually, the DEPFET is a charge integrating type of device. Without external influence, the internal gate will collect all bulk-generated electrons, which continuously lower its potential. To clear the internal gate of the collected charge and to reset the DEPFET structure to a defined state, the clearFET, a dedicated NMOS structure merged with the PMOS, is used. The clearFET uses a separate cleargate contact as gate and an implanted n + contact, the clear contact, as drain, while the internal gate acts as source. In normal biasing conditions (Clear OFF state), the clear contact is shielded from the bulk and the internal gate by the barrier created by the cleargate and an additional deep-p shielding implant. When the internal gate keeps collecting charge, its potential will eventually reach the potential of the barrier towards the clear. This point defines the CHC of the DEPFET, as from now on, excess charge will start leaking to the clear, and the potential of the internal gate will remain constant.

By applying a positive voltage to the cleargate, the potential barrier between internal gate and clear is lowered, and by setting the

clear to a positive voltage as well, electrons from the internal gate are extracted by the clear. If clear and cleargate potentials are set appropriately (Clear ON state), all collected charge is removed from the internal gate.

During operation, clear pulses with a fixed frequency are applied to the pixels. In the phase between the clear pulses, the DEPFETs internal gate will integrate and store bulk-generated charge. Each clear pulse will completely empty the internal gate, and the height of the step produced by the clear pulse is used as a measure for the total amount of charge collected in the internal gate during the charge integration phase. This is done by comparing the signal levels of the DEPFET in the (partially) filled state before the clear and the empty state after the clear. The signal step is evaluated using CDS with dedicated preamplifier/shaper ASICs. Sometimes, e.g., in cases requiring extremely short readout times, no individual baseline, only one pixel-individual reference value is measured for the complete acquisition cycle. Thus, the same value is subtracted from all full state signal levels, compromising the noise figure in favor of the readout speed.

An important point to mention is that the DEPFET can be used as an integrating type of detector using readout on demand. Its internal gate can store charge regardless from the presence of a transistor current. This allows the detector pixel to reside in a low power biasing state, while all the pixels are still sensitive.

3.3.3 DEPFET readout

As mentioned before, the charge in the DEPFETs internal gate influences the channel conductivity in the PMOS channel. To evaluate the channel conductivity, two common methods are in use. Like the FirstFET of a pnCCD it can be operated as a source follower. Changes in the charge collected in the internal gate will then be converted into a voltage step at the source of the DEPFET, which, in first order approximation, is

$$\partial V_s \approx - \frac{g_q}{g_m} \cdot \partial q_{int}$$

as the external gate tries to keep the source voltage at the old level. To circumvent the intrinsic limitations of the source follower, an increasing number of applications use the so-called drain readout, where the DEPFET is biased with constant potentials on drain and source, and changes in the internal gate charge are converted to a current step, which is evaluated by means of a Trans-Impedance Amplifier (TIA):

$$\partial I_{DS} = g_q \cdot \partial q_{int}$$

Main advantage of this approach is the higher speed. For the source follower the source node has to be recharged to its new potential during signal formation. In this process, the capacitance of the source in combination with the g_m of the DEPFET limits the rise time of the source follower, especially as, to optimize the source follower gain, the DEPFET was on purpose designed with a low g_m . This intrinsic speed limitation does not apply to the drain readout case, as the potentials on the DEPFET drains are kept constant by the TIA, and the achievable readout speed is higher. This is why drain readout is used for most applications requiring higher readout speeds. The design of the input stage of the TIA, however, must assure stability even for the high load capacitance values encountered for large matrix devices.

3.3.4 DEPFET matrix operation

Due to their nature as sideways depleted devices, DEPFETs can be utilized as sensitive elements for pixels in a variety of detector configurations, covering a wide range of pixel sizes from the mm scale, e.g., if surrounded by driftings, down to few tens of μm if highly dense, coupled designs are chosen. They also provide for a huge variety of options for in-pixel functionality, like electronic shutters, multiple storages, repetitive sampling and many more [35, 36]. In the simplest configuration, however, the sensor surface is subdivided in a rectangular array of pixels, each of which is furnished with a DEPFET as central detector/amplifier element.

Conceptually, the DEPFET provides for a high degree of flexibility concerning pixel interconnection and readout, as all relevant terminals are individually accessible. Arbitrary patterns of pixels could be flexibly addressed and read out on demand, with a readout speed adapted to the respective signal. In practice, however, only two different readout configurations have found their way into application. The first approach is the fully parallel readout, where all pixels are connected to the same control- and bias traces, and all readout terminals are connected to their own individual readout electronics channel. This approach has been implemented for DEPFETs for the DSSC, and also for high-speed spectroscopy. Fully parallel readout offers highest framerate and readout speed capabilities, especially if combined with drain readout. This approach is, however, also most challenging for large area matrices due to, e.g., the required high integration density of the readout electronics, fast timing requirements, and high demands on the driver electronics for the control terminals.

The second approach is the so-called n -fold parallel readout, where different subsets of pixels share their readout electronics, and only one subset of pixels is read out at a time. A topologically simple subdivision here, which is also the most widely used, is connecting all pixels of one, or sometimes a number of n , rows to be read out in parallel. The associated control electronics only needs to access all pixels within a subset, but with the option to access all subsets sequentially. This simplifies pixel interconnect on a matrix and a much lower number of readout channels is required compared to the fully parallel readout case. Consequently, the overall complexity and integration density of such a system is substantially reduced compared to a full parallel readout system. The readout of such a system happens in a rolling-shutter mode, where the subsets, typically rows or numbers of rows, are switched on and read out sequentially, with the process starting over once all subsets have been read. This mode also offers the feature to read subsets of the matrix, windows or regions of interests (ROIs) with elevated frame rate and time resolution, if the matrix control circuitry allows to flexibly address the pixels to be read out.

3.3.5 DEPFETs with nonlinear amplification

It was stated above that the nonlinear characteristics needed to match the accuracy of counting error can also be achieved by modifying the response of the detector itself, rather than by dedicated front-end electronics. The DEPFET structure is an ideal platform for the implementation of such an approach due to its operation principle. The DEPFETs gain characteristics can be precisely modified using g_q engineering.

The DEPFET translates the presence of signal charge in the internal gate into a modulation of the PMOS current. The

mechanism here is that the charge carriers in the internal gate induce mirror charge in the PMOS channel, which contribute to the overall channel conductivity, in the same way as charge on the external gate creates the channel in the first place. To achieve maximum influence of charge in the internal gate, the design must minimize stray capacitance towards adjacent electrodes other than the PMOS channel. To achieve this, a weak *deep-n* implant creates a space charge region, which confines the signal charge in a certain depth and position right below the external gate.

Within a typical pixel environment, potential barriers from source and drain and the potential barrier to the clear surround the internal gate. The potential barrier to the clear usually defines the pixels CHC, as charge starts leaking to the clear once the internal gate becomes negative enough to allow electrons to overcome this barrier.

The idea behind the technique of g_q -engineering is to create regions within the pixel area other than the clear allowing charge to overflow to once the internal gate reaches a certain fill level. This overflow must occur underneath a region with a potential more positive than the clear barrier in order to prevent charge loss. For a normal pixel layout, the region below the source, as the most positive electrode nearby, is the obvious overflow target. There are various ways to extend the internal gate to the overflow regions. One option, first realized for the DEPFETs for DSSC [37, 38], uses additional *deep-n* implants. In this variant, the first overflow region is created by a second weak *deep-n* implant extending also below the source. It needs to be noted that a slightly different source implantation, the so-called *p_{son}*, is used to define the source region wherever an overflow region is implemented below. This is because the normal p + source implant would compensate large parts of the *deep-n*.

Once the threshold fill level is reached, charge will start to extend to the overflow region, which will see a much higher parasitic capacitance towards the source, so its influence on the PMOS current will be much lower. The parameters of the *deep-n* overflow implant, mainly dose and size, adjust both the fraction of source area the overflow will extend to and the threshold value at which the overflow will occur. In effect, these parameters determine the onset point (commonly referred to as “kink”) and the effective g_q for charge in the overflow region. In the same way, even more overflow regions can be added by means of even further *deep-n* implants (see Figure 10).

In this way, a DEPFET device with graded internal gate is generated, where different sections in the dynamic range show different g_q , each with its unique onset point. The actual nonlinear characteristics depends on a large variety of input parameters like requested primary g_q , gain function shape, overall dynamic range or requested pixel size. In addition, the final gain function will smoothen out due to diffusion processes in the semiconductor manufacturing process. Manufacturing technology and actual pixel design are verified using best practice design optimization methods, which utilize process simulations and 3D device simulations. The overflow regions also increase the CHC of the sensor pixels.

An example of DEPFET characteristics with nonlinear gain manufactured for the DSSC at the MPG HLL is shown in Figure 11. The DSSC uses large modules with hexagonal DEPFET pixels with circular geometry similar to Figure 9B. Gate and source concentrically surround the drain, which is located in the

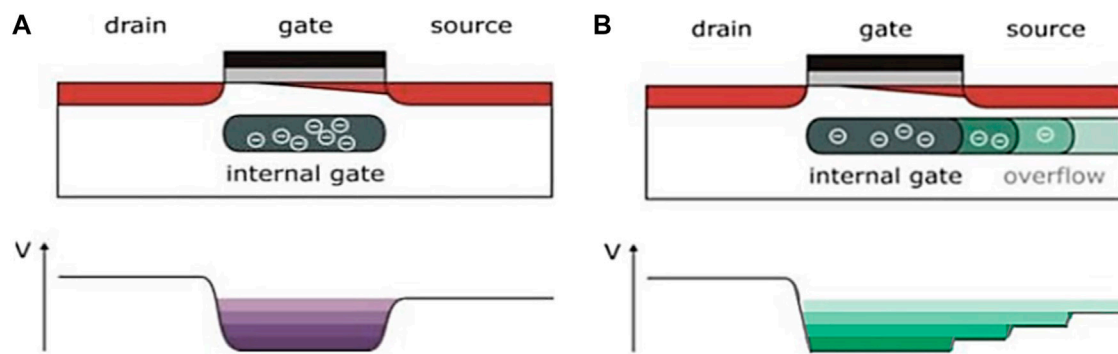


FIGURE 10

Principle of nonlinear response with a DEPFET device. Standard DEPFET Devices (A) have their internal gate underneath the external gate. Mirror charge is induced only in the PMOS gate, and stray capacitance is minimized. For a DEPFET with nonlinear response (B), dedicated overflow regions under the source are created. Above a certain fill level, signal charge starts to spread underneath the source. Here, the parasitic induction to the source is much larger, effecting a much lower g_q than for the base charge. The overflow regions can be created using “graded” internal gate implantations. The superposition of multiple deep-n implants creates an internal gate extending underneath the source in several steps. This creates fill level sections with different g_q each.

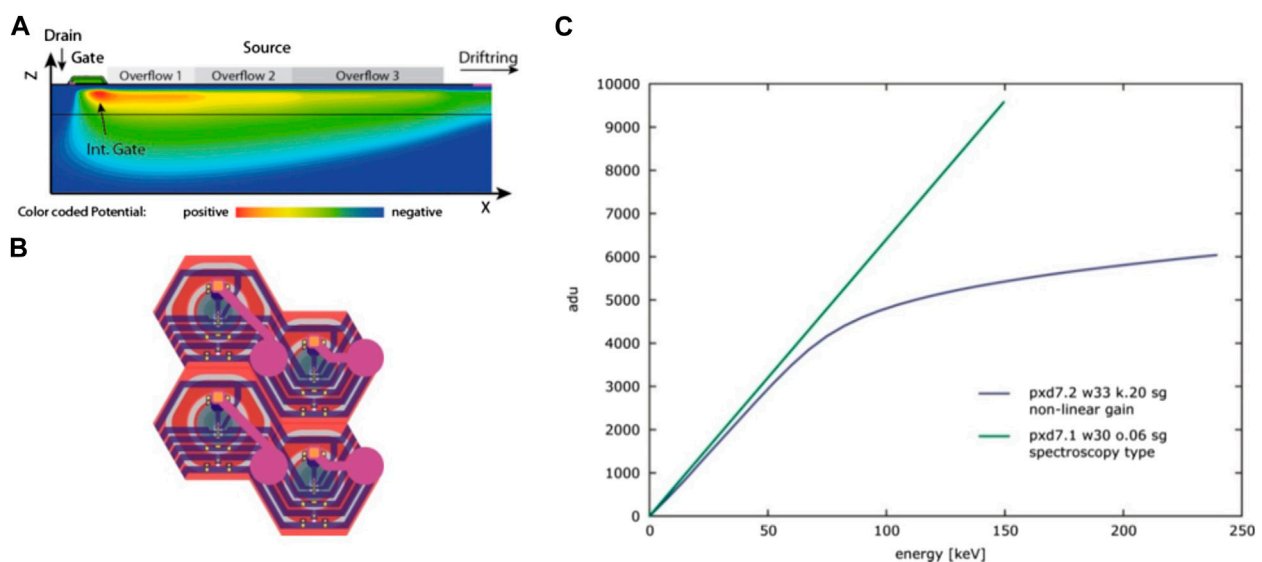


FIGURE 11

Implementation example for nonlinear characteristics. Inlay (B) shows the layout of nonlinear DEPFET pixels for DSSC in hexagonal geometry with 200 μm pixel size. Here, the graded internal gate potential is created by an overlay of three different deep-n implants. Inlay (A) shows a radial cut from the center of the pixel to its circumference. The potential distribution in the pixel area is color-coded. The three overflow regions of the internal gate are indicated by yellow to red colors, indicating the potential becoming more negative for regions closer to the pixel center. Inlay (C) shows the resulting gain curve is shown in blue. For comparison, the green curve shows the (linear) gain curve of a traditional spectroscopy grade DEPFET device.

center of the pixel, only sparing the clear region. Accordingly, the various overflow regions extend concentrically underneath the part of the source region defined by the p_{son} . Outside the overflow regions, the conventional p implant in form of a concentric ring reaching out to the first drift electrode defines the source region. To fill the pixel area of $200 \times 200 \mu\text{m}^2$, a drift structure with two drift rings surrounds the DEPFET readout node. The DEPFETs operate in pulsed clear operation mode, where all pixels are read in parallel with an overall maximum framerate of 5 MHz. The preamplifiers contact the sensor pixels via bump bonding.

The hexagonal approach for DSSC works well, but is of limited scalability, as far as pixel size is concerned. The smaller the pixel size, the more difficult it gets to superimpose the various deep-n implants underneath the limited source area. For smaller pixels, a different approach providing for easier scalability has been developed [39, 40]. Here, the g_q is engineered not by graded deep-n implantations for the internal gate, but by the geometrical design of the source implant itself. This type of layout was developed for the so-called EDET project [41], and is based on heritage from the very compact vertex detector designs for the pixel detector of the BELLE II silicon vertex tracker. A layout example is shown in Figure 12A.

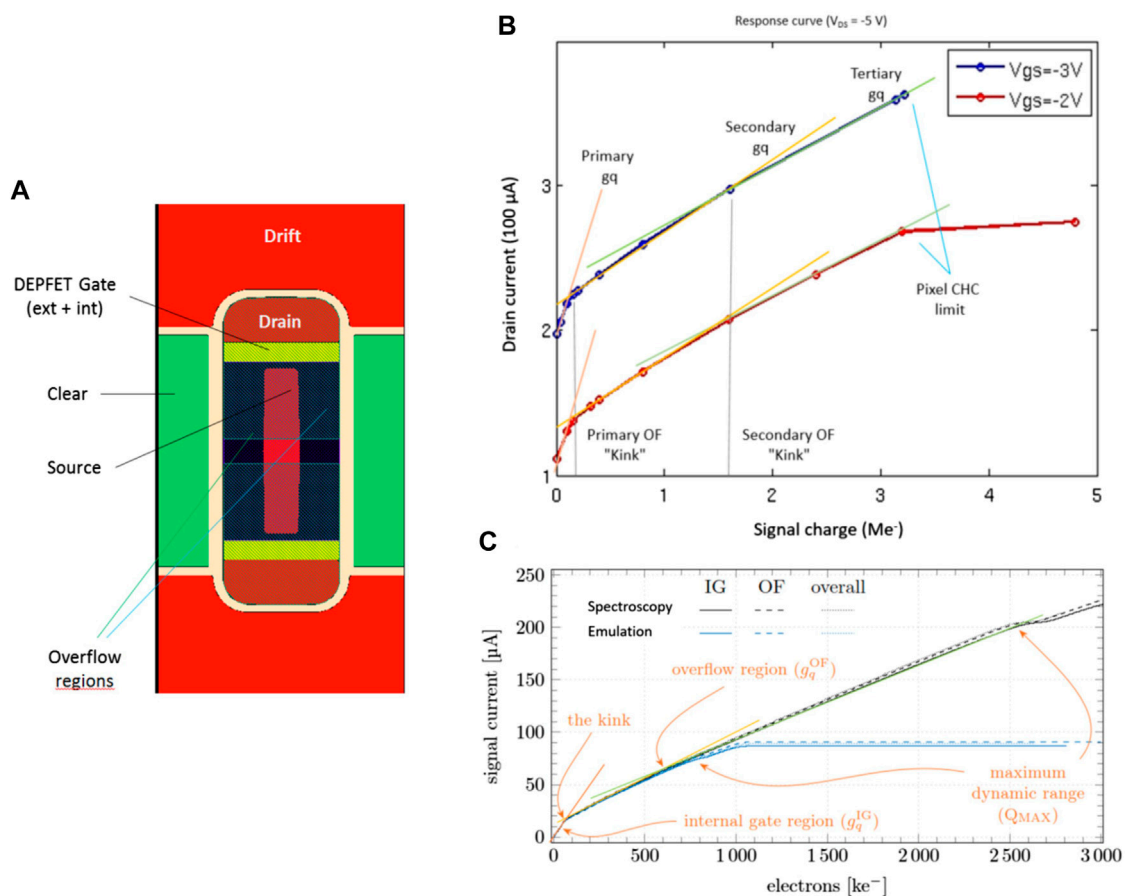


FIGURE 12

Implementation examples for nonlinear characteristics. Inlay (A) shows the pixel design for the EDET DH80 k direct electron detector for TEM imaging with $60 \times 60\text{ }\mu\text{m}^2$ pixel size. Here, the graded potential was created by geometrically tailoring the p + implant for the source, which partially compensates the deep-n for the internal gate. This has a similar effect on the internal gate as the graded deep-n approach, but is technologically simpler. The layout generates the gain curves as shown in (B, C) for different operating conditions. The wedge shape on the right side of the source implant creates the second kink, or overflow (OF). Inlay (B) shows the simulated performance curves, inlay (C) the performance as measured. The upper plot shows the simulated performance curves, the lower one the performance as measured on prototype devices. The main features from the simulation can be found in the measured performance curves as well, although the secondary kink seems to be at slightly lower charge values, and the difference between secondary and tertiary g_q appears to be smaller than expected. The blue curves correspond to simulated operating conditions after receiving severe radiation damage and show the expected decrease in CHC. This effect can be mitigated by annealing.

Unlike DSSC, the target application for the EDET detector is not photon counting, but intensity measurement in the focal plane of a TEM. Here, the transition of the TEM primary electrons deposits the signal charge in the detector, and the intensity is measured by counting the number of primaries. To avoid compromising the spatial resolution by multiple scattering, the detectors pixels are integrated on a thinned SOI substrate, with a substrate thickness of only 50 to 30 μm . The target application uses a primary electron energy of around 300 keV, depositing 8,000 electrons (MPV) each in the 50 μm thick substrate. To achieve sufficient contrast, a pixel should be able to store the signal charge of up to 100 primaries, which yields a pixel CHC larger than 800×10^3 electrons. It needs to be noted that for this application, the stimulus error ΔS_0 is of course given the quadratic sum of the error on the energy deposition of the primaries in the sensor and the corresponding Fano contribution. The error on the energy deposition is much higher than for photons and can, due to the thin substrates, be estimated in good approximation by the noise RMS of a Vavilov distribution.

Due to the spatial constraints the design has to be more compact compared to the circular DSSC sensor with its isolated $200 \times 200\text{ }\mu\text{m}^2$ pixels. The basic building block of the EDET sensors is a so-called couple of two pixels with linear geometry, with $60 \times 60\text{ }\mu\text{m}^2$ area each. The couple partners are rotated by 180° with respect to each other. The pixels within one couple share source, clear, cleargate and gate contact, so the pixels have to be read out in parallel. Each couple partner is read out via its individual drain contact. The final matrix layout places couples next to each other in a row, in such a way that they share the common clear areas. Negatively biased p + regions, the so-called drift electrodes, separate the different rows of couples and act as a focusing implant for the bulk generated charge similar to the drift rings for DSSC.

The main improvement compared to the DSSC technology is that the geometry of the source electrode generates the nonlinear gain. As for DSSC, two implants define the source area, the *p_{son}* and the *p₊*. But as the *p₊* compensates the *deep-n* in large parts, and the *p_{son}* does not, the overflow region extends only underneath the

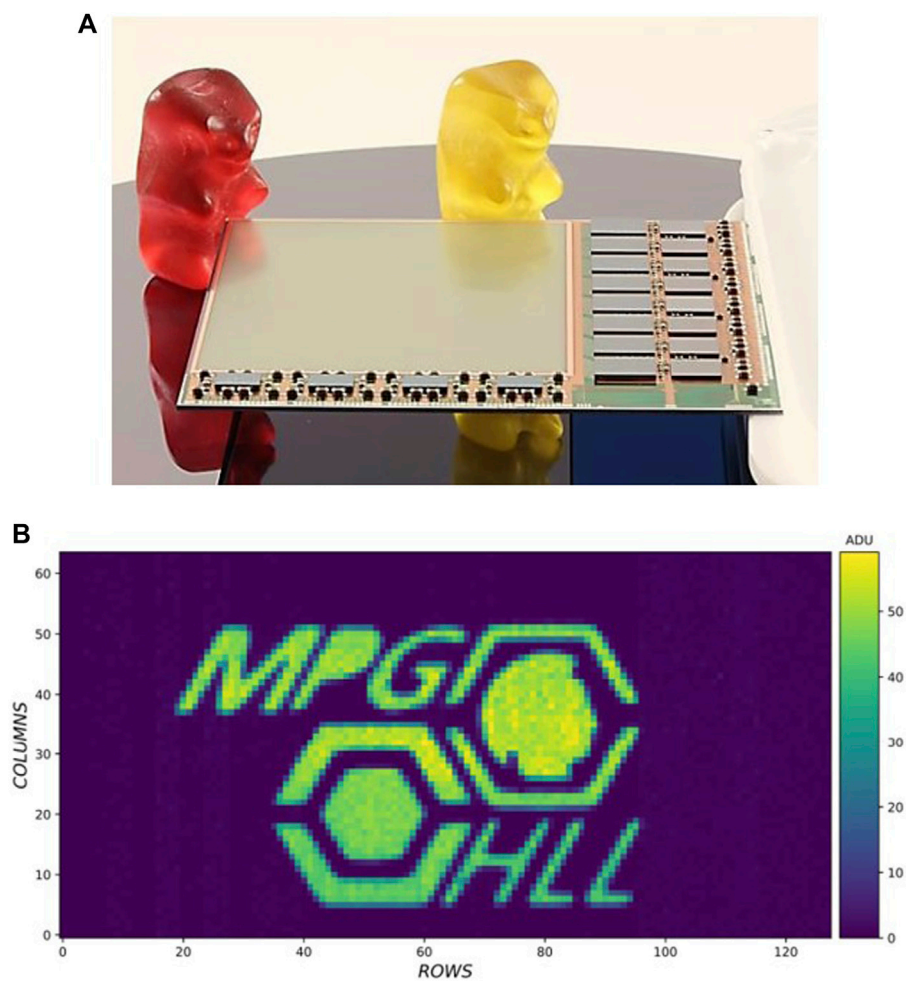


FIGURE 13
(A) Photo of an EDET Quadrant matrix with 512×512 pixels of $60 \times 60 \times 60 \mu\text{m}^2$ size, integrated on an ASM together with passives and front end electronics. **(B)** Example image taken with a prototype matrix on a TEM using a shadow mask.

pson. Thus, the layout of the $p+$ implant modifies both size and potential of the overflow region underneath the *pson*. By modifying the details of the $p+$ shape, e.g., adding a wedge or curvature, the effective g_q of the overflow region can be modeled to achieve the desired characteristics. In contrast to DSSC, only one *deep-n* implant is required, which significantly reduces the degree of complexity for the technology. Again, the actual design of the $p+$ shape is verified using design optimization methods, process simulations and 3D device simulations.

The CHC of the pixels has two limitations. As soon as the overflow region of a pixel is filled to a level at which the potential of the internal gate becomes more negative than the potential barrier to the neighboring couple partner, charge will start to spill over to the internal gate of the couple partner (primary CHC). The couple partner will then start to see the charge injected into its neighbor. This “couple spillover” will continue as long as the pixels have different fill levels. Once the couple partner is filled up as well, the couple will resume being filled up until the clear barrier is overcome and charge loss occurs (secondary CHC). The layout of a pixel couple is what is shown in Figure 12, alongside with simulated and measured performance curves. An initial CHC of 3×10^6 electrons

has been verified. CHC will suffer in the long run under irradiation mainly due to cleargate voltage shifts, but this effect can be mitigated by annealing. In addition, radiation hard oxides for the DEPFET have been successfully prototyped.

For the intended application [41], direct bump-bonding of the front end electronics on the sensitive area is not the preferred solution, as the materials of the front end ICs and the bumps cause backscattering, which spoils the imaging performance. Therefore, the signals are routed to the edge of the sensor die, where the handle wafer below the thin sensor substrate has not been removed, providing for a solid frame around the sensitive region. This so-called service balcony integrates the complete front-end and control electronics as well as passives for voltage decoupling, forming a so-called all-silicon-module (ASM). The ASM interfaces the peripheral readout electronics as well as the thermomechanical support of the sensors. A photo of a quadrant ASM is shown in Figure 13, alongside with an image taken with a shadow mask onto the TEM using a prototype matrix.

The sensors are read out using readout-on-demand, reading a total number of four pixel rows, i.e., two couple rows, in parallel with an advanced timing of 100 ns per readout. In this way, a complete

sensor matrix of 512×512 pixels can be read in $12.8 \mu\text{s}$. The system uses drain readout, using the so-called DCD-E IC as front-end, a 256 channel TIA/digitizer IC with a resolution of 8 bit. Each DCD hands its digitized data over to one digital sequencer/buffer IC of type DMC, which stores the data before it is transferred to the peripheral data acquisition system using a 1.6 GHz AURORA link. In this way, bursts of up to 50 images with maximum time resolution can be acquired. Using selective readout of window ROIs, both (window) framerate and number of images per burst increase by the same factor.

4 Summary

Three different approaches for dealing with the increasing demands on photon science instrumentation, currently under development at the MPG HLL were presented. Especially challenging is meeting the requirements on dynamic range at low and high end simultaneously. Optimum solutions here require sophisticated front-end electronics and detector structures capable of applying customized nonlinear gain characteristics.

The newly developed MARTHA structures, a new type of APD arrays operated in low gain, overcome traditional drawbacks of APD arrays and have the potential to increase the low energy detection limit by applying avalanche multiplication. If combined with multilinear front-end electronics, they could serve as drop-in replacement for existing pad or pixel detector based solutions for hybrid pixel detector systems.

New operation modes for pnCCDs allow to drastically increase their CHC up to the point where they become interesting for high contrast photon science measurement. They can offer a technologically simple solution for covering large areas in applications, where framerate is not the main driver. To accommodate for the high dynamic range, DEPFET based EOC amplifiers with nonlinear gain are used to customize the dynamic range to the requirements. The use of DEPFET based EOCs here also helps to increase readout speed of the CCDs.

Finally, DEPFET active pixel arrays offer the potential to combine nonlinear gain, implemented in each pixel, with multiple possible readout modes. Systems implementing full parallel readout as well as four-fold multiparallel readout have already been implemented. In addition to its potentially higher readout speed, the DEPFET also offers higher flexibility in terms of readout modes, e.g., reading of windows and ROIs, and pixel sizes. In case of a one- or twofold multiparallel readout, however, a pnCCD using the DEPFET only as EOC amplifier could be an

alternative, if no advanced readout modes and pixel sizes within a certain range are requested.

Although not discussed in this paper, modularity is another important prerequisite for photon science detectors. Many experiments benefit from large solid angles, so assembling large composite focal planes from multiple smaller detector modules with minimum dead area is an important capability. Most hybrid pixel detectors are build that way, but also pnCCDs and DEPFET pixel arrays can be used for building highly integrated, compact, 4-side buttable sensor tiles. The required advanced mechanical and thermal support structures have been successfully prototyped.

Author contributions

JN: Conceptualization, Writing–original draft, Writing–review and editing. AB: Conceptualization, Investigation, Methodology, Writing–original draft, Writing–review and editing. RR: Conceptualization, Investigation, Methodology, Writing–original draft, Writing–review and editing. JT: Conceptualization, Investigation, Methodology, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Hart M, Angelsen C, Burge S, Coughlan J, Halsall R, Koch A, et al. Development of the LPD, a high dynamic range pixel detector for the European XFEL. In: Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC); November 2012; Anaheim, CA, USA (2012). p. 534–7. doi:10.1109/NSSMIC.2012.6551165
- Allahgholi A, Becker J, Bianco L, Delfs A, Dinapoli R, Goettlicher P, et al. AGIPD, a high dynamic range fast detector for the European XFEL. *J Instrumentation* (2015) 10: C01023. doi:10.1088/1748-0221/10/01/C01023
- Photonics Online. Silicon avalanche photodiode (APD) arrays (2024). Available at: <https://www.photonicsonline.com/doc/silicon-avalanche-photodiode-apd-arrays-0001>.
- Pellegrini G, Fernandez-Martinez P, Baselga M, Fleta C, Flores D, Greco V, et al. Technology developments and first measurements of low gain avalanche detectors (LGAD) for high energy physics applications. *Nucl Instrum Methods A* (2014) 765:12–6. doi:10.1016/j.nima.2014.06.008
- Cartiglia N, Arcidiacono R, Boscardin M, Dalla Betta GF, Cartiglia N, Cenna F, et al. Design optimization of ultra-fast silicon detectors. *Nucl Instrum Meth A* (2015) 796: 141–8. doi:10.1016/j.nima.2015.04.025
- Sadrozinski H-W, Baselga M, Ely S, Fadeyev V, Galloway Z, Ngo J, et al. Ultra-fast silicon detectors. *Nucl Instrum Meth A* (2014) 765:226–31. doi:10.1016/j.nima.2013.06.033
- Paternoster G, Borghi G, Arcidiacono R, Boscardin M, Cartiglia N, Centis Vignali M, et al. Novel strategies for fine-segmented low gain avalanche diodes. *Nucl Instrum Meth A* (2021) 987:164840. doi:10.1016/j.nima.2020.164840

8. Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of LGAD sensors with a thin entrance window for soft X-ray detection. *JINST* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/C11011
9. Cartiglia N, Arcidiacono R, Borghi G, Boscardin M, Costa M, Galloway Z, et al. LGAD designs for future particle trackers. *Nucl Instrum Meth A* (2020) 979:164383. doi:10.1016/j.nima.2020.164383
10. Zhao Y, Ayyoub S, Chen W, Gee C, Islam R, Mazza M, et al. A new approach to achieving high granularity for silicon diode detectors with impact ionization gain. *J Phys Conf Ser* (2022) 2374:012171. doi:10.1088/1742-6596/2374/1/012171
11. McIntyre RJ. Recent developments in silicon avalanche photo-diodes. *Measurement* (1985) 3:146–52. doi:10.1016/0263-2241(85)90024-7
12. Synopsys TCAD. TCAD solutions (2024). Available at: www.synopsys.com/manufacturing/tcad.html.
13. van Overstraeten R, DeMan H. Measurement of the ionization rates in diffused silicon p-n junctions. *Solid-State Electro* (1970) 13:583–608. doi:10.1016/0038-1101(70)90139-5
14. Ebermayer S, Andritschke R, Elbs J, Meidinger N, Strüder L, Hartmann R, et al. Quantum efficiency measurements of eROSITA pnCCDs. *Proc SPIE* (2010) 7742:77420U. doi:10.1117/12.856939
15. Currás E, Fernández M, Moll M. Gain reduction mechanism observed in low gain avalanche diodes. *Nucl Instrum Meth A* (2022) 1031:166530. doi:10.1016/j.nima.2022.166530
16. Jansen F, Lumb D, Altieri B, Clavel J, Ehle M, Erd C, et al. XMM-Newton observatory. *A&A* (2001) 365:L1–6. doi:10.1051/0004-6361:20000036
17. Strüder L, Brie U, Dennerl K, Hartmann R, Kendziorra E, Meidinger N, et al. The European photon imaging camera on XMM-Newton: the pn-CCD camera. *A&A* (2001) 365:L18–26. doi:10.1051/0004-6361:20000066
18. Merloni A, Predehl P, Becker W, Böhringer H, Boller T, Brunner H. eROSITA science book: mapping the structure of the energetic universe (2012). Available at: <https://arxiv.org/abs/1209.3114>.
19. Erk B, Müller JP, Bomme C, Boll R, Brenner G, Chapman HN, et al. CAMP@FLASH: an end-station for imaging, electron and ion-spectroscopy, and pump-probe experiments at the FLASH free-electron laser. *J Synchrotron Radiat* (2018) 25:1529–40. doi:10.1107/S1600577518008585
20. Osipov T, Bostedt C, Castagna J-C, Ferguson KR, Bucher M, Montero SC, et al. The LAMP instrument at the linac coherent light source free-electron laser. *Rev Scientific Instr* (2018) 89:035112. doi:10.1063/1.5017727
21. Strüder L, Epp S, Rolles D, Hartmann R, Holl P, Lutz G, et al. Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. *NIMA* (2010) 614:483–96. doi:10.1016/j.nima.2009.12.053
22. Schmidt J, Hartmann R, Holl P, Huth M, Lutz G, Soltau J, et al. Dynamic range of fully depleted pnCCDs: modeling and experimental confirmation. *Proc SPIE* (2014) 9154. doi:10.1117/12.2055343
23. Meidinger N, Bonerz S, Eckhardt R, Englhauser J, Hartmann R, Hasinger G, et al. First measurements with a frame store PN-CCD X-ray detector. *NIMA* (2003) 512:341–9. doi:10.1016/S0168-9002(03)01912-0
24. Hermann S, Buttler W, Hartmann R, Meidinger N, Porro M, Strueder L, et al. CAMEX readout ASICs for pnCCDs. In: Proceedings of the 2008 IEEE Nuclear Science Symposium Conference Record; October 2008; Dresden, Germany (2008). doi:10.1109/NSSMIC.2008.4774983
25. Porro M, Bianchi D, De Vita G, Hartmann R, Hauser G, Hermann S, et al. VERITAS: a 128-channel asic for readout of pnCCDs and DEPFET arrays for X-ray imaging, spectroscopy and XFEL applications. *IEEE TNS* (2013) 60(1):446–55. doi:10.1109/TNS.2012.2228410
26. Chapman H, Fromme P, Barty A, White TA, Kirian RA, Aquila A, et al. Femtosecond X-ray protein nanocrystallography. *Nature* (2011) 470:73–7. doi:10.1038/nature09750
27. Kemmer J, Lutz G. New detector concepts. *NIMA* (1987) 253:365–77. doi:10.1016/0168-9002(87)90518-3
28. Prereira DE. DePFET active pixel sensors for the vertex detector of the Belle-II experiment. *JINST* (2013). doi:10.1088/1748-0221/9/03/C03004
29. Ye H, Abudinen F, Ackermann K, Ahlburg P, Albalawi M, Alonso O, et al. Commissioning and performance of the Belle II pixel detector. *NIM A* (2021) 987:164875. doi:10.1016/j.nima.2020.164875
30. Meidinger N, Eder J, Fürmetz M, Nandra K, Pietschner D, Strecker R, et al. Development of the wide field imager for athena. *Proc SPIE* (2015) 9601. doi:10.1117/12.2187012
31. Meidinger N, Barbera M, Emberger V, Fürmetz M, Manhart M, Müller-Seidlitz J, et al. The wide field imager instrument for athena. *Proc SPIE* (2017) 10397. doi:10.1117/12.2271844
32. BunceMartindale EJA, Lindsay S, Muinonen K, Rothery DA, Pearson J, McDonnell I, et al. The BepiColombo mercury imaging X-ray spectrometer: science goals, instrument performance and operations. *Space Sci Rev* (2020) 216:126. doi:10.1007/s11214-020-00750-2
33. Majewski P, Andricek L, Christensen U, Hilchenbach M, Lauf T, Lechner P, et al. DEPFET macropixel detectors for MIXS: first electrical qualification measurements. *IEEE Trans Nucl Sci* (2010) 57(4):2389–96. doi:10.1109/TNS.2010.2053557
34. Maffessanti S, Hansen K, Aschauer S, Castoldi A, Erdinger F, Fiorini C, et al. A 64k pixel CMOS-DEPFET module for the soft X-rays DSSC imager operating at MHz-frame rates. *Sci Rep* (2023) 13:11799. doi:10.1038/s41598-023-38508-9
35. Andricek L, Bähr A, Lechner P, Ninkovic J, Richter RH, Schopper F, et al. DEPFET -recent developments and future prospects. *Front Phys Sec Radiat Detectors Imaging* (2022) 10. doi:10.3389/fphy.2022.896212
36. Bähr A, Feller A, Lechner P, Ninkovic J, Richter R, Schopper F, et al. Advanced DEPFET concepts: quadropix. *NIM A* (2018) 912:70–4. doi:10.1016/j.nima.2017.10.048
37. Lechner P, Andricek L, Aschauer S, Bähr A, De Vita G, Hermenau K, et al. DEPFET active pixel sensor with non-linear amplification. In: Proceedings of the 2011 IEEE Nuclear Science Symposium Conference Record; October 2011; Valencia, Spain (2011). doi:10.1109/NSSMIC.2011.6154112
38. Aschauer S, Majewski P, Lutz G, Soltau H, Holl P, Hartmann R, et al. First results on DEPFET Active Pixel Sensors fabricated in a CMOS foundry—a promising approach for new detector development and scientific instrumentation. *JINST* (2017) 12:P11013. doi:10.1088/1748-0221/12/11/P11013
39. Predikaka M, Bähr A, Koffmane C, Ninkovic J, Prinker E, Richter R, et al. EDET DH80k - characterization of a DEPFET based sensor for TEM direct electron imaging. *NIM A* (2020) 958:162544. doi:10.1016/j.nima.2019.162544
40. Predikaka M. *Characterization of a novel DEPFET based sensor for an ultra-fast Transmission Electron Microscopy application*. PhD Thesis. Munich, Germany: Technische Universität München (2022). Available at: <https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20220627-1656556-1-5>.
41. Dourki I. *Development of a novel electron detector with 80 kHz frame rate for imaging applications: simulation of the detector response and first experimental results*. PhD Thesis. Hamburg, Germany: University of Hamburg (2019). Available at: <https://ediss.sub.uni-hamburg.de/handle/ediss/6325>.



OPEN ACCESS

EDITED BY

Iain Sedgwick,
Rutherford Appleton Laboratory,
United Kingdom

REVIEWED BY

Marcus French,
Rutherford Appleton Laboratory,
United Kingdom
Arkadiusz Dawiec,
Soleil Synchrotron, France

*CORRESPONDENCE

Erik Fröjd, erik.frojd@psi.ch

RECEIVED 30 September 2023

ACCEPTED 09 January 2024

PUBLISHED 05 February 2024

CITATION

Fröjd E, Bergamaschi A and Schmitt B (2024),
Single-photon counting detectors for
diffraction-limited light sources.
Front. Phys. 12:1304896.
doi: 10.3389/fphy.2024.1304896

COPYRIGHT

© 2024 Fröjd, Bergamaschi and Schmitt. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Single-photon counting detectors for diffraction-limited light sources

Erik Fröjd*, Anna Bergamaschi and Bernd Schmitt

Paul Scherrer Institut, Villigen, Switzerland

When first introduced, single-photon counting detectors reshaped crystallography at synchrotrons. Their fast readout speed enabled, for example, shutter-less data collection and fine slicing of the rotation angle and boosted the development of new experimental techniques like ptychography. Under optimal conditions, single-photon counting detectors provide an unlimited dynamic range with image noise only limited by the Poisson statistics of the incoming photons. Counting the pulses from individual photons, essentially what made the detectors so successful, also causes the main drawback, which is the loss of efficiency at high photon fluxes due to pulse pileup in the analog front end. To fully take advantage of diffraction-limited light sources, the next-generation single-photon counters need to improve their count rate capabilities in the same order of magnitude as the increased flux. Moreover, fast frame rates (a few kHz) are required to cope with the shorter dwell time achievable, thanks to the higher flux. Detector architecture with multiple comparators and counters can open new possibilities for energy-resolved imaging, while interpixel communication can overcome the issues arising from charge sharing and reduce the loss of efficiency at the pixel corners. Coupling single-photon counting detectors to high-Z sensors for hard X-ray detection (> 20 keV) and to low-gain avalanche diodes (LGADs) for soft X-rays is also necessary to make use of the increased coherence of the new light sources over the full radiation spectrum. In this paper, we present possible strategies to improve the performance of single-photon counting detectors at the fourth-generation synchrotron sources and compare them to charge integrating detectors.

KEYWORDS

hybrid detectors, x-ray detectors, pixel detectors, single-photon counting, synchrotron

1 Introduction

The idea of hybrid pixel detectors originated around the end of 1980s from within the particle physics community [1]. The first fully functional hybrid pixel detector was tested in the Omega-Ion experiment (WA94) in 1991 [2], and larger installations soon followed [3; 4]. Scientists realized early that hybrid detectors also would be a good fit for X-ray imaging [5; 6], and several groups [7; 8] including PSI Brönnimann et al. [9] started working on dedicated detectors. The development of single-photon counting (SPC) hybrid pixel detectors at PSI was motivated by the needs of diffraction applications at the Swiss Light Source, and most of the know how came from working on the original pixel chip for the CMS experiment at the LHC [10]. The PILATUS 1-M detector was the first large-

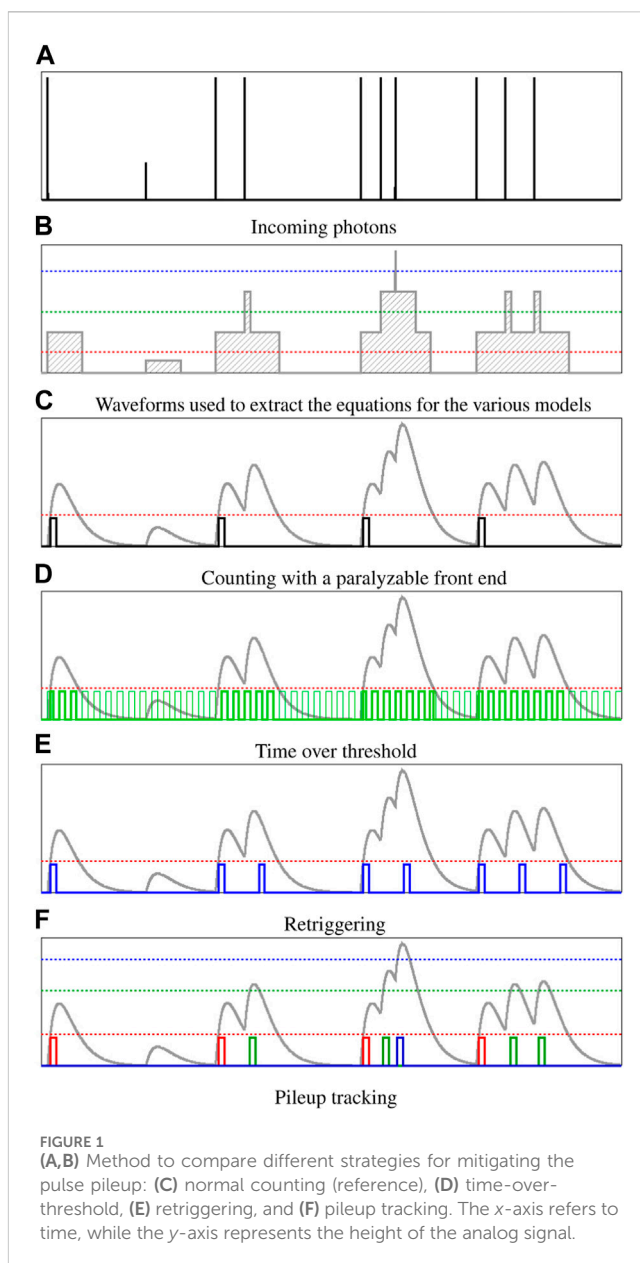
area SPC detector dedicated to macromolecular crystallography [11]. It comprised 18 multi-chip modules for a total of 288 readout chips covering $21 \times 24 \text{ cm}^2$.

Since their introduction, single-photon counting detectors have transformed data collection at synchrotrons. Replacing slower and less-sensitive detectors, they enable, for example, fine slicing of the rotation angle and shutter-less data collection [12; 13; 14] and now underpin many modern measurement techniques like ptychography [15]. Currently, most scattering beamlines at synchrotrons rely on SPC detectors like PILATUS [16; 17], EIGER [18; 19], XPAD3S [20], UFXC32k [21; 22], MEDIPIX [23; 24; 25], and MYTHEN [26]. SPC detectors can also be used for electron detection, for example, in low-energy electron microscopy and photoemission electron microscopy [27].

A pixel in a typical single-photon counting detector consists of a charge-sensitive preamplifier, shaper, comparator, and counter. When the analog signal exceeds a certain threshold, it is counted as a photon. If the threshold is high enough compared to the electronic noise (at least five times the RMS) and the photon energy is similarly higher than the threshold, it is possible to detect photons with a high efficiency and extremely good noise rejection [28]. Given that when a photon is absorbed between two pixels, its charge is shared between them, the threshold should be set at half of the photon energy in order to maximize the number of detected X-rays while avoiding double counts [16]. The electronic noise defines the minimum detectable energy at about 10 times the RMS of the electronic noise [26].

Another important parameter is the threshold dispersion, i.e., the accuracy of tuning the threshold at the same energy level for all the channels. Threshold equalization methods must be implemented to compensate for mismatches between channels in the analog chain, affecting mainly the gain and the baseline level. Since the threshold is normally set as a voltage, the threshold equalization circuitry is then implemented by adding an additional *trim* threshold to the global threshold common to all channels of the chip (or of the detector module). The *trim* digital-to-analog converter (DAC) has a resolution ranging between 3 and 6 bits, and usually, the range of the voltage available for trimming can be tuned by using an external voltage, as in [29]. Recently, trimming architectures capable of tuning the gain and offset of the analog chain by acting at the shaper level have been demonstrated by [30]. Threshold equalization strategies can either be implemented by optimizing the count dispersion using a flat illumination, as in [31], or by equalizing the gain at a certain energy, as in [32]. The energy resolution of the detector, i.e., the accuracy to define the threshold to discriminate photons of different energies, is given by the quadratic sum of the electronic noise and of the threshold dispersion [33]. This can be exploited, for example, in the case of fluorescence radiation emitted by the sample, which can be rejected by setting the threshold between the main beam energy and the fluorescence line.

Since SPC detectors do not provide any information about the energy of the photons, their performance is not ideal in case of a polychromatic radiation spectrum. However, also in this case, by assigning the same weight to each detected photon, their performance outdoes charge-integrating detectors, where the weight of the photons is proportional to their energy, resulting in a reduction of the image contrast [34] or in stronger high harmonic contamination in the case of diffraction applications.



The readout of the detector is completely digital, and it does not add any noise to the data. Therefore, in ideal conditions, SPC detectors provide *noiseless* data, where the image quality is only limited by the Poisson fluctuations of the number of incident photons. SPC detectors can provide perfect linearity and virtually infinite dynamic range, only limited by the exposure time since it is possible to sum frames without adding noise. Lower bit depths or partial readout of the counters allow extremely high frame rates, for example, 20 kfps for EIGER [35] and 56 kfps for UFXC32k [21], while data compression can help achieve frame rates higher than 100 kHz [36]. A further improvement in speed can be achieved by defining a region of interest [37].

Using an electronic shutter, SPC detectors can be gated with a time resolution usually of a few tens of ns. This is of particular interest for pump-probe experiments since the signal from multiple probes can be accumulated at high frequencies without the need to

read out the detector in between the gates. With the hybrid fill pattern of synchrotrons, it is possible to gate the isolated bunch and obtain a time resolution limited by the duration of the bunch and by its jitter [21; 38].

Despite their success, single-photon counters still suffer from one major inherent weakness since the pulse processing front-end makes them susceptible to a loss of efficiency due to the pileup (Figure 1C). For this reason, SPC detectors require significant improvements in order to best use the increased brilliance of the fourth-generation synchrotron sources. In the following section, we will discuss the main ideas behind the development of the next-generation SPC detectors and look at their expected performance compared to charge-integrating detectors, initially developed for X-ray free-electron lasers (XFELs) but increasingly also used at synchrotrons. We will investigate the count rate capability (i.e., the linearity of the detector as a function of the number of photons per second), the spatial resolution, and strategies for extending the detectable energy range toward both lower- and higher-energy photons using advanced sensors.

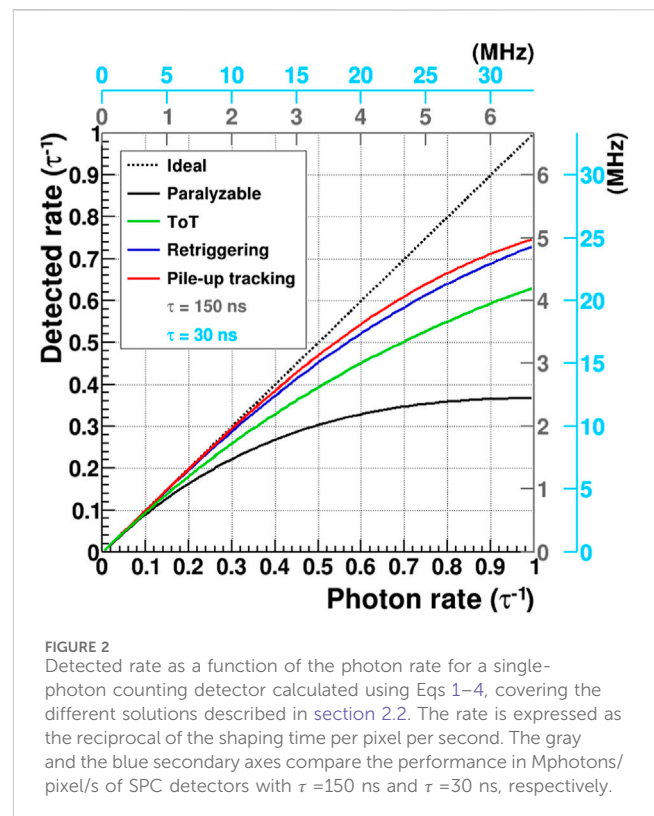
2 Discussion

The readout electronics of hybrid pixel detectors benefits from the advances in CMOS technology. By exploiting more advanced, smaller technology nodes with a higher transistor density, it is possible to integrate more functionalities in the pixel and target smaller pixel pitches, still with an acceptable power consumption. Moreover, by implementing advanced logic on-chip, taking advantage of the synthesized digital circuitry, the fully digital data can be processed in hardware to reduce the data throughput or speed up the readout.

In this section, we discuss the major trends in SPC detector development toward diffraction-limited light sources, namely, the possibility of having multiple comparators and counters per pixel with independent thresholds and enabling signals; the development of methods to reduce the loss of efficiency at high count rates; the goal of achieving a high spatial resolution either using smaller pixels or exploiting advanced inter-communication between pixels; and the possibility to combine SPC detectors with novel sensor technologies in order to cover an energy range spanning from soft to hard X-rays. Finally, we compare SPC with charge-integrating detectors, highlighting the strengths and weaknesses of both readout modes, and discuss the challenge of handling the data produced by fast large-area detectors.

2.1 Multiple comparators and counters

Thanks to the miniaturization of electronic components using advanced CMOS technologies, several comparators, each with an individual threshold, trimbits, and counter with independent gates can be allocated in the same pixel. The comparators and counters can be connected using the logic of varying complexity, and even inter-pixel communication can be implemented. The independent thresholds give access to different energy bins, the independent gates to different time windows.



The main application exploiting the independent thresholds is energy binning [39]. It can be used at polychromatic X-ray sources and is exploited in clinical CT systems [40]. However, energy binning can also be used at synchrotron beamlines for high harmonic and fluorescence suppression or detection. As long as the spectroscopic capabilities do not degrade at high fluxes, the energy discrimination capabilities can be used to operate in the pink beam mode, i.e., using a larger bandwidth without monochromatizing optics, to isolate the full undulator harmonic. The energy resolution depends not only on the noise and on the threshold dispersion but also on charge sharing. Therefore, larger pixel sizes and charge-sharing suppression methods (see section 2.3) are an advantage for energy binning. The multiple thresholds can also be exploited to improve the count rate capability, as described in section 2.2.3. Complex digital circuitry and inter-pixel communication can be used, for example, for interpolation to achieve a sub-pixel resolution, as highlighted in section 2.3.

The independent gates enabling the counters can be used to perform measurements in different time windows. Given that the reaction being studied happens on the same time scale, or slower, than the shortest gate that can be applied (normally tens of nanoseconds), it is possible to probe at multiple times and thus reduce the duration of the experiment proportionally to the number of counters compared to a single probe. Moreover, it is possible to acquire pumped and unpumped data alternating in time and, therefore, correct for possible low-frequency drifts of the system [41]. Ideally, a single comparator is connected to the multiple independent counters to minimize mismatches between different probes since the changes in the sample can be subtle ($\sim 1\%$ or less).

2.2 Count rate capability

Single-photon counting detectors present a paralyzable behavior due to the pileup of the signal of photons arriving close to each other in time. This means that after detecting a photon, the detector is insensitive for a defined time (dead time), and a photon detected during this interval will not be counted and will also restart the dead time. As a consequence, with the increasing rate, the detector will reach a saturation point where it will be incapable of recording any event at all. To fully capitalize on the increased brilliance of the fourth-generation light sources, count rate capabilities up to 100 Mphotons/pixel/second are required. In the following section, we discuss and compare the new approaches that have been proposed to extend the usability of SPC detectors at next-generation synchrotron sources. They rely on the idea that when two or more photons pile up, the analog signal has a larger amplitude and longer duration. The working principle of various methods is shown in Figure 1, while Figure 2 shows the comparison of their performance in terms of count rate capability. The secondary axes quantify the performance in terms of impinging photon rate for 150-ns and 30-ns shaping times.

Usually, the count rate capability of an SPC detector is modeled according to the parameter τ , which is proportional to the shaping time of the analog chain. It scales inversely with the shaping time of the analog signal, which is usually a compromise between the noise and gain of the detector. Therefore, the count rate performance tends to improve at high energies, where a lower gain can be used, and a loss of signal due to ballistic deficit can also be afforded. Another approach to obtain a fast shaping is active reset, as demonstrated by [42]. Faster shaping time and active reset cause a high power consumption, which is sometimes unacceptable for large readout chips with tens of thousands of pixels. Current single-photon counters exhibit a τ of 30–150 ns, which translates to approximately 2–10 Mcounts/pixel/s. The KITE ASIC from DECTRIS has been designed for electron microscopy with extremely short signal pulses of 6 ns FWHM [36]. This is possible thanks to the fast settings and low gain allowed by the extremely large signal generated by high-energy electrons, as well as the small size with no buttability, which allows high power consumption and optimal power distribution.

It is important to highlight that the count rate capability of SPC detectors depends on the photon distribution, i.e., on the filling pattern of the light source [43]. The actual time structure varies between synchrotrons and can, in many cases, be tuned, from a few isolated bunches to the quasi continuous mode, in order to optimize for certain experiments. Filling patterns with fewer bunches spaced more than the shaping time of the analog signal usually allow better performance, the rate correction present requires a simpler calibration of the parameters, and the methods explained in the following are more effective since they are less subjected to statistical fluctuations in the photon time distribution. However, in this case, the equations reported for the count rate corrections must be modified due to the different photon distribution.

For modeling the count-rate corrections presented in this section, a Poisson-like photon distribution is assumed, with no charge sharing and τ -wide rectangular pulse shape, where τ approximates the time-over-threshold of the analog pulse, as shown in Figure 1B. For a simple SPC detector, the paralyzable

detector model applies for converting from the impinging photon rate ϕ to the measured flux φ [44]:

$$\varphi = \phi e^{-\phi\tau} \quad (1)$$

Despite being a simplified description of the pileup, this approach allows to model analytically the corrections needed to convert the detected signal into the impinging photon rate. The real corrections will depend also on other factors, such as the fill pattern of the photon source, charge sharing, and—in particular—on the shape of the analog signal, which should be optimized depending on the strategy chosen to improve the count rate capability.

2.2.1 Time-over-threshold

The time-over-threshold (ToT) readout measures the time duration of the analog signal above the threshold by providing a clock to each pixel and incrementing the counter while the signal is above the threshold (Figure 1D). This method is usually applied to measure the energy of the impinging particle, as in the TIMEPIX detectors [45]. However, to improve the energy resolution, slow shaping times are usually implemented, in contrast with the requirements to obtain a high count rate capability in SPC detectors. ToT, as a means to improve the count rate capability at high fluxes, has been demonstrated with an improvement of a factor 3–6, depending on the settings [46]. The count-rate correction equation is expressed as follows:

$$\varphi = \phi \frac{1 - e^{-\phi\tau}}{\phi\tau} \quad (2)$$

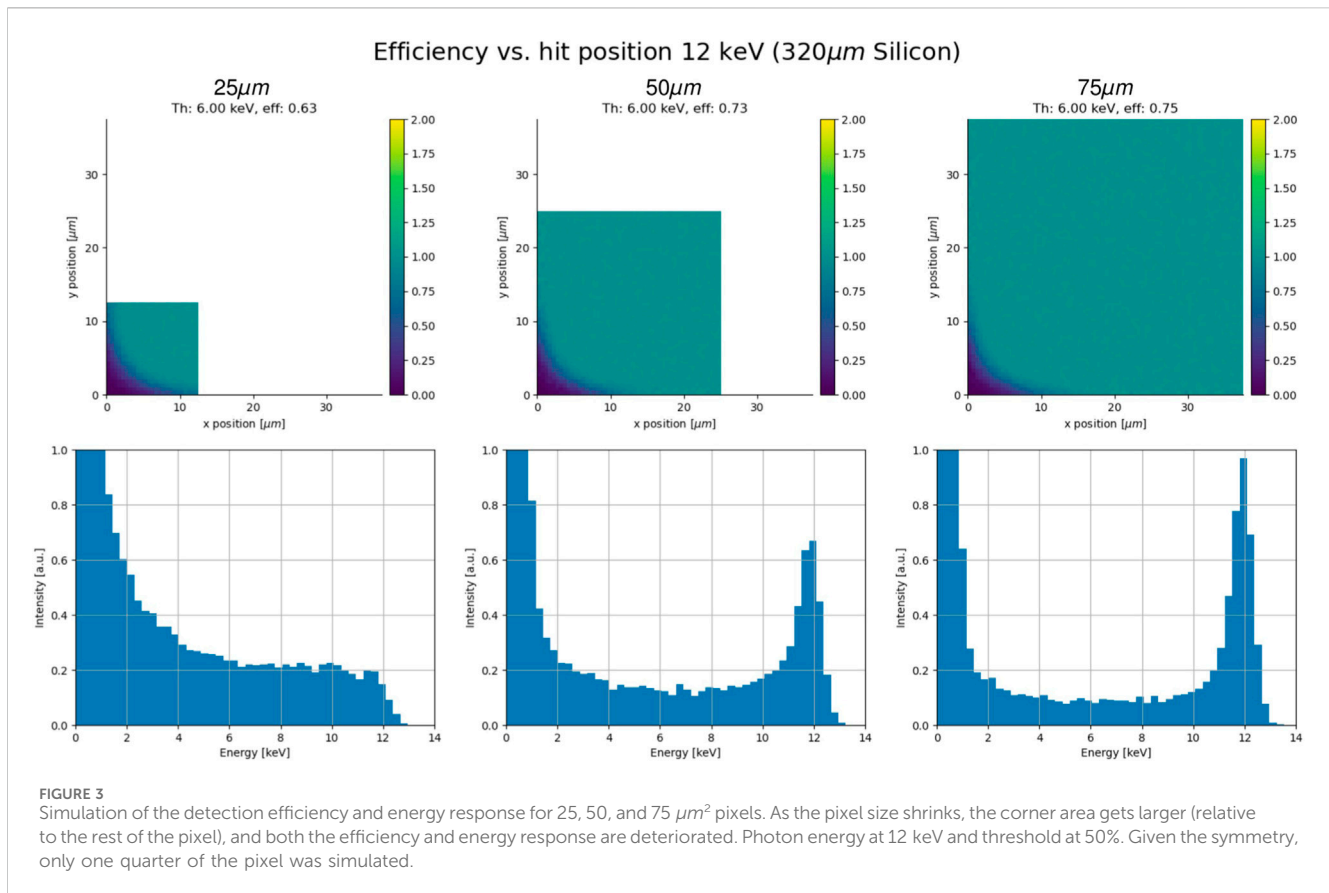
with little dependence on the frequency of the clock, as long as it is $\geq 5\tau^{-1}$. It is important to point out that in order to obtain the number of photons, it is necessary to normalize the counter value by the average number of counts per photon of that energy and threshold. The main disadvantage of the ToT approach is the distribution of the 10–100-MHz clock over the whole pixel matrix, which requires power and can generate the digital-to-analog crosstalk, increasing the noise. Optimized clock distribution solutions, such as the digital delay-locked loop (dDLL) [47], can help overcome these bottlenecks.

2.2.2 Retriggering

Retriggering was introduced using DECTRIS in the PILATUS3 detector [17] as a measure to increase the count rate capability and avoid the ambiguities of a paralyzable counter at extremely high photon rates. As the name implies, retriggering works by triggering an additional count after a certain time delay. The delay is started on the crossing of the threshold which gives retriggering a clear advantage over the non-synchronized ToT clock. Retriggering relies on knowing the pulse width for a single photon. The rate correction model is described in [48]:

$$\varphi = \frac{\phi}{e^{-\phi\tau_r} + \phi\tau_r} \quad (3)$$

where τ_r is the time after which the retrigger is evaluated and an additional signal eventually generated. Ideally $\tau_r = \tau$, but τ_r is usually larger to avoid double counts. The KITE ASIC from DECTRIS has been designed for electron detection, and it can support up to ~ 70 Mcounts/pixel/s with a 43-keV



threshold for 200-keV electrons with a 10% loss of counting efficiency compared to approximately 20 Mcounts/pixel/s without retriggering [36].

2.2.3 Pileup tracking

Given a sufficient dynamic range of the front-end, another way to mitigate the loss of efficiency is to use multiple thresholds to count the pileup by placing the additional thresholds above the photon energy (e.g., 1.3 and 1.7 $\times E_{\text{photon}}$, as shown in Figure 1F). This method is known as the *pileup trigger method* [49; 50] or *pileup tracking* [51]. Finding the right threshold depends on the analog shape of the signal and sensor geometry (charge sharing). Ideally, also, the higher thresholds are proportional to the energy, similar to the one at half energy. The count rate correction model can be found in [52]:

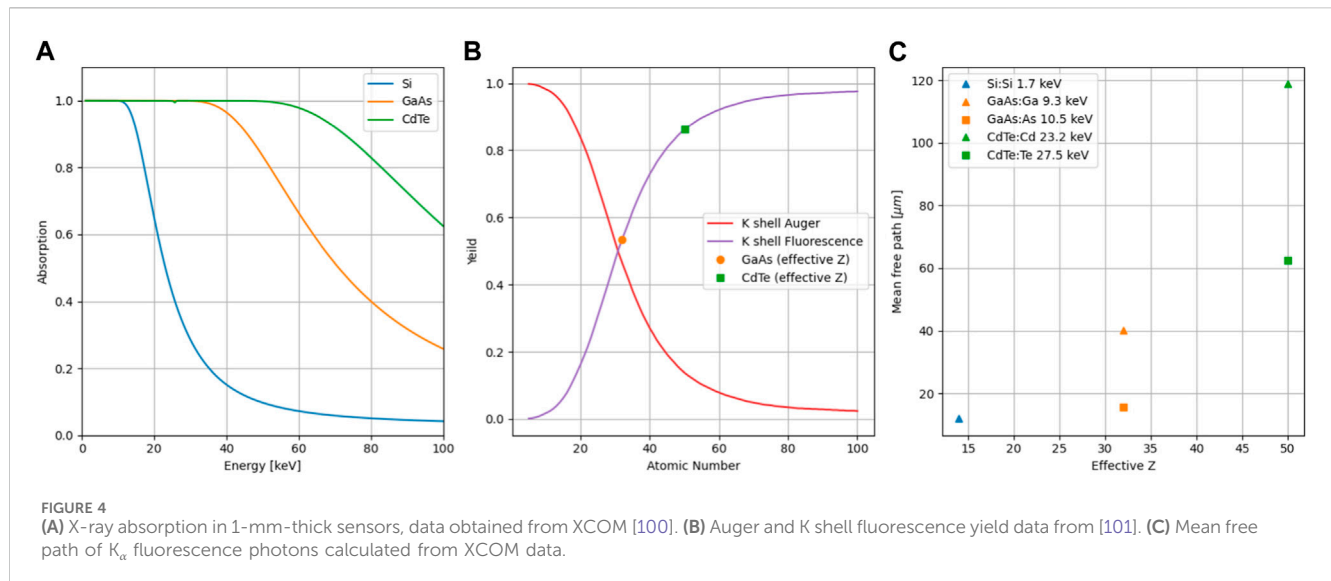
$$\varphi = \phi \sum_{i=0}^N (1 - e^{-\phi\tau})^i \quad (4)$$

where N is the number of comparators and counters. The improvement in the count rate capability is very pronounced up to three counters. For additional counters, however, there are diminishing returns, and one also has to consider issues like saturation of the preamplifier and available space in the pixel. The plot shown in Figure 2 refers to three counters. With MYTHEN3 we see an improvement of 4–6x in the count rate at 90% efficiency with three counters. For the new MATTERHORN detector, which is currently at the prototyping stage at PSI, we plan to use this approach with four counters.

2.3 Spatial resolution

Reducing the pixel size is an effective way to limit the incoming photon rate per pixel and improve the spatial resolution but only up to a certain extent. Single-photon counting detectors with relatively large pixels show an ideal MTF [53], when operated with the threshold at half the photon energy. As the pixel size shrinks, the charge cloud, which—in a typical 300- μ m-thick silicon sensor—is on the order of 10–20 μ m [54], becomes comparable to the pixel size, leading to worse energy response and a loss of detection efficiency in the corners of the pixel. Figure 3 shows a simulation of the energy response and detection efficiency as a function of the interaction position in the pixel for 25, 50, and 75 μ m² pixels performed using GEANT4 [55] with custom drift-diffusion [56] implementation. As the pixel size becomes smaller, the corner area, where the charge generated by a photon is shared between four pixels, gets larger relative to the rest of the pixel, and both energy response and detection efficiency are degraded. Photons absorbed close to the pixel corner, whose signal does not exceed the 50% energy threshold in any of the four neighboring pixels, will not be detected. For pixels smaller than 30–40 μ m, it is not possible to work in the single-photon counting mode without a solution to compensate for charge sharing. Due to the absence of a threshold, charge-integrating detectors do not suffer from the corner effect, giving a flat response throughout the pixel [57].

For high-Z sensor materials (section 2.4.1), reducing the pixel size is even more difficult since the often thicker sensors provide a larger charge cloud, and in addition, X-ray fluorescence in the



material itself further spreads the charge. The fluorescence yield for CdTe is above 80%, and the mean free paths of the characteristic photons are 111 and 58 μm [58]. This leads to distortions in the recorded energy spectrum and additional count rate load since, in many cases, the fluorescence photon is counted separately. Using a charge summing architecture, many of these disadvantages can be overcome [59; 60] but at the cost of the count rate capability [61]. It is possible to handle the corner effect digitally by using multiple comparators and evaluating the coincidence between neighboring pixels. This approach affects the count rate capability similarly to charge summation, but it does not affect the performance of the analog chain.

The best possible spatial resolution of hybrid detectors is reached using a charge-integrating detector under sparse illumination. In this case, charge sharing can be exploited to interpolate between neighboring pixels, reaching a spatial resolution in the micrometer range [62]. For single-photon counting detectors, processing has to happen in the pixel (or strip) since later, no energy information is available. With the MYTHEN3 SPC microstrip detector, it has been demonstrated in 1D that it is possible to use the digital circuitry of a single-photon counting detector with multiple comparators to obtain a spatial resolution better than the physical strip pitch [63]. Similar approaches for pixels are being evaluated, with the additional challenge of a more complex 2D coincidence logic and the need to use a very advanced technology node to shrink the whole circuitry within the necessarily small pixel size [64].

2.4 Advanced sensors

The pulse processing front-end in single-photon counting is more forgiving in terms of leakage current than a charge-integrating front-end and allows the use of sensors with a high and/or variable leakage current since the fluctuations are “filtered” out. This is helpful for detection of both low- and high-energy photons since low-gain avalanche diodes (LGADs) and many high-Z sensors exhibit this behavior. However, to allow full flexibility on the

sensor choice, the analog chain of the SPC detector must be designed to work both for hole collection, which is the standard for silicon and inverse LGADs, and electron collection, which is usually required for high-Z sensors and standard LGAD technologies.

2.4.1 High-Z materials

Since silicon becomes almost transparent above 20 keV (see Figure 4A), there is a strong need for sensor materials with higher atomic numbers (i.e., high-Z sensor materials) and thus increased photon cross section. Moving away from silicon, we are faced with intrinsic problems like increased X-ray fluorescence yield and range in the sensor material (Figures 4B,C) but also material defects coming from the fact that it is harder to grow high-quality crystals of compound semiconductors which additionally have not benefited from the massive investment from the electronics industry like silicon.

Over time, GaAs, CdTe, and CZT emerged as the most promising materials, and specifically, CdTe is applicable both for medical imaging [65; 66] and at synchrotrons (among others [67; 68; 69]). Larger-area detectors have also started to appear with, for example, the 16-M CdTe EIGER2X from DECTRIS at the P14 beamline at PETRA III. There was some interest in using germanium for SPC detectors [70], but due to the small band gap, they need to be cooled to $\lesssim -100^\circ\text{C}$, complicating the operation. As sensor materials have improved, the focus has shifted from understanding defects like tellurium inclusions [71] to optimizing for high flux and understanding dynamic effects [72; 73]. [74] and [75] offered a good overview of common sensor materials and their use within the Medipix community. At PSI, we have, for example, studied GaAs with JUNGFRÄU, probing an effective pixel size and understanding the negative signals observed when used with charge-integrating detectors [76].

In terms of new materials, perovskites [77] have attracted a lot of attention, showing a combination of high atomic number and good mobility-lifetime product. The production cost could be orders of magnitude lower than for CZT/CdTe due to cheap base material and a simpler manufacturing procedure. Although some experiments

have been done with photon counting [78], we are still far from observing large-area perovskite detectors at synchrotrons. One potential issue with the current materials is the relatively low mobility [79], which could cause problems in terms of ballistic deficit at the fast shaping times needed for high count rates.

Looking at high-Z sensor material from the perspective of the readout ASIC, we observe a need for electron collection since most high-Z sensor materials have better transport properties for electrons than holes. Leakage current compensation is also important, given the higher leakage currents in most materials, for example, > 200 pA per $75 \times 75 \mu\text{m}^2$ pixel at room temperature using GaAs, as shown in [80]. The pixel size is also limited by the long range of X-ray fluorescence in the sensor layer, for example, $110 \mu\text{m}$ in CdTe, requiring relatively large pixels even with interpixel communication for an optimal response. If the pixel size goes much below $75 \mu\text{m}$, one has to consider events where the fluorescent photon deposits the energy in a pixel that could be several pixels away from the initial interaction.

2.4.2 LGADs

Soft X-ray detection using hybrid detector technology is challenged by the shallow absorption and by the small charge generated by low-energy photons, resulting in low quantum efficiency and low signal-to-noise ratio, respectively. In particular, while charge-integrating detectors can be used for soft X-ray detection even without reaching single-photon resolution, SPC detectors require a $\text{SNR} \geq 10$ for single photons, and they are consequently limited to energies above ~ 2 keV when using standard silicon sensors [28]. The quantum efficiency below 2 keV can be improved by optimizing the entrance window of the silicon sensor, obtaining a performance comparable to state-of-the-art CCDs or CMOS imagers [81].

The signal-to-noise ratio can be improved by exploiting the internal amplification of LGAD sensors recently developed for high-energy physics application. They consist of a silicon sensor with a highly doped p-n junction, where charge carriers are multiplied by impact ionization, thanks to the high electric field. The main goal in particle physics is to use the fast avalanche to improve the timing performance and use the time of arrival of the particle for 4D tracking at high-luminosity colliders [82]. However, the requirements for photon science require additional developments, including small pixels $\leq 100 \mu\text{m}$ and full sensitivity at the entrance window, in contrast to HEP detectors, with large pads and a thin sensitive layer on a thick substrate. Various LGAD fabrication technologies are described in detail in [83]. A few feasibility studies have been dedicated to X-ray detection [84; 85; 51], aimed at reducing the effective noise, thanks to the multiplication gain, while maximizing the fill factor, which is limited in most LGAD technologies due to the presence of regions without multiplication between the pixels. The performance of LGAD sensors combined with SPC readout electronics does not suffer much from the high leakage current due to the internal amplification, which, however, limits the maximum exposure time acceptable for charge-integrating detectors with LGAD sensors. Recently, the inverse LGAD sensors with optimized entrance window developed by [86] have been combined with the EIGER single-photon counting detector, allowing, for the first time, a SPC pixel detector to reach energies below 1 keV. The low

noise and high dynamic range allowed achieving unprecedented data quality in magnetic contrast soft X-ray ptychography, as described by [87].

LGADs can also be exploited for tender and hard X-ray detection to improve the count rate capability using a faster shaping of the analog signal since gain is provided in the sensor, and therefore, the signal can be sacrificed using fast analog settings and accepting a more ballistic deficit.

2.5 Single-photon counting *versus* charge integrating

With the introduction of charge-integrating detectors with dynamic gain switching (like JUNGFRU Mozzanica et al. [88]), it is possible to measure with an electronic noise below the Poisson limit from a single photon throughout the full dynamic range (10^4 12 keV photons/pixel). The maximum 2 kHz frame rate of JUNGFRU translates to a count rate capability of approximately 20 Mphotons/pixel/s at 12 keV. Since it is an integrating detector, the maximum supported flux scales inversely with the photon energy, providing even higher numbers at lower energies (e.g., 10^5 photons/pixel/frame and 200 Mphotons/pixel/s at 1.2 keV). Different from SPC detectors, JUNGFRU is linear throughout the dynamic range independent of the photon rate (e.g., 40 Mphotons/pixel/s at 6 keV). [89] showed that data quality remained high even at full beamline transmission (thaumatin crystal at 6 keV) in contrast to previously published results with an EIGER 1-M detector [14].

For a charge-integrating detector, the highest continuous flux is determined by the dynamic range multiplied by the frame rate. We are currently developing the second-generation JUNGFRU detector with a target frame rate of 10 kHz, which, given the same dynamic range, could cope with 100 Mphotons/pixel/s at 12 keV. This would provide a solution for the most extreme fluxes at the cost of dealing with 12.5 GB/s per 500 kpixel module. However, since the performance of charge-integrating detectors degrades above a few hundred microseconds integration time (higher noise and lower dynamic range), longer exposure times must be achieved by summing up multiple images, with a consequent increase in the electronic noise, which scales with the square root of the number of frames.

At the other extremes, charge-integrating detectors have their place in photon-starved applications since with a low enough flux, it is possible to measure the charge deposited per pixel per photon, enabling both spectroscopic measurements and interpolation (Bergamaschi et al. [90]). However, in this application, the maximum supported flux is limited by the frame rate of the detector, which is typically of the order of a few kHz with one outstanding exception being the HEXITEC_{MHz} [91]. For the same applications, the implementation of energy binning or a sub-pixel resolution in a single-photon counting chip could increase the maximum supported flux of up to three orders of magnitude since it would be limited by the ~ 100 ns shaping time rather than by the $\sim 500 \mu\text{s}$ readout time [63].

An ideal detector should perform like a single-photon counter under low illumination (noiseless, stable, and low frame rate possible) but be capable of supporting high fluxes when

necessary. Charge-integrating detectors with a single-photon resolution converting the analog data to the number of photons on-the-fly and summing up frames to achieve long exposure times promise to achieve this goal, as shown in Leonarski et al. [92], but compared to SPC detectors, they require more developments in terms of chip design, firmware development, and data backend. Moreover, the integration of the leakage current requires the acquisition of frequent dark images to follow possible thermal drifts and the effect of radiation damage, which are usually filtered out in SPC detectors.

2.5.1 Mixed mode

An approach for obtaining ideal behavior from a large detector is to equip the more illuminated area of the detector (e.g., the central region for coherent scattering) with charge-integrating detector modules and a high-performance data back-end system and substitute single-photon counting modules in regions with lower illumination, synchronizing the acquisition of all the modules. Some challenges and artefacts might still arise when combining data from such different detector systems.

Charge removal architectures are a hybrid between single-photon counting and charge-integrating detectors [93; 94]. They have a fully digital readout, and ideally, they can provide a high gain (and low noise) on the whole virtually infinite dynamic range. High-flux detection is possible, only limited by the speed of charge removal (e.g., they are not usable at XFELs), while a slow frame rate readout is still feasible, with the disadvantage of large pedestal corrections due to the integration of the leakage current. Charge removal architectures might suffer from high noise levels at high intensities due to charge injection fluctuations during charge removal.

The possibility to statically configure the detector in the counting or integrating mode depending on the application (or on the illumination level) would add great flexibility, and users could benefit from the advantages of two detector systems, without the need to change the detector. This is possible by adding to the feedback circuit of the preamplifier of a SPC detector a reset switch. When the reset is open and the field-effect transistor (FET) controlling the feedback resistor is active, the detector can be operated in the SPC mode, feeding the preamplifier output to an additional shaper and comparator [95]. When the FET of the feedback resistor is off, the reset can be operated to integrate the charge on the feedback capacitor during the given exposure time and then sample on a storage capacitance. The comparator used in the SPC mode could be used in order to implement dynamic gain switching, while the shaper could be operated similar to the preamplifier for correlated double sampling (CDS). Such an architecture has been implemented and successfully tested in a MYTHEN3 microstrip prototype and will be adapted to pixel detectors. Ideally, the detector should dynamically adapt its behavior depending on the illumination, similar to dynamic gain switching, but this comes with additional challenges in the conversion of the signal into number of photons.

2.6 More photons, more data

Single-photon counting detectors present some advantages from the data handling point of view compared to charge-integrating

detectors since the detector readout already consists of the number of detected photons, without the need for additional processing. Moreover, the fully digital readout simplifies the data compression (e.g., zero suppression), making it possible already on the readout board or even on-chip.

Still, the higher frame rates driven by the increased brilliance will be a huge challenge for the beam lines. Already the previous generation detectors kicked off the “data deluge” [96], and with new detectors coming out with unprecedented data rates like MATTERHORN with 100 Gbit/s per 500k pixel module or TIMEPIX4 [97] with 16×10 Gbit/s serializers per chip (448 Å— 512 pixels) saving (or even receiving), all raw data on the computing infrastructure will not be possible. The LEAPS¹ data strategy [98] outlines the importance of involving central IT services to take some of the burden off experimental groups which might not be used for large datasets, but with the explosion of data, we believe that measures also need to be taken to reduce data at the source.

We see two different approaches, of which one is to build a custom receiving system using FPGAs and GPUs to perform local processing like in [92] before streaming out the data, and the other is to move data reduction onto the readout board or even into the ASIC itself. Reducing the data closer to where it is produced has the added benefit that it lowers the demands on subsequent network and computing infrastructure at the cost of flexibility. For now, it remains an open but highly important field, and we need to see what works best in practice.

3 Conclusion

SPC detectors have been the most used position-sensitive detectors at synchrotron facilities for more than a decade, but the higher brilliance of the fourth-generation light sources requires a new generation with improved performance. There are several single-photon counting detectors under development which will be used at diffraction-limited sources including SPHIRD [94], MEDIPIX4 [99], and our own MATTERHORN. All these projects aim to significantly improve their count rate capability compared to existing systems, and in section 2.2, we presented some of the solutions which are being investigated to achieve this goal. Of the compared methods, pileup tracking shows potential to give the highest count rate, but one also has to consider other features of the detector like calibration, correction of mismatches between channels, spectral response, and radiation hardness. Therefore, the best solution will be the one which provides the most accurate data, something that could vary between applications and facilities.

Additional features like the possibility of using multiple comparators and counters and extending the usable energy range from soft X-rays to high-energy photons will open new possibilities for SPC detectors and can eventually pioneer novel experimental

¹ League of European Accelerator based Photon Sources <https://leaps-initiative.eu/>

techniques. Despite not being the optimal solution for targeting high-resolution imaging applications, SPC hybrid detectors with intercommunication between pixels at the analog or at the digital level can further push the spatial resolution or at least improve some of the current flaws of SPC detectors with small pixels, like the corner effect.

Despite the challenge brought by the increased flux, we think that SPC detectors will have a bright future at diffraction-limited light sources and remain the workhorse detector in the near to mid-term future, thanks to their reliability and ease of use. For the highest rate applications, charge-integrating detectors will complement photon counting, until the mixed mode readout combining the advantage of both architecture and other novel readout methods will be available.

Author contributions

EF: conceptualization and writing—original draft. AB: conceptualization and writing—original draft. BS: writing—review and editing.

References

- Heijne E, Jaron P, Olsen A, Redaelli N. The silicon micropattern detector: a dream? *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1988) 273:615–9. doi:10.1016/0168-9002(88)90065-4
- Anghinolfi F, Aspell P, Bass K, Beusch W, Bosio L, Boutonnet C, et al. A 1006 element hybrid silicon pixel detector with strobed binary output. *IEEE Trans Nucl Sci* (1992) 39:654–61. doi:10.1109/23.159682
- Alexeev G, Andersen E, Andrichetto A, Antinori F, Armenise N, Ban J, et al. First results from the 1994 lead beam run of wa97. *Nucl Phys A* (1995) 590:139–46. doi:10.1016/0375-9474(95)00232-2
- Becks K, Borghi P, Brunet J, Caccia M, Clemens J, Cohen-Solal M, et al. The delphi pixels. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1997) 386:11–7. doi:10.1016/S0168-9002(96)01089-3
- Heijne E, Antinori F, Beker H, Batignani G, Beusch W, Bonvicini V, et al. Development of silicon micropattern pixel detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1994) 348:399–408. doi:10.1016/0168-9002(94)90768-4
- Da Via C, Bates R, Bertolucci E, Bottigli U, Campbell M, Chesi E, et al. Gallium arsenide pixel detectors for medical imaging. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1997) 395:148–51. doi:10.1016/S0168-9002(97)00631-1
- Campbell M, Heijne E, Meddler G, Pernigotti E, Snoeys W. A readout chip for a 64/spl times/64 pixel matrix with 15-bit single photon counting. *IEEE Trans Nucl Sci* (1998) 45:751–3. doi:10.1109/23.682629
- Delpierre P, Berar J, Blanquart L, Caillot B, Clemens J, Mouget C. X-ray pixel detector for crystallography. *IEEE Trans Nucl Sci* (2001) 48:987–91. doi:10.1109/23.958710
- Brönnimann C, Baur R, Eikenberry E, Kohout S, Lindner M, Schmitt B, et al. A pixel read-out chip for the pilatus project. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2001) 465:235–9. doi:10.1016/S0168-9002(01)00396-5
- Kästli H, Barbero M, Erdmann W, Hörmann C, Horisberger R, Kotlinski D, et al. Design and performance of the cms pixel detector readout chip. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2006) 565:188–94. doi:10.1016/j.nima.2006.05.038
- Brönnimann C, Eikenberry EF, Henrich B, Horisberger R, Huelsen G, Pohl E, et al. The PILATUS 1M detector. *J Synchrotron Radiat* (2006) 13:120–30. doi:10.1107/S0909049505038665
- Pflugrath JW. The finer things in x-ray diffraction data collection. *Acta Crystallogr Section D* (1999) 55:1718–25. doi:10.1107/s090744499900935x
- Förster A, Brandstetter S, Schulze-Briesse C. Transforming x-ray detection with hybrid photon counting detectors. *Philos Trans R Soc A* (2019) 377:20180241. doi:10.1098/rsta.2018.0241
- Casas A, Warshamane R, Finke AD, Panepucci E, Olieric V, Nöll A, et al. Eiger detector: application in macromolecular crystallography. *Acta Crystallogr* (2016) D72:1036–48. doi:10.1107/s2059798316012304
- Pfeiffer F. X-ray ptychography. *Nat Photon* (2018) 12:9–17. doi:10.1038/s41566-017-0072-5
- Kraft P, Bergamaschi A, Broennimann C, Dinapoli R, Eikenberry EF, Henrich B, et al. Performance of single-photon-counting PILATUS detector modules. *J Synchrotron Radiat* (2009) 16:368–75. doi:10.1107/S0909049509000911
- Loeliger T, Brönnimann C, Donath T, Schneebeli M, Schnyder R, Trüb P. The new pilatus3 asic with instant retrigger capability. In: Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC); November 2012; Anaheim, CA, USA (2012). p. 610–5. doi:10.1109/NSSMIC.2012.6551180
- Johnson I, Bergamaschi A, Buitenhuis J, Dinapoli R, Greiffenberg D, Henrich B, et al. Capturing dynamics with Eiger, a fast-framing X-ray detector. *J Synchrotron Radiat* (2012) 19:1001–5. doi:10.1107/S0909049512035972
- Donath T, Šišak Jung D, Burian M, Radicci V, Zamboni P, Fitch AN, et al. EIGER2 hybrid-photon-counting X-ray detectors for advanced synchrotron diffraction experiments. *J Synchrotron Radiat* (2023) 30:723–38. doi:10.1107/S160057752300454X
- Medjoubi K, Bucaille T, Hustache S, Bérar J-F, Boudet N, Clemens J-C, et al. Detective quantum efficiency, modulation transfer function and energy resolution comparison between CdTe and silicon sensors bump-bonded to XPAD3S. *J Synchrotron Radiat* (2010) 17:486–95. doi:10.1107/S0909049510013257
- Bachiller-Perea D, Abiven Y-M, Bisou J, Fertey P, Grybos P, Jarnac A, et al. First pump-probe-hard X-ray diffraction experiments with a 2D hybrid pixel detector developed at the SOLEIL synchrotron. *J Synchrotron Radiat* (2020) 27:340–50. doi:10.1107/S1600577520000612
- Nakaye Y, Sakumura T, Sakuma Y, Mikusu S, Dawiec A, Orsini F, et al. Characterization and performance evaluation of the XSPA-500k detector using synchrotron radiation. *J Synchrotron Radiat* (2021) 28:439–47. doi:10.1107/S1600577520016665
- Ponchut C, Rigal JM, Clément J, Papillon E, Homs A, Petitdemange S. Maxipix, a fast readout photon-counting x-ray area detector for synchrotron applications. *J Instrumentation* (2011) 6:C01069. doi:10.1088/1748-0221/6/01/C01069
- Pennicard D, Lange S, Smoljanin S, Becker J, Hirsemann H, Eppel M, et al. Development of lambda: large area medipix-based detector array. *J Instrumentation* (2011) 6:C11009. doi:10.1088/1748-0221/6/11/C11009
- Tartoni N, Dennis G, Gibbons P, Gimenez E, Horswell I, Marchal J, et al. Excalibur: a three million pixels photon counting area detector for coherent diffraction imaging based on the medipix3 asic. In: Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC); November 2012; Anaheim, CA, USA (2012). p. 530–3. doi:10.1109/NSSMIC.2012.6551164
- Bergamaschi A, Cervellino A, Dinapoli R, Gozzo F, Henrich B, Johnson I, et al. The MYTHEN detector for X-ray powder diffraction experiments at the Swiss Light Source. *J Synchrotron Radiat* (2010) 17:653–68. doi:10.1107/S0909049510026051

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Open access funding was provided by the PSI—Paul Scherrer Institute.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

27. Tinti G, Marchetto H, Vaz CAF, Kleibert A, Andrä M, Barten R, et al. The EIGER detector for low-energy electron microscopy and photoemission electron microscopy. *J Synchrotron Radiat* (2017) 24:963–74. doi:10.1107/S1600577517009109
28. Andrä M, Barten R, Bergamaschi A, Brückner M, Casati N, Cervellino A, et al. Towards mythen iii - prototype characterisation of mythen iii.0.2. *J Instrumentation* (2019) 14:C11028. doi:10.1088/1748-0221/14/11/C11028
29. Kraft P, Bergamaschi A, Bronnimann C, Dinapoli R, Eikenberry EF, Graafsma H, et al. Characterization and calibration of pilatus detectors. *IEEE Trans Nucl Sci* (2009) 56:758–64. doi:10.1109/TNS.2008.2009448
30. Maj P, Grybos P, Kmon P, Szczygiel R. 23552-channel ic for single photon counting pixel detectors with 75 μ m pitch, enc of 89 e rms, 19 e rms offset spread and 3 spread. In: Proceedings of the ESSCIRC 2014 - 40th European Solid State Circuits Conference (ESSCIRC); September 2014; Venice Lido, Italy (2014). p. 147–50. doi:10.1109/ESSCIRC.2014.6942043
31. Ponchut C, Visschers J, Fornaini A, Graafsma H, Maiorino M, Mettievier G, et al. Evaluation of a photon-counting hybrid pixel detector array with a synchrotron x-ray source. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2002) 484:396–406. doi:10.1016/S0168-9002(01)02029-0
32. Niederlöhner D, Bert C, Giersch J, Pfeiffer K-F, Anton G. Threshold characterisation of the medipix1 chip. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 509:138–45. doi:10.1016/S0168-9002(03)01562-6
33. Bergamaschi A, Cervellino A, Dinapoli R, Gozzo F, Henrich B, Johnson I, et al. Photon counting microstrip detector for time resolved powder diffraction experiments. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2009) 604:136–9. doi:10.1016/j.nima.2009.01.092
34. Swank RK. Absorption and noise in x-ray phosphors. *J Appl Phys* (1973) 44:4199–203. doi:10.1063/1.1662918
35. Hocine S, Van Swyngheoven H, Van Petegem S, Chang CST, Maimaitiylili T, Tinti G, et al. Operando x-ray diffraction during laser 3d printing. *Mater Today* (2020) 34:30–40. doi:10.1016/j.mattod.2019.10.001
36. Zambon P, Bottinelli S, Schnyder R, Musarra D, Boye D, Dudina A, et al. Kite: high frame rate, high count rate pixelated electron counting asic for 4d stem applications featuring high-z sensor. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1048:167888. doi:10.1016/j.nima.2022.167888
37. König H-H, Pettersson NH, Durga A, Van Petegem S, Grolimund D, Chuang AC, et al. Solidification modes during additive manufacturing of steel revealed by high-speed x-ray diffraction. *Acta Materialia* (2023) 246:118713. doi:10.1016/j.actamat.2023.118713
38. Burian M, Marmioli B, Radeticchio A, Morello C, Naumenko D, Biasiol G, et al. Picosecond pump-probe X-ray scattering at the Elettra SAXS beamline. *J Synchrotron Radiat* (2020) 27:51–9. doi:10.1107/S1600577519015728
39. Ballabriga R, Alozy J, Campbell M, Frojdh E, Heijne E, Koenig T, et al. Review of hybrid pixel detector readout asics for spectroscopic x-ray imaging. *J Instrumentation* (2016) 11:P01007. doi:10.1088/1748-0221/11/01/P01007
40. Lowe C. *WhitePaper_Draft10.pdf* (2021). doi:10.6084/m9.figshare.13541051.v1
41. Naumenko D, Burian M, Marmioli B, Haider R, Radeticchio A, Wagner L, et al. Implication of the double-gating mode in a hybrid photon counting detector for measurements of transient heat conduction in GaAs/AlAs superlattice structures. *J Appl Crystallogr* (2023) 56:961–6. doi:10.1107/S1600576723004302
42. Grybos P, Kleczek R, Kmon P, Otfinowski P, Fajardo P, Magalhães D, et al. Sphird-single photon counting pixel readout asic with pulse pile-up compensation methods. *IEEE Trans Circuits Syst Express Briefs* (2023) 70:3248–52. doi:10.1109/TCSII.2023.3267859
43. Trueb P, Dejoie C, Kobas M, Pattison P, Peake DJ, Radici V, et al. Bunch mode specific rate corrections for PILATUS3 detectors. *J Synchrotron Radiat* (2015) 22:701–7. doi:10.1107/S1600577515003288
44. Leo WR. *Techniques for nuclear and particle physics experiments: a how-to approach*. 2nd ed. Berlin: Springer (1994). doi:10.1007/978-3-642-57920-2
45. Llopart X, Ballabriga R, Campbell M, Tlustos L, Wong W. Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2007) 581:485–94. doi:10.1016/j.nima.2007.08.079
46. Bergamaschi A, Dinapoli R, Greiffenberg D, Henrich B, Johnson I, Mozzanica A, et al. Time-over-threshold readout to enhance the high flux capabilities of single-photon-counting detectors. *J Synchrotron Radiat* (2011) 18:923–9. doi:10.1107/S0909049511034480
47. Llopart X, Alozy J, Ballabriga R, Campbell M, Egidio N, Fernandez J, et al. Study of low power front-ends for hybrid pixel detectors with sub-ns time tagging. *J Instrumentation* (2019) 14:C01024. doi:10.1088/1748-0221/14/01/C01024
48. Zambon P. Dead time model for x-ray photon counting detectors with retrigger capability. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2021) 994:165087. doi:10.1016/j.nima.2021.165087
49. Kappler S, Hölzer S, Kraft E, Stierstorfer K, Flohr T. Quantum-counting CT in the regime of count-rate paralysis: introduction of the pile-up trigger method. *Proc SPIE - Int Soc Opt Eng* (2011) 7961:79610T. doi:10.1117/12.877939
50. Kraft E, Glasser F, Kappler S, Niederloehner D, Villard P. Experimental evaluation of the pile-up trigger method in a revised quantum-counting CT detector. In: Proceedings of the Medical Imaging 2012: Physics of Medical Imaging; San Diego, California (2012). p. 83134A. doi:10.1117/12.911231
51. Andrä M, Zhang J, Bergamaschi A, Barten R, Borca C, Borghi G, et al. Development of low-energy X-ray detectors using LGAD sensors. *J Synchrotron Radiat* (2019) 26:1226–37. doi:10.1107/S1600577519005393
52. Andrä M. *The MYTHEN III Detector System - a single photon-counting microstrip detector for powder diffraction experiments*. PhD thesis. Zürich, Switzerland: ETH Zürich (2021). Available at: <http://hdl.handle.net/20.500.11850/462676>.
53. Donath T, Brandstetter S, Cibik L, Commichau S, Hofer P, Krumrey M, et al. Characterization of the pilatus photon-counting pixel detector for x-ray energies from 1.75 keV to 60 keV. *J Phys Conf Ser* (2013) 425:062001. doi:10.1088/1742-6596/425/6/062001
54. Bergamaschi A, Cartier S, Dinapoli R, Greiffenberg D, Jungmann-Smith J, Mezza D, et al. Looking at single photons using hybrid detectors. *J Instrumentation* (2015) 10:C01033. doi:10.1088/1748-0221/10/01/C01033
55. Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. Geant4—a simulation toolkit. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 506:250–303. doi:10.1016/S0168-9002(03)01368-8
56. Schübel A, Krapohl D, Fröjd E, Fröjd C, Thungström G. A geant4 based framework for pixel detector simulation. *J Instrumentation* (2014) 9:C12018. doi:10.1088/1748-0221/9/12/C12018
57. Leonarski F, Redford S, Mozzanica A, Lopez-Cuenca C, Panepucci E, Nass K, et al. Fast and accurate data collection for macromolecular crystallography using the jungfrau detector. *Nat Methods* (2018) 15:799–804. doi:10.1038/s41592-018-0143-7
58. Tlustos L. *Performance and limitations of high granularity single photon processing X-ray imaging detectors* (2005).
59. Gimenez EN, Ballabriga R, Campbell M, Horswell I, Llopart X, Marchal J, et al. Study of charge-sharing in medipix3 using a micro-focused synchrotron beam. *J Instrumentation* (2011) 6:C01031. doi:10.1088/1748-0221/6/01/C01031
60. Koenig T, Zuber M, Hamann E, Cecilia A, Ballabriga R, Campbell M, et al. How spectroscopic x-ray imaging benefits from inter-pixel communication. *Phys Med Biol* (2014) 59:6195–213. doi:10.1088/0031-9155/59/20/6195
61. Frojdh E, Ballabriga R, Campbell M, Fiederle M, Hamann E, Koenig T, et al. Count rate linearity and spectral response of the medipix3rx chip coupled to a 300 μ m silicon sensor under high flux conditions. *J Instrumentation* (2014) 9:C04028. doi:10.1088/1748-0221/9/04/C04028
62. Cartier S, Kagi M, Bergamaschi A, Wang Z, Dinapoli R, Mozzanica A, et al. Micrometer-resolution imaging using MÖNCH: towards G₂-less grating interferometry. *J Synchrotron Radiat* (2016) 23:1462–73. doi:10.1107/S1600577516014788
63. Bergamaschi A, Andrä M, Barten R, Baruffaldi F, Brückner M, Carulla M, et al. First demonstration of on-chip interpolation using a single photon counting microstrip detector. *J Instrumentation* (2022) 17:C11012. doi:10.1088/1748-0221/17/11/C11012
64. Krzyzanowska A, Szczygiel R, Grybos P, Koczwar J, Cisko K, Stasiak A. Charge sharing simulations and measurements for digital algorithms aiming at subpixel resolution in photon counting pixel detectors. *J Instrumentation* (2023) 18:C02024. doi:10.1088/1748-0221/18/02/C02024
65. Flohr T, Petersilka M, Henning A, Ulzheimer S, Ferda J, Schmidt B. Photon-counting ct review. *Physica Med* (2020) 79:126–36. doi:10.1016/j.ejpm.2020.10.030
66. Holmes TW, Liu LP, Shapira N, McVeigh E, Litt HI, Pourmortez A, et al. Mixed coronary plaque characterization with the first clinical dual-source photon-counting CT scanner: a phantom study. In: Stayman JW, editor. Proceedings of the 7th International Conference on Image Formation in X-Ray Computed Tomography. International Society for Optics and Photonics SPIE (2022). p. 123041U. doi:10.1117/12.2647009
67. Brombal L, Donato S, Brun F, Delogu P, Fanti V, Oliva P, et al. Large-area single-photon-counting CdTe detector for synchrotron radiation computed tomography: a dedicated pre-processing procedure. *J Synchrotron Radiat* (2018) 25:1068–77. doi:10.1107/S1600577518006197
68. Krause L, Tolborg K, Grønbech TBE, Sugimoto K, Iversen BB, Overgaard J. Accurate high-resolution single-crystal diffraction data from a Pilatus3X CdTe detector. *J Appl Crystallogr* (2020) 53:635–49. doi:10.1107/S1600576720003775
69. Aslandukov A, Aslandukova A, Laniel D, Koemets I, Fedotenko T, Yuan L, et al. High-pressure yttrium nitride, y₅n₁₄, featuring three distinct types of nitrogen dimers. *The J Phys Chem C* (2021) 125:18077–84. doi:10.1021/acs.jpcc.1c06210
70. Pennicard D, Struth B, Hirsemann H, Sarajlic M, Smoljanin S, Zuvic M, et al. A germanium hybrid pixel detector with 55m pixel size and 65,000 channels. *J Instrumentation* (2014) 9:P12003. doi:10.1088/1748-0221/9/12/P12003
71. Bolotnikov AE, Abdul-Jabbar NM, Babalola OS, Camarda GS, Cui Y, Hossain AM, et al. Effects of te inclusions on the performance of cdznte radiation detectors. *IEEE Trans Nucl Sci* (2008) 55:2757–64. doi:10.1109/TNS.2008.2003355
72. Baussens O, Ponchut C, Ruat M, Bettelli M, Zanettini S, Zappettini A. Characterization of high-flux cdznte with optimized electrodes for 4th generation

- synchrotrons. *J Instrumentation* (2022) 17:C11008. doi:10.1088/1748-0221/17/11/C11008
73. Iniewski K. Czt detector technology for medical imaging. *J Instrumentation* (2014) 9:C11001. doi:10.1088/1748-0221/9/11/C11001
74. Pennicard D, Pirard B, Tolbanov O, Iniewski K. Semiconductor materials for x-ray detectors. *MRS Bull* (2017) 42:445–50. doi:10.1557/mrs.2017.95
75. Fiederle M, Procz S, Hamann E, Fauler A, Fröjdh C. Overview of gaas und cdte pixel detectors using medipix electronics. *Cryst Res Tech* (2020) 55:2000021. doi:10.1002/crat.202000021
76. Greiffenberg D, Andrä M, Barten R, Bergamaschi A, Busca P, Brückner M, et al. Characterization of gaas:cr sensors using the charge-integrating jungfrau readout chip. *J Instrumentation* (2019) 14:P05020. doi:10.1088/1748-0221/14/05/P05020
77. Wei H, Huang J. Halide lead perovskites for ionizing radiation detection. *Nat Commun* (2019) 10:1066. doi:10.1038/s41467-019-08981-w
78. Pan L, He Y, Klepov VV, De Siena MC, Kanatzidis MG. Perovskite cspbbr3 single crystal detector for high flux x-ray photon counting. *IEEE Trans Med Imaging* (2022) 41:3053–61. doi:10.1109/TMI.2022.3176801
79. Wei H, DeSantis D, Wei W, Deng Y, Guo D, Savenije TJ, et al. Dopant compensation in alloyed ch3nh3pbbr3-x cl x perovskite single crystals for gamma-ray spectroscopy. *Nat Mater* (2017) 16:826–33. doi:10.1038/nmat4927
80. Greiffenberg D, Andrä M, Barten R, Bergamaschi A, Brückner M, Busca P, et al. Characterization of chromium compensated gaas sensors with the charge-integrating jungfrau readout chip by means of a highly collimated pencil beam. *Sensors* (2021) 21:1550. doi:10.3390/s21041550
81. Carulla M, Vignali MC, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, et al. Study of the internal quantum efficiency of fbk sensors with optimized entrance windows. *J Instrumentation* (2023) 18:C01073. doi:10.1088/1748-0221/18/01/C01073
82. Cartiglia N, Arcidiacono R, Borghi G, Boscardin M, Costa M, Galloway Z, et al. Lgad designs for future particle trackers. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2020) 979:164383. doi:10.1016/j.nima.2020.164383
83. Giacomini G. Lgad-based silicon sensors for 4d detectors. *Sensors* (2023) 23:2132. doi:10.3390/s23042132
84. Moffat N, Bates R, Bullough M, Flores L, Maneuski D, Simon L, et al. Low gain avalanche detectors (lgad) for particle physics and synchrotron applications. *J Instrumentation* (2018) 13:C03014. doi:10.1088/1748-0221/13/03/c03014
85. Galloway Z, Gee C, Mazza S, Ohldag H, Rodriguez R, Sadrozinski H-W, et al. Use of “lgad” ultra-fast silicon detectors for time-resolved low-keV x-ray science. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 923:5–7. doi:10.1016/j.nima.2019.01.050
86. Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of lgad sensors with a thin entrance window for soft x-ray detection. *J Instrumentation* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/C11011
87. Butcher TA, Phillips NW, Chiu C-C, Wei C-C, Ho S-Z, Chen Y-C, et al. *Ptychographic nanoscale imaging of the magnetoelectric coupling in freestanding bifeo3* (2023).
88. Mozzanica A, Bergamaschi A, Brueckner M, Cartier S, Dinapoli R, Greiffenberg D, et al. Characterization results of the jungfrau full scale readout ASIC. *J Instrum* (2016) 11:C02047. doi:10.1088/1748-0221/11/02/C02047
89. Leonarski F, Redford S, Mozzanica A, Lopez-Cuenca C, Panepucci E, Nass K, et al. Fast and accurate data collection for macromolecular crystallography using the jungfrau detector. *Nat Commun* (2018) 15:799–804. doi:10.1038/s41592-018-0143-7
90. Bergamaschi A, Andrä M, Barten R, Borca C, Brückner M, Chirioti S, et al. The mÖnch detector for soft x-ray, high-resolution, and energy resolved applications. *Synchrotron Radiat News* (2018) 31:11–5. doi:10.1080/08940886.2018.1528428
91. Cline B, Banks D, Bell S, Church I, Cross S, Davis A, et al. Characterisation of hexitecmhz - a 1 mhz continuous frame rate spectroscopic x-ray imaging detector system. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168718. doi:10.1016/j.nima.2023.168718
92. Leonarski F, Brückner M, Lopez-Cuenca C, Mozzanica A, Stadler H-C, Matěj Z, et al. JungfrauJoch: hardware-accelerated data-acquisition system for kilohertz pixel-array X-ray detectors. *J Synchrotron Radiat* (2023) 30:227–34. doi:10.1107/S1600577522010268
93. Weiss JT, Shanks KS, Philipp HT, Becker J, Chamberlain D, Purohit P, et al. High dynamic range x-ray detector pixel architectures utilizing charge removal. *IEEE Trans Nucl Sci* (2017) 64:1101–7. doi:10.1109/TNS.2017.2679540
94. Fajardo P, Busca P, Erdinger F, Fischer P, Ruat M, Schimansky D. Digital integration: a novel readout concept for x-ray detector for the next generation of synchrotron radiation sources. *J Instrumentation* (2020) 15:C01040. doi:10.1088/1748-0221/15/01/C01040
95. Shi X, Tinti G, Dinapoli R, Mozzanica A, Schmitt B. *Dual mode detector, european patent application ep4042204a1* (2020).
96. Wang C, Steiner U, Sepe A. Synchrotron big data science. *Small* (2018) 14:1802291. doi:10.1002/smll.201802291
97. Llopert X, Aloyz J, Ballabriga R, Campbell M, Casanova R, Gromov V, et al. Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning. *J Instrumentation* (2022) 17:C01044. doi:10.1088/1748-0221/17/01/C01044
98. Götz A, le Gall E, Konrad U, Kourousias G, Knodel O, Matalgah S, et al. Leaps data strategy. *The Eur Phys J Plus* (2023) 138:617. doi:10.1140/epjp/s13360-023-04189-6
99. Sriskaran V (2022). *Medipix4, a high granularity four sides buttable pixel readout chip for high resolution spectroscopic x-ray imaging at rates compatible with medical ct scans*. doi:10.5075/epfl-thesis-8617
100. Berger MJ, H Hubbell J, Olsen K. *Xcom: photon cross section database (version 1.5)* (2010).
101. Krause MO. Atomic radiative and radiationless yields for K and L shells. *J Phys Chem Reference Data* (1979) 8:307–27. doi:10.1063/1.555594



OPEN ACCESS

EDITED BY

Piernicola Oliva,
University of Sassari, Italy

REVIEWED BY

Gabriel Blaj,
Stanford University, United States
Antonino Miceli,
Argonne National Laboratory (DOE),
United States

*CORRESPONDENCE

David Pennicard,
✉ david.pennicard@desy.de

RECEIVED 30 August 2023

ACCEPTED 25 January 2024

PUBLISHED 05 February 2024

CITATION

Pennicard D, Rahmani V and Graafsma H (2024),
Data reduction and processing for photon
science detectors.
Front. Phys. 12:1285854.
doi: 10.3389/fphy.2024.1285854

COPYRIGHT

© 2024 Pennicard, Rahmani and Graafsma. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Data reduction and processing for photon science detectors

David Pennicard^{1*}, Vahid Rahmani¹ and Heinz Graafsma^{1,2}

¹Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ²STC research centre, Mid Sweden University, Sundsvall, Sweden

New detectors in photon science experiments produce rapidly-growing volumes of data. For detector developers, this poses two challenges; firstly, raw data streams from detectors must be converted to meaningful images at ever-higher rates, and secondly, there is an increasing need for data reduction relatively early in the data processing chain. An overview of data correction and reduction is presented, with an emphasis on how different data reduction methods apply to different experiments in photon science. These methods can be implemented in different hardware (e.g., CPU, GPU or FPGA) and in different stages of a detector's data acquisition chain; the strengths and weaknesses of these different approaches are discussed.

KEYWORDS

photon science, detectors, X-rays, data processing, data reduction, hardware acceleration, DAQ

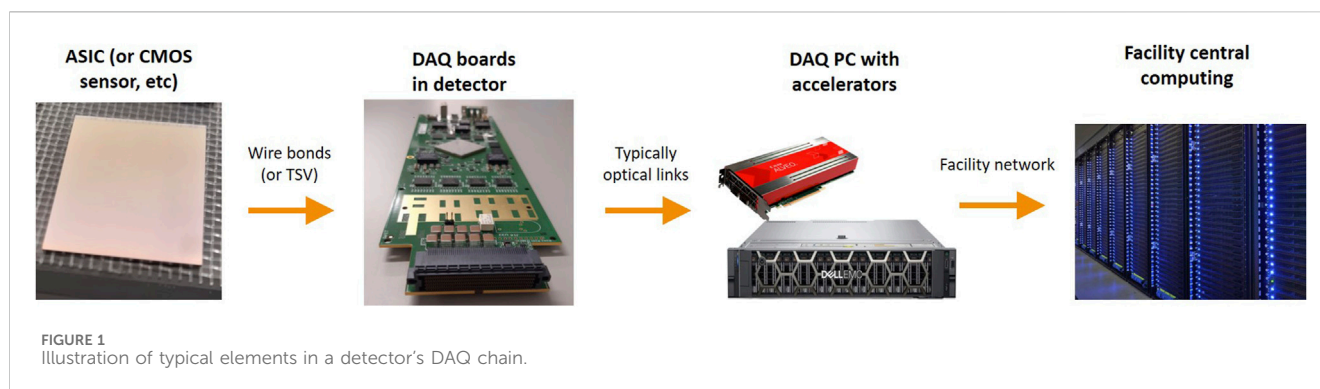
1 Introduction

Developments in photon science sources and detectors have led to rapidly-growing data rates and volumes [1]. For example, experiments at the recently-upgraded ESRF EBS can potentially produce a total of a petabyte of data per day, and future detectors targeting frame rates over 100 kHz will have data rates (for raw data) exceeding 1 Tbit/s [2].

These improvements not only allow a much higher throughput of experiments, but also make new measurements feasible. For example, by focusing the beam and raster-scanning it across a sample at high speed, essentially any X-ray technique can be used as a form of microscopy, obtaining atomic-scale structure and chemical information about large samples. But naturally, these increasing data rates pose a variety of challenges for data storage and analysis. In particular, there is increasing demand for data reduction, to ensure that the volume of data that needs to be transferred and stored is not unfeasibly large. From the perspective of detector developers, there are two key issues that need to be addressed.

Firstly, the raw data stream from a detector needs to be converted into meaningful images, and this becomes increasingly challenging at high data rates. This conversion process is detector-specific, so implementing it requires detailed knowledge of the detector's characteristics. At the same time, since the correction process is relatively fixed, there's a lot of potential to optimize it for performance. In addition, the complexity of this conversion process depends on the detector design, so this is something that should be considered during detector development.

Secondly, it can be beneficial to perform data reduction on-detector, or as part of the detector's DAQ system. For a variety of reasons, the useful information in a dataset can be captured with a smaller number of bits than the original raw data size; for example, by taking advantage of patterns or redundancy in the raw data, or by rejecting non-useful images in the dataset. In the DAQ system, the data will typically pass through a series of stages, as illustrated in [Figure 1](#); firstly, from the ASIC or monolithic sensor to a custom



board in the detector, then to a specialized DAQ PC (or similar hardware), and finally to a more conventional computing environment. By performing data reduction early on in this chain, it is possible to reduce the bandwidth required by later stages. Not only can this reduce the cost and complexity of later stages (especially the cost of data storage), it can also potentially enable the development of faster detectors by overcoming data bandwidth bottlenecks. Performing this early data reduction often ties together with the process of converting the raw data stream to real images, since real images can be easier to compress. Conversely, though, performing data reduction early in the chain can be more challenging, since the hardware in these early stages tends to offer less flexibility, and there are constraints on space and power consumption within the detector.

This paper presents an overview of image correction and data reduction in photon science, with a particular focus on how the characteristics of different detectors and experiments affect the choice of data reduction methods. [Section 2](#) discusses algorithms for converting raw data into corrected images. [Section 3](#) addresses algorithms for data compression, followed by other methods of data reduction (e.g., rejecting bad images) in [Section 4](#). [Section 5](#) provides an overview of standard data-processing hardware such as CPUs, GPUs and FPGAs that can be built into detectors and DAQ systems. Finally, [Section 6](#) brings these elements together, by discussing how data correction and reduction can be implemented at different stages of a detector's readout and processing chain, and the pros and cons of different approaches.

2 Detector data correction

The raw data stream produced by a detector generally requires processing in order to produce a meaningful image. The steps will of course depend on the design; here, two common cases of photon counting and integrating pixel detectors are considered. Although it can be possible to compress data before all the corrections are applied, corrected data can be more compressible—for example, correcting pixel-to-pixel variations can result in a more uniform image.

Firstly, we want the pixels in an image to follow a straightforward ordering; typically this is row-by-row or column-by-column, though some image formats represent images as a series of blocks for performance reasons. However, data streams from detectors often have a more complex ordering. One reason for this is

that data is typically read out from an ASIC or monolithic detector in parallel across multiple readout signal lines which can result in interleaving of data. In addition, in detectors composed of multiple chips or modules, there may be gaps in the image, or some parts of the detector may be rotated—this is illustrated for the AGIPD 1M detector [3] in [Figure 2](#). So, data reordering is a common first step.

In the case of photon counting detectors, the value read out from each pixel is an integer that corresponds relatively directly to the number of photons hitting the pixel. Nevertheless, at higher count rates losses occur due to pulse pileup, when photons hit a pixel in quick succession and only one pulse is counted. So, pileup correction is needed, where the hit rate in each pixel is calculated and a rate-dependent multiplicative factor applied. Although there are two well-known models for pileup—paralyzable and non-paralyzable—in practice the behaviour of pixel detectors can be somewhere in-between, and the pulsed structure of synchrotron sources can also affect the probability of pileup [4]. In addition to this some modern detectors have additional pileup compensation, for example, by detecting longer pulses that would indicate pileup [5]. So, the pileup correction model can vary between different detectors.

In integrating detectors, each pixel's amplifier produces an analog value, which is then digitized. This digitized value then needs to be converted to the energy deposited in the pixel. In many detectors, this is a linear relationship. In experiments with monochromatic beam, each photon will deposit the same amount of energy in the sensor, so it is possible to calculate the corresponding number of photons. The correction process typically consists of the following steps [6, 7]:

- **Baseline correction/dark subtraction.** As part of the calibration process, dark images are taken with no X-ray beam present and averaged. Then, during image taking, this is subtracted from each new image. This corrects for both pixel-to-pixel variations in the amplifiers' zero level, and also for the effects of leakage current integrated during image taking. Since this integrated leakage current can vary with factors like integration time, operating temperature and radiation damage in the sensor, new dark images often need to be taken frequently, e.g., at the start of each experiment.
- **Common-mode correction.** In some detectors, there may be image-to-image variation that is correlated between pixels, for example, due to supply voltage fluctuations. One method for correcting this is to have a small number of pixels that are either masked from X-rays or unbonded, use these to measure

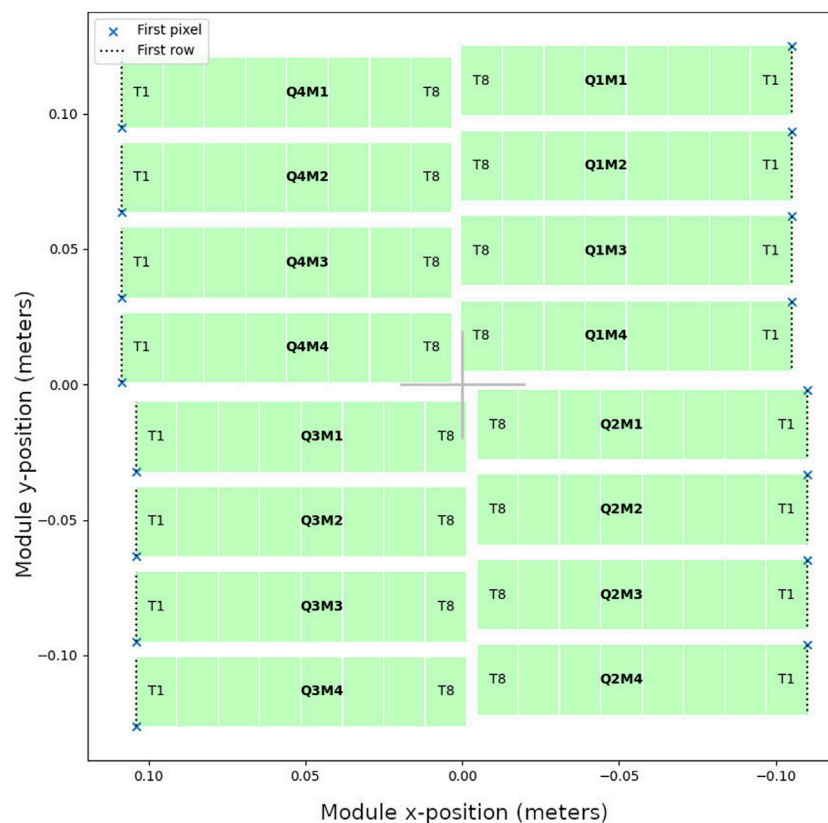


FIGURE 2
Module layout of the AGIPD 1M detector, which has a central hole for the beam, gaps between modules, and different module orientations (with the first pixel of each module indicated by a cross).

this common-mode noise in each image, and then subtract this from all the pixels. Another related effect is crosstalk, where pixels are affected by their neighbours; see for example, Ref. [8] for a discussion of this.

- Gain correction. While integrating detectors are typically designed to have a linear relationship between deposited energy and amplifier voltage, the gain will vary from pixel to pixel, and thus must be measured and corrected for by multiplication.
- “Photonization”. If the incoming X-ray beam is monochromatic, then the measured energy may then be converted to an equivalent number of photons. Up to this point, the correction process is typically performed with floating point numbers, but as discussed later it can be advantageous for compression to round this to an integer number of photons (or alternatively fixed point).

In some integrating detectors designed for large dynamic range, the response may not be a simple linear one. For example, in dynamic gain switching detectors [3, 9] each pixel can adjust its gain in response to the magnitude of the incoming signal, and the output of each pixel consists of a digitized value plus information on which gain setting was used. This increases the number of calibration parameters needing to be measured and corrected, since for each gain setting there will be distinct baseline and gain corrections.

Furthermore, there are a variety of ways in which detectors may deviate from the ideal response, and additional corrections may be required. For example, the response of a detector may not be fully linear, and more complex functions may be needed to describe their response. It is also common to treat malfunctioning pixels, for example, by setting them to some special value.

An additional aspect of detector data processing is combining the image data with metadata, i.e., contextual information about images such as their format, detector type, and the experimental conditions under which they were acquired. Some metadata may be directly incorporated into the detector’s data stream. For example, in Free Electron Laser (FEL) experiments the detector needs to be synchronised with the X-ray bunches, and bunches can vary in their characteristics, so each image will be accompanied by a bunch ID, fed to the detector from the facility’s control system. Other metadata may be added later. For example, the NeXus data format [10] has been adopted by many labs; this is based on the HDF5 format [11], and specifies how metadata should be structured in experiments in photon science and other fields.

3 Data compression and photon science datasets

In general, data compression algorithms reduce the number of bits needed to represent data, by encoding it in a way that takes

advantage of patterns or redundancy in the data. These algorithms can be subdivided into lossless compression [12], where the original data can be reconstructed with no error, and lossy compression [13], where there is some difference between the original and reconstructed data. Currently, lossless compression is often applied to photon science data before storage, whereas lossy compression is uncommon, due to concerns about degrading or biasing the results of later analysis. However, lossy compression can achieve higher compression ratios.

The compressibility of an image naturally depends on its content, and in photon science this can vary depending on both the type of experiment and the detector characteristics. In particular, since random noise will not have any particular redundancy or pattern, the presence of noise in an image will reduce the amount of compression that can be achieved with lossless algorithms.

Compared to typical visible light images, X-ray images from pixel detectors can have distinctive features that affect their compressibility, as discussed further below. Firstly, individual X-ray photons have much greater energy and thus can more easily be discriminated with a suitable detector. Also, X-ray diffraction patterns have characteristics that can make lossless compression relatively efficient. However, imaging experiments using scintillators and visible light cameras produce images more akin to conventional visible light imaging, which do not losslessly compress well.

3.1 Noise in X-ray images, and its effects on compression

Random noise in an image can potentially come from different sources; firstly, electronic noise introduced by the detector, and secondly, inherent statistical variation in the experiment itself.

Any readout electronics will inevitably have electronic noise. However, the signal seen by a detector consists of discrete X-ray photons, and the inherent discreteness of our signal makes it possible to reject the electronic noise, provided that it is small enough to ensure that noise fluctuations are rarely mistaken for a photon. In a silicon sensor without gain, ionizing radiation will generate on average one electron-hole pair per 3.6 eV of energy deposited, e.g., a 12 keV photon will generate approximately 3,300 electrons. In turn, if, for example, we assume the noise is Gaussian with a standard deviation corresponding to 0.1 times the photon energy (330 electrons here) then the probability of a pixel having a noise fluctuation corresponding to 0.5 photons or more would be approximately 1 in 1.7 million [14]. (Common noise sources such as thermal and shot noise discussed below are Gaussian, or approximately so, though other noise sources such as fluctuations in supply voltage may not be.)

When using an integrating detector to detect monochromatic X-rays, then it is possible to convert the integrated signal to an equivalent number of photons in postprocessing as described above. Given sufficiently low noise, this value can be quantized to the nearest whole number of photons to eliminate electronic noise. In photon counting detectors, a hit will be recorded in a pixel if the pulse produced by the photon exceeds a user-defined threshold; once again, if noise fluctuations exceeding the threshold are rare, we will have effectively noise-free counting.

In both cases, the electronic noise in a pixel will be dependent on the integration time for an image, or the shaping time in the case of a photon-counting detector. There are two main competing effects here. On the one hand, for longer timescales, the shot noise due to fluctuations in leakage current will be larger. Conversely, to achieve shorter integration times or shaping times, a higher amplifier bandwidth is required, and this will increase the amount of thermal noise detected [15]. Thermal noise is the noise associated with random thermal motion of electrons, which is effectively white noise.

In addition to electronic noise, however, the physics of photon emission and interaction are inherently probabilistic, and so even if an experiment were repeated under identical conditions there would be statistical fluctuations in the number of photons impinging on each pixel. These fluctuations follow Poisson statistics, and if the expected number of photons arriving in a pixel during a measurement is N , then the standard deviation of the corresponding Poisson distribution will be \sqrt{N} . On the one hand, this can make it easier to develop “low noise” detectors; provided the detector noise for a given photon flux is significantly smaller than \sqrt{N} , then it will have little effect on the final result. However, this makes the data less compressible with lossless algorithms, since random noise introduces variation in the image that is not patterned or redundant. (As discussed later, quantizing pixel values with a variable step size, increasing with \sqrt{N} , can be a way of achieving lossy compression.)

3.2 Applying lossless compression to diffraction data

As noted above, data compression relies on patterns or redundancy in data, and this will vary depending on the experiment. Diffraction patterns differ from conventional images in a variety of respects, with the pattern being mathematically related to the Fourier transform of the object. An example of a diffraction pattern from macromolecular crystallography is shown in Figure 3.

Diffraction patterns tend to have various features that are advantageous for lossless compression:

- Hybrid pixel detectors with sensitivity to single photons are the technology of choice for these experiments, making the measurement effectively free from electronic noise as described above. (Detectors for X-ray diffraction require high sensitivity, but pixel sizes in the range of 50–200 μm are generally acceptable.)
- The intensity values measured in the detector typically have a nonuniform distribution, with most pixels measuring relatively low or even zero photons, and a small proportion of pixels having higher values. This is partly because the diffracted intensity drops rapidly at higher scattering angles, and partly because of interference phenomena that tend to produce high intensity in certain places, e.g., Bragg spots or speckles, and low intensity elsewhere.
- Depending on the experiment, nearby pixels will often have similar intensities; for example, in single-crystal diffraction experiments with widely-spaced Bragg spots, the background signal between the spots tends to be smoothly varying.

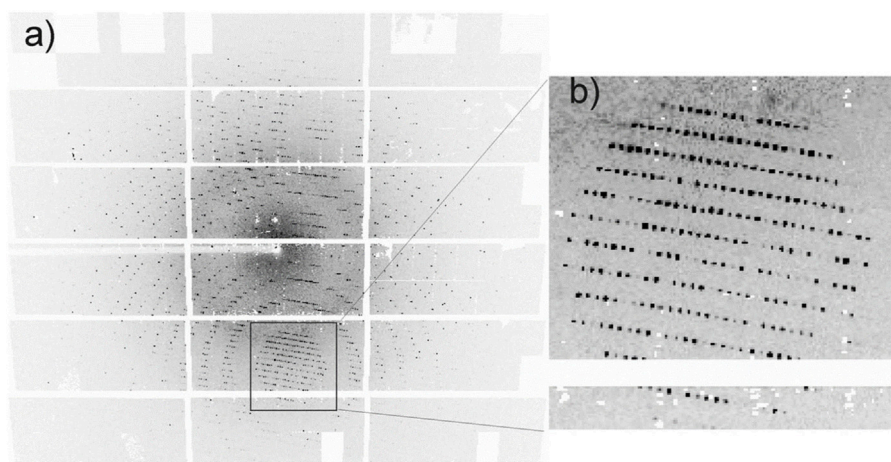


FIGURE 3
A diffraction pattern from a thaumatin crystal, recorded with a Pilatus 1M detector, showing (A) the full detector and (B) a zoom-in of Bragg spots. Reproduced from [16] with permission of the International Union of Crystallography.

- Combining these points, in images with fewer photons, there can be patches in the image with many neighbouring zero-value pixels.

Empirically, a range of lossless compression algorithms can achieve good results with diffraction data, especially in high-frame-rate measurements where the number of photons per image will tend to be lower. Experiments applying the DEFLATE [17] algorithm used in GZIP to datasets acquired with photon counting detectors running at high speed showed compression ratios of 19 for high-energy X-ray diffraction, 70 for ptychography and 350 for XPCS experiments with dilute samples (where most pixel values are zero) [18]. Experiments with the Jungfrau integrating detector, applying different compression algorithms to the same data, found that multiple compression methods such as GZIP, LZ4 with a bitshuffle filter and Zstd gave similar compression ratios to one another, but varied greatly in speed, with GZIP being a factor of 10 slower [19].

3.2.1 Example—lossless compression with DEFLATE (GZIP) and Bitshuffle/LZ4

As illustrative examples, we consider the DEFLATE algorithm used in GZIP [20], and the Bitshuffle/LZ4 algorithm [21], firstly to discuss how they take advantage of redundancy to compress diffraction data, and secondly how algorithms with different computational cost can achieve similar performance.

DEFLATE [17] consists of two stages. Firstly, the LZ77 algorithm [22] is applied to the data. This searches for recurring sequences of characters in a file (such as repeated words or phrases in text, or long runs of the same character) and encodes them efficiently, as follows: in the output, the first instance of a sequence of characters is written in full, but then later instances are replaced with special codes referring back to the previous instance. In diffraction datasets, long recurring sequences of characters are generally rare, but there is one big exception; long runs of zeroes. So, the pattern-matching in LZ77 will efficiently compress long runs of zeroes, but the computational work the

algorithm does to find more complex recurring sequences is largely wasted.

Secondly, DEFLATE takes the output of the LZ77 stage, and applies Huffman coding [23] to it. Normally, different characters in a dataset (e.g., integers in image data) are all represented with the same number of bits. Huffman coding looks at the frequencies of different characters in the dataset, and produces a new coding scheme that represents common characters with shorter sequences of bits and rare ones with longer sequences. This is analogous to Morse code, where the most common letter in English, “E”, is represented by a single dot, whereas rare letters have longer sequences. For diffraction data, this stage will achieve compression due to the nonuniform statistics of pixel values, where low pixel values are much more common than high ones.

As a contrasting example, the Bitshuffle LZ4 algorithm [21] implicitly takes advantage of the knowledge that the higher bits of pixel values are mostly zero and strongly correlated between neighbouring pixels. In the first step, Bitshuffle, the bits in the data stream are rearranged so that the first bit from every pixel are all grouped together, then the second bit, etc. After this regrouping, the result will often contain long runs of bytes with value zero, as illustrated in Figure 4. The LZ4 algorithm, which is similar to LZ77, will then efficiently encode these long runs of 0 bytes. As mentioned above, this algorithm gives similar performance to GZIP for X-ray diffraction data while being less computationally expensive.

3.3 Applying lossy compression to imaging data

While lossless compression can achieve good compression ratios for diffraction data, there is increasing demand for lossy data compression. Firstly, data from some experiments such as imaging do not losslessly compress well, due to noise, and secondly, the increasing data volumes produced by new detectors and experiments mean that higher compression ratios are desirable. The key challenge of lossy compression is ensuring that the

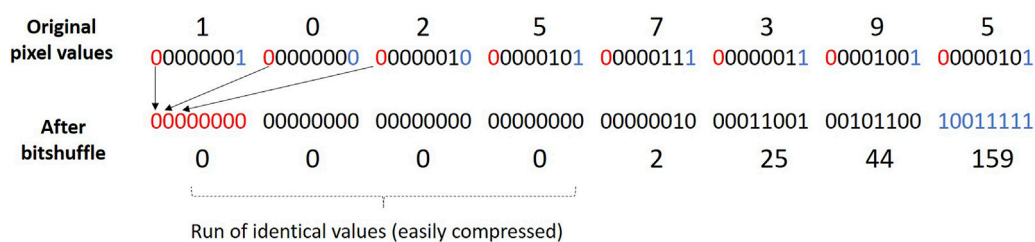


FIGURE 4

Illustration of the Bitshuffle process, used in Bitshuffle/LZ4. This reorders the bits in the data stream so that the LZ4 compression stage can take advantage of the fact that the upper bits in diffraction data are mostly zero, and strongly correlated between neighbouring pixels.. (Eight pixels with 8-bit depth are shown here, but the procedure may be applied to different numbers of pixels and bits.)

compression does not significantly change the final result of analysis. This requires evaluating the results of the compression with a variety of datasets, either by directly comparing the compressed and uncompressed images with a metric of similarity, or by performing the data analysis and applying some appropriate metric of quality to the final result.

In X-ray imaging and tomography experiments, the detector measures X-ray transmission through the sample, and perhaps also additional effects such as enhancement of edges through phase contrast. This means that most pixels will receive a reasonably high X-ray flux, in contrast to X-ray diffraction where most pixels see few or even zero photons. So, noise due to Poisson statistics will be relatively high in most pixels. Likewise, a key requirement for detectors in these experiments is a small effective pixel size (particularly for micro- and nano-tomography), whereas single photon sensitivity is less critical. To achieve this, a common approach is to couple a thin or structured scintillator to a visible light camera with small pixels such as a CMOS sensor or CCD [24]. Magnifying optics may be used to achieve a smaller effective pixel size [25]. This means that the detector noise is also non-negligible.

Since an X-ray transmission image is a real-space image of an object, and broadly resembles a conventional photograph (unlike a diffraction pattern), widely-used lossy compression algorithms for images can potentially be used for compression. In particular, JPEG2000 [26] is already well-established in medical imaging, and in tomography at synchrotrons it has been demonstrated to achieve a factor of three to four compression without significantly affecting the results of the reconstruction [27]. To test this, the reconstruction of the object was performed both before and after compression, and the two results compared using the Mean Structural Similarity Index Measure (MSSIM) metric. A compression factor of six to eight was possible in data with a high signal-to-noise ratio.

3.3.1 Example—lossy compression with JPEG2000

One common approach to both lossy and lossless compression is apply a transform to the data that results in a sparse representation, i.e., most of the resulting values are low or zero. (The choice of transform naturally depends on the characteristics of the data.) The sparse representation can then be compressed efficiently.

In the case of JPEG2000 [26], the Discrete Wavelet Transform [28] is used; in effect, this transformation represents the image as a sum of wave packets with different positions and spatial frequencies. This tends to work well for real space images, which tend to consist of a

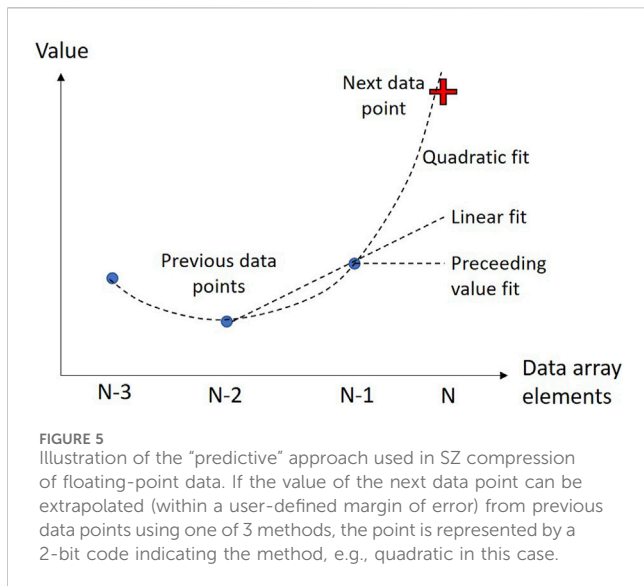
combination of localized objects and smooth gradients, which can be found on different length scales. After applying the transform, the resulting values are typically rounded off to some level of accuracy, allowing for varying degrees of lossy compression. (By not applying this rounding, lossless compression is also possible.) Then, these values are encoded by a method called arithmetic coding, which is comparable to Huffman coding; it achieves compression by taking advantage of the nonuniform statistics of the transformed values, which in this case are mostly small or zero.

JPEG2000 can be compared with the older JPEG standard, where the transformation consists of splitting the image into 8×8 blocks, and applying the Discrete Cosine Transform to each block, thus taking advantage of the fact that images tend to be locally smooth. This is computationally cheaper than most JPEG2000 implementations, but one key drawback is that this can lead to discontinuities between these 8×8 blocks after compression. Additionally, JPEG is limited to 8-bit depth (per colour channel) which is unsuitable for many applications, whereas JPEG2000 allows different bit depths. Recently, a high-throughput implementation of JPEG2000 has been developed, HTJ2K [29], with similar speed performance to JPEG. (This is compatible with the JPEG2000 standard but lacks certain features that are not important to scientific applications.)

3.4 Novel methods for lossy compression

As mentioned previously, increasing data volumes mean there is demand for achieving increasing compression, even for experiments such as X-ray diffraction where lossless compression works reasonably well. Naturally, this can be approached by testing a variety of well-established lossy compression algorithms on data, and experimenting with methods such as rounding the data to some level of accuracy. However, there are also new lossy compression methods being developed specifically for scientific data. One particular point of contrast is that most image compression algorithms focus on minimizing the perceptible difference to a human viewer, whereas for scientific data other criteria can be more important, such as imposing limits on the maximum error allowed between original values and compressed values.

One simple example is quantizing X-ray data with a step size smaller than the Poisson noise, which is proportional to \sqrt{N} . For ptychography, for example, it has been demonstrated that



quantizing pixel values with a step size of $0.5 \times \sqrt{N}$ [30] does not degrade the quality of the reconstruction.

A more sophisticated example of error-bounded lossy compression is the SZ algorithm [31], which is a method for compressing a series of floating-point values. For each new element in the series, the algorithm checks if its value can be successfully “predicted” (within a specified margin of error) using any one of three methods: directly taking the previous value; linear extrapolation from the previous two values; or quadratically extrapolating from the previous three values. If so, then the pixel value is represented by a 2-bit code indicating the appropriate prediction method. If not, then this “unpredictable” value needs to be stored explicitly. This is illustrated in Figure 5. After the algorithm runs, further compression is applied to the list of unpredictable values.

Reference [32] reports on applying SZ to serial crystallography data from LCLS, where the images consist of floating point values obtained from an integrating detector. In this method, lossless compression is applied to regions of interest consisting of Bragg peaks detected in the image, while binning followed by lossy compression is applied to the rest of the image, with the goal of ensuring that any weak peaks that went undetected will still have their intensities preserved sufficiently well. It is reported that when using this strategy, using SZ for the lossy compression can achieve a compression ratio of 190 while still achieving acceptable data quality, whereas other lossy compression methods tested gave a compression ratio of 35 at best.

4 Other forms of data reduction

The data compression methods discussed thus far work by representing a given file using fewer bits; the original file can be reconstructed from the compressed file, albeit with some inaccuracy in the case of lossy compression. More broadly, though, there are other methods of data reduction which rely on eliminating non-useful data entirely, or which transform or process the data in a non-recoverable way. These methods have the potential to greatly reduce

the amount of data needing to be stored, though of course it is crucial to establish that these methods are reliable before putting them into practice. In recent years, there has been increasing research into using machine learning for both rejecting non-useful data and for data processing. This includes supervised learning methods, where an algorithm learns to perform a task using training data consisting of inputs and the corresponding correct output, and unsupervised learning methods, which can discover underlying patterns in data.

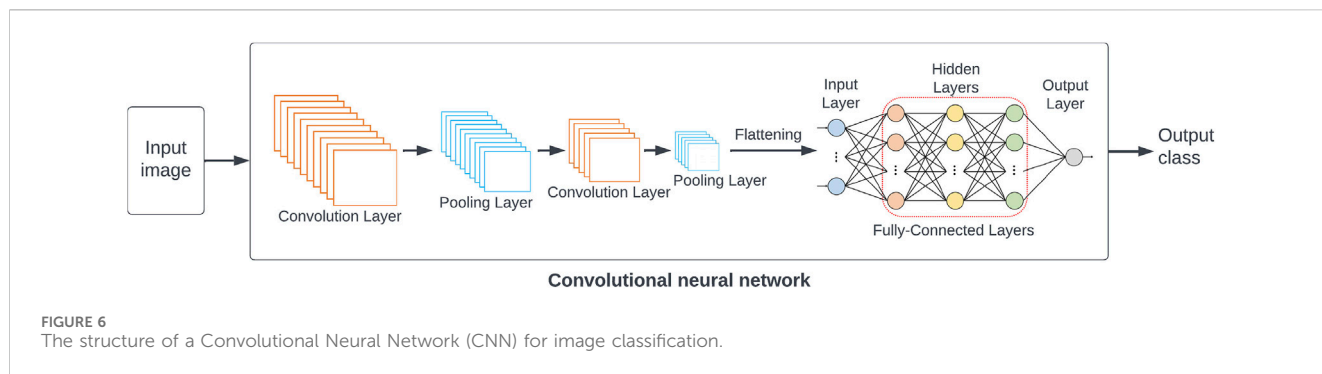
4.1 Data rejection/vetoing

In some experiments, a large fraction of the data collected does not contain useful information. For example, in experiments at FELs such as serial crystallography and single particle imaging, the sample can consist of a liquid jet containing protein nanocrystals or objects such as viruses passing through the path of the beam. However, only a small fraction of X-ray pulses (in some cases, of order 1%) actually hit a sample to produce an image containing a useful diffraction pattern. So, in these kind of experiments, data volumes can be greatly reduced by rejecting images where the beam did not hit the target.

A variety of methods have been developed for doing this. In serial crystallography, images where the beam hit a crystal will contain Bragg peaks, whereas miss images will only have scattering from the liquid jet. So, a common approach is to search for Bragg peaks in each image, and only accept images where the total number of peaks exceeds some threshold value; this approach is used in software such as Cheetah [33] and DIALS [34]. Since X-ray diffraction intensity varies with scattering angle, these methods often rely on estimating the background signal in the surrounding area of the image when judging whether a peak is present or not.

Machine learning techniques have also been applied to this task, where supervised learning is used to distinguish between good and bad images [35–38]. Supervised learning relies on having training data consisting of input images and the corresponding correct output. This data is used to train a model, such as a neural network; after this, the model can be used to classify new images. In some cases, the training data may consist of simulated data, or images from experiments that have been classified as good or bad by an expert. In this case, the appeal of machine learning is that it makes it possible to distinguish between good and bad images in cases where there is no known algorithm for doing so. Alternatively, the training data may consist of known good and bad images from previous experiments that have been classified by a pre-existing algorithm. In this case, machine learning can potentially improve on the algorithm by being less computationally intensive or by requiring less fine-tuning of algorithm parameters.

As an example, one common approach to image classification is deep learning using convolutional neural networks (CNNs) [39]. For neural networks, the training process consists of feeding input examples into the network, comparing the resulting outputs to the expected correct outputs, and adjusting the network’s parameters through a process called backpropagation in order to improve the network’s performance. CNNs are a special type of network suitable for large, structured inputs like images. As shown in Figure 6, they have a series of convolutional layers; each layer consists of an array of identical neurons, each looking at small patch



of the image. These layers learn to recognise progressively higher-level features in the image. Finally, fully-connected layers of neurons use these high-level features to classify the image. An example of using CNNs to categorize images in serial crystallography found in Ref. [36].

While machine learning is a rapidly-growing field with great potential, these techniques also have limitations. Firstly, substantial amounts of labelled training data are needed. Secondly, systematic differences between training data and new data may result in a failure to generalise to the new data; as a result, it is often necessary to test and retrain models. Furthermore, even if a machine learning algorithm works successfully, it is often not transparent how it works, making it difficult to understand and trust.

The strategy of vetoing uninteresting data is already well-established in particle physics collider experiments, where only a tiny fraction of collisions will produce particles that are interesting to the experimenter. The data from each collision is temporarily buffered, a subset of this data is read out, and this is used to make a decision on whether to read out the full event [40]. The need for triggering has driven the development of real-time data processing with short latencies [41] and the development of machine learning methods [42]; photon science can potentially benefit from this. However, compared to particle physics, vetoing in photon science faces the challenge that there can be a lot of variation between different beamlines and user experiments, and there is generally less *a priori* knowledge or simulation of what interesting or uninteresting data will look like.

4.2 Data reduction by processing

Often, the final result of data analysis is much more compact than the original dataset; in determining a protein structure, for example, the data may consist of tens of thousands of diffraction images, while the resulting structure can be described as a list of atomic positions in the molecule.

In some cases the full analysis of a dataset by the user may take months or years, and a lot of this analysis is very experiment-specific, so this is not suitable for achieving fast data reduction close to the detector. Nevertheless, there can be initial processing steps that are common to multiple experiments and which could be applied quickly as part of the DAQ system.

One example of this is azimuthal integration. In some X-ray diffraction experiments, such as powder diffraction, the signal on the

detector has rotational symmetry, and so the data can be reduced to a 1-D profile of X-ray intensity as a function of diffraction angle. For a multi-megapixel-sized detector, this corresponds to a factor of 1,000 reduction in data size. Hardware-accelerated implementations of azimuthal integration have been developed for GPUs [43], and more recently for FPGAs [44]. Another example is the calculation of autocorrelation functions in XPCS. In this technique, the dynamics of fluids can be studied from the fluctuations in speckle patterns produced by a coherent X-ray beam. Data processing consists first of calculating a per-pixel-correlation function from a large stack of images over time, and then averaging as a function of scattering angle. FPGA implementations are discussed in Ref. [45].

This is also an area where there is the potential to use machine learning. Firstly, data processing algorithms can be computationally expensive and time-consuming, especially if they involve an iterative reconstruction process. Supervised machine learning can potentially find a computationally-cheaper way of doing this, using experimental data and the results of the existing reconstruction method as the training data. For example, a neural network has been developed for extracting the structure of FEL pulses from gas-based streaking detectors, which would normally require iteratively solving a complex system of equations [46]. In this particular case, the neural network was implemented in an FPGA, in order to achieve low latency.

Additionally, unsupervised machine learning could potentially be used to extract information from X-ray data. These methods require training data, but unlike supervised learning only input data is needed, not information on the “correct” output. For example, variational autoencoders [47] learn to reduce complex data such as images to a vector representing their key features. This approach has been applied to X-ray diffraction patterns obtained from doped crystals, and it was found that the features extracted by the autoencoder could be used to determine the doping concentration [48].

4.3 Connections between data reduction and on-the-fly analysis

At light sources, there is a move towards performing on-the-fly data analysis, where data from the experiment is streamed to a computing cluster and analysed immediately [49]. This can provide immediate feedback to the user to guide the experiment and make data taking more efficient; for example, by allowing users to quickly

identify the most useful samples to study, or regions of interest within a sample that can be investigated in more detail [50].

If data reduction close to the detector involves performing analysis steps on data (e.g., azimuthal integration) or makes the data easier to process (e.g., making the dataset smaller by rejecting bad data) then this synergises with on-the-fly data processing, by reducing the workload on the computing cluster.

Typically, these sort of data reduction methods have the drawback that if some or all of the original data is discarded, then any errors in the processing cannot be corrected. For example, performing azimuthal integration correctly requires accurate calibration of the centre of the diffraction pattern and any tilt of the detector relative to the beam.

However, since on-the-fly data analysis is useful to users, the trustworthiness of these methods can be established in stages. For example, as a first stage all the raw data may be saved to disk for later analysis, with on-the-fly processing providing quick feedback during the experiment. Once the reliability of the automatic processing is better-established, raw data might only be stored for a shorter time before deletion (or transfer to cheaper storage such as tape) to give the opportunity for the user to cross-check for correctness. Finally, once the on-the-fly data processing is fully trusted, it may be possible to only save a small fraction of the raw data for validation purposes.

5 Overview of data processing hardware

“Off the shelf” data processing hardware such as CPUs, GPUs and FPGAs play a key role in building data acquisition systems for detectors. These components can be used in various places in the detector and DAQ system, and in some cases a combination of these can be used. For example, even custom hardware such as a circuit board for detector control will likely incorporate a microcontroller with a CPU, an FPGA, or a System-On-Chip containing both of these.

Here, we give an overview of the distinctive features of CPUs, GPUs and FPGAs for data processing. In particular, a major aspect of high-performance computing is parallelization of work, where a data processing task is divided between many processing units, but how this is achieved varies between devices, which can affect the best choice of hardware for the task. (How these different types of hardware can be incorporated into a detector system is discussed later, in [Section 6](#).) As an illustrative example [51], compares CPUs, GPUs and FPGAs for image processing tasks. For all the tasks studied, either the GPU, the FPGA or both offered around an order of magnitude higher speed than CPU, but which one was better depended on the task; for example, GPUs performed much better than FPGAs for image filtering with small filter size, whereas FPGAs were better for stereo vision.

In addition to these, a variety of new hardware accelerators aimed at machine learning have been developed. These vary in architecture, but some examples of these are discussed in [Subsection 5.4](#). Lastly, in recent years there have also been efforts to incorporate data processing directly into detector ASICs. This is discussed later on, in [6.1](#), since this is detector-specific processing.

5.1 CPUs—central processing units

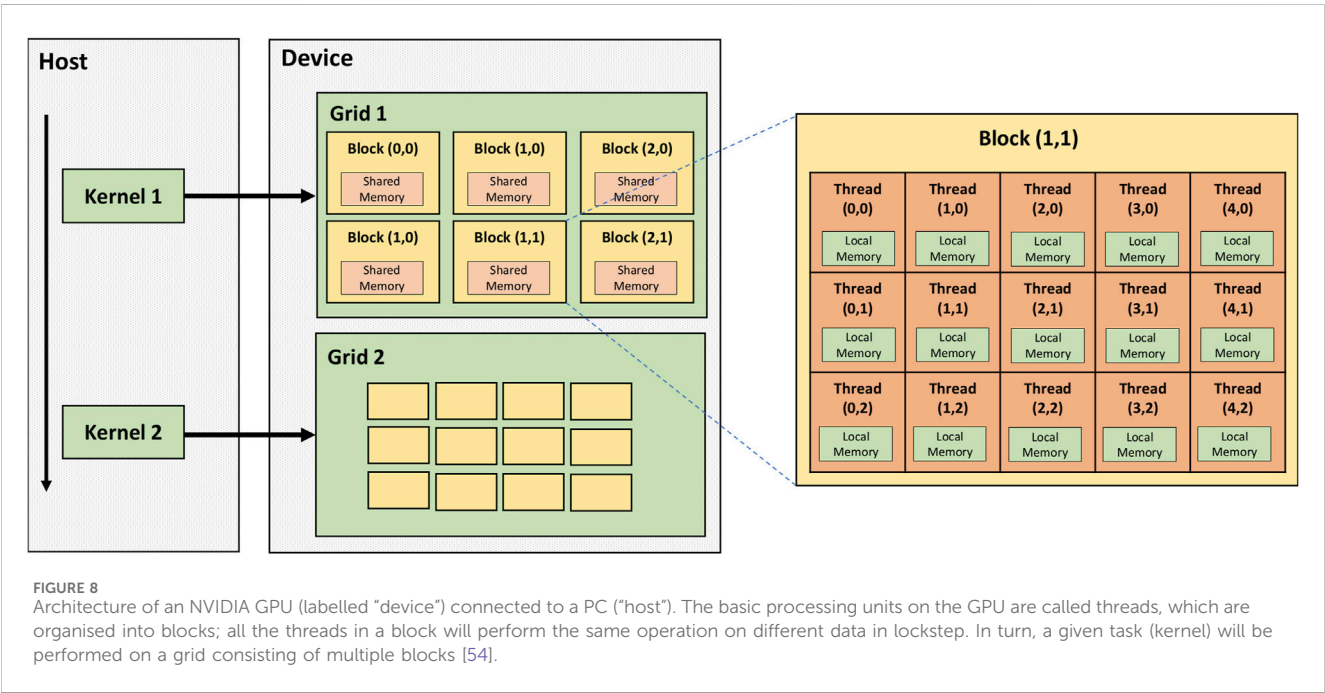
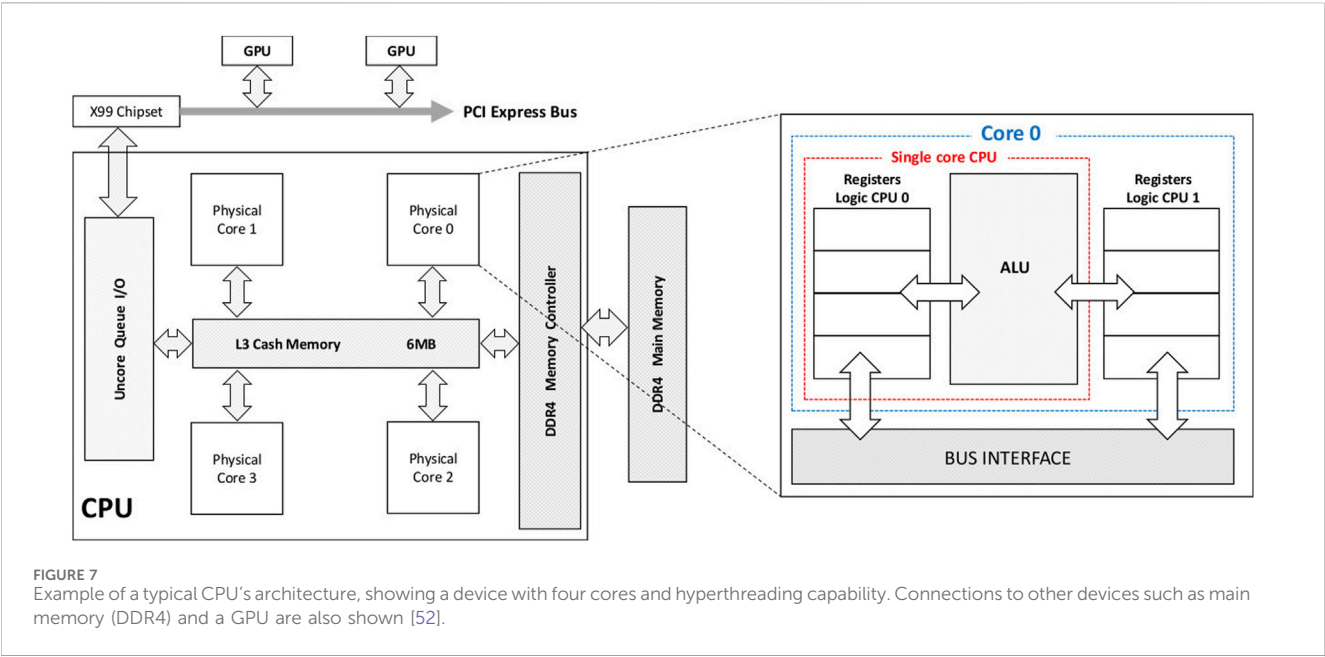
The architecture of a typical CPU [52] is shown in [Figure 7](#). A key feature of modern CPUs is that they’re designed to be able to run many different programs in a way that appears simultaneous to the user. Most modern CPUs are composed of a number of cores; for example, CPUs in the high-end AMD EPYC series have from 32 to 128, though desktop PCs typically have 2–6. Each core has a control unit, which is able to independently interpret a series of instructions in software and execute them, along with an arithmetic unit for performing operations on data, and cache memory for temporarily storing data. Furthermore, CPUs cores can very rapidly switch between executing different tasks, so even a single CPU core can perform many tasks in a way that seems parallel to the user. In most cases, the CPU will be running an operating system, which handles the sharing of resources such as memory and peripherals between the tasks in a convenient way for programmers. High performance computing with CPUs generally involves parallelization of work across multiple cores; in doing this, the different cores can operate relatively independently. For illustration, a detailed example of optimizing CPU code for the widely-used Fast Fourier Transform algorithm can be found in [Ref. \[53\]](#), which compares three different multithreaded packages for this.

Compared to GPUs and FPGAs, CPUs are generally the easiest to use and program. Indeed, GPUs are virtually always used as accelerator cards as part of a system with a CPU, and it is common for systems based on FPGAs to make use of a CPU (e.g., as part of a System On Chip). In addition, if a task is inherently sequential and cannot be parallelized, then a CPU core may give better performance than a GPU or FPGA. However, CPUs generally have much less parallel processing power than GPUs or FPGAs.

5.2 GPUs—graphics processing units

GPUs are designed to allow massively parallel processing for tasks such as graphics processing or general-purpose computing. The basic units of GPUs are called CUDA cores (in NVIDIA GPUs) or stream processors (in AMD GPUs). These are much more numerous than CPU cores, numbering in the thousands, allowing much greater parallelism and processing power. However, rather than each core being able to independently interpret a stream of instructions, these cores are organised into blocks, with all the cores in a block simultaneously executing the same instruction on different data elements - for example, applying the same mathematical operation to different elements in an array [54]. The structure of a typical GPU is shown in [Figure 8](#).

GPU programming relies on specialized frameworks such as CUDA [55], which was specifically developed for NVIDIA GPUs, or OpenCL [56], an open framework supporting a range of devices that also includes CPUs and FPGAs. GPUs are generally used as accelerator cards in a system with a CPU, and these frameworks firstly allow the CPU to control the GPU’s operation (by passing data to and from it, and starting tasks) and secondly to program functions that will run on the GPU. Typically, GPU code contains additional instructions controlling how different GPU cores in a block access data; e.g., when performing an operation on an array, the first core in a block might perform operations on the first



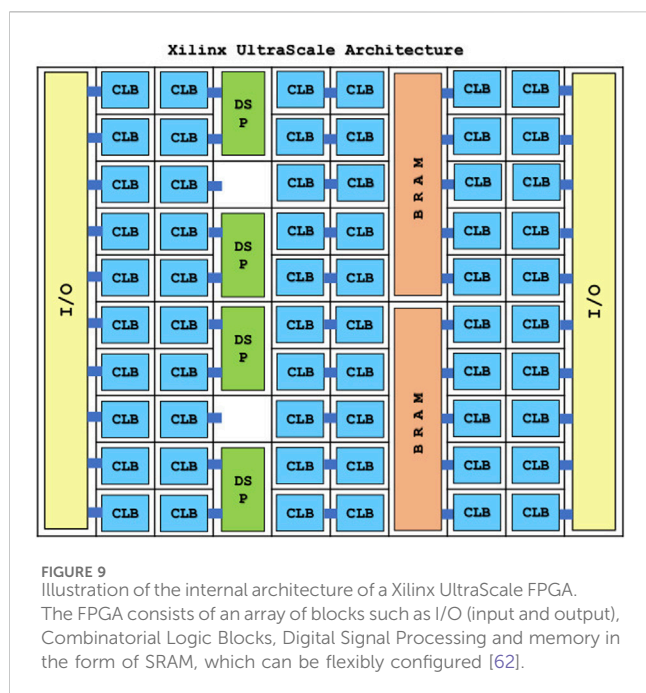
element, and so forth. This can be relatively easy to apply to a lot of image processing tasks, where each processing step can be applied to pixels relatively independently, but GPU implementation can be more challenging for some compression algorithms that act sequentially on data.

GPUs have been shown to perform well for many commonly-used algorithms such as matrix multiplication and Fourier transforms—for example, achieving an order of magnitude speed improvement in performing a Fast Fourier Transform [57]. In particular, GPUs are widely used for training and inference in machine learning in photon science [58–60]. Some newer models

of GPU are specially optimized for machine learning applications [61]. For example, neural network calculations typically consist of multiplying and adding matrices, and new GPUs can have specialised cores for this.

5.3 FPGAs—field programmable gate arrays

FPGAs are a type of highly configurable hardware. As shown in Figure 9, within an FPGA there are many blocks providing functionality such as combinatorial logic, digital signal processing



and memory, with programmable interconnects between them [62]. While software for a GPU or CPU consists of a series of instructions for the processor to follow, an FPGA's firmware configures the programmable interconnects to produce a circuit providing the required functionality. The large number of blocks in an FPGA can then provide highly parallel processing and high throughput of data.

Traditionally, firmware is developed by using a hardware description language to describe the required functionality; a compiler then finds a suitable configuration of blocks and interconnects to achieve this. While hardware description languages differ substantially from conventional programming languages, high-level-synthesis tools for FPGAs have been more recently developed to make it possible to describe an algorithm with a more conventional programming language such as C++, with special directives in the code being used to indicate how parallelism can be achieved with the FPGA.

A distinctive form of parallelism in FPGAs is pipelining. A series of processing steps will often be implemented as a series of blocks forming a pipeline, with data being passed along from one block to another. This is analogous to a production line in a factory, where each station carries out a particular fabrication step, and goods pass from one station to another. As with the production line, at any given moment there will be multiple data elements in different stages of processing being worked on simultaneously. In turn, to make algorithms run efficiently in an FPGA, they need to be designed to make effective use of this pipeline parallelism. This means that FPGAs tend to be better-suited to algorithms that operate on relatively continuous streams of data, rather than algorithms that rely on holding large amounts of data in memory and accessing them in an arbitrary way. As well as achieving a high throughput, algorithms on FPGA can give a reliably low latency, which can be important in cases where fast feedback is required. However, it is generally more challenging to program FPGAs than CPUs or GPUs.

In photon science, FPGAs have been used, for example, for performing autocorrelation in XPCS experiments [45] and azimuthal integration [44]. For machine learning, libraries are becoming available to make it easier to perform inference with neural networks on FPGAs [63]. Additionally, FPGA vendors have developed devices aimed at machine learning [64]; provides an evaluation of Xilinx's Versal platform, which, for example, includes "AI engines" for running neural networks.

5.4 AI accelerators

In recent years, a range of accelerators have been developed for machine learning. Examples of these include Intelligence Processing Units from Graphcore, and the GroqChip from Groq; an overview of some of these can be found in Ref. [65].

While these have varied architectures, they typically have certain features aimed at implementing neural networks efficiently. A neuron's output is a function of the weighted sum of its inputs, so AI accelerators are designed to perform many multiply-and-accumulate operations efficiently. They typically have large amounts of on-chip memory, in order to accommodate the high number of weights in a big neural network. Additionally, these accelerators can use numerical formats that are optimised for machine learning. For example, BFLOAT16 [66] is a 16-bit floating point format which covers the same range as standard IEEE 32-bit floating point but with lower precision; this is still sufficient for many neural networks, but makes calculations computationally cheaper.

These kinds of accelerator have recently been investigated for photon science applications. For example, in Ref. [67] a neural network for processing FEL pulse structure information from a streaking detector was implemented both for NVidia A100 GPUs and Graphcore IPUs. This was a convolutional network with encoder and decoder stages that could be used both to denoise the data and to extract latent features. Inference on the IPUs was roughly an order of magnitude faster than on GPUs, and training time was also significantly improved.

6 Data transfer and processing chains in detectors

A typical detector and DAQ system consist of series of stages, as illustrated previously in Figure 1. At each stage, data transfer is needed, and we face a series of bottlenecks which can limit the possible data rate. By doing data reduction earlier on in this chain, the data transfer demands on later stages are reduced, and we can take advantage of this to achieve higher detector speeds. Conversely, the hardware used in earlier stages of this chain is typically more "custom" or specialized, making the implementation of data reduction more challenging, and typically there is also less flexibility in these stages.

The stages in the readout and processing chain are as follows, and the potential for data reduction at each stage will be discussed in more detail in the following section.

- The readout chip (hybrid pixel) or monolithic sensor.

- Readout electronics within the detector itself, which typically incorporate an FPGA, microcontroller or similar.
- Detector-specific data processing hardware. Typically, this consists of one or more PCs, which can include peripherals such as FPGAs or GPUs, but it is also possible to build more specialized systems, such as crates of FPGA boards.
- Generic high-performance-computing hardware, typically in the facility's computing cluster.

Please note that this discussion focuses on data processing and reduction “close to the detector”, rather than a facility's full data analysis and storage system. Further information on this broader topic is available, for example, in publications on the newly-built LCLS-II data system [68] and an overview of big data at synchrotrons [69]. Furthermore, the level of data reduction required depends strongly on the costs of data processing and storage downstream. For example, for LCLS-II the goal is to reduce the data volume by at least a factor of 10 for each experiment.

6.1 On-chip data reduction

The first data bottleneck encountered in the detector is transferring data out of the hybrid pixel readout chip or monolithic sensor. In a pixel detector, we have a 2D array of pixels generating data, and within the chip it is possible to have a high density of signal routing; it is common to have a data bus for each pixel column. However, data output takes place across a limited number of transceivers (for example, 16 for the recent Timepix4 chip [70]), which are typically located at the periphery of the chip. These transceivers usually connect to a circuit board via wire bonds, though there is increasing effort in using Through Silicon Vias—TSVs—for interconnect [71]. In turn, further bottlenecks are faced in connecting these to the rest of the readout system. Many X-ray experiments require large continuous detector areas. Typically, these large detectors have a modular design, with each module's electronics placed behind it. In this situation, achieving high detector frame rates requires building readout electronics with high bandwidth per unit area; the achievable bandwidth per unit area then depends both on the performance and physical size of PCB traces and other components [72, 73]. So, performing on-chip data reduction can make it possible to build faster detectors without being limited by these bottlenecks.

Naturally, any circuitry for on-chip compression must be designed to not occupy excessive amounts of space in the pixels or periphery. Additionally, readout chips are developed in technology scales that are relatively large compared to commercial data processing hardware like GPUs and FPGAs, due to the high cost of smaller nodes. Typically, rather than implementing general-purpose processing logic into detector chips, specific algorithms are implemented.

Reference [18] presents a design for on-chip data compression for photon counting detectors, using techniques similar to those discussed previously in Section 3. Firstly, within the pixels, count values are encoded with a varying step size, with the step size getting larger for larger pixel values such that the step does not exceed the \sqrt{N} Poisson noise. This is lossy compression, but the additional

noise introduced by this encoding should be less than the Poisson noise. Then, in the chip's periphery during readout, a lossless compression scheme is applied that bit shuffles the data (much as described for bitshuffle LZ4) then encodes runs of zeroes efficiently. Applied to example datasets, this computationally-cheap approach achieved a compression ratio of around 6 for XRD data, compared to 19 obtained with GZIP.

More recently, lossy on-chip compression methods have been developed using two different algorithms - principal components analysis and an autoencoder [74]. The logic for doing this was distributed throughout the chip, with the analog pixel circuitry consisting of “islands” of 2×2 pixels. Applying these algorithms to an image requires prior knowledge about the content of a typical image; in the case of an autoencoder, which is a type of neural network, the neural weights are learned from training data. In this implementation, training was done using ptychography data, and the resulting weights were hard-coded in the chip.

One important limitation of on-chip data compression is that in a large tiled detector composed of multiple chips, the compressibility of the data can vary a great deal between different chips. In X-ray diffraction experiments in particular, the X-ray intensity close to the beam is much higher than at large scattering angles, leading to less compression. So, chips close to the beam may encounter data bottlenecks even if a high overall level of compression is achieved for the detector as a whole.

Another example of on-chip reduction is data vetoing by rejecting bad images, as discussed previously. The Sparkpix-ED chip is one of a family of chips with built-in data reduction being developed by SLAC [75]. It is an integrating pixel detector with two key features. Firstly, each pixel has built-in memory which is used to store recent images. Secondly, there is summing circuitry that can add together the signals in groups of 3×3 pixels, to produce a low-resolution image with 1/9 of the size. During operation, each time a new image is acquired the detector will store the full-resolution image in memory and send out the low-resolution image to the readout system. The readout system will then analyse the low-resolution images to see if an interesting event occurred (e.g., diffraction from a protein crystal in a serial crystallography experiment). If so, the detector can be triggered to read out the corresponding full-resolution image from memory; if not, the image will be discarded. A small prototype has demonstrated reading out low-resolution images at 1 MHz frame rate, and full-resolution images at 100 kHz.

6.2 On-detector processing and compression

Data processing and compression can potentially be done within the detector, before data is transferred to the control system (typically over optical links). Once again, this can reduce the bandwidth required for this data transfer.

Typically, on-detector electronics are needed both to control the detector's operation, and to perform serialization and encoding of data into some standard format so that it can be sent efficiently back to the control system and received using off-the-shelf hardware. (In some cases, it is also necessary to interface to additional components such as on-board ADCs.) One particular benefit of serialization is

that individual transceivers on a readout chip typically run at a lower data rate than can be achieved by modern optical links, so serialization can allow these links to be used more efficiently; for example, even the fastest on-chip transceivers typically do not have data rates above 5 Gbit/s, whereas data sent using 100 Gigabit Ethernet with QSFP28 transceivers consists of 4 channels with a data rate of 25 Gbit/s each. These tasks of control and serialization are typically implemented in an FPGA, or a System-On-Chip (SoC) with an FPGA fabric. So, the most common approach to on-detector data processing is to use the FPGA's processing resources.

As discussed in more detail earlier, FPGAs can provide highly-parallelized data processing, but firmware development can be challenging and time consuming. So, detector-specific processing routines that will always need to be performed on the data are better-suited to FPGA implementation than ones that vary a lot between experiments.

For example, FPGAs in the EIGER photon counting detector perform count-rate correction and image summing [76]. Since the counters in the pixel have a depth of 12 bits, this makes it possible for the detector to acquire images with a depth of up to 32 bits by acquiring a series of images and internally summing them, rather than needing to transfer many 12-bit images to the DAQ system.

One drawback of on-detector processing is that if the FPGA is used for control, serialization and data processing, then it can become more difficult to change the data processing routines. When FPGA firmware is compiled, the compiler will route together blocks in the FPGA to produce the desired functionality. So, changing the data processing routines can change the routing of other functionality in the FPGA, potentially affecting reliability. So, careful re-testing is required after changing the data processing. This is another reason why on-detector processing with FPGAs is mostly used for fixed, detector-specific routines.

6.3 Data acquisition hardware such as PCs with accelerator cards

Data sent out from a detector module will normally be received by either one or more DAQ PCs, or more specialized hardware, located either at the beamline or in the facility's computing centre. These parts of the DAQ system typically have a range of functions:

- Detector configuration and acquisition control, which requires both monitoring the detector's state and data output, and receiving commands from the control system.
- Ensuring reliable data reception. It is common for a detector's output to be a continuous flow of data, transferred by a simple protocol like UDP without the capability to re-send lost packets [77]. So the DAQ system is required to reliably receive and buffer this data; this typically requires, for example, having dedicated, high performance network or receiver cards.
- Data correction and reduction.
- Transferring data to where it is needed, for example, the facility's storage system, online processing and/or a user interface for feedback. This can include tasks like adding metadata and file formatting.

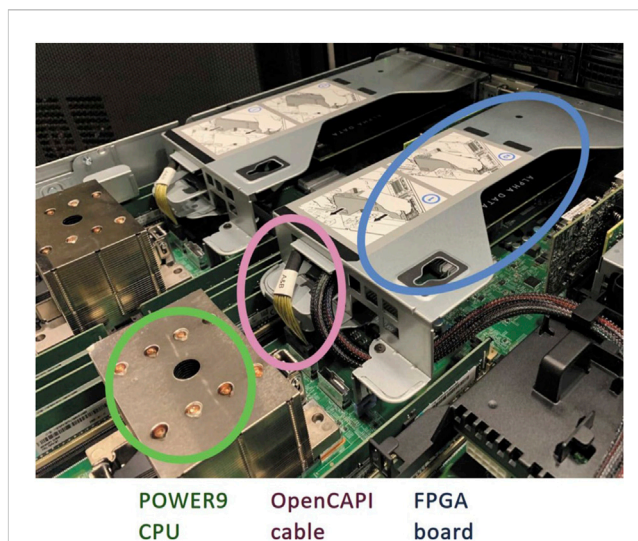


FIGURE 10

A photograph of the IC922 server (IBM) used in Jungfrauoch, showing the FPGA board used for data reception and processing, and the OpenCAPI link allowing coherent access to the CPU's memory. Reproduced from [78] with permission of the International Union of Crystallography.

Compared to detectors themselves, these systems tend to be built with relatively off-the-shelf hardware components - for example, standard network cards or accelerator cards. This use of off-the-shelf hardware can reduce the costs of implementing data correction and reduction. In addition, it can be easier to upgrade this hardware over time to take advantage of improvements in technology. These DAQ systems may do processing with conventional CPUs, hardware accelerators such as GPUs and FPGAs, or a mixture of these.

One recent example of this approach is the Jungfrauoch processing system, developed by PSI for the Jungfrau 4-megapixel detector [78]. (A similar approach is taken by the CITIUS detector [79].) The Jungfrauoch system is based on an IBM IC922 server PC, equipped with accelerator cards, and can handle 17 GB/s data when the detector is running at 2 kHz frame rate.

The Jungfrauoch server is shown in Figure 10 (reproduced from [78] with permission of the International Union of Crystallography). Firstly, it is equipped with two "smart network cards" from Alpha Data, each with a Xilinx Virtex Ultrascale + FPGA and a 100 Gigabit Ethernet network link. These receive data directly from the detector over UDP. The FPGA then converts the raw data into images, as described in Ref. [9]. Jungfrau is an integrating detector with gain switching, so the raw data consists of ADC values, and the process of converting this to either photons or energy values includes subtracting the dark current and scaling by an appropriate gain factor. The FPGA also performs the Bitshuffle algorithm, which is the first step in Bitshuffle LZ4 compression described previously; since this algorithm involves reordering bits in a data stream, this is well-suited to implementation on FPGA. The FPGA is primarily programmed by using High Level Synthesis (HLS) based on C++.

The data from each FPGA is transferred to the host server PC using OpenCAPI interconnects, which both allows high speed data transfer at 25 GB/s, and makes it simpler for the FPGA to access

memory on the server. The CPU in the server performs the LZ4 part of the Bitshuffle LZ4 lossless compression algorithm, then forwards this data to the file system using ZeroMQ so it can be written to the file system as an HDF5 file. There is also a GPU in the system that can perform tasks such as spot finding in macromolecular crystallography, and monitoring aspects of the detector's behaviour such as the dark current. (These tasks are well suited to GPUs, since they involve performing operations on full images in parallel.)

One promising technological development in this field is that FPGA and GPU vendors are increasingly developing accelerator cards with built-in network links for the datacenter market. For example, Xilinx have the Alveo line of FPGA cards with two 100 Gigabit Ethernet links and optional high-bandwidth memory, and likewise NVIDIA is currently developing GPU cards with network links. So, this will increase the availability of powerful, standardized hardware for building DAQ systems.

6.4 Computing clusters

Modern light sources are generally supported by large computer clusters for data processing and large storage systems [68]. Data from DAQ PCs or similar hardware at the beamlines can be transferred to them over the facility's network. Data processing in computing clusters can be divided into online analysis, where the data is transferred directly to the cluster to provide fast results to the experiment, and offline analysis where data is read back from storage for processing at a later point. (Naturally, there may not be a sharp dividing line between these two approaches.) In addition to server PCs with CPUs, computing clusters can also incorporate GPUs or less commonly FPGAs.

In a computing cluster, there is generally a scheduling system that controls how tasks are assigned to the computers. This has the advantage of allowing sharing of resources between different experiments and users according to needs, whereas dedicated hardware installed at a beamline may be idle much of the time. Computing clusters are also much better-suited to allowing users to remotely access to their data and providing the software tools required to analyse it. However, this flexibility in task scheduling and usage can make it more difficult to ensure that we can reliably receive and process images from the detector at the required rate; this is one reason for having dedicated DAQ computers at the beamline for receiving detector data.

7 Conclusion

The increasing data rates of detectors for photon science mean there is a need for high-speed detector data correction, and data reduction.

There are strong ties between these tasks. On the one hand, data reduction can often yield better results on properly-corrected data, since the correction process can reduce spurious variation in images (e.g., pixel-to-pixel variation in response) and make it easier to exploit redundancy in the data. Conversely, after lossy data reduction it is impossible to perfectly recover the original data, so it is crucial to ensure that the quality of the data correction is as good as possible.

Both of these tasks can benefit from making better use of hardware accelerators such as GPUs and FPGAs for highly parallel processing. The increasing use of accelerators in other areas, such as datacenters, means that we can take advantage of improvements in both their hardware and in tools for programming them. In particular, as FPGAs and GPUs with built-in network links become available “off the shelf”, this increases the potential for different labs to build their DAQ systems with compatible hardware, and share the algorithms they develop. Although this paper emphasizes data processing hardware, it is also important to note that well-designed and coded algorithms can deliver much better performance.

Data reduction is a growing field in photon science. To date, lossless compression has mainly been used, since this ensures that there is no loss in data quality, and in some experiments lossless methods can achieve impressive compression ratios. However, lossless compression is relatively ineffective for methods such as imaging, and as data volumes increase there is demand for even greater data reduction in diffraction experiments. So, there will be an increasing need for lossy compression and other methods of reduction. In doing so, it is crucial to use appropriate metrics to test that the data reduction does not significantly reduce the quality of the final analysis. By using well-established methods of lossy compression, for example, image compression with JPEG2000, it is easier to incorporate data reduction into existing data analysis pipelines. However, novel methods of data reduction tailored to photon science experiments have the potential for better performance.

Data reduction can be incorporated into different stages of a detector's DAQ chain; generally, performing data reduction earlier in the chain is more challenging and less flexible, but has the advantage of reducing the bandwidth requirements of later stages, and can enable greater detector performance by overcoming bottlenecks in bandwidth. The development of on-chip data reduction is a particularly exciting development for enabling higher-speed detectors, though this would typically require the development of different chips for different classes of experiment. As the demand for data reduction increases, we can expect detectors and experiments to incorporate a series of data reduction steps, beginning with simpler or more generic reduction early in the processing chain, and then more experiment-specific data reduction taking place in computer clusters.

Author contributions

DP: Conceptualization, Investigation, Writing—original draft, Writing—review and editing. VR: Visualization, Writing—review and editing. HG: Funding acquisition, Supervision, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Additional funding for work on data reduction has been provided by multiple sources: LEAPS-INNOV, which has received funding from the European Union's Horizon 2020 research and innovation

programme under grant agreement no. 101004728; Helmholtz Innovation Pool project Data-X; and HIR3X—Helmholtz International Laboratory on Reliability, Repetition, Results at the most Advanced X-Ray Sources.

Acknowledgments

Thanks to members of the LEAPS consortium and LEAPS-INNOV project, particularly those who have presented and discussed their work on data reduction at events; this has been a valuable source of information in putting together this review. Thanks also to Thorsten Kracht (DESY) for providing feedback on the paper. The authors acknowledge support from DESY, a member of the Helmholtz Association HGF.

References

- Rao R. Synchrotrons face a data deluge. *Phys Today* (2020). doi:10.1063/PT.6.2.20200925a
- Marras A, Klugev A, Lange S, Laurus T, Pennicard D, Trunk U, et al. Development of CoRDIA: an imaging detector for next-generation photon science X-ray sources. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1047:167814. doi:10.1016/j.nima.2022.167814
- Allahgholi A, Becker J, Bianco L, Delfs A, Dinapoli R, Goettlicher P, et al. AGIPD, a high dynamic range fast detector for the European XFEL. *J Instrumentation* (2015) 10: C01023. doi:10.1088/1748-0221/10/01/C01023
- Trueb P, Sobott BA, Schnyder R, Loeliger T, Schneebeli M, Kobas M, et al. Improved count rate corrections for highest data quality with PILATUS detectors. *J Synchrotron Radiat* (2012) 19:347–51. doi:10.1107/S0909049512003950
- Hsieh SS, Iniewski K. Improving paralysis compensation in photon counting detectors. *IEEE Trans Med Imaging* (2021) 40:3–11. doi:10.1109/TMI.2020.3019461
- Mezza D, Becker J, Carraresi L, Castoldi A, Dinapoli R, Goettlicher P, et al. Calibration methods for charge integrating detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1024:166078. doi:10.1016/j.nima.2021.166078
- Blaj G, Caragiulo P, Carini G, Carron S, Dragone A, Freytag D, et al. X-Ray detectors at the linac coherent light source. *J Synchrotron Radiat* (2015) 22:577–83. doi:10.1107/S1600577515005317
- van Driel TB, Herrmann S, Carini G, Nielsen MM, Lemke HT. Correction of complex nonlinear signal response from a pixel array detector. *J Synchrotron Radiat* (2015) 22:584–91. doi:10.1107/S1600577515005536
- Redford S, Andr   M, Barten R, Bergamaschi A, Br  ckner M, Dinapoli R, et al. First full dynamic range calibration of the JUNGFR  U photon detector. *J Instrumentation* (2018) 13:C01027. doi:10.1088/1748-0221/13/01/C01027
- K  nnecke M, Akeroyd FA, Bernstein HJ, Brewster AS, Campbell SI, Clausen B, et al. The NeXus data format. *J Appl Crystallogr* (2015) 48:301–5. doi:10.1107/S1600576714027575
- The HDF Group. *Hierarchical data format version 5* (2000–2010).
- Sayood K. *Lossless compression handbook*. Elsevier (2002).
- Al-Shaykh OK, Mersereau RM. Lossy compression of noisy images. *IEEE Trans Image Process* (1998) 7:1641–52. doi:10.1109/83.730376
- Becker J, Greiffenberg D, Trunk U, Shi X, Dinapoli R, Mozzanica A, et al. The single photon sensitivity of the adaptive gain integrating pixel detector. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2012) 694:82–90. doi:10.1016/j.nima.2012.08.008
- Ballabriga R, Aloyo J, Campbell M, Frojdh E, Heijne E, Koenig T, et al. Review of hybrid pixel detector readout ASICs for spectroscopic X-ray imaging. *J Instrumentation* (2016) 11:P01007. doi:10.1088/1748-0221/11/01/P01007
- Broennimann C, Eikenberry EF, Henrich B, Horisberger R, Huelsen G, Pohl E, et al. The PILATUS 1M detector. *J Synchrotron Radiat* (2006) 13:120–30. doi:10.1107/S0909049505038665
- Deutsch LP. *DEFLATE compressed data format specification version 1.3*. No. 1951 in *request for comments* (RFC editor) (1996). doi:10.17487/RFC1951
- Hammer M, Yoshii K, Miceli A. Strategies for on-chip digital data compression for X-ray pixel detectors. *J Instrumentation* (2021) 16:P01025. doi:10.1088/1748-0221/16/01/P01025
- Leonarski F, Mozzanica A, Br  ckner M, Lopez-Cuenca C, Redford S, Sala L, et al. JUNGFR  U detector for brighter x-ray sources: solutions for IT and data science challenges in macromolecular crystallography. *Struct Dyn* (2020) 7:014305. doi:10.1063/1.5143480
- Gailly J, Adler M. *GZIP documentation and sources* (1993). available as gzip-*.tar in ftp://prep.ai.mit.edu/pub/gnu.
- Masui K, Amiri M, Connor L, Deng M, Fandino M, H  fer C, et al. A compression scheme for radio data in high performance computing. *Astron Comput* (2015) 12: 181–90. doi:10.1016/j.ascom.2015.07.002
- Ziv J, Lempel A. A universal algorithm for sequential data compression. *IEEE Trans Inf Theor* (1977) 23:337–43. doi:10.1109/tit.1977.1055714
- Huffman DA. A method for the construction of minimum-redundancy codes. *Proc IRE* (1952) 40:1098–101. doi:10.1109/jrproc.1952.273898
- Olsen U, Schmidt S, Poulsen H, Linnros J, Yun S, Di Michiel M, et al. Structured scintillators for x-ray imaging with micrometre resolution. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2009) 607:141–4. doi:10.1016/j.nima.2009.03.139
- Mittone A, Manakov I, Broche L, Jarnias C, Coan P, Bravin A. Characterization of a sCMOS-based high-resolution imaging system. *J Synchrotron Radiat* (2017) 24: 1226–36. doi:10.1107/S160057751701222X
- Skodras A, Christopoulos C, Ebrahimi T. The JPEG 2000 still image compression standard. *IEEE Signal Process. Mag* (2001) 18:36–58. doi:10.1109/79.952804
- Marone F, Vogel J, Stamparoni M. Impact of lossy compression of X-ray projections onto reconstructed tomographic slices. *J Synchrotron Radiat* (2020) 27: 1326–38. doi:10.1107/S1600577520007353
- Shensa MJ, et al. The discrete wavelet transform: wedding the a trous and Mallat algorithms. *IEEE Trans signal Process* (1992) 40:2464–82. doi:10.1109/78.157290
- Taubman D, Naman A, Smith M, Lemieux PA, Saadat H, Watanabe O, et al. High throughput JPEG 2000 for video content production and delivery over IP networks. *Front Signal Process* (2022) 2. doi:10.3389/frsip.2022.885644
- Huang P, Du M, Hammer M, Miceli A, Jacobsen C. Fast digital lossy compression for X-ray ptychographic data. *J Synchrotron Radiat* (2021) 28:292–300. doi:10.1107/S1600577520013326
- Di S, Cappello F. Fast error-bounded lossy HPC data compression with SZ. In: *2016 IEEE international parallel and distributed processing symposium*. IPDPS (2016). p. 730–9. doi:10.1109/IPDPS.2016.11
- Underwood R, Yoon C, Gok A, Di S, Cappello Froibin- SZ. ROIBIN-SZ: fast and science-preserving compression for serial crystallography. *Synchrotron Radiat News* (2023) 36:17–22. doi:10.1080/08940886.2023.2245722
- Barty A, Kirian RA, Maia FR, Hantke M, Yoon CH, White TA, et al. Cheetah: software for high-throughput reduction and analysis of serial femtosecond x-ray diffraction data. *J Appl Crystallogr* (2014) 47:1118–31. doi:10.1107/S1600576714007626
- Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, et al. *DIALS: implementation and evaluation of a new integration package*. *Acta Crystallogr Section D* (2018) 74:85–97. doi:10.1107/S2059798317017235
- Rahmani V, Nawaz S, Pennicard D, Setty SPR, Graafsma H. Data reduction for X-ray serial crystallography using machine learning. *J Appl Crystallogr* (2023) 56: 200–13. doi:10.1107/S1600576722011748

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

36. Ke TW, Brewster AS, Yu SX, Ushizima D, Yang C, Sauter NK. A convolutional neural network-based screening tool for X-ray serial crystallography. *J synchrotron Radiat* (2018) 25:655–70. doi:10.1107/s1600577518004873
37. Blaj G, Chang CE, Kenney CJ. Ultrafast processing of pixel detector data with machine learning frameworks. In: *AIP conference proceedings*, 2054. AIP Publishing (2019).
38. Chen L, Xu K, Zheng X, Zhu Y, Jing Y. Image distillation based screening for x-ray crystallography diffraction images. In: *2021 IEEE intl conf on parallel & distributed processing with applications, big data & cloud computing, sustainable computing & communications, social computing & networking (ISPA/BDCloud/SocialCom/SustainCom)*. IEEE (2021). p. 517–21.
39. Behar R, Mitrea DA, Vancea F, Marita T, Nedeveschi S, Lupsor-Platon M, et al. Comparison of deep-learning and conventional machine-learning methods for the automatic recognition of the hepatocellular carcinoma areas from ultrasound images. *Sensors* (2020) 20:3085. doi:10.3390/s20113085
40. Baehr S, Kempf F, Becker J. Data reduction and readout triggering in particle physics experiments using neural networks on FPGAs. In: *2018 IEEE 18th international conference on nanotechnology (IEEE-NANO)*. IEEE (2018). p. 1–4.
41. Ryd A, Skinnari L. Tracking triggers for the HL-LHC. *Annu Rev Nucl Part Sci* (2020) 70:171–95. doi:10.1146/annurev-nucl-020420-093547
42. Skambraks S, Abudínén F, Chen Y, Feindt M, Frühwirth R, Heck M, et al. A z-vertex trigger for Belle II. *IEEE Trans Nucl Sci* (2015) 62:1732–40.
43. Kieffer J, Wright J. PyFAI: a python library for high performance azimuthal integration on GPU. *Powder Diffraction* (2013) 28:S339–50. doi:10.1017/S0885715613000924
44. Matěj Z, Skovhede K, Johnsen C, Barczyk A, Weninger C, Salnikov A, et al. Azimuthal integration and crystallographic algorithms on field-programmable gate arrays. *Acta Crystallogr Section A* (2021) 77:C1185. doi:10.1107/S0108767321085263
45. Madden T, Niu S, Narayanan S, Sandy A, Weizeorick J, Denes P, et al. Real-time MPI-based software for processing of XPCS data. In: *2014 IEEE nuclear science symposium and medical imaging conference (NSS/MIC)* (2014). p. 1–5. doi:10.1109/NSSMIC.2014.7431129
46. Therrien AC, Herbst R, Quijano O, Gattón A, Coffee R. Machine learning at the edge for ultra high rate detectors. In: *2019 IEEE nuclear science symposium and medical imaging conference (NSS/MIC)* (2019). p. 1–4. doi:10.1109/NSS/MIC42101.2019.9059671
47. Kingma DP, Welling M. An introduction to variational autoencoders. *Foundations Trends Machine Learn* (2019) 12:307–92. doi:10.1561/22000000056
48. Utimula K, Yano M, Kimoto H, Hongo K, Nakano K, Maezono R. Feature space of XRD patterns constructed by an autoencoder. *Adv Theor Simulations* (2023) 6:2200613. doi:10.1002/adts.202200613
49. Blaschke JP, Wittwer F, Enders B, Bard D. How a lightsource uses a supercomputer for live interactive analysis of large data sets. *Synchrotron Radiat News* (2023) 36:10–6. doi:10.1080/08940886.2023.2245700
50. Nikitin V, Shevchenko P, Deriy A, Kastengren A, Carlo FD. Streaming collection and real-time analysis of tomographic data at the APS. *Synchrotron Radiat News* (2023) 36:3–9. doi:10.1080/08940886.2023.2245693
51. Asano S, Maruyama T, Yamaguchi Y. Performance comparison of FPGA, GPU and CPU in image processing. In: *2009 international conference on field programmable logic and applications (IEEE)* (2009). p. 126–31.
52. Vajda A. *Multi-core and many-core processor architectures*. Programming Many-Core Chips (2011). p. 9–43.
53. Khokhriakov S, Manumachu RR, Lastovetsky A. Performance optimization of multithreaded 2d fast fourier transform on multicore processors using load imbalancing parallel computing method. *IEEE Access* (2018) 6:64202–24. doi:10.1109/access.2018.2878271
54. Kirk D. NVIDIA CUDA software and GPU parallel computing architecture. *ISMM* (2007) 7:103–4.
55. Nvidia. *Vingelmann P, Fitzek FH. CUDA, release: 10.2.89* (2020).
56. Stone JE, Gohara D, OpenCL SG. OpenCL: a parallel programming standard for heterogeneous computing systems. *Comput Sci Eng* (2010) 12:66–73. doi:10.1109/mcse.2010.69
57. Lin JM. Python non-uniform fast fourier transform (PyNUFFT): an accelerated non-Cartesian MRI package on a heterogeneous platform (CPU/GPU). *J Imaging* (2018) 4:51. doi:10.3390/jimaging4030051
58. Becker D, Streit A. A neural network based pre-selection of big data in photon science. In: *2014 IEEE fourth international conference on big data and cloud computing*. IEEE (2014). p. 71–6.
59. Souza A, Oliveira LB, Hollatz S, Feldman M, Olukotun K, Holton JM, et al. *Deepfreak: learning crystallography diffraction patterns with automated machine learning* (2019). *arXiv preprint arXiv:1904.11834*.
60. Branco S, Ferreira AG, Cabral J. Machine learning in resource-scarce embedded systems, fpgas, and end-devices: a survey. *Electronics* (2019) 8:1289. doi:10.3390/electronics8111289
61. Choquette J, Gandhi W, Giroux O, Stam N, Krashinsky R. NVIDIA A100 tensor core GPU: performance and innovation. *IEEE Micro* (2021) 41:29–35. doi:10.1109/MM.2021.3061394
62. Abuowaimar Z, Maarouf D, Martin T, Foxcroft J, Gréwal G, Areibi S, et al. GPlace3.0: routability-driven analytic placer for Ultrascale FPGA architectures. *ACM Trans Des Automation Electron Syst (Todaes)* (2018) 23:1–33. doi:10.1145/3233244
63. Kathail V. Xilinx Vitis unified software platform. In: *Proceedings of the 2020 ACM/SIGDA international symposium on field-programmable gate arrays (New York, NY, USA: association for computing machinery)*. FPGA '20 (2020). p. 173–4. doi:10.1145/3373087.3375887
64. Perryman N, Wilson C, George A. Evaluation of Xilinx Versal architecture for next-gen edge computing in space. In: *2023 IEEE aerospace conference* (2023). p. 1–11. doi:10.1109/AERO55745.2023.10115906
65. Emani M, Xie Z, Raskar S, Sastry V, Arnold W, Wilson B, et al. A comprehensive evaluation of novel AI accelerators for deep learning workloads. In: *2022 IEEE/ACM international workshop on performance modeling, benchmarking and simulation of high performance computer systems (PMBS)* (2022). p. 13–25. doi:10.1109/PMBS56514.2022.00007
66. [Dataset] Kalamkar D, Mudigere D, Mellempudi N, Das D, Banerjee K, Avancha S, et al. *A study of BFLOAT16 for deep learning training* (2019).
67. Kraus M, Layad N, Liu Z, Coffee R. EdgeAI: machine learning via direct attached accelerator for streaming data processing at high shot rate x-ray free-electron lasers. *Front Phys* (2022) 10. doi:10.3389/fphy.2022.957509
68. Thayer J, Damiani D, Ford C, Dubrovin M, Gaponenko I, O'Grady CP, et al. Data systems for the linac coherent light source. *Adv Struct Chem Imaging* (2017) 3:3. doi:10.1186/s40679-016-0037-7
69. Wang C, Steiner U, Sepe A. Synchrotron big data science. *Small* (2018) 14:1802291. doi:10.1002/smll.201802291
70. Llopert X, Aloyz J, Ballabriga R, Campbell M, Casanova R, Gromov V, et al. Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning. *J Instrumentation* (2022) 17:C01044. doi:10.1088/1748-0221/17/01/C01044
71. Hügging F, Owtscharenko N, Pohl DL, Wermes N, Ehrmann O, Fritzsche T, et al. Advanced through silicon vias for hybrid pixel detector modules. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 936:642–3. doi:10.1016/j.nima.2018.08.067
72. Doering D, Kwiatkowski M, Kamath UR, Tamma C, Rota L, Ruckman L, et al. Readout system for ePixHR x-ray detectors: a framework and case study. In: *2020 IEEE nuclear science symposium and medical imaging conference (NSS/MIC)* (2020). p. 1–4. doi:10.1109/NSS/MIC42677.2020.9507754
73. Pennicard D, Smoljanin S, Pithan F, Sarajlic M, Rothkirch A, Yu Y, et al. LAMBDA 2M GaAs—a multi-megapixel hard x-ray detector for synchrotrons. *J Instrumentation* (2018) 13:C01026. doi:10.1088/1748-0221/13/01/C01026
74. B Valentin M, Di Guglielmo G, Noonan D, Dilip P, Huang P, Quinn A, et al. In-pixel AI for lossy data compression at source for X-ray detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168665. doi:10.1016/j.nima.2023.168665
75. Rota L, Perez AP, Habib A, Dragone A, Miceli A, Markovic B, et al. X-ray detectors for LCLS-II with real-time information extraction: the SparkPix family. In: *24th international workshop on radiation imaging detectors (IWORID 2023)*. Oslo, Norway (2023).
76. Bruckner M, Bergamaschi A, Cartier S, Dinapoli R, Frojdh E, Greiffenberg D, et al. *A multiple 10 Gbit Ethernet data transfer system for EIGER*. Padova, Italy: 20th IEEE Real Time Conference (2016).
77. Gottlicher P, Sheviakov I, Zimmer M. 10G-Ethernet prototyping for 2-D X-Ray detectors at the XFEL. In: *2009 16th IEEE-NPSS real time conference* (2009). p. 434–7. doi:10.1109/RTC.2009.5321620
78. Leonarski F, Brückner M, Lopez-Cuenca C, Mozzanica A, Stadler HC, Matěj Z, et al. Jungfrau-joch: hardware-accelerated data-acquisition system for kilohertz pixel-array X-ray detectors. *J Synchrotron Radiat* (2023) 30:227–34. doi:10.1107/S1600577522010268
79. Grimes M, Pauwels K, Schüll TU, Martin T, Fajardo P, Douissard PA, et al. Bragg coherent diffraction imaging with the CITIUS charge-integrating detector. *J Appl Crystallogr* (2023) 56:1032–7. doi:10.1107/S1600577523004314



OPEN ACCESS

EDITED BY

Jianguo Zhang,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY

Valeria Radicci,
Dectris AG, Switzerland
Kirsty Paton,
Paul Scherrer Institut (PSI), Switzerland

*CORRESPONDENCE

Paolo Busca,
✉ paolo.busca@esrf.fr

RECEIVED 29 September 2023

ACCEPTED 23 January 2024

PUBLISHED 13 February 2024

CITATION

Collonge M, Baussens O, Busca P, Fajardo P, Fischer P, Martin T, Ritzert M, Ruat M, Schimansky D and Williams M (2024), Probing the potential of CdZnTe for high-energy high-flux 2D X-ray detection using the XIDer incremental digital integrating readout. *Front. Phys.* 12:1304570. doi: 10.3389/fphy.2024.1304570

COPYRIGHT

© 2024 Collonge, Baussens, Busca, Fajardo, Fischer, Martin, Ritzert, Ruat, Schimansky and Williams. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Probing the potential of CdZnTe for high-energy high-flux 2D X-ray detection using the XIDer incremental digital integrating readout

Marin Collonge^{1,2}, Oriane Baussens¹, Paolo Busca^{1*}, Pablo Fajardo¹, Peter Fischer², Thierry Martin¹, Michael Ritzert², Marie Ruat¹, David Schimansky² and Morag Williams¹

¹Detector and Electronics Group, European Synchrotron Radiation Facility, Grenoble, France, ²Institut für Technische Informatik, Universität Heidelberg, Heidelberg, Germany

The latest synchrotron radiation sources have the capability to produce X-ray beams with a photon flux that can be up to three orders of magnitude higher than previous-generation facilities, and that are not manageable by the currently available 2D photon-counting pixel detectors. The construction of new detectors that exceed the limitations of existing devices is a critical strategic need. Developing such detectors is a challenge in terms of readout electronics as well as sensor material, particularly in the case of devices intended to operate at X-ray energies above 30 keV. The approach adopted at the ESRF to deal with this major difficulty is twofold: the use of a novel semiconductor material with improved electrical properties, high-flux CdZnTe, and the investigation of a specific readout scheme, incremental digital integration, via the XIDer project in collaboration with the University of Heidelberg. Incremental digital integration is a method intended to be less sensitive to variations of the dark current than the conventional charge integration readout. However, this readout scheme requires that the leakage current from the sensor material stays below a certain threshold to reduce the leakage contributions. This paper introduces the ESRF strategy and few examples of the methods employed to evaluate the performance and leakage current behavior of high-flux CdZnTe pixelated sensors. These examples illustrate the first results obtained with this material under moderate to very high X-ray irradiation fluxes of up to 10^{12} photons/mm²/s.

KEYWORDS

X-ray hybrid pixel detectors, charge integrating detectors, high-Z sensors, CdZnTe, high dynamic range, high-brilliance synchrotron beams

1 Introduction

1.1 Exploiting 4th generation synchrotron sources: limitations of current detectors

The proper exploitation of the latest generation of synchrotron radiation storage rings presents non-trivial challenges. In addition to the increased brilliance of these very low emittance accelerator-based sources, the possibility of operating with reduced gap undulators and the use of more efficient optics allow the delivery of X-ray beams of unprecedented intensity [1]. The actual gain in photon flux over previous sources is highly dependent on the particular applications and experimental setups, but in many cases can be as much as two to three orders of magnitude [2]. The photon fluxes are expected to increase further with the development of optimized insertion devices and improved X-ray optics. In this scenario, the construction of new X-ray detectors that exceed the limits of current devices by several orders of magnitude is a critical strategic need. In the case of the EBS (Extremely Brilliant Source) [3], the new ESRF storage ring and the first fourth-generation high-energy synchrotron facility in operation, the monochromatic X-ray beams can reach 10^{16} photons per second. In addition, the challenge of building a new generation of detectors is further increased due to the need to cover a significant number of applications with photon energies at or above 30 keV.

The most suitable 2D detectors currently available for diffraction and scattering experiments at the ESRF and other synchrotron storage rings are hybrid pixel detectors, built with pixelated semiconductor sensors, and operating in photon-counting mode [4, 5]. Thanks to the direct conversion of X-rays into charge, these devices provide the single-photon sensitivity required for photon-counting operation. And the signal-processing scheme in a photon counting device produces output data with extremely low readout noise that can be negligible in most practical cases [6, 7]. Another, less obvious, but extremely important feature of photon-counting readout is the possibility of engineering signal-processing electronics that are insensitive to fluctuations of the dark current generated in the semiconductor sensor. This aspect, which is relevant for silicon-based devices, becomes absolutely critical when it comes to high-energy photon detectors built with pixelated compound semiconductors. It has been instrumental in the successful deployment of 2D photon-counting CdTe detectors at the ESRF, as well as at a number of high-energy experimental stations at other synchrotron radiation facilities [8–10].

However, although photon-counting hybrid pixel technology is the current state of the art for 2D detection in a wide range of synchrotron applications, the maximum photon flux per pixel is limited by pulse pile-up [11]. When the average photon flux per pixel approaches a certain X-ray hit rate, the pulse discrimination circuitry begins to miss counts and the detector suffers from count losses. In practice, with the most advanced current photon-counting detectors, when the flux per pixel approaches few million photon hits per second, the pile-up count losses exceed 10% of the average incident hit rate [8, 12]. This is a major limitation of existing detectors, which prevents the efficient use of new generation of synchrotrons. Two different strategies are being investigated at the ESRF and other facilities to overcome the above limitation: (i) the design of new photon-counting detectors with one to two orders of

magnitude higher count rate capability, (ii) the implementation of special charge-integrating detectors suitable for time-resolved experiments and able to operate with continuous beams at high photon energies at storage rings.

1.2 The need and challenge of building high-energy charge-integrating detectors

It is possible to increase the count rate capabilities of photon-counting detectors by developing faster front-end readout electronics with pile-up compensation techniques or other strategies such as reducing the pixel size to increase the number of counting channels per pixel area [13–15]. However, the maximum achievable photon flux with photon-counting devices would still not be sufficient for applications dealing with very strong signals, such as those found in many diffraction experiments in material science. In these experiments, the samples under study are usually made of inorganic materials and the diffraction signals at the detector can be very intense. The energy of the X-ray photons must be high enough to penetrate and properly probe the regions of the samples being investigated. This results in illuminated sample volumes that are usually well above 100 μm in size, which are significantly larger than in other types of diffraction or scattering experiments with thin samples as it is the case, for instance, when studying organic or biological materials. This circumstance strongly influences the minimum required pixel size as, although the spatial distribution of the measured pattern is dependent on the distance from the sample to the active plane of the detector, the highest spatial resolution of the pattern is limited, among other parameters, by the size of the illuminated volume in the sample which is indeed the effective diffracting source. This consideration is not applicable to coherent scattering experiments, but it is fully valid for conventional diffraction techniques such as single-crystal or powder diffraction and has as a consequence that the required detector pixel size for this type of experiments is not very small. In practice, in most material science experiments, a pixel size in the range of 100 μm –200 μm is well adapted to properly record the diffraction patterns that can be very intense, reaching, or exceeding in some cases, 10^9 photons per second per pixel.

The photon statistics in experiments with such intense signals are sufficient to investigate physical phenomena on the micro- and sub-microsecond time scale, including irreversible processes. Such short time domains, that were unattainable with the previous generation of synchrotron sources, can be reached with detectors able to sample the diffraction patterns at MHz rates while coping with the high incident photon flux. This requirement corresponds quite closely to the capabilities of the high dynamic range charge-integrating hybrid pixel detectors developed over the last decade for free-electron lasers [16, 17] and more recently for spectroscopy applications [18]. These detectors are designed to operate by integrating the charge coming from the pixelated sensors during a very short time window for which the undesired signal component resulting from the integration of the sensor leakage current is negligible. Although these last generation X-ray integrating detectors are very well suited for very intense single pulse measurements or short integration times, they do not offer the versatility required for applications at storage rings, in which the duration of the integration intervals, the acquisition duty cycle and the operation conditions vary considerably from experiment to experiment.

Furthermore, the use of high-Z compound semiconductor sensor materials, which is essential to ensure good absorption efficiency above 30 keV, results in the degradation of the performance of detector systems due to charge trapping and to the associated polarization effects [19]. The polarization effects correspond to over-time changes in the electric field distribution in the sensor due to charge build-up. The build-up can be caused both by the bias applied to the sensor and by high incident X-ray doses. The bias-induced polarization effect specifically affects Schottky devices by screening the electric field near the cathode [20]. This not only results in a reduction of the charge collection efficiency, but also in a reduction of the Schottky barrier height, which causes the leakage current in the sensor to increase. Moreover, the bias-induced polarization effect can take up to hours to stabilize, which means that leakage current is unstable during that time. High and unstable values of leakage current make it challenging to implement an effective leakage current compensation scheme. And without a proper compensation, the leakage current is integrated along with the signal of interest. For intermediate or long integration times, the leakage dominates the signal of interest, thus degrading the dynamic range of the integrating system. As for the irradiation-induced polarization, it affects both ohmic and Schottky devices [21]. This effect depends on many factors, including the energy and intensity of the X-ray beam, the sensor type and the temperature. Like in the case of bias-induced polarization, the charge build-up created can screen the electric field in the sensor and hinder charge collection. Moreover, since in this case the charge build-up is tied to the irradiation, stopping the irradiation leads to the detrapping of the built-up charge. This results in a residual current, the aftersignal, that decreases over time with characteristic time constants linked to the traps levels implicated in the initial build-up. This aftersignal component, also referred to as “afterglow” or “lag” in the literature [22–24], is particularly detrimental for applications requiring pulsed irradiation. In the case of semiconductor compounds of the CdTe family, the polarization effect is caused by the very poor charge transport properties of holes and it makes conventional charge-integration readout schemes inapplicable in practice.

1.3 An ESRF approach to mitigate the leakage current contributions

As introduced in the previous section, a major challenge faced when building high-energy charge-integrating detectors for storage rings is finding a sufficiently effective method to reduce the contributions of the leakage current from the sensor material. The approach adopted follows two complementary paths. On the one side, there is the use of a semiconductor material with improved electrical properties and minimum leakage current variations: the so called high-flux CdZnTe (HF-CZT) [25, 26]. On the other side, the implementation of a specific readout scheme, the incremental digital integration, which uses quantization in the signal conversion process to partially cancel undesired contributions coming from the integration of the sensor leakage current. The investigation of incremental digital integration is a central part of the XIDeR project, a collaboration of the ESRF and the University of Heidelberg, and has been described and discussed in previous

publications [27, 28]. The results presented in this work are examples of the approach followed at ESRF for the evaluation of the behavior of HF-CZT sensors at high X-ray incident fluxes with both standalone discrete devices and with pixelated sensors hybridized with XIDeR readout ASICs.

2 Materials and methods

2.1 High-Z sensors and hybrids

HF-CZT material is a specific variant of CdZnTe, produced by Redlen Technologies [29], that has the particularity of presenting significantly enhanced hole transport properties compared to more conventional spectroscopic CZT and to CdTe [25, 26]. Thanks to this particularity, HF-CZT is much less subject to the build-up of positive space charge in the sensor volume and therefore less sensitive to both the bias-driven and radiation-driven polarization phenomena. Previous work has shown that, at 20 keV, HF-CZT is indeed more resilient than its standard counterpart, at least up to 10^{10} ph/mm²/s [30, 31]. Although the material used in this work was grown by Redlen Technologies, the samples under study were reprocessed by IMEM-CNR with their own contact technology and diced to produce sensors with the desired size and electrode layout. All the sensors are terminated with quasi-ohmic contacts at both sides and have a continuous cathode and a structured anode. The anodes are made of sputtered platinum electrodes with layouts that include pixelated structures surrounded by a full metal contact that is used as a guard ring. For operation, the cathodes of the sensors are biased to negative voltages while the guard rings are grounded and the pixel contacts at the anodes are connected to the measurement devices.

This work presents results obtained with two types of HF-CZT sensors: standalone samples measured with discrete readout electronics, and devices that were hybridized with the XIDeR readout ASICs.

The standalone samples studied in this paper are 1.5 mm thick and have a total external size, after reprocessing, of approximately 5×5 mm². The anode layout, as seen in Figure 1, includes two single pixels and a 2×2 pixel matrix surrounded by the full guard ring contact. All the pixels are 500 μ m square. On the cathode side, sample 1 has a gold electrode while sample 2 has a platinum electrode. These electrodes were manufactured via electroless and sputtering processes, respectively. The samples were mounted onto PCBs with bicomponent epoxy glue and bonded using copper wires and conductive silver epoxy (Figure 1). These samples were originally part of a larger study, testing all possible electrode configurations with gold and platinum [32]. In this previous study, it was found that a platinum electrode is needed on the anode (Cd-face) to act as a hole blocking layer, but that both gold and platinum act as electron blocking layers on the cathode (Te-face). Essentially, both samples discussed in the present study exhibited similar response under dark conditions. In addition, the response of these samples under irradiation has previously been published [31] for lower incident X-ray fluxes than those presented in this paper. It was found that both samples gave similar responses under irradiation.

In the current paper, all of the measurements presented were performed, at room temperature, on the single pixels. During measurements of each sample, they were placed in a custom-

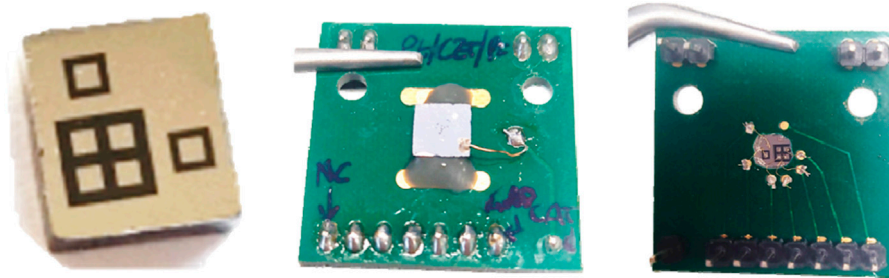


FIGURE 1
Standalone HF-CZT samples: Electrode layout and pictures of one of the samples mounted on the test PCB. [left] Pixelated anode layout. [center] Top side of the test PCB. [right] Bottom side of the test PCB.

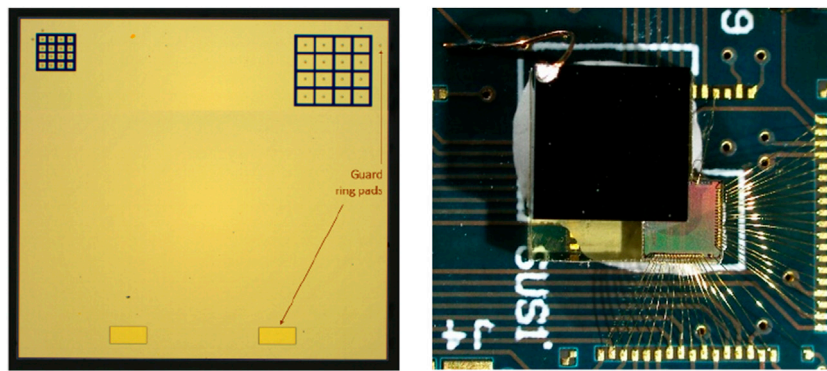


FIGURE 2
Hybridized HF-CZT devices: [left] $4 \times 4 \text{ mm}^2$ sensor before hybridization, showing the $100 \text{ }\mu\text{m}$ and $200 \text{ }\mu\text{m}$ pitch 16-pixel matrices designed to be bonded to XIDer ASICs. [right] Picture of the $200 \text{ }\mu\text{m}$ pitch sensor-ASIC assembly mounted on the test PCB.

made metal box with a Kapton tape entrance window and their cathodes were biased at -1000 V (6.7 kV/cm) using a high voltage power supply ISEG SHQ 224. The current collected by the single pixel was then measured using a picoammeter Keithley 6,485.

The devices with integrated readout are assemblies built with small HF-CZT pixelated sensors hybridized to XIDer readout ASICs. These readout chips are designed in CMOS 65 nm TSMC technology and implement the incremental digital integration readout [27, 28]. The HF-CZT sensors are 2 mm thick in this case and include two 4×4 pixel minimatrices of $100 \text{ }\mu\text{m}$ and $200 \text{ }\mu\text{m}$ pitch, as shown in the picture at the left of Figure 2. The cathode consists of a continuous platinum contact. In the assemblies used for this work only the $200 \text{ }\mu\text{m}$ minimatrix was connected to the 16-channel readout ASIC. The bonding was performed by Polymer Assembly Technology [33] by applying a low-temperature flip-chip process with gold studs and conductive epoxy [34]. The picture on the right shows one of the hybridized assemblies mounted on the test PCB with the connection to the top cathode electrode that, in operation, was biased at -1000 V .

For comparison purposes, the results obtained with the HF-CZT hybrids are presented in the next section together with results obtained with CdTe sensors from the company Acrorad [35]. These sensors are also terminated with ohmic platinum contacts, at both the anode and the cathode, and have the same size and layout

that is depicted in Figure 2. However, as the CdTe sensors are 1 mm thick, thinner than the HF-CZT devices, they were biased at a proportionally lower voltage of -500 V to operate with a comparable internal electric field.

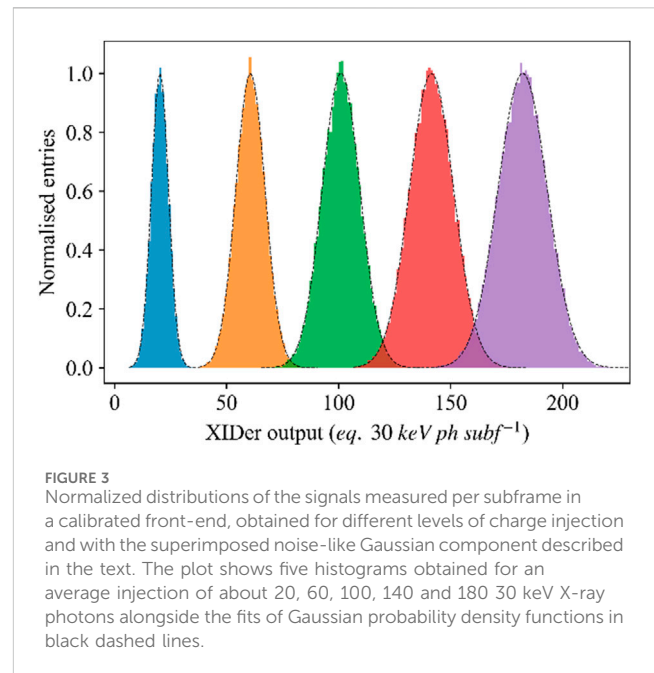
Note that the electric fields used to characterize the standalone samples (6.7 kV/cm) and the hybridized samples (5 kV/cm) are different. This difference comes from the fact that the measurements were planned separately for each type of samples. However, in both cases, the electric field was chosen to ensure that the samples were operated in full charge-collection regime. This regime occurs around 1 kV/cm for both CZT and CdTe samples.

2.2 Application of the incremental digital integration readout

The basic principle of incremental digital integration is the division of the total exposure time in a number N of subintervals, subframes in the XIDer terminology, where for each subframe the input signal in each pixel is integrated and digitized [27]. The resulting pixel intensity is the sum of the N digital values produced during the total integration time, an operation that is fully performed by the pixel electronics in the readout chip. This scheme, operating with very short subframes in the micro- or sub-

microsecond range, implements an in-pixel fast analog-to-digital conversion with a certain level of rejection of the dark current contribution. The rejection is effective if the total leakage charge collected during each subframe is well below a certain threshold that, in the case of the XIDer circuit implementation, is determined by the resolution of the analog-to-digital converter (ADC). When working with monochromatic radiation, the least significant bit of the ADC in the pixel is typically set to the equivalent of one X-ray photon and the threshold to 50% of that value. For example, with 30 keV monochromatic irradiation, the discrimination threshold would be set at an equivalent 15 keV, which in CdZnTe corresponds to a constant 260 pA/pixel leakage current being integrated during a 2 μ s window. The incremental digital integration readout then behaves at low photon fluxes like a photon-counting readout, and at higher photon fluxes, when quantization error becomes negligible, digital integration is functionally equivalent to a conventional charge-integrating scheme [36].

In the work presented in this paper, the XIDer readout ASIC was used as a tool to investigate the variation of the leakage current of the HF-CZT devices in the microsecond scale. For this purpose, the devices were calibrated to operate with 30 keV photons, the lowest photon energy they were designed to handle, and configured to acquire single subframes. The calibration is an important step which consists in trimming four adjustable parameters of the internal circuitry of the ASIC on a per-pixel basis. This includes the trimming of the quantization steps and the discrimination thresholds with respect to the energy of the incident photons in the pixel front-end electronics, for each of the two pipelined stages, coarse and fine [28]. The calibration procedure relies on the internal charge injection circuitry and each parameter was trimmed by analyzing the effect of a sweep of their respective DAC in the chip, under the injection of a specific amount of charge, *e.g.*, the charge equivalent to the integration of one 30 keV photon. The equivalence between photon energy and injected charge was derived from the theoretical sensor average ionizing energy, 4.64 eV per electron-hole pair for CZT [37], and the nominal values of the injection circuitry components. Discussing the procedure in more detail is beyond the scope of this paper but it is relevant to point out that, although the absolute calibration accuracy is limited by the manufacturing tolerance of the charge injection circuitry, it is possible and important to check the consistency of the trimming of the four internal parameters in each pixel. The proper relative matching of the pixel values can be assessed by using the charge injector itself as any deviation in the response of the front-end electronics reveals mismatches in the calibration of the coarse and fine stages. For that purpose, any arbitrary but controlled signal distribution can be injected into the front-end to verify the correspondence of the data produced by the ASIC. In the example below, a waveform generator was used to inject a signal corresponding to a predefined number of 30 keV photons with a superimposed Gaussian random component. For each number of equivalent photons, a large number of subframes was acquired, with the amplitude of the noise-like component adjusted to make the injection approximately match the signal distribution that would be produced by a Poissonian X-ray source illuminating uniformly the device. Examples of retrieved distributions of the acquired subframe values are shown in Figure 3 as normalized histograms alongside their fits to



Gaussian functions plotted with dashed black lines. Qualitatively, the Gaussian shapes of the input charge distributions are properly retrieved by the front-end electronics, and the resulting distributions closely match their respective Gaussian fit, with no apparent defects. This is a neat indication of the consistency in the calibration of the front-end.

2.3 Photocurrent generation by LED illumination

The study of the response of semiconductor sensors to strong irradiation transients in the microsecond regime requires a readout system able to operate in such a time scale but also the possibility of modulating the incident photon flux with very short switching times, ideally shorter than the measurement intervals. When using X-rays, the very intense photon fluxes required for the measurements cannot be produced by conventional laboratory sources and the only option is to use pulsed beams from either X-ray free-electron lasers or from synchrotron radiation storage rings operating in timing modes. During the investigation of sensor response for the XIDer project, the limited availability of suitable beamtime as well as the lack of flexibility in selecting the time structure of the irradiation patterns with synchrotron beams has motivated the development of an alternative scheme to generate photocurrent time patterns using visible light.

The scheme implemented is the use of a 730 nm LED device mounted to produce a uniform illumination over the entire active surface of the sensor, and whose intensity and temporal structure can be modulated electronically. At this wavelength, the transmission of the light through a 20 nm thick platinum contact is close to 4% and the photon energy is above the bandgap of detector grade CdZnTe (1.57 eV, 790 nm) [37]. Therefore, a small but non-negligible part of the light can penetrate in the

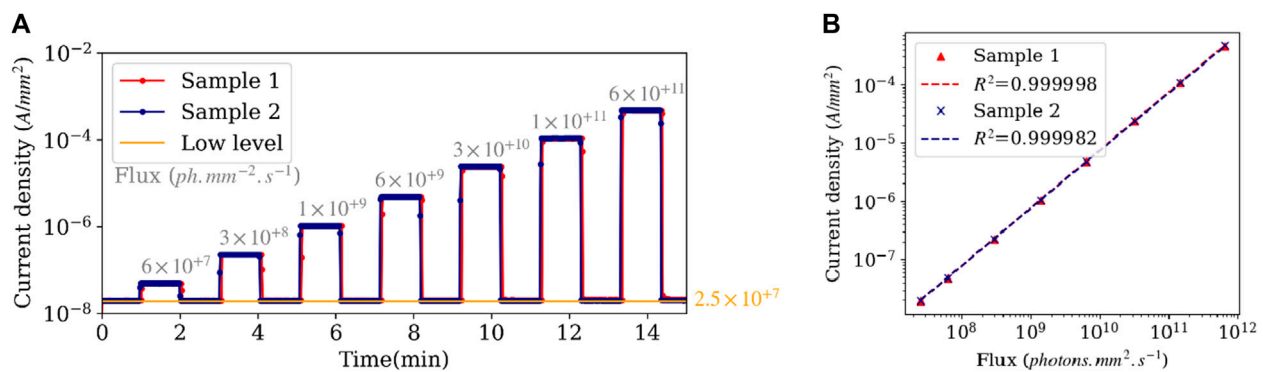


FIGURE 4

(A) Cycled irradiation of standalone HF-CZT samples for increasing incident X-ray fluxes. The measured current values are normalized by the pixel area to be expressed in A/mm². In grey, above each step, the corresponding incident X-ray flux in ph/mm²/s. In orange, the flux with highest beam attenuation that was used as the low-level reference of the irradiation cycles (B) Response of the samples with respect to the incident X-ray flux. This figure was produced with the data from (A). The dots correspond to the measured average intensity for the low-level reference and for each of the higher irradiation steps. The dotted lines correspond to the linear regression of the data with the resulting R^2 displayed in the legend.

semiconductor material through the thin cathode metal and generate electron-hole pairs near the surface of the device, the absorption length of 730 nm light in CdTe being approximately 0.5 μm [38]. These conditions differ from the photocurrent generation by X-rays because of the different penetration depth. However, given that, at 20 keV, the attenuation length in Cd(Zn)Te is on the order of 80 μm , the absorption of X-rays in the low end of the energy range of interest can still be considered rather superficial when compared to the 1–2 mm thickness of the devices under study. Moreover, other observations made by the authors during preliminary measurements, although with not properly calibrated devices, have shown that the high-Z sensors have a similar response in stability measurements under LED illumination and under low-energy X-ray irradiation, as well as similar aftersignal time components. Those observations support our assumption that 730 nm LED illumination can appropriately emulate the illumination by superficially-absorbed low-energy X-rays.

3 Results

3.1 Linearity and stability of the HF-CZT standalone samples

The linearity and stability under quasi-static high-flux irradiation of the two standalone HF-CZT samples were investigated at the BM05 beamline of the ESRF. A multilayer monochromator was used to produce a 20 keV monochromatic X-ray beam reaching a maximum flux of 10^{12} ph/mm²/s on the samples. The beam was collimated to match the pixel size (500 μm \times 500 μm) and aligned on one of the single pixels. The measurement consisted in subjecting the samples to 2-min long X-ray irradiation cycles with a progressively increasing maximum incident photon flux. A set of aluminum attenuators with 191 thickness combinations in steps of 60 μm was used to reduce the incident X-ray intensity and provide fast transitions between flux values. The attenuators could be changed remotely using a pneumatically-operated system with a response time of around 100 ms.

The incident flux as a function of the number of filters was obtained prior to measuring the samples from the photocurrent measured in a 500 μm thick silicon photodiode inserted in the beam path. The lowest flux achievable with this setup, using 11.46 mm of aluminum, was 2.5×10^7 ph/mm²/s (in orange in Figure 4), and it was taken as the reference of the irradiation cycles. The samples were subjected to seven cycles with flux values ranging from 6×10^7 to 6×10^{11} ph/mm²/s (1.5×10^7 to 1.5×10^{11} photons per pixel per second). Each cycle consisted of 1 minute of irradiation at a specific photon flux followed by 1 minute of irradiation at the low-level reference. During all the sequence, the photocurrent collected by the irradiated pixel was recorded every second using the picoammeter. As the irradiation sequences and the photocurrent measurements were not synchronized, there is usually one data point for each rising and falling edge that was measured during the change of attenuators. These data points are disregarded in the discussion. Figure 4A shows the measured current obtained for both standalone samples.

The two standalone devices exhibit similar behavior under cycled irradiation as well as an excellent linearity with the incident X-ray flux. This is depicted in Figure 4B, where the average measured intensity of the low-level reference and of each of the higher intensity irradiation steps are displayed as a function of the incident flux. In terms of stability, the rms variations of the measured photocurrent did not exceed the 0.4% of the mean value during the 1-min irradiation intervals, variations that are difficult to fully dissociate from the fluctuations of the intensity of the incident X-ray beam. Previous measurements, taken during longer intervals of 10 min, and for fluxes up to 10^{10} ph/mm²/s, have shown variations below 1 nA/mm² rms [31].

3.2 Transient response of HF-CZT standalone samples

The measurements in Figure 4A were not optimized to evaluate the transient response because the low-level reference did not correspond to the true dark current baseline and the recording

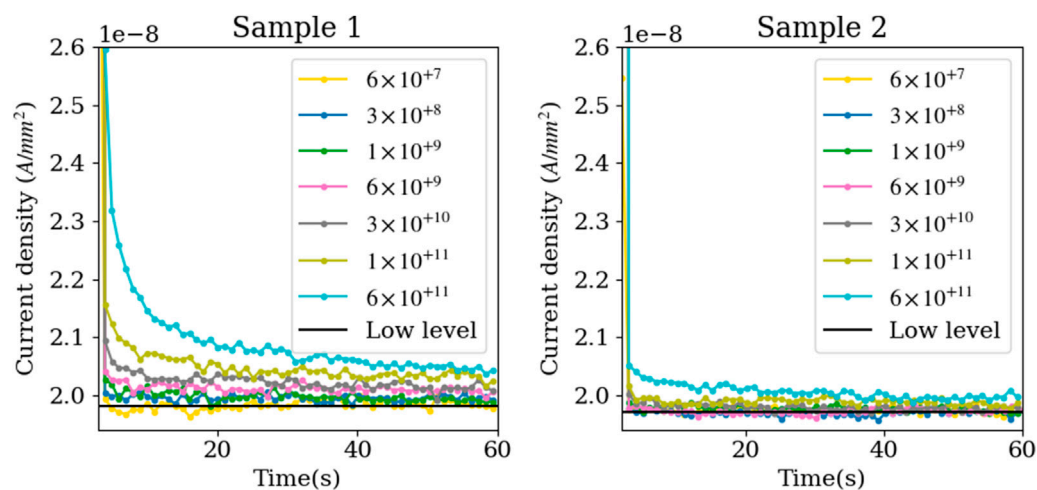


FIGURE 5

Zoom on the falling tails of the photocurrent steps presented in Figure 4A showing the aftersignal tails for sample 1 (left) and sample 2 (right). The flux values in the legends are given in $\text{ph}/\text{mm}^2/\text{s}$. The black horizontal line represents the low-level reference that corresponds to an incident flux of $2.5 \times 10^{17} \text{ ph}/\text{mm}^2/\text{s}$.

time resolution was limited to 1 s. However, these measurements still provide some useful information about the transient responses. Firstly, Figure 4A shows sharp rising edges for the pulses, regardless of the incident X-ray flux. Secondly, the falling edges exhibit some visible aftersignal, for instance after the last pulse of the sequence in the case of sample 1. The decay of the measured current for each sample is presented in more detail in Figure 5, in which the tails of the falling edges for all the irradiation levels are superimposed.

The data shown in Figure 5 gives a first insight into the aftersignal behavior of the sensors within the first 60 s after the end of the irradiation cycles. The aftersignal amplitude at a given time appears to increase with the incident X-ray intensity. In the case of sample 1, this behavior is clearly visible above $6 \times 10^9 \text{ ph}/\text{mm}^2/\text{s}$. In the case of sample 2, this behavior is only noticeable above $1 \times 10^{11} \text{ ph}/\text{mm}^2/\text{s}$.

Before discussing these results further, another set of data, measured 1 year before the data shown in Figure 5, will be introduced to give more elements for the analysis. This older dataset, that is already published [31], comes from a similar measurement campaign, conducted with the same HF-CZT samples, at the same X-ray energy of 20 keV. However, there were some differences in the way the experiment was carried out. Firstly, the irradiation times were longer, 10 and 20 min for sample 1 and sample 2 respectively. Secondly, the low level of the cycles corresponded to the dark current baseline, which was about $100 \text{ pA}/\text{mm}^2$. This was feasible because the maximum flux used for those measurements, $8 \times 10^9 \text{ ph}/\text{mm}^2/\text{s}$, was two orders of magnitude lower than in the case of Figure 5, so that the strongest attenuation (using 11.46 mm of aluminum) corresponded to dark conditions. And finally, the photocurrent was sampled at a 4 Hz rate. In the Supplementary Figure S1 displays the photocurrent and aftersignal tails of both datasets side by side to facilitate the comparison between the two.

The superimposed tails of the current decay of the samples under cycled irradiation for the older dataset [31] are displayed in Figure 6. All the aftersignal values in the plots of the figure are the

result of subtracting the dark current component from the measured current.

The results of Figure 6 are similar for both samples. Aftersignal tails are visible and the aftersignal amplitude at a given time increases with the intensity of the preceding X-ray irradiation. The aftersignal presents a fast component, that cannot be measured with the low time resolution of this setup but is the subject of the following section, and a slow component with a decay time in the range of seconds. In order to illustrate the dependency of the slow component of the aftersignal with the irradiation, Figure 7 presents the levels observed in Figure 6 for both samples as a function of the photon flux. The values in Figure 7A are the aftersignal levels obtained 1 s and 60 s after the end of the irradiation cycle, while the chart in Figure 7B compares the aftersignal 1 s after the end of the irradiation with the average photocurrent measured during the irradiation interval.

The measured aftersignal tends to increase with the incident X-ray flux. The aftersignal values at 1 s for sample 2 are consistently smaller than for sample 1. However, this difference subsides, and is even reversed, for longer times. Even if the aftersignal amplitude increases with incident X-ray flux, the ratio between the aftersignal and the photocurrent produced during the irradiation phase remains small. After 1 s, this ratio is below 0.15%, even for the most intense irradiation conditions. And, after 60 s, the ratio drops below 0.01%.

Going back to Figure 5, we observe much larger aftersignal values for sample 1 than for sample 2. A possible explanation could lie in the difference of cathode contacts. Even if both gold and platinum act as electron blocking contacts, the Schottky barriers might react differently to the density of charge created by the irradiation near the cathode [21]. Following this line of thought, a medium-flux irradiation could introduce the small difference observed in the response of the samples in Figure 6 and Figure 7, and a high-flux irradiation would amplify this difference further, as observed in Figure 5. However, there is an alternative explanation for the large difference observed between samples in Figure 5, and it resides in the history of the samples. As

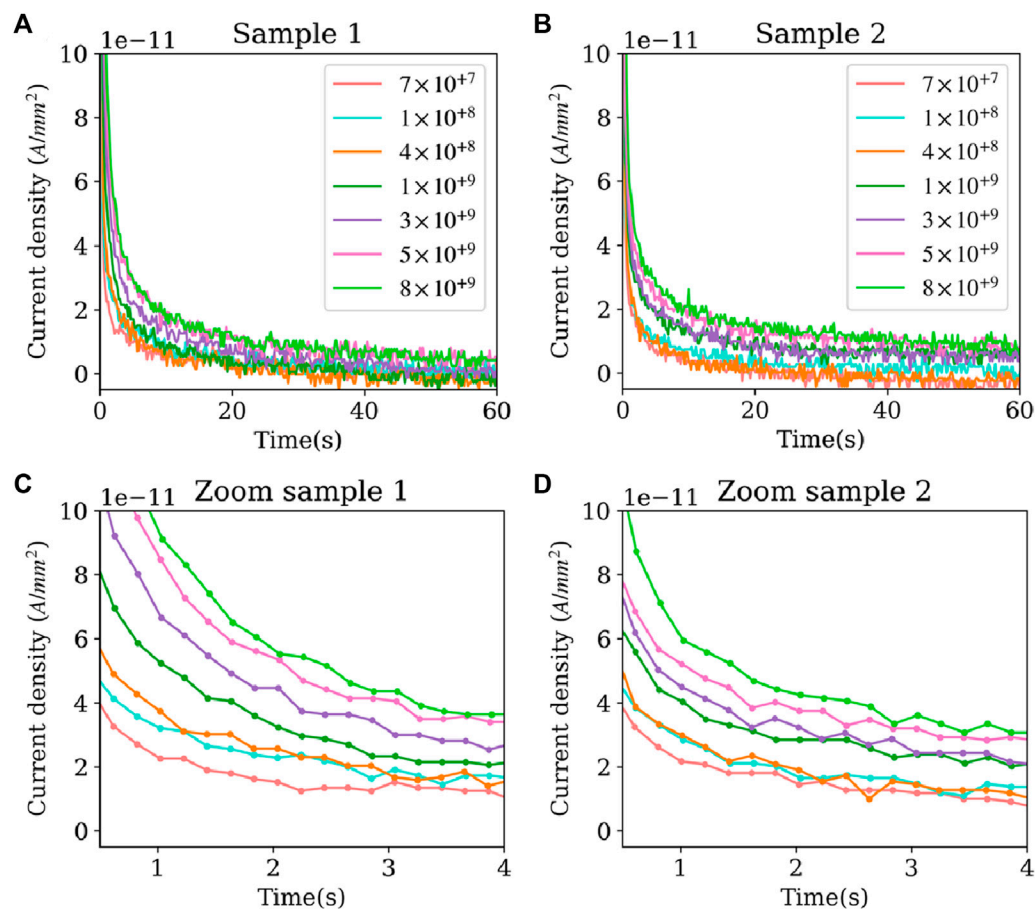


FIGURE 6
Baseline corrected aftersignal tails of the photocurrent measured in the two standalone HF-CZT samples after the cycled irradiation sequences described in [31] for different photon fluxes, in the legend, in $\text{ph}/\text{mm}^2/\text{s}$. (C) and (D) are zoomed in plots of figure (A) and (B) respectively.

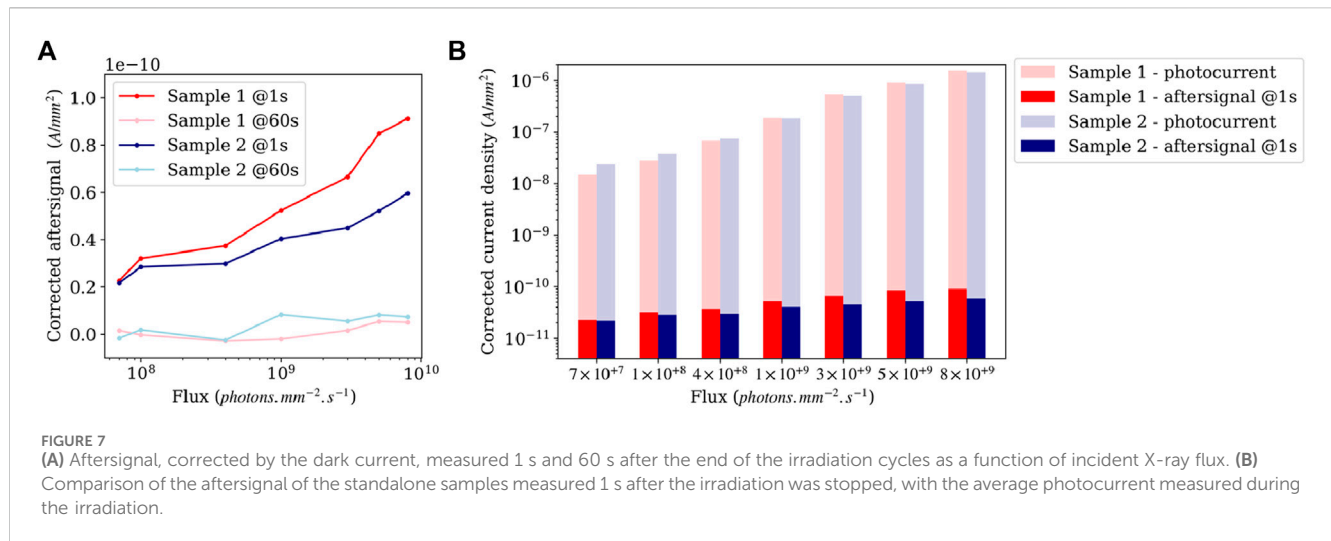
mentioned above, the dataset used in Figure 6 was measured 1 year prior to the one used in Figure 5. During this time period, sample 1 suffered an electrical breakdown on one of its corners. This corner was sanded down to remove the damaged part. Following this, the sensor remained functional but its transient response might have been affected by the introduction of new mechanically induced trapping centers. It is worth noting, however, that the preservation of the functionality of the sensor after such a crude removal of the damaged area can be taken as a strong indication of the resilience of the material.

To finish with, in Figure 5, no aftersignal is visible below $6 \times 10^9 \text{ ph}/\text{mm}^2/\text{s}$ in the case of sample 1, and below $1 \times 10^{11} \text{ ph}/\text{mm}^2/\text{s}$ in the case of sample 2. This is due to a combination of two factors. Firstly, there is a significant difference between the transients observed in Figure 5 and those observed in Figure 6. On the one hand, in Figure 5, the transients correspond to the transition between a trapping-detrapping equilibrium at high flux and a trapping-detrapping equilibrium at a lower flux because, as previously mentioned, the low-level reference is not the dark current baseline. On the other hand, in Figure 6, the transients correspond to the transition between a trapping-detrapping equilibrium under high flux and the equilibrium under dark

conditions. This means that the trap states being emptied to reach equilibrium might not be the same, or at least, the concentration of traps emptied might not be the same. This could explain the faster transient response observed in Figure 5 compared with Figure 6 at similar incident X-ray fluxes. Secondly, because the low-level reference in Figure 5 is three orders of magnitude higher than the dark current baseline, the lowest picoammeter measurement range that could be used for the current measurement was 20 nA. This measurement range has an accuracy of 0.4%, so the smallest variations that could be recorded were $3 \times 10^{-10} \text{ A}/\text{mm}^2$. This is too large to measure any of the aftersignal responses observed in Figure 6.

3.3 Fast response of hybridized detectors

This section presents an example that illustrates the implementation of the methods introduced in Section 2 and their application to the evaluation of the fast components of the photocurrent aftersignal transients of HF-CZT and CdTe pixelated sensors. For this purpose, the following results were obtained from 200 μm pixel pitch devices connected to XIDer readout ASICs as described in 2.3 and the measurements were



performed by illuminating the assemblies with LED light emulating low-energy X-rays as discussed in 2.4. As the XIDer devices were carefully pre-calibrated to provide the value of the integrated charge as an integer number of 30 keV equivalent photons, the adopted approach is to express the LED illumination also in the same units by assuming that under continuous illumination conditions, the readout values provided by XIDer are a good representation of the equivalent incident flux. This assumption would not be fully valid to provide accurate absolute flux values if the charge generation or the collection efficiency in the sensors is not ideal, but it can be taken as an adequate approximation to investigate the dynamic performance of the leakage current and the aftersignal effects. Therefore, despite the use of visible light, all intensity values for both incident and measured intensities in this example are referenced to 30 keV equivalent photons.

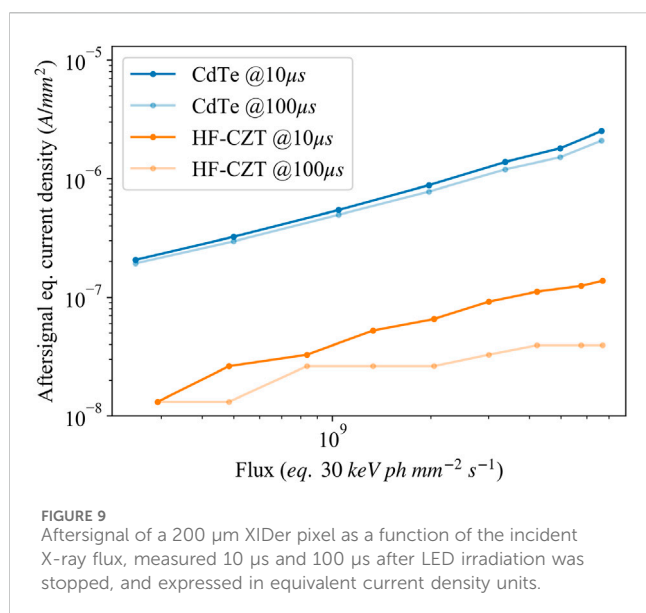
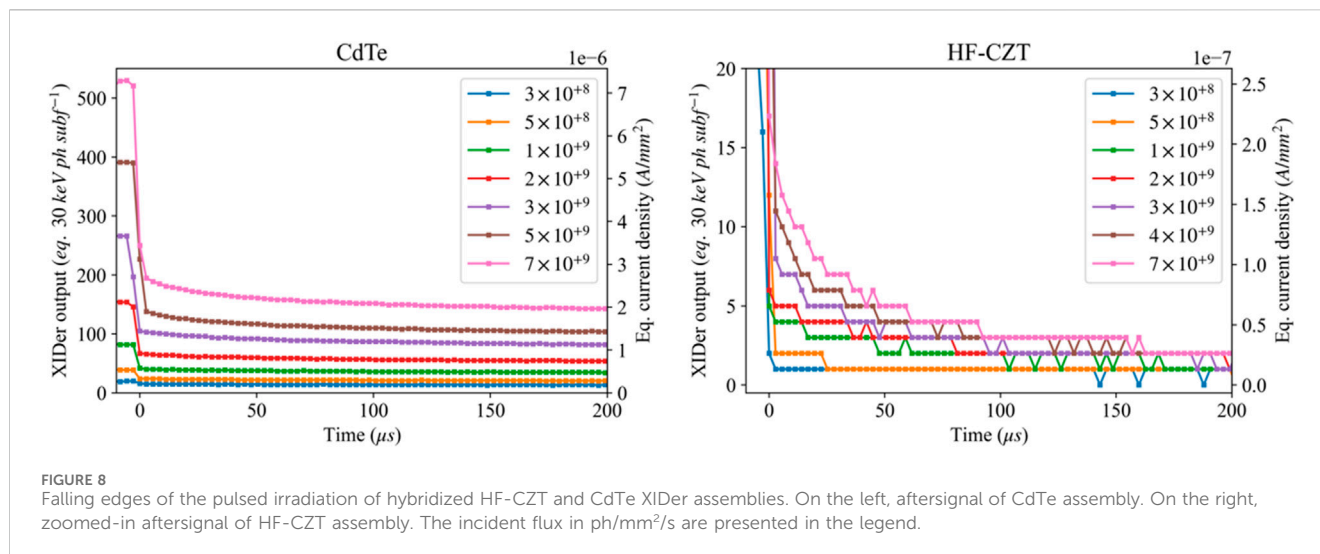
During the measurements, the assemblies were illuminated with a periodic 5 Hz square waveform of 50% duty cycle that corresponds to light pulses of 100 ms duration. The XIDer readout was set to acquire single subframes with a repetition period of 2.8 μs . This value is somehow arbitrary for the purposes of these measurements but it corresponds to the period of the orbit of the ESRF storage ring and is one of the standard subframe times foreseen for the regular final operation of XIDer at the ESRF. In the conditions of these measurements, the effective integration window within each 2.8 μs subframe period was of 1.97 μs , followed by a dead time of 0.83 μs . The LED source was independently characterized with a silicon photodiode to verify the proper shape of the light waveform and that the light emission fully decayed to zero in less than 1 μs at the falling edges of the illumination cycles.

Figure 8 shows the falling edges of the sensor current measured with the CdTe and HF-CZT assemblies under the pulsed illumination, for several light intensities ranging from 20 to 530 equivalent 30 keV photons per pixel per subframe. For the given pixel pitch and integration window, these values correspond to fluxes from 3×10^8 to 7×10^9 ph/mm²/s. The figure shows the response of one of the 200 μm pixels of each assembly once the

steady state has been reached after several cycles of pulsed illumination. The plots present the pixel signal integrated for each 2.8 μs subframe during the time interval immediately following the falling edge of the incident illumination. They illustrate the level and temporal evolution of the aftersignal from the sensors. The vertical axis in the plot for the HF-CZT assembly has been expanded to reveal the aftersignal decay, as the measured values are between one and two orders of magnitude lower than in the case of CdTe. Because the subframe acquisition was not synchronized with the periodic 5 Hz illumination pulses, the first subframe recorded after the initial signal drop is likely to include partial illumination and should be disregarded.

The weaker HF-CZT aftersignals drop with a much smaller time constant when compared to CdTe. After a few hundred microseconds, the HF-CZT aftersignal integrated during one subframe drops below the 15 keV equivalent photon preset discrimination threshold, corresponding to half the least significant bit photon level, resulting in an effective null readout value. This coincidentally illustrates how the analog-to-digital conversion in the XIDer readout can completely discard residual signals by exploiting the intrinsic quantization of the X-ray illumination. This is the key feature of the incremental digital integration readout that makes possible in-pixel aggregation of multiple subframes without accumulating leakage current and similarly undesired contributions, such as the aftersignal, as long as they remain below the discrimination threshold.

The variation with the intensity of the illumination and the differences between sensor materials is more clearly appreciated in Figure 9, which presents the aftersignal level measured 10 and 100 μs after the end of the LED illumination pulses as a function of the injected signal expressed as equivalent X-ray flux. For comparison with the results in Figure 7, the measured aftersignal values in the vertical axis have been converted into equivalent current density units by using the pair creation energies tabulated in [13]. Please note again that the electric fields used for the standalone samples and the hybridized samples are different due to differences in thickness. This has to be considered when comparing the results from Figure 7, 9.



4 Summary and outlook

This paper introduces the strategy followed at the ESRF for the development of 2D X-ray detectors suitable for experiments with the high brilliance beams produced by the last generation of synchrotron storage rings such as the ESRF-EBS source. The most demanding applications require charge-integrating detectors able to operate at high X-ray energies up to 100 keV and with photon fluxes that may reach 10^{11} ph/mm²/s at the detector. A major challenge for building such a type of detectors is the proper management of the fluctuations and instabilities of the leakage current from the compound semiconductor sensors required to operate with high-energy X-rays. The presented approach is based on the use of the incremental digital integration readout scheme that is being developed within the XIDer project. This type of readout can mitigate the impact of sensor leakage current variations, as long they remain under certain limits [27, 28]. The availability of high-Z sensors with good performance under intense photon fluxes and

with low leakage current is therefore essential, and today, high-flux CZT is the only material candidate to fulfil those requirements.

The characterization of HF-CZT, in particular its ability to operate with high X-ray fluxes and the behavior of the leakage current is a crucial aspect. Understanding and quantifying the effects of the irradiation on the sensor performance including effects such as leakage or aftersignal variations are fundamental. The paper presents examples of the work recently started at the ESRF with HF-CZT sample sensors and some first promising results that confirm the much better performance of this material with respect to CdTe as it has already been observed under different conditions in several other published works [26]. The examples also illustrate how the XIDer readout can be used as a tool to evaluate the leakage variations in the micro- and sub-microsecond time scale. The results show good linearity and stability when the HF-CZT sensors are irradiated under quasi-static conditions with very high incident fluxes, well above 10^{11} ph/mm²/s, showing no apparent drop in charge collection efficiency. The measurements also reveal the presence of weak but observable aftersignal components, i.e., a decaying residual current when the incident illumination on the sensor is sharply stopped or attenuated.

Our study is at an early stage, the measurements with HF-CZT have so far been performed at relatively low X-ray energies (20 keV) or with visible light illumination, conditions in which the charge carriers are generated close to the surface of the devices. An important aspect will be the investigation of the behavior of the HF-CZT sensors under higher-energy irradiation, up to 100 keV, which penetrates deeper into the material. A good understanding of all the dark or leakage components, not just the aftersignal, is essential. Some preliminary observations, still under investigation and that we plan to present in future publications, suggest an increase of the sensor leakage current that is induced by the irradiation itself, which adds to the generated photocurrent. If this effect is confirmed, the aftersignal component could be wholly or partly due to this additional leakage current component rather than only to charge detrapping processes in the sensor. Another important aspect that needs thorough study is the impact of the bias voltage and the operating temperature. So far, all the measurements have been carried out with HF-CZT sensors at room temperature. With the present knowledge of the leakage and charge

transport processes in these devices, it is difficult to predict the temperature dependence of the sensor response. However, it is not excluded that one could find operating conditions with significantly improved sensor performance.

Therefore, the deeper investigation of the HF-CZT sensors and a better understanding of their electro-optical response under high incident flux conditions is a key objective as it will have a decisive impact on the final design, construction, and operation of charge-integrating X-ray hybrid pixel detectors for high-energy applications with intense photon beams. In a more general way, the establishment by the detector community of a strong knowledge base on materials suitable for high-energy X-rays and high-flux applications is absolutely essential to overcome the challenges posed by the latest generation of synchrotron radiation sources. We expect that the ongoing work on the characterization of HF-CZT within the XIDER project will contribute in this direction.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

MC: Writing–original draft, Writing–review and editing, Conceptualization, Methodology, Visualization, Validation, Software, Formal analysis, Data curation, Investigation. OB: Writing–original draft, Writing–review and editing, Conceptualization, Methodology, Visualization, Validation, Software, Formal analysis, Data curation, Investigation. PB: Writing–review and editing, Funding acquisition, Project administration, Supervision, Resources, Conceptualization, Methodology, Visualization, Validation, Software, Investigation. PFa: Writing–original draft, Writing–review and editing, Funding acquisition, Project administration, Supervision, Conceptualization, Visualization, Validation. PFi: Writing–review and editing, Funding acquisition, Project administration, Supervision, Conceptualization. TM: Writing–review and editing, Funding acquisition, Project administration, Supervision. MRi: Writing–review and editing, Software, Resources. MRu: Writing–review and editing, Funding acquisition, Project administration, Supervision, Resources, Conceptualization, Methodology, Visualization, Validation. DS:

Writing–review and editing, Software, Resources. MW: Writing–review and editing, Methodology.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to acknowledge Phil Cook for his support at the ESRF BM05 beamline, Silvia Zanettini of Due2Lab for her expertise in the reprocessing of the HF-CZT samples, and James E. Clayton of Polymer Assembly Technology Inc. for his contribution to the project in bonding the hybridized assemblies.

Conflict of interest

Author MC, OB, PB, PFa, TM, MRu, and MW were employed by Detector and Electronics Group.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphy.2024.1304570/full#supplementary-material>

References

- Susini J, Barrett R, Chavanne J, Fajardo P, Götz A, Revol JL, et al. New challenges in beamline instrumentation for the ESRF upgrade programme phase II. *J Synchrotron Rad* (2014) 21(5):986–95. doi:10.1107/S1600577514015951
- Farvacque L. *EBS storage ring technical report*. France: ESRF (2018).
- Raimondi P. ESRF-EBS: the extremely brilliant source project. *Synchrotron Radiat News* (2016) 29(6):8–15. doi:10.1080/08940886.2016.1244462
- Rossi L, Fischer P, Rohe T, Wermes N. Pixel detectors: from fundamentals to applications. In: *Particle acceleration and detection*. Berlin, Heidelberg: Springer (2006). doi:10.1007/3-540-28333-1
- Brönnimann C, Trüb P. Hybrid pixel photon counting X-ray detectors for synchrotron radiation. In: Jaeschke EJ, Khan S, Schneider JR, Hastings JB, editors. *Synchrotron light sources and free-electron lasers: accelerator Physics, instrumentation and science applications*. Cham: Springer International Publishing (2020). p. 1191–223. doi:10.1007/978-3-030-23201-6_36
- Ballabriga R, Blaj G, Campbell M, Fiederle M, Greiffenberg D, Heijne EHM, et al. Characterization of the Medipix3 pixel readout chip. *J Inst* (2011) 6(01):C01052. doi:10.1088/1748-0221/6/01/C01052
- Maj P, Taguchi T, Nakaye Y. Tests of UFXC32k chip with CdTe pixel detector. *J Inst* (2018) 13(02):C02014. doi:10.1088/1748-0221/13/02/C02014
- Donath T, Šišak Jung D, Burian M, Radicci V, Zambon P, Fitch AN, et al. EIGER2 hybrid-photon-counting X-ray detectors for advanced synchrotron diffraction experiments. *J Synchrotron Rad* (2023) 30(4):723–38. doi:10.1107/S160057752300454X
- Medjoubi K, Hustache S, Picca F, Bélar JF, Boudet N, Bompard F, et al. Performance and Applications of the CdTe- and Si-XPAD3 photon counting 2D detector. *J Inst* (2011) 6(01):C01080. doi:10.1088/1748-0221/6/01/C01080

10. Tsigaridas S, Ponchut C. X-ray imaging with high-Z sensors for the ESRF-EBS Upgrade. *J Inst* (2019) 14(04):C04009. doi:10.1088/1748-0221/14/04/C04009
11. Becker W. Advanced time-correlated single photon counting techniques. In: *Springer series in chemical Physics*, 81. Berlin, Heidelberg: Springer (2005). doi:10.1007/3-540-28882-1
12. Loeliger T, Brönnimann C, Donath T, Schneebeli M, Schnyder R, Trüb P. The new PILATUS3 ASIC with instant retrigger capability. In: 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC); October 29–November 3, 2012; Anaheim, California, USA (2012). p. 610–5. doi:10.1109/NSSMIC.2012.6551180
13. Andrä M, Dinapoli R, Bergamaschi A, Barten R, Brückner M, Chirioti Alvarez S, et al. Towards MYTHEN 3: characterization of prototype chips. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 936:383–5. doi:10.1016/j.nima.2018.11.026
14. Grybos P, Kleczek R, Kmon P, Otfinowski P, Fajardo P, Magalhães D, et al. SPHIRD—Single photon counting pixel readout ASIC with pulse pile-up compensation methods. *IEEE Trans Circuits Syst* (2023) 70(9):3248–52. doi:10.1109/TCSII.2023.3267859
15. Magalhães D, Fajardo P, Grybos P, Kleczek R, Ruat M. Very high rate X-ray photon counting 2D detectors with small pixels - the SPHIRD project. *IEEE NSS-MIC-RTSD Conf Proc* (2022). doi:10.1109/NSS/MIC44845.2022.10399295
16. Veale MC, Adkin P, Booker P, Coughlan J, French M, Hart M, et al. Characterisation of the high dynamic range Large Pixel Detector (LPD) and its use at X-ray free electron laser sources. *J Inst* (2017) 12(12):P12003. doi:10.1088/1748-0221/12/12/P12003
17. Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Rad* (2019) 26(1):74–82. doi:10.1107/S1600577518016077
18. Cline BD, Banks D, Bell S, Church I, Cross S, Davis A, et al. Characterisation of HEXITECMHz - a 1 MHz continuous frame rate spectroscopic X-ray imaging detector system. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168718. doi:10.1016/j.nima.2023.168718
19. Grill R, Belas E, Franc J, Bugar M, Uxa S, Moravec P, et al. Polarization study of defect structure of CdTe radiation detectors. *IEEE Trans Nucl Sci* (2011) 58(6):3172–81. doi:10.1109/TNS.2011.2165730
20. Cola A, Farella I. The polarization mechanism in CdTe Schottky detectors. *Appl Phys Lett* (2009) 94(10):102113. doi:10.1063/1.3099051
21. Cola A, Farella I. CdTe X-ray detectors under strong optical irradiation. *Appl Phys Lett* (2014) 105(20):203501. doi:10.1063/1.4902188
22. Ricq S, Glasser F, Garcin M. CdTe and CdZnTe detectors behavior in X-ray computed tomography conditions. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2000) 442(1):45–52. doi:10.1016/S0168-9002(99)01198-5
23. Overdick M, Baumer C, Engel KJ, Fink J, Herrmann C, Kruger H, et al. Status of direct conversion detectors for medical imaging with X-rays. *IEEE Trans Nucl Sci* (2009) 56(4):1800–9. doi:10.1109/TNS.2009.2025041
24. DECTRIS. *DECTRIS EIGER2 CdTe Technical specifications* (2018). Available: https://media.dectris.com/BR_EIGER2_X_CdTe_Jun2018_web.pdf (Accessed November 16, 2023).
25. Iniewski K. CZT sensors for Computed Tomography: from crystal growth to image quality. *J Inst* (2016) 11(12):C12034. doi:10.1088/1748-0221/11/12/C12034
26. Thomas B, Veale MC, Wilson MD, Seller P, Schneider A, Iniewski K. Characterisation of redlen high-flux CdZnTe. *J Inst* (2017) 12(12):C12045. doi:10.1088/1748-0221/12/12/C12045
27. Fajardo P, Busca P, Erdinger F, Fischer P, Ruat M, Schimansky D. Digital integration: a novel readout concept for XIDER, an X-ray detector for the next generation of synchrotron radiation sources. *J Inst* (2020) 15(01):C01040. doi:10.1088/1748-0221/15/01/C01040
28. Schimansky D, Busca P, Erdinger F, Fajardo P, Fischer P. Concepts for the XIDER readout ASIC incorporating a pipelined ADC with very low dead time. *J Inst* (2021) 16(03):P03023. doi:10.1088/1748-0221/16/03/P03023
29. Redlen. *Redlen Technologies website* (2023). Available: <https://www.redlen.com/>.
30. Koch-Mehrin KAL, Bugby SL, Lees JE, Veale MC, Wilson MD. Charge sharing and charge loss in high-flux capable pixelated CdZnTe detectors. *Sensors* (2021) 21(9):3260. doi:10.3390/s21093260
31. Baussens O, Ponchut C, Ruat M, Bettelli M, Zanettini S, Zappettini A. Characterization of High-Flux CdZnTe with optimized electrodes for 4th generation synchrotrons. *J Inst* (2022) 17(11):C11008. doi:10.1088/1748-0221/17/11/C11008
32. Bettelli M, Zanettini S, Abbene L, Casoli F, Nasi L, Trevisi G, et al. High performance platinum contacts on high-flux CdZnTe detectors. *Sci Rep* (2023) 13(1):17963. Art. no. 1. doi:10.1038/s41598-023-45331-9
33. Polymer Assembly Technology. *Polymer assembly technology website* (2023). Available: <https://www.polymerassemblytech.com/>.
34. Clayton JE, Chen CMH, Cook WR, Harrison FA. Assembly technique for a fine-pitch, low-noise interface; Joining a CdZnTe pixel-array detector and custom VLSI chip with Au stud bumps and conductive epoxy. In: 2003 IEEE Nuclear Science Symposium. Conference Record; 19–25 October, 2003; Portland, Oregon, USA, 5 (2003). p. 3513–7. doi:10.1109/NSSMIC.2003.1352668
35. Acrorad. *Acrorad website* (2023). Available: https://www.acrorad.co.jp/index_en.html.
36. Collonge M, Busca P, Fajardo P, Williams M. Monte Carlo simulations for XIDER, a novel digital integration X-ray detector for the next generation of synchrotron radiation sources. *J Inst* (2022) 17(01):C01037. doi:10.1088/1748-0221/17/01/C01037
37. Owens A, Peacock A. Compound semiconductor radiation detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2004) 531(1):18–37. doi:10.1016/j.nima.2004.05.071
38. Cola A, Farella I, Anni M, Martucci MC. Charge transients by variable wavelength optical pulses in CdTe nuclear detectors. *IEEE Trans Nucl Sci* (2012) 59(4):1569–74. doi:10.1109/TNS.2012.2194509



OPEN ACCESS

EDITED BY

Cornelia B. Wunderer,
Helmholtz Association of German Research
Centres (HZ), Germany

REVIEWED BY

Xiangyu Xie,
Paul Scherrer Institut (PSI), Switzerland
Shuming Jiao,
Peng Cheng Laboratory, China
Leszek Grzanka,
Polish Academy of Sciences, Poland

*CORRESPONDENCE

R. T. Chen,
✉ chenrt@austin.utexas.edu
Zhehui Wang,
✉ zwang@lanl.gov

RECEIVED 06 November 2023

ACCEPTED 06 February 2024

PUBLISHED 22 February 2024


CITATION

Lin S, Ning S, Zhu H, Zhou T, Morris CL,
Clayton S, Cherukara MJ, Chen RT and Wang Z
(2024), Neural network methods for radiation
detectors and imaging.
Front. Phys. 12:1334298.
doi: 10.3389/fphy.2024.1334298

COPYRIGHT

© 2024 Lin, Ning, Zhu, Zhou, Morris, Clayton,
Cherukara, Chen and Wang. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Neural network methods for radiation detectors and imaging

S. Lin^{1,2}, S. Ning², H. Zhu², T. Zhou³, C. L. Morris¹, S. Clayton¹,
M. J. Cherukara ⁴, R. T. Chen^{2,5,6*} and Zhehui Wang^{1*}

¹Los Alamos National Laboratory, Los Alamos, NM, United States, ²Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, United States, ³Center for Nanoscale Materials, Argonne National Laboratory, Lemont, IL, United States, ⁴Advanced Photon Source, Argonne National Laboratory, Lemont, IL, United States, ⁵Microelectronics Research Center, The University of Texas at Austin, Austin, TX, United States, ⁶Omega Optics Inc, Austin, TX, United States

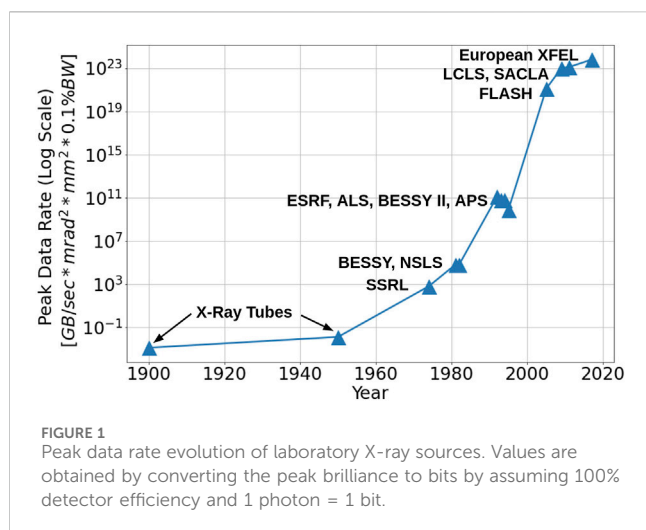
Recent advances in image data processing through deep learning allow for new optimization and performance-enhancement schemes for radiation detectors and imaging hardware. This enables radiation experiments, which includes photon sciences in synchrotron and X-ray free electron lasers as a subclass, through data-endowed artificial intelligence. We give an overview of data generation at photon sources, deep learning-based methods for image processing tasks, and hardware solutions for deep learning acceleration. Most existing deep learning approaches are trained offline, typically using large amounts of computational resources. However, once trained, DNNs can achieve fast inference speeds and can be deployed to edge devices. A new trend is edge computing with less energy consumption (hundreds of watts or less) and real-time analysis potential. While popularly used for edge computing, electronic-based hardware accelerators ranging from general purpose processors such as central processing units (CPUs) to application-specific integrated circuits (ASICs) are constantly reaching performance limits in latency, energy consumption, and other physical constraints. These limits give rise to next-generation analog neuromorphic hardware platforms, such as optical neural networks (ONNs), for high parallel, low latency, and low energy computing to boost deep learning acceleration (LA-UR-23-32395).

KEYWORDS

neural networks, ONN, radiation detectors, radiographic imaging and tomography, AI, edge computing

1 Introduction

X-rays produced by synchrotrons and free electron lasers (XFELs), together with high-energy photons above 100 keV, which are often generated using high-current (kA) electron accelerators and lately high-power lasers, are widely used as radiographic imaging and tomography (RadIT) tools to examine material properties and their temporal evolution [1–3]. Spatial resolution (δ) down to atomic dimensions is possible by using diffraction-limited X-rays, $\delta \sim \lambda/2$, corresponding to Abbe's diffraction limit for X-ray wavelength λ [4,5]. The overall object size that X-rays can probe readily reaches a length (L) greater than 1 mm, which is limited by the X-ray attenuation length and is X-ray energy dependent. In room-temperature water, for example, $L = 0.19, 1.2, 5.9$, and 14.1 cm for $1/e$ -attenuation length of 10 keV, 20 keV, 100 keV, and 1 MeV X-rays, respectively. The temporal resolution has now approached a few femtoseconds by using XFELs, where an XFEL experiment can be repeated for many hours in a pump-probe configuration [6,7]. In other words, the spatial



dynamic range (i.e., for 10 keV X-rays, $L \sim 1$ mm) is $2L/\lambda > 10^7$ and temporal dynamic range is $> 10^{18}$. Such ultra-wide-dynamic-range abilities of X-ray and photon techniques to connect elementary atomic and molecular processes, which are described by ultra fast (sub-nanosecond) quantum physics, with emergent macroscopic material properties and functions, which are usually treated classically through continuum approximations, make them extremely valuable in a wide range of applications. A few applications include medicine (i.e., new drug discovery), high-energy density battery development, and applications in materials exposed to high-temperature, high radiation, and other harsh or ‘extreme’ conditions. Additional applications include the optimization of chemical catalysis and the development of new superconductors and other quantum materials for information technology, accelerated computing, and artificial intelligence (AI).

The enormous spatial and temporal dynamic ranges give rise to “big data” in X-ray imaging, tomography, and photon science. Theoretically, 1 mm³ of water contains about 5.6×10^{-5} mol of water molecules ($N = 3.3 \times 10^{19}$). If the position of every molecule were recorded, the memory size would be $N \log_2 N$ ($\log_2 N$ is the bit length for a binary data system) or 2.2×10^{21} bits. In experiments, explosive data growth in X-ray and other forms of RadIT is built upon steady progress for more than 120 years in X-ray and radiation sources, detectors, computation, and lately data science. Figure 1 shows the evolution of the peak data rate due to the increasing X-ray source brilliance over the years [8]. The fourth generation synchrotrons such as APS-U [9] and PETRA IV [10] will have a significant reduction in emittance and a brilliance increase by a factor about 10^3 over the parameters of the third generation synchrotrons such as APS and PETRA III. XFELs, which are many of orders of magnitude brighter than synchrotrons, will run at a higher repetition rate up to 1 MHz [11]. The original LCLS, in comparison, operates at 120 Hz. However, the upgraded LCLS-II greatly increased the repetition rate to 1 MHz. High-speed detectors with frame rate frequencies above 1 MHz are commercially available. The combination of high-repetition-rate experiments with a mega-pixel and larger recording system leads to high data rates, exceeding 1 TB/s (1 TB = 10^{12} bytes), as we discuss further in Sec. 2.1.

Big data not only presents a significant challenge to data handling in terms of computing speed, computing power, short- and long-term computer memory, and computer energy consumption, which all together is called “computational resources”, but also offer a transformative approach to process and interpret data, i.e., machine learning (ML) and AI through data-enabled algorithms. Such algorithms, including deep learning (DL) [12,13], are distinctive from traditional physics, statistical, and other forward-model- or domain-knowledge-driven algorithms. Traditional algorithms are based on the domain knowledge, such as physics and statistics, and applicable to both small or large ensembles of data. In contrast, data-driven models may only rely on data explicitly for model training (tuning), model validation and use, with no domain knowledge required. In practice, domain knowledge always helps, partly due to the fact that some aspects of data models, such as the model architecture and other hyper-parameters, are chosen pragmatically and do not depend on the data. The amount of data required for data model training depends on the number of model parameters such as weights, activation functions, the number of nodes, etc. It is not uncommon that a deep neural network (DNN) may contain billions of tunable free parameters, which require a commensurate amount of data for training. Hybrid approaches to ML and AI [14,15], which merge data and domain knowledge, are increasingly popular. Hybrid models not only supplement data-driven models with domain knowledge and reduce the amount of data required for training, but also accelerate the computational speed of traditional forward models by 10 to more than 100 times by bypassing some detailed and time-consuming computations [16–18].

We may differentiate two approaches to ML and AI by the computational resources involved and how the resources are distributed. In the *centralized* approach, data are collected from distributed locations or different data acquisition instruments through the internet. The data are then stored in a data center, and processed by high-performance computers or mainframes. Cloud computing and data centers are now widely used to process ‘big data’ in industry, healthcare, and research institutions. However, using cloud computing to process data generated at the network edge is not always efficient. One limiting factor is the limited network bandwidth for data transportation due to increasing data generation rates. For example, in 2017, CERN had to install a third 100 Gigabit per second fiber optic line to increase their network capacity and bandwidth [19]. Other factors include the scalability and privacy issues of data transmission to the cloud [20]. Through the cloud computing and data center approach, data generation and data processing tasks can be separated, which can mitigate the computation and data processing burden on people who generate data. In the *distributed* or *edge* approach, ML and AI, together with the computing hardware, are deployed at the individual device or instrument level. Distributed computing now pairs with distributed data. Through an internet of ML/AI-enhanced instruments, each ML/AI-enhanced instrument can be optimized for a specific purpose such as data reduction and real-time data processing. Shown later in Table 1, detection cameras used at various synchrotron and XFEL facilities can generate data at a rate of > 1 GB/s and over > 12.5 GB/s for state-of-the-art cameras. This results in high costs of memory storage as well as high energy costs

TABLE 1 A comparison of different camera data rates and specifications for individual integration modules for each detector. Additional details and examples may be found in [1]. Note that this table tabulates select cameras to illustrate the data rates and their uses in light sources or X-ray measurements. The state of the art is >10 Gpixel/s (>12.5 GB/s assuming 10-bit data) in continuous mode imaging. The burst mode imaging is > 1 Tpixel/s (>1250 GB/s assuming 10-bit data) [218].

Detector (camera)	Facility (particle/ photon)	Det. Mode (Direct/ inDirect)	Array format (voxel size, μm^3 /pixel size, μm^2)	Frame-rate (fps/Hz)	Data bits	Data rate (GB/s)
AGIPD [39]	Eu-XFEL	D	512×128^a	16 k/6.5 M ^b	14	1.85
	(12.4 keV)		$(200^2 \times 500)$			
CS-PAD [219]	LCLS	D	194×370^c	120	14	0.02
	(8.3 keV)		$(110^2 \times 500)$			
DSSC [220]	Eu-FXEL	D	1024×1024	1–5 M	8	1050–5240
	(0.5–20 keV)		$(204 \times 236 \times 450)$			
ePix100 [221]	LCLS	D	384×352^d	120	14	0.03
	(8.3 keV)		$(50^2 \times 500)$	(≤ 240)		
ePix10k	LCLS	D	384×352^e	120	14	0.03
	(8.3 keV)		$(100^2 \times 500)$	($\leq 10^3$)		
EIGER2 (Dectris)	APS & others	D	1028×512^f	2.25 k	16	2.37
			$(75^2 \times 450)$	(4.5 k)	(8)	
HEXITEC [222]	DIAMOND	D	80^2	6.3–8.9 k	14	0.07–0.1
	(2–200 keV)		$(250^2 \times 1000^g)$			
Icarus [223]	NIF, Z	D	1024×512	≥ 250 M ^h	10	163,840
	(0.7–10 keV)		$(25^2 \times 25)$			
(Advanced hCMOS Sys.) JUNGFRAU [224]	SwissFEL/SLS	D	1024×512^i	2.2 k	16	2.31
	(0.25–12 keV)					
MM-PAD [225]	CHESS	D	128^2	10 k/100 M ^k	14	0.3–2867
	(> 20 keV) ^l		$(150^2 \times 500)$			
SOPHIAS	SACLA	D	891×2157	60	12	0.17
			$(30^2 \times 500)$			
HPV-X2 [226]	APS & others	inD	400×250	7.8 k/5 M ^m	10	0.98–625
(Shimadzu)	(10–40 keV)		(32^2)			
Kraken [227]	NNSS	inD	800×800	20 M ⁿ	12	19,200
			(30^2)			
MX170-HS	LCLS	inD	3840^2	2.5 ^o	16	0.07
(Rayonix)	(8–12 keV)		(44^2)			
PI MAX 4	APS	inD	1024^2	26 ^p	16	0.05
(Teledyne)	(10–40 keV)		(12.8^2)			

^aAGIPD is deployed as mega-pixel/voxel cameras through tiling.
^bBurst mode for 352 stored frames.
^cCS-PAD is deployed as tiled 2, 8, and 32 modules with up to 2.3 M voxels.
^dePix100 is deployed as tiled 4 modules with about 0.5 M voxels.
^eePix10K replaces CS-PAD, and is deployed as a single, or tiled 16 modules with about 2.2 M voxels.
^fEiger2 is deployed as a single, or tiled modules with more than 10 M voxels.
^gAlso 2 mm CdZnTe.
^hIn burst mode for 4 frames.
ⁱArray size of individual modules. Multiple modules can be tiled to create larger detector configurations [228]. For example, the JUNGFRAU 4M consists of eight modules.
^jCan resolve single photons down to 800 eV or lower by combining the readout chip with LGAD sensors [229].
^kIn burst mode for 8 frames.
^lWhen using 750 μm CdTe as sensor.
^mIn burst mode for 128 stored frames; or 10 M Hz frame rate and 256 stored frames possible by reducing the number of pixels by half.
ⁿIn burst mode for 8 frames. Read noise 157 e^- , Full Well $4.0 \times 10^5 e^-$. Buttable to larger array 2×2 .
^oHigher frame rate can be obtained through pixel binning, at 10×10 binning, the frame rate increases to 120 Hz.
^pHigher frame rate can be obtained through pixel binning, at 4×4 binning, the frame rate increases to 95 Hz.

for data transmission and memory access during data processing. The large volumes, varieties, and generation rate of X-ray data motivate automated processing and reduction in light sources, such as synchrotrons and XFELs, to reduce the memory requirement, minimize latency related to data transmission and processing, and lower energy and power consumption. Later in the paper, [Section 3.1](#) will discuss more details.

We will give an overview of DL methods for real-time radiation image analysis as well as hardware solutions for DL acceleration at the edge. We note that while not all scientific applications may require real-time image analysis, it is possible to offload some computing and preprocessing steps to an edge device. The edge device can preprocess the acquired data in real-time before sending the processed data to upstream processing centers for heavier computations. This paper is organized as follows. In [Section 2](#), we discuss different radiation detectors and imaging devices, the resulting big data generation at photon sources, and the motivations for edge computing and DL. In [Section 3](#), we present an overview of popular neural network architectures and several image processing tasks that have potential to be performed on edge devices. In addition, we discuss examples of DL-based methods for each. In [Section 4](#), we present an overview of hardware solutions for DL acceleration and recent works that have applied them for computing at the edge. Lastly, [Section 5](#) concludes this paper.

2 Experimental data generation at photon sources

Data science at light sources is centered around scientific data generation and processing. Scientific data at synchrotrons and XFEL sources consist of experimental data, simulation and synthetic data, and meta data, such as detector calibration data, material properties of objects and sensors, and point spread functions of the detectors. Methods (imaging modalities) and detectors to collect experimental data are driven by the light sources, which continue to improve in source brightness, repetition rate, source coherence, photon energy, and spectral tunability. Computing hardware and algorithms are used to process experimental data and for data visualization. Computing hardware and algorithms are also used to simulate the experiments and produce synthetic data as close to the experimental data as possible for experimental data interpretation. Diversity of the materials to be integrated and imaged, together with the photon source and detector improvement have demanded continued improvements in computing hardware and algorithms towards real-time data processing, reductions in data transmission over long distances, and reducing data storage volumes.

2.1 Radiation detectors and imaging for photon science

Complementary metal-oxide semiconductor (CMOS) pixelated detectors, including hybrid CMOS, are now widely used for X-ray photon science, replacing charge-coupled devices (CCDs) as the primary digital imaging technology, see [Figure 2](#). CMOS technology is rapidly catching up to CCD cameras, with recent developments

such as Sony's STARVIS which can offer better sensitivity than traditional CCD sensors [21]. In addition, CMOS sensors are much cheaper than CCD sensors, making them more cost efficient while achieving matching performance. The latest trend is smart CMOS technology to enable edge computing and neural networks on CMOS sensors [22,23]; see [Section 2.3](#) for more details.

CMOS sensors are used in many state-of-the-art radiation applications. For example, CMOS-based back-thinned monolithic active pixel sensors (MAPS) are the state-of-the-art detectors used for cryo-electron microscopy applications. MAPS detectors are CMOS sensors that combine the photodetectors and readout electronics on the same silicon layer, while backthinning reduces the electron scattering within pixels. MAPS detectors are also being developed for high-energy physics [24], cryogenic electron microscopy (cryo-EM), cryo-ptychography, integrated differential phase contrast (iDPC), and liquid cell imaging applications [25]. Meanwhile, hybrid CMOS detectors such as the AGIPD, ePix, and MM-PAD (see [Table 1](#) for more detectors) are popularly used at facilities for photon science applications. Hybrid detectors are composed of a sensor array and pixel electronics readout layer that are interconnected through bump bonding, while the sensor frontend can be fabricated using different semiconductor materials. The thickness and material properties of the sensor array is dependent on the active absorbing layer design requirements and given X-ray energy to obtain high quantum efficiency. For example, high-Z sensors use materials with high atomic numbers such as Gallium Arsenide (GaAs), Cadmium Telluride (CdTe) and Cadmium Zinc Telluride (CZT) [26]. The hybrid design architecture allows for independent optimization of the quantum efficiency of the sensor array and pixel electronics functionality to meet imaging and measurement performance requirements [27]. Currently, hybrid CMOS detectors are the most widely used image sensors for high energy physics experiments [24]. Another family of image sensor is called the low-gain avalanche detector (LGAD), a silicon sensor fabricated on thin substrates to deliver fast signal pulses to achieve enhanced time resolution [28], as well as to increase the X-ray signal amplitudes and the signal-to-noise ratio to achieve single photon resolution [29]. As a result, LGADs are popularly used in experiments that require fast time resolution and good spatial resolution such as 4D tracking [30] and for soft X-ray applications in low energy diffraction, spectro-microscopy and imaging experiments such as the resonant inelastic X-ray scattering experiments [29]. In summary, radiation pixel detectors aim to capture incident photons and convert the accumulated charges in the pixel into an output image. We also mention that CMOS image sensors, including hybrid CMOS, may also be extended to neutron imaging by converting incident neutrons to visible photons through neutron capture reactions [31].

The particle nature of photons motivates digitized detectors for photon counting. Hybrid CMOS detectors are one of the most popular detectors that use the photon counting mode of operation, where individual photons are detected by tuning the discriminator threshold and the energy value of each incident photon is recorded as electronic signals. However, several factors complicate photon counting implementation in high-luminosity X-ray sources. The intensity of the sources can be too high to count individual photons one by one. The amount of X-ray photon-induced charge in CMOS detectors, which is the basis of X-ray photon counting, is not

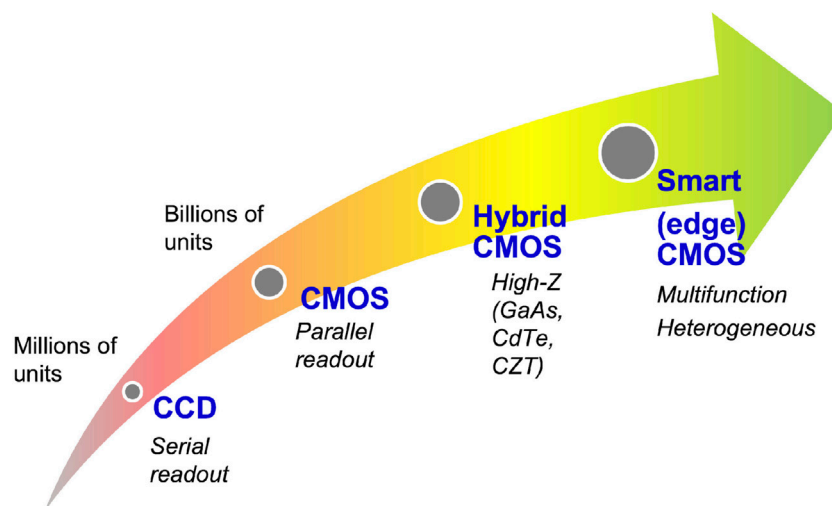


FIGURE 2
Evolution of digital image sensor technology, which started with the introduction of the charge-coupled device (CCD) in the late 1960s. The latest trend is smart multi-functional CMOS image sensors enabled by three-dimensional (3D) integration in fabrication, innovations in heterogeneous materials and structures, neural networks, and edge computing.

constant for the same X-ray energy. Furthermore, the detectors suffer from a charge sharing issue when a photon interacts on the border between neighboring pixels. The source energy is not monochromatic, especially in imaging applications. Inelastic scattering of mono-chromatic X-rays can result in a broad distribution of X-ray photon energies after scattering by the object. When an optical camera is used together with a X-ray scintillator, the energy resolution of individual X-rays based on the photon detection is worse than direct detection when the X-ray directly deposits its energy in a silicon photo-diode. See Table 1 for a comparison of different direct and indirect detection cameras and their data rates. Note that Table 1 tabulates the specifications of select cameras to illustrate their data rates as well as their uses in light sources or X-ray measurements. To overcome the issue of the too high photon flux rate, hybrid detectors are developed to operate under a charge integration mode, where the signal intensity is obtained by integrating over the exposure time. The current generation of hybrid CMOS detectors are capable of different modes of operation (i.e., photon counting and charge integration) for direct photon detection [32].

2.2 Imaging modalities

X-ray microscopy uses X-ray lenses, zone plates, mirrors and other optics to modulate the X-ray field to form images [33]. As the X-ray intensities generated by synchrotrons and XFELs continue to increase, the advances in computational imaging modalities and lens-less X-ray modalities are increasingly used in synchrotrons and XFELs. In some cases, lens-less modalities may be preferred to avoid damages to X-ray lenses and mirrors. Lens-less modalities may also avoid aberration, diffraction due to imperfect X-ray lens, defects in zone plates, and other optics. The simplest lens-less X-ray imaging setup is radiography or projection imaging, pioneered by Röntgen. Röntgen's lens-less radiographic imaging modality directly

measures attenuated X-ray intensity due to absorption. Synchrotrons and XFELs also allow a growing number of phase contrast imaging, see Ref. [1] and references therein. Other modalities include in-line holography [34] and coherent diffractive imaging [35]. Additional phase and intensity modulation using pinholes, coded apertures, and kinoforms are also possible. Combinatorial X-ray modalities have also been introduced. For example, X-ray ptychography microscopy combines raster scanning X-ray microscopy with coherent diffraction imaging [36]. Compton scattering, usually ignored in the synchrotron and XFEL setting, may offer some additional information about the samples and potentially reduce the dose required [37]. The versatility of modalities requires different off-line and real-time data processing techniques. Background reduction is a common issue for all X-ray modalities. Real-time data processing, including energy-resolving detection, is highly desirable to distinguish different sources of X-rays since the detector pixel may simultaneously collect X-ray photons from different sources of X-ray attenuation and scattering.

2.3 Real time in-pixel data-processing

When an X-ray photon is detected directly or indirectly through the use of a scintillator, charge-hole pairs are created through photo-to-electric conversion, or the photoelectric effect, within pixels of a camera or a pixelated array. CCD cameras, CMOS cameras, and LGAD arrays are now available for synchrotron and XFEL applications. Unlike a CCD camera, a CMOS image sensor collects charge and stores it in capacitors in pixels in parallel. Parallel charge collection and capacitor voltage digitization, which turns analog voltage signals into digitized signals, allow CMOS image sensors to operate at a much higher frame rate than CCDs. Charge and voltage amplification, in LGADs and sometimes in CMOS image sensors, are also used to improve

signal-to-noise ratio. Any source of charge or voltage modulation not related to the photoelectric effect is a potential source of noise. The photoelectric effect itself can lead to so-called Poisson noise due to the probabilistic process of photo-to-electric conversion. Other sources of noise include thermal noise or dark current, salt'n'pepper noise (due to charge migration in and out of pixel defects and traps), and readout noise.

Automated real-time in-pixel signal and data processing are therefore required in CMOS and other pixelated array sensors for noise rejection and noise reduction for charge and voltage amplification controls, and for charge sharing corrections. Figure 3 illustrates a generic approach on in-pixel neural network processing for optimized and real-time data processing. Common approaches process the data by transmitting it to a separate processor and storing the data in memory. However, the data transmission and memory access actions are known to be among the most power hungry in imaging systems [38]. As a result, it is desirable to optimize the end-to-end processes of sensing, data transmission, and processing tasks. One solution is to utilize in-pixel processing to directly extract features of the input pixels which can significantly reduce system bandwidth and power consumption of data transmission, memory management, and downstream data processing. In recent years, a number of works have been proposed to implement image sensors with in-pixel neural network processing; see [22,23] and references therein. This motivates real-time image processing for image sensors for various image processing tasks including noise removal. If uncorrected, noise can corrupt the image information and make it hard for post processing or misleading for data interpretation. Charge and voltage amplification may lead to nonlinear distortion between the X-ray flux and voltage signal. When the X-ray flux is too high, the so-called plasma effect may also need correction. Charge-sharing happens when an X-ray photon arrives at a pixel border and the electron-hole pairs created are spread across multiple neighboring pixels.

By using transistor circuits, correlated double sampling (CDS) is an extremely successful example in noise reduction. Adaptive gain control circuits have been implemented in the AGIPD high-speed camera [39,40]. While real-time pixel-level signal processing by novel transistor circuits is important, there is also room for novel data-processing approaches that do not require hardware modifications to the pixels. As a recent example [41], a physics-informed neural network was demonstrated to improve spatial resolution of neutron imaging. Other novel applications of neural networks and their integration with hardware, see Figure 3, may offer new possibilities in noise reduction and image corrections. Integrated hardware and software (neural networks are emphasized here) approaches for optimal performance also need to take into account of the complexity of the workflow [42–44], or computational cost, power consumptions, constrained by the frame rate and other metrics. For example, the computational cost of an $n \times n$ matrix is $O(n^3)$ [45].

3 Deep learning for image processing

In recent years, deep learning (DL) has contributed significantly to the progress in computer vision, especially in different areas of image processing tasks including but not limited to image denoising,

segmentation, super-resolution, and classification. DL is a sub-field of ML and AI that utilize neural networks (NNs) and their superior nonlinear approximation capabilities to learn underlying structures and patterns within high-dimensional data [12]. In other words, DL aims to learn multiple levels of representation, corresponding to a hierarchy of features or concepts, where higher-level features are defined from lower-level ones and lower-level ones can help build up higher-level features.

For DL algorithms to extract underlying features and to obtain accurate predictions, it is important to understand the workflow of the DL process. In general, the DL process can be broken into several stages: i) data acquisition, ii) data preprocessing, iii) model training, testing, and evaluation, iv) model deployment and monitoring [46,47]. The first step to ML and DL problems is to collect large amounts of data from sources including but not limited to sensors, cameras, and databases. Next, the collected data needs to be preprocessed into useful features as inputs into the DL model. At a high level, the preprocessing step aims to prepare the raw data (e.g., data cleaning, outlier removal, data normalization, etc.) and to allow data analysts to preform data exploration (i.e., identifying data structure, relevance, type, and suitability). The preprocessed data is split for model training, testing, and evaluation. The appropriate DL training algorithm, model, and ML problem are dependent on the nature of the application. The model is trained on the training dataset to tune the model hyperparameters and is evaluated using unseen data, also known as the test dataset. This process is reiterated until a desired accuracy performance or stopping criteria has been achieved. Last, the trained model is deployed and monitored for further retraining and redeployment. See [46,48,49] for comprehensive details on the basics of DL.

3.1 Centralized and decentralized learning

Recall that ML and AI, and thus DL, can be differentiated into two approaches, namely, centralized and decentralized approaches. In the centralized approach, the data collected from network edge detectors are transmitted to and stored in a data center and then processed by high-performance computers. Very large traditional, ML, or hybrid algorithms can be deployed in the data center, which also requires correspondingly large memory, energy and power consumption. Estimated global data center electricity consumption in 2022 was 240–340 TWh [50], or around 1%–1.3% of global final electricity demand from data centers and data transmission networks [51]. To put this value in a better perspective, it is estimated that Bitcoin alone consumes around 125 TWh per year [52] and that the combination of Bitcoin and Ethereum consumed around 190 TWh (0.81% of the world energy consumption) in 2021 [53]. Furthermore, centralized approaches and large ML models are commonly executed by a large team of people. A Meta AI research team recently introduced the model called Segment Anything Model (SAM) and a dataset of more than 1 billion masks on 11 Million images [54]. Nvidia unveiled Project Clara at its recent GTC conference, showing early results using DL post-processing to dramatically enhance existing, often grainy and indistinct echocardiograms (sonograms of the heart). Clara motivates acceleration in research being done on several fronts that exploits explosive growth in DL computational capability to

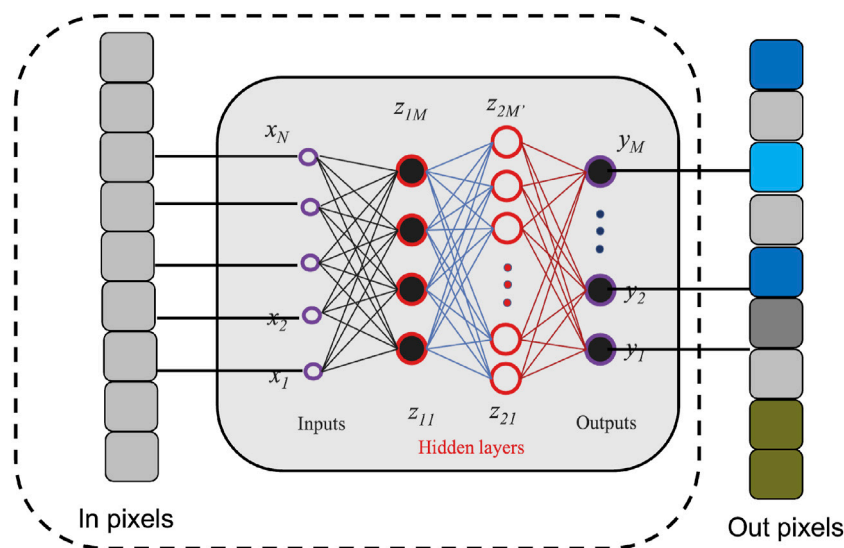


FIGURE 3

A generic illustration of in-pixel neural network processing for optimized and real-time end-to-end data processing and reductions. The neural network is directly implemented on the imaging sensor. For this specific example, the network illustrated is a fully connected neural network (FCNN). The network takes in the sensor pixel values as inputs x (in pixels) then feeds the input into hidden layers z with the number of neurons per layer denoted by M . The processed pixels (out pixels) is the output of the neural network denoted by y .

perform analysis that was previously impossible or far too costly. One technique is called 3D volumetric segmentation that can accurately measure the size of organs, tumors or fluids, such as the volume of blood flowing through arteries and the heart. Nvidia claims that a recent implementation, an algorithm called V-Net, “would’ve needed a computer that cost \$10 million and consumed 500 kW of power [15 years ago]. Today, it can run on a few Tesla V100 GPUs” [55,56]. This claim accentuates the rapid advancements made in the hardware industry to accommodate DL computational requirements. For example, a work by [57] implemented and trained V-Net for 48 h on a workstation equipped with an NVIDIA GTX 1080 with 8 GB of video memory.

However, data processing using cloud computing and data centers are inefficient due to factors including limited network latency, scalability, and privacy [20]. To address these challenges, edge computing, or the distributed approach, offloads computing resources to the edge devices to improve network latency, to enable real-time services, and to address data privacy challenges by directly analyzing data generated by the source. In addition, edge computing will help reduce the high costs of memory storage as well as high energy costs for data transmission and memory access during data processing. For example, CERN’s Large Hadron Collider (LHC) and non-LHC experiments generate over 100 petabytes of data each year and CERN’s main data center had an energy consumption of about 37 GWh over the year of 2021 [58]. Assuming that the average cost of electricity is \$0.15/kWh, then the cost of using 37 GWh is \$5.55 million. CERN’s new data center in Prévessin aims to have a power usage effectiveness (PUE) below 1.1 (ideal PUE is 1.0), and in future data centers, CERN aims to implement ML based approaches to key computing tasks to help reduce the amount of computing resources and energy consumption [58].

There are some other successes using ML and AI in areas such as HEP experiments (i.e., Higgs boson discovery) and electron

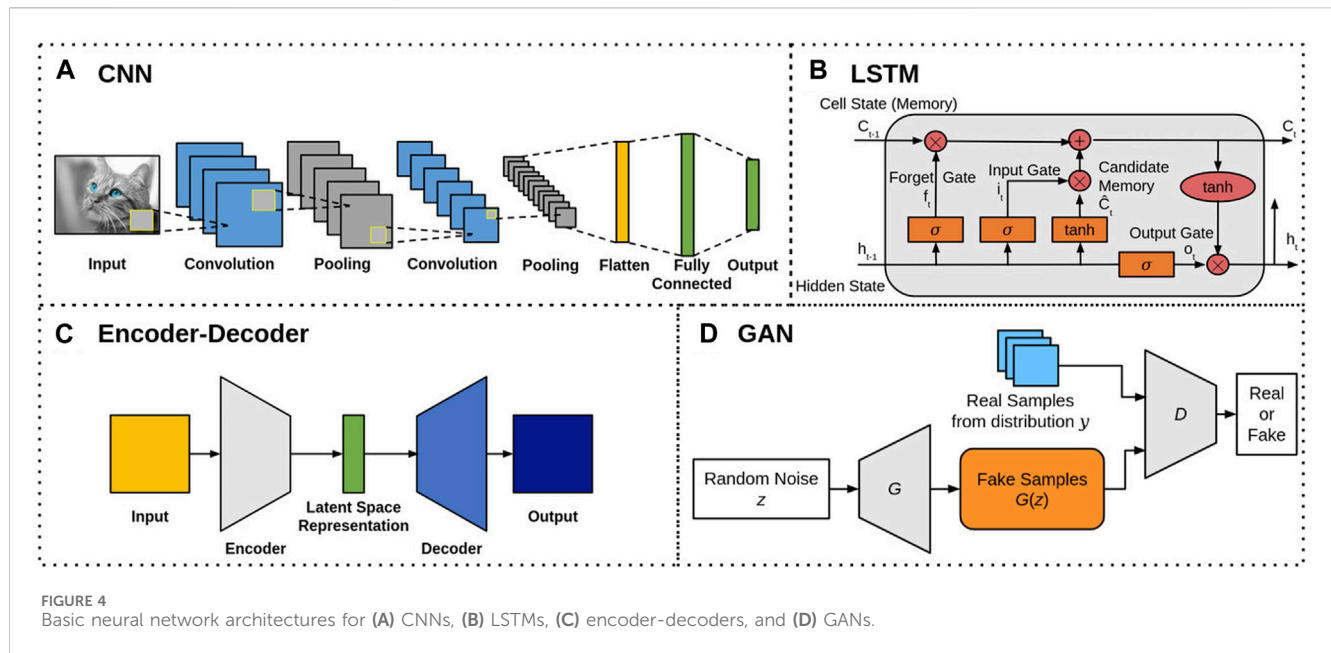
microscopy. The discovery of the Higgs boson was a major challenge in HEP and can be setup as a classification problem. Many ML methods such as decision trees, logistic regression, and DL algorithms have been applied to solve the signal separation problem [59,60]. Meanwhile, ML and AI in electron microscopy are proposed to enable autonomous experimentation. Specifically, the automation of routine operations including but not limited to probe conditioning, guided exploration of large images, optimized spectroscopy measurements, and time-intensive and repetitive operations [61]. Edge ML and edge AI have already attracted a lot of attention in medicine. The fusion of DL and medical images creates dramatic improvements [56]. The concept is similar to techniques like high-dynamic range (HDR) photography, digital remastering of recordings or even film colorization in that one or more original sources of data are post-processed and enhanced to bring out additional detail, remove noise or improve aesthetics.

3.2 Neural network architectures

This section provides an overview of different popular deep neural network (DNN) architectures used for image processing tasks. These widely used architectures include but are not limited to convolutional neural networks (CNNs), long short-term memory (LSTM), encoder-decoder networks, and generative adversarial networks (GANs). Due to space limitations, other DNN architectures such as transformers [62], restricted boltzmann machines [63], and extreme learning [64] will not be covered here.

3.2.1 Convolutional neural networks (CNNs)

CNNs are one of the most widely used architectures in DL, especially for image processing tasks, due to their inherent spatial invariance property. The built-in convolutional layers allow the



network to naturally reduce the high dimensionality of the input data, i.e., images, without information loss. Figure 4A shows the basic architecture of CNNs, which usually consists of 3 types of layers: i) convolutional layers, ii) pooling layers, and iii) fully connected layers. The convolutional layer uses various kernels to convolve the entire input image, including intermediate feature maps, and generate new feature maps. There are 3 major advantages of the convolutional operation [65]: i) the number of parameters is reduced by using weight sharing mechanisms, ii) the correlation among neighboring pixels are easily learned through local connectivity, and iii) the location of objects are fixed due to spatial invariance. Generally following a convolutional layer, the pooling layer is used to further reduce the dimensions of feature maps and network parameters. The average pooling and max pooling methods are commonly used, and their theoretical performances have been evaluated by [66,67], where max pooling is shown to achieve faster convergence and improved CNN performance. Lastly, the fully connected layer follows the last pooling or convolutional layer to convert the 2D feature maps into a 1D vector for additional feature mapping, i.e., labels. A few of the well known CNN models are the AlexNet [68], VGG [69], GoogLeNet [70], and ResNet [71], where all the models were top 3 finishers in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Discussed later in Section 3.3, CNNs are popularly used in many image processing tasks such as image restoration (e.g., denoising, deblurring, and super resolution), segmentation, classification, and 3D reconstruction. A few examples from photon sciences include but are not limited to denoising synchrotron computed tomography images, deblurring neutron images, segmentation of inertial confinement fusion radiographs, and 3D reconstruction of coherent diffraction imaging.

3.2.2 Long short-term memory (LSTM)

LSTMs [72] are a special type of recurrent neural network (RNN) that is commonly used to process sequential datasets, such as audio recordings, videos, and time-series data. Figure 4B

shows the basic structure of a LSTM block, which consists of 3 gates (the forget gate f_t , the input gate i_t , and output gate o_t), as well as the candidate memory (new information) \tilde{C}_t , that regulate the stored memory and information flow within the block. Note that the subscript t denotes the variable state at time t . In summary, the forget gate decides what information to remove from the cell state, the input gate decides what information to update in the cell state by selecting the candidate memory values, and the output gate decides what information to output in the current cell state. The multiple gate architecture of LSTMs is specifically designed to capture long-term dependencies in the data as well as to avoid the vanishing gradient problem of vanilla RNNs [72]. The vanishing gradient problem in RNNs results from the backpropagation of gradients through time, which can result in very small gradient values and short-term memory behaviors. Meanwhile, the gating mechanisms of LSTMs control the flow of gradients through time during backpropagation, and thus effectively addresses the vanishing gradient problem and allows the network to learn long-term memory behaviors. Other LSTM architectures are derived from the basic architecture in Figure 4B such as LSTM without a forget gate, LSTM with peephole connections, the gated recurrent unit (GRU), and other variants [72]. Due to their ability to process sequential data, LSTMs can be used to process time-series experimental data such as videos, where video-based processing techniques can be applied.

3.2.3 Encoder-decoders

Encoder-decoder neural networks, also known as sequence-to-sequence networks, are a type of network that learns to map the input domain to a desired output domain [13]. As shown in Figure 4C, the network consists of two main components: an encoder network which uses an encoder function $h = f(x)$ to compress the input x into a latent-space representation h , and a decoder network $y = g(h)$ that produces a reconstruction y from h . The latent-space representation h prioritizes learning the important aspects of the input x which are useful in reconstructing the output y .

A special case of encoder-decoder models, autoencoders are networks in which the input and output domains are the same. These networks are popularly used in DL applications involving sequence-to-sequence modeling such as natural language processing [73], image captioning [74], and speech recognition [75]. In image processing, encoder-decoder networks are popularly used for image denoising, segmentation, compression, and 3D reconstruction. For example, one popular encoder-decoder model for image segmentation is U-Net [76]. Discussed later in Section 3.3, U-Net is used for image segmentation of inertial confinement fusion images and modified versions of the U-Net architecture are used in many works for image processing tasks. A few examples include but are not limited to the denoising and super resolution of synchrotron and X-ray computed tomography images.

3.2.4 Generative adversarial networks (GANs)

GANs [77] are increasingly popular DL frameworks for generative AI models. Classical GANs consist of 2 different networks, a generator and a discriminator, as shown in Figure 4D. The generator network G aims to generate data $G(z)$ that is indistinguishable from the real data by learning a mapping from an input noise distribution z to a target distribution y of the real data. Meanwhile, the discriminator network D takes as input the real and generated data, and aims to correctly classify them as “real” or “fake” (generated). The GAN learning objective takes on a game-theoretic approach as a two player minimax game between G and D . Let \mathcal{L} denote the loss function and the GAN objective as $\min_G \max_D \mathcal{L}(G, D)$. Intuitively, D aims to minimize its own classification error, which maximizes $\mathcal{L}(G, D)$. Meanwhile, G aims to maximize the classification error of D , which minimizes $\mathcal{L}(G, D)$. This adversarial loss function allows both models to be trained simultaneously and in competition with each other. Other GAN architectures are derived from the basic architecture in Figure 4D such as conditional GANs, GANs with inference models, and adversarial autoencoders [78]. Similar to CNNs, GANs are popularly used in many image processing tasks such as image restoration, compression, and 3D reconstruction. For example, a GAN-based image denoising method was proposed to denoise low-dose X-ray computed tomography images. In addition, GANs can be used to generate synthetic data similar to experimental data. A recent work [79] uses a GAN-based model to generate synthetic inertial confinement fusion radiographs.

3.3 Image processing techniques

This section provides an overview of several image processing tasks that have potential to be performed on edge devices. In addition, this section surveys different works that have applied the DL-based image processing techniques to radiographic image processing.

3.3.1 Restoration

Image restoration is the process of adjusting the quality of digital images such that the enhanced image can facilitate further image analysis. Common enhancement operations include histogram-based equalization, brightness, and contrast adjustment. However, these operations are very elemental and advanced operations are

necessary to further improve the perceptual quality. These advanced operations include image denoising, deblurring, and super-resolution (SR); see [80–83] for examples of images before and after processing.

3.3.1.1 Denoising

One of the fundamental challenges in image processing, image denoising aims to estimate the ground-truth image by suppressing internal and external noise factors such as sensor and environmental noise, as discussed in Section 2.3. Sources of noise include but are not limited to Poisson noise due to photo-electric conversion, camera thermal noise or dark current, salt’n’pepper noise, camera readout noise, and shot noise for low-dose X-ray imaging conditions. Conventional methods including but not limited to adaptive nonlinear filters, Markov random field (MRF), and weighed nuclear norm minimization (WNNM), have achieved good performance in image denoising [84], however, they suffer from several drawbacks [85]. Two major drawbacks are the need to manually set parameters as the proposed methods are non-convex and the high computational cost for the optimization problem for the test phase. To overcome these challenges, DL methods are applied for image denoising problems to learn the underlying noise distribution. Various neural network architectures, such as CNNs, encoder-decoders, and GANs, have been proposed for image denoising in recent years; see [84] for details.

An example application that uses image denoising is in X-ray computed tomography (CT). X-ray CT imaging is a common noninvasive imaging technique that allows for reconstructing the internal structure of objects by using 3D reconstruction from 2D projection images; see Section 3.3.6 on 3D reconstruction. The spatial resolution of XFEL-based and synchrotron-based X-ray CT images can range from tens of microns to a few nanometers, while higher resolutions can be obtained by using higher radiation doses. However, some experiments may require short exposure times or low radiation dosage to avoid damaging the sample. The low-dose image conditions results in noisy 2D projection images, which in turn impacts the quality of the 3D reconstructed image. To address this issue [86], developed a GAN-based image denoising method called TomoGAN. TomoGAN is a conditional GAN model where the generator G conditionally uses the noisy reconstruction as input and outputs enhanced (denoised) reconstructions. Furthermore, the generator network architecture adopts a modified U-Net [76] architecture, popularly used for image segmentation. Meanwhile, the discriminator D is trained to classify reconstructions of the enhanced reconstructions and reconstructions of normal dose projections [86]. Evaluates the effectiveness of TomoGAN on two experimental (shale sample) datasets. TomoGAN outperforms conventional methods in noise reduction and reports a higher structural similarity (SSIM) value. In addition, TomoGAN is demonstrated to be robust to images with dynamic features from faster experiments, e.g., collecting fewer projections and/or using shorter exposure times.

Denoising has also been applied to synchrotron radiation CT (SR-CT) in a recent work by [87], which developed a CNN-based image denoising method called Sparse2Noise. Similar to the previous work for TomoGAN, this work presents a low-dose imaging strategy and utilizes paired normal-flux CT images (sparse-view) and low-flux CT image (full-view) to train

Sparse2Noise. In addition, Sparse2Noise also adopts a modified U-Net architecture for its performance of removing image degradation factors such as noise and ring artifacts. The Sparse2Noise network takes as input the normal-flux CT images into the modified U-Net architecture and outputs the enhanced image. During training, the network is trained in a supervised fashion using the low-flux CT images. The loss function to update the network weights is defined to minimize the difference between the enhanced image and the reconstructed low-flux CT image [87]. Evaluates the effectiveness of Sparse2Noise on one simulated and two experimental datasets. Furthermore, Sparse2Noise is compared to simultaneous iterative reconstruction technique (SIRT), unsupervised deep imaging prior (DIP), and supervised training algorithms Noise2Inverse [88] and Noise2Noise [89]. For the simulated dataset, Sparse2Noise outperforms all methods by achieving the highest SSIM and peak signal to noise ratio (PSNR) values, and in terms of removing image degradation factors such as noise and ring artifacts. For the experimental datasets, Sparse2Noise also achieves the best performance in terms of noise and ring artifact removal. Most importantly, however, Sparse2Noise can achieve excellent performance for low-dose experiments (0.5 Gy per scan).

3.3.1.2 Deblurring

Image deblurring aims to recover a sharp image from a blurred image by suppressing blur factors such as lack of focus, camera shake, and target motion. Some blur factors are application specific such as multiple Coulomb scattering and chromatic aberration in proton radiography [90]. A blurred image can be modeled mathematically as $B = K * I + N$, where B denotes the blurred image, K the blur kernel, I the sharp image, N the additive noise, and $*$ the convolution operation. The blur kernel K is typically modeled as a convolution of blur kernels that are spatially invariant or varying [82]. Conventional methods aim to solve the inverse filtering problem to estimate K , however, this is an ill-posed problem as the sharp image I needs to be estimated as well. To address this issue, prior-based optimization approaches, also known as maximum *a posteriori* (MAP)-based approaches, have been proposed to define priors for K and I [91]. While these approaches are shown to achieve good results for image deblurring, deep learning approaches can further improve the accuracy of the blur kernel estimation or even skip the kernel estimation process altogether by using end-to-end methods. Various neural network architectures, such as CNNs, LSTMs, and GANs, have been proposed for image deblurring; see [82,91] for details.

One example application that uses image deblurring is in neutron imaging restoration (NIR), a non-destructive imaging method. However, the neutron images suffer from noise and blur artifacts due to the neutron source and the digital image system. The low quality of raw neutron images limits their applications in research, and thus image denoising and deblurring techniques are necessary to produce sharp images. To address these issues [92], proposes a fast and lightweight neural network called DAUNet. DAUNet consists of three main blocks: a feature extraction block (FEB), multiple cascaded attention U-Net blocks (AUB), and a reconstruction block (RB). First, DAUNet takes as input a degraded neutron image and feeds it into the FEB to extract

important underlying features. Next, the AUB inputs the extracted feature maps into a modified U-Net with an attention mechanism, which allows U-Net to focus on harder to address features such as texture and structure information, and outputs a restored image. Last, the RB block outputs the enhanced image by reconstructing the restored image. To evaluate DAUNet, its performance is compared with several popular DNN image restoration methods such as DnCNN [93] and RDUNet [94]. Due to the lack of available neutron imaging datasets, the networks are trained on X-ray images that are similar to the neutron imaging principle; specifically, the X-ray images are obtained from the SIXray dataset [95], where 4699 and 23 images are used as the training and test set respectively. In addition, seven clean neutron images are added to the test set. Results show that DAUNet can effectively improve the image quality by removing noise and blurring artifacts, while achieving quality close to the large network with faster running times and a smaller number of network parameters.

3.3.1.3 Super-resolution (SR)

Image SR is the process of reconstructing high-resolution images from low-resolution images. It has been widely applied in many real-world applications, especially in medical imaging [96] and surveillance [97], where the spatial resolution of captured images are not sufficient due to limitations such as hardware and imaging conditions. A variety of DL-based methods for SR have been explored, ranging from CNN-based methods (e.g., SRCNN [98]) to more recent GAN-based methods (e.g., SRGAN [99]). In addition to utilizing different neural network architectures, DL-based SR algorithms also differ in other major aspects such as their loss functions and training approaches [83,100]. These differences result from various factors that contribute to the degradation of image quality including but not limited to blurring, sensor noise, and compression artifacts. Intuitively, one can think of the low-resolution image as the output of a degradation function with an input high-quality image. In the most general case, the degradation function is unknown and an approximate mapping is learned through deep learning. These degradation factors influence the design of loss function, and thus training approaches. A detailed discussion of the various loss functions, SR network architectures, and learning frameworks is out of scope for this paper; however, see [100] for details.

An example application that applies super resolution is for X-ray CT imaging. As mentioned earlier, CT imaging has many factors that impact the resulting image quality such as radiation dose and slice thickness. In addition, 3D image reconstruction may require heavy computational power due to the number of slices or projection views taken, where thicker slices results in lower image resolution, and slower operational speed, which increases with the number of slices. To address this issue, it is desirable to obtain higher-resolution (thin-slice) images from low-resolution (thick-slice) ones [101]. Develops an end-to-end super-resolution method based on a modified U-Net. The network takes as input the low-resolution image and outputs the high-resolution one. The network is trained on slices of brain CT images obtained from a 65 clinical positron emission tomography (PET)/CT studies for Parkinson's disease. The low-resolution images are generated as the moving average of five high resolution slices and the

ground-truth image is taken as the middle slice. The performance of the proposed method is compared with the Richardson-Lucy (RL) deblurring algorithm using the PSNR and normalized root mean square error (NRMSE) metrics. The results show that the proposed method achieves the highest PSNR and lowest NRMSE values compared to the RL algorithm. In addition, the noise level of the enhanced images are reported to be lower than that of the ground-truth.

Super resolution has also been applied to transmission and cryogenic electron microscopy (cryo-EM) imaging applications for sub-pixel electron event localization [25,102]. In transmission electron microscopy, electron events are captured using pixelated detectors as a 2D projection track of the energy deposition [102]. Conventional reconstruction methods, such as the weighted centroid method and the furthest away method (FAM), require an event analysis procedure to extract electron track events. However, these classical algorithms are unable to separate overlapping electron event tracks, and do not take into consideration the statistical behavior of the electron movement and energy deposition. To address this issue [102], used a U-Net-based CNN to learn a mapping from input electron track image to an output probability map that indicates the probability of the point of entry for each pixel. The network is trained using a labeled dataset generated through Monte Carlo simulations, and tested on simulated data and experimental data from a pnCCD [103]. The performance of the proposed CNN model is compared with FAM. The results show that the proposed method achieves superior localization performance compared to FAM by reducing the distribution spread of the Euclidean distance on the simulated dataset, while achieving a modulation transfer function closer to the ground truth on the experimental dataset. For cryo-EM, a CNN model was applied in a similar manner for electron event localization, but with a slightly different dataset. Cryo-EM experiments popularly use MAPS detectors to directly detect electron events, where each captured electron results in a pixel cluster on the captured image. In [25], a CNN model is designed to output a sub-pixel incident position given an input pixel cluster image and the corresponding time over threshold values.

3.3.2 Segmentation

Image segmentation is the process which segments an image or video frames into multiple regions or clusters, where each pixel can be represented by a mask or be assigned a class [104]. This task is essential in a broad range of computer vision problems, especially for visual understanding systems. A few applications that utilize image segmentation include but are not limited to medical imaging for organ and tumor localization [105], autonomous vehicles for surface and object detection, and video footage for object and people detection and tracking [106]. Numerous techniques for image segmentation have been proposed throughout the years, ranging from early techniques based on region splitting or merging such as thresholding and clustering algorithms, to newer algorithms based on active contours and level sets such as graph cuts and Markov random fields [104,107]. Although these conventional methods have achieved acceptable performance for some applications, image segmentation still remains a challenging task due to various image degradation factors such as noise, blur, and contrast. To address these issues, numerous deep learning methods have been

developed and have been shown to achieve remarkable performance. This is due to the powerful feature learning capabilities of DNNs, which allows DNNs to have reduced sensitivity to image degradation factors compared to the conventional methods. Popular neural network architectures used for DL-based segmentation includes CNNs, encoder-decoder models, and multiscale architectures; see [107] for details. Two popular DNN architectures used for image segmentation problems are U-Net [76] and SegNet [106]; see [107] for examples of images before and after processing.

Image segmentation is an important step in analyzing X-ray radiographs from, for example, inertial confinement fusion (ICF) experiments [79]. ICF experiments typically use single or double shell targets which are imploded as the laser energy or laser-induced X-rays rapidly compress the target surface. X-ray and neutron radiographs of the target provide insight to the shape of target shells during the implosion. Contour extraction methods are used to extract the shell shape to conduct shot diagnostics such as quantifying the implosion and kinetic energy, identifying shell shape asymmetries, and determining instability information [79]. Uses U-Net [76], a CNN architecture for image segmentation, to output a binary masked image of the outer shell in ICF images. The shell contour is then extracted from the masked image using edge detection and shape extraction methods. Due to the limited number of actual ICF images, a synthetic dataset consisting of 2000 experimental-like radiographs is used to train the U-Net. In addition, the synthetic dataset provides ground-truth ICF image-mask pairs, which are required to train U-Net. The trained U-Net is tested on experimental images and has successfully extracted the binary mask of high-signal-to-noise ratio ICF images as shown in Figure 5.

Another example of X-ray image segmentation is for the Magnetized Linear Inertial Fusion (MagLIF) experiments at Sandia National Laboratory's Z-facility [108]. The MagLIF experiments compresses a cylindrical beryllium tube, also known as a liner, filled with pure deuterium fuel using a very large electric current on the order of $O(20MA)$. Before compression, the deuterium fuel is pre-heated and an axially oriented magnetic field is applied. The electric current causes the liner to implode and compresses the deuterium fuel in a quasi-adiabatic implosion. The magnetic field flux is also compressed which aids in the trapping of charged fusion particles at stagnation. X-ray radiographs are taken during the implosion process for diagnostics and to analyze the resulting plasma conditions and liner shape. To better analyze the implosion, a CNN model is proposed to segment the captured X-ray images into fuel strand and background. The CNN is trained using synthetically generated and augmented dataset of 10,000 X-ray images and their corresponding binary masks. The trained CNN is tested on experimental images where the results generally demonstrate excellent fuel-background segmentation performance. The worst segmentation performance is due to factors such as excessive background noise and X-ray image plate damage.

3.3.3 Image classification and object detection

Image classification, a fundamental problem in computer vision, aims to assign labels or categories to images or specific regions in images. It is known to form the basis of other computer vision tasks including segmentation and object detection. Traditional approaches to solve the classification problem typically use a two stage approach, where handcrafted features extracted from the

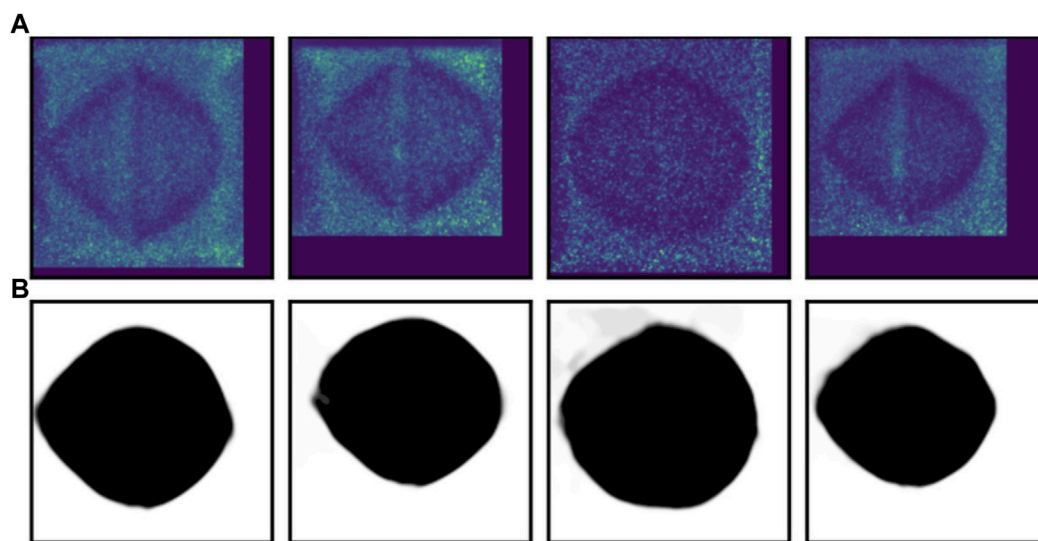


FIGURE 5
Results of image segmentation using U-Net on experimental ICF images (A) and the corresponding output masks (B). Reproduced with permission from [79].

image are used to train a classifier. The traditional approaches suffer from low classification accuracy due to the heavy dependence on the design of the handcrafted features. DL approaches can easily overcome this challenge by exploiting neural network layers for automated feature extraction, transformation, and pattern analysis. CNNs are the most popular neural network architecture used for image classification [109,110] due to their capability of reducing the high dimensionality of images without information loss, as discussed in Section 3.2.1. In addition, recall in Section 3.2.1, the CNN architectures AlexNet, VGG, GoogLeNet, and ResNet were top 3 finishers in the ILSVRC. The ILSVRC is an annual software contest where algorithms compete to correctly classify images in the ImageNet database.

Object detection builds upon image classification by estimating the location of the object in an input image in addition to classifying the object. As a result, the workflow for traditional detection algorithms can be broken down into informative region selection, feature extraction, and classification. For informative region selection, a multiscale sliding window (bounding box) is used to scan the image to determine regions of interest. Feature extraction is used on the selected region, which is then used for object classification. However, traditional methods are time consuming and robust algorithms are difficult to design. For example, a large number of candidate sliding windows need to be considered or the algorithm may return bad regions of interest. In addition, the imaging conditions can vary significantly due to factors such as lighting conditions, backgrounds, and distortion effects. Again, DL algorithms can overcome these challenges due to their capability of learning complex features using robust training algorithms [111–113]. Popular DL object detection models are Fast R-CNN [114,115] which jointly optimizes the classification and bounding box regression tasks, You Only Look Once (YOLO) [116] which uses a fixed-grid regression, and Single Shot MultiBox Detector (SSD) [117] which improves upon YOLO using multi-reference and multi-resolution techniques.

3.3.4 Compression

Image compression is the process of reducing the file size, in bytes, without reducing the quality of the image below a threshold. This process is important in order to save memory storage space and to reduce the memory bandwidth to transmit data, especially for running image processing algorithms on edge devices. The fundamental principle of compression is to reduce spatial and visual redundancies in images, by exploiting inter-pixel, psycho-visual, and coding redundancies. Conventional methods commonly leverage various quantization and entropy coding techniques [118]. Popularly used conventional methods for lossy and lossless compression includes but are not limited to JPEG [119], JPEG2000, wavelet, and PNG. While conventional methods are widely used for both image and video compression, their performance is not the most optimal for all types of image and video applications. DL approaches can achieve improved compression results due to several factors. DNNs can learn non-linear mappings to capture the compression process as well as extract the important underlying features of the image through dimensionality reduction. For example, an encoder network or CNN can extract important features into a latent feature space for compact representation. In addition, DNNs can implement direct end-to-end methods using networks such as encoder-decoders to directly obtain the compressed image from an input sharp image. Furthermore, once a DNN is trained, the inference time is much faster. For DL-based image compression methods, the most commonly used neural network architectures are CNNs, encoder-decoders, and GANs [118].

3.3.5 Sparse sampling

A closely related process to image compression is sparse sampling. While compression aims to reduce the file size, sparse sampling, also known as compressed sensing (CS), aims to efficiently acquire and reconstruct a signal by solving underdetermined linear systems. It has been shown in CS theory that a signal can be

recovered from sampling fewer measurements than required by the Nyquist-Shannon sampling theorem [120]. As a result, both memory storage space and data transmission bandwidth can be reduced. In conventional methods, CS algorithms need to overcome two main challenges: the design of the sampling and reconstruction matrices. Numerous methods have been proposed including but not limited to random and binary sampling matrices and reconstruction methods using convex-optimization and greedy algorithms [121]. However, these conventional methods suffer from long computational times or low quality reconstruction. DL approaches allow for fast inference (reconstruction) times for a trained network, as well as learning non-linear functions for higher quality signal reconstruction [121,122]; see [123] for examples of images before and after processing.

Neural network (NN) models that learn to invert X-ray data have also been shown to significantly reduce the sampling requirements faced by traditional iterative approaches. For example, in ptychography, traditional iterative phase retrieval methods require at least 50% overlap between adjacent scan positions to successfully reconstruct sample images as required by Nyquist-Shannon sampling. In contrast, Figure 6B shows image reconstructions obtained from PtychoNN when sampled at $25\times$ less than required for conventional phase retrieval methods [124]. Figure 6A shows the probe positions and intensities, there is minimal overlap between probes. Through use of inductive bias provided through online training of the network [125], PtychoNN is able to reproduce most of the features seen in the sample even when provided extremely sparse data. Figure 6C shows the same region reconstructed using an oversampled dataset and traditional iterative phase retrieval. Furthermore [125], demonstrated live inference performance during a real experiment using an edge device and running the detector at its maximum frame rate of 2 kHz.

In the previous example, DL is used to reduce sampling requirements but not to alter the sampling strategy. In other words, the scan proceeds using a conventional acquisition strategy, but using fewer points along that trajectory than traditionally required. In contrast, active learning approaches are being developed that use data-driven priors to direct the acquisition strategy. Typically, this is treated as a Bayesian optimization (BO) problem using Gaussian processes (GPs). This method has been applied to a variety of characterization modalities including scanning probe microscopy [126], X-ray scattering [127], and neutron characterization [128]. A downside to such approaches is that the computational complexity typically increases as $O(N^3)$ with the action space [129], making real-time decision a challenge. To address these scaling limitations which are critical especially in fast scanning instruments, recent work has demonstrated the use of pre-trained NNs to make such control decisions [130,131]. Figure 7 shows the workflow and results from the Fast Autonomous Scanning Toolkit applied to a scanning diffraction X-ray microscopy measurement of a WSe₂ sample. Starting from some quasi-random initial measurements, FAST generates an estimate of the sample morphology, predicts the next batch of 50 points to sample from, triggers acquisition on the instrument, analyzes the image after the next set of points has been acquired and continues the process until the improvement in sample image is minimal. Figures 7B, A, C, and E show the predicted image after 5%, 15%, and 20% sampling while Figure 7 B, D, and F shows the points preferentially selected by the AI. The AI has learned to prioritize

acquisition where the expected information gain is maximum, e.g., around contrast features on the sample.

3.3.6 3D reconstruction

Image-based 3D reconstruction is the process of inferring a 3D structure from a single or multiple 2D images, and is a common topic in the fields of computer vision, medical imaging, and virtual reality. This problem is well known to be an ill-posed inverse problem. Conventional methods attempt to formulate a mathematical formula for the 3D to 2D projection process, use prior models, 2D annotations, and other techniques [132,133]. In addition, high quality reconstruction typically requires 2D projections from multiple views or angles, which may be difficult to calibrate (i.e., cameras) or time consuming to obtain (i.e., CT) depending on the application. DL techniques and the increasing availability of large datasets motivates new advances in 3D reconstruction by address challenges found in conventional methods. The popular networks used for image-based 3D reconstruction are CNNs, encoder-decoder, and GAN models [132]; see [132] for examples on 2D to 3D reconstruction.

X-ray phase information is now available for 3D reconstruction in the state-of-the-art X-ray sources such as synchrotrons and XFELs. In contrast to iterative phase retrieval methods that incorporate NNs through a DIP or other means, single-shot phase retrieval NNs provide sample images from a single pass through a trained NN. The inference time on a trained NN is minimal and such methods are hundreds of times faster than conventional phase retrieval [134,135]. Figures 8A, B compare AutoPhaseNN and traditional phase retrieval for 3D coherent image reconstruction, respectively [136]. AutoPhaseNN is trained to invert 3D coherently scattered data into sample image in a single shot. Once trained AutoPhaseNN is $>100\times$ faster than iterative phase retrieval with some reduction in accuracy. The prediction from AutoPhaseNN can also be used to seed phase retrieval, i.e., provide an initial estimate which can be refined by a few iterations of phase retrieval. This combined approach of NN + phase retrieval is shown to be both faster and more accurate than iterative phase retrieval.

A recent work by Scheinker and Pokharel [137] developed an adaptive CNN-based 3D reconstruction method for coherent diffraction imaging (CDI), a non-destructive X-ray imaging technique that provides 3D measurements of electron density with nanometer resolution. The CDI detectors record only the intensity of the complex diffraction pattern of the incident object. However, all phase information is lost in this detection method, and thus results in an ill-posed inverse Fourier transform problem to obtain the 3D electron density. Conventional methods encounter many challenges including expert knowledge, sensitivity to small variations, and heavy computation requirements. While DL methods currently cannot completely substitute conventional methods, they can speed up the 3D reconstruction speed given an initial guess, and can be fine-tuned using conventional methods to achieve better performances. For CDI 3D reconstruction, Scheinker and Pokharel [137] proposes a 3D CNN architecture with model-independent adaptive feedback agents. The network takes in 3D diffracted intensities as inputs and outputs a vector of spherical harmonic coefficients, which describe the surface of the 3D incident object. The adaptive feedback agents take as input the

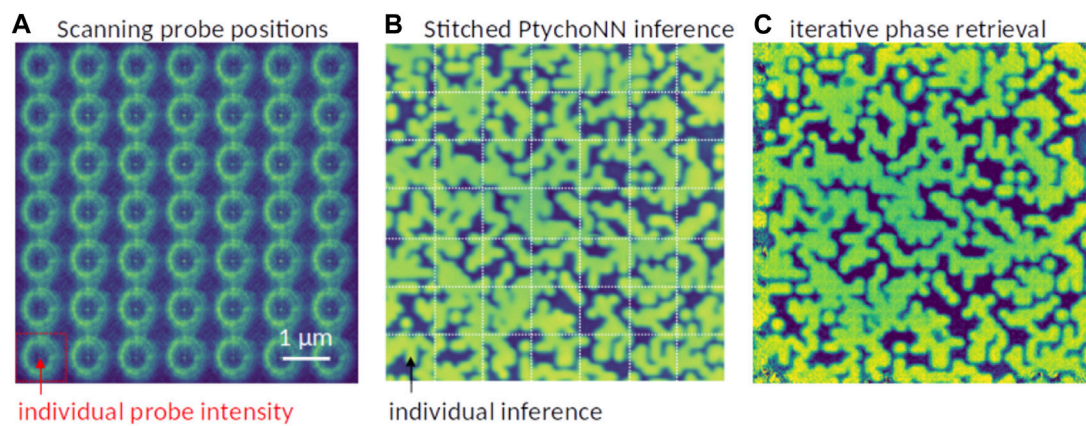


FIGURE 6
Sparse-sampled single-shot PtychoNN reconstruction using PtychoNN. **(A)** Scanning probe positions with minimal overlap. **(B)** Single-shot PtychoNN predictions on $25 \times$ sub-sampled data compared to **(C)** ePIE reconstruction of the full resolution dataset.

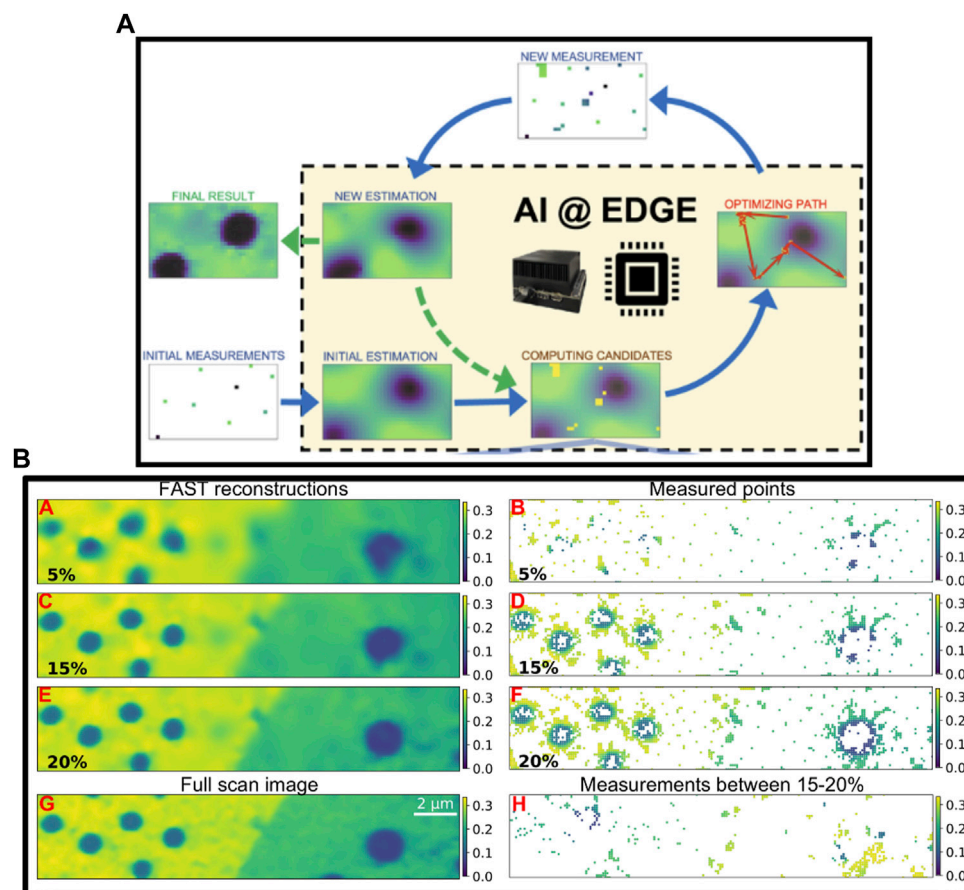


FIGURE 7
FAST framework for autonomous experimentation. **(A)** shows the workflow that enables real-time steering of scanning microscopy experiments. **(B)** shows reconstructed images at 5%, 15% and 20% sampling along with the corresponding locations from which they were sampled. In addition, the full-grid pointwise scan and corresponding points sampled between 15% and 20% is also shown. Reproduced with permission from [131].

spherical harmonics to adaptively adjust the intensities, positions, and decay rates of a collection of radial basis functions. The 3DCNN is trained using a synthetic dataset consisting of 500,000 training set

of 49 sampling coefficients as well as the spherical surface and volume of each in order to perform a 3D Fourier transform. An additional 100 random 3D shapes and their corresponding 3D

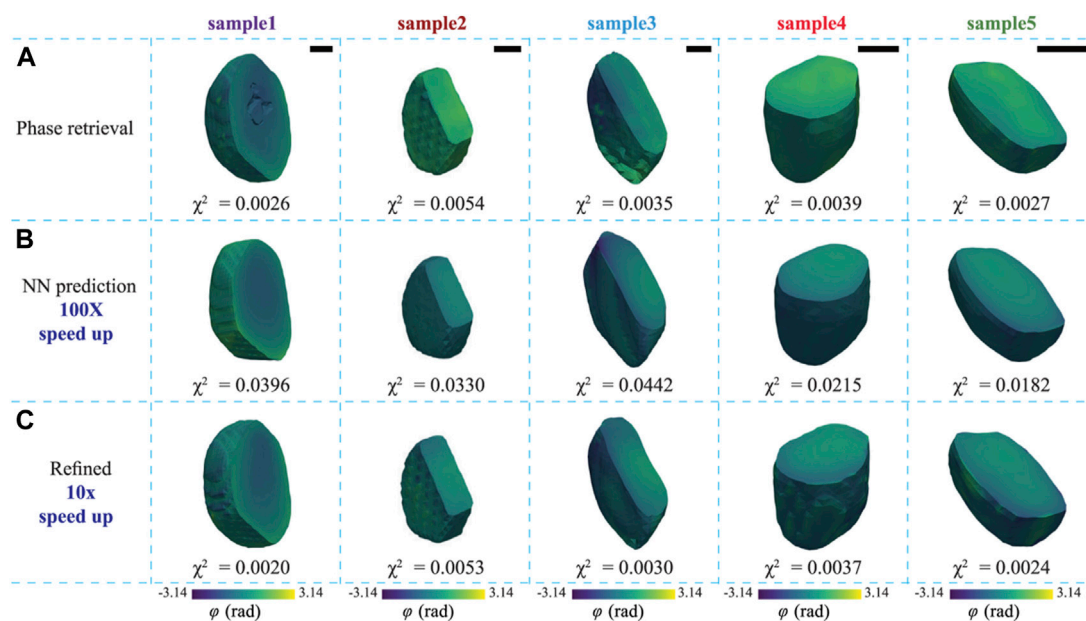


FIGURE 8
Comparison of 3D sample images obtained by (A) phase retrieval, (B) AutoPhaseNN, and (C) AutoPhaseNN + phase retrieval. Reproduced with permission from [136].

Fourier transforms are used to test the adaptive model-independent feedback algorithm, with the CNN output as its initial guess. Last, the robustness of the trained 3DCNN is tested on the experimental data of a 3D grain from a polycrystalline copper sample measured using high-energy diffraction microscopy. Results show that the 3DCNN provides an initial guess that captures the average size and a rough estimate of the shape of the grain. The adaptive feedback algorithm uses the 3DCNN initial guess to fine-tune the harmonic coefficients to match and converge the generated and measured diffraction patterns of the grain.

4 Hardware solutions for deep learning

DNNs have been implemented for many imaging processing tasks ranging from enhancement to generation as discussed above. To achieve good performance, these algorithms use very deep networks which can be very computationally intensive during training and inference in their own ways. During training, DNNs are fed large amounts of data and a large number of computations must be performed to update network weights to achieve accurate predictions. For example, AlexNet [68] took five to six days to train on two NVIDIA GTX 580 graphical processing units. As a result, powerful computing hardware is needed to accelerate DNN training. Meanwhile, during inference, larger networks require more computing power and memory storage space, and thus results in higher energy consumption and latency to obtain predictions in real-time. For very large networks such as AlexNet, a single forward pass may require millions of multiply and accumulate (MAC) operations, thus making DNNs both computationally and energy costly. For real-time data processing in imaging devices, DNN algorithms need to be executed with low

latency, limited energy, and other design constraints. Hence, there is a need to develop cost and energy efficient hardware solutions for DL applications.

Interestingly, neural network algorithms are known to have at least two types of inherent parallelism, namely, model and data parallelism [138]. Model parallelism refers to the partitioning of the neural network weights for MAC operations for parallel execution as there are no data dependencies. Data parallelism refers to processing the data samples in batches rather than a single sample at a time. Hardware accelerators can exploit these characteristics by implementing parallel computing paradigms. This section presents different hardware accelerators used for DL applications. Note that the best hardware solution is dependent on the application and corresponding design requirements. For example, edge computing devices such as cameras and sensors may require small chip area with limited power consumption.

4.1 Electronic-based accelerators

The electronic-based hardware solutions for DL are broad, ranging from general purpose processors such as central processing units (CPUs) and graphical processing units (GPUs), field-programmable gate arrays (FPGAs), to application-specific integrated circuits (ASICs). The circuit architecture design typically follows either temporal or spatial architectures [139] as shown in Figures 9A, D. The architectures are similar in using multiple processing elements (PEs) for parallel computing, however, there are differences in control, memory, and communication. The temporal architecture features a centralized control for simple PEs, consisting of only arithmetic logic units (ALUs), which can only access data from the centralized memory. Meanwhile, the spatial

architecture features a decentralized control scheme with complex PEs, where each unit can have its own local memory or register file (RF), ALU, and control logic. The decentralized control scheme forms interconnections between neighboring PEs to exchange data directly, allowing for dataflow processing techniques.

4.1.1 Temporal architectures: CPUs and GPUs

CPUs and GPUs are general purpose processors that typically adopt the temporal architecture as shown in Figures 9B, C. Modern CPUs can be realized as vector processors, which adopt the single-instruction multiple-data (SIMD) model to process a single instruction on multiple ALUs simultaneously. In addition, CPUs are optimized for instruction-level parallelism in order to accelerate the execution time of serial algorithms and programs. Meanwhile, modern GPUs adopt the single-instruction multiple threads (SIMT) model to process a single instruction across multiple threads or cores. Different from CPUs, GPUs are made up of more specialized, parallel, and smaller cores than CPUs to efficiently process vector data with high performance and reduced latency. As a result, GPU optimization relies on software defined parallelism rather than instruction-level parallelism [140]. Both the SIMD and SIMT execution models for CPUs and GPUs, respectively, allow for parallel MAC operations for accelerated computations.

Nonetheless, CPUs are not the most used processor for DNN training and inference. Compared to GPUs, CPUs have a limited number of cores, and thus a limited number of parallel executions. For example, one of Intel's server-grade CPUs is the Intel Xeon Platinum 8280 processor which can have up to 28 cores, 56 threads, 131.12 GB/s maximum memory bandwidth, and 2190 Giga-floating point operations per second (GFLOPS) for single-precision compute power. In addition, AMD's server-grade EPYC 9645 features 96 cores, 192 threads, and a memory bandwidth of 460.8 GB/s. In comparison, NVIDIA's GeForce RTX 2080 Ti is a desktop-grade GPU with 4352 CUDA cores, 616.0 GB/s memory bandwidth, and 13450 GFLOPS single-precision compute power. Furthermore, a recently released NVIDIA RTX 4090 desktop-grade GPU features 16,834 CUDA cores, 1008 GB/s memory bandwidth, and 82.85 TFLOPS single-precision compute power. Therefore, GPUs outperform CPUs in terms of parallel computing.

For DL at the edge, the hardware industry has developed embedded platforms for AI. One popular platform is the NVIDIA Jetson for next-generation embedded computing. The Jetson processor features a heterogeneous CPU-GPU architecture [141] where the CPU accelerates the serial instructions and the GPU accelerates the parallel neural network computation. Furthermore, the Jetson is designed with a small form factor, size, and power consumption. A broad survey by [142] presents different works using the Jetson platform for DL applications such as medical, robotics, and speech recognition. Several surveyed works have used the Jetson platform to implement imaging processing tasks including segmentation, object detection, and classification.

Also using the NVIDIA Jetson platform, a work by [143] investigates the performance of the Jetson TX2 for edge deployment for TomoGAN [86], an image denoising technique using generative adversarial networks (GANs) for low-dose X-ray images. The training and testing datasets consist of 1024 pairs of images of size 1024×1024 with each image pair consisting of a noisy image and its corresponding ground truth. The pre-trained

TomoGAN network is deployed and tested on the Jetson TX2 and a laptop with an Intel Core i7-6700HQ CPU @2.60GHz with 32GB RAM. The laptop CPU achieves an average inference performance of 1.537 s per image, while the TX2 achieves an inference performance of 0.88 s per image, approximately 1.7× faster than the laptop CPU.

A recent work by [144] investigates the classification accuracy of tuberculosis detection from chest X-ray images using MobileNet [145], ShuffleNet [146], SqueezeNet [147], and their proposed E-TBNet. In addition, they further investigate the inference time during testing of each network on the NVIDIA Jetson Xavier and a laptop with Intel Core i5-9600KF CPU and NVIDIA Titan V GPU. The dataset consists of 800 chest X-ray images scaled to size $512 \times 512 \times 3$. The MobileNet network achieves the highest accuracy at 90% while their proposed E-TBNet achieves 85%. However, the inference time for E-TBNet is the fastest for all investigated networks with an inference time of 0.3 m and 3 m per image when deployed on the laptop with Titan GPU and Jetson Xavier, respectively. The slowest reported inference time for the Jetson Xavier is 6 m per image for the ShuffleNet. Although the inference time for the Xavier is an order of magnitude slower, classification inference can be achieved in real-time with smaller hardware footprint for edge deployment.

4.1.2 Spatial architectures: FPGAs and ASICs

Field-programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs) typically adopt the spatial architecture as shown in Figures 9E, F. FPGAs and ASICs are specialized hardware that are tailored for specific applications due to their design process. FPGAs can be configured to perform any function as it is made up of programmable logic modules and interconnecting switches as shown in Figure 9E. The FPGA software is used to directly build the logic and data flow directly into the chip architecture. On the other hand, ASICs are designed and optimized for a single application, and cannot be reconfigured. Nonetheless, the spatial architecture of FPGAs and ASICs makes them well suited for neural network computations as the mathematical operations of each layer are fixed and known *a priori*. As a result, FPGAs and ASICs can attain highly optimized performance.

As shown in Figure 9D, the spatial architecture consists of an array of PEs interconnected with a Network-on-Chip (NoC) design, allowing for custom data flow schemes. Although not shown in Figure 9D, the memory hierarchy consists of three levels. The lowest level consists of the RF in each PE, which is used to locally store data for inter-PE data movement or local accumulation operations. The middle level consists of a global buffer (GB) that holds the neural network weights and inputs to feed the PEs. The highest level is the off-chip memory, usually a DRAM, to store the weights and activations of the whole network. MAC operations need to be performed on large data sets. Hence, the major bottleneck is the high latency and energy costs of DRAM accesses. A comparison between DianNao and Cambricon-X, two CNN accelerators, show that DRAM accesses consume more than 80% of the total energy consumption [148]. In addition [149], reports that the energy cost of DRAM access is approximately 200× more than a RF access. Therefore, energy efficiency can be greatly improved through the reduction of DRAM accesses, commonly done by exploiting the idea of data reuse.

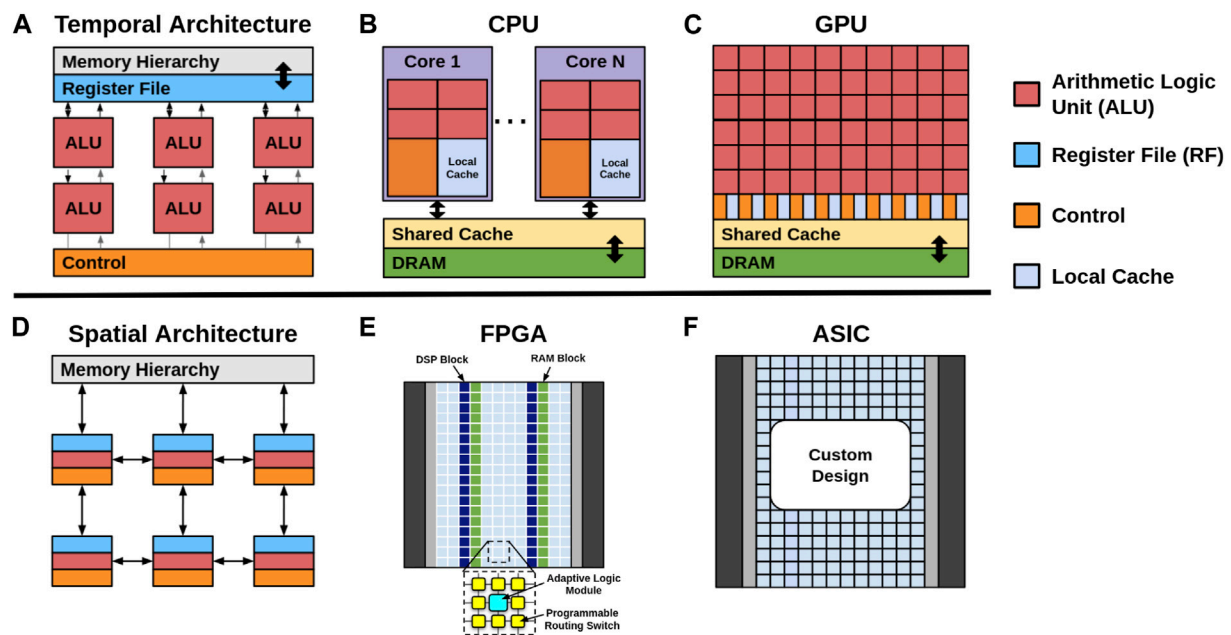


FIGURE 9
Basic models of the (A) temporal, (B) CPU, (C) GPU, (D) spatial, (E) FPGA, and (F) ASIC architectures.

The focus of data reuse is to utilize the data already stored in RFs and the GB as often as possible. This gives rise to the investigations of efficient data flow paradigms in both the spatial and temporal operations of PEs. For example, in fully connected layers, the input reuse scheme is popular since the input vector is dot multiplied by each row of the weight matrix to compute the layer output. For convolutional layers, the weight reuse scheme is popular as the weight kernel matrix is used for multiple subsets of the input feature map. In addition for convolutional layers, convolutional reuse can be applied by exploiting the overlapping region of the sliding window of kernel weights and the input feature map. Additional data reuse schemes are the weight stationary, output stationary, row stationary, and no local reuse schemes. A detailed discussion of the data reuse schemes is out of scope for this paper. However, for a comprehensive review, see details in [139,150,151]. In summary, optimizing the data flow is crucial for FPGAs and ASICs to attain high energy efficiency.

Nonetheless, it is important to note the challenges faced by FPGAs, and in turn ASICs, have in implementing DL networks. A few challenges include but are not limited to memory storage requirements, memory bandwidth, and large computational requirements on the order of Giga-operations per second (GOPS). For example, AlexNet requires 250 MB of memory with 32-bit representation to store 60 million model parameters and 1.5 GOPS for each input image [152], while VGG has 138 million model parameters and requires 30 GOPS per image [153]. Commercial FPGAs do not have enough memory storage space and thus requires external memory to store model parameters, which needs to be transmitted to the FPGA during computation. One way to address this issue, is to compress the neural network by reducing its size through methods such as compression and quantization [154,155]. For example, SqueezeNet [147] can be thought of a compressed AlexNet with 50× fewer model

parameters and <0.5 MB model size. On the other hand, quantization reduces the number of data bits or transforms floating point data to fixed point data to reduce the computational burden. As a result, FPGA-based implementation of DL models will suffer a degree of accuracy loss.

The energy efficiency and massive parallelism of FPGA and ASIC-based accelerators make them desirable for edge computing. A recent work [156] develops a lightweight CNN architecture called SparkNet for image classification tasks. SparkNet features approximately 3× less parameters compared with the SqueezeNet, and approximately 150× less parameters than AlexNet. In addition, a comprehensive design is presented to map all layers of the network onto an Intel Arria 10 GX1150 FPGA platform with each layer mapped to a its own hardware unit to achieve simultaneously pipelined work, increasing throughput. SparkNet is tested on 4 benchmark image classification datasets, i.e., MINIST, CIFAR-10, CIFAR-100 and SVHN. The performance and average time for the Intel FPGA, NVIDIA Titan X GPU, and Intel Xeon E5 CPU to process 10,000 $32 \times 32 \times 3$ is reported. The FPGA-based accelerator achieves a processing time of 11.18 μ s, which is 41× and 9× faster than the CPU and GPU, respectively. Furthermore, the FPGA average power consumption is 7.58 W with a performance of 337.2 Giga operations per second (GOP/s), making the FPGA more energy and computationally efficient compared to the CPU (95 W, 8.2 GOP/s) and GPU (250 W, 39.4 GOP/s).

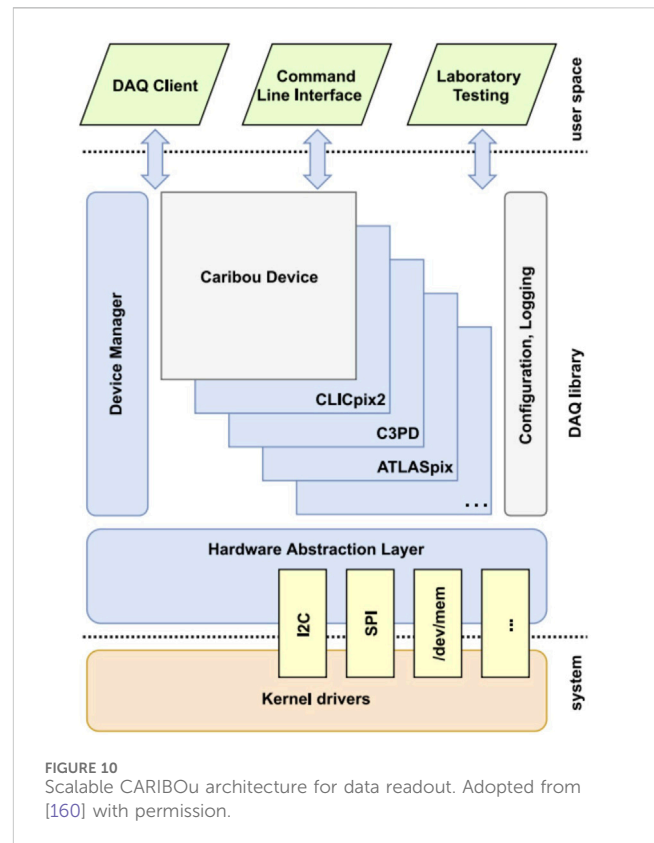
Another recent work [157] uses FPGAs to deploy MobileNet for face recognition in a video-based face tracking system. The work further integrates the FPGA with CPUs and GPUs to build a heterogeneous system with a delay-aware energy-efficient scheduling algorithm to achieve reduced execution time, latency, and energy cost. The face tracking experiment is run using an Intel Gold 5118 CPU, NVIDIA Tesla P100 GPU, and the Intel Arria 10 GX 900 and Intel Stratix 10 GX1100 FPGAs. The reported

experimental results evaluate the computing speed and power efficiency of the FPGA-based accelerator compared to the CPU and GPU, as well as the efficiency of the combined detection system with CPU/GPU/FPGA. The FPGA accelerators achieve a computational speed that is approximate to or better than the GPU, while achieving superior power efficiency in GOP/s/W. The difference in performance of the FPGAs is due to their hardware specifications, where the hardware richer Intel Stratix will out perform the Intel Arria. Lastly, the experimental results report that the CPU/GPU/FPGA system can achieve optimal performance in comparison to using only one or a combination of two different accelerators. This is due to the energy efficient scheduling algorithm to optimally pipeline tasks to the different accelerators. The idea of utilizing a heterogeneous system and scheduling algorithm to improve computational and energy efficiency can be explored to address challenges of edge computing.

The Google Edge Tensor Processing Unit (TPU) platform [158] is a general purpose ASIC designed and built by Google for inference at the edge. One example product is the Dual Edge TPU which features an area footprint of $22 \times 30 \text{ mm}^2$, peak performance of 8 trillion operations per second (TOPS), and power consumption of 2 TOPS/W. Other hardware options are available for ASIC prototyping and deployment for edge devices. A survey by [159] presents works that use the Edge TPU platform for DL applications such as image classification, object detection, and image segmentation.

The previously discussed work [143], which deployed TomoGAN on the NVIDIA Jetson platform for X-ray image denoising, also deployed it on the Edge TPU. The work presents a quantized model of TomoGAN to address limitations of the Edge TPU, such as output size. A fine tuning model is also presented to improve the output quality of the quantized model. The Edge TPU's average inference time is 0.554 s per image, which is faster than the Jetson TX2 inference time of 0.88 s per image. In addition, the power consumption is reduced to 2 W compared to Jetson TX2's 7.5 W.

In addition, there is interest in the development of software and firmware for modular and scalable implementation of energy efficient algorithms on FPGA platforms. One such example is for high-speed readout systems for pixel detectors. Oak Ridge National Laboratory (ORNL), through the support of the Department of Energy (DOE) in High Energy Physics (HEP) and Nuclear Physics (NP), is leading the design of a new generic readout system for pixel detectors based on the successful first-generation system, the CARIBOu 2.0 [160]. The CARIBOu 2.0 system, shown in Figure 10, will be the proposed architecture for the platform. The concept of the system is to provide a generic framework for the readout of ASIC detectors for research and development and scalable to larger detector arrays. CARIBOu 2.0 shares knowledge and code to provide the community with a convenient platform that maximizes reusability and minimizes overhead when developing such systems. ORNL will initially implement the readout firmware and software specific to the Timepix4 or to commercial CMOS image sensors, SMALLGAD, Photon-to-Digital Converters (PDCs), and the interconnect for the assemblies. The hardware platform is based on Xilinx Ultrascale + FPGA, that provides resources for CPU and FPGA side data processing at high speed. Using the resources of this modern FPGA, software and firmware will be developed to flexibly implement data processing and reduction, edge computing,



by using conventional and ML algorithms running in the FPGA. For larger data rates, firmware will be developed to move the data to a FELIX card, which can handle up to 24 CARIBOu 2.0 systems and transmit data via a high-speed network interface to a data center or process them locally via GPU and CPU in the FELIX host machine. As a result, the system can be scaled up to the readout of large smart sensor stack arrays.

Furthermore, as advancements in ASIC technology have enabled greater integration of digital functionalities for scientific applications, there has been growing interest in incorporating compression capabilities directly within ASIC detectors to enhance data processing speed. ASIC architectures capable of frame rates approaching 1 MHz have been designed, providing a viable solution for enhancing the speed of various diffraction techniques employed at X-ray light sources, including those relying on coherent imaging methodologies like ptychography [161]. Developing ASIC compression strategies that exploit the structure in detector data enables high compression performance while requiring lower computational complexity than commonly used lossless compression methods like LZ4 [161].

4.1.3 Summary and limitations

We have presented an overview of 4 different electronic-based accelerators and a few works applying them to DL applications at the edge. Figure 11A shows that there is a clear trade-off between programmability and efficiency. To attain higher performance and power efficiency, FPGAs and ASICs require more design complexity to optimize data flow, while ASICs need further hardware optimization. Correspondingly, the time-to-market increases with design complexity. DL algorithms can be deployed

at the edge using these existing electronic-based hardware accelerators.

However, in recent years, these electronic-based accelerators are constantly reaching performance limits in latency, energy consumption, high interconnect cost, excessive heat, and other physical constraints [162]. Figure 11B illustrates the past 50 years of CPU trends in regards to the number of transistors, single-thread performance, frequency, typical power consumption, and number of cores. The trends show that the number of transistors and correspondingly the power consumption continues to grow. Furthermore, the trends indicate that CPU clock frequency has plateaued since around 2005 while single-thread performance and number of cores are slowly tapering. On the other hand, GPU performance has not been limited and is the most popularly used hardware for deep learning training. The single-precision computational throughput of GPUs continues to grow [163]; summarizes the increasing peak performance trends of Nvidia GPUs in GFLOPS from 2006 to 2018. As a result of the high computational speeds, GPUs consume more power [163]. Shows that the power consumption of GPUs increases with the computational throughput in GFLOPS and [164] shows that GPUs have higher computational speed than FPGAs and ASICs at the cost of higher power consumption. Meanwhile, FPGAs and ASICs can achieve good computational performance with lower energy consumption at the cost of design time to develop data flow algorithms and to optimize hardware.

In addition to the hardware performances, the unit prices of each hardware should also be taken into consideration. Recall in Section 4.1.1 we compared the parallel computing performance between server-grade CPUs (Intel Xeon Platinum 8280 and AMD EPYC 9645) and desktop-grade GPUs (NVIDIA RTX 2080 and 4090), where the GPUs outperform CPUs due to the higher number of computing cores. Not only can desktop GPUs perform better than server-grade CPUs, they are also more price efficient. The NVIDIA RTX 2080 and 4090 have starting prices around \$1,000 and \$1,600, respectively, while the Xeon Platinum 8280 and AMD EPYC 9645 cost over \$5,000. The unit prices of FPGAs can vary from as low as a few USD to thousands of USD depending on various factors such as the manufacturer, the number of units, the number of configurable logic blocks, the number of input/output connections, and the amount of available RAM [165]. On the other hand, the unit prices of custom designed ASICs can be lower than that of FPGAs, but only when purchased in large quantities [166]. The starting cost of ASICs is easily over \$1,000 as it suffers from very high non-recurring engineering (NRE) costs [167], while FPGAs have no NRE costs. Nonetheless, Google offers prototyping products using the Edge TPU starting at \$60.

At any rate, electronic accelerators are traditionally designed to follow von Neumann architecture where the processor and memory units are connected by buses [168], which inherently increases data transfer and power consumption during computation [148]. Demonstrates that more than 75% of the energy utilized by processors comes from DRAM accesses. These limits in electronic based computing gives rise to a shift in focus to analog neuromorphic computing and non-von Neumann architectures such as optical neural networks and bio-inspired spiking neural networks for high-speed, energy-efficient, and parallel computing [169,170].

4.2 Neuromorphic hardware outlook

This section presents two emerging neuromorphic hardware solutions, namely, optical neural networks (ONNs) and spiking neural networks (SNNs), as promising architectures for highly energy efficient and parallel processing.

4.2.1 Optical neural networks

ONNs have emerged as a promising avenue for achieving high-performance and energy-efficient computing, given their compute-in-light speed, ultra-high parallelism, and near-zero computation energy [171–174]. Series of photonic tensor cores (PTCs) are designed to enhance the execution of linear matrix operations, the fundamental operations in AI and signal processing, with coherent photonic integrated circuits [171,175], micro-ring resonators [176], photonic phase-change materials [177–179], and diffractive optics [180–182].

Comparing metal wire connections, optical signals modulated at different wavelengths, can be concurrently processed using wavelength-division multiplexing (WDM) within the waveguide and photonic tensor cores [175,183]. Besides, waveguides are free from inductance, which means frequency-dependent signal distortions are negligible for the extended connections in neural interconnects. Hence, given the extensive parallel signal fan-out and fan-in requirements in neural networks, the physical implementation based on PTCs offers distinct advantages. On the basis of the linear optical computing paradigms, ONNs have been constructed for various machine learning tasks such as image classification [182,184,185], vowel recognition [171], and edge detection [186]. Photonic computing methods [187] also feature great potential for supporting advanced Transformer models. Furthermore, ONNs holds significant promise for real-time image processing, where they process image signals directly in light fields, as opposed to after digitalization [188–191]. For instance, recent advancements include the proposal of an image sensor with an ONN encoder [188], which filters relevant information within a scene using an energy-efficient ONN decoder before detection by image sensors.

Despite their advantages, PTCs face significant challenges related to cross-domain signal conversion energy overhead, specifically in analog-to-digital (A/D) and digital-to-analog (D/A) conversion. Moreover, the physical layout constraints of PTCs, manufacturing complexities, and elevated costs have made scalability a primary obstacle in the broad adoption of ONNs. For example, Mach-Zehnder interferometer (MZI)-based PTCs [171] require $O(m^2 + n^2)$ bulky MZIs and approximately $\sim (m + n)$ cascaded MZIs within a single optical path to implement an n -input, m -output layer. Current state-of-art ONNs therefore employ time-division multiplexing with WDM, trading bandwidth and chip complexity.

Efficient analog-to-digital conversion solution [192] and various hardware-software co-design methodologies [186] have been investigated to reduce signal conversion overhead by reducing precision and energy per conversion. In pursuit of enhancing the scalability and efficiency of ONNs, researchers have delved into innovative optimizations at both the architecture and device levels. One noteworthy approach at the architecture level is the introduction of optical subspace neural networks (OSNNs),

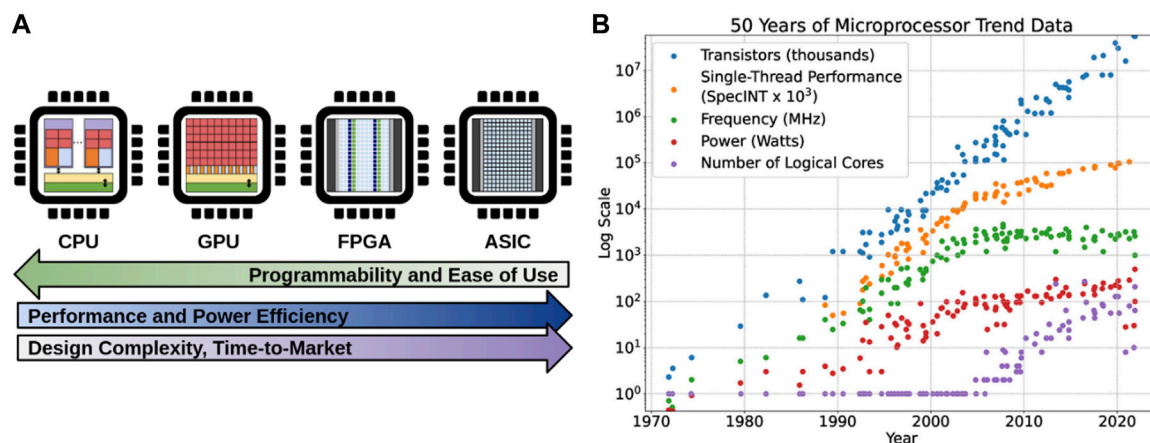


FIGURE 11
(A) Summary of electronic-based hardware comparison. (B) 50 years of CPU processor trend data from [217].

which make a trade-off between weight representation universality and the reduction of optical component usage, area costs, and energy consumption. For example, a butterfly-style OSNN as shown in Figure 12A, which achieved a remarkable reduction of 7 times in trainable optical components compared to GEMM-based ONNs, was reported and demonstrated a measured accuracy of 94.16% in image recognition tasks [184]. Without sacrificing much model expressiveness, OSNNs can reduce footprints, often ranging from one to several orders of magnitude less than previous MZI-based ONN [171].

At the device level, employing compact custom-designed PTCs, such as multi-operand optical neurons (MOON) [185,193,194], enables the consolidation of matrix operations into arrays of optical components. Figures 12B, C shows a customized multi-operand MZI-based and microring resonator-based PTCs, respectively. Instead of performing a single math operation (e.g., scalar product) per device, MOON fuses a tensor operation in the single device. Crucially, this approach retains the capability to represent general matrices while still maintaining an exceptionally compact layout, in contrast to prior compact tensor designs like star couplers and metasurfaces [195]. One specific achievement in MOON is the development of multi-operand MZI-based (MOMZI) ONN [194], which has realized a two-orders-of-magnitude reduction in propagation loss, delay, and total footprint without losing matrix expressivity. The customized ONN demonstrated an 85.89% measured accuracy in the street view house number (SVHN) recognition dataset with 4-bit control precision. The combined progress in architecture, device design, and optimization techniques is pivotal in advancing the capabilities of ONNs, making themselves efficient, scalable, and practical for AI applications.

4.2.2 Spiking neural networks

In addition to photonic neuromorphic computing, extensive research has been done for other neuromorphic computing architectures. Due to the bottleneck seen in von Neumann architectures, these computing paradigms aim to greatly reduce data movement between memory and PEs to attain high energy efficiency and parallel processing. Taking a unique approach to improve energy efficiency, neuromorphic computing architectures

are inspired by the human brain's neurons and synapses. The human brain is extremely energy efficient, where in terms of computing terminology, it is estimated to have a computing power of 1 exaFLOPS while only consuming 20 W. In recent years, there is a rise in interest to explore brain-inspired neural network computing architectures, better known as SNNs [196,197].

SNNs are a special type of artificial neural network (ANN) that closely mimics biological neural networks. While ANNs are traditionally modeled after the brain, there are still many fundamental differences between them such as neuron computation and learning rules. In addition, one major difference is the propagation of information between neurons. Biological neurons, shown in Figure 13A, transmit information to downstream neurons using a spike train of signals, or a time-series of delta functions. The individual spikes (delta functions) are known to be sparse in time and have high information content. Therefore, SNNs are designed to convey information by utilizing the spike timings and spike rates [198,199] as shown in Figure 13C. Furthermore, the advantages of the spiking event sparsity can be exploited in special hardware to reduce energy consumption while maintaining the transmission of high information content [200].

The hardware industry as well as academia are striving to develop unique solutions for neuromorphic computing chips. Intel's Loihi [201] features 128 neuromorphic cores with 1024 spiking neural units per core. A recent work [202] surveys different works that utilize Loihi as a computing platform for applications such as event-based sensing and perception, odor recognition, closed-loop control for robotics, and simultaneous localization and mapping. For medical image analysis [203], uses Loihi to implement a SNN for brain cancer MRI image classification. IBM developed TrueNorth [204], a neuromorphic chip featuring 4096 neuromorphic cores, 1 million spiking neurons and 256 million synapses. A work by [205] uses the TrueNorth computing platform to detect and count cars from input images by mapping CNNs, such as AlexNet and VGG-16, onto TrueNorth. A few other well-known SNN hardware are Neurogrid [206], BrainScaleS [207], and SpiNNaker [208], which all adopt different solutions to emulate spiking neurons. For a comprehensive review, see details in [209,210].

Due to its low power consumption, SNN hardware is a potential platform for edge computing. A work by [211] presents preliminary

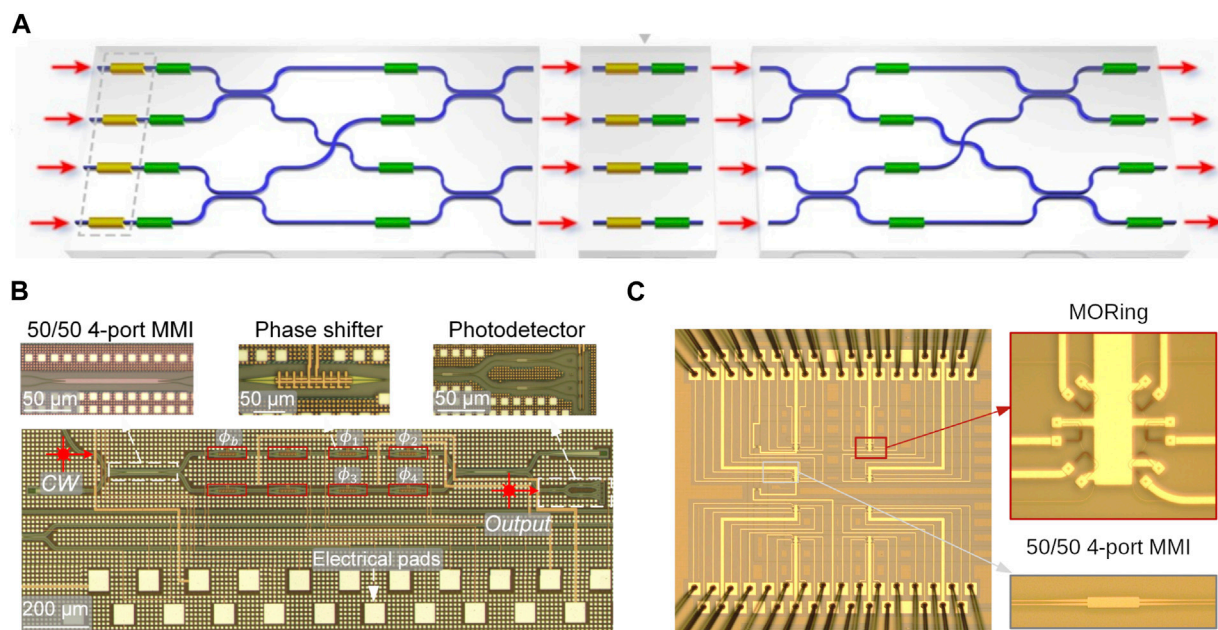


FIGURE 12 Integrated photonic chips for optical neural networks. (A) a butterfly-style PTCs to reduce the optical components from an architecture level [184]. (B,C) are customized multi-operand MZI-based and microring resonator-based PTCs, respectively, which improve scalability and efficiency at the device level [185,193,194].

results for implementing SNN on a mixed analog digital memresistive hardware for classifying neutrino scattering data collected at Fermi National Accelerator Laboratory using the MINERvA detector [212]. Two different SNNs, the neuroscience-inspired dynamic architecture (NIDA) [213] and a memresistive dynamic adaptive neural network array (mrDANNA) [214], were trained and tested on the MINERvA dataset's X view. The training and testing datasets consisted of 10,000 and 90,000 synthetic instances, respectively, generated by a Monte Carlo generator. The NIDA network was trained on the Oak Ridge Leadership Computing Facility's Titan using 10,000 computing nodes, and achieved a classification accuracy of 79.11% on the training set. Meanwhile, the mrDANNA was trained on a desktop and achieved a classification accuracy of 76.14% and 73.59% on the training and combined training and testing dataset, respectively. Both networks can attain an accuracy close to the state-of-art CNN accuracy of 80.42% while using far less neurons and synapses. In addition, the energy consumption was computed for the mrDANNA network and is estimated to be $1.66 \mu\text{J}$ per calculation. Although there is an accuracy drop using the smaller SNN networks, the energy consumption per calculation is very small, and thus can be deployed in edge devices.

A recent work (R [215]) implemented an SNN algorithm for filtering data from edge electronics in high energy collider experiments conducted at the High Luminosity Large Hadron Collider (HL-LHC), in order to reduce large data transfer rate or bandwidth (on the order of a few petabytes per second) to downstream electronics. In collider experiments, the collision events of charged particles with energy greater than 2 GeV is of significant interest. However, the high energy charged particles only comprise of approximately 10% of all recorded collision events. Therefore, filtering out low energy particle track clusters will greatly reduce data collection rate at edge devices. A synthetic dataset is used to train and test the SNN.

The full synthetic dataset consists of 4 million charged particle interactions in a silicon pixel sensor. The training dataset is limited to the particle interactions in a 13×21 pixel sub-region of the silicon sensor, with binary classification labels indicating high or low energy. The SNN is realized on Caspian [216], a neuromorphic development platform, and achieved a signal classification accuracy of 91.89%, very close to a prototyped full-precision DNN accuracy of 94.8%. In addition to accuracy, the SNN achieves good performance using nearly half of the number of DNN parameters. The reduced size and improved power efficiency of the SNN model makes it a good candidate for deployment on edge devices which have limited memory and power constraints.

5 Summary

Experimental data generation at photon sources are rapidly increasing due to the advancements in light sources, detectors, and more efficient methods or modalities to collect data. As tabulated in Table 1, detectors can achieve frame-rates on the order of thousands and millions of frames per second in continuous and burst mode, respectively. Each frame can consist of thousands to millions of pixels, depending on the size of the pixel array format, with at least 10-bit data resolution. As a result, the detectors can achieve data rates over 1 GB/s in continuous mode, and orders of magnitude higher data rate in burst mode. Specifically, the state-of-the-art detectors with a 10-bit data format have demonstrated a data rate above 12.5 GB/s in continuous mode and 1.25 TB/s in burst mode. The high data rate is very costly in terms of data storage and transmission over long distances. These issues motivates the use of edge computing on detectors for real-time data processing and for reducing data transmission latency and storage volumes.

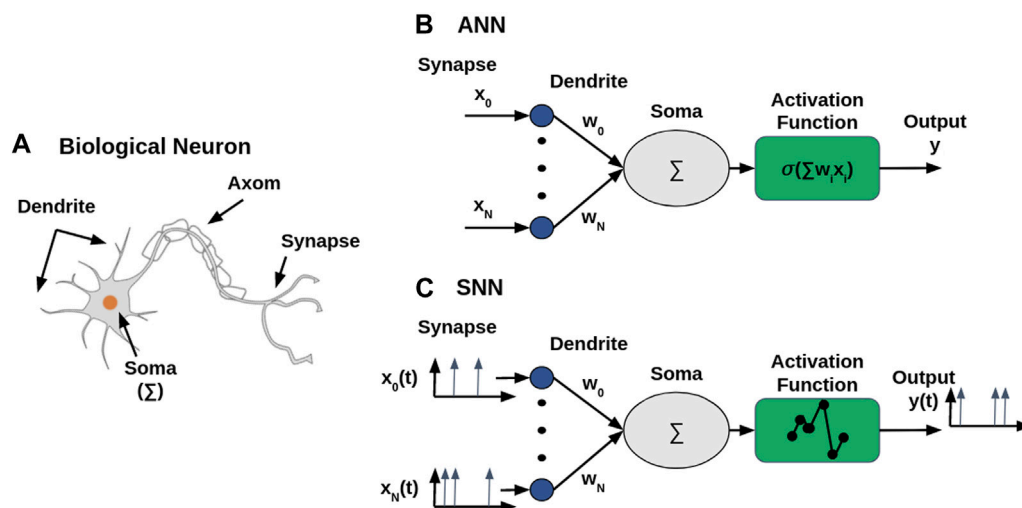


FIGURE 13
A comparison among (A) a biological neuron, (B) artificial network neuron [46], and (C) spiking network neuron [197].

Deep learning approaches have achieved significant progress in image processing tasks including but not limited to restoration, segmentation, compression, and 3D reconstruction. Their superior nonlinear approximation capabilities allow them to learn complex underlying structures and patterns in high dimensional data. The state-of-the-art methods for each image processing task achieve superior performance compared to conventional methods, while also overcoming the issues of conventional methods such as computational burdens associated with explicit programming for each data processing steps. Furthermore, once trained, deep learning methods can achieve very fast inference speeds for real-time computation.

While deep learning approaches are widely used for many applications, they require deep networks to achieve good performance, and thus require heavy computational power and high energy consumption. This is critical hurdle for edge computing devices which have design constraints such as latency and energy. To address this issue, hardware accelerators now exist that leverage the model and data parallelism characteristics of neural network algorithms to implement parallel computing paradigms. Electronic-based hardware accelerators such as CPUs, GPUs, FPGAs, and ASICs are popularly used platforms for deep learning. However, the electronic-based solutions are constantly reaching performance limitations in clock speed, energy consumption, and other physical constraints. This gives rise to research in analog neuromorphic computing paradigms such as ONNs and SNNs to achieve high-speed, energy-efficient, and high-parallel computing, with significant potential for radiation detection and applications in photon science. Nonetheless, note that the power constraint can be alleviated if the experimental space can accommodate the installation of larger processing centers such as workstations or servers, as well as the necessary data transmission networks. Furthermore, a larger processing center allows for the deployment of heavier DL models with improved accuracy for experiments that do not necessarily require real-time processing. To help alleviate data transmission, it is possible to offload simple computing and preprocessing steps to downstream edge devices.

Author contributions

SL: Supervision, Writing—original draft, Writing—review and editing. SN: Writing—original draft, Writing—review and editing. HZ: Writing—original draft, Writing—review and editing. TZ: Writing—original draft, Writing—review and editing. CM: Writing—review and editing. SC: Writing—review and editing. MC: Writing—original draft, Writing—review and editing. RC: Conceptualization, Writing—review and editing. ZW: Conceptualization, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. LANL work was performed under the auspices of the U.S. Department of Energy (DOE) by Triad National Security, LLC, operator of the Los Alamos National Laboratory under Contract No. 89233218CNA000001, including LANL Laboratory Directed Research and Development (LDRD) Program. This work is also supported in part by AFOSR MURI research center on Energy-efficient Optical Interconnects and Computing (Cont. No. FA9550-17-1-0071 managed by Dr. Gernot Pomrenke). Work performed at the Center for Nanoscale Materials and Advanced Photon Source, both U.S. Department of Energy Office of Science User Facilities, was supported by the U.S. DOE, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. MJC also acknowledges support from Argonne LDRD 2021-0090—AutoPycho: Autonomous, Sparse-sampled Ptychographic Imaging.

Acknowledgments

SL and ZW wish to thank Dr. Alice Bean from University of Kansas for reviewing the hardware section. SL and ZW wish to thank

Dr. Mathieu Benoit from Oak Ridge National Laboratory for the CARIBOu discussions. MJC also acknowledges support from Argonne LDRD 2021-0090—AutoPtycho: Autonomous, Sparse-sampled Ptychographic Imaging.

Conflict of interest

Author RC was employed by Omega Optics Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial

relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Wang Z, Leong AF, Dragone A, Gleason AE, Ballabriga R, Campbell C, et al. Ultrafast radiographic imaging and tracking: an overview of instruments, methods, data, and applications. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168690. doi:10.1016/j.nima.2023.168690
- Young L, Ueda K, Gühr M, Bucksbaum PH, Simon M, Mukamel S, et al. Roadmap of ultrafast x-ray atomic and molecular physics. *J Phys B: At Mol Opt Phys* (2018) 51: 032003. doi:10.1088/1361-6455/aa9735
- Nature Reviews Physics. *The next decade of XFELs* (2020). doi:10.1038/s42254-020-0206-4
- Russo P. *Handbook of X-ray imaging: physics and technology*. Boca Raton, Florida: CRC Press (2017).
- Weisenburger S, Sandoghdar V. Light microscopy: an ongoing contemporary revolution. *Contemp Phys* (2015) 56:123–43. doi:10.1080/00107514.2015.1026557
- Lu W, Friedrich B, Noll T, Zhou K, Hallmann J, Ansaldi G, et al. Development of a hard x-ray split-and-delay line and performance simulations for two-color pump-probe experiments at the european xfel. *Rev Scientific Instr* (2018) 89:063121. doi:10.1063/1.5027071
- Inoue I, Inubushi Y, Sato T, Tono K, Katayama T, Kameshima T, et al. Observation of femtosecond x-ray interactions with matter using an x-ray-x-ray pump-probe scheme. *Proc Natl Acad Sci* (2016) 113:1492–7. doi:10.1073/pnas.1516426113
- Eberhardt W. Synchrotron radiation: a continuing revolution in x-ray science—diffraction limited storage rings and beyond. *J Electron Spectrosc Relat Phenomena* (2015) 200:31–9. doi:10.1016/j.elspec.2015.06.009
- Dooling J, Borland M, Berg W, Calvey J, Decker G, Emery L, et al. Collimator irradiation studies in the argonne advanced photon source at energy densities expected in next-generation storage ring light sources. *Phys Rev Accel Beams* (2022) 25:043001. doi:10.1103/physrevaccbeams.25.043001
- Schroer CG, Agapov I, Brefeld W, Brinkmann R, Chae Y-C, Chao H-C, et al. PETRA IV: the ultralow-emittance source project at DESY. *J Synchrotron Radiat* (2018) 25:1277–90. doi:10.1107/S1600577518008858
- Huang N, Deng H, Liu B, Wang D, Zhao Z. Features and futures of x-ray free-electron lasers. *The Innovation* (2021) 2(2):100097. doi:10.1016/j.xinn.2021.100097
- LeCun Y, Bengio Y, Hinton G. Deep learning. *nature* (2015) 521:436–44. doi:10.1038/nature14539
- Goodfellow I, Bengio Y, Courville A. *Deep learning*. MIT Press (2016). <http://www.deeplearningbook.org> (Accessed October 10, 2023).
- Wu J-L, Kashinath K, Albert A, Chirila D, Prabhat Xiao H. Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems. *J Comput Phys* (2020) 406:109209. doi:10.1016/j.jcp.2019.109209
- Lin YT, Tian Y, Livescu D, Anghel M. Data-driven learning for the mori-zwanzig formalism: a generalization of the koopman learning framework. *SIAM J Appl Dynamical Syst* (2021) 20:2558–601. doi:10.1137/21m1401759
- Kochkov D, Smith JA, Alieva A, Wang Q, Brenner MP, Hoyer S. Machine learning-accelerated computational fluid dynamics. *Proc Natl Acad Sci* (2021) 118: e2101784118. doi:10.1073/pnas.2101784118
- O'Driscoll L, Nichols R, Knott PA. A hybrid machine learning algorithm for designing quantum experiments. *Quan Machine Intelligence* (2019) 1:5–15. doi:10.1007/s42484-019-00003-8
- Li J, Lopez S. *Machine learning accelerated photodynamics simulations* (2023).
- CERN. *CERN Data Centre passes the 200-petabyte milestone* (2017). <https://home.cern/news/news/computing/cern-data-centre-passes-200-petabyte-milestone> (Accessed December 15, 2023).
- Chen J, Ran X. Deep learning with edge computing: a review. *Proc IEEE* (2019) 107:1655–74. doi:10.1109/jproc.2019.2921977
- Kumar P. CMOS vs CCD: why CMOS sensors are ruling the world of embedded vision (2023). Available at: <https://www.edge-ai-vision.com/2023/04/cmos-vs-ccd-why-cmos-sensors-are-ruling-the-world-of-embedded-vision/> (Accessed December 9, 2023).
- Tabrizchi S, Nezhadi A, Angizi S, Roohi A. Appcip: energy-efficient approximate convolution-in-pixel scheme for neural network acceleration. *IEEE J Emerging Selected Top Circuits Syst* (2023) 13:225–36. doi:10.1109/jetcas.2023.3242167
- So HM, Bose L, Dudek P, Wetzstein G (2023). Pixelrnn: in-pixel recurrent neural networks for end-to-end-optimized perception with neural sensors. *arXiv preprint arXiv:2304.05440*
- Snoeys W. Monolithic cmos sensors for high energy physics—challenges and perspectives. In: *Nuclear instruments and methods in physics research section A: accelerators, spectrometers, detectors and associated equipment* (2023). 168678.
- van Schayck JP, Zhang Y, Knoops K, Peters PJ, Ravelli RB. Integration of an event-driven timepix3 hybrid pixel detector into a cryo-em workflow. *Microsc Microanalysis* (2023) 29:352–63. doi:10.1093/micmic/ozac009
- Tsigaridas S, Ponchut C. High-z pixel sensors for synchrotron applications. In: *Advanced X-ray detector technologies: design and applications* (2022). p. 87–107.
- Porter J, Looker Q, Claus L. Hybrid cmos detectors for high-speed x-ray imaging. *Rev Scientific Instr* (2023) 94:061101. doi:10.1063/5.0138264
- Carulla M, Doblas A, Flores D, Galloway Z, Hidalgo S, Kramberger G, et al. 50 μ m thin low gain avalanche detectors (lgad) for timing applications. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 924: 373–9. doi:10.1016/j.nima.2018.08.041
- Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of lgad sensors with a thin entrance window for soft x-ray detection. *J Instrumentation* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/c11011
- Giacomini G. Lgad-based silicon sensors for 4d detectors. *Sensors* (2023) 23:2132. doi:10.3390/s23042132
- Pietropaolo A, Angelone M, Bedogni R, Colonna N, Hurd A, Khaplanov A, et al. Neutron detection techniques from μ ev to gev. *Phys Rep* (2020) 875:1–65. doi:10.1016/j.physrep.2020.06.003
- Graafsma H. Hybrid pixel array detectors for photon science. In: *Semiconductor radiation detection systems*. Boca Raton, Florida: CRC Press (2018). p. 229–48.
- Niemann B, Rudolph D, Schmahl G. X-ray microscopy with synchrotron radiation. *Appl Opt* (1976) 15:1883–4. doi:10.1364/ao.15.001883
- Spanne P, Raven C, Snigireva I, Snigirev A. In-line holography and phase-contrast microtomography with high energy x-rays. *Phys Med Biol* (1999) 44:741–9. doi:10.1088/0031-9155/44/3/016
- Miao J, Ishikawa T, Robinson IK, Murnane MM. Beyond crystallography: diffractive imaging using coherent x-ray light sources. *Science* (2015) 348:530–5. doi:10.1126/science.aaa1394
- Pfeiffer F. X-ray ptychography. *Nat Photon* (2018) 12:9–17. doi:10.1038/s41566-017-0072-5
- Villanueva-Perez P, Bajt S, Chapman H. Dose efficient compton x-ray microscopy. *Optica* (2018) 5:450–7. doi:10.1364/optica.5.000450
- Gomez J, Patel S, Sarwar SS, Li Z, Capoccia R, Wang Z, et al. *Distributed on-sensor compute system for ar/vr devices: a semi-analytical simulation framework for power estimation* (2022). *arXiv preprint arXiv:2203.07474*.
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Graafsma H, et al. Megapixels@ megahertz—the agipd high-speed cameras for the european xfel. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 942:162324. doi:10.1016/j.nima.2019.06.065

40. Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the european xfel. *J synchrotron Radiat* (2019) 26:74–82. doi:10.1107/s1600577518016077
41. Lin S, Baldwin J, Blatnik M, Clayton S, Cude-Woods C, Currie S, et al. Demonstration of sub-micron ucn position resolution using room-temperature cmos sensor. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2023) 1057:168769. doi:10.1016/j.nima.2023.168769
42. Orponen P. Computational complexity of neural networks: a survey. *Nordic J Comp* (1994) 94–110.
43. Maji P, Mullins R. On the reduction of computational complexity of deep convolutional neural networks. *Entropy* (2018) 20:305. doi:10.3390/e20040305
44. Freire PJ, Srivallapanondh S, Napoli A, Prilepsky JE, Turitsyn SK. *Computational complexity evaluation of neural network applications in signal processing* (2022). arXiv: 2206.12191.
45. Golub GH, Van Loan CF. *Matrix computations*. 4 edn. Baltimore, MD: Johns Hopkins University Press (2013).
46. Meedeniya D. *Deep learning: a beginners' guide*. Boca Raton, Florida: CRC Press (2023).
47. Yip W. *Lifecycle of machine learning models* (2024). <https://www.oracle.com/a/ocom/docs/data-science-lifecycle-ebook.pdf> (Accessed December 9, 2023).
48. Patterson J, Gibson A. *Deep learning: a practitioner's approach*. Sebastopol, CA: O'Reilly Media, Inc. (2017).
49. Zhang A, Lipton ZC, Li M, Smola AJ. *Dive into deep learning*. Cambridge University Press (2023).
50. Masanet E, Shehabi A, Lei N, Smith S, Koomey J. Recalibrating global data center energy-use estimates. *Science* (2020) 367:984–6. doi:10.1126/science.aba3758
51. IEA. Data centres and data transmission networks (2023). <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks> (Accessed October 15, 2023).
52. RMI. Cryptocurrency's energy consumption problem (2023). <https://rmi.org/cryptocurrencies-energy-consumption-problem/> (Accessed December 5, 2023).
53. Kohli V, Chakravarty S, Chamola V, Sangwan KS, Zeadally S. An analysis of energy consumption and carbon footprints of cryptocurrencies and possible solutions. *Digital Commun Networks* (2023) 9:79–89. doi:10.1016/j.dcan.2022.06.017
54. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. Segment anything (2023). arXiv 2304 <https://segment-anything.com/> (Accessed December 5, 2023).
55. Morse J. NVIDIA's Project Clara is creating game-changing technology for medical imaging (2018). <https://community.radrounds.com/profiles/blogs/nvidia-s-project-clara-is-creating-game-changing-technology-for> (Accessed December 5, 2023).
56. Marko K. AI-enhanced instrumentation - the fusion of deep learning and medical sensors creates dramatic improvements (2018). <https://diginomica.com/ai-enhanced-instrumentation-fusion-deep-learning-medical-sensors-creates-dramatic-improvements> (Accessed December 5, 2023).
57. Milletari F, Navab N, Ahmadi S-A V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV); Stanford, CA (2016). 3DV. IEEE. 565–71.
58. Cern. *Environmental awareness: the challenges of CERN's IT infrastructure* (2022). <https://home.cern/news/news/cern/environmental-awareness-challenges-cerns-it-infrastructure> (Accessed December 18, 2023).
59. Adam-Bourdarios C, Cowan G, Germain C, Guyon I, Kégl B, Rousseau D. The Higgs boson machine learning challenge. In: NIPS 2014 Workshop on High-energy Physics and Machine Learning; Montreal, Canada. PMLR (2015). p. 19–55.
60. Azhari M, Abarda A, Ettaki B, Zerouaoui J, Dakkon M. Higgs boson discovery using machine learning methods with pyspark. *Proced Comp Sci* (2020) 170:1141–6. doi:10.1016/j.procs.2020.03.053
61. Kalinin SV, Ziatdinov M, Hinkle J, Jesse S, Ghosh A, Kelley KP, et al. Automated and autonomous experiments in electron and scanning probe microscopy. *ACS nano* (2021) 15:12604–27. doi:10.1021/acsnano.1c02104
62. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: a survey. *ACM Comput Surv (Csur)* (2022) 54:1–41. doi:10.1145/3505244
63. Zhang N, Ding S, Zhang J, Xue Y. An overview on restricted Boltzmann machines. *Neurocomputing* (2018) 275:1186–99. doi:10.1016/j.neucom.2017.09.065
64. Cao J, Lin Z. Extreme learning machines on high dimensional and large data applications: a survey. *Math Probl Eng* (2015) 2015(3):1–13. doi:10.1155/2015/103796
65. Zeiler MD. *Hierarchical convolutional deep learning in computer vision*. Ph.D. thesis. Ann Arbor, Michigan: New York University Proquest (2013).
66. Boureau YL, Ponce J, LeCun Y. A theoretical analysis of feature pooling in visual recognition. In: Proceedings of the 27th international conference on machine learning (ICML-10) (2010). p. 111–8.
67. Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. In: International conference on artificial neural networks. Springer (2010). 92–101.
68. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* (2012) 25. doi:10.1145/3065386
69. Simonyan K, Zisserman A. *Very deep convolutional networks for large-scale image recognition* (2014). arXiv preprint arXiv:1409.1556.
70. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2015). p. 1–9.
71. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2016). p. 770–8.
72. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: lstm cells and network architectures. *Neural Comput* (2019) 31:1235–70. doi:10.1162/neco_a_01199
73. Bahdanau D, Cho K, Bengio Y. *Neural machine translation by jointly learning to align and translate* (2014). arXiv preprint arXiv:1409.0473.
74. Herdade S, Kappeler A, Boakye K, Soares J. Image captioning: transforming objects into words. *Adv Neural Inf Process Syst* (2019) 32.
75. Chiu C-C, Sainath TN, Wu Y, Prabhavalkar R, Nguyen P, Chen Z, et al. State-of-the-art speech recognition with sequence-to-sequence models. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE (2018). p. 4774–8.
76. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Proceedings, Part III 18 Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference; October 5–9, 2015; Munich, Germany. Springer (2015). p. 234–41.
77. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. *Adv Neural Inf Process Syst* (2014) 27. doi:10.1145/3422622
78. Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: an overview. *IEEE Signal Processing Magazine* (2018) 35:53–65. doi:10.1109/msp.2017.2765202
79. Falato M, Wolfe B, Nguyen N, Zhang X, Wang Z. *Contour extraction of inertial confinement fusion images by data augmentation* (2022). arXiv preprint arXiv: 2211.04597.
80. Wali A, Naseer A, Tamoor M, Gilani S. Recent progress in digital image restoration techniques: a review. *Digital Signal Process.* (2023) 141:104187. doi:10.1016/j.dsp.2023.104187
81. Fan L, Zhang F, Fan H, Zhang C. Brief review of image denoising techniques. *Vis Comput Industry, Biomed Art* (2019) 2:7–12. doi:10.1186/s42492-019-0016-7
82. Zhang K, Ren W, Luo W, Lai W-S, Stenger B, Yang M-H, et al. Deep image deblurring: a survey. *Int J Comp Vis* (2022) 130:2103–30. doi:10.1007/s11263-022-01633-5
83. Yang W, Zhang X, Tian Y, Wang W, Xue J-H, Liao Q. Deep learning for single image super-resolution: a brief review. *IEEE Trans Multimedia* (2019) 21:3106–21. doi:10.1109/tmm.2019.2919431
84. Tian C, Fei L, Zheng W, Xu Y, Zuo W, Lin C-W. Deep learning on image denoising: an overview. *Neural Networks* (2020) 131:251–75. doi:10.1016/j.neunet.2020.07.025
85. Lucas A, Iliadis M, Molina R, Katsaggelos AK. Using deep neural networks for inverse problems in imaging: beyond analytical methods. *IEEE Signal Process. Mag* (2018) 35:20–36. doi:10.1109/msp.2017.2760358
86. Liu Z, Bicer T, Kettimuthu R, Gursoy D, De Carlo F, Foster I. Tomogan: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion. *JOSA A* (2020) 37:422–34. doi:10.1364/josaa.375595
87. Duan X, Ding XF, Li N, Wu F-X, Chen X, Zhu N. Sparse2noise: low-dose synchrotron x-ray tomography without high-quality reference data. *Comput Biol Med* (2023) 165:107473. doi:10.1016/j.combiomed.2023.107473
88. Hendriksen AA, Bührer M, Leone L, Merlini M, Vigano N, Pelt DM, et al. Deep denoising for multi-dimensional synchrotron x-ray tomography without high-quality reference data. *Scientific Rep* (2021) 11:11895. doi:10.1038/s41598-021-91084-8
89. Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala M, et al. *Noise2noise: learning image restoration without clean data* (2018). arXiv preprint arXiv:1803.04189.
90. Morris CL, King N, Kwiatkowski K, Mariam F, Merrill F, Saunders A. Charged particle radiography. *Rep Prog Phys* (2013) 76:046301. doi:10.1088/0034-4885/76/4/046301
91. Biyouki SA, Hwangbo H. *A comprehensive survey on deep neural image deblurring* (2023). arXiv preprint arXiv:2310.04719.
92. Yang J, Zhao C, Qiao S, Zhang T, Yao X. Deep learning methods for neutron image restoration. *Ann Nucl Energy* (2023) 188:109820. doi:10.1016/j.anucene.2023.109820
93. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: residual learning of deep cnn for image denoising. *IEEE Trans Image Process* (2017) 26:3142–55. doi:10.1109/tip.2017.2662206
94. Gurrola-Ramos J, Dalmau O, Alarcón TE. A residual dense u-net neural network for image denoising. *IEEE Access* (2021) 9:31742–54. doi:10.1109/access.2021.3061062

95. Miao C, Xie L, Wan F, Su C, Liu H, Jiao J, et al. Sixray: a large-scale security inspection x-ray benchmark for prohibited item discovery in overlapping images. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (2019). p. 2119–28.
96. Li Y, Sixou B, Peyrin F. A review of the deep learning methods for medical images super resolution problems. *Irbm* (2021) 42:120–33. doi:10.1016/j.irbm.2020.08.004
97. Jiang J, Wang C, Liu X, Ma J. Deep learning-based face super-resolution: a survey. *ACM Comput Surv (Csur)* (2021) 55:1–36. doi:10.1145/3485132
98. Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Trans pattern Anal machine intelligence* (2015) 38:295–307. doi:10.1109/tpami.2015.2439281
99. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2017).
100. Wang Z, Chen J, Hoi SC. Deep learning for image super-resolution: a survey. *IEEE Trans pattern Anal machine intelligence* (2020) 43:3365–87. doi:10.1109/tpami.2020.2982166
101. Park J, Hwang D, Kim KY, Kang SK, Kim YK, Lee JS. Computed tomography super-resolution using deep convolutional neural network. *Phys Med Biol* (2018) 63:145011. doi:10.1088/1361-6560/aacdd4
102. Eckert B, Aschauer S, Holl P, Majewski P, Zabel T, Strüder L. Electron imaging reconstruction for pixelated semiconductor tracking detectors in transmission electron microscopes using the approach of convolutional neural networks. *IEEE Trans Nucl Sci* (2022) 69:1014–21. doi:10.1109/tns.2022.3169281
103. Ryll H, Simson M, Hartmann R, Holl P, Huth M, Ihle S, et al. A pncccd-based, fast direct single electron imaging camera for tem and stem. *J Instrumentation* (2016) 11:P04006. doi:10.1088/1748-0221/11/04/p04006
104. Szeliski R. *Computer vision: algorithms and applications*. Springer Nature (2022).
105. Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK. Medical image segmentation using deep learning: a survey. *IET Image Process* (2022) 16:1243–67. doi:10.1049/ipr2.12419
106. Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans pattern Anal machine intelligence* (2017) 39:2481–95. doi:10.1109/tpami.2016.2644615
107. Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: a survey. *IEEE Trans pattern Anal machine intelligence* (2021) 44:3523–42. doi:10.1109/TPAMI.2021.3059968
108. Lewis WE, Knapp PF, Harding EC, Beckwith K. Statistical characterization of experimental magnetized liner inertial fusion stagnation images using deep-learning-based fuel-background segmentation. *J Plasma Phys* (2022) 88:895880501. doi:10.1017/s0022377822000800
109. Rawat W, Wang Z. Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput* (2017) 29:2352–449. doi:10.1162/neco_a_00990
110. Pak M, Kim S A review of deep learning in image recognition. In: 2017 4th international conference on computer applications and information processing technology (CAIPT) (2017). IEEE. 1–3.
111. Zhao Z-Q, Zheng P, Xu S-t, Wu X. Object detection with deep learning: a review. *IEEE Trans Neural networks Learn Syst* (2019) 30:3212–32. doi:10.1109/tnnls.2018.2876865
112. Sharma VK, Mir RN. A comprehensive and systematic look up into deep learning based object detection techniques: a review. *Comp Sci Rev* (2020) 38:100301. doi:10.1016/j.cosrev.2020.100301
113. Jiang X, Hadid A, Pang Y, Granger E, Feng X. *Deep learning in object detection and recognition* (2019).
114. Girshick R. Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision (2015). p. 1440–8.
115. Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. *Adv Neural Inf Process Syst* (2015) 28.
116. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2016). p. 779–88.
117. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. Ssd: single shot multibox detector. In: Proceedings, Part I Computer Vision–ECCV 2016: 14th European Conference, October 11–14, 2016; Amsterdam, The Netherlands, 14. Springer (2016). p. 21–37.
118. Mishra D, Singh SK, Singh RK. Deep architectures for image compression: a critical review. *Signal Process*. (2022) 191:108346. doi:10.1016/j.sigpro.2021.108346
119. Wallace GK. The jpeg still picture compression standard. *IEEE Trans consumer Electron* (1992) 38. xviii–xxxiv. doi:10.1109/30.125072
120. Donoho DL. Compressed sensing. *IEEE Trans Inf Theor* (2006) 52:1289–306. doi:10.1109/tit.2006.871582
121. Shi W, Jiang F, Liu S, Zhao D. Image compressed sensing using convolutional neural network. *IEEE Trans Image Process* (2019) 29:375–88. doi:10.1109/tip.2019.2928136
122. Machidon AL, Pejović V. Deep learning for compressive sensing: a ubiquitous systems perspective. *Artif Intelligence Rev* (2023) 56:3619–58. doi:10.1007/s10462-022-10259-5
123. Qiao M, Meng Z, Ma J, Yuan X. Deep learning for video compressive sensing. *Appl Photon* (2020) 5. doi:10.1063/1.5140721
124. Cherukara MJ, Zhou T, Nashed Y, Enfedaque P, Hexemer A, Harder RJ, et al. Ai-enabled high-resolution scanning coherent diffraction imaging. *Appl Phys Lett* (2020) 117. doi:10.1063/5.0013065
125. Babu AV, Zhou T, Kandel S, Bicer T, Liu Z, Judge W, et al. *Deep learning at the edge enables real-time streaming ptychographic imaging* (2022). arXiv preprint arXiv:2209.09408. doi:10.1038/s41467-023-41496-z
126. Vasudevan RK, Kelley KP, Hinkle J, Funakubo H, Jesse S, Kalinin SV, et al. Autonomous experiments in scanning probe microscopy and spectroscopy: choosing where to explore polarization dynamics in ferroelectrics. *ACS nano* (2021) 15:11253–62. doi:10.1021/acsnano.0c10239
127. Noack MM, Yager KG, Fukuto M, Doerk GS, Li R, Sethian JA. A kriging-based approach to autonomous experimentation with applications to x-ray scattering. *Scientific Rep* (2019) 9:11809. doi:10.1038/s41598-019-48114-3
128. Venkatakrishnan S, Fancher CM, Ziatdinov M, Vasudevan R, Saleeby K, Haley J, et al. Adaptive sampling for accelerating neutron diffraction-based strain mapping. *Machine Learn Sci Tech* (2023) 4:025001. doi:10.1088/2632-2153/ac512
129. Liu H, Ong Y-S, Shen X, Cai J. When Gaussian process meets big data: a review of scalable gps. *IEEE Trans Neural networks Learn Syst* (2020) 31:4405–23. doi:10.1109/tnnls.2019.2957109
130. Scholz M, Müller J, Pekin TC, Van den Broek W, Madsen J, Susi T, et al. Deep reinforcement learning for data-driven adaptive scanning in ptychography. *Scientific Rep* (2023) 13:8732. doi:10.1038/s41598-023-35740-1
131. Kandel S, Zhou T, Babu AV, Di Z, Li X, Ma X, et al. Demonstration of an ai-driven workflow for autonomous high-resolution scanning microscopy. *Nat Commun* (2023) 14:5501. doi:10.1038/s41467-023-40339-1
132. Han X-F, Laga H, Bannamoun M. Image-based 3d object reconstruction: state-of-the-art and trends in the deep learning era. *IEEE Trans pattern Anal machine intelligence* (2019) 43:1578–604. doi:10.1109/tpami.2019.2954885
133. Fu K, Peng J, He Q, Zhang H. Single image 3d object reconstruction based on deep learning: a review. *Multimedia Tools Appl* (2021) 80:463–98. doi:10.1007/s11042-020-09722-8
134. Guan Z, Tsai EH, Huang X, Yager KG, Qin H. *Ptychonet: fast and high quality phase retrieval for ptychography*. Upton, NY (United States): Brookhaven National Lab (2019). Tech. rep.
135. Cherukara MJ, Nashed YS, Harder RJ. Real-time coherent diffraction inversion using deep generative networks. *Scientific Rep* (2018) 8:16520. doi:10.1038/s41598-018-34525-1
136. Yao Y, Chan H, Sankaranarayanan S, Balaprakash P, Harder RJ, Cherukara MJ. Autophasenn: unsupervised physics-aware deep learning of 3d nanoscale bragg coherent diffraction imaging. *npj Comput Mater* (2022) 8:124. doi:10.1038/s41524-022-00803-w
137. Scheinker A, Pokharel R. Adaptive 3d convolutional neural network-based reconstruction method for 3d coherent diffraction imaging. *J Appl Phys* (2020) 128. doi:10.1063/5.0014725
138. Gholami A, Azad A, Jin P, Keutzer K, Buluc A. Integrated model, batch, and domain parallelism in training neural networks. In: Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures (2018). p. 77–86.
139. Sze V, Chen Y-H, Yang T-J, Emer JS. Efficient processing of deep neural networks: a tutorial and survey. *Proc IEEE* (2017) 105:2295–329. doi:10.1109/jproc.2017.2761740
140. Intel. *Compare benefits of CPUs, GPUs, and FPGAs for different oneAPI compute workloads* (2022). <https://www.intel.com/content/www/us/en/developer/articles/technical/comparing-cpus-gpus-and-fpgas-for-oneapi.html> (Accessed September 25, 2023).
141. Mittal S, Vetter JS. A survey of cpu-gpu heterogeneous computing techniques. *ACM Comput Surv (Csur)* (2015) 47:1–35. doi:10.1145/2788396
142. Mittal S. A survey on optimized implementation of deep learning models on the nvidia jetson platform. *J Syst Architecture* (2019) 97:428–42. doi:10.1016/j.sysarc.2019.01.011
143. Abeykoon V, Liu Z, Kettimuthu R, Fox G, Foster I. Scientific image restoration anywhere. In: 2019 IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP). IEEE (2019). p. 8–13.
144. An L, Peng K, Yang X, Huang P, Luo Y, Feng P, et al. E-tbnet: light deep neural network for automatic detection of tuberculosis with x-ray dr imaging. *Sensors* (2022) 22:821. doi:10.3390/s22030821

145. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2018). p. 4510–20.
146. Zhang X, Zhou X, Lin M, Sun J. ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2018). p. 6848–56.
147. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. *SqueezeNet: alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size* (2016). *arXiv preprint arXiv:1602.07360*.
148. Li J, Yan G, Lu W, Jiang S, Gong S, Wu J, et al. Smartshuttle: optimizing off-chip memory accesses for deep learning accelerators. In: 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE (2018). 343–8.
149. Chen Y-H, Emer J, Sze V. Eyeriss: a spatial architecture for energy-efficient dataflow for convolutional neural networks. *ACM SIGARCH Comput architecture News* (2016) 44:367–79. doi:10.1145/3007787.3001177
150. Capra M, Bussolino B, Marchisio A, Masera G, Martina M, Shafique M. Hardware and software optimizations for accelerating deep neural networks: survey of current trends, challenges, and the road ahead. *IEEE Access* (2020) 8:225134–80. doi:10.1109/access.2020.3039858
151. Dhillewarao P, Boppu S, Manikandan MS, Cenkeramaddi LR. *Efficient hardware architectures for accelerating deep neural networks: survey*. IEEE Access (2022).
152. Suda N, Chandra V, Dasika G, Mohanty A, Ma Y, Vruthula S, et al. Throughput-optimized opencl-based fpga accelerator for large-scale convolutional neural networks. In: Proceedings of the 2016 ACM/SIGDA international symposium on field-programmable gate arrays (2016). p. 16–25.
153. Qiu J, Wang J, Yao S, Guo K, Li B, Zhou E, et al. Going deeper with embedded fpga platform for convolutional neural network. In Proceedings of the 2016 ACM/SIGDA international symposium on field-programmable gate arrays (2016).
154. Shawahna A, Sait SM, El-Maleh A. Fpga-based accelerators of deep learning networks for learning and classification: a review. *IEEE Access* (2018) 7:7823–59. doi:10.1109/access.2018.2890150
155. Wu R, Guo X, Du J, Li J. Accelerating neural network inference on fpga-based platforms—a survey. *Electronics* (2021) 10:1025. doi:10.3390/electronics10091025
156. Xia M, Huang Z, Tian L, Wang H, Chang V, Zhu Y, et al. Sparknoc: an energy-efficiency fpga-based accelerator using optimized lightweight cnn for edge computing. *J Syst Architecture* (2021) 115:101991. doi:10.1016/j.sysarc.2021.101991
157. Liu X, Yang J, Zou C, Chen Q, Yan X, Chen Y, et al. Collaborative edge computing with fpga-based cnn accelerators for energy-efficient and time-aware face tracking system. *IEEE Trans Comput Soc Syst* (2021) 9:252–66. doi:10.1109/tcss.2021.3059318
158. Cass S. Taking ai to the edge: Google's tpu now comes in a maker-friendly package. *IEEE Spectr* (2019) 56:16–7. doi:10.1109/mspec.2019.8701189
159. Sun Y, Kist AM. *Deep learning on edge tpus* (2021). *arXiv preprint arXiv:2108.13732*.
160. Liu H, Benoit M, Chen H, Chen K, Di Bello F, Iacobucci G, et al. Development of a modular test system for the silicon sensor r&d of the atlas upgrade. *J Instrumentation* (2017) 12:P01008. doi:10.1088/1748-0221/12/01/p01008
161. Stremper S, Zhou T, Yoshii K, Hammer M, Babu A, Bycul D, et al. A lightweight, user-configurable detector asic digital architecture with on-chip data compression for mhz x-ray coherent diffraction imaging. *J Instrumentation* (2022) 17:P10042. doi:10.1088/1748-0221/17/10/p10042
162. Waldrop MM. More than moore. *Nature* (2016) 530:144–8. doi:10.1038/530144a
163. Feng X, Jiang Y, Yang X, Du M, Li X. Computer vision algorithms and hardware implementations: a survey. *Integration* (2019) 69:309–20. doi:10.1016/j.vlsi.2019.07.005
164. Le Kernec J, Fioranelli F, Ding C, Zhao H, Sun L, Hong H, et al. Radar signal processing for sensing in assisted living: the challenges associated with real-time implementation of emerging algorithms. *IEEE Signal Process. Mag* (2019) 36:29–41. doi:10.1109/msp.2019.2903715
165. Curtis H. *FPGA programming and its cost comparison* (2024). <https://hillmancurtis.com/fpga-programming-and-its-cost-comparison/> (Accessed December 6, 2023).
166. Singh R. *FPGA vs ASIC: differences between them and which one to use* (2018).
167. Sigenics. *Asics* (2024). <https://sigenics.com/page/asics-c> (Accessed December 6, 2023).
168. Heuring VP, Murdocca MJ. *Principles of computer architecture* (2021).
169. Ganguly A, Muralidhar R, Singh V. Towards energy efficient non-von neumann architectures for deep learning. In: 20th international symposium on quality electronic design (ISQED). IEEE (2019). p. 335–42.
170. Sui X, Wu Q, Liu J, Chen Q, Gu G. A review of optical neural networks. *IEEE Access* (2020) 8:70773–83. doi:10.1109/access.2020.2987333
171. Shen Y, Harris NC, Skirlo S, Prabhu M, Baehr-Jones T, Hochberg M, et al. Deep learning with coherent nanophotonic circuits. *Nat Photon* (2017) 11:441–6. doi:10.1038/nphoton.2017.93
172. Shastri BJ, Tait AN, Ferreira de Lima T, Pernice WH, Bhaskaran H, Wright CD, et al. Photonics for artificial intelligence and neuromorphic computing. *Nat Photon* (2021) 15:102–14. doi:10.1038/s41566-020-00754-y
173. Feng C, Ning S, Gu J, Zhu H, Pan DZ, Chen RT. Integrated photonics for computing and artificial intelligence. In: 2023 IEEE Photonics Society Summer Topicals Meeting Series (SUM). IEEE (2023). p. 1–2.
174. Gu J, Feng C, Zhu H, Chen RT, Pan DZ. Light in ai: toward efficient neurocomputing with optical neural networks—a tutorial. *IEEE Trans Circuits Syst Express Briefs* (2022) 69:2581–5. doi:10.1109/tcsii.2022.3171170
175. Wu J, Lin X, Guo Y, Liu J, Fang L, Jiao S, et al. Analog optical computing for artificial intelligence. *Engineering* (2022) 10:133–45. doi:10.1016/j.eng.2021.06.021
176. Tait AN, De Lima TF, Zhou E, Wu AX, Nahmias MA, Shastri BJ, et al. Neuromorphic photonic networks using silicon photonic weight banks. *Scientific Rep* (2017) 7:7430. doi:10.1038/s41598-017-07754-z
177. Feldmann J, Youngblood N, Karpov M, Gehring H, Li X, Stappers M, et al. Parallel convolutional processing using an integrated photonic tensor core. *Nature* (2021) 589:52–8. doi:10.1038/s41586-020-03070-1
178. Ríos C, Youngblood N, Cheng Z, Le Gallo M, Pernice WH, Wright CD, et al. In-memory computing on a photonic platform. *Sci Adv* (2019) 5. 5759. eaa5759. doi:10.1126/sciadv.aau5759
179. Zhu H, Gu J, Feng C, Liu M, Jiang Z, Chen RT, et al. Elight: toward efficient and aging-resilient photonic in-memory neurocomputing. *IEEE Trans Computer-Aided Des Integrated Circuits Syst* (2022) 42:820–33. doi:10.1109/tcad.2022.3180969
180. Lin X, Rivenson Y, Yardimci NT, Veli M, Luo Y, Jarrahi M, et al. All-optical machine learning using diffractive deep neural networks. *Science* (2018) 361:1004–8. doi:10.1126/science.aat8084
181. Yan T, Wu J, Zhou T, Xie H, Xu F, Fan J, et al. Fourier-space diffractive deep neural network. *Phys Rev Lett* (2019) 123:023901. doi:10.1103/physrevlett.123.023901
182. Ashtiani F, Geers AJ, Aflatouni F. An on-chip photonic deep neural network for image classification. *Nature* (2022) 606:501–6. doi:10.1038/s41586-022-04714-0
183. Huang C, Sorger VJ, Miscuglio M, Al-Qadasi M, Mukherjee A, Lampe L, et al. Prospects and applications of photonic neural networks. *Adv Phys X* (2022) 7:1981155. doi:10.1080/23746149.2021.1981155
184. Feng C, Gu J, Zhu H, Ying Z, Zhao Z, Pan DZ, et al. A compact butterfly-style silicon photonic–electronic neural chip for hardware-efficient deep learning. *ACS Photon* (2022) 9:3906–16. doi:10.1021/acsp Photonics.2c01188
185. Feng C, Tang R, Gu J, Zhu H, Pan DZ, Chen RT. Optically-interconnected, hardware-efficient, electronic-photonic neural network using compact multi-operand photonic devices. In: Proceedings Volume PC12427, Optical Interconnects XXIII, San Francisco, California. SPIE (2023). p. 1242702.
186. Gu J, Zhao Z, Feng C, Zhu H, Chen RT, Pan DZ. Roq: a noise-aware quantization scheme towards robust optical neural networks with low-bit controls. In: 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE (2020). p. 1586–9.
187. Zhu H, Gu J, Wang H, Tang R, Zhang Z, Feng C, et al. Lightning-transformer: A dynamically-operated optically-interconnected photonic transformer accelerator. In *IEEE International Symposium on High-Performance Computer Architecture (HPCA)* (2024). Available at: <https://arxiv.org/abs/2305.19533>.
188. Wang T, Sohoni MM, Wright LG, Stein MM, Ma S-Y, Onodera T, et al. Image sensing with multilayer nonlinear optical neural networks. *Nat Photon* (2023) 17: 408–15. doi:10.1038/s41566-023-01170-8
189. Zhou T, Wu W, Zhang J, Yu S, Fang L. Ultrafast dynamic machine vision with spatiotemporal photonic computing. *Sci Adv* (2023) 9:eadg4391. doi:10.1126/sciadv. adg4391
190. Yamaguchi T, Arai K, Niyama T, Uchida A, Sunada S. Time-domain photonic image processor based on speckle projection and reservoir computing. *Commun Phys* (2023) 6:250. doi:10.1038/s42005-023-01368-w
191. Huang L, Mukherjee S, Tanguy Q, Frösch J, Majumdar A. Photonic advantage of optical encoders. In: 2023 Conference on Lasers and Electro-Optics (CLEO). IEEE (2023). p. 1–2.
192. Zhu H, Zhu K, Gu J, Jin H, Chen RT, Incorvia JA, et al. Fuse and mix: macam-enabled analog activation for energy-efficient neural acceleration. In: Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design (2022). p. 1–9.
193. Gu J, Feng C, Zhu H, Zhao Z, Ying Z, Liu M, et al. Squeezelight: a multi-operand ring-based optical neural network with cross-layer scalability. *IEEE Trans Computer-Aided Des Integrated Circuits Syst* (2022) 42:807–19. doi:10.1109/tcad.2022.3189567
194. Feng C, Gu J, Zhu H, Tang R, Ning S, Hlaing M, et al. *Integrated multi-operand optical neurons for scalable and hardware-efficient deep learning* (2023). *arXiv preprint arXiv:2305.19592*.
195. Zhu H, Zou J, Zhang H, Shi Y, Luo S, Wang N, et al. Space-efficient optical computing with an integrated chip diffractive neural network. *Nat Commun* (2022) 13: 1044. doi:10.1038/s41467-022-28702-0
196. Nunes JD, Carvalho M, Carneiro D, Cardoso JS. Spiking neural networks: a survey. *IEEE Access* (2022) 10:60738–64. doi:10.1109/access.2022.3179968

197. Roy K, Jaiswal A, Panda P. Towards spike-based machine intelligence with neuromorphic computing. *Nature* (2019) 575:607–17. doi:10.1038/s41586-019-1677-2
198. Tavanaei A, Ghodrati M, Kheradpisheh SR, Masquelier T, Maida A. Deep learning in spiking neural networks. *Neural networks* (2019) 111:47–63. doi:10.1016/j.neunet.2018.12.002
199. Schuman CD, Kulkarni SR, Parsa M, Mitchell JP, Date P, Kay B. Opportunities for neuromorphic computing algorithms and applications. *Nat Comput Sci* (2022) 2: 10–9. doi:10.1038/s43588-021-00184-y
200. Pfeiffer M, Pfeil T. Deep learning with spiking neurons: opportunities and challenges. *Front Neurosci* (2018) 12:774. doi:10.3389/fnins.2018.00774
201. Davies M, Srinivasa N, Lin T-H, China Y, Cao Y, Choday SH, et al. Loihi: a neuromorphic manycore processor with on-chip learning. *Ieee Micro* (2018) 38:82–99. doi:10.1109/mm.2018.112130359
202. Davies M, Wild A, Orchard G, Sandamirskaya Y, Guerra GAF, Joshi P, et al. Advancing neuromorphic computing with loihi: a survey of results and outlook. *Proc IEEE* (2021) 109:911–34. doi:10.1109/jproc.2021.3067593
203. Getty N, Brettin T, Jin D, Stevens R, Xia F. Deep medical image analysis with representation learning and neuromorphic computing. *Interf Focus* (2021) 11: 20190122. doi:10.1098/rsfs.2019.0122
204. Merolla PA, Arthur JV, Alvarez-Icaza R, Cassidy AS, Sawada J, Akopyan F, et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* (2014) 345:668–73. doi:10.1126/science.1254642
205. Shukla R, Lipasti M, Van Essen B, Moody A, Maruyama N. Remodel: rethinking deep cnn models to detect and count on a neurosynaptic system. *Front Neurosci* (2019) 13:4. doi:10.3389/fnins.2019.00004
206. Benjamin BV, Gao P, McQuinn E, Choudhary S, Chandrasekaran AR, Bussat J-M, et al. Neurogrid: a mixed-analog-digital multichip system for large-scale neural simulations. *Proc IEEE* (2014) 102:699–716. doi:10.1109/jproc.2014.2313565
207. Schemmel J, Brüderle D, Grübl A, Hock M, Meier K, Millner S. A wafer-scale neuromorphic hardware system for large-scale neural modeling. In: 2010 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE (2010). 1947–50.
208. Furber SB, Galluppi F, Temple S, Plana LA. The spinnaker project. *Proc IEEE* (2014) 102:652–65. doi:10.1109/jproc.2014.2304638
209. Bouvier M, Valentian A, Mesquida T, Rummens F, Reyboz M, Vianello E, et al. Spiking neural networks hardware implementations and challenges: a survey. *ACM J Emerging Tech Comput Syst (Jetc)* (2019) 15:1–35. doi:10.1145/3304103
210. Basu A, Deng L, Frenkel C, Zhang X. Spiking neural network integrated circuits: a review of trends and future directions. In: 2022 IEEE Custom Integrated Circuits Conference (CICC). IEEE (2022). 1–8.
211. Schuman CD, Potok TE, Young S, Patton R, Perdue G, Chakma G, et al. Neuromorphic computing for temporal scientific data classification. In: Proceedings of the Neuromorphic Computing Symposium (2017). p. 1–6.
212. Aliaga L, Bagby L, Baldin B, Baumbaugh A, Bodek A, Bradford R, et al. Design, calibration, and performance of the minerva detector. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2014) 743:130–59. doi:10.1016/j.nima.2013.12.053
213. Schuman CD, Birdwell JD, Dean ME. Spatiotemporal classification using neuroscience-inspired dynamic architectures. *Proced Comp Sci* (2014) 41:89–97. doi:10.1016/j.procs.2014.11.089
214. Cady NC. *Development of a memristive dynamic adaptive neural network array (mrdanna)* (2019). of Nanoscale Science, S. P. C., and Engineering.
215. Kulkarni R, Young A, Date P, Rao Miniskar N, Vetter J, Fahim F, et al. On-sensor data filtering using neuromorphic computing for high energy physics experiments. In: Proceedings of the 2023 International Conference on Neuromorphic Systems (2023). p. 1–8.
216. Mitchell JP, Schuman CD, Patton RM, Potok TE. Caspian: a neuromorphic development platform. In: Proceedings of the 2020 Annual Neuro-Inspired Computational Elements Workshop (2020). p. 1–6.
217. Rupp K. *Microprocessor trend data* (2022). <https://github.com/karlrupp/microprocessor-trend-data/tree/master> (Accessed September 25, 2023).
218. Turchetta R (2017). Towards gfps cmos image sensors. In In Workshop on Computational Image Sensors and Smart Cameras; May, 2017; Barcelona, Spain, WASC. vol. 2017.
219. Philipp HT, Hromalik M, Tate M, Koerner L, Gruner SM. Pixel array detector for x-ray free electron laser experiments. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2011) 649:67–9. doi:10.1016/j.nima.2010.11.189
220. Porro M, Andricek L, Bombelli L, De Vita G, Fiorini C, Fischer P, et al. Expected performance of the depfet sensor with signal compression: a large format x-ray imager with mega-frame readout capability for the european xfel. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2010) 624:509–19. doi:10.1016/j.nima.2010.02.254
221. Carini G, Alonso-Mori R, Blaj G, Caragiulo P, Chollet M, Damiani D, et al. epix100 camera: use and applications at lcls. In: Proceedings of the 12th International Conference on Synchrotron Radiation Instrumentation – SRI2015; New York, NY, 1741. AIP Publishing (2016).
222. Veale M, Seller P, Wilson M, Liotti E. Hexitec: a high-energy x-ray spectroscopic imaging detector for synchrotron applications. *Synchrotron Radiat News* (2018) 31: 28–32. doi:10.1080/08940886.2018.1528431
223. Claus L, England T, Fang L, Robertson G, Sanchez M, Trotter D, et al. Design and characterization of an improved, 2 ns, multi-frame imager for the ultra-fast x-ray imager (uxi) program at sandia national laboratories. In: Proceedings Volume 10390, Target Diagnostics Physics and Engineering for Inertial Confinement Fusion VI; San Diego, California, 10390. SPIE (2017). 16–26.
224. Leonarski F, Nan J, Matej Z, Bertrand Q, Furrer A, Gorgisyan I, et al. Kilohertz serial crystallography with the jungfrau detector at a fourth-generation synchrotron source. *IUCrJ* (2023) 10:729–37. doi:10.1107/s20522523008618
225. Gadkari D, Shanks K, Hu H, Philipp H, Tate M, Thom-Levy J, et al. Characterization of 128 × 128 MM-PAD-2.1 ASIC: a fast framing hard x-ray detector with high dynamic range. *J Instrumentation* (2022) 17:P03003. doi:10.1088/1748-0221/17/03/p03003
226. Tochigi Y, Hanzawa K, Kato Y, Kuroda R, Mutoh H, Hirose R, et al. A global-shutter cmos image sensor with readout speed of 1-tpixel/s burst and 780-mpixel/s continuous. *IEEE J Solid-State Circuits* (2012) 48:329–38. doi:10.1109/jssc.2012.2219685
227. Lewis A, Baker S, Corredor A, Fegenbush L, Fitzpatrick Z, Jones M, et al. New design yields robust large-area framing camera. *Rev Scientific Instr* (2021) 92:083103. doi:10.1063/5.0049110
228. Mozzanica A, Andrä M, Barten R, Bergamaschi A, Chirioti S, Brückner M, et al. The jungfrau detector for applications at synchrotron light sources and xfels. *Synchr Rad News* (2018) 31:16–20. doi:10.1080/08940886.2018.1528429
229. Hinger V, al Haddad A, Barten R, Bergamaschi A, Brückner M, Carulla M, et al. Advancing the jungfrau detector toward low-energy x-ray applications. *J Instrumentation* (2022) 17:C09027. doi:10.1088/1748-0221/17/09/c09027



OPEN ACCESS

EDITED BY

Ki Hyun Nam,
Kookmin University, Republic of Korea

REVIEWED BY

Carlo Fiorini,
Polytechnic University of Milan, Italy
Giovanni Di Domenico,
University of Ferrara, Italy

*CORRESPONDENCE

Jolanta Sztuk-Dambietz,
✉ jolanta.sztuk@xfel.eu

[†]Previously known as Ivana Klačková

[†]Previously known as Alexander Klyuev

[†]PRESENT ADDRESS

Stephan Stern,
X-Spectrum GmbH, Hamburg, Germany

RECEIVED 28 October 2023

ACCEPTED 18 December 2023

PUBLISHED 26 February 2024

CITATION

Sztuk-Dambietz J, Rovensky V, Klujev A, Laurus T, Trunk U, Ahmed K, Meyer O, Möller J, Parenti A, Raab N, Shayduk R, Sikorski M, Ansaldi G, Bösenberg U, Luis LM, Muenich A, Preston TR, Schmidt P, Stern S, Bean R, Madsen A, Gelisio L, Hauf S, Gessler P, Wrona K, Graafsma H and Turcato M (2024), Operational experience with Adaptive Gain Integrating Pixel Detectors at European XFEL.

Front. Phys. 11:1329378.

doi: 10.3389/fphy.2023.1329378

COPYRIGHT

© 2024 Sztuk-Dambietz, Rovensky, Klujev, Laurus, Trunk, Ahmed, Meyer, Möller, Parenti, Raab, Shayduk, Sikorski, Ansaldi, Bösenberg, Luis, Muenich, Preston, Schmidt, Stern, Bean, Madsen, Gelisio, Hauf, Gessler, Wrona, Graafsma and Turcato. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Operational experience with Adaptive Gain Integrating Pixel Detectors at European XFEL

Jolanta Sztuk-Dambietz^{1*}, Vratko Rovensky^{1†}, Alexander Klujev^{2§}, Torsten Laurus², Ulrich Trunk², Karim Ahmed¹, Olivier Meyer¹, Johannes Möller¹, Andrea Parenti¹, Natascha Raab¹, Roman Shayduk¹, Marcin Sikorski¹, Gabriele Ansaldi¹, Ulrike Bösenberg¹, Lopez M. Luis¹, Astrid Muenich¹, Thomas R. Preston¹, Philipp Schmidt¹, Stephan Stern^{2†}, Richard Bean¹, Anders Madsen¹, Luca Gelisio¹, Steffen Hauf¹, Patrick Gessler¹, Krzysztof Wrona¹, Heinz Graafsma^{2,3} and Monica Turcato¹

¹European X-Ray Free-Electron Laser Facility, Schenefeld, Germany, ²Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, ³Mid Sweden University, Sundsvall, Sweden

The European X-ray Free Electron Laser (European XFEL) is a cutting-edge user facility that generates per second up to 27,000 ultra-short, spatially coherent X-ray pulses within an energy range of 0.26 to more than 20 keV. Specialized instrumentation, including various 2D X-ray detectors capable of handling the unique time structure of the beam, is required. The one-megapixel AGIPD (AGIPD1M) detectors, developed for the European XFEL by the AGIPD Consortium, are the primary detectors used for user experiments at the SPB/SFX and MID instruments. The first AGIPD1M detector was installed at SPB/SFX when the facility began operation in 2017, and the second one was installed at MID in November 2018. The AGIPD detector systems require a dedicated infrastructure, well-defined safety systems, and high-level control procedures to ensure stable and safe operation. As of now, the AGIPD1M detectors installed at the SPB/SFX and MID experimental end stations are fully integrated into the European XFEL environment, including mechanical integration, vacuum, power, control, data acquisition, and data processing systems. Specific high-level procedures allow facilitated detector control, and dedicated interlock systems based on Programmable Logic Controllers ensure detector safety in case of power, vacuum, or cooling failure. The first 6 years of operation have clearly demonstrated that the AGIPD1M detectors provide high-quality scientific results. The collected data, along with additional dedicated studies, have also enabled the identification and quantification of issues related to detector performance, ensuring stable operation. Characterization and calibration of detectors are among the most critical and challenging aspects of operation due to their complex nature. A methodology has been developed to enable detector characterization and data correction, both in near real-time (online) and offline mode. The calibration process optimizes detector performance and ensures the highest quality of experimental results. Overall, the experience gained from integrating and operating the AGIPD detectors at the European XFEL, along

with the developed methodology for detector characterization and calibration, provides valuable insights for the development of next-generation detectors for Free Electron Laser X-ray sources.

KEYWORDS

MHz X-ray detectors, 2D detectors, FEL instrumentation, calibration, adaptive gain

1 Introduction

This paper examines the integration, functioning, characterization, and calibration of the first generation of AGIPD detectors, and discusses how these findings can be used to shape the development of new detectors for the European XFEL.

1.1 Overview of European XFEL facility

The European X-ray Free Electron Laser (European XFEL) [1, 2] is an international user research facility located in the Hamburg area that started operation in 2017. It currently features three free-electron laser x-ray sources, providing spatially coherent X-rays for seven experimental stations [3–8] in the energy range of approximately 260 eV to 25 keV. The sources can deliver up to 2,700 pulses with a repetition rate of up to 4.5 MHz in 10 equidistant X-ray pulse trains per second.

Various important scientific applications at the European XFEL, such as serial crystallography, single particle and material science experiments, require specific detectors that can cope with the MHz repetition rate of the machine and the unique time structure of the European XFEL, as well as a wide dynamic range of up to 10^4 photon/pixel/pulse whilst at the same time providing single-photon sensitivity at the same energy [9]. In order to address these challenges, three detector consortia successfully developed 2D area detectors for the European XFEL. Out of these, two have focused on detectors optimized for the hard X-ray energy range, delivering optimal performance at photon energies exceeding 10 keV: the Large Pixel Detector (LPD) [10] and the Adaptive Gain Integrating Pixel Detector (AGIPD) [11]. The DEPFET Sensor with Signal Compression (DSSC) [12] is designed for experiments utilizing lower-energy X-rays, down to a few hundred eV.

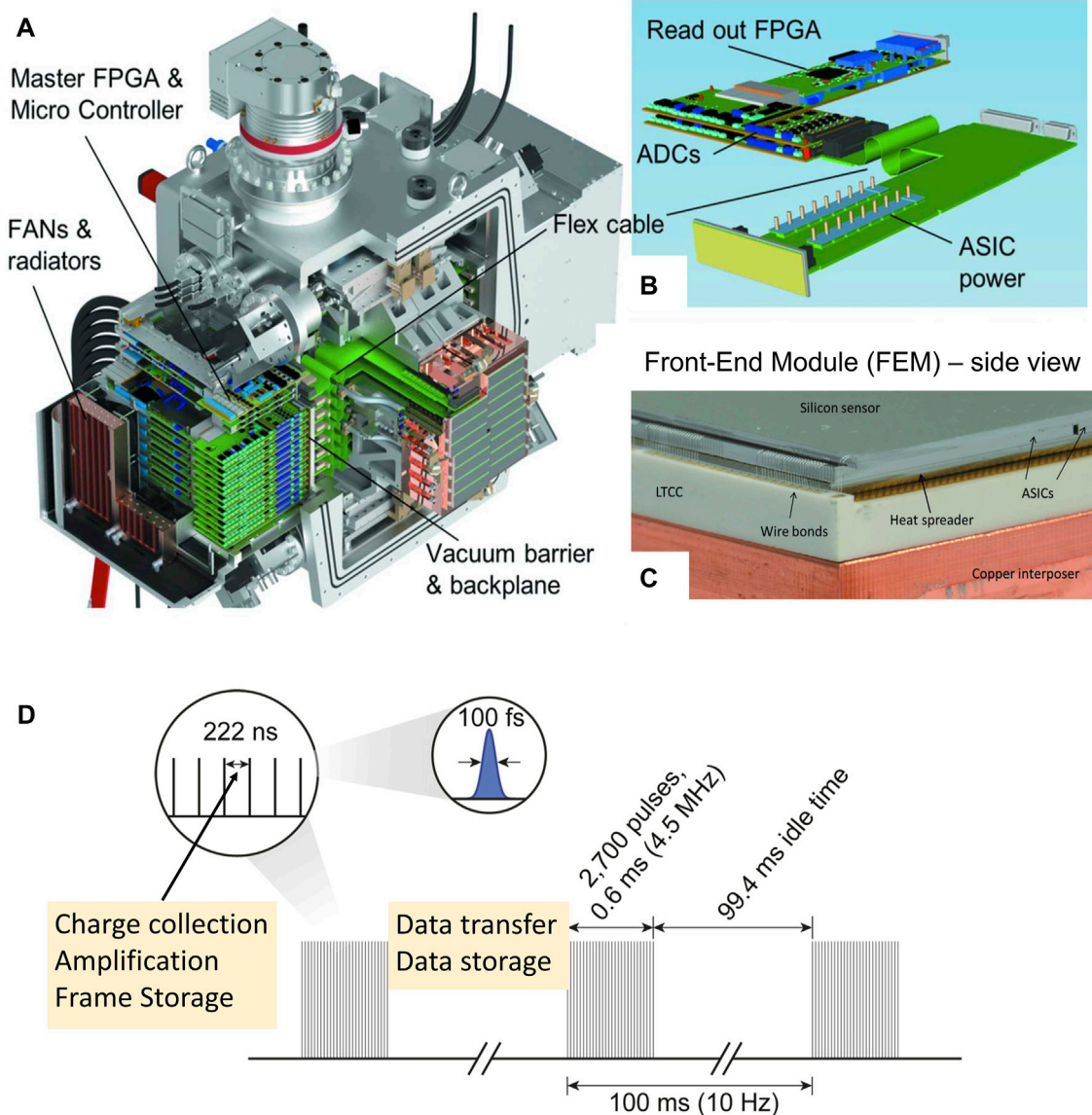
1.2 The AGIPD detectors for the European XFEL

The AGIPD detector [13], was developed for the European XFEL by an international consortium led by DESY, in collaboration with partners from renowned international institutions, including the Paul Scherrer Institute, University of Bonn, and Hamburg University. It features a classical hybrid pixel array with readout ASICs bump-bonded to a 500 μm thick silicon sensor [14]. The ASIC [15, 16] is designed using 130 nm CMOS technology and employs an adaptive gain switching technique to cover a wide dynamic range: from single photon to 10^4 photons per pixel per pulse at $E = 10$ keV. To achieve such a high dynamic range, each pixel utilizes a charge-sensitive preamplifier with three gain settings

that dynamically switch during the charge integration process. A comparator monitors the preamplifier's output voltage, which corresponds to the detected charge level. The preamplifier starts with its highest gain, and when the output voltage reaches the threshold, the comparator triggers gain switching by introducing an additional capacitor into the preamplifier's feedback loop. This results in a lower gain setting and higher noise. By progressively adding a maximum of two more capacitors to the initial one, the system allows for three gain settings: high (HG), medium (MG), and low (LG) gain. It also utilizes analog memory storage cells to store recorded images during the 0.6 ms duration of the pulse train. These images are subsequently read out and digitized during the 99.4 ms interval between pulse trains arriving at 10 Hz as it is shown in Figure 1D. The analog memory comprises two types of storage cells, one for amplitude values and the other for encoded gain settings. It is designed to store 352 images, equivalent to 352 samples per pixel, each pixel having a size of $200\ \mu\text{m} \times 200\ \mu\text{m}$. The storage cell matrix consists of 11 rows and 32 columns and occupies approximately 80% of the pixel area. Therefore, the number of storage cells is a compromise between the size of the pixels and the number of X-ray pulses that AGIPD can record. To optimize the use of this limited storage depth by overwriting unwanted images, the memory operates in random access mode. Furthermore, both the sensor and the ASIC components of the detector are optimized to withstand exposure to X-ray radiation [17].

Each ASIC is composed of a matrix of 64×64 pixels; 16 ASICs are bump-bonded to a 512×128 pixel silicon sensor, forming the sensitive hybrid assembly unit of the AGIPD detector. This hybrid assembly is then glued to the Low Temperature Co-fired Ceramic (LTCC) board, which is thermally and mechanically connected to the copper interposer, forming the fundamental detector unit - the Front-End Module (FEM). The FEM is connected by means of 500-pin SAMTEC connector to the back-end electronics. A photograph of the edge of a Front-End Module is shown in Figure 1C.

The one-megapixel detector consists of 16 FEMs, grouped into four independently moving quadrants, and is designed to operate in a vacuum environment. Figure 1A shows the CAD model of the AGIPD1M detector with cuts to reveal the arrangement of the electronics both inside and outside the vacuum vessel. From a control point of view, the AGIPD1M detector forms two electronically independent halves (called “wings”). The back-end electronics of each half consist of the ADC boards and the control and data IO board of each module (one set each per module, 8 modules per wing); a vacuum backplane board, which acts as a vacuum barrier and routes signals in and out of the vacuum vessel; a micro controller board for slow control; and a master FPGA board. These boards are located outside the vacuum chamber in a thermally sealed, water-cooled housing. The two master FPGA boards, one for each side, provide the interface to the European XFEL timing system and control the detector FEMs. Another part of the back-end are the

**FIGURE 1**

(A) CAD model of the AGIPD1M detector with sections cut away to reveal the arrangement of the electronics, both inside and outside of the vacuum vessel (B) The electronics of a single detector module. (C) A photograph of the edge of a Front-End Module (FEM) including annotations of the main components. Figures are sourced from [11]. (D) Pulse structure of the European XFEL and its impact on the requirements for detector data collection.

boards located in vacuum, they provide power to ASICs and IO signal connectivity between the FEM and the backplane. More details can be found in [11].

Two 1 Megapixel AGIPD detectors are used as primary detectors for experiments at the Single Particles, Clusters, and Biomolecules and Serial Femtosecond Crystallography (SPB/SFX) [3] and Material Imaging and Dynamics (MID) [5] instruments of the European XFEL. The SPB/SFX AGIPD1M system has been installed since the start of operation in 2017, while the MID AGIPD1M followed in November 2018. Both AGIPD1M detectors at the European XFEL are based on the same hardware and firmware, currently using the AGIPD 1.1 ASIC version [16].

It should be mentioned that the second generation of the AGIPD detector is under development and is already in use at the European

XFEL as the AGIPD500K prototype, which was installed in September 2020 at the HED Instrument [18]. The AGIPD500K comprises eight modules and is operated in air. This second-generation AGIPD features a new version of electronic boards, new back-end electronics architecture, and uses the new AGIPD1.2 ASIC version [19]. Unlike the AGIPD1M systems, which used complex configurations of several boards with a single function on each, as shown in Figure 1B, the new, more compact readout board incorporates a streamlined design that houses both the analog-to-digital converter (ADC) and a new FPGA, along with all-optical communication via multi-fiber Gbit transceivers. Despite being a prototype, the AGIPD500K detector at HED has been employed in a number of successful user experiments [20, 21] and more are planned.

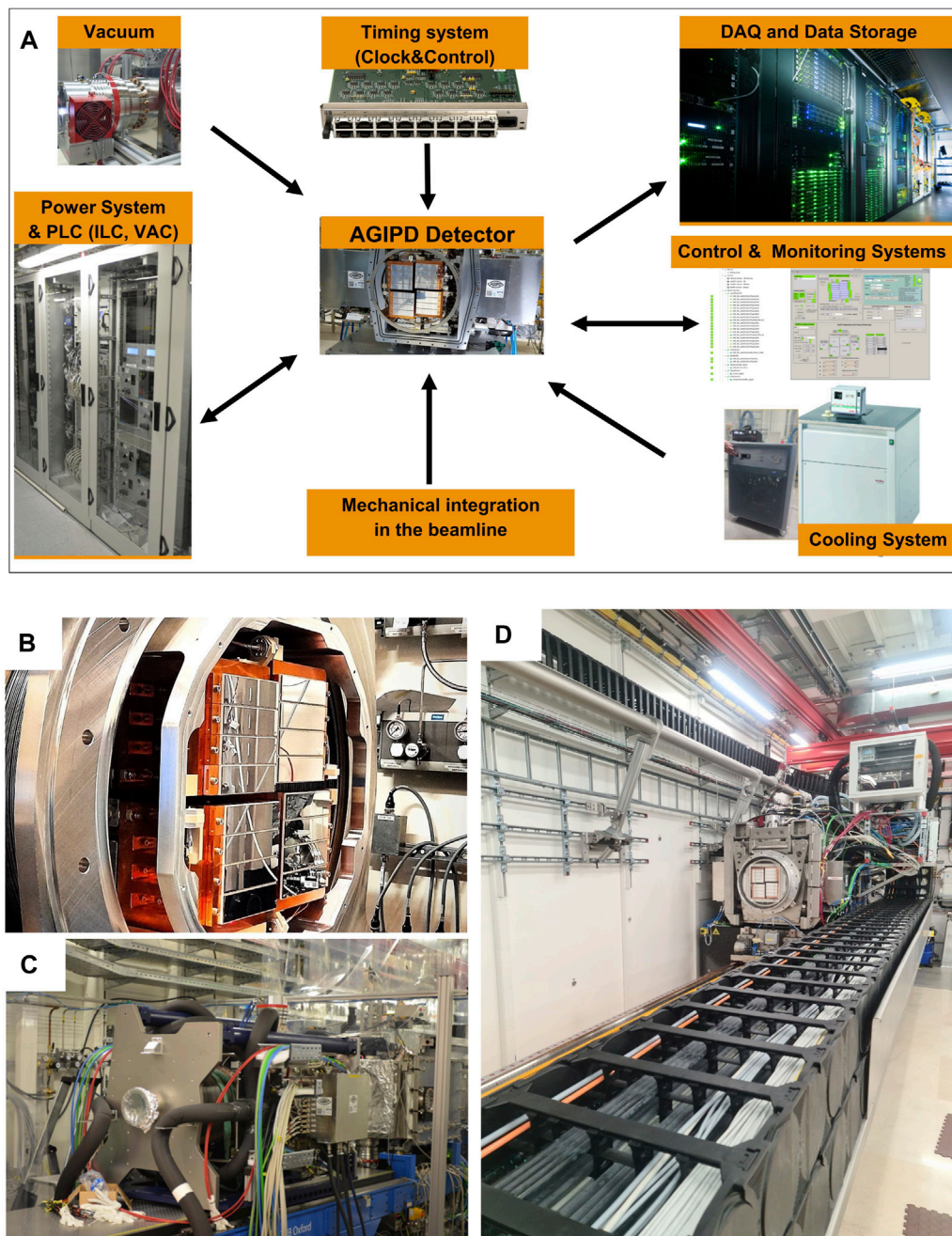


FIGURE 2

(A) Schematic overview of the required infrastructure for AGIPD detector integration. (B) Close-up image of AGIPD1M installation at SPB, showcasing the Front-End Modules (FEMs) in detail. (C) Back view of AGIPD1M installation at SPB, featuring the detector with all cables and pipes connected to the vessel. (D) AGIPD1M installation at MID, with a focus on the drag chain in the foreground, highlighting the detector cables.

2 Detector integration and infrastructure

The complexity of integrating a detector into an instrument and the amount of infrastructure required to operate it are illustrated in Figure 2A. The mechanical integration of the detector as well as necessary infrastructure (i.e., cooling, vacuum and powering) was carefully considered to match each instrument's requests. Furthermore, the detectors have been integrated into the

European XFEL control and data acquisition system from a software perspective. The framework of the User Facility generates additional constraints in terms of operability: the detectors are used continuously during long periods (up to 6 days in a row) and run by various teams. Those constraints imply to make the detector setup reliable, secure (interlock systems), maintainable (access to the consumable components) and easy to operate (specific control interfaces). Before being placed at the instrument, the detector was installed in the laboratory for comprehensive

testing. This enabled the identification and solving of initial problems and a more efficient installation in the instrument.

2.1 Mechanical integration into the instrument

Both AGIPD1M detectors were required to operate in a vacuum environment. These considerations also factored in the impact of vacuum forces and aimed to ensure the isolation of the detector vacuum from the instrument sample chamber, minimizing the need for frequent system rewarming. Given the multitude of cables required for power, cooling, motors, data acquisition (DAQ), and control, comprehensive support structures were essential. Additionally, accessibility to the Front-End Modules (FEMs) needed to be as straightforward as possible.

Preparation for the mechanical integration of the detector into the beamline began in parallel with detector development. This dual-track approach required adjustments on both fronts to meet the instrument requirements and the needs of the detector development team. The outcome of this integration process is illustrated in Figures 2B–D.

In the current installations, it is relatively straightforward to access the front of the detector, where the FEMs are installed, and this process typically takes only a few hours to 1 day. This is the result of a design decision, since the FEMs are the detector components deemed by far most likely to experience damage—from radiation, by cooling and vacuum accidents, and due to mechanical mishaps. However, gaining access to the rear of the detector poses a significant challenge in terms of both time and the increased risk of damaging critical detector or beamline components. If access to the vacuum boards within the vacuum vessel becomes necessary—for maintenance, repairs or replacements due to malfunctions—a considerable effort is required to access and then open the rear flange of the detector.

2.2 Power system

The AGIPD1M detector electronics is powered by a set of WIENER [22] power supplies, which include Low-Voltage (LV) and High-Current MVP8016 and MVP8008 MPOD multichannel power supplies, providing 16V/5A and 8V/8A power output, respectively. The sensor bias voltage (ranging from 300 to 500 V) is supplied by a High-Voltage (HV) and Low-Current ISEG unit. These power modules are installed and controlled via MPOD crates.

The assignment of MPOD voltage channels (including LV sense) to the detector-head input connectors and cables is carried out near the terminal block, often referred to as patch-panel. Moreover, the terminal block serves the purpose of connecting interlock signals from the power supplies to a PLC (Programmable Logic Controller) and establishing connections for relevant interlock lines between the detector head and the PLC.

In particular, there are no internal DC/DC converters implemented at the detector level in the first generation systems. Instead, each detector board and sensor receive dedicated power from an external power supply. It is worth mentioning that no operational or stability problems have been observed with this

power system. The MPOD high-precision supplies meet the electrical requirements (voltage, current, and interlock functionality) of the detectors and can handle the lengths of the cables (30–40 m) used, which is significant. This approach has certain limitations, mainly due to the large number of cables (>100) required to operate AGIPD1M, including 56 power cables, 20 control cables, 16 data cables, and 16 cooling pipes. The thickness of the power cables ranges from 1.5 to 2 cm each, while the cooling lines are composed of pipes ranging from 2 to more than 10 cm in diameter. Therefore, effective cable management and mechanical support are essential when moving the detector within the hutches, as shown in Figure 2D. A more compact design, such as the implementation of DC/DC converters and the reduction of the distance between the detector and dedicated power supplies, would enhance the flexibility of the system and reduce the risk of damaging the detector or power cables during movement.

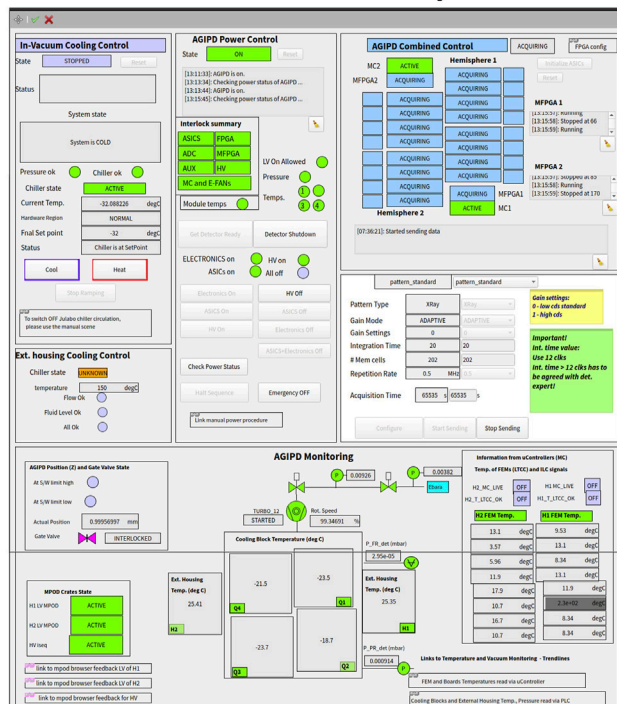
2.3 Cooling system

The AGIPD1M systems require two cooling circuits: one for the electronics located outside vacuum (with a power consumption of more than 2 kW) and another for the electronics in vacuum (with a power consumption of less than 0.5 kW). For cooling the detector components outside vacuum, we employ a water-based cooling system utilizing commercially available dedicated chillers. No issues have been observed with the cooling of electronics outside vacuum, and the monitored temperatures on the electronic boards remain consistently below 40°C.

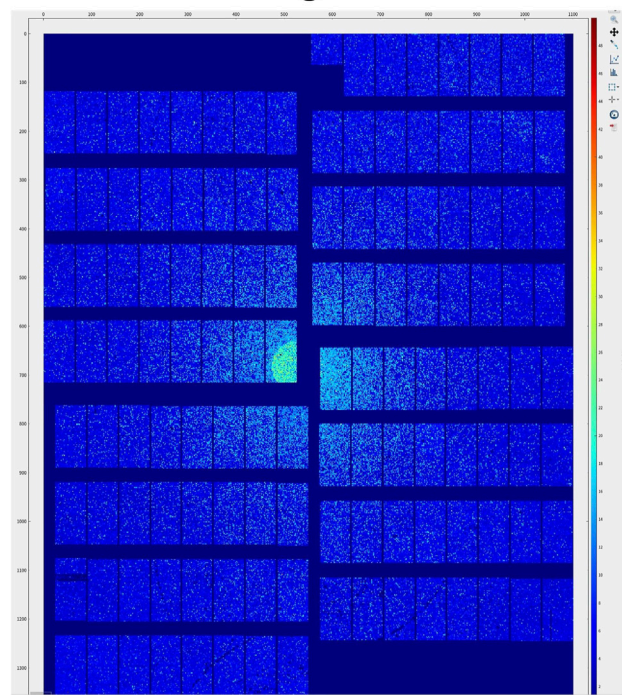
Cooling the in-vacuum part of the detector presents considerably greater challenges. In-vacuum electronics are cooled using a customized Julabo water-cooled chiller with silicone oil serving as coolant. The chiller is located in the experiment rack room and requires long (30–40 m) well-insulated cooling pipes to deliver the coolant to the detector cooling blocks. It is essential to note that silicone oil can be difficult to clean and poses a potential hazard to other components of the beamline if it leaks. Fortunately, over the years, no evidence of silicone oil leakage in vacuum has emerged. However, there have been some leaks in the cooling system, particularly in the oil distribution between the chiller and the detector.

The temperature difference between the coolant in the chiller (−32°C) and the temperature of the detector cooling blocks, when no power is applied to the detector head, exceeds 5°C. When the ASICs are active, the temperature of the cooling blocks increases further to above −24°C. FEM temperature measurements, obtained using PT100 sensors installed on the back side of the LTCC board, indicate temperature values ranging from −3°C to +15°C. These readings suggest that the cooling efficiency is limited and that the target temperature (−20°C) cannot be achieved. The gradient between the coolant temperature at the chiller and the attained temperatures on the FEMs is quite substantial, reaching up to 50°C, as illustrated in Figure 3C. The primary cause of the thermal gradient is the limited heat transfer efficiency from the ASICs, which are the main source of heat (approximately 30–40 W per module), in the FEM assembly to the cooling block. The current design, as illustrated in Figure 1C, includes several stacked thermal interfaces, resulting in multiple layers of thermal resistances.

A AGIPD main control panel



B Corrected image – online view



C AGIPD cooling control and monitoring

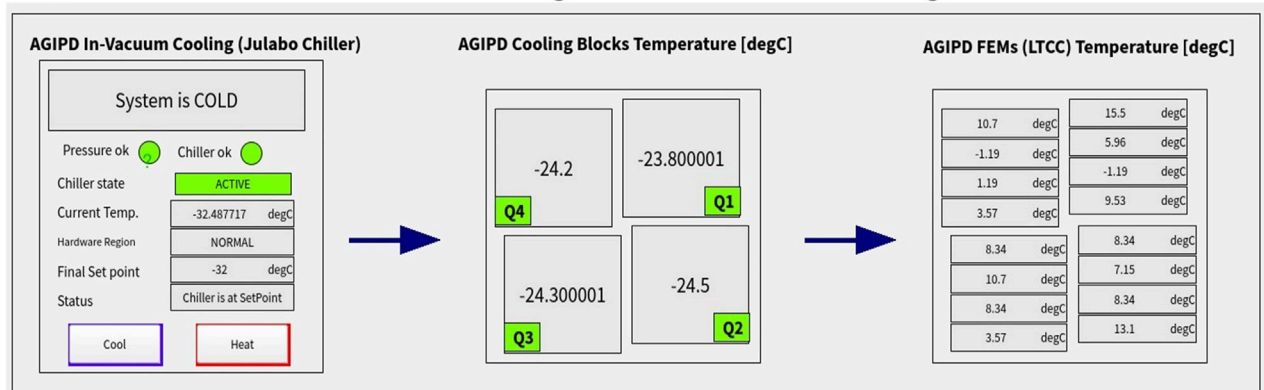


FIGURE 3 (A) Primary AGIPD1M control panel. (B) Illustration of an online-corrected image displayed in the Karabo GUI. (C) Real-time control and monitoring of in-vacuum cooling through the control system, providing temperature information for the cooling blocks and Front-End Modules (FEMs).

Additionally, a large 500-pin connector on the back of the LTCC board hinders heat transfer efficiency, especially in the vacuum environment. Lowering the temperature of the coolant is not feasible, as it can lead to overcooling and malfunction of certain components of the electronic boards installed in the vacuum vessel. Nevertheless, the temperature of the FEMs stabilizes after several minutes and remains constant without fluctuations that could impact the detector characteristics (i.e., the temperature change over time remains below 1°C).

To address this problem in the next-generation of detectors, a dedicated R&D program has been initiated. The goal is to explore more efficient cooling methods, such as micro-channel cooling [23], to enhance overall cooling performance. The advantage of this

cooling method lies in the positioning of micro-channels directly under the readout ASICs, optimizing heat dissipation.

2.4 Interlock system for detector protection

The interlock system for the AGIPD1M detectors is based mainly on programmable logic controllers (PLCs) [24]. These PLCs monitor the vacuum quality and cooling efficiency, reading out temperatures of the detector cooling blocks; pressure values from the sensors installed at the detector vessel and in the connected sample chamber; chiller conditions; and internal detector conditions evaluated by the detector slow-control board, such as hardware

TABLE 1 Interlock triggers and actions.

Input	Interlock trigger	Interlock action
Vacuum status	Pressure $p > 10^{-3}$ mbar or pump failure	Warm up detector to room temperature, Switch off HV, Close relevant valves
Cooling blocks temperature	Temperature $T > 0^{\circ}\text{C}$	Switch off power for components in vacuum (HV, ASICs, vacuum boards)
Electronics temperature (outside vacuum)	Temperature $T > 35^{\circ}\text{C}$	Switch off power for all components (except MicroController)
MicroController signal (2nd level interlock)	FEM Temperature $> 50^{\circ}\text{C}$	Switch off power for all components (except MicroController)

conditions and temperatures of the FEMs. If necessary, the PLCs take appropriate actions. The interlock logic implemented for the AGIPD1M detectors is shown in Table 1.

Over several years of practical experience, this system has proven to be highly effective in protecting the detector against unexpected failures, such as pump or chiller failures, power outages, or inadvertent cooling shutdowns. This level of protection is absolutely indispensable for smooth detector operation within a user facility like the European XFEL. Ideally, the detector system has to possess self-protective capabilities.

Another category of incidents refers to radiation damage resulting from the primary or scattered/diffracted beam. Such incidents can occur, and components primarily affected by radiation, such as ASICs and sensors, should be explicitly designed to be radiation-hard, as protecting them from such events is generally challenging. In most cases, radiation damage is caused by high-intensity X-ray radiation generated in a single shot, which leaves no time to react. However, there is a class of incidents that can be prevented or significantly minimized. An example is the formation of ice on the sample injection nozzle. This can be monitored, and if ice is observed, the beam must be attenuated or the shutter must be closed. Such a protection system has recently been developed and is planned to be used at the SPB/SFX instrument.

2.5 Control and monitoring system for AGIPD1M

The control system for the AGIPD1M detector, together with its associated infrastructure including power supplies, cooling and interlock systems, has been integrated into the European XFEL control framework known as Karabo [25]. Its core relies on transparently distributable servers, offering functionality through pluggable “devices”. These devices can function as control interfaces for tasks such as motor control or detector control, monitoring interfaces for measurements such as pressure readings, or computational devices for data processing.

Karabo also supports complex procedures, such as detector startup, by employing “middle-layer devices”. These devices do not directly interact with hardware but instead coordinate with other control devices, consolidate incoming data, and execute the necessary steps. These high-level interfaces enable detector operators, including instrument scientists, to efficiently manage these systems independently. In addition, they serve as a second-level safety mechanism which will not allow to perform actions that can be dangerous for the detector hardware (i.e., powering the detector when the temperatures or pressure are not in the expected range). These interfaces facilitate various tasks, including power management, configuration adjustments, and the efficient acquisition of calibration

data needed for the production of calibration constants used for raw data corrections. As an example, Figure 3A displays the AGIPD main control panel, while Figure 3B presents an online-corrected image.

Interactivity is provided through either a command line interface (CLI) using IPython with automatic command completion or a PyQt-based graphical user interface (GUI). The GUI consolidates parameters, control and data flows, state information, and error feedback, enabling the creation of complex control and processing configurations, including data analysis pipelines such as the data correction pipeline.

Any newly developed operational features undergo rigorous integration and testing procedures before being deployed for user operations. Consequently, the control system for the AGIPD1M detectors demonstrates impressive stability and is free of significant issues or instabilities.

2.6 Data collection and storage infrastructure

Acquiring meaningful data from the detector requires precise synchronization with the European XFEL beam, a feature accomplished through the Clock and Control (C&C) system [26]. This system plays a crucial role in ensuring synchronization by providing essential components such as clocks, bunch, and train-related information. Importantly, the C&C system can also issue veto signals to discard undesirable bunches, although it is pertinent to mention that AGIPD currently does not employ this functionality. The information from the C&C system is first received by the AGIPD1M Master FPGA (MFPGA), and then is distributed to all detector modules. Following this, raw data from the detector, including the train Id information, is transmitted via the UDP protocol to the European XFEL Data Acquisition PC Layer and written in HDF5 data format [27]. Initially, this data is stored within the online GPFS (General Parallel File System) [28] cluster, while the metadata is concurrently made accessible through the metadata catalogue myMdc [29]. myMdc offers the ability to evaluate the data before copying it from the online cluster to the offline GPFS cluster. Moreover, in addition to the raw data, the corrected data, referred to as “processed data,” is also retained.

3 Data quality: characterization and calibration

3.1 Calibration strategy

Accurate calibration and comprehensive characterization of detectors are critical aspects to ensure the successful operation of

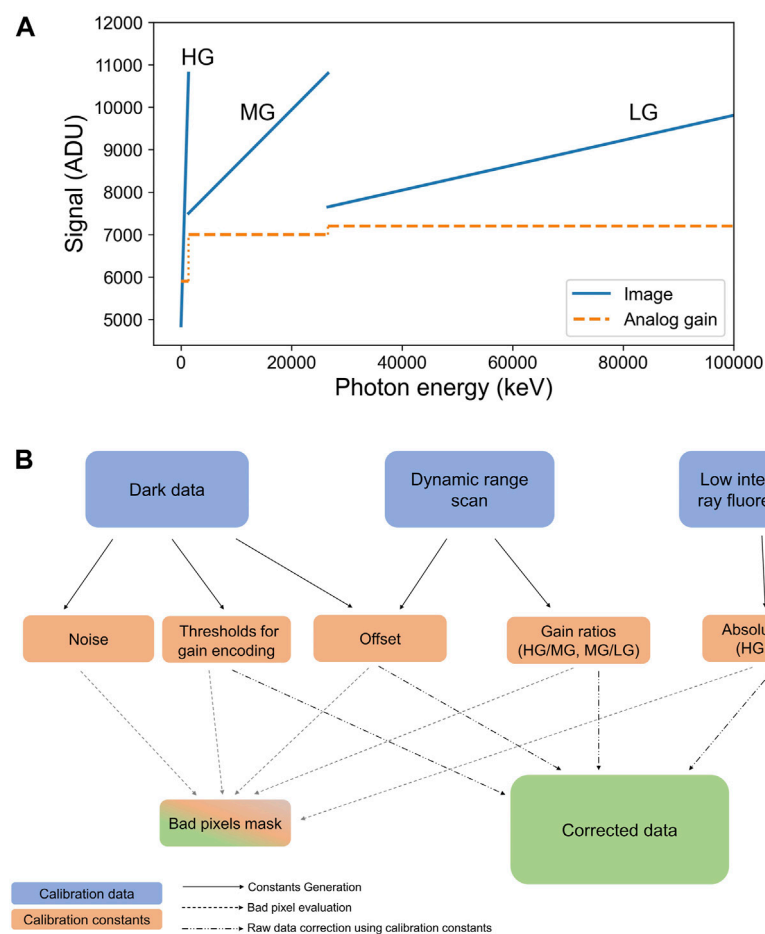


FIGURE 4

(A) Visualization of AGIPD's dynamic range, showcasing three adaptive gain settings: High Gain (HG), Medium Gain, and Low Gain (LG). (B) Flowchart illustrating AGIPD's calibration strategy.

instruments employing these detectors. Equally important is the development of user-friendly procedures and tools that facilitate recalibration during the operational phase. In this section, the general concept of AGIPD calibration, shown in Figure 4, is explored, and a more detailed explanation of the calibration process is provided.

To fully calibrate the AGIPD detector, three distinct data sets are required:

1. **Dark data** (without any external stimulus) to determine offset and noise values for each gain setting and two thresholds for gain encoding.
2. **Dynamic range scan** of available gain settings with internal calibration sources to provide ratios between different gain settings as well as offsets for each gain.
3. **Low-intensity fluorescence data** are used to determine the absolute gain value.

All three gain settings, as illustrated in Figure 4A, require characterization for each memory cell within each pixel. This characterization involves 11 parameters, including:

1. **Three offsets**, one for each gain setting.
2. **Two thresholds** for gain setting encoding.
3. **One absolute gain value**.
4. **Two gain ratios**: HG/MG and HG/LG.
5. **Three bad pixel maps** for each gain setting using information from the data sets mentioned above, and it is kept as a 32 bit mask to prevent the source of the bad pixel from being lost.

This results in more than 4×10^9 unique parameter values for an AGIPD1M detector. The calibration constant values strongly depend on the operating mode, including also factors such as acquisition rate, the number of memory cells used, and integration time. Therefore, each operational scenario requires dedicated calibration data collection and analysis to derive calibration constants tailored to that specific operating mode. The constants are generated using a dedicated calibration software, which is run on the Maxwell - HPC cluster [30] at DESY. Table 2 shows the estimated time needed for the collection and processing of the calibration data to obtain a complete set of calibration constants for the AGIPD1M detector and one operation mode. It also includes how frequently the

TABLE 2 Summary of data collection and processing times and sizes, as well as frequency of generation for one complete set of calibration constants for the AGIPD1M detector at the European XFEL. The presented values are based on the operation mode with all memory cells and do not include preparatory time required for the measurements.

Data type	Data size (TB)	Measurement time (mins)	Data processing time (mins)	Frequency
Dark Data	2.2	5	~ 10	at least once per shift
Dynamic Range Scan - Pulsed Capacitor	8.2	20	~ 100	6 months
Dynamic Range Scan - Current Source	21	65	~ 180	6 months
Fluorescence Data	15–20	25–30	up to 720	6–12 months

constants are generated for each of the operation modes. Multiple detector operation modes were implemented to enhance detector performance for specific types of experiments and address observed problems, as detailed in Section 4.

The generated constants are injected into the calibration database [31] and automatically retrieved to correct the raw data [32]. It is quite a challenging process because of the amount of data that has to be collected and processed to generate the full set of calibration constants. Another challenge is caused by detector artifacts, which require additional data treatment, as described in more detail in Section 4.2.

3.2 Detector baseline

The detector baseline (offset) is defined as a dark signal measured in the absence of external stimuli, such as X-ray photons. In the ideal scenario, the baseline measured under well-defined conditions (e.g., temperature, readout speed, integration time) should enable the subtraction of signals not originating from incoming photons. It should also be independent of the intensity value and remain stable over time.

The dark data is acquired using a dedicated configuration, which allows the detector to be forced into Medium or Low gain settings. The collected data is then processed, and the new versions of constants (offsets and noise maps) that are generated can be used for data correction. The offset O is calculated as the median of the dark signal (D_s) over a certain number of trains t for a given gain setting (gs: 0-HG, 1-MG, 2-LG), pixel (x, y) and memory cell c . The noise N is calculated as the standard deviation σ of the dark signal ($O_{x,y,c} = \text{median}(D_s)_t$, $N_{x,y,c} = \sigma(D_s)_t$).

The calibration constants from dark data are generated on a regular basis (i.e., at least once a day during user experiments). To ensure the frequent generation of the constants, the acquisition and processing of dark data is automated, facilitating the task for the beamline operator.

3.3 Gain setting identification

As mentioned in Section 1.2, the beam's time structure at the European XFEL does not allow for a continuous readout of single frames during pulse trains. Therefore, each frame in the train, up to a maximum of 352, must be stored in pixel and is read out in the 99.4 ms gap between pulse trains. Since both the information

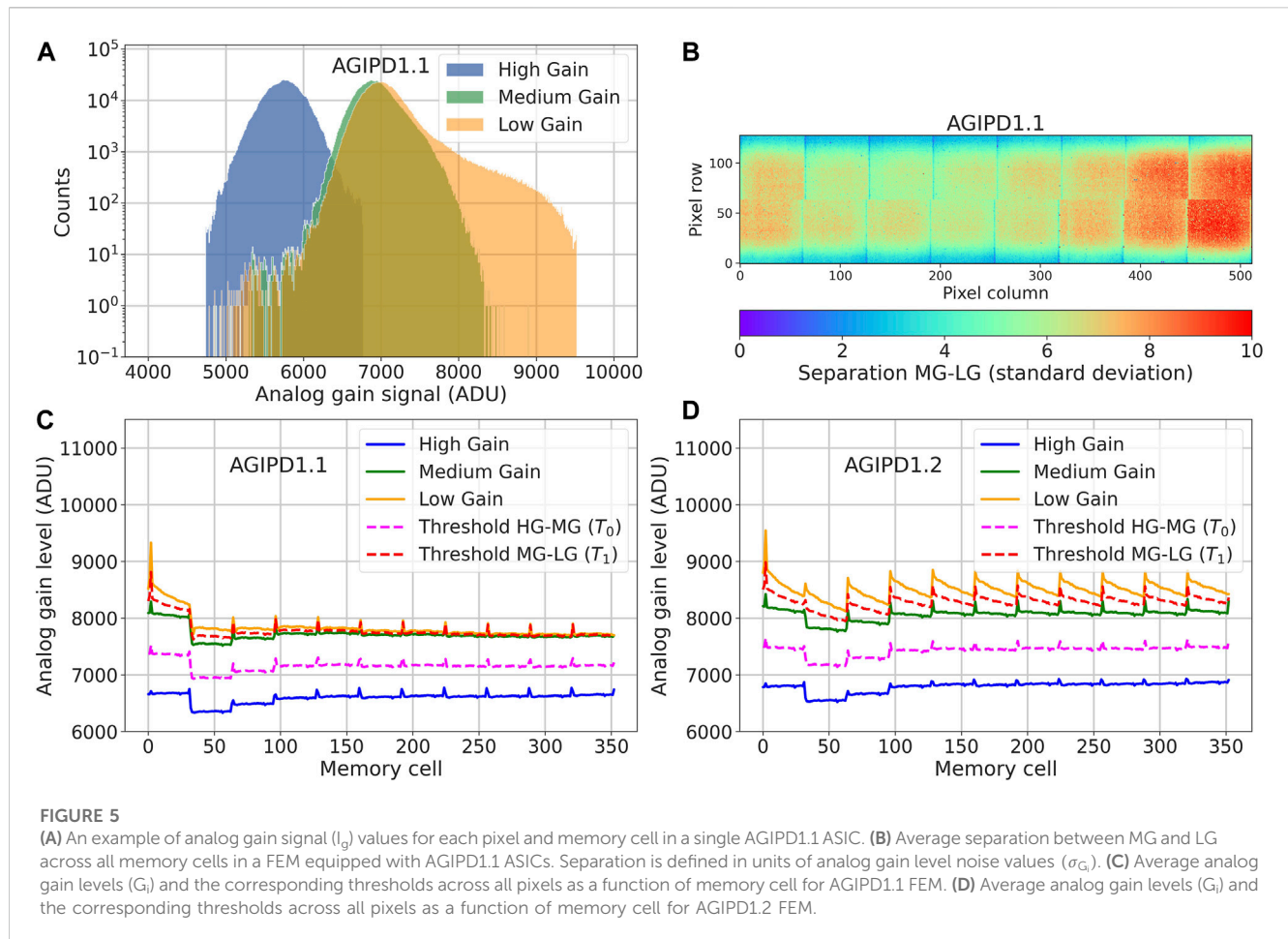
regarding the pulse height of the charge-integrated signal and the gain status are essential for extracting incident X-ray intensities on the detector, each AGIPD pixel has two storage cell matrices, each consisting of 352 capacitors. One matrix stores analog information about the height of the X-ray signal, and the second one stores the signal related to the gain setting (i.e., $V = 0.7$ V for HG, 1 V for MG and 1.5 V for LG) that is being used (I_g).

The dark data is used not only for noise and offset determination, but also for generation of constants that are needed for the identification of gain settings, so-called thresholds. Analog gain levels (G_i) for each gain setting (where $i = \text{HG, MG, and LG}$) are calculated as the median of the dark gain signal (median (D_{gs})) over a certain number of trains t for a given gain setting (gs: 0-HG, 1-MG, 2-LG), pixel (x, y) and memory cell c . The noise associated with the gain level is quantified as the standard deviation of σ_{G_i} of the dark gain signal. It is essential that the signal level indicating the gain setting is clearly distinguishable from the analog gain signal noise, ideally surpassing a limit of at least 5 standard deviations (σ_{G_i}) for different gain settings. To encode the gain setting, the gain signal (I_g) of each pixel and memory cell is evaluated against two defined thresholds, T_0 and T_1 , which categorize the pixel as follows:

- High gain if $I_g \leq T_0$
- Medium gain if $T_0 < I_g \leq T_1$
- Low gain if $I_g > T_1$

Where the thresholds T_0 and T_1 are determined as the mean value between individual analog gain levels (G_i), that is, $T_0 = \frac{G_{\text{HG}} + G_{\text{MG}}}{2}$ and $T_1 = \frac{G_{\text{MG}} + G_{\text{LG}}}{2}$.

The original idea was to use two thresholds per chip (64×64 pixels) since the three gain values were expected to be well-separated and have similar values for each pixel and memory cell. A study [16] carried out on a single AGIPD1.1 ASIC in a dedicated test system showed that up to 0.5% of pixels can have an incorrect gain assignment if only two threshold values are applied per chip. Analysis of data collected with the full-scale system (AGIPD1M at the European XFEL) revealed that applying 'sanity cuts' to filter out outlier values of analog gain (i.e., selecting gain values within ± 5 standard deviations from the average gain value for all pixels and memory cells) resulted in the removal of at least 0.5% of the pixels. Analog gain values for medium and low gain in a single ASIC overlap, as shown in Figure 5A. Therefore, using only two thresholds per chip is not feasible if we want to maintain a sufficiently low probability of



‘false’ encoding (preferably below 0.1%). Studies show that the gain separation MG-LG strongly decreases for pixels close to the chip periphery (Figure 5B) and exhibits a strong temperature dependency, pointing towards a leakage current. Another observed issue, as shown in Figure 5C, is that the difference in the gain level value for medium and low gain degrades along the memory cell’s number in the read-out sequence, and it is not always possible to determine if the signal is collected at medium or low gain.

The observed effect is caused by a bug in the chip design, which, during the data readout of the 352 image and 352 gain memory cells, results in a partial discharge of the gain signal in a particular gain memory cell if the encoded gain level stored in that cell corresponded to LG. As gain cells are read out sequentially from cell0 to cell351, the cells that are read out later have already been deprived of a significant amount of charge if they are in LG, causing their voltage levels to reduce to levels encoding MG. Therefore, for cells read out later, no MG-LG distinction would be possible anymore, even if individual thresholds are applied to each memory cell in each pixel. A new version 1.2 of the AGIPD ASIC mitigates this issue (Figure 5D). However, both AGIPD1M systems at the European XFEL are currently equipped with the AGIPD1.1 ASIC version. Therefore, low gain is not used for scientific analyses and is still not fully characterized. The installation of FEMs equipped with AGIPD1.2 ASICs is planned in the near future.

3.4 Dynamic range

To establish a relationship between different gain settings and effectively calibrate the entire gain of the detector, it is essential to perform an intensity scan throughout the entire dynamic range of the detector. However, using the European XFEL beam, conducting a comprehensive intensity scan for every memory in every pixel within the entire dynamic range proved to be unfeasible. Therefore, X-ray photons are exclusively used to determine absolute gain factors in HG, as is elaborated in Section 3.5.

To perform a dynamic range scan of the detector, we rely on the available internal “on-chip” calibration sources. Alternative methods, as described in [33], such as dynamic range scans with sensor backside pulsing, IR pulsed laser, pulsed monoenergetic proton beams, or LED light, are not viable for the full-scale AGIPD1M system for several reasons. Primary factors include the need for specialized installations and interfaces to the ASICs, which are absent in AGIPD1M and would require a complete disassembly of the detector to execute the scan. Some of these methods are suited only for single-pixel scans, for example, the laser and pulsed monoenergetic proton beam, which are not applicable to the detector module or system. Therefore, in our specific case, we rely on the internal calibration sources implemented at the ASIC level.

The AGIPD detector provides two options for injecting test charges into pixels. One approach for scanning the dynamic range involves employing the Pulsed Capacitor (PC), a circuitry implemented at the input of the preamplifier of each pixel, which

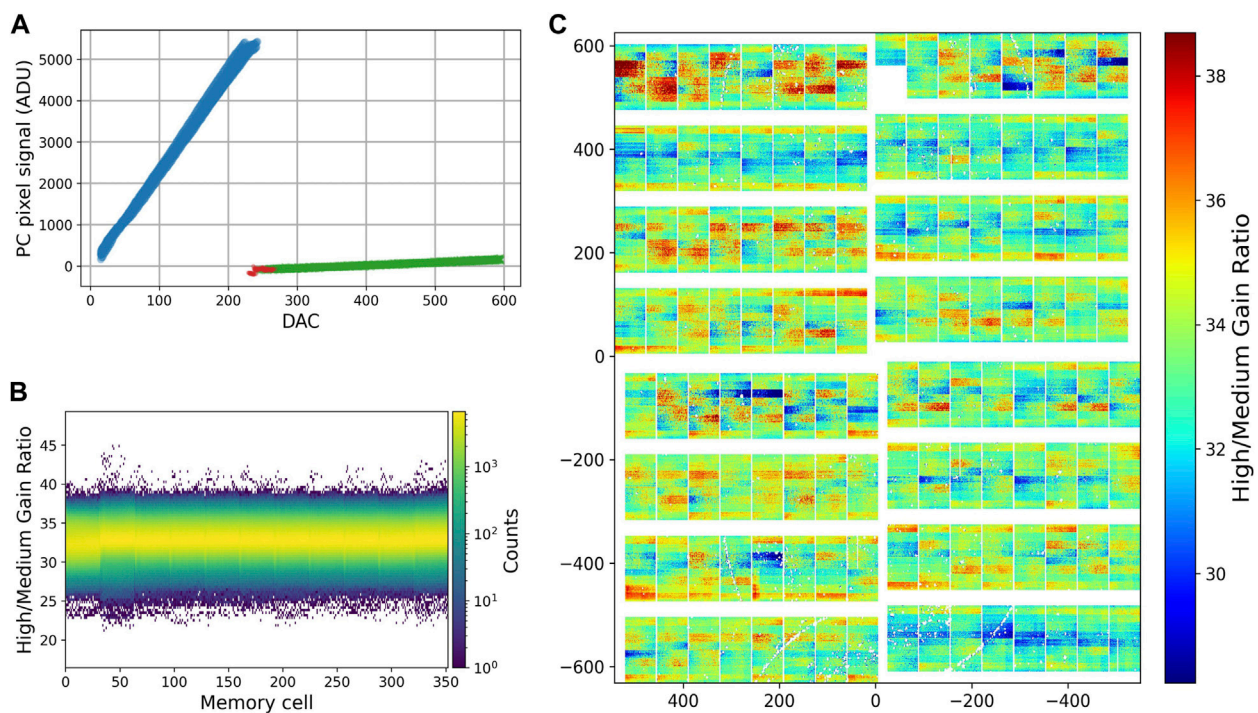


FIGURE 6

(A) Dynamic range scan example using the Pulsed Capacitor (PC). The blue region corresponds to the high gain setting, red represents the transition region between the high and medium gain, and green represents the medium gain region. (B) High gain to medium gain ratio values for all pixels of a single module, plotted as a function of the memory cell index. (C) High gain to medium gain ratio map for AGIPD1M, averaged across all memory cells.

allows for applying voltage steps to a capacitor to introduce a defined, albeit small, charge. The second approach utilizes an on-chip current source (CS), which permits the injection of a constant current into the preamplifier's input. By increasing the integration time, the amount of charge injected will increase in proportion, allowing for a thorough scan of the detector's dynamic range. The current source implemented on the AGIPD ASIC is programmable, allowing its current value to be tailored to specific ASIC instances to encompass the entire dynamic range.

At the European XFEL, we primarily utilize the PC source for calibration purposes. The CS source is also implemented, and the first set of calibration constants has been generated. However, the application of LG in the current version of AGIPD1M is restricted due to the issue mentioned in Section 3.3. Consequently, the full characterization and validation of this internal calibration source remains a work in progress. To prepare the calibration, we first scan the dynamic range by injecting test charges into the pixels. The calibration process then involves two steps. First, we subtract offsets derived from the dark data from the scan data. Next, we model the data distribution using a fitting method to determine the relationship between different gain settings. An example of a dynamic range scan with Pulsed Capacitor is shown in Figure 6A.

For the initial part of the distribution, a linear function is fitted to the data, allowing us to establish the gain slope and offset for high gain settings, expressed as $y = mx + l$. For the subsequent regions, we employ a composite function: $y = A \cdot e^{-(x-o)/C} + mx + l$, which enables us to cover both the transition region and the medium gain slope and offset. However, due to the inherent challenges posed by the transition region between high and medium gain settings, as discussed in more detail in

Section 4.2.2, and the difficulty in reliably applying these models for the correction of detector data (since distinguishing between pixels in the transition region and valid MG values can be challenging), we opted to exclude from the analysis the scan intensities corresponding to the transition region. Instead, we fit only the linear part of this function to generate the calibration constants. The results of the analysis are presented in Figures 6B, C. In Figure 6B, the gain ratio values for all pixels in a single AGIPD FEM module are shown as a function of memory cell index. Figure 6C shows the average gain ratio (HG/MG) map for AGIPD1M at SPB/SFX, computed across all memory cells.

The procedure for obtaining calibration constants from the dynamic range scan data collected with CS is similar. In this case, we fit three linear functions to each gain region and derive the gain setting ratio from the slopes of the fit.

It is important to note that while both sources are valuable tools, they do come with limitations: the PC source can cover only HG and a relatively small part of MG, typically around 10%, primarily up to a few hundred 10 keV photons. On the other hand, CS spans all gain settings, but is not compatible with the European XFEL timing. In both cases, certain artifacts may be visible in the data that are not always reproducible when collecting X-rays.

3.5 Absolute gain conversion factors from low intensity fluorescence data

The absolute calibration in the high gain (HG) region, quantified in terms of the conversion factor ADU/keV, is

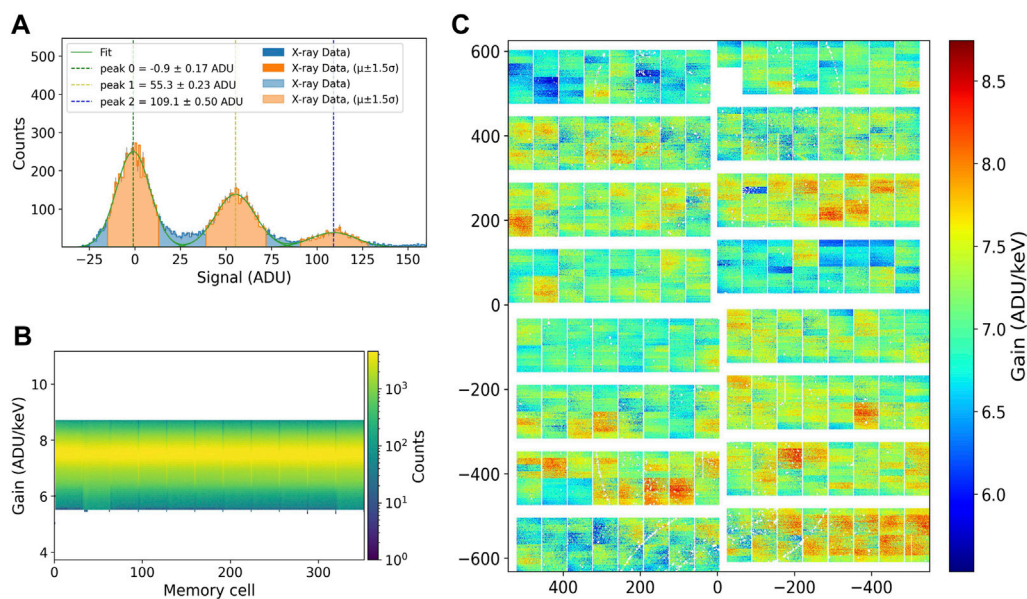


FIGURE 7

(A) An example of low-intensity Cu fluorescence ($E_\gamma = 8.05$ keV) spectrum for single pixel and memory cell. Multi-Gaussian function was fitted to the data. (B) Absolute gain values of high gain setting as a function of memory cell index for one of FEMs installed in AGIPD1M detector at SPB/SFX. (C) Absolute gain map of high gain setting for AGIPD1M installed at SPB/SFX, averaged across all memory cells.

executed using low-intensity, flat-field fluorescence photon data. These measurements are conducted right at the beamline, eliminating the need to disassemble the detector system from the instrument. Typically, a copper (Cu) foil serves as the medium for these measurements, although at higher energies, such as with Yttrium Y), the measurements are adapted accordingly. To prevent rapid degradation of the material under the intense XFEL beam, the X-ray beam size on the foil is intentionally defocused to approximately $100 \mu\text{m} \times 100 \mu\text{m}$. Additionally, the intensity of the incident X-ray beam is finely tuned to yield a detectable signal of around half a photon per pixel per pulse. The distance between the detector and the interaction point is adjusted to ensure uniform illumination of the entire detector simultaneously. This uniformity can be achieved at distances on the order of 100 cm and beyond.

The collected data (approximately 15,000 events for each memory cell in every pixel) after offset subtraction is used to determine the absolute gain factors. This determination is made by evaluating the separation between the photon peak positions, including the 0-photon (noise) peak. The position of these peaks is derived by fitting a multi-Gaussian function to the single-photon spectral distributions collected for each memory cell within each pixel, as shown in Figure 7A. This process involves performing a total of 352 million fits for AGIPD1M. Figure 7B shows the absolute gain factors for high gain as a function of the memory cell index for one of the FEM modules installed in AGIPD1M at SPB/SFX. Additionally, an example of an absolute gain map for high gain, averaged across all memory cells in a pixel, is presented in Figure 7C.

The absolute gain values, in conjunction with the gain ratios obtained from the dynamic range scans described in Section 3.4, provide a comprehensive calibration of the detector gain.

3.6 Bad pixels determination

The accurate identification and classification of all detector channels, in this case pixels, is essential for any scientific analysis. This process involves recognizing and documenting information regarding pixels that do not perform optimally. The primary objective of this identification is to facilitate the exclusion of problematic pixels from the analysis. This information is encapsulated within what is commonly referred to as a “bad pixel mask”.

Problematic pixels can exhibit various issues. Some may be non-functional, a condition commonly referred to as “dead pixels.” Others may not respond to X-ray stimuli, exhibit excessive noise, or display parameter values outside the expected range.

The identification of these problematic pixels is based on the evaluation of calibration constants derived from the calibration data mentioned above and provided to users with the corrected data.

1. As an initial step, detection of abnormal or “dead” pixels involves evaluation of the offset and noise values derived from dark data. A pixel is flagged as “bad” if one of these values exceeds or falls below predefined thresholds. Two threshold settings are available: one establishing absolute limits for noise and offset values and the other determined by the standard deviation calculated from mean offset and noise values across all memory cells and pixels. Typically, pixels with values exceeding ± 5 standard deviations from the mean are classified as “bad pixels.” These threshold settings are adjustable, allowing for more leniency or stringency depending on the specific requirements of the scientific analysis. For example, when dealing with sparse XPCS data, a more conservative definition

- of “bad pixels” may be necessary to minimize the detection of “false positive” single photons.
- Identifying pixels that might exhibit normal behavior in high gain but struggle to transition to medium or low gain settings is the next step. To recognize these problematic pixels, internal charge injection data is used. Pixels are classified as ‘bad’ if their gain ratios exceed or fall below predefined thresholds, which are determined in the same manner as for offset and noise values derived from dark data.
 - The final refinement of problematic pixel detection is achieved using data obtained with X-rays. These data allow for the identification of pixels that are insensitive to X-rays and pixels with abnormal absolute gain values that exceed predefined thresholds, which are defined in the same manner as previously described for other calibration constants.

4 Operational aspects—performance and reliability of AGIPD detectors

4.1 Scientific outcome

The AGIPD detector is used in different kinds of experiments. The following sections give an overview of the main experimental techniques used with AGIPD and highlight some of the published results.

4.1.1 Serial femtosecond crystallography

High peak brilliance, combined with the MHz repetition rate of X-ray pulses, makes the European XFEL exceptionally attractive for serial femtosecond crystallography (SFX), a method routinely used to study both the structure and dynamics of proteins at room temperature [34]. SFX employs the “diffract before destruction” principle. This means that, due to the short duration of the incident X-ray pulses, a diffraction signal is collected before the sample is obliterated by the X-rays. The SPB/SFX instrument at the European XFEL is purpose-built to facilitate these SFX-type measurements. The AGIPD1M detector is an integral component of the in-vacuum interaction region at SPB/SFX and among its crucial attributes for SFX are the high dynamic range and MHz acquisition rate. In a typical experiment, sub-micron crystals suspended in a low-viscosity buffer medium are injected into the X-ray beam in the form of a liquid jet. Resolving a protein structure requires around ten thousand randomly oriented diffraction patterns. Taking into account an average crystal hit rate of approximately 1%, this requires the collection of approximately a million images in the shortest possible time frame. The high repetition rate of X-ray pulses at the European XFEL allows the recording of diffraction data more than an order of magnitude faster than previously achievable [35]. The most common configurations are 3,510 images per second at a detector acquisition rate of 1.13 MHz or approximately 2000 images per second at a detector acquisition rate of 0.56 MHz. The reason for collecting fewer images at 0.56 MHz compared to 1.13 MHz is due to the constraints on the number of pulses available from the XFEL accelerator at this repetition rate. Specifically, the RF window is limited to a maximum duration of 600 μ s, and these pulses are distributed among three beamlines simultaneously. An example of an experimental setup used for SFX is provided in Figure 8A, while

Figure 8B presents a single-crystal diffraction image captured with the AGIPD1M [37]. Multiple experiments have convincingly demonstrated that data recorded with the AGIPD1M and calibrated with the European XFEL calibration routine yield high-quality results [36, 41, 42].

4.1.2 Single particle imaging

Single particle imaging (SPI) is a technique oriented towards resolving the structure of individual particles or molecules based on a multitude of interactions between X-ray pulses and the sample. The underlying principle of this experiment is illustrated in Figure 8C. The SPI technique uses high repetition-rate X-ray pulses provided by the European XFEL to capture 2D diffraction patterns from a repeatable sample in random orientations. It relies on statistical sampling to gather diffraction patterns, which are subsequently analytically combined to reconstruct the 3D electron density of the sample. In the case of SPI, low noise and the capability to resolve a single photon are among the key requirements for the detector. In practice, tens of thousands of good quality patterns are necessary to complete the measurement. During SPI experiments, the AGIPD1M typically operates at its highest frame rate. For investigations involving weakly scattering samples, AGIPD offers the flexibility to adjust the gain of the Correlated Double Sampling (CDS) stage [15] of the pixels. This adjustment enhances single-photon resolution at the cost of reduced dynamic range.

The AGIPD1M detector has demonstrated performance, allowing the successful execution of single particle imaging experiments at SPB/SFX. Detailed results from a study on gold nanoparticles are presented in [38], while the findings of studies involving Iridium Chloride (IrCl_3) and Mimivirus are presented in [39]. Examples of scattering patterns from these experiments are shown in Figure 8D.

4.1.3 X-ray photon correlation spectroscopy

X-ray Photon Correlation Spectroscopy (XPCS) is a technique used to investigate the dynamics and kinetics in hard and soft condensed matter samples. At synchrotrons, it conventionally enables the probing of dynamics at timescales ranging from milliseconds to hours. However, the MHz repetition rate of the European XFEL and the AGIPD detector permits exploration of structural dynamics at much shorter timescales, in the sub-microsecond range. This is particularly relevant since sub-microsecond and microsecond timescales are natural for the diffusion of biophysical systems and nanoparticles in their aqueous environments [40]. An illustration of an XPCS experiment and the average scattering intensity obtained from the analysis of the data collected with the AGIPD1M detector are shown in Figures 8E, F respectively.

XPCS exploits the coherence of the XFEL beam, by recording speckle patterns of typically non-crystalline samples, which encode the spatial arrangement of the scatterer. The dynamics can be obtained from a series of such speckle patterns by calculating the temporal auto-correlation function. The AGIPD was designed primarily for experiments such as imaging or femtosecond crystallography. However, it can also be utilized for X-ray Photon Correlation Spectroscopy (XPCS) under certain conditions. One main challenge is to detect the speckle pattern with sufficient spatial

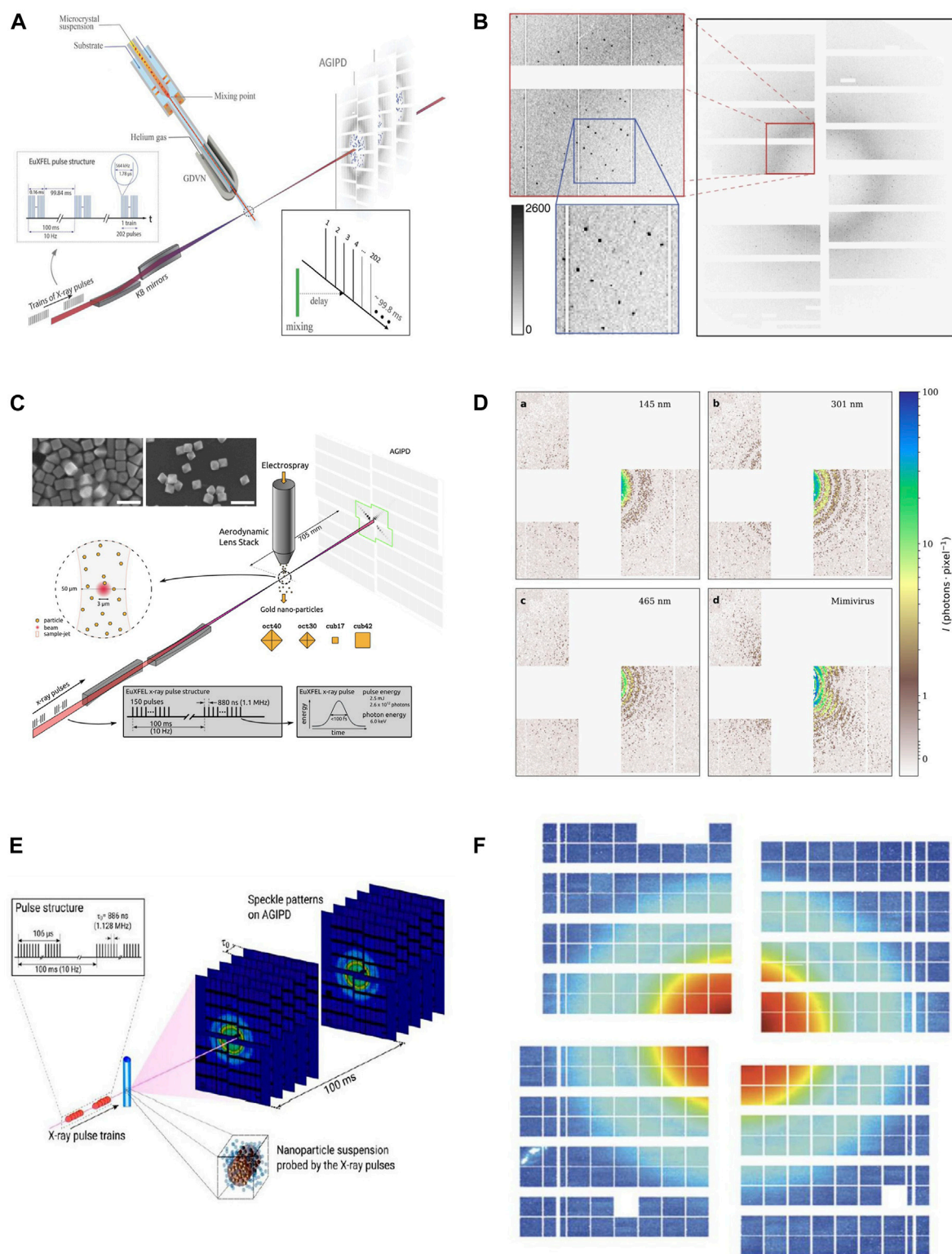


FIGURE 8

(A) Experimental setup for SFX measurements (originally published in [36]). (B) Example serial crystallography data taken using AGIPD1M during early user experiments at SPB/SFX. Note the well-defined Bragg peaks that span a large fraction of the detector dynamic range (published in [37]). (C) Experimental setup for SPI experiments (originally published in [38]). (D) Examples of scattering patterns from IrCl_3 and Mimivirus recorded by AGIPD1M (originally published in [39]). (E) Experimental setup for XPCS experiments published in [40]). (F) Example of the mean scattering intensity of an aqueous silica nanoparticle solution recorded by AGIPD1M (originally published in [5]).

resolution. The speckle visibility (or speckle contrast in XPCS experiments) is diminished if the pixel size of the detector is large compared to the speckle size, with an optimum Signal-to-Noise ratio if both quantities are equal. In user experiments probing nanometer length scales in small-angle (SAXS) configuration this criterion can readily be achieved by sufficient focusing of the X-ray beam. Several user publications have used this configuration successfully and it is offered as standard configuration at the MID instrument [5]. Early exploratory studies [5, 43–45] were built on a previous experiment at SPB/SFX [40], and have since been followed by more application-oriented experiments on protein aggregation and diffusion [46], functional nanoparticle self-assembly, or temperature-jump swelling-deswelling kinetics of PNIPAM nanogels [47].

Extending these measurements to atomic length scales in wide-angle (WAXS) geometry has proven to be more challenging. The finite longitudinal coherence length leads to a further reduction of speckle contrast in addition to the one due to the large detector pixels. Additionally, the scattering intensity at larger angles is typically much reduced, making it comparable to the inherent noise level of the detector. To accommodate the sources of noise in AGIPD that can negatively impact XPCS data analysis even at higher intensities, a special data treatment had to be developed, as reported later in 4.3. This has proven to work well for most of the XPCS experiments in SAXS geometry; however, for WAXS experiments the noise characteristics at lower intensities as well as the large pixel size at smaller speckle contrasts remain challenging.

4.2 Performance challenges and continuous improvements

The data obtained with all AGIPD detectors installed at the European XFEL have been used to generate a considerable amount of scientific output. More than 20 publications have been produced on the basis of the data collected with these detectors. The experience with the commissioning, calibration and operation of the AGIPD has allowed us to identify several shortcomings that affected the quality of the data. The European XFEL and AGIPD Consortium have made a significant effort to understand the behavior of the detector and to successfully improve its performance. This involved hardware enhancements, optimization of detector configurations, and software data processing that are described in the following sections.

4.2.1 Baseline shift dependence of incoming X-ray intensities

It was observed that certain pixels in the images exhibit negative intensity values, and this effect becomes more pronounced as the overall intensity in this FEM module increases. To quantify this effect, a ‘mask’ was installed in front of the AGIPD1M detector at SPB/SFX, shadowing some parts of all FEMs, and an intensity scan was performed. An example image is shown in Figure 9A. This data was further analyzed and yielded the following observation: the common baseline signal of a module shifts to lower values, and the shift is linearly proportional to the incoming X-ray intensity causing a $\approx 20\%$ decrease in the level of measured photon signal.

The issue is caused by the depletion of the sensor, which then also acts as a capacitance, and the resistor connecting it to the high-voltage (HV) input, in this case exacerbated by the parallel capacitances of a Pi-filter. The purpose of a resistor at this point is to limit the current, once the HV has to be turned off. This current limitation also affects sensor signals: Once the sensor absorbs X-rays, it cannot be re-charged immediately due to the finite RC – time constant, and in turn the high voltage drops by a finite amount, registered as a shifted baseline.

These resistors were exchanged on all AGIPD1M vacuum boards, leading to a reduction in the baseline shift to $<2\%$ of the deposited signal. The baseline shift before and after hardware modification is plotted as a function of the integrated intensity in Figure 9B. This effect is most pronounced for pixels in high gain as the aforementioned resistors are in series with the input impedance of the preamplifier. Since the latter is reduced by switching to MG or LG, the RC – time constant is further lowered, and consequently the baseline shift is further suppressed.

4.2.2 Gain continuity close to the transition region

In scattering experiments, the presence of randomly fluctuating pixels exhibiting unusually high signal values, commonly referred to as “snowy pixels”, was observed. These snowy pixels tend to manifest when incident X-ray intensities approach the transition region between high and medium gain stages. Although the gain setting of these pixels is classified as medium gain based on their analog gain signals, a closer examination reveals that the signal levels of these pixels do not align with expected intensities. They appear to exhibit values that are inadequately low for high gain and excessively high in medium gain settings. An example of snowy pixels is shown in Figure 10A. These pixels, depending on context, may look like valid events, e.g., in a Bragg peak, and actually disturb analysis.

The appearance of “snowy pixels” within the detector can be attributed to a late gain switching effect. In the AGIPD system, a voltage representing the collected charge is written to the analog memory at the end of the integration time. However, when gain switching occurs too close to the end of this period, the CDS stage, which has the lowest bandwidth in the ASIC, does not have enough time to settle, resulting in an excessively high stored voltage. Late gain switching can be influenced by several factors. One is the relatively short integration time, despite the XFEL pulses being extremely short, i.e., tens of fs. The charge collection time for the sensor is several tens of nanoseconds, and the profile of this charge collection approximates a falling exponential curve. Consequently, the tail of the signal charge may overlap with the end of the integration time, potentially causing incomplete stabilization of the signal before integration is completed, as illustrated in Figure 10C. Furthermore, noise can be a contributing factor, as random noise can trigger gain switching for signals near the threshold. Additionally, the leakage current becomes relevant for signal levels close to the threshold, which eventually leads to gain switching.

To mitigate the issue of “snowy pixels,” two potential solutions have been identified that do not require a re-design of the ASIC. Both solutions aim to optimize the detector configuration to enhance the quality of output data for specific experimental methods. In situations where ensuring single-photon sensitivity is crucial while maintaining a high dynamic range, and considering

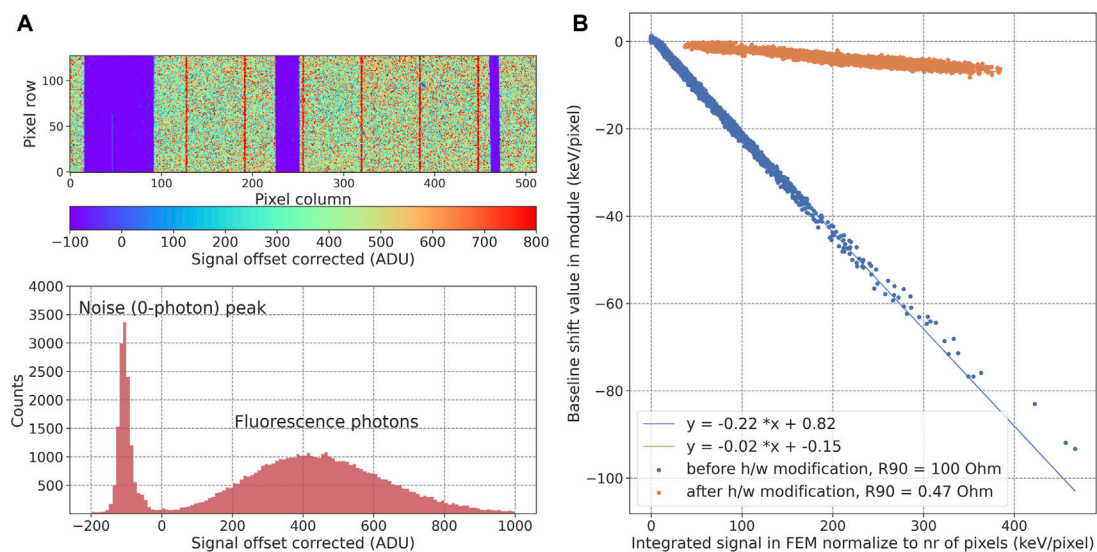


FIGURE 9

(A) A single image illuminated with high-intensity Cu fluorescence photons after offset subtraction. The three areas in purple color represent the part of the module covered with Ta stripes. The histogram illustrates the shift of the noise peak toward negative values resulting from the baseline shift. (B) Baseline shift value as a function of the integrated signal in FEM, normalized to the number of pixels, before modification of the vacuum board hardware (depicted in blue) and after modification of the hardware (depicted in orange). Linear functions were fitted to each dataset to quantify the baseline shift effect before and after hardware modification, resulting in an order of magnitude reduction of the effect.

that most experiments at SPB/SFX and MID typically involve pulses at repetition rates below 4.5 MHz (i.e., 2.25 MHz or lower), extending the integration time is an option to consider. In this mode, the pulse is recorded at the beginning of the integration window and the window itself is prolonged. Operation of the detector with an extended integration time results in a substantial reduction in snowy pixels, reducing their occurrence from around 10% to below 0.01% for transition region intensities. This improvement was verified through a dedicated measurement at the SPB/SFX instrument, specifically with water jet scattering and two integration times of 120 ns and 200 ns. A comparison of the results is shown in Figures 10A, B, which highlights the significant improvement in data quality when integrating over 200 ns. On the other hand, if single-photon sensitivity is not required, the AGIPD can be operated in a fixed medium gain mode, where the detector gain is set to a predefined value (i.e., medium gain), and dynamic gain switching is disabled. This approach prevents transitions between gain stages. However, it is not suitable for low-intensity data that require single-photon sensitivity, as it leads to increased noise from approximately 1.3 keV to more than 40 keV.

4.3 Dealing with very sparse data

In experiments with low photon intensities, particularly when dealing with sparse data (i.e., below 0.1 photon per pixel per pulse), a special operation mode is implemented to minimize noise and increase the single-photon sensitivity. This mode effectively reduces noise levels from 1.3 keV to approximately 0.9 keV, resulting in a narrower dynamic range limited to only a few tens of photons. This operation mode has become the standard mode of choice of SPI 4.1.2, XPCS 4.1.3, or fluorescence correlation imaging [48].

In addition to the standard correction using calibration constants, specific common-mode corrections are additionally applied. Common-mode noise denotes a type of signal variation that affects groups of read-out channels in a synchronized manner. It can arise from various sources, including common electromagnetic interference or voltage fluctuations. Common-mode noise not only contributes to overall noise levels, but can also potentially introduce artificial hit patterns. The extent of its impact varies depending on specific detector components, such as ASICs, and the surrounding environmental conditions. The precise spectrum of common-mode noise is typically not known in advance. In certain cases, the contribution of common-mode noise can be estimated either on a per-image basis or for groups of read-out channels. This process involves subtracting the offset from the raw signals, excluding the channels with the actual signal, and computing the average signal observed on these channels. This calculated average signal provides a reasonable estimate of the contribution of common-mode noise, which can then be subtracted from the signals.

For the AGIPD1M detectors, two sources of common-mode noise have been identified: one at the ASIC level and another along the memory cell rows (due to the cells being grouped into 11×32 matrices, as described in Section 1.2). Correcting for common-mode noise significantly improves data quality, especially when dealing with sparse intensities, as demonstrated in Figure 11. However, it is essential to note that common-mode corrections may not fully address all issues related to offset instabilities. In specific instances, it has been observed that a small fraction of pixels (less than 0.01%) experience “jumps” in offsets within blocks of 32 storage cells, but for several consecutive trains. While the magnitude of these “jumps” is typically on the order of a single photon, synchronous jumping of several storage cells in one pixel can introduce significant artifacts in correlation analyses, such as XPCS where different storage cells are correlated with each other. To address these issues, a tailored correction

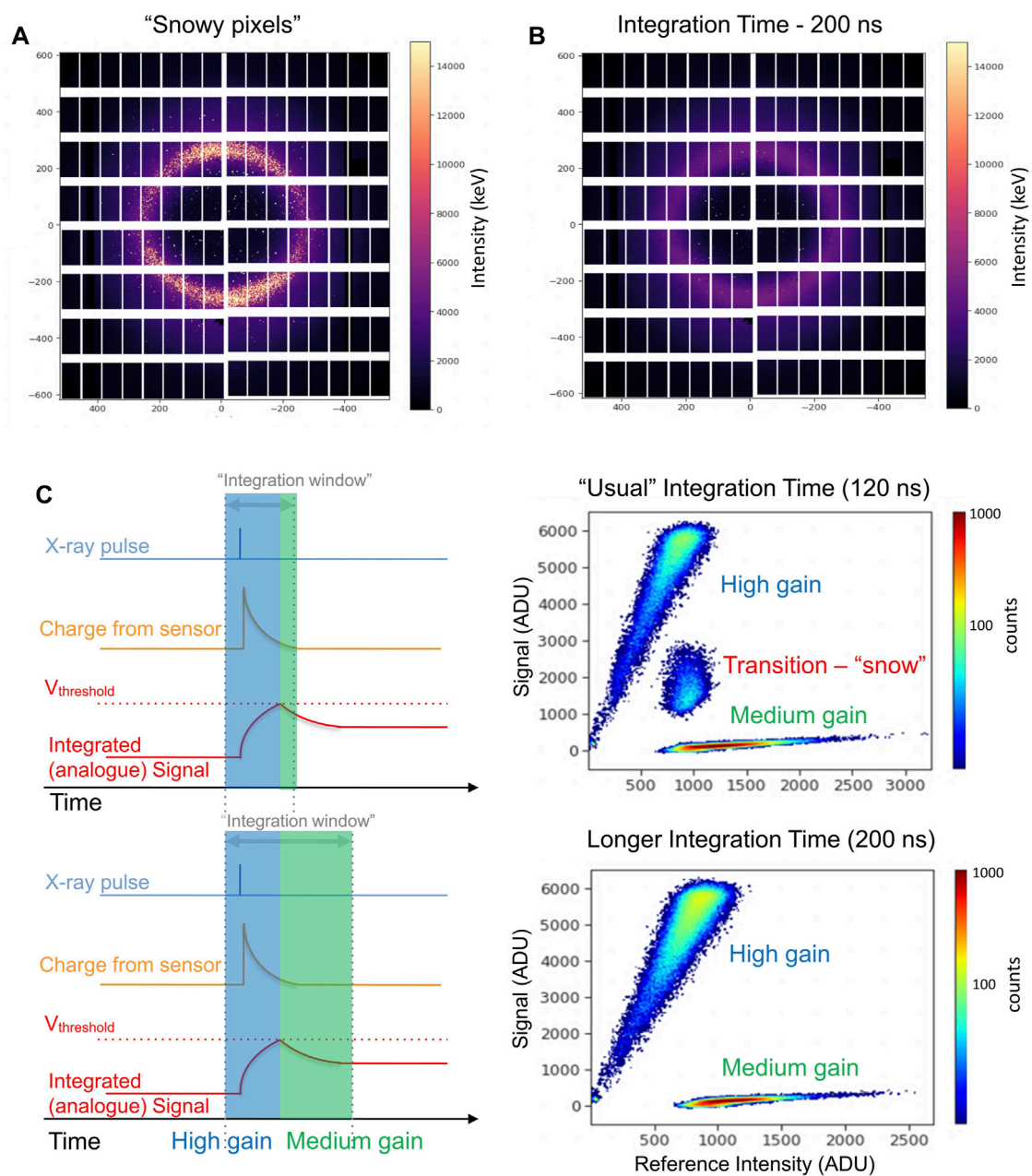


FIGURE 10

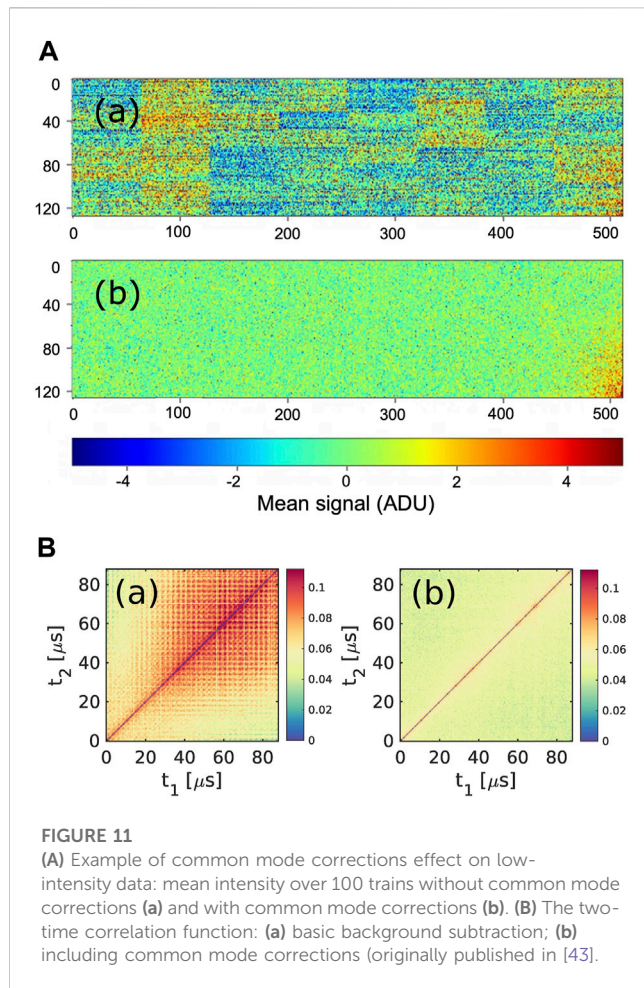
(A) Image of a water jet ring collected with the AGIPD1M detector at SPB/SFX with a 120 ns integration time. Snowy pixels are identified as those with unexpected high-intensity values. (B) Image of a water jet ring taken with the AGIPD1M detector using a 200 ns integration time. The presence of snowy pixels is to a large extent suppressed. (C) Illustration of the late gain switching mechanism (left plot) and intensity scans for a single pixel (right plot), including the transition region between high and medium gain visible for the integration time of 120 ns.

approach becomes necessary as part of data analysis. As an example, the corrections to mitigate the 'jumping' pixel issue were developed by MID and the XPCS user community and were successfully used during data analysis [5, 43]. This method involves calculating not only the temporal autocorrelation function within one train (along the storage cell dimension) but also cross-correlation terms involving data from different trains. Since the measured speckle pattern should be uncorrelated from train to train, the resulting correlation matrices should primarily contain terms originating from jumping-pixel contributions. This approach has been shown to mitigate this

artifact down to low intensities of 10^{-2} – 10^{-1} photons per pixel per pulse, but was not sufficient to correct XPCS data at even lower intensities.

4.4 Performance degradation due to radiation damage

The AGIPD1M detectors at European XFEL are at risk of radiation damage due to the intense X-ray beams used in the

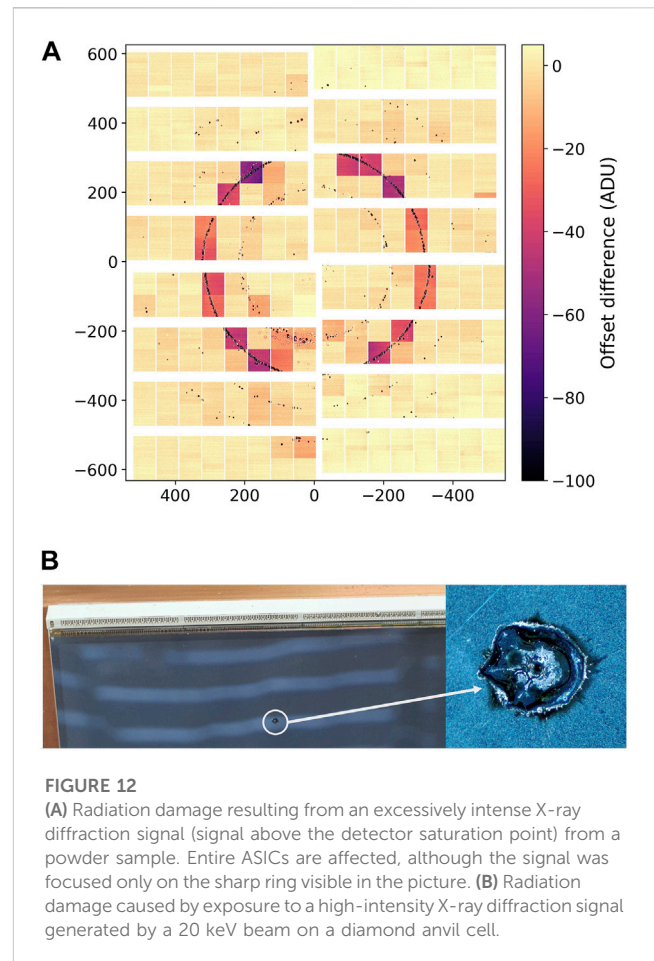


experiments. This damage can result in a decrease in detector performance over time, leading to issues such as changes in offset, increased noise, or reduced capacity to detect photons. Notably, the impact of radiation damage is more pronounced in the detector's electronic components, particularly the ASICs, than in its sensor.

Despite safety measures, accidents can occur, such as the exposure of the detector to the direct X-ray beam or to exceptionally intense X-ray diffraction signal during experiments. In the case of the AGIPD1M detector used in serial femtosecond crystallography, such intense diffraction signals that exceed the detector's radiation damage threshold can occur when ice, which can form at the nozzle of the sample injection system, is hit by the X-ray beam. As mentioned in Section 2.4, this can result in permanent damage to specific regions of the detector.

Furthermore, exposure of the AGIPD1M detector to high-energy X-rays above 20 keV, for which the 500 μm silicon sensor is almost transparent, can lead to ASIC damage occurring more frequently than when the detector is used in its optimal energy range of 8–12 keV.

In instances of radiation damage, it becomes necessary to recalibrate the affected portions of the AGIPD1M detector. When damaged pixels are critical for scientific analysis, it may be necessary to replace the entire FEM for optimal performance restoration. Examples of radiation damage are shown in Figure 12.



4.5 Data volume

The AGIPD detectors at the European XFEL are among the most frequently used MHz detectors at the facility, contributing to more than 80% of the total raw data. This significant contribution is mainly attributed to the detectors' MHz operation. Furthermore, the unique design of the AGIPD detectors, which incorporates analog gain information alongside the images, significantly increases the data volume. The gain encoding process, due to the reasons discussed in Section 3.3, is performed "offline" and not in the detector back-end electronics, involving FPGAs as originally planned.

The data recorded at the European XFEL continues to accumulate rapidly and is approaching the 100 petabyte mark, necessitating a robust data volume management and data reduction strategy. To address this challenge of data volume, the European XFEL has devised a comprehensive strategy [49]. This approach encompasses early data reduction planning within the Data Management Plan (DMP), the development of operation-specific and technique-specific data reduction methods, enabling real-time data reduction for new experiments, and facilitating retroactive data reduction for previously collected data in collaboration with users. It also underscores the importance of integrating data reduction capabilities closer to the detector head, including within the ASIC or detector back-end electronics, as a critical consideration for the design of future detectors.

5 Insights for next-generation detector development

Over years of operation, the AGIPD detectors have consistently demonstrated their ability to deliver high-quality scientific results, as discussed in [Section 4.1](#). Concurrently, data collection and dedicated studies have brought to light challenges that impact detector performance and scientific measurements. These challenges, although demanding, have provided invaluable experience and insights. This knowledge will be instrumental in shaping future detector projects for the European XFEL, offering valuable lessons that extend beyond the facility itself.

5.1 Effective scheduling and collaboration

Accurate project scheduling in research and development can present significant challenges. In 2006, the European XFEL issued a call for proposals to develop a MHz 2D Imaging Detector, leading to three projects, one of which was AGIPD. The first AGIPD1M detector became operational in mid-2017. Comparable timelines were observed in other detector projects at the European XFEL, including LPD and DSSC. Therefore, prudent planning that involves contingencies and realistic expectations regarding the development timeline, potentially spanning around a decade, is advisable.

Early collaboration with instrument experts, detector end users, and system integration, control, and data acquisition teams is of paramount importance. This collaboration helps ensure harmonious alignment of scientific and technical requirements, covering aspects such as infrastructure and control systems, between the detector developers and the ultimate users.

5.2 Detector system integration and operation

During the integration of detectors, such as AGIPD1M, a complex process unfolds, with a strong emphasis on convenience and ease of operation. Considerations in designing the detector's mechanical interfaces to the rest of the instrument, as well as detector internal/external components, should ensure accessibility to cables, connectors, and internal detector components, simplifying replacement and repair processes. In particular, components exposed to X-rays are susceptible to damage, underscoring the importance of spare parts availability and straightforward repair procedures.

To streamline integration, efforts to simplify the power supply system and optimize the power consumption of the detector can significantly reduce complexity. This includes minimizing the number and thickness of power cables, enhancing the flexibility of the system, and mitigating the risk of potential damage during detector movement. Although this aspect has been partially addressed in the design of new-generation AGIPD detectors, more work can be done.

Efforts to reduce power consumption can also positively affect detector cooling requirements, as discussed in [Section 2.3](#). Achieving optimal temperatures for detector components within a vacuum environment remains a challenge. Future detector developments

should consider careful optimization of cooling system design, electronics design, and power dissipation to ensure desirable outcomes.

Furthermore, optimizing the design of electronic boards is essential to maximize the active area of the detector that relates to the entire system, particularly in instruments with limited space.

It should be noted that the adoption of a modular system with minimal critical components for the entire detector assembly can prove beneficial. If such functionality is necessary, the components should ideally be located outside the detector assembly vessel. This approach enables efficient component replacement in a short time frame, typically minutes to hours, without extensive effort. In the case of AGIPD1M detectors, composed of two electronically independent halves, the control of each half is managed through the master FPGA. However, this concept presents some disadvantages, such as the loss of half of the detector's functionality in the case of a board failure. The design of the next-generation of AGIPD detectors, which addresses these concerns through redesigned back electronics, is promising.

Another vital but often overlooked aspect is the detector safety system. The interlock system developed for the AGIPD1M detectors effectively protects them against unexpected failures (e.g., vacuum or cooling failure) and potential human errors. Over years of operation at the European XFEL, this system has prevented more than ten accidents, including damage from water condensation on electronics due to vacuum failure or component overheating due to cooling issues. However, as mentioned in [Section 2.4](#), the AGIPD1M interlock system relies on external PLC-based interlocks that monitor temperature and pressure sensors within the detector vessel and act accordingly. This increases the infrastructure of the detector, requiring a more self-protecting system. Therefore, ensuring detector safety must begin at the design level, implementing a simple, reliable and ideally self-protecting system to minimize the risk of detector damage.

However, there are cases where protection is not always feasible, as exemplified by radiation damage, discussed in [Section 4.4](#). Therefore, the mitigation of radiation damage in the ASIC is a critical aspect of future detector development for high-intensity X-ray sources.

Finally, the significance of data quality cannot be overstated in determining the performance of the detector. Achieving high-quality data is highly dependent on the precise calibration and characterization of the detectors. This process is not a one-time activity, but must be conducted regularly, even during user experiments, to address incidents that may impact data quality. The utilization of internal calibration sources, such as the Pulsed Capacitor (PC) and Current Source (CS), facilitates routine dynamic range scans. Nevertheless, it is worth noting that both sources possess limitations, as discussed in [Section 3.4](#). Therefore, prioritizing the development of a reliable, “*in-situ*” calibration mechanism, particularly at the ASIC level, becomes imperative for the progression of future detector generations.

5.3 Testing the system under real conditions

Comprehensive testing of a detector system can only be achieved under actual operational conditions. This necessitates evaluating the

detector at its full size, within the final installation infrastructure, and, in our case, in the presence of the European XFEL beam. The absence of the beam during the detector's development phase left certain features unidentified and unaddressed until the detector was fully assembled and installed. Consequently, it is advisable for any new detector project to test its prototype under conditions that closely resemble the final operational environment, with a focus on considering scalability from the project's inception.

When delving into particulars of the chip design, several key observations were brought to light that could be beneficial for forthcoming development projects.

1. Data quality is influenced by the utilization of an analog memory matrix for storing image data on pixel before it can be transferred from the detector head. Challenges are posed by variations in offset and gain, which depend on the index (position) of the memory cell in the pixel and the leakage into neighboring memory cells that leads to the issue of gain encoding as described in [Section 3.3](#). To enhance data quality, the requirement of generating calibration constants for each memory cell in the pixel should be reconsidered. In the next-generation of detectors, it is proposed that exploration of options such as digital storage cells or alternatives to existing storage cells be undertaken, with the potential to enable real-time data transfer from Front-End Modules (FEMs).
2. The achievement of a high dynamic range through an adaptive gain mechanism is accompanied by challenges related to linearity and data quality, particularly in the transition region between gain settings. Comparable issues have been noted in other detectors that employ adaptive gain mechanisms, such as JUNGFRU [50]. To address these concerns, the exploration of alternative implementations and solutions aimed at realizing a high dynamic range or a more robust design of the adaptive gain switching mechanism, preventing the aforementioned "late gain switching" and avoiding the transition region between different gain settings, is deemed to be imperative. This is of paramount importance, as the dynamic range remains a critical parameter to be realized in the next-generation of detectors.

6 Conclusion and outlook

The AGIPD detector systems deployed at the European XFEL instruments have demonstrated their reliability and made a significant contribution to the generation of valuable scientific data, as evidenced by numerous scientific publications. This underscores their ability to support experiments with specific demands, provided that their characteristics and limitations are well understood.

A commitment to continuous improvement and development in the operation of the AGIPD detectors has resulted in substantial benefits, leveraging the expertise of detector specialists and beamline instrument scientists with a range of backgrounds. Through hardware optimization, enhanced detector characterization, and advanced data processing techniques, we have effectively accommodated diverse experiment requests and acquired reliable scientific data. Regular maintenance, updates, and innovative data

processing have collectively increased the quality of AGIPD-generated data.

However, it is essential to acknowledge that AGIPD optimization remains an ongoing process. Collaborative efforts that draw upon the expertise of individuals with diverse backgrounds remain essential components in optimizing detector performance and addressing observed issues.

Looking ahead, the second generation of AGIPD detectors, slated for installation in 2024 at the HED instrument (one-megapixel system) and the SPB/SFX instrument (four-megapixel system), represents a significant step forward. A prototype of the new detectors is already in use at HED, with the first results from user experiments already published. Although certain challenges discussed in this publication have been addressed in the second-generation AGIPD, not all have been resolved. As emphasized in this paper, the redesign of key detector components, such as the ASIC, is essential to effectively address these challenges.

The operational insights gained from AGIPD detectors will play a valuable role in shaping the future of the detector development program at the European XFEL. The integration of the initial detector generation into the European XFEL instruments was a formidable undertaking, revealing infrastructure challenges for integration and operation. To address these challenges, we emphasize the need for more compact and efficient power and cooling designs with standardized interfaces.

Our experiences underscore the importance of ease of operation and reliability, accessible detector components for maintenance and replacement, and the critical role of hardware interlocks, including the consideration of self-protecting detectors. Furthermore, the management of the substantial volume of generated data requires early design-level data reduction strategies.

Data quality remains the paramount measure of detector performance, highlighting the necessity of evaluating methods to achieve high dynamic range, exploring alternative solutions for storing data at pixel level instead of relying on analog memory cells, and designing calibration-friendly systems with reliable *in-situ* calibration sources.

Although the scientific requirements for the initial years of instrument operation are met by current detectors to the largest extent, plans for the next-generation of detectors are well underway. The demand for smaller pixel sizes, extended dynamic range, and the ability to operate at MHz rates while recording even more pulses per second are the primary driving factors behind the search for innovative technological solutions. Transitioning from 130 to 65 or 28 nm CMOS technology and exploring three-dimensional integrated electronic circuits may offer promising avenues for addressing these challenges. Additionally, the development of edgeless sensors holds potential for minimizing inactive detector areas. Furthermore, the imperative of maintaining high quantum efficiency, even at energies exceeding 20 keV, motivates the exploration of materials with atomic numbers higher than those of silicon.

With the prospect of developing novel detectors for the post-2030 European XFEL operation, the insights and experiences gained from the AGIPD detector operation, in conjunction with the ever-evolving scientific requirements, constitute indispensable foundations for shaping the future of detector technology for the European XFEL.

Author contributions

JS-D: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing—original draft. VR: Investigation, Methodology, Validation, Writing—review and editing. AK: Investigation, Methodology, Validation, Writing—review and editing. TL: Investigation, Methodology, Validation, Writing—review and editing. UT: Conceptualization, Investigation, Methodology, Writing—review and editing, Validation. KA: Software, Visualization, Writing—review and editing. OM: Validation, Writing—review and editing. JM: Methodology, Validation, Investigation, Writing—review and editing. AP: Software, Writing—review and editing. NR: Investigation, Validation, Writing—review and editing. RS: Validation, Writing—review and editing. MS: Investigation, Validation, Writing—review and editing. GA: Conceptualization, Validation, Writing—review and editing. UB: Investigation, Validation, Writing—review and editing. LL: Conceptualization, Validation, Writing—review and editing. AM: Software, Writing—review and editing. TP: Validation, Writing—review and editing. PS: Software, Writing—review and editing. SS: Investigation, Validation, Writing—review and editing. RB: Resources, Writing—review and editing. AM: Resources, Writing—review and editing. LG: Resources, Writing—review and editing. SH: Methodology, Resources, Software, Writing—review and editing. PG: Resources, Writing—review and editing. KW: Resources, Writing—review and editing. HG: Project administration, Resources, Writing—review and editing. MT: Resources, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, et al. A mhz-repetition-rate hard x-ray free-electron laser driven by a superconducting linear accelerator. *Nat Photon* (2020) 14:391–7. doi:10.1038/s41566-020-0607-z
- Tschentscher T. Investigating ultrafast structural dynamics using high repetition rate x-ray free electron radiation at European XFEL. *The Eur Phys J Plus* (2023) 138:274. doi:10.1140/epjp/s13360-023-03809-5
- Mancuso AP, Aquila A, Batchelor L, Bean RJ, Bielecki J, Borchers G, et al. The single particles, clusters and biomolecules and serial femtosecond crystallography instrument of the European XFEL: initial installation. *J Synchrotron Radiat* (2019) 26:660–76. doi:10.1107/S1600577519003308
- Galler A, Gawelda W, Biednov M, Bomer C, Britz A, Brockhauser S, et al. Scientific instrument Femtosecond X-ray Experiments (FXE): instrumentation and baseline experimental capabilities. *J Synchrotron Radiat* (2019) 26:1432–47. doi:10.1107/S1600577519006647
- Madsen A, Hallmann J, Ansaldi G, Roth T, Lu W, Kim C, et al. Materials imaging and dynamics (MID) instrument at the European X-ray free-electron laser facility. *J Synchrotron Radiat* (2021) 28:637–49. doi:10.1107/S1600577521001302
- Zastrau U, Appel K, Baetz C, Baehr O, Batchelor L, Berghäuser A, et al. The high energy density scientific instrument at the European XFEL. *J Synchrotron Radiat* (2021) 28:1393–416. doi:10.1107/S1600577521007335
- Tschentscher T, Bressler C, Grünert J, Madsen A, Mancuso AP, Meyer M, et al. Photon beam transport and scientific instruments at the European XFEL. *Appl Sci* (2017) 7:592. doi:10.3390/app7060592
- Mazza T, Baumann TM, Boll R, De Fanis A, Grychtol P, Ilchen M, et al. The beam transport system for the Small Quantum Systems instrument at the European XFEL:

Acknowledgments

The authors thank the AGIPD consortium for developing the AGIPD detectors for the European XFEL. We acknowledge European XFEL in Schenefeld, Germany, for provision of X-ray free-electron laser beamtime at Scientific Instrument SPB/SFX (Single Particles, Clusters, and Biomolecules and Serial Femtosecond Crystallography) and Scientific Instrument MID (Materials Imaging and Dynamics) and would like to thank the staff for their assistance. We also acknowledge support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF. Furthermore, the authors express their gratitude to the Data Department groups, including Controls, IT and Data Management, and Electronics and Electrical Engineering, for their significant contributions to detector integration and Data Analysis for providing the software tools for detector data analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- optical layout and first commissioning results. *J Synchrotron Radiat* (2023) 30:457–67. doi:10.1107/S1600577522012085
- Graafsma H. Requirements for and development of 2 dimensional x-ray detectors for the European X-ray free electron laser in Hamburg. *J Instrumentation* (2009) 4: P12011. doi:10.1088/1748-0221/4/12/p12011
- Veale MC, Adkin P, Booker P, Coughlan J, French MJ, Hart M, et al. Characterisation of the high dynamic range large pixel detector (lpd) and its use at x-ray free electron laser sources. *J Instrumentation* (2017) 12:P12003. doi:10.1088/1748-0221/12/12/P12003
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Radiat* (2019) 26:74–82. doi:10.1107/S1600577518016077
- Porro M, Andricek L, Aschauer S, Castoldi A, Donato M, Engelke J, et al. The MiniSDD-based 1-mpixel camera of the DSSC project for the European XFEL. *IEEE Trans Nucl Sci* (2021) 68:1334–50. doi:10.1109/TNS.2021.3076602
- Shi X, Dinapoli R, Henrich B, Mozzanica A, Schmitt B, Mazzocco R, et al. Challenges in chip design for the agipd detector. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2010) 624:387–91. doi:10.1016/j.nima.2010.05.038
- Schwandt J, Fretwurst E, Klanner R, Kopsalis I, Zhang J. Design and first tests of a radiation-hard pixel sensor for the European x-ray free-electron laser. In: 2013 14th European Conference on Radiation and Its Effects on Components and Systems (RADECS); 23–27 September 2013; Oxford, UK (2013). p. 1–8. doi:10.1109/RADECS.2013.6937446
- Mezza D, Allahgholi A, Arino-Estrada G, Bianco L, Delfs A, Dinapoli R, et al. Characterization of agipd1.0: the full scale chip. *Nucl Instr Methods Phys Res Section A:*

Acc Spectrometers, Detectors Associated Equipment (2016) 838:39–46. doi:10.1016/j.nima.2016.09.007

16. Mezza D, Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, et al. Characterization of the agipd1.1 readout chip and improvements with respect to agipd1.0. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 945:162606. doi:10.1016/j.nima.2019.162606

17. Zhang J, Fretwurst E, Graafsma H, Klanner R, Kopsalis I, Schwandt J. Study of x-ray radiation damage in the agipd sensor for the european xfel. *J Instrumentation* (2014) 9:C05022. doi:10.1088/1748-0221/9/05/C05022

18. Husband RJ, Strohm C, Appel K, Ball OB, Briggs R, Buchen J, et al. A MHz X-ray diffraction set-up for dynamic compression experiments in the diamond anvil cell. *J Synchrotron Radiat* (2023) 30:671–85. doi:10.1107/S1600577523003910

19. Allahgholi A, Becker J, Delfs A, Dinapoli R, Göttlicher P, Graafsma H, et al. Megapixels @ megahertz – the agipd high-speed cameras for the european xfel. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 942:162324. doi:10.1016/j.nima.2019.06.065

20. Ball OB, Prescher C, Appel K, Baetz C, Baron MA, Briggs R, et al. Dynamic optical spectroscopy and pyrometry of static targets under optical and x-ray laser heating at the european xfel. *J Appl Phys* (2023) 134:055901. doi:10.1063/5.0142196

21. Siska E, Smith GA, Villa-Cortes S, Conway LJ, Husband RJ, Cleave JV, et al. Ultra-fast yttrium hydride chemistry at high pressures via non-equilibrium states induced by x-ray free electron laser (2023).

22. Wiener-D. *MPOD modules* (2023). Available from: <https://www.wiener-d.com/modules/mpod-modules/> (Accessed on September 28, 2023).

23. Flaschel N, Ariza D, Diez S, Gerboles M, Gregor IM, Jorda X, et al. Thermal and hydrodynamic studies for micro-channel cooling for large area silicon sensors in high energy physics experiments. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2017) 863:26–34. doi:10.1016/j.nima.2017.05.003

24. Coppola N, Tolkiehn J, Youngman C. Control using beckhoff distributed rail systems at the european xfel. In: Proceedings of ICALEPCS 2013; October 6–11, 2013; San Francisco, CA, USA (2013). p. TUPPC046.

25. Hauf S, Heisen B, Aplin S, Beg M, Bergemann M, Bondar V, et al. The Karabo distributed control system. *J Synchrotron Radiat* (2019) 26:1448–61. doi:10.1107/S1600577519006696

26. Rehlich K, Aghababayan A, Kay H, Petrosyan G, Petrosyan L, Petrosyan V, et al. *The new timing system for the european xfel* (2013).

27. Koranne S. *Hierarchical data format 5: HDF5*. Boston, MA: Springer US (2011). p. 191–200. doi:10.1007/978-1-4419-7719-9_10

28. Schmuck F, Haskin R. Gpfs: a shared-disk file system for large computing clusters. In: Proceedings of the FAST'02 Conference on File and Storage Technologies (Monterey, California, US: USENIX); January 28–30, 2002; Monterey, California, USA (2002). p. 231–44. Retrieved 2008-01-18.

29. European XFEL. *Metadata catalogue* (2023).

30. DESY DES. *Maxwell cluster* (2023).

31. European XFEL. *Calibration catalogue* (2023). Available from: <https://in.xfel.eu/calibration>.

32. Schmidt P, Ahmed K, Danilevski C, Hammer D, Rosca R, Kluyver T, et al. Turning european xfel raw data into user data. *Front Phys* (2023) 11. doi:10.3389/fphy.2023.1321524

33. Mezza D, Becker J, Carraresi L, Castoldi A, Dinapoli R, Goettlicher P, et al. Calibration methods for charge integrating detectors. *Nucl Instr Methods Phys Res*

Section A: Acc Spectrometers, Detectors Associated Equipment (2022) 1024:166078. doi:10.1016/j.nima.2021.166078

34. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, et al. Femtosecond x-ray protein nanocrystallography. *Nature* (2011) 470:73–7. doi:10.1038/nature09750

35. Kirkwood HJ, de Wijn R, Mills G, Letrun R, Kloos M, Vakili M, et al. A multi-million image serial femtosecond crystallography dataset collected at the european xfel. *Scientific Data* (2022) 9:161. doi:10.1038/s41597-022-01266-w

36. Pandey S, Calvey G, Katz AM, Malla TN, Koua FHM, Martin-Garcia JM, et al. Observation of substrate diffusion and ligand binding in enzyme crystals using high-repetition-rate mix-and-inject serial crystallography. *IUCrJ* (2021) 8:878–95. doi:10.1107/S2052252521008125

37. Wiedorn MO, Oberthür D, Bean R, Schubert R, Werner N, Abbey B, et al. Megahertz serial crystallography. *Nat Commun* (2018) 9:4025. doi:10.1038/s41467-018-06156-7

38. Ayyer K, Xavier PL, Bielecki J, Shen Z, Daurer BJ, Samanta AK, et al. 3d diffractive imaging of nanoparticle ensembles using an x-ray laser. *Optica* (2021) 8:15–23. doi:10.1364/OPTICA.410851

39. Sobolev E, Zolotarev S, Giewekemeyer K, Bielecki J, Okamoto K, Reddy HKN, et al. Megahertz single-particle imaging at the european xfel. *Commun Phys* (2020) 3:97. doi:10.1038/s42005-020-0362-y

40. Lehmkuhler F, Dallari F, Jain A, Sikorski M, Möller J, Frenzel L, et al. Emergence of anomalous dynamics in soft matter probed at the european xfel. *Proc Natl Acad Sci* (2020) 117:24110–6. doi:10.1073/pnas.2003337117

41. Grünbein ML, Bielecki J, Gorel A, Stricker M, Bean R, Cammarata M, et al. Megahertz data collection from protein microcrystals at an x-ray free-electron laser. *Nat Commun* (2018) 9:3487. doi:10.1038/s41467-018-05953-4

42. Grünbein ML, Bielecki J, Gorel A, Stricker M, Bean R, Cammarata M, et al. Mhz data collection of a microcrystalline mixture of different jack bean proteins. *Scientific Data* (2019) 6:18. doi:10.1038/s41597-019-0010-0

43. Dallari F, Reiser M, Lokteva I, Jain A, Möller J, Scholz M, et al. Analysis strategies for mhz xpcs at the european xfel. *Appl Sci* (2021) 11:8037. doi:10.3390/app11178037

44. Dallari F, Jain A, Sikorski M, Möller J, Bean R, Boesenberg U, et al. Microsecond hydrodynamic interactions in dense colloidal dispersions probed at the European XFEL. *IUCrJ* (2021) 8:775–83. doi:10.1107/S2052252521006333

45. Dallari F, Lokteva I, Möller J, Jain A, Roseker W, Westermeier F, et al. Coherence properties from speckle contrast analysis at the european xfel. *J Phys Conf Ser* (2022) 2380:012085. doi:10.1088/1742-6596/2380/1/012085

46. Reiser M, Girelli A, Ragulska A, Das S, Berkowicz S, Bin M, et al. Resolving molecular diffusion and aggregation of antibody proteins with megahertz x-ray free-electron laser pulses. *Nat Commun* (2022) 13:5528. doi:10.1038/s41467-022-33154-7

47. Dallari F, Lokteva I, Möller J, Roseker W, Goy C, Westermeier F, et al. *Real-time swelling-deswelling kinetics of nanogels driven by xfel pulses*. in preparation (2023).

48. Trost F, Ayyer K, Prasciolu M, Fleckenstein H, Barthelmeß M, Yefanov O, et al. Imaging via correlation of x-ray fluorescence photons. *Phys Rev Lett* (2023) 130:173201. doi:10.1103/PhysRevLett.130.173201

49. Sobolev E, Schmidt P, Malka J, Hammer D, Boukhelef D, Dall'Antonia F, et al. Reducing big data at the european xfel. *Front Phys* (2023). Paper in preparation.

50. Mozzanica A, Andrä M, Barten R, Bergamaschi A, Chirioti S, Brückner M, et al. The JUNGFRU detector for applications at synchrotron light sources and XFELs. *Synchrotron Radiat News* (2018) 31:16–20. doi:10.1080/08940886.2018.1528429



OPEN ACCESS

EDITED BY

Iain Sedgwick,
Rutherford Appleton Laboratory,
United Kingdom

REVIEWED BY

Emmanuel Farhi,
Soleil Synchrotron, France
Andras Wirth,
Diamond Light Source, United Kingdom

*CORRESPONDENCE

Egor Sobolev,
✉ egor.sobolev@xfel.eu

†PRESENT ADDRESS

Hans Fangohr, Max Planck Institute for the
Structure and Dynamics of Matter, Hamburg,
Germany

RECEIVED 31 October 2023

ACCEPTED 06 February 2024

PUBLISHED 27 February 2024

CITATION

Sobolev E, Schmidt P, Malka J, Hammer D,
Boukhelef D, Möller J, Ahmed K, Bean R,
Bermúdez Macías IJ, Bielecki J, Bösenberg U,
Carinan C, Dall'Antonia F, Esenov S, Fangohr H,
Ferreira de Lima DE, Ferreira Maia LG, Firoozi H,
Flucke G, Gessler P, Giovanetti G, Koliyadu J,
Madsen A, Michelat T, Schuh M, Sikorski M,
Silenzi A, Sztuk-Dambietz J, Turcato M,
Turkot O, Wrigley J, Aplin S, Hauf S, Wrona K and
Gelasio L (2024), Data reduction activities at
European XFEL: early results.
Front. Phys. 12:1331329.
doi: 10.3389/fphy.2024.1331329

COPYRIGHT

© 2024 Sobolev, Schmidt, Malka, Hammer,
Boukhelef, Möller, Ahmed, Bean, Bermúdez
Macías, Bielecki, Bösenberg, Carinan,
Dall'Antonia, Esenov, Fangohr, Ferreira de Lima,
Ferreira Maia, Firoozi, Flucke, Gessler,
Giovanetti, Koliyadu, Madsen, Michelat, Schuh,
Sikorski, Silenzi, Sztuk-Dambietz, Turcato,
Turkot, Wrigley, Aplin, Hauf, Wrona and Gelasio.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Data reduction activities at European XFEL: early results

Egor Sobolev^{1*}, Philipp Schmidt¹, Janusz Malka¹,
David Hammer¹, Djelloul Boukhelef¹, Johannes Möller¹,
Karim Ahmed¹, Richard Bean¹, Ivette Jazmín Bermúdez Macías¹,
Johan Bielecki¹, Ulrike Bösenberg¹, Cammille Carinan¹,
Fabio Dall'Antonia¹, Sergey Esenov¹, Hans Fangohr^{1,2†},
Danilo Enoque Ferreira de Lima¹, Luís Gonçalo Ferreira Maia¹,
Hadi Firoozi¹, Gero Flucke¹, Patrick Gessler¹,
Gabriele Giovanetti¹, Jayanath Koliyadu¹, Anders Madsen¹,
Thomas Michelat¹, Michael Schuh¹, Marcin Sikorski¹,
Alessandro Silenzi¹, Jolanta Sztuk-Dambietz¹, Monica Turcato¹,
Oleksii Turkot¹, James Wrigley¹, Steve Aplin¹, Steffen Hauf¹,
Krzysztof Wrona¹ and Luca Gelasio¹

¹European XFEL, Schenefeld, Germany, ²University of Southampton, Southampton, United Kingdom

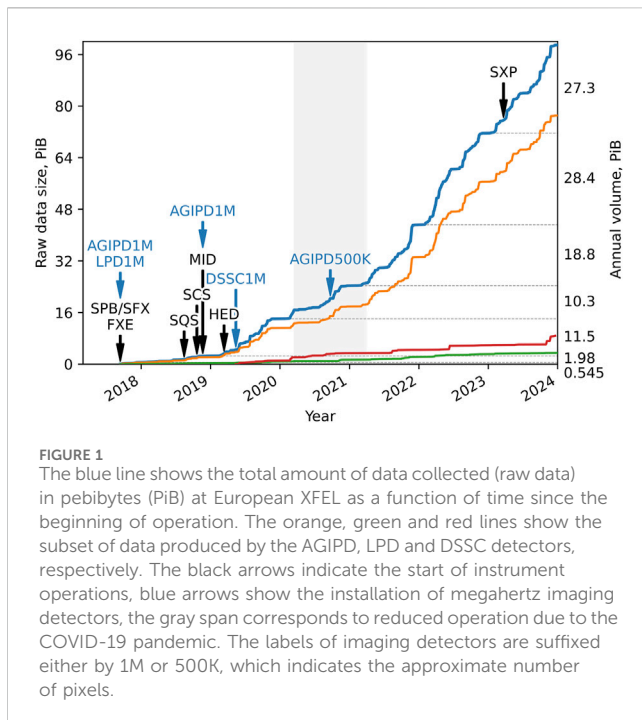
The European XFEL is a megahertz repetition-rate facility producing extremely bright and coherent pulses of a few tens of femtoseconds duration. The amount of data generated in the context of user experiments can exceed hundreds of gigabits per second, resulting in tens of petabytes stored every year. These rates and volumes pose significant challenges both for facilities and users thereof. In fact, if unaddressed, extraction and interpretation of scientific content will be hindered, and investment and operational costs will quickly become unsustainable. In this article, we outline challenges and solutions in data reduction.

KEYWORDS

data reduction, photon science, X-ray free electron laser, data management, high-repetition rate, big data

1 Introduction

Scientific ambition pushes the progress of modern X-ray light sources. As a result of the steady evolution in accelerator and detector technology, as well as increased levels of automation and improved quasi-real-time feedback, the amount of experimental data produced by photon sources is increasing at unprecedented rates [1, 2]. In particular, the continuous development and improvement of X-ray imaging detectors (see, e.g., [3], and references therein) is instrumental to enable the scientific exploitation of the exceptional brightness characteristic of fourth-generation light sources [4] and X-ray free electron lasers (XFELs) (see, e.g., [5, 6]). State-of-the-art pixelated X-ray detectors, custom-made or commercially available, can routinely collect hundreds to a few thousands images per second [3, 7–14]. Among modern facilities, the European XFEL is a MHz-repetition-rate X-ray free electron laser providing extremely bright, spatially coherent pulses, which are characterized by a temporal duration of tens of femtoseconds or less [15, 16]. The facility



operates in a so-called “burst mode,” delivering 10 Hz trains of up to 2,700 X-ray pulses. The intra-pulse separation can be as low as 222 ns, equivalent to a repetition rate of 4.5 MHz.

The detector data rates at the European XFEL can exceed one hundred gigabits per second, resulting in the production of several petabytes of data for a single experiment with a typical duration of two to 6 days. The MHz-capable detectors at the European XFEL are the DEPFET Sensor with Signal Compression (DSSC) [12], the Adaptive Gain Integrating Pixel Detector (AGIPD) [10, 11], and the Large Pixel Detector (LPD) [8, 9], each of which has up to 1024×1024 pixels. When integrated into our infrastructure, the maximum data rates are 134 Gbit/s, 118 Gbit/s, and 86 Gbit/s, respectively. Additionally, digitizers’ data rates can approach several gigabits per second, and multiple of these devices may be employed during a single experiment. As a result, the volume of scientific data collected during user experiments has steadily increased since operation began in 2017. This is illustrated by Figure 1, where the total amount of raw data (its precise definition is given in Section 2.1)—about 100 PiB as of today—is shown as a function of time. Apart from a deceleration caused by a reduced number of experiments during the most acute phase of the COVID-19 pandemic, the rate of data collection is ever-growing. This can be explained by the asynchronous start of the seven scientific instruments, as highlighted in Figure 1, as well as the continuous enhancement in operational efficiency—from accelerator and instrument performances to data systems reliability, from procedure optimisation and automation to advances in sample delivery. While operational efficiency cannot increase indefinitely, future facility upgrades will inevitably result in higher data throughput. In the short term, for example, an AGIPD of 3.7 Mpx, with a theoretical data rate approaching half a terabit per second will be installed. In the medium

term, upgrades of the accelerator will increase its duty factor, and in turn the number of X-ray pulses delivered each second [16, 17]. Additionally, the current scientific data policy¹ defines that scientific data at European XFEL shall be curated for at least five years although striving for ten.

Storage systems are expensive and limited in lifetime, they consume electric energy, increase CO₂ emissions, and require dedicated personnel for their maintenance and operation. The resulting non-negligible economical and environmental footprint must be urgently addressed. This means that, altogether, a continuous expansion of the storage system is not sustainable.

While storage-related issues are the most evident, the enormous data rates and volumes pose other challenges, both from a technical and a scientific point of view. In fact, the complex solutions required to handle the enormous data rates often necessitate using leading-edge technology. This is expensive and requires deep expert knowledge to keep the systems stable. In operation, these systems may be prone to instabilities—like the degradation of their performances—which in turn, could potentially disrupt data acquisition and near real-time monitoring of the experiments. The data coming from the MHz-capable detectors is not trivial to interpret, and the European XFEL has been developing the so-called correction pipeline to transform it into physics content [18]. This pipeline is typically triggered automatically as soon as data is collected, and results in additional data transfer, processing, and storage requirements. As exemplified by the correction pipeline, the analysis of large amounts of data typically requires software that can exploit multiple computational nodes, and cope with latencies inherent to ingesting data at high rates. Accordingly, distilling scientific content can be considerably more challenging with a larger data volume, and can be potentially compromised if data is not pre-processed by specialized tools developed by experts. More complex analysis methods, in turn, increase latency, which is particularly detrimental when using analysis results to steer the running experiments. The additional complexity can even represent an insurmountable barrier for inexperienced users, which makes the facility less accessible to test new scientific methodologies and ideas.

The only solution to the aforementioned issues is to *reduce* the amount of data, while maximizing its scientific value. Generally, several reduction operations can be performed during processing and evaluation of collected data. These are either *data selections*—rarely the entirety of collected data is used—or *data transformations*, e.g., dimensionality reduction through integration along some variable.

While the previous discussion revolved around the use case of the European XFEL, the issues encountered are by no means specific to our facility. Large-scale high-energy particle physics facilities, for instance, have embedded their data reduction strategy as part of their original technical design decades ago (see, for example, Refs. [19–21]). Additionally, most of the modern X-ray photon sources are exploring and developing data reduction strategies [22], also owing to initiatives of policymakers such as European Union’s

¹ Scientific Data Policy of the European X-Ray Free-Electron Laser Facility GmbH (Version 1). European XFEL, Schenefeld, Germany (2017). doi: 10.22003/XFEL.EU-TR-2017-003.

Horizon 2020 LEAPS-INNOV, and the results of individual research groups [23–30]. Topics explored vary from lossless and lossy compression [31] to artificial intelligence [25, 26, 32, 33], from dedicated hardware solutions to FAIR data [34].

Even though the benefits are clear to both facilities and users, several open questions and challenges remain from a technical, scientific, and social point of view. Overall, the risk of data reduction introducing bias in the results must be minimized, and the ratio of scientific content over collected data maximized. That is, only the data contributing to the answer of a specific scientific question should ideally be curated. Scientists must be given control of the reduction pipeline, including access to detailed validation metrics.

Reducing data is not avoidable anymore at the European XFEL. It is our duty to provide tools that enable users to perform data reduction, thereby maximizing the scientific outcome of the experiments and minimizing the pressure on our infrastructure. These tools need to be as transparent and automated as possible, and their output must be corroborated through extensive validation. We finally aim at providing extensive and reliable information to support users' decisions during and after experiments.

This manuscript aims to serve as an entry-point for our users as it reports on developed solutions, future plans, as well as strategies for data reduction and curation at the European XFEL. It furthermore details challenges and opportunities intrinsic to data reduction. Further documentation and continuously updated information are, and will be, made available in Ref. [35]. In Section 2, we provide an overview of the data infrastructure of the European XFEL, and its upgrade to enable integration of data reduction techniques. In Section 3, we present and discuss selected data reduction workflows, and their applications to data reduction. In Section 4, the current state and future plans are discussed.

2 Methods

We define data reduction as the act of applying selection and transformation techniques to experimental data with the goal of maximizing the density of scientific content. Different quality criteria or filtering of particular event types can potentially be used to distinguish valuable and disposable data. An example of a quality criterion is the X-ray pulse energy being measured above a given threshold, while a possible event type includes the identification of photons scattered by a sample. Similarly, interesting detector regions can be identified, and data outside these regions can be ignored. Possible data transformations include dimensionality reduction, change of representation, compression, and additional data analysis methods. Dimensionality reduction may be achieved by discarding a portion of the parameter space, or by integrating data along certain variables. Common operations include binning, averaging of several data sets, or integration of images along, e.g., the azimuthal angle.

Different experimental techniques leverage various sets of physical observables, which are analysed according to the scientific goal of the experiment. European XFEL offers a wide spectrum of such techniques, hence requiring a flexible choice of the data reduction method. We refer to such data reduction methods as *technique-specific*. In contrast, *operation-specific* methods depend on particular experiment modalities, e.g., a specific detector configuration.

Examples of technique-specific data reduction can be found in serial femtosecond crystallography (SFX) [36, 37] and single particle imaging (SPI) [36, 38], where a significant fraction of the collected data does not capture a scattering event. The procedure of identifying whether or not the X-ray beam scattered off the sample is referred to as hit finding, and can be an important trigger for selecting data. Another example of a technique-specific data reduction method is found in small- and wide-angle X-ray scattering experiments [39], where rotation invariant scattering data can be azimuthally integrated and reduced to a one-dimensional curve. Likewise, in X-ray photon correlation spectroscopy (XPCS) [40, 41] and X-ray cross-correlation analysis (XCCA) [42], the 2D detector data may be reduced to intensity-intensity correlation functions.

Technique-specific methods are often more challenging, as they rely on a proper selection and configuration of the analysis pipeline. Facility users are often the most experienced with the latter task, driving the scientific analysis for a given experiment. Thus, we aim at offering full control over the reduction pipeline, alongside detailed metrics to continuously monitor the reduction outcome.

Finally, it is worth mentioning that data reduction can be implemented at different stages of the experiment, that is, during the data acquisition (online) or after data have been stored to disk (offline), implying different requirements and limitations. Decisions can either be automatic and irreversible, or manual and assisted through detailed event-based annotation.

2.1 Overview of the data infrastructure at the European XFEL

The European XFEL's storage and computing systems [43] are separated into *online* and *offline* storage and processing infrastructure (see Figure 2). The online storage is a performant cache capable of ingesting scientific data produced during experiments. In order to be able to effectively steer and control experiments, the online computing cluster is used to process data streams provided during data acquisition, and with minimal latency. After collection, data identified as potentially interesting in the data management portal myMdC [44] is copied from the online storage to a second high-performance layer, implemented using the IBM Elastic Storage System building blocks and presented as a unified IBM Spectrum Scale (a.k.a. GPFS) filesystem [45]. Here, file-based processing is performed using the offline computing cluster Maxwell [46]. This storage system is used during the experiment up to a few months after it. The third layer is mass storage based on the middleware system dCache [47], which extends the capacity of the high-performance system. Both the high-performance and mass storage systems together are often referred to as offline storage. The last layer is the tape archive, which provides resources for long-term data preservation.

The supervisory control and data acquisition (SCADA) system Karabo [48, 49], which is developed in-house, plays a key role in data ingestion, and experiment and beamline control. Karabo implements an event-driven paradigm, built around a central message broker. Functionalities—either hardware integration or high-level procedures—can be easily added to the core system via plugins called devices. Devices can access any information in the

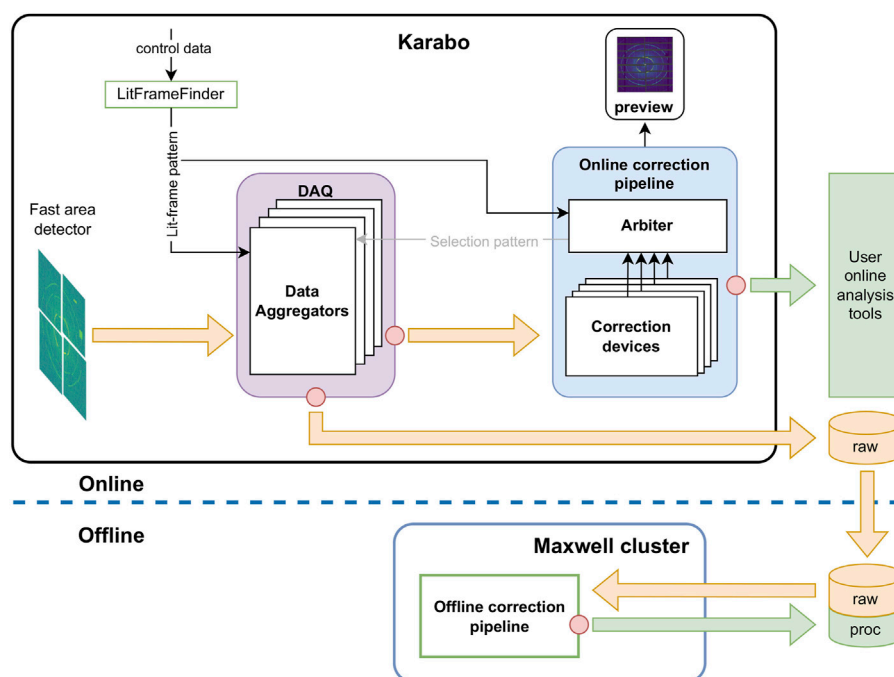


FIGURE 2

Simplified schematic of the European XFEL data acquisition and processing. Cylinders represent data storage elements. Processes above the dashed blue line are part of the online infrastructure. After the raw data is stored, it is copied to the offline Maxwell cluster, at which point the offline correction pipeline transforms raw data into corrected data, both of which are made available as part of the offline infrastructure. The online infrastructure is implemented in the Karabo framework, in which fast detector data is fed to the data acquisition system, that aggregates data from several sources. The online correction pipeline not only corrects data for online inspection: it also contains an arbiter, which decides which data to store, based on the configured data reduction methods. In this schematic, red circles represent data reduction points.

distributed Karabo system, e.g., control parameters or detector readings, which can be readily used also for data reduction purposes.

The data acquisition (DAQ) system is also implemented in Karabo [50, 51]. It aggregates data from any selected device in the distributed Karabo system, including area detectors, and matches the data by event (train) index, before storing it on the disk. This data is termed raw data. In addition, the DAQ outputs and streams the data online for monitoring purposes. Data acquisition and recording are implemented in so-called data aggregators.

On the online cluster, the monitoring data stream from the DAQ is sent to the online correction pipeline [18]. The latter, also implemented in Karabo, transforms raw into usable data with latencies up to a few seconds. The correction pipeline can be extended through computational kernels—say, custom data analysis procedures—implemented as add-ons. For big area detectors, consisting of multiple sensor modules, processed data is aggregated and dispatched for further online processing.

After raw data stored on the online cluster is copied to the offline cluster, the offline correction pipeline produces a corrected copy of this data, which is stored as so-called corrected data. This typically doubles the volume of data collected from area detectors in the context of an experiment. However, the lifetime of processed data can be arbitrarily short, as it can be reproduced from the corresponding raw dataset at any point in time.

2.2 Data reduction points in the data system and associated risks

Data reduction can be applied at different points in the data system (see Figure 2), with different implications. In particular, the earlier and closer the point to the source of the data, the higher the impact on the system.

As previously introduced, the DAQ defines which raw data will be stored or transferred to the correction pipeline. Therefore, any reduction at the DAQ-level is irreversible and can hence only be applied when the associated risk is minimal. Furthermore, at this stage, it is difficult to include complex processing of detector data, due to the strict latency requirements, and the dependencies on the scientific methodology or detailed data analysis, which are difficult to automate. Therefore, decisions are only based on operating conditions that are readily available in the Karabo environment. Any reduction at the DAQ level maximizes the impact on the downstream data system. As of today, this point is used solely to filter detector frames not exposed to X-rays (see Section 3.1.1), but other reduction techniques will be implemented, including module or region of interest selection or gain bit suppression.

The next reduction point is at the output of the online correction pipeline. Here, the data has been modified for the benefit of downstream online analysis tools, which receive filtered or pre-processed and simplified data. Owing to this, data reduction at this point can decrease feedback latency,

thereby enhancing response times in experiment steering. Additionally, the load on the network and computing infrastructure decreases. Examples of add-ons implemented within the correction pipeline include a peak-finding algorithm (for SFX), a lit-pixel counter (mainly for SPI) and a per-detector-module estimator of average intensity. These cover typical imaging- and event-based experimental techniques. We foresee that our users will be able to fully exploit this reduction point by contributing additional data processing code in the future. Decision criteria at this stage are potentially much more complex than the ones at the DAQ-level, and might require some degree of parameter tuning either by experts or algorithms as the experiment progresses. The time budget for these optimizations, however, is extremely limited owing to the ephemeral nature of data streams. That is, decisions must be taken before the next data batch. Therefore, there is a certain risk of biasing the downstream analysis due to inaccurate data reduction, with consequences on experiment steering and, thus, experiment outcome. In the future, the filtering applied at this point will also be fed back into the DAQ, which can either annotate or reduce raw data before storage. The latter case is more risky, as data is irreversibly discarded. Furthermore, the tuned configuration parameters will be stored such that they can be later considered as part of offline correction, or retroactively by users.

Further downstream, the next data reduction point is at the end of the offline correction pipeline. In this case, only processed data is affected, while raw data remains unaffected. As the former can be fully reproduced (see Ref. [18]) from the latter, this is a minimal risk data reduction. However, similar to its online counterpart, reduction can bias analysis, thus affecting the quality of the extracted scientific content. At this point, further reduction decisions can be taken that will be applied to the processed data immediately, or can be used to either annotate or reduce stored raw data as well.

Finally, data reduction methods can be applied to offline raw or processed data retroactively by users of the facility or sophisticated algorithms. The reduced data sets can be produced (i) by a tool provided by the European XFEL, (ii) by one of the said tools taking into account decisions derived from user input (e.g., list of hits for SFX or SPI experiments), or (iii) by user tools (perhaps to be integrated into our data system for the benefit of a larger community), provided the data format is compatible with the EuXFEL's.³

To facilitate the reduction of existing data and ensure its compatibility with the facility's data format, the `exdf-tools` package [52] has been developed. This is implemented via small plugins which allow for the usage of several data reduction operations, such as removing a train or pulse for specific sources and keys. All such operations may be collected and applied while rewriting the input data into new files, and serve as a detailed record of how the data was modified.

3 Results

Below, selected examples of data reduction methods implemented at the European XFEL are introduced, and their impact is discussed. First, operation-specific methods are presented. As is evident from Figure 1, to date, AGIPD detectors have produced the majority of data. Therefore, developing methods specific to this detector has been of the highest priority. In the second part of this section, technique-specific methods are discussed.

3.1 Operation-specific methods

Below, we describe the operation-specific methods currently implemented at European XFEL. As mentioned previously, operation-specific methods are technique-independent and related to instrument operation itself. As ideally, no analysis is required to decide on the data, these methods are robust, low risk, and the feedback latency is compatible with online requirements. All methods except for the module selection are fully automated.

3.1.1 Lit frames selection

Fast area detectors collect data frames in batch mode upon triggering at 10 Hz. Within such a batch, called a train, X-ray pulses can be delivered in arbitrary patterns, according to experimental conditions and requirements. For instance, the intra-train repetition rate can be lowered so as to allow the sample delivery system to replenish the interaction region before the next pulse. In other cases, a complex pulse pattern can be used to probe particular sample dynamics. As a result, some detector images might be recorded in absence of X-rays, and therefore are called dark frames. Megahertz imaging detectors at European XFEL were designed to implement veto mechanisms to reuse memory cells and avoid recording dark frames.

However, given the complexity of current detectors, vetoing might potentially affect data quality or complicate operation. This is particularly true for the AGIPD, which requires an individual set of calibration constants for different veto patterns. Covering all possibilities in calibration is infeasible, and thus, the AGIPD is usually operated with a fixed veto pattern, rather than one that acts on the dynamic changes in the X-ray pulse pattern. To mitigate this, we have implemented a Karabo device which aggregates relevant accelerator and AGIPD settings and annotates collected data accordingly. This information can be used to select data at any reduction point. This is a low risk method, and its reduction factor is the ratio of selected and collected frames.

The selection of lit-frames is routinely applied as part of the offline correction pipeline to the corrected data at the MID [53] and SPB/SFX [54] scientific instruments. Furthermore, this method has been applied at the DAQ reduction point, so far for testing purposes. Owing to the application of this method, in 2023 the storage of 0.65 PiB of raw data and 1.7 PiB of processed data has been avoided (see Table 1). In the latter case, the corresponding raw data can also be retroactively reduced.

3.1.2 Gain data suppression or compression

X-ray detectors at the European XFEL use different mechanisms to increase the dynamic range of the detected signal so as to extend

2 Scientific Data Policy of the European X-Ray Free-Electron Laser Facility GmbH (Version 2). European XFEL, Schenefeld, Germany (2025). doi: 10.22003/XFEL.EU-TR-2025-001.

TABLE 1 Examples of application of reduction methods to AGIPD data.

Reduction method	Type	Instrument	Since	Experiments	Original data size, PiB	Reduction factor
Applied reductions (avoided storage of 7.4 PiB)						
Lit-frame selection	raw	SPB/SFX	1 month	2	0.88	3.8
	corr	SPB/SFX	3 months	12	3.8	1.2
		MID	1 year	10	5.8	2.5
Conversion to ph. and compression	corr	MID	1 year	10	5.8	17
Train selection	corr	HED	1 year	4	0.52	19
Candidate to retroactive reduction (17 PiB expected to be freed)						
Lit-frame selection	raw	SPB/SFX	1 year	27	9	1.11
		MID	2 years	23	14	1.9
Gain information suppression	raw	SPB/SFX	1 year	5	1.2	2
		MID	2 years	12	7.4	2
Train selection	raw	HED	1 year	4	0.52	19
Module selection	raw	MID	2 years	5	2.3	5
SPI hit finding	raw	SPB/SFX	2 years	4	5.5	19

The table reports the reduction method (“Reduction method”); the type of data, that is “raw” or “corr” for raw and corrected data, respectively (“Type”); the scientific instrument (“Instrument”); the time period the reduction has been applied (“Since”); the number of experiments that have been reduced (“Experiments”); the unreduced data volume (“Original data size”); the average reduction factor (“Reduction factor”).

their ability to acquire a trustworthy and physically meaningful signal. Among them, the AGIPD uses an adaptive gain method: it stores two 16-bit integers for every pixel, one for the signal amplitude and one encoding the gain stage, which identifies the amplification factor. Under certain illumination conditions, achieved typically during XPCS or SPI experiments, a single gain stage is used. Furthermore, for some experimental techniques, pinning the gain stage is desirable so as to achieve a simpler detector response. In such cases, the gain information could be substituted without risk by a fixed value for the entire data acquisition period.

The reduction factor corresponding to the suppression of the gain data is two. The method is currently available for retroactive reduction. To avoid any modification of the data format, we exploit a HDF5 [55] feature which allows to keep the original raw data dimensions without allocating storage for the gain information. An implementation compatible with the online correction pipeline is also being developed.

In 2023, 1.2 PiB of data has been collected with fixed (medium) gain at the SPB/SFX scientific instrument, and 3.1 PiB with special settings useful to amplify low-intensity signal at the MID instrument. Retroactive suppression of the gain data therefore will allow to release 2.2 PiB of storage (see Table 1).

Complementary to this strategy, we are evaluating the replacement of a 16-bit gain signal with a unique integer representing the gain stage. Ideally such a step would not represent a loss of data. However, the impact of noise in the original gain signal may lead to a data quality loss, especially when close to a gain transition. Therefore we are currently evaluating that impact and establishing contingencies. For AGIPD, the original 16-bit values are converted into three

possible values. Accordingly, the expected benefit from a lossless compression of the gain information, even when using a standard algorithm such as *Deflate* [56], is sizeable.

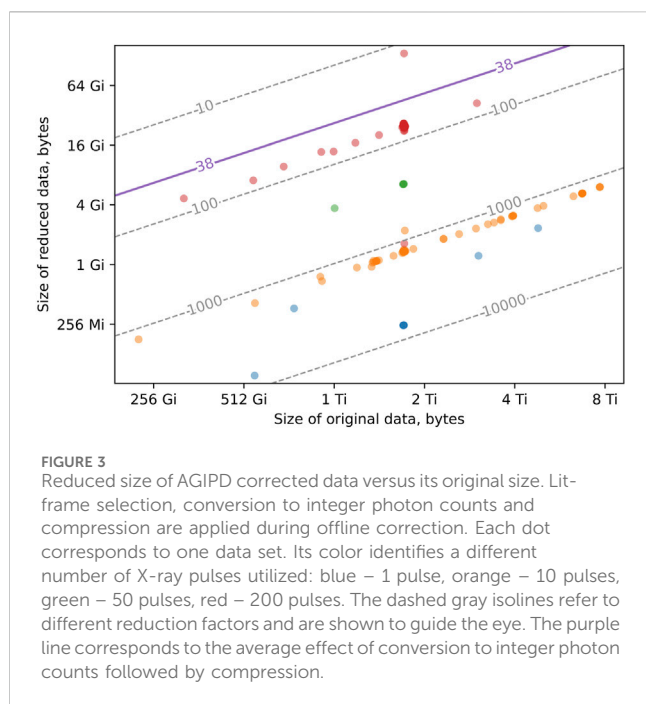
3.1.3 Train selection

In some cases, it may be meaningful to select a subset of the trains to record. This is particularly relevant to the HED scientific instrument [57], where a shutter wheel is used to mechanically filter X-rays and illuminate the sample only with a specific pulse train. An incorrect selection poses a large risk of losing the relevant data. Thus, the DAQ stores several adjacent trains in addition to the selected one. This allows for validation of the train selection reduction method, and minimizes the risk of data loss in case of an incorrect setting. The information on selected trains is available in Karabo, which controls the wheel, and can be readily exploited by the offline correction pipeline upon validation, or retroactively. In the future, we aim to incorporate this reduction method at the DAQ level.

The reduction factor equals the ratio of DAQ-acquired and selected trains, and, depending on the DAQ settings and the applied procedure, it can be of the order of hundreds. In 2023, 0.6 PiB of data, including disposable trains, has been collected at the HED scientific instrument. The train selection method has been applied to these data at the offline correction pipeline stage, reducing it by a factor of 19 (see Table 1). We plan to retrospectively reduce the corresponding raw data as well.

3.1.4 Module and region-of-interest selection

For certain experiments, the relevant signal is confined to a well-defined region-of-interest (ROI) on the detector. Most European XFEL X-ray imaging detectors are modular, and therefore, only a



few modules may intersect with the ROI. For technical reasons, the data from each module is saved in a different file. Hence, it is straightforward to select only files corresponding to the relevant modules to obtain a significant reduction of offline data.

Bragg coherent diffraction imaging (BCDI) is one of the experimental techniques that can benefit from this reduction method. Since 2022, five BCDI experiments were performed at the MID scientific instrument, and for half of them only one detector module (out of sixteen) contained data of interest. By retroactively removing data corresponding to the other modules, the initial volume can be reduced from 2.3 PiB to 0.46 PiB, that is a reduction factor of five (see Table 1).

At the time of writing, this method can be employed by manually selecting relevant modules using the DAQ interface. In the near future, a graphical user interface will be available to configure the DAQ and perform validation by monitoring the signal on the entire detector.

The reduction factor equals the ratio of the total number of detector modules to the number of selected modules. In the future, the possibility of storing defined regions of interest within modules will also be exploited.

3.2 Technique-specific methods

Technique-specific methods require processing of collected data, which typically involves fine tuning of certain analysis parameters so as to ensure accurate results. As such, associated risks of discarding meaningful data are generally higher compared to operation-specific methods, and present challenges for automation. Furthermore, appropriate pre-processing of the data—e.g., handling of detector artifacts such as pixels with erroneous readings (perhaps damaged), or accurately mapping data according to the physical detector

layout—as well as extensive validation are required. For several experimental techniques, the scientific community has developed specialized methods and software tools, which can provide feedback on data content and quality. These may need to be integrated into the European XFEL computing and control environments to fully leverage data reduction opportunities through automation and implementation of fast feedback loops.

3.2.1 Effective compression by decreasing entropy: conversion to integer photon counts and down-sampling of collected intensities

The effectiveness of compression methods increases with a lower Shannon entropy of collected data [58]. To reduce entropy, we have evaluated the application of physics-motivated techniques to compress detected intensities. The risk associated to these is particularly low if analysis techniques rely on a sizable ensemble of individual measurements.

In some cases, it may be advantageous to convert measured intensities into the absolute number of photons and represent them as integers. This applies to experiments where the scattering signal is rather weak, such as SPI and XPCS, or the strong signal is localized and the background is sparse and weak, such as BCDI. Through this quantization procedure the entropy may be significantly reduced and detector images become extremely compressible with lossless methods. We use the *Deflate* algorithm as implemented in the *HDF5* library, making access to data transparent to users.

In particular, at the MID scientific instrument the combination of selection of lit-frames, conversion to integer photon counts and compression is routinely applied at the offline correction reduction point. Owing to this, 5.7 PiB of processed data were not stored in 2023, corresponding to an overall reduction factor of approximately 42 (see Table 1). Figure 3 illustrates the effect of applying the above-mentioned reduction chain to a Bragg XPCS experiment. Here the overall reduction factor was 97. In the figure, the reduced size of AGIPD corrected data is shown as a function of its original size. Each point refers to the data set of an individual measurement, a so-called run. Points are colored depending on the number of X-ray pulses utilized for the measurement, and each dashed gray line, shown to guide the eye, is an isoline indicating the same reduction factor.

The position of each point is given by the combined effect of selecting lit-frames, converting intensities to an integer number of photons and, finally, compressing this data. The average reduction factor due to conversion to integer photon counts followed by compression, shown as the purple line, corresponds to 38. This value depends on illumination conditions, which, in this experiment, are similar for most data sets. Additionally, the lit-frames selection contributes, with a reduction factor equal to the ratio between total number of collected frames to X-ray pulses, varying between 1.76 and 352.

For experiment techniques exploiting a large intensity range, such as SFX, rounding to a given number of significant bits reduces the distribution of pixel values [24], and Shannon entropy accordingly. This also makes images more compressible with lossless methods. The quality of the final result depends on the rounding settings, which have to be balanced with respect to the potential for desired reduction.

This method is typically reliable upon validation, which can be based, for example, on comparing subsets of results, and its reliability increases with the number of repetitions of a certain measurement. The risk of using such methods is mitigated by applying them only to the corrected data.

Both methods discussed are available for offline usage. Conversion to integer photon counts and subsequent compression is integrated into the offline correction pipeline and used in operations.

3.2.2 Azimuthal integration of rotation invariant data: small- and wide-angle scattering

The first processing step of rotation invariant data is typically the azimuthal integration of detector frames. The transformation of two-dimensional images into one-dimensional radial profiles is an operation which scales with the square-root of the number of pixels, and thus yields a reduction factor of about 1,000 for megapixel images.

Proof-of-principle automatic azimuthal integration after detector data correction has already been employed at European XFEL. The automated pipeline processes frames in batch mode and can exploit the *pyFAI* library [59]. We are further improving it by enabling parallel data reads from disk, and signal integration in parallel on GPUs. The pipeline output can replace processed data containing corresponding two-dimensional images. Furthermore, we are developing an azimuthal integrator add-on for the online correction pipeline.

Current research in validation includes a reliable and automatic correction for potential displacements of the X-ray beam during data collection. In fact, several experimental techniques require a precise estimation of the X-ray axis. Its erroneous assessment degrades the quality of azimuthally integrated data, for example, as the integration axis does not coincide with the X-ray axis, which is a symmetry one.

3.2.3 Hit finding: serial femtosecond crystallography and single particle imaging

For experimental techniques like SFX and SPI, the X-ray beam interacts with the sample with a certain probability, known as the hit rate. In fact, “hits” and “non-hits” are defined as detector frames either containing signal scattered from the sample, or only background photons. The number of hits compared to the total amount of delivered pulses, the hit rate, is typically rather modest, of the order of 0.1%–10%, depending on the sample and the injection method. As a result, a considerable amount of detector images have to be acquired during the experiment for successful data analysis, and the potential for data reduction by discarding all non-hits is significant.

The first step of the SFX data analysis pipeline consists of the identification of Bragg peaks in a detector frame. If the number of peaks exceeds a user-defined threshold, that frame is considered to be a hit. If the next step of the analysis pipeline, indexing, is considered as well, the reduction factor can be potentially increased further at the cost of higher complexity. Different software tools provided by the scientific community implement such complete analysis pipelines [60, 61] (a description can be found, e.g., in [62] and references therein). Among these, we have integrated the *CrystFEL* suite [60] into the European XFEL

infrastructure. We provide the latter through the *EXtra-Xwiz* tool [62, 63], so as to abstract certain complications specific to our data structure and computing environment.

When processing SPI data, the number of pixels on a detector frame characterized by a signal intensity above a certain threshold is initially evaluated. Hits satisfy the condition that the number of such lit pixels exceeds another threshold. Also in this case, data analysis tools provided by scientific communities exist [64].

Strategies for validation of the hit finding output include the (graphical) provision of key indicators. These can be statistical views – such as mean, variance, or detected outliers – of retained and discarded data, or more sophisticated feedback calculated at different stages of the data analysis pipelines. For example, a pseudo-powder diffraction pattern can be calculated from SFX data as the sum of extracted Bragg peaks, and relates to the crystalline structure of the sample. Similarly, cell parameters or quality metrics can be extracted at the indexing step.

At the time of writing, hit finders for SFX and SPI experiments are implemented as add-ons in the online correction pipeline, with reduction decisions taken at the arbiter (cfr. Figure 2). These have been tested in production, and satisfy the stringent latency requirements of online analysis. Furthermore, implementations for the offline correction pipeline are in progress, and information on hits produced by diverse tools can be used to retroactively reduce data. Certain SPI experiments have been already identified for retroactive data reduction, corresponding to a raw data volume of 5.5 PiB and with hit rates ranging between 0.1% and 13.5% (see Table 1). Therefore, in this case the average reduction factor is roughly 19 or lower, depending on the need for non-hits which can be used for background estimation. Corrected data can be reduced with the same ratio or better, if conversion to integer photon counts and compression are further applied.

While in this section we present examples of binary classification, concepts introduced here can be extended to more complicated use cases and generalized as data clustering.

3.2.4 Physics reconstruction: reaction microscopy

Reaction microscopy (REMI) [65], also called cold target recoil ion momentum spectroscopy (COLTRIMS), is a momentum imaging technique employed at the SQS scientific instrument [66]. Up to two delay line detectors placed on opposite sides at the end of time-of-flight spectrometers are used to record the kinetic energy and momentum of electrons and ions in coincidence, potentially allowing for a full reconstruction of the scattering process in the molecular frame of reference. On the detector side, the raw data consists of digitized voltage levels acquired at gigahertz sample rates across multiple channels reaching rates up to 4 Gbit/s. From a scientific perspective, only the correlated pairs of position and time for each particle impact are relevant. These have a much lower bandwidth on the order of a few Mbit/s.

An automated reconstruction process from raw data to detector hits is available via the facility offline correction pipeline and allows for efficient data reduction by a factor of approximately 1,000. Validation is provided in the form of reports, which contain statistics and document signal correlations to allow for rapid assessments of reconstruction quality.

3.2.5 Correlation functions: X-ray photon correlation spectroscopy

X-ray photon correlation spectroscopy (XPCS) [40, 41] is a technique used to measure the dynamics of a sample on various time scales. Most XPCS experiments at European XFEL measure dynamics on a timescale of microseconds. Central to the data analysis for XPCS is the calculation of two-time correlation functions (TTCFs). For microsecond XPCS, calculating the TTCF's requires correlating regions of the detector across pulses within a single train.

A library to streamline and optimize the analysis of XPCS data [67] is being developed at the European XFEL. This approach reduces the data that users need to deal with from (on average) 1 TB of detector data to approximately 1 GB of TTCF data, that is a factor of 1,000.

4 Discussion

Although the data reduction activities at the European XFEL are still at their infancy, several methods and strategies have been identified and integrated as part of the data acquisition and analysis systems. Consequently, an initial portfolio of tools has been made available to users of the facility. By routinely applying low-risk methods to processed data, we have already avoided the storage of approximately 7 PiB of data in 2023, thereby reducing the expected volume of processed data to about 70%. Furthermore, the same tools can be applied retroactively with minimal risk, potentially making approximately 17 PiB of additional storage available. An overview of applying the discussed reduction methods to selected AGIPD data is shown in Table 1.

Risks have been assessed for each considered method to ensure minimal impact on the scientific activities, as reiterated throughout this paper (see in particular, Section 2.2). Data reduction is intrinsically associated to the risk of compromising scientific throughput, as a consequence of discarding valuable and non-redundant information, or of applying inaccurate transformations, for example, due to unreliable parameters. Additionally, an incorrect application of data reduction methods to the online data stream would lead to degraded online analysis feedback. Consequently, the experiment steering quality and the beamtime efficiency are compromised. At the other end of the spectrum, if raw data are erroneously reduced, scientific content is potentially irretrievably destroyed.

To mitigate risks, and in addition to extensive user support, we aim to provide our users with information which is as complete and reliable as possible. This includes the production of extensive quality and validation metrics, transparent and comprehensive documentation of the reduction workflow (including any parameter involved), powerful interfaces, as well as various statistics on data usage, which will support user decisions. To be effective, validation metrics must be interpretable, and offer feedback on the effect of any parameter involved. Validation is particularly critical when technique-specific methods are used, as they rely on data assessment and might require tuning. At the time of writing, we offer an initial set of metrics for certain reduction techniques,

both in the form of online feedback, and offline reports. These include, for example, time-averaged online views of retained and discarded data, or monitoring of the signal on the entire detector when only a region of interest is selected.

Additionally, we systematically organize workshops, in which presentations and tutorials are shown to aid users even before they access the facility. Invaluable user feedback is also obtained at such meetings, and helps us adapt to the user needs and address their concerns. This integrates with the extensive documentation and training material that will be made available.

A complementary strategy to reduce risks relies on the development of sophisticated algorithms to decide on reduction methods. Such algorithms work with a clear optimization strategy and with several interpretable metrics embedded in them, to allow for monitoring and control. Although this leads to a more abstract decision process, the availability of meaningful validation systems empowers users to disengage such methods or reconfigure them as required. An example of such an automated process under development, is an application based on mathematical modeling of the parameter optimization procedure for SFX data analysis [68]. This method empowers users by providing high-quality information on collected data, aiding them in steering the experiment, and preventing accordingly the acquisition of low-quality data. Another procedure under development includes the clustering of data as it arrives in the data stream, which allows users to rapidly assess similarities in the data collected, discover patterns and establish low-quality data.

In addition to the technical and scientific aspects of data reduction, another essential enabling step is a corresponding legal framework, which establishes a contract between all parties and defines their responsibilities in this process. The scientific data policy of the European XFEL has undergone a major upgrade to provision this. The upgrade process involved all the stakeholders inside the facility, including groups responsible for data management and analysis, legal specialists, instrument scientists, as well as external advisory committees, and users of the European XFEL. The latter have either been approached individually, or through dedicated events in the context of European XFEL user meetings.

This inclusive process has contributed to assess and address the sociological aspect intrinsic in the paradigm of data reduction. This is overall a new paradigm in photon science, and as such concerns might originate both from (i) the risk of scientific data loss, and (ii) the burden associated to selection of viable data reduction methods, the decision process itself, as well as further analysis downstream of the reduced data sets. Mitigation strategies for the former have already been discussed: we are convinced that, in addition to involving users early on in the process, these strategies will increase users' confidence when applying reduction methods. For the latter, we aim at providing simple interfaces to aid in the decision process, as well as software allowing users to transparently access any kind of data produced at the facility. Furthermore, to support and advise on reduction opportunities as well as to train on available methods and tools we provide, we will take advantage of internal experts as contacts. Overall, users will be assisted in data reduction activities with the provision of tools, information and expertise.

Another measure to increase user involvement in the data reduction process has been the establishment of a data

management plan. Such a data management plan, required by the updated scientific data policy³ for each user proposal, would contain a detailed overview of the reduction solutions applicable to the data collected, formalize requirements and document decisions. This procedure establishes a clear bidirectional communication pathway between the users and the facility from proposal acceptance onward, with the aim of increasing users' trust in the data reduction processes.

To summarize, our aim is to empower users to extract valuable scientific content from collected data. Data reduction is in the users' benefit: for instance, it allows them to achieve a faster turnaround when analysing the experiment's result and to simplify their analysis methodology. Importantly, users shall be responsible for selecting methods and reduction points, balancing risks and benefits, or the retroactive reduction of collected data, such that within a defined amount of time (six months at the time of writing) the size of their data is within the constraints defined by the facility (up-to-date information is available in Ref. [69]).

5 Conclusion

Reducing collected data to its scientific content brings significant advantages to users, the environment, and facilities. For users, the scientific outcome of the experiment is potentially improved due to better decision making, as well as simpler and more effective data analysis. The environment benefits from a decreased energy footprint in processing and storing the data, and in turn facilities profit from the reduced initial investment and operation costs, therefore improving operational sustainability.

To support this effort, we have developed an initial portfolio of data reduction methods. A few of these have been deployed and are already routinely used in operation at the European XFEL. We show that, by applying these, we avoided storage of about one third of the expected volume of processed data in 2023. Additionally, we have started developing technique-specific methods, some of which have been already employed for online data analysis and reduction. In parallel, the data system has been upgraded to include reduction points, and technique-specific data reduction methods have been investigated. To further develop and validate effective solutions for the latter, the considerable domain knowledge of our users is required.

Data reduction activities are a clear priority of European XFEL. Their development involves a multitude of actors, inside and outside the facility, which exemplifies the need for the diverse expertise intrinsic to data reduction.

Our strategy to maximize the impact of reduction activities is founded on increasing this synergy between facility experts and users. The deep understanding of infrastructure, software practices, detection systems, and methodologies that facility staff can provide needs to be paired with the knowledge of the scientific domain users bring in. Facility-side we aim at offering information, interpretable metrics, efficient interfaces and expertise that support our users in making effective decisions on the data reduction strategy for their experiments.

To conclude, in this paper we report on our vision for data reduction at the European XFEL, as well as selected preliminary results. We are convinced that the collaboration and co-design of reduction tools with our users will simultaneously ensure excellent scientific results and a sustainable operation. On this note, our early experience with users that volunteered to apply reduction methods at

the European XFEL resulted in critical feedback that is contributing to shape ideas and develop tools. The systematic implementation of streamlined data reduction methods as part of the data acquisition, analysis and storage can result in a paradigm shift in photon science concerning data handling and processing.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Author contributions

ES: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Project administration, Validation, Visualization, Writing—original draft, Writing—review and editing. PS: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. JaM: Formal Analysis, Investigation, Methodology, Writing—review and editing. DH: Methodology, Software, Writing—review and editing. DB: Methodology, Software, Writing—review and editing. JoM: Conceptualization, Formal Analysis, Investigation, Validation, Writing—review and editing. KA: Software, Writing—review and editing. RB: Resources, Writing—review and editing. IB: Validation, Writing—review and editing. JB: Validation, Writing—review and editing. UB: Validation, Writing—review and editing. CC: Software, Writing—review and editing. FD'A: Validation, Writing—review and editing. SE: Software, Writing—review and editing. HFa: Conceptualization, Funding acquisition, Project administration, Supervision, Visualization, Writing—review and editing. DFdL: Conceptualization, Visualization, Writing—original draft, Writing—review and editing. LM: Software, Writing—review and editing. HFi: Investigation, Writing—review and editing. GF: Software, Writing—review and editing. PG: Resources, Writing—review and editing. GG: Software, Writing—review and editing. JK: Validation, Writing—review and editing. AM: Resources, Writing—review and editing. TM: Software, Writing—review and editing. MiS: Validation, Writing—review and editing. MaS: Validation, Writing—review and editing. AS: Software, Writing—review and editing. JS-D: Validation, Writing—review and editing. MT: Validation, Writing—review and editing. OT: Formal Analysis, Investigation, Software, Writing—original draft, Writing—review and editing. JW: Formal Analysis, Investigation, Software, Writing—original draft, Writing—review and editing. SA: Conceptualization, Funding acquisition, Supervision, Writing—review and editing. SH: Conceptualization, Software, Writing—original draft, Writing—review and editing. KW: Conceptualization, Project administration, Supervision, Writing—original draft, Writing—review and editing. LG: Conceptualization, Funding acquisition, Project administration, Supervision, Visualization, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Data

reduction activities at the European XFEL have been partially funded through the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101004728.

Acknowledgments

The authors truly acknowledge the users of the European XFEL which are bravely assisting us in validating and pushing forward data reduction activities. We acknowledge the participants in the Horizon 2020 LEAPS-INNOV WP7, for the constructive discussions over the last few years. We acknowledge countless colleagues at the European XFEL and DESY that have contributed in different ways to this project, and in particular the participants of the working group on scientific data policy and data reduction. We particularly acknowledge the members of the European XFEL Detector Advisory Committee for the continuous support in this activity, and the members of the Data Analysis group at the European XFEL for the endless assistance and feedback. We are indebted to Thomas Tschentscher for the steady

encouragement, insightful discussions and critical reading of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Wang C, Steiner U, Sepe A. Synchrotron big data science. *Small* (2018) 14:1802291. doi:10.1002/smll.201802291
- Götz A, le Gall E, Konrad U, Kourousias G, Knodel O, Matalgah S, et al. LEAPS data strategy. *The Eur Phys J Plus* (2023) 138:617. doi:10.1140/epjp/s13360-023-04189-6
- Hatsui T, Graafsma H. X-ray imaging detectors for synchrotron and XFEL sources. *IUCr* (2015) 2:371–83. doi:10.1107/S205225251500010X
- Chapman HN. Fourth-generation light sources. *IUCr* (2023) 10:246–7. doi:10.1107/S2052252523003585
- Pellegrini C. The history of X-ray free-electron lasers. *The Eur Phys J H* (2012) 37: 659–708. doi:10.1140/epjh/e2012-20064-5
- Chapman HN. X-ray free-electron lasers for the structure and dynamics of macromolecules. *Annu Rev Biochem* (2019) 88:35–58. PMID: 30601681. doi:10.1146/annurev-biochem-013118-110744
- Blaj G, Caragiulo P, Carini G, Dragone A, Haller G, Hart P, et al. Future of ePix detectors for high repetition rate FELs. *AIP Conf Proc* (2016) 1741:040012. doi:10.1063/1.4952884
- Hart M, Angelsen C, Burge S, Coughlan J, Halsall R, Koch A, et al. Development of the LPD, a high dynamic range pixel detector for the European XFEL. In: *2012 IEEE nuclear science symposium and medical imaging conference record (NSS/MIC)* (2012). p. 534–7. doi:10.1109/NSSMIC.2012.6551165
- Veale M, Adkin P, Booker P, Coughlan J, French M, Hart M, et al. Characterisation of the high dynamic range Large Pixel Detector (LPD) and its use at X-ray free electron laser sources. *J Instrumentation* (2017) 12:P12003. doi:10.1088/1748-0221/12/12/P12003
- Allahgholi A, Becker J, Bianco L, Bradford R, Delfs A, Dinapoli R, et al. The adaptive gain integrating pixel detector. *J Instrumentation* (2016) 11:C02066. doi:10.1088/1748-0221/11/02/C02066
- Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Radiat* (2019) 26:74–82. doi:10.1107/S1600577518016077
- Porro M, Andricke L, Aschauer S, Castoldi A, Donato M, Engelke J, et al. The MiniSDD-Based 1-Mpixel Camera of the DSSC Project for the European XFEL. *IEEE Trans Nucl Sci* (2021) 68:1334–50. doi:10.1109/TNS.2021.3076602
- Johnson I, Bergamaschi A, Billich H, Cartier S, Dinapoli R, Greiffenberg D, et al. Eiger: a single-photon counting x-ray detector. *J Instrumentation* (2014) 9:C05032. doi:10.1088/1748-0221/9/05/C05032
- Hatsui T. *CITIU: a 17400 frames/s x-ray imaging detector*. Tenth Intl. Workshop on Pixel Detectors for Particles and Imaging (2022). <https://indico.cern.ch/event/829863/contributions/4479490/> (Accessed January 23, 2024).
- Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, et al. A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nat Photon* (2020) 14:391–7. doi:10.1038/s41566-020-0607-z
- Tschentscher T. Investigating ultrafast structural dynamics using high repetition rate x-ray FEL radiation at European XFEL. *Eur Phys J Plus* (2023) 138:274. doi:10.1140/epjp/s13360-023-03809-5
- Sekutowicz J, Ayvazyan V, Barlak M, Branlard J, Cichalewski W, Grabowski W, et al. Research and development towards duty factor upgrade of the European X-Ray Free Electron Laser linac. *Phys Rev ST Accel Beams* (2015) 18:050701. doi:10.1103/PhysRevSTAB.18.050701
- Schmidt P, Ahmed K, Danilevski C, Hammer D, Rosca R, Kluyver T, et al. Turning European XFEL raw data into user data. *Front Phys* (2024) 11. doi:10.3389/fphy.2023.1321524
- ATLAS level-1 trigger: technical design report. Tech. rep., CERN, Geneva (1998).
- Jenni P, Nessi M, Nordberg M, Smith K. ATLAS high-level trigger, data-acquisition and controls: technical Design Report. In *Tech. rep.* Geneva: CERN (2003).
- Bayatyan GL, Grigorian N, Khachatryan VG, Margarian AT, Sirunyan AM, Stepanian S, et al. CMS TriDAS project: technical design report. In: *The trigger systems. Tech. Rep.*, 1. CERN (2000).
- Thayer JB, Carini G, Kroeger W, O'Grady C, Perazzo A, Shankar M, et al. Building a data system for LCLS-II. In: *2017 IEEE nuclear science symposium and medical imaging conference (NSS/MIC)* (2017). p. 1–4. doi:10.1109/NSSMIC.2017.8533033
- Hadian-Jazi M, Sadri A, Barty A, Yefanov O, Galchenkova M, Oberthuer D, et al. Data reduction for serial crystallography using a robust peak finder. *J Appl Crystallogr* (2021) 54:1360–78. doi:10.1107/S1600576721007317
- Galchenkova M, Tolstikova A, Yefanov O, Chapman H. Data reduction in protein crystallography. *Acta Crystallogr Section A* (2022) 78:e266. doi:10.1107/S2053273222094517
- Nawaz S, Rahmani V, Pennicard D, Setty SPR, Klauedel B, Graafsma H. Explainable machine learning for diffraction patterns. *J Appl Crystallogr* (2023) 56: 1494–504. doi:10.1107/S1600576723007446
- Rahmani V, Nawaz S, Pennicard D, Setty SPR, Graafsma H. Data reduction for X-ray serial crystallography using machine learning. *J Appl Crystallogr* (2023) 56: 200–13. doi:10.1107/S1600576722011748
- Kieffer J, Petitdemange S, Vincent T. Real-time diffraction computed tomography data reduction. *J Synchrotron Radiat* (2018) 25:612–7. doi:10.1107/S1600577518000607
- Kieffer J, Brennich M, Florial JB, Oscarsson M, De Maria Antolinos A, Tully M, et al. New data analysis for BioSAXS at the ESRF. *J Synchrotron Radiat* (2022) 29: 1318–28. doi:10.1107/S1600577522007238
- Kieffer J, Coquelle N, Santoni G, Basu S, Debionne S, Homs A, et al. Real-time pre-processing of serial crystallography. *Acta Crystallogr Section A* (2022) 78:e263. doi:10.1107/S2053273222094530
- Zhang Q, Dufresne EM, Nakaye Y, Jemian PR, Sakumura T, Sakuma Y, et al. 20 μ s-resolved high-throughput X-ray photon correlation spectroscopy on a 500k pixel detector enabled by data-management workflow. *J Synchrotron Radiat* (2021) 28: 259–65. doi:10.1107/S1600577520014319
- Zhao K, Di S, Lian X, Li S, Tao D, Bessac J, et al. SDRbench: scientific data reduction benchmark for lossy compressors. In: *2020 IEEE international conference on big data (big data)*. Los Alamitos, CA, USA: IEEE Computer Society (2020). p. 2716–24. doi:10.1109/BigData50022.2020.9378449

32. Wang C, Florin E, Chang HY, Thayer J, Yoon CH. SpeckleNN: a unified embedding for real-time speckle pattern classification in X-ray single-particle imaging with limited labeled examples. *IUCrJ* (2023) 10:568–78. doi:10.1107/S205252523006115
33. Sun Y, Brockhauser S, Hegedűs P, Plückthun C, Gelisio L, Ferreira de Lima DE. Application of self-supervised approaches to the classification of X-ray diffraction spectra during phase transitions. *Scientific Rep* (2023) 13:9370. doi:10.1038/s41598-023-36456-y
34. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* (2016) 3:160018. doi:10.1038/sdata.2016.18
35. European XFEL. *User documentation for data reduction at European XFEL* (2024). Available at: <https://rtd.xfel.eu/docs/data-reduction-user-documentation/en/latest/> (Accessed January 23, 2024).
36. Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* (2000) 406:752–7. doi:10.1038/35021099
37. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, et al. Femtosecond X-ray protein nanocrystallography. *Nature* (2011) 470:73–7. doi:10.1038/nature09750
38. Bogan MJ, Benner WH, Boutet S, Rohner U, Frank M, Barty A, et al. Single particle X-ray diffractive imaging. *Nano Lett* (2008) 8:310–6. doi:10.1021/nl072728k
39. Graewert MA, Svergun DI. Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). *Curr Opin Struct Biol* (2013) 23:748–54. doi:10.1016/j.sbi.2013.06.007
40. Lehmkuhler F, Dallari F, Jain A, Sikorski M, Möller J, Frenzel L, et al. Emergence of anomalous dynamics in soft matter probed at the European XFEL. *Proc Natl Acad Sci* (2020) 117:24110–6. doi:10.1073/pnas.2003337117
41. Reiser M, Girelli A, Ragulskaya A, Das S, Berkowicz S, Bin M, et al. Resolving molecular diffusion and aggregation of antibody proteins with megahertz X-ray free-electron laser pulses. *Nat Commun* (2022) 13:5528. doi:10.1038/s41467-022-33154-7
42. Altarelli M, Kurta RP, Vartanyants IA. X-ray cross-correlation analysis and local symmetries of disordered systems: general theory. *Phys Rev B* (2010) 82:104207. doi:10.1103/PhysRevB.82.104207
43. Malka J, Aplin S, Boukhelef D, Dietrich S, Filippakopoulos K, Gasthuber M, et al. Data management infrastructure for European XFEL. In: *Proceedings of ICALEPCS2023*; Geneva, Switzerland: JACoW Publishing (2024). p. 952–957. doi:10.18429/JACoW-ICALEPCS2023-WE1BCO02
44. European XFEL. *Metadata catalogue* (2023). Available at: <https://in.xfel.eu/metadata> (Accessed January 23, 2024).
45. Schmuck F, Haskin R. GPFS: a shared-disk file system for large computing clusters. In: *Proceedings of the 1st USENIX conference on file and storage technologies (USA: USENIX association)* (2002). FAST '02, 19–es.
46. Deutsches Elektronen-Synchrotron. *Maxwell cluster* (2023). <https://confluence.desy.de/display/MXW/Maxwell+Cluster> (Accessed January 23, 2024).
47. Ernst M, Fuhrmann P, Gasthuber M, Mkrtychyan T, Waldman C. dCache, a distributed storage data caching system. In: *Proceedings of computing in high energy physics*. Beijing (China): Science Press (2001). Available from China Nuclear Information Centre.
48. Hauf S, Heisen B, Aplin S, Beg M, Bergemann M, Bondar V, et al. The Karabo distributed control system. *J Synchrotron Radiat* (2019) 26:1448–61. doi:10.1107/S1600577519006696
49. Göries D, Ehsan W, Flucke G, Annakkappala N, Bondar V, Costa R, et al. The Karabo SCADA system at the European XFEL. *Synchrotron Radiation News* (2023). 36, 40–46. doi:10.1080/08940886.2023.2277650
50. Esenov S, Wrona K, Youngman C. *Technical design report: European XFEL DAQ and DM computing – 2009 public version*. Schenefeld, Germany: Tech. Rep. XFEL/UE TR-2009-001, European XFEL (2009). doi:10.3204/XFEL/UE/TR-2009-001
51. Boukhelef D, Szuba J, Wrona K, Youngman C. Software development for high speed data recording and processing. In: *Proceedings of ICALEPCS2013*; Geneva, Switzerland: JACoW Publishing (2014). p. 665–668.
52. European XFEL. *EXDF-tools: tools to work with EXDF HDF5 files* (2023). <https://git.xfel.eu/dataAnalysis/exdf-tools> (Accessed January 23, 2024).
53. Madsen A, Hallmann J, Ansaldi G, Roth T, Lu W, Kim C, et al. Materials Imaging and Dynamics (MID) instrument at the European X-ray Free-Electron Laser Facility. *J Synchrotron Radiat* (2021) 28:637–49. doi:10.1107/S1600577521001302
54. Mancuso AP, Aquila A, Batchelor L, Bean RJ, Bielecki J, Borchers G, et al. The Single Particles, Clusters and Biomolecules and Serial Femtosecond Crystallography instrument of the European XFEL: initial installation. *J Synchrotron Radiat* (2019) 26: 660–76. doi:10.1107/S1600577519003308
55. Koziol Q. *HDF5*. Boston, MA: Springer US (2011). p. 827–33. doi:10.1007/978-0-387-09766-4_44
56. Deutsch LP. *DEFLATE compressed data format specification version 1.3* (1996). doi:10.17487/RFC1951
57. Zastra U, Appel K, Baetz C, Baehr O, Batchelor L, Berghäuser A, et al. The High Energy Density Scientific Instrument at the European XFEL. *J Synchrotron Radiat* (2021) 28:1393–416. doi:10.1107/S1600577521007335
58. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* (1948) 27: 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
59. Kieffer J, Valls V, Blanc N, Hennig C. New tools for calibrating diffraction setups. *J Synchrotron Radiat* (2020) 27:558–66. doi:10.1107/S1600577520000776
60. White T, Kirian R, Martin A, Aquila A, Nass K, Barty A, et al. CrystFEL: a software suite for snapshot serial crystallography. *J Appl Cryst* (2012) 45:335–41. doi:10.1107/S0021889812002312
61. Brewster AS, Waterman DG, Parkhurst JM, Gildea RJ, Young ID, O'Riordan LJ, et al. Improving signal strength in serial crystallography with DIALS geometry refinement. *Acta Crystallogr Section D* (2018) 74:877–94. doi:10.1107/S2059798318009191
62. Turkot O, Dall'Antonia F, Bean RJ, E J, Fangohr H, Ferreira de Lima DE, et al. Extra-xviz: a tool to streamline serial femtosecond crystallography workflows at European XFEL. *Crystals* (2023) 13:1533. doi:10.3390/cryst13111533
63. Turkot O, Dall'Antonia F, Bean RJ, E J, Fangohr H, Ferreira de Lima DE, et al. Towards automated analysis of serial crystallography data at the European XFEL. In: Tschentscher T, Patthey L, Tiedtke K, Zangrando M, editors. *X-ray free-electron lasers: advances in source development and instrumentation VI*, 12581. Bellingham, WA: International Society for Optics and Photonics (2023). p. 125810M. doi:10.1117/12.2669569
64. Barty A, Kirian RA, Maia FRNC, Hantke M, Yoon CH, White TA, et al. *Cheetah*: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J Appl Crystallogr* (2014) 47:1118–31. doi:10.1107/S1600576714007626
65. Ullrich J, Moshhammer R, Dorn A, Dörner R, Schmidt LPH, Schmidt-Böcking H. Recoil-ion and electron momentum spectroscopy: reaction-microscopes. *Rep Prog Phys* (2003) 66:1463–545. doi:10.1088/0034-4885/66/9/203
66. Boll R, Schäfer JM, Richard B, Fehre K, Kastirke G, Jurek Z, et al. X-ray multiphoton-induced coulomb explosion images complex single molecules. *Nat Phys* (2022) 18:423–8. doi:10.1038/s41567-022-01507-0
67. Dallari F, Reiser M, Lokteva I, Jain A, Möller J, Scholz M, et al. Analysis strategies for MHz XPCS at the European XFEL. *Appl Sci* (2021) 11:8037. doi:10.3390/app11178037
68. Ferreira de Lima D, Davtyan A, Turkot O, Yefanov O, White T, Galchenkova M, et al. Automatic online data analysis optimization: application to serial femtosecond crystallography. In: *preparation* (2024).
69. European XFEL. *Quality of data services* (2025). Available at: https://www.xfel.eu/sites/sites_custom/site_xfel/content/e51499/e141242/e141245/xfel_file2334455/Quality_of_data_services_01.2024_draft_eng.pdf (Accessed January 23, 2024).



OPEN ACCESS

EDITED BY

Lodovico Ratti,
University of Pavia, Italy

REVIEWED BY

Carlo Fiorini,
Polytechnic University of Milan, Italy
David Pennicard,
Helmholtz Association of German Research
Centres (HZ), Germany
Dionisio Doering,
Stanford University, United States

*CORRESPONDENCE

Viktoria Hinger,
✉ viktoria.hinger@psi.ch

RECEIVED 07 December 2023

ACCEPTED 07 February 2024

PUBLISHED 28 February 2024

CITATION

Hinger V, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, Brückner M, Carulla M, Centis Vignali M, Dinapoli R, Ebner S, Ficorella F, Fröjd E, Greiffenberg D, Hammad Ali O, Hasanaj S, Heymes J, King T, Kozłowski P, Lopez-Cuenca C, Mezza D, Mozzanica A, Moustakas K, Paternoster G, Paton KA, Ronchin S, Ruder C, Schmitt B, Sieberer P, Thattil D, Xie X and Zhang J (2024), Resolving soft X-ray photons with a high-rate hybrid pixel detector.
Front. Phys. 12:1352134.
doi: 10.3389/fphy.2024.1352134

COPYRIGHT

© 2024 Hinger, Barten, Baruffaldi, Bergamaschi, Borghi, Boscardin, Brückner, Carulla, Centis Vignali, Dinapoli, Ebner, Ficorella, Fröjd, Greiffenberg, Hammad Ali, Hasanaj, Heymes, King, Kozłowski, Lopez-Cuenca, Mezza, Mozzanica, Moustakas, Paternoster, Paton, Ronchin, Ruder, Schmitt, Sieberer, Thattil, Xie and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Resolving soft X-ray photons with a high-rate hybrid pixel detector

Viktoria Hinger^{1*}, Rebecca Barten¹, Filippo Baruffaldi¹, Anna Bergamaschi¹, Giacomo Borghi², Maurizio Boscardin², Martin Brückner¹, Maria Carulla¹, Matteo Centis Vignali², Roberto Dinapoli¹, Simon Ebner¹, Francesco Ficorella², Erik Fröjd¹, Dominic Greiffenberg¹, Omar Hammad Ali², Shqipe Hasanaj¹, Julian Heymes¹, Thomas King¹, Pawel Kozłowski¹, Carlos Lopez-Cuenca¹, Davide Mezza¹, Aldo Mozzanica¹, Konstantinos Moustakas¹, Giovanni Paternoster², Kirsty A. Paton¹, Sabina Ronchin², Christian Ruder¹, Bernd Schmitt¹, Patrick Sieberer¹, Dhanya Thattil¹, Xiangyu Xie¹ and Jianguo Zhang¹

¹Paul Scherrer Institut, Villigen, Switzerland, ²Fondazione Bruno Kessler, Trento, Italy

Due to their high frame rates and dynamic range, large area coverage, and high signal-to-noise ratio, hybrid silicon pixel detectors are an established standard for photon science applications at X-ray energies between 2 keV and 20 keV. These properties also make hybrid detectors interesting for experiments with soft X-rays between 200 eV and 2 keV. In this energy range, however, standard hybrid detectors are limited by the quantum efficiency of the sensor and the noise of the readout electronics. These limitations can be overcome by utilizing inverse Low-Gain Avalanche Diode (iLGAD) sensors with an optimized X-ray entrance window. We have developed and characterized a prototype soft X-ray iLGAD sensor bonded to the charge integrating 75 μm pixel JUNGFRUAU chip. Cooled to -22°C , the system multiplication factor of the signal generated by an impinging photon is ≥ 11 . With this gain, the effective equivalent noise charge of the system is ≤ 5.5 electrons root-mean-square at a 5 μs integration time. We show that by cooling the system below -50°C , single photon resolution at 200 eV becomes feasible with a signal-to-noise ratio better than 5.

KEYWORDS

hybrid detector, instrumentation for FEL, synchrotron radiation, LGAD, X-ray detector, low noise detector

1 Introduction

Photon science at soft X-ray energies (~ 200 eV–2 keV) at Synchrotron Radiation (SR) sources and Free-Electron Lasers (FELs) fosters a diverse spectrum of research. Experiments exploit the presence of the K and L-edges of light elements and 3d transition metals to study, for example, photosynthetic water splitting [1] and fundamental excitations that govern phenomena such as magnetic ordering [2] and high-temperature superconductivity [3] in correlated materials. For organic samples, experimental techniques that access the energy range between the carbon and oxygen K-edges, the so-called “water-window” (277–525 eV), are particularly relevant. However, while the technological capabilities of next-generation

X-ray sources are evolving at a rapid pace, detector systems are struggling to match the needs of experimental stations operating in the soft X-ray domain.

For many applications in the range of hard X-rays between ~ 2 – 20 keV, silicon hybrid detectors are the state-of-the-art. The high data quality and count rates of single photon counting systems such as PILATUS [4] and EIGER [5] allowed unprecedented advancements in terms of throughput and sensitivity for experiments at SR facilities. Charge integrating hybrid detectors that implement dynamic gain switching on the pixel level such as AGIPD [6], ePIX [7], and JUNGFRÄU [8], on the other hand, provide the high sensitivity and dynamic range necessary for FEL-based diffraction, spectroscopy, and imaging experiments.

In the soft X-ray range, it has thus far been difficult to exploit hybrid detector technology. The reasons for this are twofold. First, because of the shallow absorption depth of low-energy X-ray photons, a significant percentage of photons are absorbed in the non-sensitive entrance window of the sensor, and the such generated charge does not contribute to the signal. Second, if a soft X-ray photon passes through the entrance window and generates a charge in the sensitive detector volume, the resulting signal is small compared to the electronic noise of the detector. As a result, a single photon cannot be discriminated from the noise floor.

Instead of standard hybrid detectors, scientists mainly use Charge Coupled Device (CCD) cameras [9–11] and CMOS image sensors [12–14] for soft X-ray detection. These systems can provide low noise (i.e., better than two electrons r.m.s. [15]) and high spatial resolution (down to $5\text{ }\mu\text{m}$ pixel size [16]) but are limited in terms of available detector area and frame rate capabilities. With the DSSC imager [17, 18], on the other hand, a system based on hybrid technology exists for the soft X-ray range. The detector is tailored to the high-rate bunch structure of EuXFEL but is limited in terms of spatial resolution. In light of the trend toward ultra-fast, high-throughput experiments at next-generation X-ray sources, the need for soft X-ray detector systems that match the sensitivity, spatial resolution, frame rate, dynamic range, and available detection area of hard X-ray hybrid detectors becomes especially apparent.

The recent development of inverse Low-Gain Avalanche Diode (iLGAD) sensors optimized for soft X-rays [19] opens the prospect for high-resolution hybrid detectors operating in this energy range. This approach addresses the two limitations of current hybrid technology. A thin entrance window maximizes the number of low-energy photons reaching the sensitive sensor volume, and the intrinsic gain of the iLGAD amplifies the signal of soft X-ray photons, allowing discrimination from the electronic noise floor of the hybrid detector.

In this work, we assess the capabilities of a soft X-ray hybrid detector based on the charge-integrating JUNGFRÄU readout chip. Owing to their low noise and high dynamic range, JUNGFRÄU systems are in operation at many facilities around the world, supporting a large variety of applications at FELs and high-flux SR experiments. Consequently, a soft X-ray version of the detector is of particular interest for applications at next-generation high-brilliance coherent light sources and high-rate FEL facilities. We present performance studies of a first JUNGFRÄU-iLGAD prototype. The detector response to low-energy photons was investigated using X-ray fluorescence measurements with

different metal targets. Taking these results as a baseline, we discuss the prospects of JUNGFRÄU-iLGAD systems for soft X-ray applications.

2 Materials and methods

2.1 The JUNGFRÄU detector

The hybrid silicon pixel detector JUNGFRÄU [8] was primarily designed for the high-flux, pulsed conditions at FELs but has since extended its use also to SR sources [20, 21]. The readout chip provides a maximum frame rate of 2.2 kHz for continuous image readout and combines a charge-integrating architecture with three linear, dynamically switching gains per pixel. One chip comprises 256×256 pixels with a pixel size of $75 \times 75\text{ }\mu\text{m}^2$. A JUNGFRÄU detector equipped with standard sensors and version 1.0 of the readout chip resolves single photons down to ~ 1.5 keV with an average noise level of 52 electrons (e^-) Equivalent Noise Charge (ENC) at a $5\text{ }\mu\text{s}$ integration time in low-noise high-gain mode (HG0). This corresponds to a Signal-to-Noise Ratio (SNR) of ~ 8 at 1.5 keV. The noise level stays well below the Poisson limit over the full dynamic range of 10^4 photons at 12 keV [8].

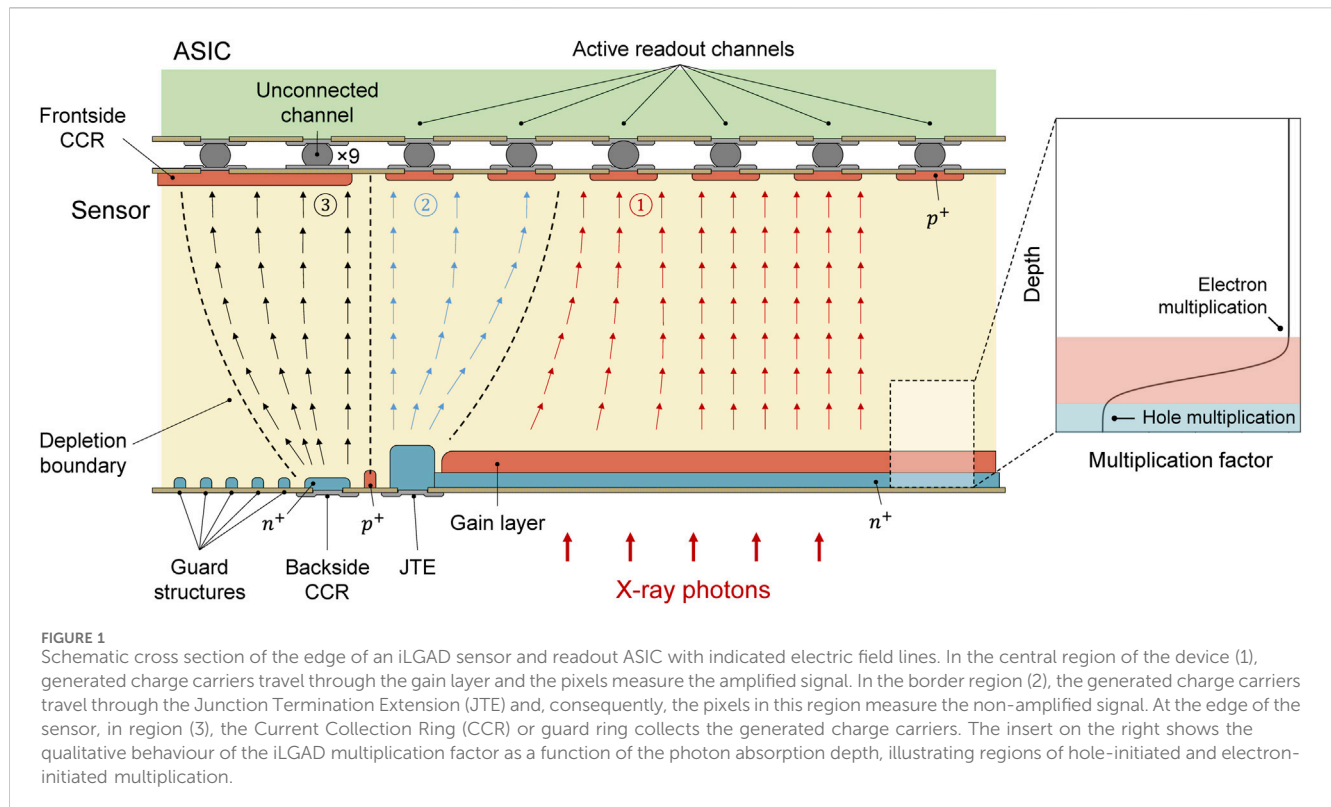
The most recent version of the readout chip, labelled JUNGFRÄU 1.1, includes additional measures to reduce the noise to facilitate soft X-ray detection [22]. It achieves an average noise of $34\text{ }e^-$ r.m.s. in HG0 at a $5\text{ }\mu\text{s}$ integration time [22]—a 35% reduction compared to version 1.0.

2.2 Inverse Low Gain Avalanche Diodes (iLGADs)

LGAD sensors [23] were initially developed for applications in high energy particle tracking that require precision timing in the range of ~ 10 ps. The sensors include an additional implant of high doping concentration ($\sim 10^{16}$ – 10^{17} cm^{-3}) below each sensor electrode. The resulting high electric field across the p - n junction promotes charge multiplication, amplifying the signal of an incoming particle or photon by a factor of 5–20 and generating an output signal proportional to the energy deposited in the sensor. This intrinsic signal amplification in the gain layer improves the SNR compared to a sensor without amplification.

However, standard LGADs for high energy physics cannot be used for photon science applications. This is due to two main reasons. To achieve better temporal resolution, the active sensors are usually thin (i.e., typically $50\text{ }\mu\text{m}$) and therefore require a non-active carrier substrate on the back, which would block X-rays. Moreover, the fill factor of standard LGADs is less than 100% [24] because of the edge termination required to prevent premature voltage breakdown [25].

To make LGAD sensors suitable for soft X-ray detection, a number of design differences are required with respect to conventional LGADs for high energy physics. Inverse LGAD (iLGAD) sensors developed for soft X-rays [19] (Figure 1) are $275\text{ }\mu\text{m}$ thick and do not require any carrier substrate. They feature an optimized thin entrance window to improve the quantum efficiency at low photon energies. The entrance window



includes a surface passivation with thin SiO_2 and Si_3N_4 layers [26]. Moreover, the gain layer is continuous and located at the back of the sensor where X-ray photons enter (i.e., “inverse” to the conventional LGAD design, where the gain layer is segmented and located on the pixelated side of the sensor). In contrast to the conventional LGAD design, the inverse design provides a 100% fill factor and, with sufficiently small pixels, interpolation is possible, enabling a spatial resolution on the order of a micron [27]. This is required for several applications, including Resonant Inelastic X-ray Scattering (RIXS) experiments [28] at FELs and SR sources.

Because the gain layer of the iLGAD is located at the entrance window side, the multiplication factor depends on the photon absorption depth. This is illustrated in Figure 1. If a photon is absorbed before the gain layer, holes drift through the gain layer and initiate charge multiplication. In contrast, if the photon is absorbed beyond the gain layer, the multiplication is initiated by electrons. For photons absorbed within the gain layer, both electrons and holes initiate multiplication. The multiplication factors for electrons and holes differ depending on the gain layer design [29]. The first iLGAD R&D batch devised by the Paul Scherrer Institute (PSI) in collaboration with Fondazione Bruno Kessler (FBK, Trento, Italy) contains iLGADs with optimized thin entrance window and different gain layer designs labelled “standard,” “shallow,” and “ultra-shallow.” For the “standard” gain layer design¹ investigated

in this work, the hole multiplication factor is about 25% of the electron multiplication factor, as has been shown with simulations [19] and also reported in [29] with measurements. As a consequence of this design, hole multiplication is the dominant effect for photons with energies below about 500 eV.

2.2.1 JUNGFRAU-iLGAD prototype

For this work, we characterized the performance of a JUNGFRAU-iLGAD prototype. The iLGAD sensor was fabricated by FBK as part of the first iLGAD R&D batch targeting iLGADs for soft X-ray detection, devised in collaboration with PSI. The total size of the sensor is $2 \times 2 \text{ cm}^2$. It is made up of $75 \times 75 \mu\text{m}^2$ pixels arranged in the standard 256×256 JUNGFRAU ASIC grid. On all sides of the grid, nine pixels are connected to the surrounding guard ring to collect the leakage current from the guard ring at the back and from the scribeline of the sensor (see Figure 1). This results in a total number of 238×238 active pixels. For the present study, a sensor with standard gain layer configuration was chosen because this design maximises the iLGAD multiplication factor. The sensor was bonded to a JUNGFRAU 1.1 readout chip (Figure 2). The quantum efficiency of the investigated iLGAD sensor and its gain layer design have been measured and reported by Liguori et al. [29].

2.3 Characterization measurements

The sensor-chip assembly was operated at a sensor bias voltage of 200 V and illuminated with fluorescence photons at a lab-based X-ray source. Four different secondary metal targets

¹ The standard variation of the gain layer features a doping profile similar to that of LGADs used for high energy particle physics, which is not specifically optimized for soft X-ray detection.

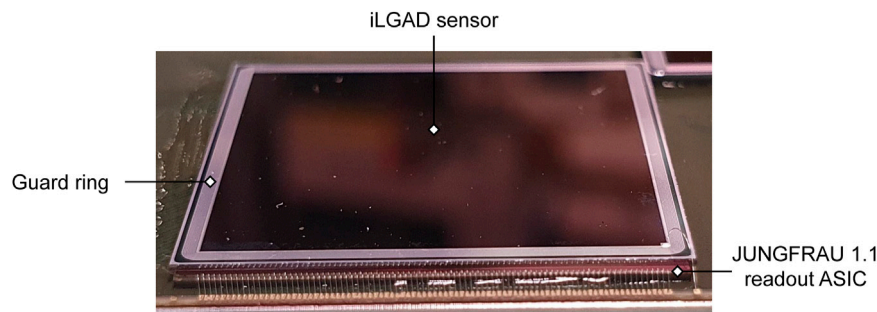


FIGURE 2
2 × 2 cm² iLGAD sensor bonded to a JUNGFRAU 1.1 readout chip.

were used to produce the fluorescence photons, namely, aluminium (K_{α} : 1.5 keV), titanium (K_{α} : 4.5 keV), chromium (K_{α} : 5.4 keV), and iron (K_{α} : 6.4 keV). These targets were chosen based on availability and considering the limitations of the lab-based setup to produce sufficient photon fluxes at lower energies. With this setup, the photon occupancy per single frame was at a level of 0.5%–4% for the investigated energies between 1.5 and 6.4 keV. The temperature of the system was regulated by means of a chiller connected to the front-end module with liquid coolant. The JUNGFRAU-iLGAD module was placed in a vacuum chamber flushed with nitrogen to a constant pressure of 10 mbar to prevent condensation at low temperatures and to minimize scattering and absorption of low-energy fluorescence photons before reaching the sensor.

Fluorescence spectra with a minimum of 10,000 images per acquisition point were taken for each metal target at eight chiller temperatures between -22°C and 15°C and at five different integration times between $5\text{ }\mu\text{s}$ and $100\text{ }\mu\text{s}$. Datasets at a $10\text{ }\mu\text{s}$ integration time were acquired with higher statistics of $> 200,000$ images per temperature point. For each configuration, 1,000 dark images were acquired ahead of illumination with X-rays. The dark images are used to determine the signal offset (pedestal) and noise of each pixel. To account for the fact that the charge of a photon hit can be shared between neighboring pixels, the data were analysed using the cluster finding algorithm detailed in [30]. For all cases, the total charge generated by a photon hit is contained within a cluster of 2×2 pixels. Consequently, for the following analyses, we use the 2×2 clustered spectra.

3 Results

3.1 Gain

The signal response of each detector pixel to a photon hit of a given energy (i.e., the total gain g_{tot} of that particular pixel) is determined by two components: the conversion gain of the JUNGFRAU readout electronics g_{JF} and the multiplication factor M of the iLGAD. From these two components, g_{tot} is given as (Eq. 1)

$$g_{\text{tot}} = g_{\text{JF}} \times M. \quad (1)$$

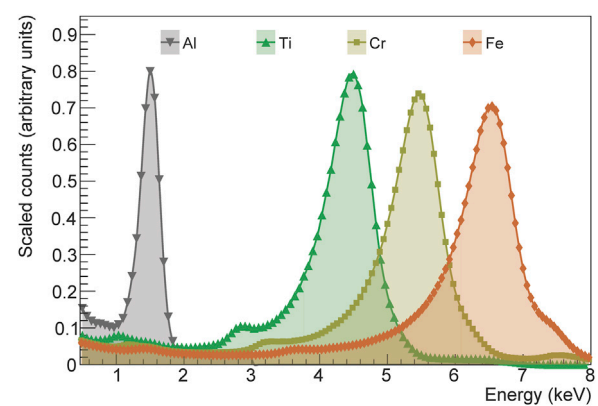


FIGURE 3
Energy calibrated fluorescence spectra (2×2 clusters) of four different metal targets (Al- K_{α} : 1.5 keV, Ti- K_{α} : 4.5 keV, Cr- K_{α} : 5.4 keV, Fe- K_{α} : 6.4 keV) acquired with the JUNGFRAU-iLGAD prototype at -22°C and $10\text{ }\mu\text{s}$ integration time. The energy calibration is performed on the pixel level with the Al- K_{α} peak as absolute reference point. The plot shows cumulative spectra across a region of 80×50 pixels.

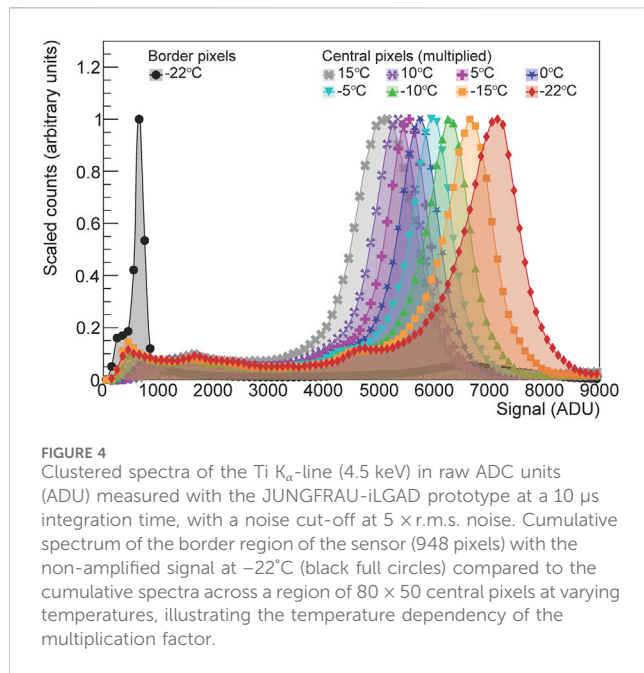
We can determine the total gain of each pixel from a Gaussian fit to the measured spectra of the respective metal K_{α} -line as (Eq. 2)

$$g_{\text{tot}} = \frac{\mu_{K_{\alpha}}}{E_{K_{\alpha}}}. \quad (2)$$

Here, $\mu_{K_{\alpha}}$ refers to the mean of the Gaussian in ADC units (ADU), and $E_{K_{\alpha}}$ is the nominal energy of the K_{α} -line. The pixel-wise calibrated fluorescence spectra obtained with the JUNGFRAU-iLGAD prototype are shown in Figure 3.

For the JUNGFRAU-iLGAD, we can separate the two components of the total gain by comparing the signal spectra of the pixels bordering the guard ring with the spectra of the central pixels. Because the gain layer terminates at the pixels next to the guard ring (see Figure 1), charge carriers generated in that region do not travel through the gain layer and the pixels detect the non-amplified signal.

Figure 4 illustrates the difference between the non-amplified spectrum of the border pixels and the amplified spectra of the central pixels at different temperatures. From the border pixel spectrum, we can calculate the conversion gain g_{JF} of the



JUNGFRU readout electronics without iLGAD multiplication equivalent to Eq. 2 from a Gaussian fit to the signal peak. We chose the Fe- K_{α} peak as reference for the gain calculation because at this energy the separation of the signal peak from the noise is sufficiently high also for the non-amplified border pixel spectra at all temperatures. However, where gain calculation using the lower energy peaks as reference is possible, the obtained values agree within statistical uncertainties. Across all border pixels, the such calculated conversion gain is constant with an average value of $g_{\text{F}} = 146.4 \pm 1.3 \text{ ADU/keV}$ for all investigated temperatures. The typical variations of the conversion gain across the full area of a JUNGFRU module, found from measurements with standard silicon sensors, are in the range of 3.5% [31]. These pixel-to-pixel

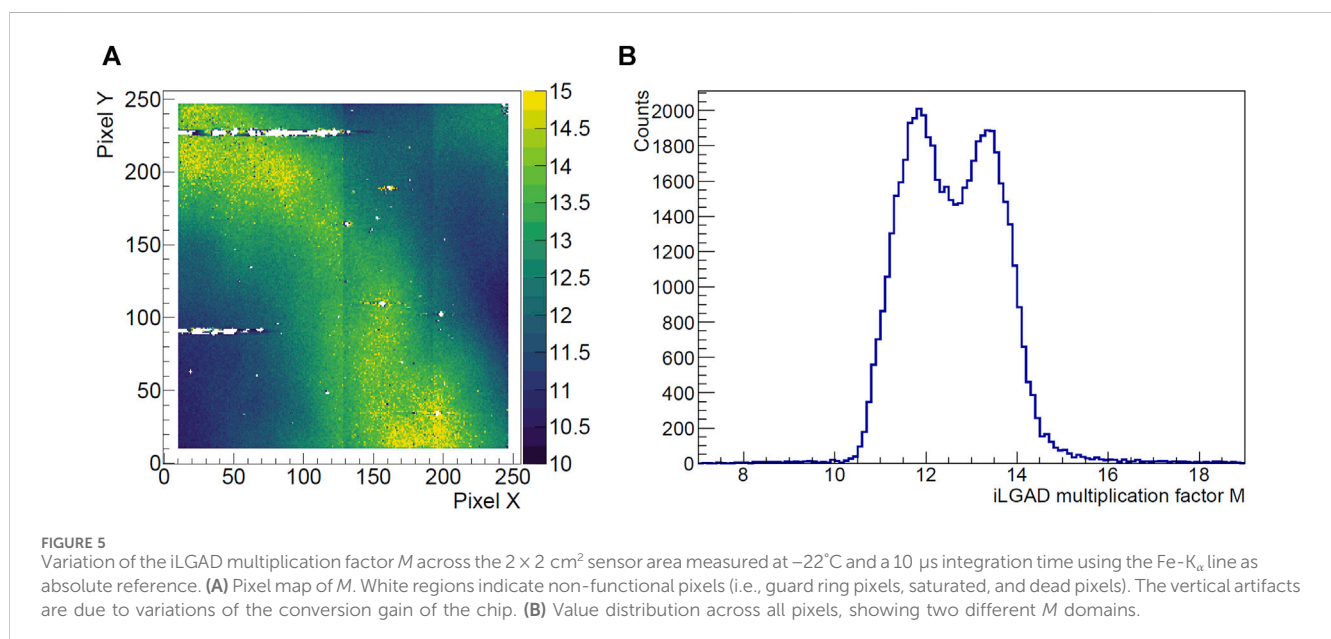
variations cannot be accounted for in the separation of conversion gain and iLGAD multiplication factor performed in this work.

To determine the iLGAD multiplication factor M , we compare the Fe- K_{α} peak positions of the border pixel spectrum and the signal spectra for each central pixel. From these measurements, M is given as (Eq. 3)

$$M = \frac{\mu_{\text{central}}}{\hat{\mu}_{\text{border}}}, \quad (3)$$

where μ_{central} is the per-pixel amplified peak position and $\hat{\mu}_{\text{border}}$ is the position of the non-amplified peak in the cumulative border pixel spectrum. The value of M varies across the area of the iLGAD (Figure 5A). In particular, we can identify two domains where the average values of M differ by about 10% (Figure 5B) with an approximate symmetry along one sensor diagonal. The dispersion of this distribution is much larger than the 3.5% variation of the JUNGFRU conversion gain measured for standard sensors, and the magnitude of the difference between domains is too large to be congruent with temperature variations across the sensor. Consequently, we attribute the variations of M to variations in the doping composition of the gain layer.

The temperature trend of the multiplication factor is shown in Figure 6. We model the temperature dependency of M using a customized program that solves the one-dimensional Poisson equation, incorporating the doping profiles of the n^+ layer, the gain layer, and the doping concentration of the silicon bulk. Boundary conditions are set at 0 and 200 V, representing the bias voltages applied to the iLGAD sensor. This simulation determines the electric field distribution within the sensor, particularly in the charge multiplication layer. Subsequently, the multiplication factor M , initiated by electrons traveling through the gain layer, is calculated at various temperatures. The calculation utilizes the impact-ionization coefficient for electrons and holes, which are functions of the electric field distribution, employing the Okuto-Crowell model with the optimized



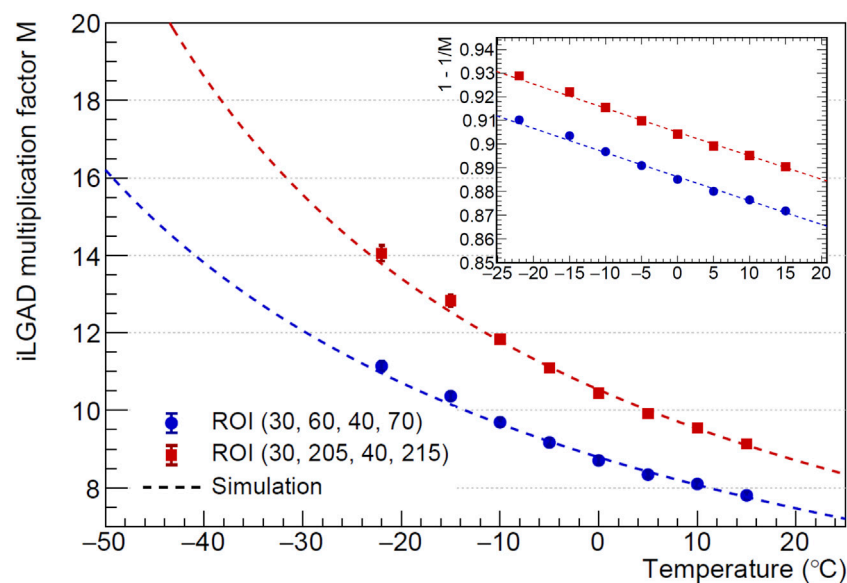


FIGURE 6

iLGAD multiplication factor M as a function of temperature. Comparison of measured values and simulation. The insert shows the linear behaviour of $1 - \frac{1}{M}$. The measured data points correspond to the mean values of M within two 11×11 pixel ROI (x_{\min} , y_{\min} , x_{\max} , y_{\max}). Error bars represent the 1σ dispersion across all pixels of the ROI. For most data points, the dispersion is smaller than the plotted marker size.

parameters extracted by Currás Rivera and Moll [32]. The doping profile of the gain layer is obtained through Silvaco² process simulation, and its integrated dose is tuned until the simulated M aligns with the measurement results. The data points shown in Figure 6 correspond to the mean values of M within two different Regions Of Interest (ROI) of 11×11 pixels. We chose these small ROI to confine the modelling of M to regions with an approximately uniform distribution. In the simulation, we model the gain variations observed in the measurements by incorporating a difference of 1.25% integrated dose in the gain layer for the two ROI. The depth profile of the gain layer used in the simulation is the same for both ROI.

3.2 Leakage current

The leakage current of the iLGAD sensor represents an important contribution to the overall detector noise. Because of the presence of the gain layer, the leakage current is also multiplied. Moreover, in a charge integrating hybrid detector such as JUNGFRUAU, in addition to the photon signal, the leakage current is integrated. Consequently, the SNR decreases with increasing integration time. Particularly for SR applications, which operate at long integration times to maximize the duty cycle of the detector, the leakage current becomes a performance limiting factor.

We determine the leakage current from dark images acquired at different temperatures and integration times. For each pixel, at each temperature, we fit a linear function to the average dark signal of

1,000 images as a function of integration time. The leakage current per pixel is then calculated from the slope of the fit for each individual pixel and the conversion gain of the JUNGFRUAU electronics g_{IF} from Eq. 4:

$$I \text{ (A)} = \frac{\text{slope (ADU/ns)}}{g_{\text{IF}} \text{ (ADU/keV)}} \times \frac{1000 \text{ (eV/keV)}}{3.6 \text{ eV}} \times q_0 \times 10^9 \text{ (ns/s)}. \quad (4)$$

Here, 3.6 eV is the average energy required to produce an electron-hole pair in silicon and q_0 is the elementary charge.

Figure 7 shows the temperature dependency of the leakage current for the same two ROI used in Section 3.1 and the average across the full sensor area. At 15°C, the average leakage current per pixel is 134 ± 18 pA and decreases to 5 ± 1 pA at -22°C. We note that the average leakage current across the full sensor area is higher than for both investigated ROI for most temperatures. This effect is due to the impact of hot pixels on the full-sensor average and is mitigated at lower temperatures. Plotted on a logarithmic scale, the leakage current displays a linear dependency on $1/T$. For a temperature difference of $\sim 20^\circ\text{C}$, the leakage current changes by one order of magnitude. When lowering the operating temperature, the hybrid detectors using iLGAD sensors benefit from the lower leakage current and, consequently, lower noise and increased gain and, thus, a better SNR. Furthermore, we find that decreasing the operating temperature of the detector improves the pixel yield. For the investigated prototype, the average amount of non-functional pixels is 1.5% at 15°C decreasing to 0.8% at -22°C.

3.3 Effective detector noise

From the pedestal fluctuations σ for each pixel, we calculate the effective Equivalent Noise Charge (ENC_{eff}) in e^- as (Eq. 5)

² <https://silvaco.com/tcad>

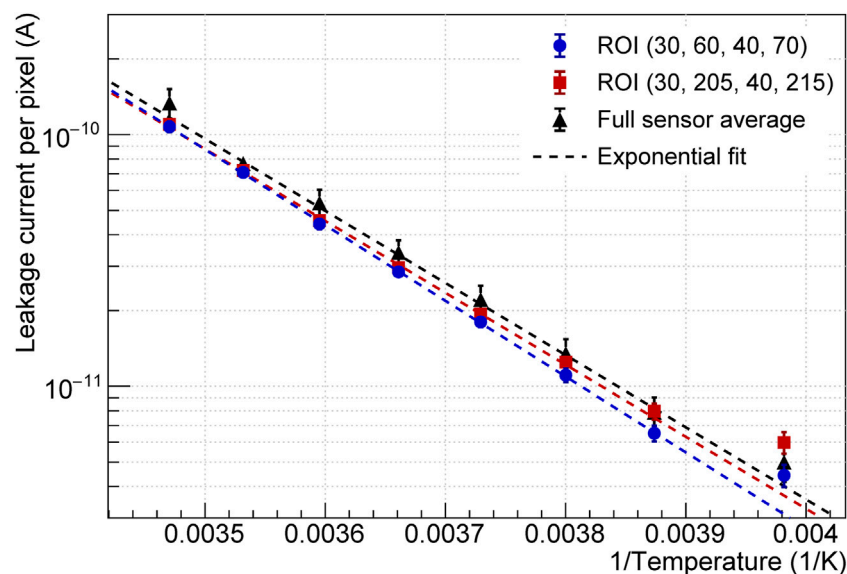


FIGURE 7

Leakage current of the JUNGFRÄU-iLGAD as a function of the reciprocal of the temperature. Comparison of measured values and exponential fit. The measured data points correspond to the mean values within two 11×11 pixel ROI ($x_{\min}, y_{\min}, x_{\max}, y_{\max}$) and across the full sensor area. Error bars represent the 1σ dispersion across all pixels of the respective ROI.

$$\text{ENC}_{\text{eff}} (e^-) = \frac{\sigma (\text{ADU})}{g_{\text{tot}} (\text{ADU/keV})} \times \frac{1000 (\text{eV/keV})}{3.6 \text{ eV}}. \quad (5)$$

We show the average ENC_{eff} as a function of temperature for different integration times in Figure 8A and the corresponding average SNR for isolated 1 keV photons in Figure 8B. Both quantities follow an exponential trend as a function of temperature. At a $5 \mu\text{s}$ integration time and a temperature of -22°C , the JUNGFRÄU-iLGAD system achieves an ENC_{eff} of $5.5 \pm 0.7 e^-$ r.m.s., which corresponds to an SNR of 51 for single 1 keV photons when using single-pixel clusters. This becomes 25.5 for 2×2 clusters as the total noise of a cluster of N pixels scales with \sqrt{N} .

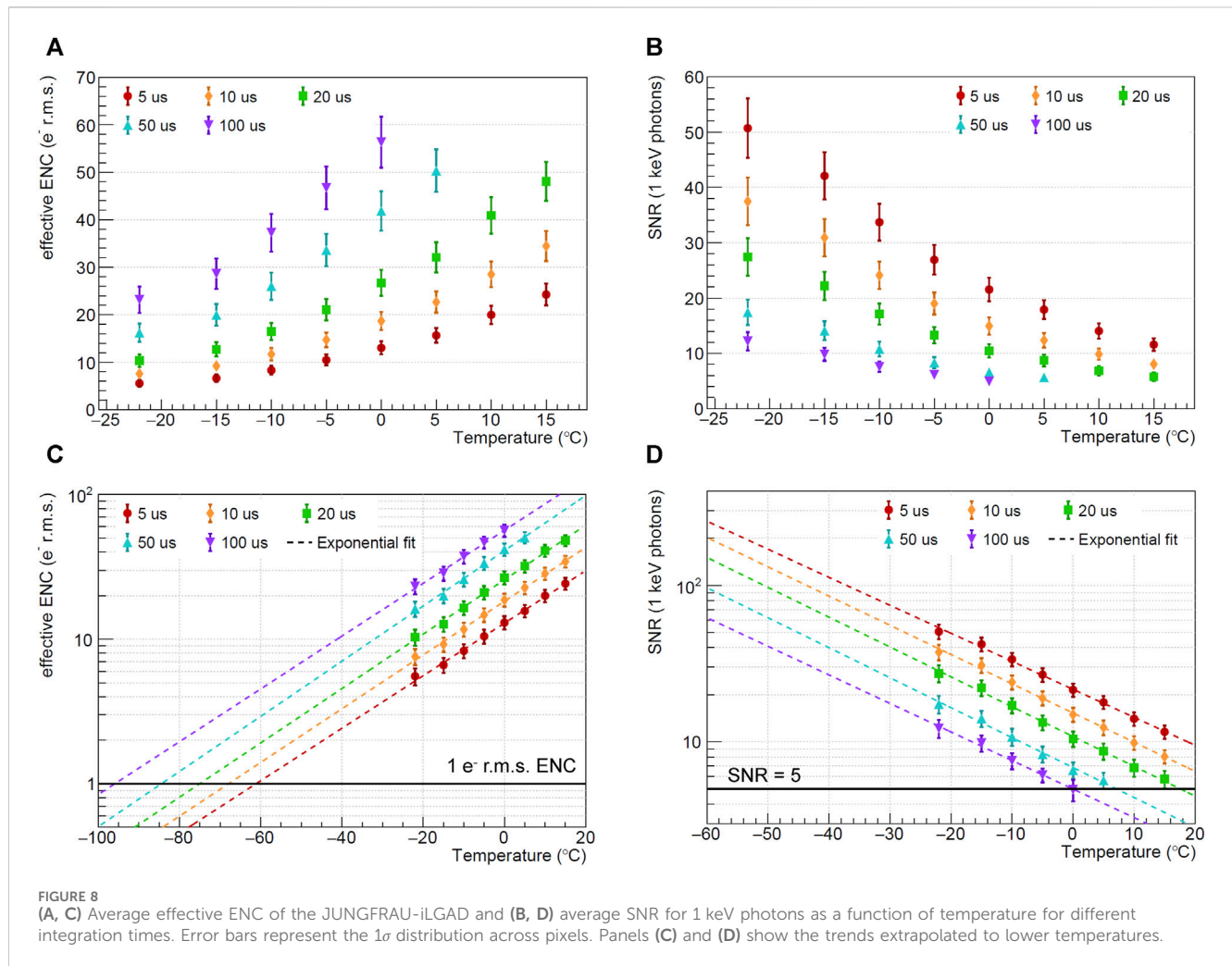
Figures 8C, D show the exponential trends extrapolated to lower temperatures. From this extrapolation, it can be estimated that in order to achieve an $\text{ENC}_{\text{eff}} \sim 1 e^-$ r.m.s. at a $5 \mu\text{s}$ integration time, the system will need to be cooled to at least -60°C . At this low temperature, the influence of the shot noise diminishes because the leakage current is reduced strongly. As a consequence, the system noise at -60°C will be dominated by the ASIC noise, and the ENC_{eff} will be determined by the magnitude of the iLGAD gain. While the noise of the JUNGFRÄU ASIC at such low temperatures needs to be verified experimentally, in a first, worst-case approximation, we can assume that it remains constant at the value of $34 e^-$ r.m.s. stated in [22]. Hence, with an iLGAD multiplication factor of $M \sim 34$, an $\text{ENC}_{\text{eff}} \sim 1 e^-$ r.m.s. would be feasible. If the trend of M shown in Figure 6 is extrapolated to lower temperatures, a value $M > 30$ is realistic below $\sim -60^\circ\text{C}$ for the high-gain ROI. To verify this estimation, direct measurements of the noise of the JUNGFRÄU-iLGAD system will need to be performed at lower temperatures.

Because of the dependency of the iLGAD multiplication factor on the photon absorption depth, the values stated above

for the ENC_{eff} are only valid for photon energies for which electron-initiated multiplication dominates. For the standard gain layer design and normal photon incidence angle, electron multiplication dominates at energies $> 500 \text{ eV}$. At lower photon energies, hole multiplication becomes important and M reduces by about a factor four [29]. For 200 eV photons at -22°C , the SNR becomes 1.25 with clustering and 2.5 without. According to the extrapolation of the SNR for 1 keV photons shown in Figure 8D, we estimate that a $\text{SNR} > 5$ for 2×2 clustered data can be obtained at 200 eV with the investigated iLGAD sensor at a $5 \mu\text{s}$ integration time by cooling below $\sim -50^\circ\text{C}$.

4 Discussion

The characterization results of the JUNGFRÄU-iLGAD prototype demonstrate that this technology will enable hybrid detectors to advance into the soft X-ray domain. The present study constitutes the first characterization of a JUNGFRÄU system employing iLGAD technology. It adds to previous investigations, which demonstrated the capabilities of the combination of an iLGAD sensor with the charge integrating, small pixel pitch MÖNCH chip [19]. With the sensor cooled to -22°C , this soft X-ray version of JUNGFRÄU achieves an effective noise of $5.5 \pm 0.7 e^-$ r.m.s. at a $5 \mu\text{s}$ integration time, which constitutes an improvement by a factor ~ 7 compared to the currently available best version of JUNGFRÄU with standard silicon sensors. Achieving an $\text{SNR} > 5$ for photon energies below 500 eV, for which hole-initiated multiplication dominates, will be possible by cooling the sensor to lower temperatures and exploiting the increase in multiplication factor with decreasing temperature.



The present study also shows that iLGAD technology for soft X-ray applications can benefit from further optimization, particularly process improvements targeting a reduction of the leakage current, an increased uniformity of the gain layer, an increase of the pixel yield, and the improvement of the SNR at low photon energies (i.e., employing a shallow gain layer design to reduce the influence of hole-initiated multiplication). These optimizations will be addressed in future iLGAD R&D batches developed by the collaboration between PSI and FBK.

A JUNGFRAU-iLGAD system for soft X-rays opens possibilities for many photon science applications at FEL and SR sources alike. The 2.2 kHz frame rate of JUNGFRAU constitutes a more than thousandfold increase in speed compared to commercial CCD cameras for soft X-rays and at least a factor of 20 improvement compared to standard CMOS image sensors. The capability of JUNGFRAU 1.0 to acquire bursts of up to 16 images with an effective frame rate of ~ 150 kHz, opens further possibilities for high-rate applications with soft X-rays at next-generation FEL facilities [33]. Furthermore, with an area of $4 \times 8 \text{ cm}^2$ for a 500 k pixel module, a JUNGFRAU system provides an active area that is eight times larger than standard CCD and CMOS devices. This can be further extended by tiling together multiple modules in

customizable configurations. Especially for photon-starved spectroscopic techniques such as RIXS, a hybrid detector with high frame rate and large area would provide a significant improvement of the ratio between measurement time and acquired statistics. Additionally, the charge integrating architecture means charge sharing between neighboring pixels can be utilized to perform position interpolation, improving the spatial resolution of the detector beyond the pixel size. With the $25 \mu\text{m}$ pitch MÖNCH chip, for instance, spatial resolution in the micron range has been demonstrated in the hard X-ray domain [27, 34]. Based on these results, we note that an iLGAD hybrid detector based on the MÖNCH chip or an adapted version of JUNGFRAU with smaller pitch in one dimension would open further possibilities for spectroscopic techniques such as RIXS, soft X-ray spectro-ptychography, and soft or tender X-ray full field imaging techniques that rely on high spatial resolution.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

VH: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review and editing. RB: Resources, Writing—review and editing. FB: Formal Analysis, Validation, Visualization, Writing—review and editing. AB: Conceptualization, Data curation, Investigation, Project administration, Supervision, Validation, Writing—review and editing. GB: Resources, Writing—review and editing. MBo: Resources, Writing—review and editing. MBr: Resources, Writing—review and editing. MCa: Resources, Validation, Writing—review and editing. MCe: Resources, Writing—review and editing. RD: Resources, Validation, Writing—review and editing. SE: Resources, Writing—review and editing. FF: Resources, Writing—review and editing. EF: Resources, Software, Validation, Writing—review and editing. DG: Validation, Writing—review and editing. OH: Resources, Writing—review and editing. SH: Resources, Writing—review and editing. JH: Validation, Writing—review and editing. TK: Resources, Writing—review and editing. PK: Writing—Review and Editing, Resources. CL-C: Resources, Writing—review and editing. DM: Resources, Writing—review and editing. AM: Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Supervision, Validation, Writing—review and editing. KM: Resources, Writing—review and editing. GP: Resources, Writing—review and editing. KP: Funding acquisition, Validation, Writing—review and editing. SR: Resources, Writing—review and editing. CR: Resources, Writing—review and editing. BS: Conceptualization, Funding acquisition, Project administration, Supervision, Validation, Writing—review and editing. PS: Resources, Writing—review and editing. DT: Resources, Writing—review and editing. XX:

Validation, Writing—review and editing. JZ: Conceptualization, Formal Analysis, Methodology, Validation, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. VH and KP received funding from MSCA PSI-FELLOW-III-3i (EU Grant Agreement No. 884104). Open access funding by PSI—Paul Scherrer Institute

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ismail ASM, Uemura Y, Park SH, Kwon S, Kim M, Elnaggar H, et al. Direct observation of the electronic states of photoexcited hematite with ultrafast 2p3d X-ray absorption spectroscopy and resonant inelastic X-ray scattering. *Phys Chem Chem Phys* (2020) 22:2685–92. doi:10.1039/C9CP03374B
- Elnaggar H, Wang RP, Lafuerza S, Paris E, Tseng Y, McNally D, et al. Magnetic contrast at spin-flip excitations: an advanced X-ray spectroscopy tool to study magnetic-ordering. *ACS Appl Mater Inter* (2019) 11:36213–20. doi:10.1021/acsami.9b10196
- Barantani F, Tran MK, Madan I, Kapon I, Bachar N, Asmara TC, et al. Resonant inelastic X-ray scattering study of electron-exciton coupling in high- T_c cuprates. *Phys Rev X* (2022) 12:021068. doi:10.1103/PhysRevX.12.021068
- Henrich B, Bergamaschi A, Broennimann C, Dinapoli R, Eikenberry E, Johnson I, et al. Pilatus: a single photon counting pixel detector for x-ray applications. *Nucl Instrum Meth A* (2009) 607:247–9. doi:10.1016/j.nima.2009.03.200
- Tinti G, Bergamaschi A, Cartier S, Dinapoli R, Greiffenberg D, Johnson I, et al. Performance of the EIGER single photon counting detector. *J Instrumentation* (2015) 10:C03011. doi:10.1088/1748-0221/10/03/C03011
- Kuster M, Boukhelef D, Donato M, Dambietz JS, Hauf S, Maia L, et al. Detectors and calibration concept for the European XFEL. *Synchrotron Radiat News* (2014) 27: 35–8. doi:10.1080/08940886.2014.930809
- Blaj G, Caragiulo P, Carini G, Carron S, Dragone A, Freytag D, et al. X-Ray detectors at the linac coherent light source. *J Synchrotron Radiat* (2015) 22:577–83. doi:10.1107/S1600577515005317
- Mozzanica A, Bergamaschi A, Brueckner M, Cartier S, Dinapoli R, Greiffenberg D, et al. Characterization results of the JUNGFRUAU full scale readout ASIC. *J Instrumentation* (2016) 11:C02047. doi:10.1088/1748-0221/11/02/C02047
- Hall DJ, Soman M, Tutt J, Murray N, Holland A, Schmitt T, et al. Improving the resolution in soft X-ray emission spectrometers through photon-counting using an Electron Multiplying CCD. *J Instrumentation* (2012) 7:C01063. doi:10.1088/1748-0221/7/01/C01063
- Kuster M, Ahmed K, Ballak KE, Danilevski C, Ekmedzić M, Fernandes B, et al. The 1-Megapixel pnCCD detector for the Small Quantum Systems Instrument at the European XFEL: system and operation aspects. *J Synchrotron Radiat* (2021) 28: 576–87. doi:10.1107/S1600577520015659
- Goldschmidt A, Grace C, Joseph J, Krieger A, Tindall C, Denes P. VeryFastCCD: a high frame rate soft X-ray detector. *Front Phys* (2023) 11. doi:10.3389/fphy.2023.1285350
- Tamma C, Rota L, Caragiulo P, Markovic B, Segal J, Kwiatkowski M, et al. ePixM: a CMOS monolithic sensor for LCLS II. In: 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); 26 October 2019 - 02; Manchester, UK (2019). p. 1–4. doi:10.1109/NSS/MIC42101.2019.9059784
- Marras A, Correa J, Lange S, Vardanyan V, Gerhardt T, Kuhn M, et al. Characterization of the Percival detector with soft X-rays. *J Synchrotron Radiat* (2021) 28:131–45. doi:10.1107/S1600577520013958
- Léveillé C, Desjardins K, Popescu H, Vondungbo B, Hennes M, Delaunay R, et al. Single-shot experiments at the soft X-FEL FERMI using a back-side-illuminated scientific CMOS detector. *J Synchrotron Radiat* (2022) 29:103–10. doi:10.1107/S1600577521012303
- Desjardins K, Medjoubi K, Sacchi M, Popescu H, Gaudemer R, Belkhou R, et al. Backside-illuminated scientific CMOS detector for soft X-ray resonant scattering and ptychography. *J Synchrotron Radiat* (2020) 27:1577–89. doi:10.1107/S160057752001262X
- Andresen N, Bakalis C, Denes P, Goldschmidt A, Johnson I, Joseph JM, et al. A low noise CMOS camera system for 2D resonant inelastic soft X-ray scattering. *Front Phys* (2023) 11. doi:10.3389/fphy.2023.1285379
- Porro M, Andricek L, Aschauer S, Castoldi A, Donato M, Engelke J, et al. The MiniSDD-based 1-mpixel camera of the DSSC Project for the European XFEL. *IEEE Trans Nucl Sci* (2021) 68:1334–50. doi:10.1109/TNS.2021.3076602
- Castoldi A, Ghisetti M, Guazzoni C, Aschauer S, Strüder L, Hansen K, et al. Qualification of the X-ray spectral performance of the DEPFET pixels of the DSSC imager. *Nucl Instrum Meth A* (2023) 1057:168686. doi:10.1016/j.nima.2023.168686

19. Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of LGAD sensors with a thin entrance window for soft X-ray detection. *J Instrumentation* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/C11011
20. Leonarski F, Redford S, Mozzanica A, Lopez-Cuenca C, Panepucci E, Nass K, et al. Fast and accurate data collection for macromolecular crystallography using the JUNGFR AU detector. *Nat Methods* (2018) 15:799–804. doi:10.1038/s41592-018-0143-7
21. Leonarski F, Nan J, Matej Z, Bertrand Q, Furrer A, Gorgisyan I, et al. Kilohertz serial crystallography with the JUNGFR AU detector at a fourth-generation synchrotron source. *IUCr* (2023) 10:729–37. doi:10.1107/S2052252523008618
22. Hinger V, al Haddad A, Barten R, Bergamaschi A, Brückner M, Carulla M, et al. Advancing the JUNGFR AU detector toward low-energy X-ray applications. *J Instrumentation* (2022) 17:C09027. doi:10.1088/1748-0221/17/09/C09027
23. Pellegrini G, Fernández-Martínez P, Baselga M, Fleta C, Flores D, Greco V, et al. Technology developments and first measurements of Low Gain Avalanche Detectors (LGAD) for high energy physics applications. *Nucl Instrum Meth A* (2014) 765:12–6. doi:10.1016/j.nima.2014.06.008
24. Andrä M, Zhang J, Bergamaschi A, Barten R, Borca C, Borghi G, et al. Development of low-energy X-ray detectors using LGAD sensors. *J Synchrotron Radiat* (2019) 26:1226–37. doi:10.1107/S1600577519005393
25. Fernández-Martínez P, Flores D, Hidalgo S, Greco V, Merlos A, Pellegrini G, et al. Design and fabrication of an optimum peripheral region for low gain avalanche detectors. *Nucl Instrum Meth A* (2016) 821:93–100. doi:10.1016/j.nima.2016.03.049
26. Carulla M, Centis Vignali M, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, et al. Study of the internal quantum efficiency of FBK sensors with optimized entrance windows. *J Instrumentation* (2023) 18:C01073. doi:10.1088/1748-0221/18/01/C01073
27. Cartier S, Kagias M, Bergamaschi A, Wang Z, Dinapoli R, Mozzanica A, et al. Micrometer-resolution imaging using MÖNCH: towards G₂-less grating interferometry. *J Synchrotron Radiat* (2016) 23:1462–73. doi:10.1107/S1600577516014788
28. Ament LJP, van Veenendaal M, Devereaux TP, Hill JP, van den Brink J. Resonant inelastic x-ray scattering studies of elementary excitations. *Rev Mod Phys* (2011) 83:705–67. doi:10.1103/RevModPhys.83.705
29. Liguori A, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Characterization of iLGADs using soft X-rays. *J Instrumentation* (2023) 18:P12006. doi:10.1088/1748-0221/18/12/P12006
30. Cartier S, Bergamaschi A, Dinapoli R, Greiffenberg D, Johnson I, Jungmann JH, et al. Micron resolution of MÖNCH and GOTTHARD, small pitch charge integrating detectors with single photon sensitivity. *J Instrumentation* (2014) 9:C05027. doi:10.1088/1748-0221/9/05/C05027
31. Redford S, Andrä M, Barten R, Bergamaschi A, Brückner M, Dinapoli R, et al. First full dynamic range calibration of the JUNGFR AU photon detector. *J Instrumentation* (2018) 13:C01027. doi:10.1088/1748-0221/13/01/C01027
32. Currás Rivera E, Moll M. Study of impact ionization coefficients in silicon with low gain avalanche diodes. *IEEE Trans Electron Devices* (2023) 70:2919–26. doi:10.1109/TED.2023.3267058
33. Sikorski M, Ramilli M, de Wijn R, Hinger V, Mozzanica A, Schmitt B, et al. First operation of the JUNGFR AU detector in 16-memory cell mode at European XFEL. *Front Phys* (2023) 11. doi:10.3389/fphy.2023.1303247
34. Chiriotti S, Barten R, Bergamaschi A, Brückner M, Carulla M, Chsherbakov I, et al. High-spatial resolution measurements with a GaAs:Cr sensor using the charge integrating MÖNCH detector with a pixel pitch of 25 μm . *J Instrumentation* (2022) 17:P04007. doi:10.1088/1748-0221/17/04/P04007



OPEN ACCESS

EDITED BY

Cinzia Da Via,
The University of Manchester, United Kingdom

REVIEWED BY

Filippo Mele,
Politecnico di Milano, Italy
Marco Povoli,
SINTEF Digital, Norway
Florian Schopper,
Halbleiterlabor of the Max-Planck-Society,
Germany

*CORRESPONDENCE

Matteo Centis Vignali,
✉ mcentisvignali@fbk.eu
Giovanni Paternoster,
✉ paternoster@fbk.eu

RECEIVED 20 December 2023

ACCEPTED 06 June 2024

PUBLISHED 24 July 2024

CITATION

Centis Vignali M and Paternoster G (2024), Low gain avalanche diodes for photon science applications.
Front. Phys. 12:1359179.
doi: 10.3389/fphy.2024.1359179

COPYRIGHT

© 2024 Centis Vignali and Paternoster. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Low gain avalanche diodes for photon science applications

Matteo Centis Vignali* and Giovanni Paternoster*

Fondazione Bruno Kessler, Center for Sensors and Devices, Trento, Italy

Low Gain Avalanche Diodes (LGADs) are silicon sensors designed to achieve an internal gain in the order of 10 through the impact ionization process. The development of LGADs was pushed forward by their application in High Energy Physics (HEP) experiments, where they will be employed to provide measurements of the time of arrival of minimum ionizing particles with a resolution of around 30 ps. The initial technological implementation of the sensors constrains their minimum channel size to be larger than 1 mm^2 , in order to reduce inefficiencies due to the segmentation of the gain structure. The gain of the sensors is kept in the order of 10 to limit the sensor shot noise and their power consumption. In photon science, the gain provided by the sensor can boost the signal-to-noise ratio of the detector system, effectively reducing the x-ray energy threshold of photon counting detectors and the minimum x-ray energy where single photon resolution is achieved in charge integrating detectors. This can improve the hybrid pixel and strip detectors for soft and tender x-rays by simply changing the sensor element of the detector system. Photon science applications in the soft and tender energy range require improvements over the LGADs developed for HEP, in particular the presence of a thin entrance window to provide a satisfactory quantum efficiency and channel size with a pitch of less than $100\text{ }\mu\text{m}$. In this review, the fundamental aspects of the LGAD technology are presented, discussing also the ongoing and future developments that are of interest for photon science applications.

KEYWORDS

LGAD, photon science, soft and tender x-rays, hybrid silicon detectors, synchrotron radiation

1 Introduction

Photon science experiments at the soft and tender x-rays energies in the region of 200–2000 eV allow for a wide variety of research topics and applications. A few examples are.

- Pharmaceutical industry: crystallography of proteins not containing heavy elements [1].
- Life science: high resolution microscopy and ptychography for cells and tissues with high contrast in the “water window” [2].
- New magnetic materials for data storage, energy production, and automotive: ptychography and diffraction for nm-scale spatial resolution [3, 4].
- High-Tc superconductors for energy transport: resonant x-ray diffraction [5].
- Microelectronics manufacturing: mask inspection using extreme ultraviolet light [6].

To enable these lines of research, x-ray detectors must provide: high spatial resolution, high frame rate, good quantum efficiency (QE), large area, and dynamic range. Different

detector technologies are employed for soft and tender x-ray measurements, like avalanche photodiodes (APDs), silicon drift detectors (SDDs), charge-coupled devices (CCDs), and CMOS sensors. These detectors have drawbacks that limit their applications. APDs offer very coarse segmentation, resulting in poor spatial resolution. SDDs showed excellent QE and rate capability, but their spatial resolution is limited. CCDs offer a better spatial resolution, however, they are limited in dynamic range and rate capability, with long readout times preventing their application in scanning measurements like ptychography. The available systems based on CMOS sensors are limited in size by the yield of the thinning and doping steps necessary to create an entrance window for the x-rays.

Currently available hybrid pixel and strip detectors excel in spatial resolution, frame rate, sensitive area, and dynamic range, making them an ideal starting point for the development of x-ray detectors. Hybrid detectors consist of a semiconductor sensor connected to readout Application Specific Integrated Circuits (ASICs), allowing for the separate optimization of the two components, and redtherefore the enhancement of the detector system for a certain application's energy range, by changing the sensor being readout by the ASIC (e.g., using CdTe sensors for hard x-rays). Examples representing the state-of-the-art in hybrid pixel and strip detectors for photon science at synchrotrons and free electron lasers can be found in [7–12]. The performance of these detectors is showcased by their spatial resolution of 1–2 μm reached by interpolation with a $25 \times 25 \mu\text{m}^2$ pixel sensor [13], their module area of up to 32 cm^2 where several ASICs are bonded to a single sensor [8, 9], their dynamic range of 10^4 12 keV photons [9], and their frame rate ranging between 2.4 kHz and 4.5 MHz [9, 11, 14].

Currently, hybrid detectors are limited to operate at energies above 2000 eV, due to their QE and signal-to-noise ratio (SNR). The SNR affects the detectors in different ways depending on their readout electronics. Photon counting detectors use a discriminator circuit with a set threshold in each pixel or strip, each time the sensor shows a signal above threshold one x-ray is counted. The threshold value of these detectors is determined by the electronics noise, whose typical value is 80–200 electrons [7, 8, 12]. The threshold is typically set to 5 times the noise level, corresponding to photon energies of 1440–3600 eV when silicon is used as sensor medium. X-ray energies below this threshold are not detected. Charge integrating detectors integrate the current measured by each pixel or strip during a frame. For these detectors the SNR determines how many photons or which photon energy is necessary to produce a measurable signal. A SNR of about 5 is necessary to achieve a sensitivity to single photons. With a noise of about 35–50 electrons [9, 10], single-photon sensitivity can be achieved for 630–900 eV x-rays using standard silicon sensors. The application of both photon counting and integrating detectors could be limited by their QE for soft and tender x-rays. A low QE limits the sensitivity of the detectors, requiring longer measurement times or higher intensity beams. To extend the operating range of hybrid detectors to lower energies, an increase in QE and system SNR is necessary.

The QE of the sensors can be improved by creating a thin entrance window (TEW) structure that minimizes the inactive layers that passivate the sensor surface, allowing the x-rays to reach the sensitive volume of the sensor and at the same time preserving the ionization generated in the sensor by reducing as much as possible all recombination mechanisms.

The SNR of the detector system can be improved by using sensors with charge multiplication like Low Gain Avalanche Diodes (LGADs). LGADs are silicon sensors with an internal gain in the order of 10 obtained via the impact ionization process. The development of LGADs was pushed forward by their application in High Energy Physics (HEP) experiments, where they will be employed to provide measurements of the time of arrival of minimum ionizing particles with a resolution of around 30 ps [15, 16]. This performance is achieved by the combination of the internal gain with silicon substrates with an active thickness of about 50 μm . The initial technological implementation of the sensors constrains their minimum channel size to be larger than 1 mm^2 , in order to reduce inefficiencies due to the segmentation of the gain structure. Subsequent improvements of the sensors addressed this problem, relying on different strategies. The gain of the sensors is kept in the order of 10 to limit the sensor shot noise and their power consumption. Recently, the use of LGADs was investigated in other fields than HEP such as beam monitoring for hadron therapy, astroparticle physics experiments, and photon science [17–20].

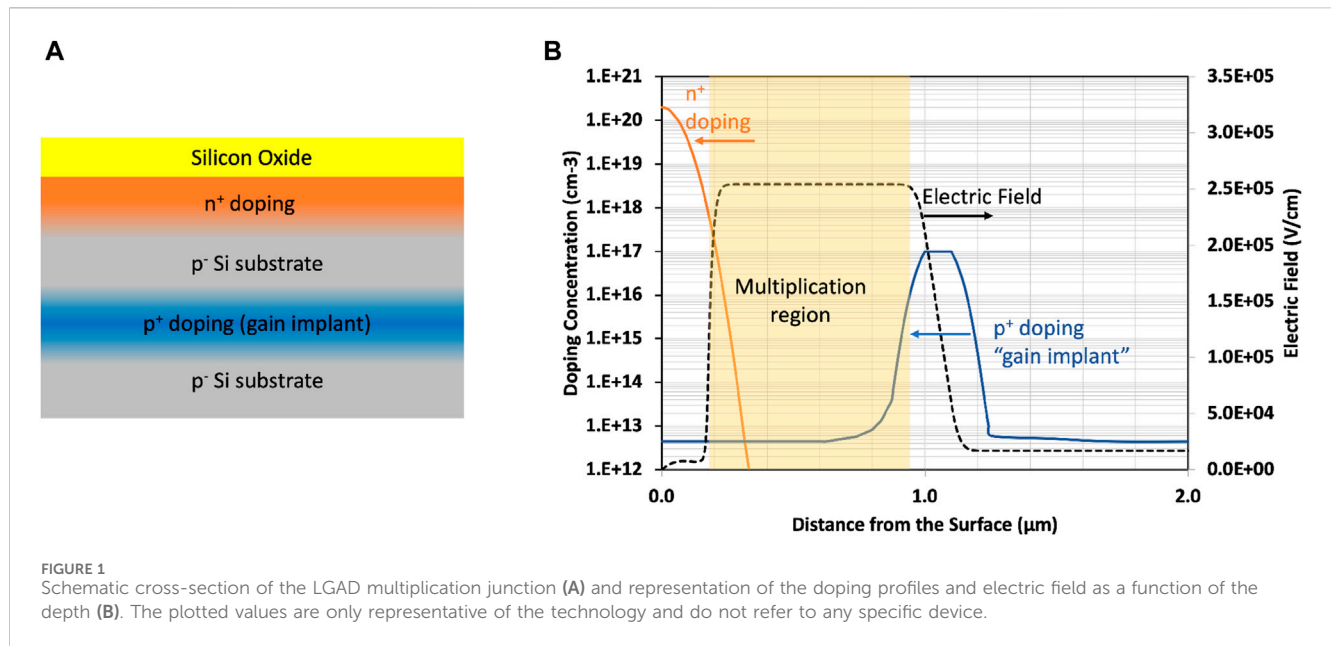
This paper provides a review of the current development of LGAD sensors for photon science applications and offers some considerations for the possible developments of the sensors to better suit the needs of this application field. Current and proposed LGAD technologies are presented in Section 2, being followed by a review of the results obtained in the characterization of LGADs using x-rays in Section 3. As the improvement in the QE of these sensors relies on the development of a TEW, the main considerations for the realization of this component are summarized in Section 4. The theory describing the gain and noise properties of the sensors due to charge multiplication is outlined in Section 5. This theoretical frame is used in Section 6 to motivate general considerations on possible approaches to realize LGAD sensors for photon science. A comparison of sensor structures, both existing and proposed, for their fitness in photon science applications is given in Section 7. Finally a summary is provided in Section 8.

2 LGAD technologies

Over the past decade, LGADs have witnessed significant technological advancements driven by three pivotal objectives identified as crucial by the HEP detector community. These objectives encompass the improvement of the sensor's spatial resolution and fill factor (FF), the enhancement of time resolution in minimum ionizing particle (MIP) detection, and the elevation of radiation hardness beyond 10^{15} neq/ cm^2 (where neq denotes 1-MeV-neutron-equivalent). To achieve these objectives, various LGAD technologies and design variations have been proposed. This section provides a brief overview and discussion of the current state-of-the-art technology for LGADs and their principal variations, emphasizing their envisioned application in photon science.

2.1 Standard LGAD

A schematic representation of the fundamental structure of an LGAD is depicted in Figure 1. The multiplication junction follows a



reach-through scheme, typical of APDs and SPADs, denoted as n⁺-p⁻-p⁺-p⁻. This scheme is obtained by incorporating two doping profiles on a p⁻ substrate: firstly, a shallow n⁺ doping using Arsenic or Phosphorous, and secondly, a deeper p-type doping (also named “gain implant”), typically achieved with Boron. Upon depletion, the electric field in the region between these two doped areas locally increases to values exceeding the impact-ionization threshold (approximately $2 \cdot 10^5 \text{ Vcm}^{-1}$), enabling the mechanism responsible for charge multiplication.

The region with the high electric field, where the impact ionization happens, is indicated as “multiplication region” or “gain layer”. The doping profiles of the multiplying junction are typically fabricated using ion implantation, even if other approaches, such as dopant diffusion by solid or gaseous sources and doping during epitaxial growth, are possible. The Boron doping profile typically peaks at 1–1.5 μm for the front surface, resulting in a multiplication region extending from a few hundred nanometers from the surface up to the falling edge of the Boron implant at about 1.5–2 μm. A review of LGADs fabrication technologies is available in [21].

The first prototypes of LGADs were manufactured by the Centro Nacional de Microelectronica (IMB-CNM, Barcelona, Spain) in 2013 [16]. Since then, the technology has undergone further developments and has reached a good technology readiness level within the CERN RD50 collaboration¹. Currently, many research facilities around the world are involved in LGAD manufacturing and R&D, such as: Fondazione Bruno Kessler (FBK, Trento, Italy), Brookhaven National Laboratory (BNL, Upton, United States) IMB-CNM, Hamamatsu (Japan), IHEP-NDL (Beijing, China), USTC-IME (China), Micron (Lancing, United Kingdom),

Teledyne e2v (Chelmsford, United Kingdom), SINTEF MiNaLab (Oslo, Norway).

The first LGADs were manufactured on approximately 300 μm thick p-type Float-zone (FZ) wafers by IMB-CNM and FBK [16, 22]. The early devices showcased the potential for signal multiplication of around 10 but the temporal resolution remained constrained, typically in the range of 100 ps. To further improve the time resolution in MIP detection, a new generation of LGADs was then produced on thinner substrates of 45–60 μm [23–25].

The advantage of using thin substrates originates from the peculiar shape of the signal generated by a MIP in LGADs: it has a rise time that is as long as the drift time of an electron traversing the entire sensor thickness. Therefore thinner substrates originate signal waveforms with a higher steepness. A detailed discussion of the signal shape for an LGAD is reported in [15]. Considering the voltage waveform after the amplifier stage, the signal steepness or “slew-rate” is indicated as $\frac{dV}{dt}$, and can be expressed as:

$$\frac{dV}{dt} \propto \frac{G}{d} \quad (1)$$

where d is the depleted (or “active”) sensor thickness and G represents the gain. This equation is only valid under the assumption of a uniform charge generation throughout the sensor thickness, and considering an ideal pad geometry (neglecting the border effects). Reducing the sensor’s active thickness improves the slew-rate but increases the junction capacitance, negatively affecting the time resolution. The optimal balance between these factors depends on the specific application and the readout electronics being used.

A similar trade-off can be found for the gain: a high gain value increases the signal slew-rate but a too high gain is detrimental to the SNR, due to higher shot noise. In addition, increasing the sensor gain by adjusting the operating voltage also results in increased current and bias voltage, leading to higher power consumption. An alternative approach is to increase the doping level of the gain

¹ RD50 - Radiation hard semiconductor devices for very high luminosity colliders. cern.ch/rd50.

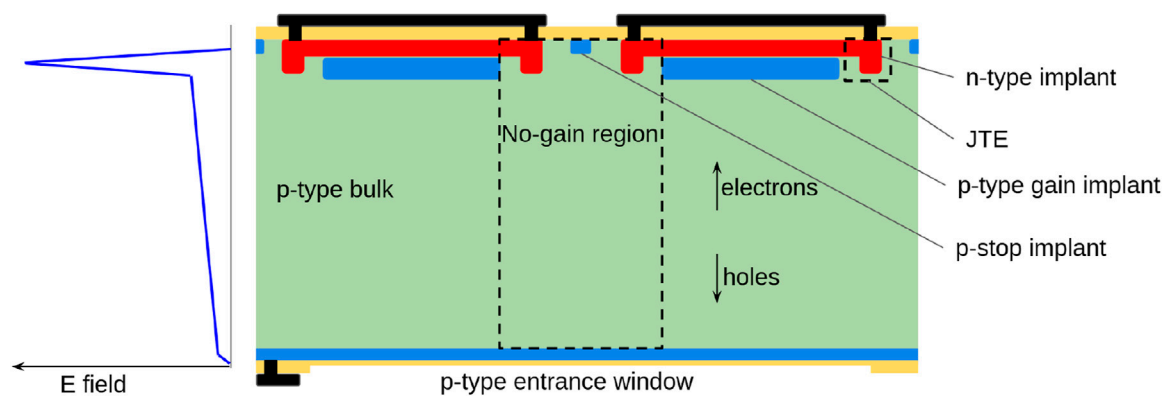


FIGURE 2

Schematic cross-section of a standard LGAD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom. The entrance window was added for the application of these sensors to x-ray detection.

implant, which allows for high gain at lower bias voltage, thereby limiting power consumption. However, this solution may not always be practical as it could prevent the electron velocity from saturating throughout the entire depleted volume, which can be detrimental to time resolution [15]. In HEP timing applications, experimental results and simulations indicate that a thickness of ~ 50 micron combined with a gain of ~ 10 provides optimum performance for channel dimensions of the order of $\sim 1 \text{ mm}^2$ [15]. Such thin LGADs, optimized for timing applications are also known as “Ultra Fast Silicon Detectors” (UFSD) and have been demonstrated to reach time resolution as low as 30 ps and 40 ps, before and after irradiation up to a fluence of 10^{15} neq/cm^2 , respectively [23, 24].

In a silicon processing clean room, $\sim 50 \text{ }\mu\text{m}$ thick substrates cannot be directly handled, and they need to be attached to a thick supporting layer. Typically the thin active silicon can be either an epitaxial layer grown over a thick low-resistivity wafer, or a thin FZ-wafer, wafer-bonded to a low-resistivity handle wafer. In both cases, the handle wafer acts as the back anode contact and the processing takes place on the front side only. Due to the low resistivity substrate on the backside of the wafer and the presence of read-out pads on the front side, thin LGADs designed for HEP applications, are not suitable for detecting photons or low-penetrating particles and x-rays. For these applications, a full-depleted bulk silicon substrate is required. In this approach, a p^+ implant and a thin passivation layer can be fabricated on the backside, serving as an entrance window for low-penetrating particles and soft x-rays.

2.1.1 LGAD segmentation

The initial strategy for creating segmented LGADs adopted the same segmentation scheme employed in standard n-i-p diode arrays. This approach entails the segmentation of the entire $n^+-p^--p^+-p^--$ multiplying junction, with the addition of an extra p-type region referred to as a “p-stop” between the pixels to ensure full electrical isolation, as illustrated in Figure 2. However, the segmentation of a sensor with internal multiplication presents additional challenges compared to conventional p-i-n diodes. Indeed, to avoid undesired high-electric field regions at the termination edge of the gain layer, an additional n-type region known as Junction Termination Extension (JTE) is typically

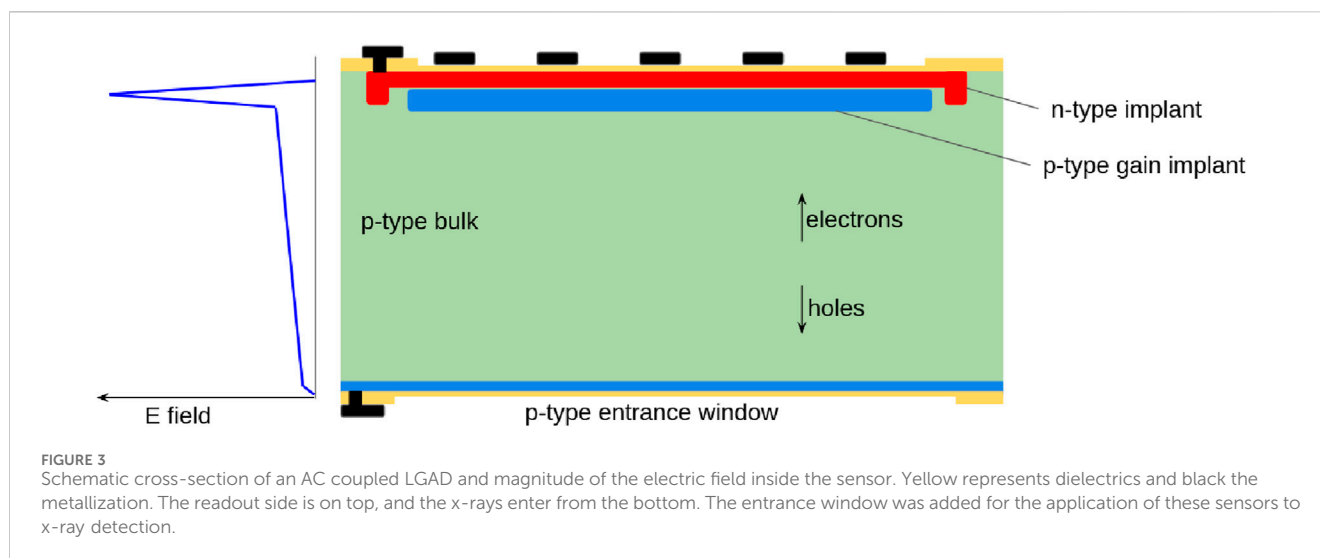
included. The JTE controls the n^+ curvature and reduces the electric field at the pixel border. The gain layer is removed in the JTE and p-stop regions and also indented from the JTE edge to prevent premature breakdown at the pixel edge (edge-breakdown). These termination structures use some of the sensor’s area and, as a consequence, introduce an inter-pixel region where the gain is suppressed [26]. This is the so-called no-gain region, defined as the distance between two adjacent gain implant regions. In standard LGADs, the no-gain region is typically in the range $50\text{--}100 \text{ }\mu\text{m}$, depending on the fabrication technology and device design. The presence of such a no-gain region reduces the fill factor (FF) of the pixel ($\text{FF} = \text{pixel area with signal multiplication} / \text{total pixel area}$).

In current HEP experiments, where the required pixel size is typically around 1 mm^2 , the FF loss is almost negligible. However, in photon science applications, where a pixel size in the range of $25\text{--}100 \text{ }\mu\text{m}$ is usually required, the wide inter-pixel region makes the standard LGAD technology inadequate.

2.2 Capacitively coupled LGAD (AC-LGAD)

The first technological solution proposed to overcome the segmentation issues in LGADs was the capacitively coupled LGAD (AC-LGAD) [27, 28]. In this architecture, the multiplying junction (both n^+ and p^+ layers) is not patterned and covers the full sensor area, without any interruption. The read-out is kept on the same side of the multiplication region and the sensor segmentation is achieved via metal pads AC-coupled to the n^+ layer via a thin dielectric film (about 100 nm thick). The n^+ layer is then connected to the ground at the sensor periphery, where the JTE is present as in the case of standard LGAD. A sketch is shown in Figure 3. In this scheme, thanks to the uniform multiplication region across the sensor, a 100% fill factor is achieved with uniform multiplication and no dead areas, independently of the channel segmentation.

Signal formation in an AC-LGAD differs significantly from the one of a standard LGAD [29]. In AC-LGADs, the signal at the readout pads is primarily induced during the charge propagation along the n^+ layer, before discharging to the ground. The resulting signal is bipolar, with a zero net integral. The positive lobe is



generated during the current lateral spread along the lossy transmission line composed of the n^+ layer, the bulk, and the AC-pads. The negative lobe accounts for the AC-pad discharge, with an RC constant that depends on the read-out input resistance, the n^+ sheet resistance and its capacitances towards the sensor backside and the metal pads.

In AC-LGADs, the signal could propagate in the n^+ layer for hundreds of microns before being completely collected by the pick-up electrodes, as a consequence, a certain amount of cross-talk exists among adjacent pixels. Signal-sharing among pixels can be used to enhance spatial resolution by interpolating signals from neighboring pixels. This resistive read-out method was first investigated by INFN and the University of Turin for AC-LGADs, which they called Resistive Silicon Detectors (RSD) [28]. They demonstrated that RSDs can detect MIPs with a similar temporal resolution of standard LGADs (approximately ~ 30 ps) and with a spatial resolution as low as 5% of the metal pads pitch, which usually ranges from $100\ \mu\text{m}$ to $500\ \mu\text{m}$ [30]. Notably, AC-LGADs (RSD) exhibit a spatial resolution surpassing that obtained in the absence of charge sharing between pixels (quoted as $\text{pitch}/\sqrt{12}$). These results pave the way for the realization of a detector that offers high temporal and spatial resolution, while also requiring fewer channels in the read-out electronics.

AC-LGAD detectors have been manufactured by various research facilities, primarily on thin epitaxial silicon substrates (approximately $50\ \mu\text{m}$) for HEP tracking. However, a prototype on a thick FZ silicon substrate was presented in [22]. If produced on thick FZ silicon wafers using a double-sided manufacturing process (see Figure 3) this device becomes capable of detecting photons and particles entering from the sensor backside. Here, a p^+ implant and a thin passivation layer can be implemented, forming an entrance window.

Despite the promising attributes, such as fine segmentation and a 100% fill factor, which make AC-LGADs appealing for photon science applications, their use in this field is now hampered by the peculiar shape of the produced signals: the fast and bipolar signals are not compatible with charge-integrating electronics or with the current single-photon counting front-end used in photon science, as the long integration time would produce zero-net-charge signals.

Some authors suggested that this effect can be attenuated in single-photon counting detectors by making the n^+ layer float through a large value resistor connected between ground and the n^+ contact [27]. Electrons collected at the n^+ layer will then discharge with a long time constant, producing a signal with negligible opposite polarity pulse amplitude. In conclusion, to fully exploit this technology in photon science applications, significant modifications of the read-out electronics or optimization of the sensor output are required.

2.3 DC-coupled resistive silicon detectors (DC-RSD)

A further development, exploiting the resistive readout employed in AC-LGADs led to the proposal of DC-coupled resistive silicon detectors (DC-RSD) by INFN and the University of Turin [31]. Similar to AC-LGADs, the multiplying junction in DC-RSD is not patterned and extends to the entire sensor area. However, in DC-RSD, the metal pads are directly connected to the n^+ layer via ohmic contacts (refer to Figure 4). Electrical isolation among different channels is achieved by designing an n^+ layer with sufficiently high sheet resistance and optimizing the inter-pads distance accordingly. The motivation behind the development of DC-RSD lies in two main objectives: 1) restricting the spread of signals beyond the nearest pads to enhance spatial resolution, and 2) generating unipolar signals instead of the bipolar signals produced by AC-LGADs. DC-RSD can be also produced using a double-sided process, allowing for the incorporation of an entrance window on the backside for applications involving photon and x-ray imaging. While this technology promises to address some of the limitations associated with AC-LGADs, its development is still ongoing, and experimental validation is still needed.

2.4 Inverted LGAD (iLGAD)

An alternative approach aimed at achieving high spatial resolution using a structure based on LGADs is known as the

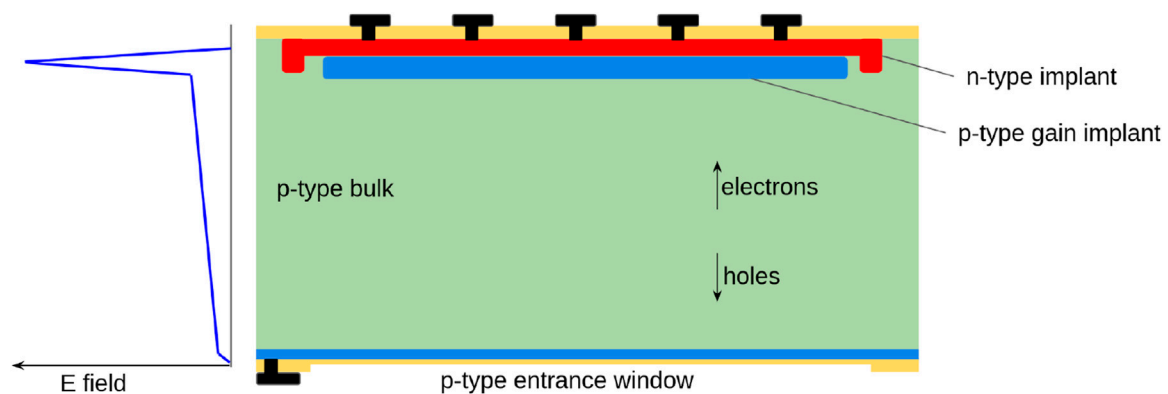


FIGURE 4
Schematic cross-section of DC-RSD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom. The entrance window was added for the application of these sensors to x-ray detection.

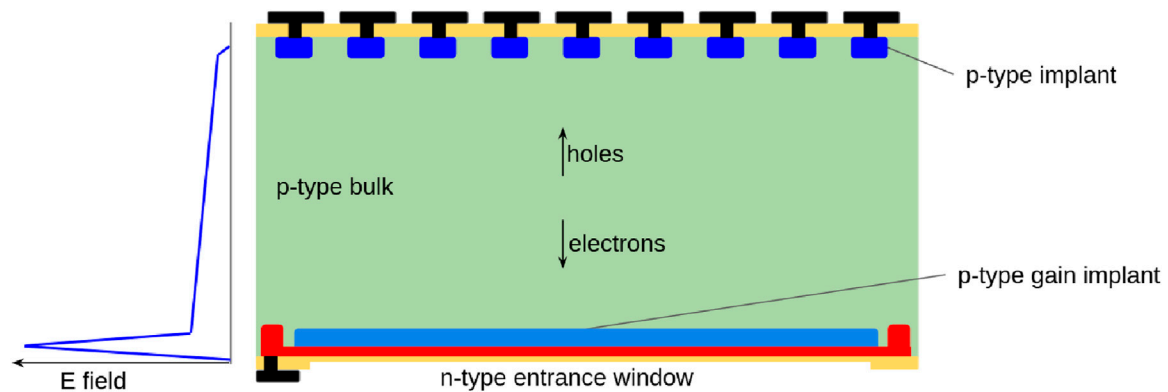


FIGURE 5
Schematic cross-section of a double-sided (inverted) LGAD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom. Figure originally published in [34].

inverted LGAD (iLGAD) [32]. Originally proposed in [33] and previously referred to as double-sided LGADs, this design features a multiplication junction formed by a uniquely unpatterned region covering the entire sensor, ensuring a 100% fill factor, while the read-out segmentation is provided by the collecting electrodes on the opposite side (ohmic or read-out side), without compromising the gain uniformity along the sensor (refer to Figure 5).

To ensure full electrical isolation among the read-out channels, the device must operate in full depletion. Additionally, as a full double-sided process is required to define both surfaces of the wafer, the latter must be approximately 200–300 μm thick. Unfortunately, such a high active thickness compromises the timing capability of iLGAD (see Eq. 1), making this technology less appealing in HEP applications where fast timing is crucial. However, due to its exceptional segmentation capabilities, iLGAD is valuable for applications where the timing information is not critical. It stands out as the predominant LGAD technology for these applications, as discussed in Section 3.

2.4.1 n-type iLGAD

In iLGADs, the integration of the read-out on the opposite side of the sensor necessitates placing the entrance window for photons and x-rays on the gain side. This feature introduces a dependence of the multiplication on the particle interaction point. In p-type LGADs, the multiplication process for low-energy particles, interacting close to the surface (in the superficial neutral region) is initiated by holes traveling through the gain layer. For particles interacting deeper in the substrate, the multiplication is electron-initiated. The latter process produces signals with both higher gain and SNR, as discussed in Section 5.

To address this limitation and enhance the efficiency of detecting low-energy photons and particles, iLGADs on n-type substrates have also been proposed in [35], and experimentally fabricated by IMB-CNM [36]. This sensor was named proton Low Gain Avalanche Detectors (pLGAD) and features inverted doping compared to traditional iLGAD (as represented in Figure 6). Such a sensor was also provided with a thin surface passivation layer and an entrance window optimized for the detection of particles with a low penetrating depth (15 keV protons in this specific case, which have a maximum range in silicon of 300 nm). In such a detector, only signals created

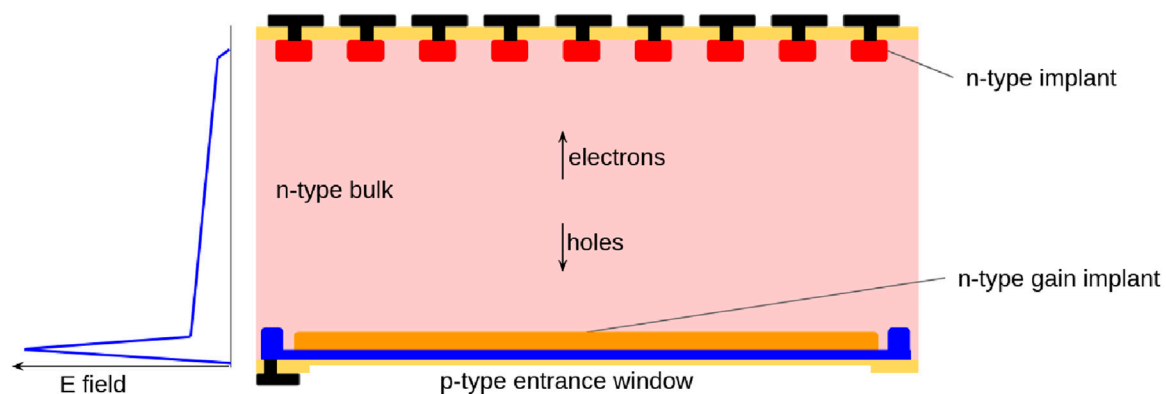


FIGURE 6
Schematic cross-section of a n-bulk inverted LGAD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom.

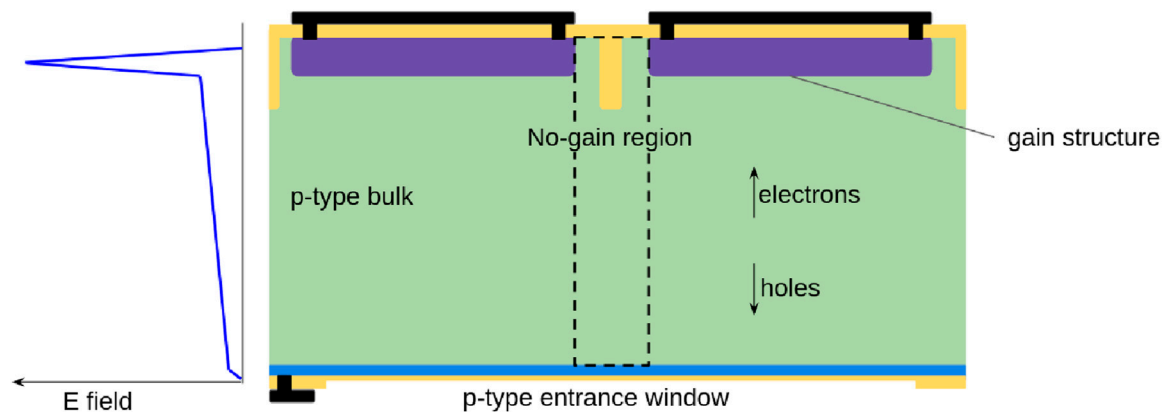


FIGURE 7
Schematic cross-section of a trench isolated LGAD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom. The entrance window was added for the application of these sensors to x-rays detection.

close to the entrance window are amplified with high gain (a gain of around 18 was reported). A new advancement in fabricating n-type i-LGADs for soft x-ray detection was also recently presented at a conference [37]. The work reported on the first LGAD prototypes equipped with a thin entrance window for photons with energy as low as 250 eV. The initial characterization showed a gain of about 7 when illuminated with 470 nm light, that corresponds to an attenuation length of $0.56\ \mu\text{m}$ in Silicon. The gain drops to 1.4 when illuminated with 940 nm photons, which interact deeper in the substrate (attenuation length equal to $54.6\ \mu\text{m}$) and generate hole-initiated multiplications.

The mentioned results make the n-type i-LGAD unfit for HEP applications but perfectly suited for low-energy physics experiments.

2.5 Trench-isolated LGAD (TI-LGAD)

An alternative technological approach to achieve fine segmentation in LGADs involves the use of narrow and shallow

trenches to isolate the pixels, as represented in Figure 7. This technology is based on etching a pattern of trenches into the silicon substrate, which are then filled with dielectric materials, such as silicon dioxide. The trenches, approximately $1\ \mu\text{m}$ wide, replace all the structures at the pixel border region, such as p-stop and JTE, while maintaining the same multiplying junction scheme ($\text{n}^+\text{-p}^-\text{-p}^+\text{-p}^-$) as standard LGADs. This design offers a clear advantage in terms of reducing the no-gain width between adjacent pixels, overcoming the technological limitations of the standard approach.

Thanks to the trench isolation technology, the nominal no-gain width can be reduced to less than $5\ \mu\text{m}$, compared to the usual few tens of microns in standard LGADs, achieving an outstanding 80% fill factor for a $50\ \mu\text{m}$ pitch sensor. TI-LGADs were proposed and initially produced by FBK based on thin epitaxial wafers [26], with a first characterization reporting an effective no-gain distance between pixels as low as $2\ \mu\text{m}$ [38].

TI-LGAD sensors manufactured with a double-sided process hold promise for soft x-rays detection. They can provide a

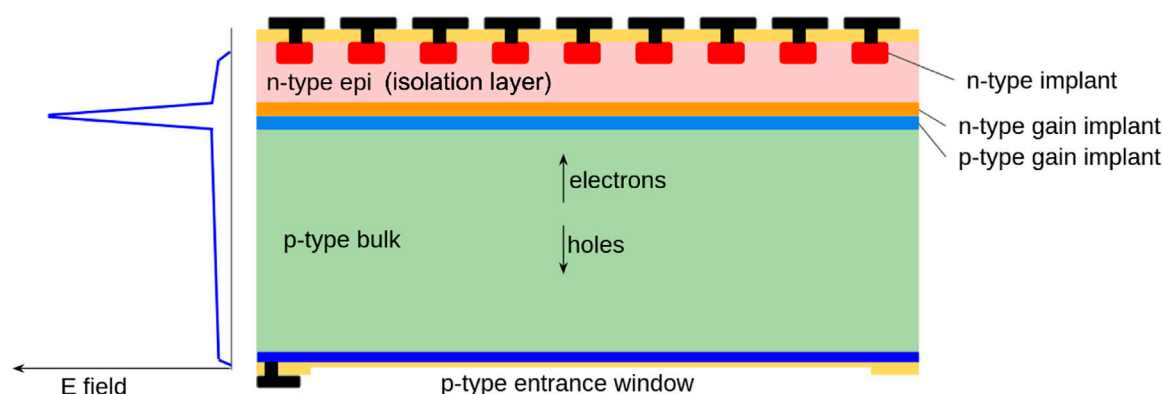


FIGURE 8
Schematic cross-section of a deep junction LGAD and magnitude of the electric field inside the sensor. Yellow represents dielectrics and black the metallization. The readout side is on top, and the x-rays enter from the bottom. The entrance window was added for the application of these sensors to x-rays detection.

reasonably high fill factor, small pixels down to $25\ \mu\text{m}$, and simultaneously offer the possibility to integrate an entrance window on the sensor backside. As of now, no experimental prototypes made on FZ thick substrates are available. Additionally, using a thicker active layer (e.g., $300\ \mu\text{m}$) necessary for a fully double-sided process could potentially be detrimental to the no-gain width. This effect is attributed to the peculiar shape of the electric field at the pixel periphery, as discussed in [26].

2.6 Deep-junction LGAD (DJ-LGAD)

Another promising LGAD technology, the deep-junction LGAD (DJ-LGAD), has been introduced and pioneered by the University of California, Santa Cruz [39]. A cross-sectional representation of the device is depicted in Figure 8. In this design, the multiplying junction forms a uniform area of n^+ and p^+ gain implants spanning the entire sensor area. Unlike other LGAD technologies, the multiplying junction is not superficial but is buried a few microns below the surface, where n^+ DC-coupled electrodes are positioned.

The bulk of the sensor comprises high-resistivity p-type silicon, and above the buried junction, a several-micron-thick layer of high-resistivity n-type material, referred to as the “isolation layer,” is present. At the operational bias voltage, both the n-type epi and p-type bulk regions are depleted. However, the multiplication region is confined between the buried n^+ and p^+ gain layers while elsewhere the electric fields are well below the threshold for impact ionization. The lateral termination of the multiplication structure can be realized by extending the n^+ gain layer to be close to a grounded guard ring while the p^+ gain layer is a few μm narrower than the n^+ layer, as suggested in [39]. Below the junction, the thicker layer of high-resistivity p-type material serves as the charge generation medium for detecting charged particles, x-rays, or heavy ions. An optical window can be included on the p-type side of the sensor, along with the back bias electrode.

Particles interacting in the n-type isolation layer undergo holes-initiated multiplication and do not experience significant multiplication. Conversely, electrons drifting from the p-type

bulk are multiplied in the gain layer and then drifted to be collected by the n^+ electrodes at the top surface. Proper tuning of the n^- and p-type regions allows for optimization of both noise and energy sensitivity for specific applications.

The unique design of this detector introduces some fabrication complexity and new challenges, primarily due to the necessity to manufacture the buried junction several microns into the substrate. One approach involves using a substrate with the two n- and p-type high-resistivity regions, obtained, for example, through a double epitaxy process or wafer-to-wafer direct bonding. High-energy ion implantation is then employed to place the two gain doping profiles at the n-p interface. However, this technique has intrinsic limitations due to the maximum achievable energy of standard ion-implantation equipment (a few MeV), restricting the junction depth to 4–5 microns from the sensor surface.

An alternative method involves implanting the two gain layers on two different wafers, which are then wafer-bonded together. The wafer carrying the n^+ gain layer is subsequently thinned down to a few microns, and its surface is processed to allocate the electrode structures. Another approach includes using a p-type substrate where both the two gain implants are done close to the surface, followed by the growth of the high-resistivity n-type layer by epitaxy.

Several institutes and research facilities are conducting R&D activities to produce first prototypes of DJ-LGADs, but preliminary results are not yet published, with only partial results presented in workshops [40].

3 Review of recent results from LGADs for X-ray detection

In the past few years LGAD sensors were tested as x-ray detectors, this section summarizes the results obtained in these studies proceeding in chronological order.

The very first measurements showed that LGAD sensors could be used as x-ray detectors using standard LGADs with a pad geometry of $5 \times 5\ \text{mm}^2$ and a thickness of $200\ \mu\text{m}$. These sensors demonstrated the detection of x-rays down to the energy of Mo

fluorescence (17.5 keV), with the lowest detectable energy being limited by the sensor capacitance [19].

The first tests with tender x-rays were done using standard LGADs developed for high energy physics, featuring an active thickness around 50 μm and a low resistivity substrate with a thickness in the order of 500 μm . The low resistivity substrate does not allow for a backside illumination of the sensors with soft x-rays, making it favourable to test them with the x-rays impinging on the front side of the sensors where a few μm of dielectrics and metal are usually present. Using a bunched x-ray beam with a repetition rate of 500 MHz, the detection of 6 keV photons was demonstrated using a sensor with pad geometry of $\sim 1 \times 1 \text{ mm}^2$. The same setup was able to distinguish the individual x-ray bunches, putting a lower limit to the rate capability of these sensors optimized for minimum ionizing particle timing [20]. In addition, an energy resolution between 8 and 15% was demonstrated for x-ray energies between 6 and 16 keV with the better resolution being achieved for the higher energy x-rays [20]. While the results presented so far used discrete component amplifiers and oscilloscopes to readout and digitize the sensors signal, the next step was to use readout ASICs developed for photon science.

Standard LGAD 5 mm long strip sensors with a pitch of 150 μm and active thickness of 50 μm were connected to both charge integrating [41, 42] and photon counting [43] readout electronics. The larger channel capacitance of the LGAD sensor (due to its strip geometry and thickness) compared to a standard silicon sensor resulted in increased noise and is detrimental to the energy resolution. Despite this effect, photons with energy of 2.1 keV were detected with the photon counting electronics, showing an improvement with respect to the 8.5 keV energy threshold achieved with a standard silicon sensor with the same electronics settings [44]. The energy resolution using the photon counting electronics was 0.31 keV for 2.1 keV x-rays, and the charge integrating electronics showed a resolution of 0.41 keV for 8.05 keV x-rays, with the latter figure representing a factor 2.7 improvement compared with a standard silicon strip sensor read out by the same charge integrating electronics [44]. The ASICs combined with a standard LGAD strip sensor resulted also in an improved energy resolution compared to the previous result obtained with standard LGADs read out using discrete component electronics. A dependence of the energy resolution on the x-ray energy is observed also in this study. Furthermore it is shown that there is an optimal operating voltage that gives the best resolution due to the rise of shot noise when further increasing the gain of the LGAD sensor [44]. The noise of the readout system when expressed in x-ray energy (keV equivalent) is shown to decrease with the LGAD operating voltage. This is due to the sensor gain that changes the conversion between collected charge and x-ray energy [44]. The results shown in [44] also highlight two main drawbacks of standard LGADs optimized for timing: the aforementioned channel capacitance with typical values around 3 pF) and the fill factor of these sensors that strongly limit their use in synchrotron radiation applications. The use of iLGADs can improve both aspects, lowering the channel capacitance and providing a 100% fill factor. A thin entrance window applied to these sensors would then allow for an improved quantum efficiency for soft x-rays.

These sensors were fabricated and considerations for their design, together with a first characterization, are shown in [45].

A $1 \times 1 \text{ cm}^2$ iLGAD pixel sensor with 25 μm pitch and 275 μm thickness was bonded to a charge integrating readout electronics [10]. The sensor leakage current and its noise were improved by cooling the sensor-ASIC assembly, resulting also in an increase of the sensor's gain [45].

Due to x-ray interactions, the use of iLGADs for synchrotron light applications causes the silicon dioxide to acquire a positive charge, that over time can impact the sensor performance. The effect of this type of radiation damage was explored for iLGADs and a modification of the sensor periphery resulted in an improved radiation hardness for x-rays [46].

The characterization of standard LGADs continues in parallel to iLGAD development. Standard LGAD pad sensors with an area of $1.3 \times 1.3 \text{ mm}^2$ and thickness of 20 and 50 μm were characterized using a 500 MHz bunched x-ray beam with energy between 5 and 70 keV [47]. The sensors were read out using discrete components electronics. The x-ray bunches could be resolved, maintaining the earlier result of the same group [20]. In this study the timing performance of the sensors was studied more accurately, showing a time resolution between 50 and 200 ps for LGADs and of ~ 80 ps for a 50 μm thick silicon sensor with the same pad geometry. The 20 μm thick LGAD performed better than the 50 μm thick ones, with the latter having a time resolution worse than 100 ps. The LGAD timing performance is significantly worse than what is observed for minimum ionizing particles (MIPs), usually around 30 ps. This is explained considering the shape of the deposited ionization cloud in the sensor: a cylinder along the whole sensor thickness for a MIP and a point-like distribution for x-rays. The shape of the ionization cloud together with the signal formation mechanism in LGADs and sensors without gain account for the different performance with the two particle types and the better resolution of a sensor without gain compared to an LGAD of the same thickness [47]. The effect of the different shapes of the ionization cloud for MIPs and x-rays also resulted in a reduction of the sensor gain for the latter [47]. A dependence of the sensor gain on the density of the ionization cloud reaching the gain layer was also studied in the detection of MIPs [48]. Both studies reach the shared conclusion that too high charge density, resulting from the multiplication, can lower the electric field in the gain layer, self-quenching the multiplication process. The energy resolution of the LGADs was studied and found to be between 6% and 20% depending on the bias voltage [47]. The sensors also show a worse resolution at higher operating voltages as shown in [44]. The linearity in the energy response of these LGAD pads was found to be better than 4% [47]. In the same study AC-LGAD strip sensors with a strip length of 5 and 10 mm were exposed to the same x-ray beam and demonstrated an energy resolution between 12% and 21% [47].

The last sensor characterization study to be mentioned in this section features the same 275 μm thick iLGADs used in [45] but with pad geometry. In the study the photocurrent generated by monochromatic x-rays with energies from 200 to 1000 eV was measured. The gain as a function of photon energy of iLGADs was measured, and their quantum efficiency (QE) was estimated using sensors without gain with the same TEW as the iLGADs [49]. The QE measurement of the sensors without gain shows a QE between 55% and 65% for 250 eV photons. Using the dependence of the QE on x-ray energy, the main parameters of the TEW were extracted, showing that the main source of inefficiency are the

dielectric layers used on the TEW, while the charge collection efficiency is close to unity [49]. The gain of the iLGAD shows a dependency on the photon energy, with higher gains being achieved at higher energies. This is explained by the distribution of the x-ray interaction positions relative to the multiplication structure [49]. The dependence of gain on x-ray energy is exploited to infer the gain as a function of the position of the x-ray interaction within the sensor. Different multiplication structures were studied, whose multiplication regions were placed at different depths in the sensors. The gain for electrons (holes) entering the multiplication structure M_n (M_p) was determined and it is shown that a thinner multiplication structure has a larger M_p/M_n ratio than a broader one [49]. The effect of the depth of interaction of the x-ray photons relative to the multiplication structure and the ratio between M_n and M_p are discussed in Section 6.

The characterization work done on LGADs using x-rays and the fabrication of sensors dedicated to photon science is a sign of interest in this technology. The performances of the sensors and the level of detail reached in their characterization are steadily evolving. These sensors are close to being employed in measurements at synchrotron light sources instead of being brought to the beamlines to be characterized. The first use of LGAD sensors for photon science measurements is mentioned in a conference [50] and documented in an article preprint [51] where an iLGAD pixel sensor was used in ptychography measurements with soft x-rays.

4 Thin entrance window for soft X-rays

Detecting soft and tender x-rays in the energy range of 200 eV to 2 keV presents new challenges, particularly in achieving high quantum efficiency. In hybrid silicon detectors, x-rays enter the sensors from the side opposite to the collecting junction where the read-out electronics is located. Similarly, in CMOS sensors, the wafer is thinned from the backside to approach the active volume of the sensor. In both cases, an entrance window must be designed and incorporated into the sensor backside to facilitate proper x-ray transmission and the collection of photogenerated charge in the sensor medium. The entrance window comprises two distinct regions: 1) an insensitive layer made of dielectric material or metal, deposited onto the silicon surface, and 2) a highly doped silicon layer (typically of the same sign as the substrate if the entrance window is opposite to the junction).

The QE for soft x-rays interacting closely to the surface may be affected by three loss mechanisms: 1) the absorption of photons in the insensitive layer; 2) the recombination of charge carriers generated by x-rays at the silicon surface, or 3) the loss of charge carriers in the neutral highly doped region via trap-assisted or Auger recombination. In simpler designs, the entrance window is typically made of a thin metal layer (usually aluminum), connected to a highly doped region. However, this design significantly affects QE at low energies due to absorption in the metal and recombination at the silicon-metal interface, where the surface recombination velocity (SRV) approaches the saturation value (approximately 10^6 cm/s). A more efficient approach involves using a thin layer of dielectric material to passivate the silicon surface and reduce SRV. Materials like thermal silicon dioxide can provide good surface passivation with SRV as low as 10 cm/s. Excellent alternatives are hydrogenated

silicon nitride (deposited by Plasma Enhanced Chemical Vapor Deposition - PECVD) [52] and aluminum oxide (deposited using Atomic Layer Deposition - ALD) [53]. Nevertheless, it is crucial to maintain the film thickness as low as possible due to the attenuation length of soft x-rays in the mentioned materials, which is comparable to that of Silicon (Figure 9).

Manufacturing of thin entrance windows (TEW) for soft x-rays poses challenges due to the strong dependencies of doping concentration profile, surface passivation, and carrier lifetime on specific fabrication processes. While it is difficult to provide a universal set of parameters, a well-passivated silicon surface via a dielectric thin-film coupled to a shallow and steep doping profile with a surface concentration in the order of 10^{19} atoms/cm³ promises good QE for energies down to 250 eV [56]. Only a few studies systematically investigated TEW for soft x-rays. One recent study investigated the manufacturing of TEW on n-type silicon sensors by using low-energy Arsenic ion implantation or Phosphorous diffusion, and a thin dielectric layer as passivation material. Using LED light they estimated a QE around 50% at 276 eV for wafers fabricated using As ion implantation annealed at high temperatures, and equal to 70% by using Phosphorous diffusion [57]. Another study on back-side illuminated CMOS sensors reported a remarkable QE above 90% in the 100–1000 eV range, corresponding to an equivalent 5 nm dead-layer [58]. This result was obtained on CMOS x-rays sensors back-thinned down to 9.5 μ m, and doped with low-energy ion implantation.

As discussed in this paper, the integration of a TEW into LGADs can occur on either the gain side or the bulk side, depending on the specific technology. For AC-LGADs, DC-RSDs, TI-LGADs, and DJ-LGADs, the TEW can be placed on the back side, opposite to the gain, following a similar approach used in standard silicon sensors. For iLGADs, where the TEW is positioned on the multiplication side of the sensors, certain constraints in tuning the doping profile and in the thermal budget have to be considered to preserve the multiplication junction functionality. One of the first studies that investigated the QE on a p-type iLGAD, reported a QE around 60% and 90% at 250 eV and 500 eV, respectively, with charge collection efficiency close to 1 in the full investigated energy range [49]. The encouraging outcomes suggest that incorporating an effective TEW in an iLGAD structure is viable, as the QE appears to be only limited by absorption in the dielectric material, and potential enhancement may be achieved by further reducing the thickness of this insensitive layer. Other studies have successfully demonstrated the integration of a TEW on n-type iLGAD using a 4 nm thick layer of aluminum oxide and 15 nm of aluminum, but no characterization with soft x-rays has been reported [36].

Recent advancements in microfabrication technologies suggested novel techniques for manufacturing TEW for low-energy particles or ultraviolet (UV) light. For instance, using low-energy (5 keV) Arsenic implantations followed by low-temperature (500°C) microwave annealing has been proposed to achieve unitary internal QE at 200 eV [56]. Other advanced techniques, such as plasma immersion ion implantation (PIII) followed by laser annealing, have already proven their effectiveness in detectors for UV light [59]. Compared to conventional ion implantation, PIII can implant at energies below 0.2 keV and when coupled with laser annealing it allows for obtaining a very shallow junction with an abrupt profile, by limiting the dopant diffusion. Alternative

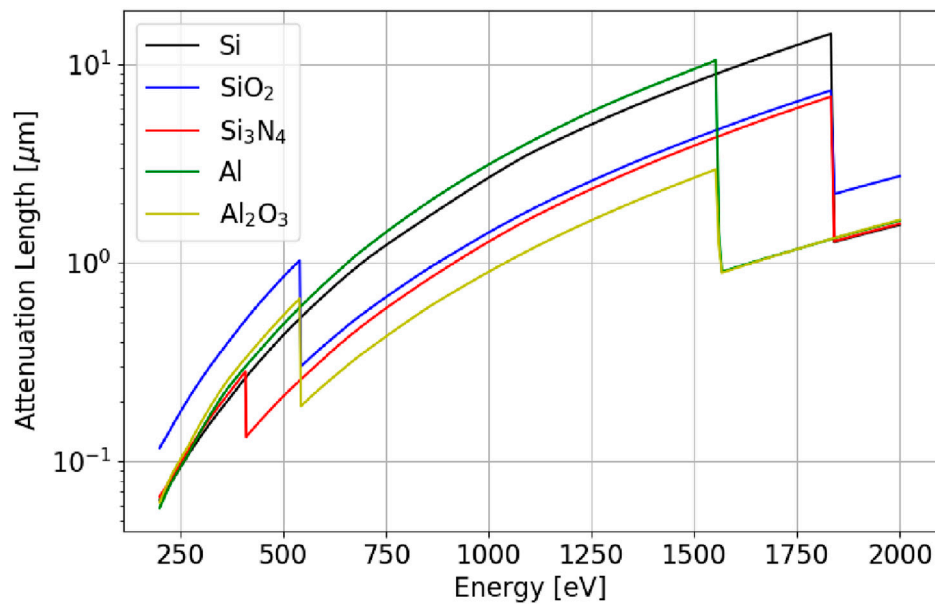


FIGURE 9

Attenuation length as a function of x-ray energy for Silicon and materials commonly used in the realization of entrance windows for x-rays. Data from [54], that is based on [55].

technologies to ion implantation have also been proposed: Pure-B deposition has been demonstrated able to produce shallow junctions for UV light detection below 220 nm [60], while superlattice-doping via Molecular Beam Epitaxy (MBE) reached record QE at deep and far UV wavelengths [61].

5 Theory summary of signal and noise of LGAD sensors

This section presents a summary of the theory used to describe the signal and noise properties of the multiplication structure of LGADs. The main references for this section use the local theory of impact ionization developed for avalanche photodiodes [62, 63]. A more accurate description takes into account the space needed by the charge carriers to reach the energy necessary for impact ionization (non-local impact ionization theory) [64, 65]. An example of the considerations presented below applied to APDs for light detection can be found in [66].

The multiplication structure of an LGAD can be represented as shown in Figure 10 where the multiplication of charge carriers happens in a region between 0 and w . In this region the ionization coefficients, denoted as α for electrons and β for holes, are larger than 0. The ionization coefficients are functions of the electric field and therefore of the position in the multiplication region. The currents traversing the system are I_n and I_p for electrons and holes respectively (the figure reports the direction of motion of the charge carriers). A generation rate $g(x)$ due to thermal effects or particle interaction is considered in the multiplication region. This general representation was used in [62] to derive expressions for gain and noise of avalanche photodiodes. This section summarizes the relevant elements for LGAD sensors. The multiplication factor $M(x)$ is the average total number of electron-hole pairs that are

generated in the multiplication region as a result of one initial pair being generated at a position x . The multiplication factor takes the form (Eq. 2)

$$M(x) = \frac{\exp\left[-\int_x^w \alpha - \beta dx'\right]}{1 - \int_0^w \alpha \exp\left[-\int_{x'}^w \alpha - \beta dx''\right] dx'} \quad (2)$$

Neglecting trapping and recombination effects, the multiplication factor varies as a function of x only in the multiplication region, remaining constant outside. For a given electric field, in silicon, $\beta < \alpha$ [67] resulting in a lower gain for holes entering in the gain region when compared to electrons ($M(0) < M(w)$). The gain of electrons and holes entering the multiplication structure is represented in the following by $M_n = M(w)$ and $M_p = M(0)$, respectively.

Considering the currents $I_p(0)$ and $I_n(w)$ being the result of thermal generation of charge carriers in the sensor, the total dark current of the device can be expressed as (Eq. 3)

$$I = I_p(0)M_p + I_n(w)M_n + \int_0^w gM dx \quad (3)$$

The multiplication affects the noise of the sensor by increasing its leakage current and due to the fluctuations of the multiplication for each charge carrier traversing the multiplication region. This in particular affects the shot noise of the sensor. For a readout bandwidth B , the current fluctuations due to shot noise are $\sigma_i = \sqrt{\phi B}$, with the noise spectral density ϕ expressed in units $[\phi] = A^2/Hz$. In [62] the shot noise spectral density takes the form

$$\phi = 2q_e \left\{ 2 \left[I_p(0)M_p^2 + I_n(w)M_n^2 + \int_0^w gM^2 dx \right] + I \left[2 \int_0^w \alpha M^2 dx - M_n^2 \right] \right\} \quad (4)$$

where q_e is the elementary charge. In LGAD sensors the width of the multiplication region is smaller than the active thickness of the

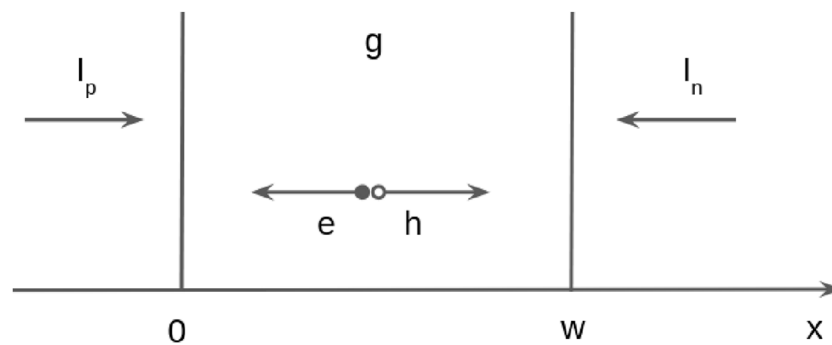


FIGURE 10

Geometry used in [61] to describe the multiplication region of LGADs. The electric field is high enough to result in multiplication between 0 and w . In the same area a generation rate g due to thermal effects or particle interaction is considered. I_p , I_n are the holes and electrons currents in the device, respectively. Note that the arrows below the currents I_p , I_n represent the direction of movement of the respective charge carriers.

sensor, representing at most a few percent of it. Assuming that generation effects related to the electric field in the multiplication region (e.g., trap assisted tunneling and field-enhanced emission) do not play a major role, the current generated in the multiplication region can be neglected in comparison to the current generated in the rest of the sensor volume. Disregarding the current generated in the multiplication region ($g = 0$) the shot noise spectral density can be expressed as $\phi = 2q_e I_p(0) M_p^2 F_p + 2q_e I_n(w) M_n^2 F_n$ where F_n , F_p are the excess noise factors for electrons and holes entering the multiplication structure, respectively. These represent the fluctuations of the multiplication and increase the noise compared to what is expected by Poisson statistics only. According to [62] the excess noise factors are given by

$$\begin{aligned} F_n &= 2 + \frac{1}{M_n} \left[2 \int_0^w \alpha M^2 dx - M_n^2 \right] \\ F_p &= 2 + \frac{1}{M_p} \left[2 \int_0^w \alpha M^2 dx - M_p^2 \right] \end{aligned} \quad (5)$$

As in silicon $M_n > M_p$, the holes excess noise factor is larger than the electrons one ($F_p > F_n$). The excess noise factors do not depend only on the values of gain at the edges of the multiplication structure, but are sensitive to the shape of the electric field in the multiplication region. In the approximation $\beta = k\alpha$ with k being constant, it is possible to show that the excess noise factors can be reduced if the multiplication of holes is kept at a minimum ($k \ll 1$) [62] and it is shown experimentally in [66]. If the gain for holes entering the multiplication region is not important for x-ray detection, reaching a given electrons gain M_n with a broader multiplication region results in lower excess noise factor values.

The shot noise can be expressed in units of electrons using a time constant τ related to the bandwidth of the readout electronics $\tau = 1/(2B)$, taking the form

$$\sigma_e = \sigma_i \frac{\tau}{q_e} A_3 \quad (6)$$

where A_3 accounts for the effect of the readout electronics, considering that electronics with the same bandwidth can result in different values of σ_e depending on their delta-pulse response function [68]. The A_3 parameter can assume values between 0.5 and

a few units [68], however, for the sake of simplicity, A_3 is set to 1 in the rest of this section. τ is also related to the time response of the readout electronics and is usually referred to as shaping time. Using Eq. 6 with $A_3 = 1$, the shot noise becomes (Eq. 7)

$$\sigma_e = \sqrt{N_n M_n^2 F_n + N_p M_p^2 F_p} \quad (7)$$

where N_n , N_p are the number of electrons and holes entering the multiplication region in a time τ , respectively. For an equal current value of electrons and holes entering the multiplication region ($N_n = N_p$), the contribution to the shot noise is larger for the electrons as $M_n > M_p$, so that $M_n^2 F_n > M_p^2 F_p$ (see Eq. 5).

The excess noise factor affects also the signal through its multiplication. An x-ray photon of a given energy creates an average of N_x electron-hole pairs in the sensor, these undergo multiplication resulting in a signal $S = N_x M$, where the value of M depends on the absorption position of the x-ray. The variance of the signal due to the multiplication of its charge carriers is given by [63] as (Eq. 8)

$$\sigma_{SM}^2 = N_x M^2 (F_{si} + F - 1) \quad (8)$$

where F_{si} is the Fano factor of silicon, and F is the excess noise factor at the absorption position of the x-ray. Eq. 5 shows the excess noise factors to be used if the x-ray interaction results in the injection of electrons or holes at the edges of the multiplication structure. If the x-ray interacts in the multiplication region at a position x , considering a point-like ionization distribution resulting from the interaction, the excess noise factor can be extracted from Eq. 4 similarly as done for F_n , F_p by imposing $g = g(x')\delta(x' - x)$ (where x' is the integration variable). The excess noise factor for $0 < x < w$ is (Eq. 9)

$$F(x) = 2 + \frac{1}{M(x)} \left[2 \int_0^w \alpha M^2 dx - M_n^2 \right] \quad (9)$$

By including all noise contributions other than the shot noise (e.g., thermal noise, readout electronics noise, etc.) in a term σ_0 , the total variance of the signal can be expressed as (Eq. 10)

$$\sigma_s^2 = N_x M^2 (F_{si} + F - 1) + \sigma_e^2 + \sigma_0^2 \quad (10)$$

The quantity $\sqrt{\sigma_e^2 + \sigma_0^2}$ represents the noise of the detector system when no x-ray interactions are present.

It is interesting to compare the signal-to-noise ratio (SNR) of a sensor with gain to the one of a sensor without gain ($M = 1$, $F = 1$). In the case of a sensor with gain, the SNR depends on the absorption position of the x-ray. For an x-ray being absorbed at a position x the SNR takes the form

$$\frac{S}{\sigma_s} = \begin{cases} N_x / \sqrt{N_x F_{Si} + N_n + N_p + \sigma_0^2} & \text{No gain} \\ N_x / \sqrt{N_x (F_{Si} + F(x) - 1) + N_n F_n \frac{M_n^2}{M^2(x)} + N_p F_p \frac{M_p^2}{M^2(x)} + \frac{\sigma_0^2}{M^2(x)}} & \text{Gain} \end{cases} \quad (11)$$

It is worth noting that the SNR of the sensor without gain is always better than the one with gain in case $\sigma_0 = 0$, that is if the only noise source is shot noise. In case $\sigma_0 > 0$, the sensor with gain reduces this term by the square of the gain at the absorption position of the x-ray. As in silicon the gain for electrons entering the multiplication region is higher than for holes ($M(w) > M(0)$), it is advantageous to have a sensor geometry where the x-rays absorption happens at $x > w$ resulting in the electrons generated by the x-ray interaction entering the multiplication region.

The conclusions that can be drawn by this summary are that, in order to achieve the best SNR, it is convenient to have the x-rays interactions in a position where the resulting electrons enter the multiplication region as they have a larger gain and lower excess noise factor compared to holes. At the same time, to reduce the shot noise, it is convenient that the thermal generation current from the sensor bulk results in holes entering the multiplication region as $M_p^2 F_p < M_n^2 F_n$. These conditions however can not be always fulfilled due to constraints in the sensor fabrication or the necessity to have a large enough thickness of silicon for the x-rays to interact. Different approaches to realize an LGAD sensor for x-rays are discussed in Section 6.

6 Technological approaches to LGADs for photon science

LGADs for x-ray detection can be realized in different ways, with different LGAD technologies being detailed in Section 2. In general, the multiplication structure can be placed on either the entrance window side of the sensor, on the segmented readout side, deep (more than a couple of μm) below the sensor surface. This section describes these technological choices and their impact on the performance and manufacturing of the sensors.

6.1 Gain structure on the readout side

In soft and tender x-ray detection, when dealing with a sensor thickness of approximately $200 \mu\text{m}$ or more, positioning the multiplication structure on the readout side of the sensor and allowing x-rays to enter from the opposite side leads to all x-ray interactions occurring in the bulk. As a result, the charge carriers undergo multiplication with a single average value of gain (M is either M_n or M_p). The polarity of the sensors bulk is usually chosen so that electrons enter the gain layer. This results in a larger value of gain for the charge carriers generated in the bulk, whether they result

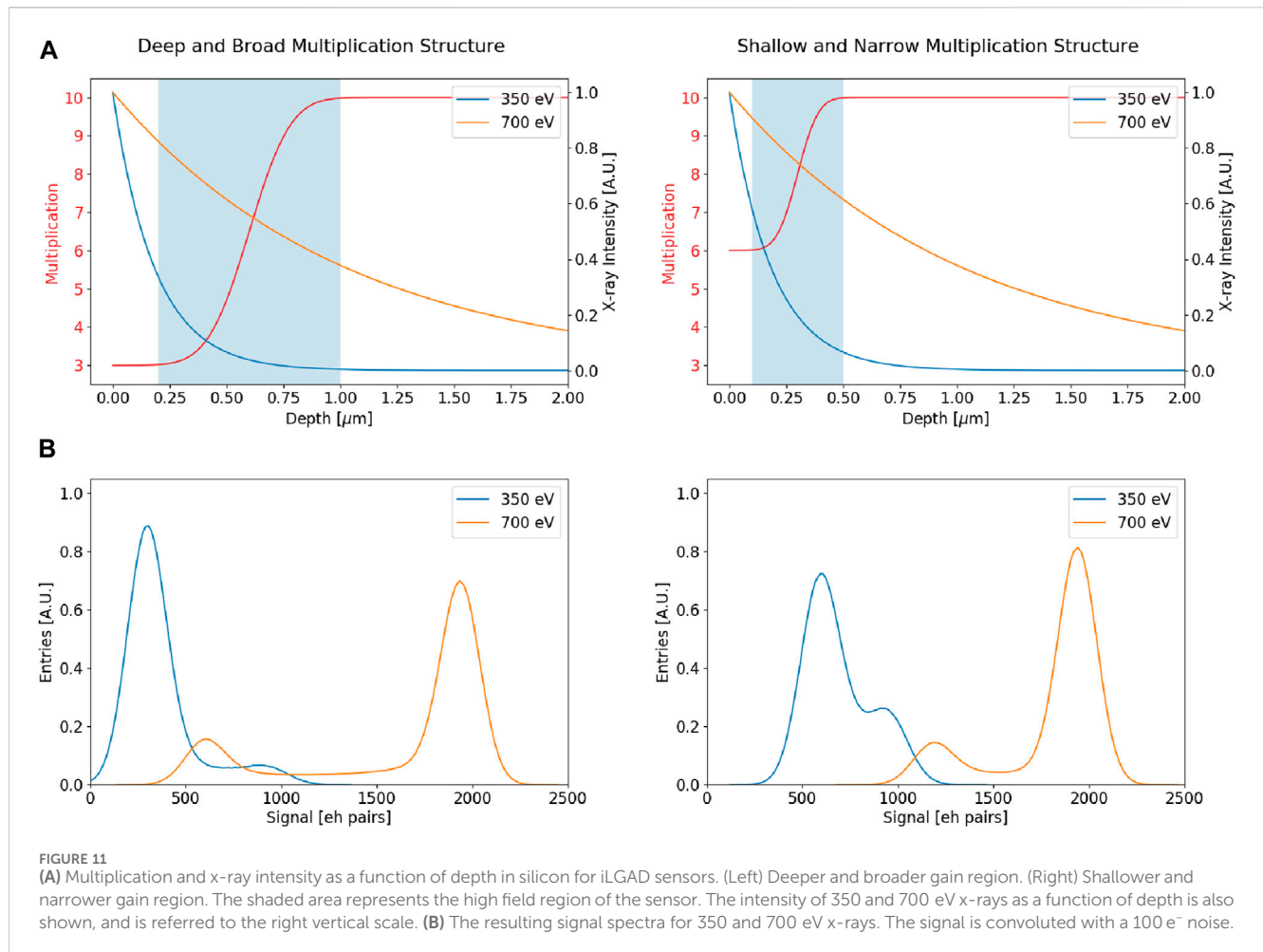
from an x-ray interaction or thermal generation. This configuration results in a better SNR compared to the choice of having holes from the bulk entering the multiplication region. This can be seen by comparing at the terms under the square root in Eq. 11 for a sensor with gain and imposing $M(x) = M_n$, $F(x) = F_n$, $N_p = 0$ for the case of a p-type bulk and $M(x) = M_p$, $F(x) = F_p$, $N_n = 0$ for an n-type bulk, with the same generation current so that N_n , N_p assume the same value when they are not 0. As the charge carriers entering from the readout side are considered to be negligible compared to the ones coming from the bulk N_n , N_p can be set to 0 depending on the bulk polarity. If electrons from the bulk are multiplied, the realization of a given gain would favour a multiplication structure as broad as possible while maintaining a bulk thick enough to absorb the x-rays. This reduces the electric field in the multiplication region and in turn the multiplication of holes, resulting in a reduction of the excess noise factors and an improved SNR.

Placing the multiplication structure on the readout side poses the challenge of the segmentation of the gain layer in order to achieve channel segmentation. This can result in volumes in the sensor where the charge carriers from an x-ray interaction do not undergo charge multiplication. This is the case for standard and trench isolated LGADS. In the case of standard LGADs, the segmentation of the gain layer is not suitable for reaching the pitch needed by pixel or strip detectors for synchrotron radiation experiments as their fill factor would be too poor. Trench isolated LGADs might fulfill the requirements, however their performance regarding the fill factor is not yet demonstrated on substrates of a suitable thickness. The resistivity of the most superficial doped layer of the multiplication region are exploited by AC coupled LGADs and the proposed DC-RSD to achieve channel segmentation without interrupting the gain layer. These sensors feature a continuous gain structure for the whole active area, avoiding regions where the charge carriers from x-ray interactions do not undergo multiplication.

Placing the multiplication structure close to the sensors surface allows to use dopant diffusion or ion implantation to realize the sensors, without the necessity of less common fabrication techniques. A further advantage of this approach is that the TEW can be developed and fabricated almost independently from the multiplication region allowing for an optimization of these components with less constraints.

6.2 Gain structure on entrance window side

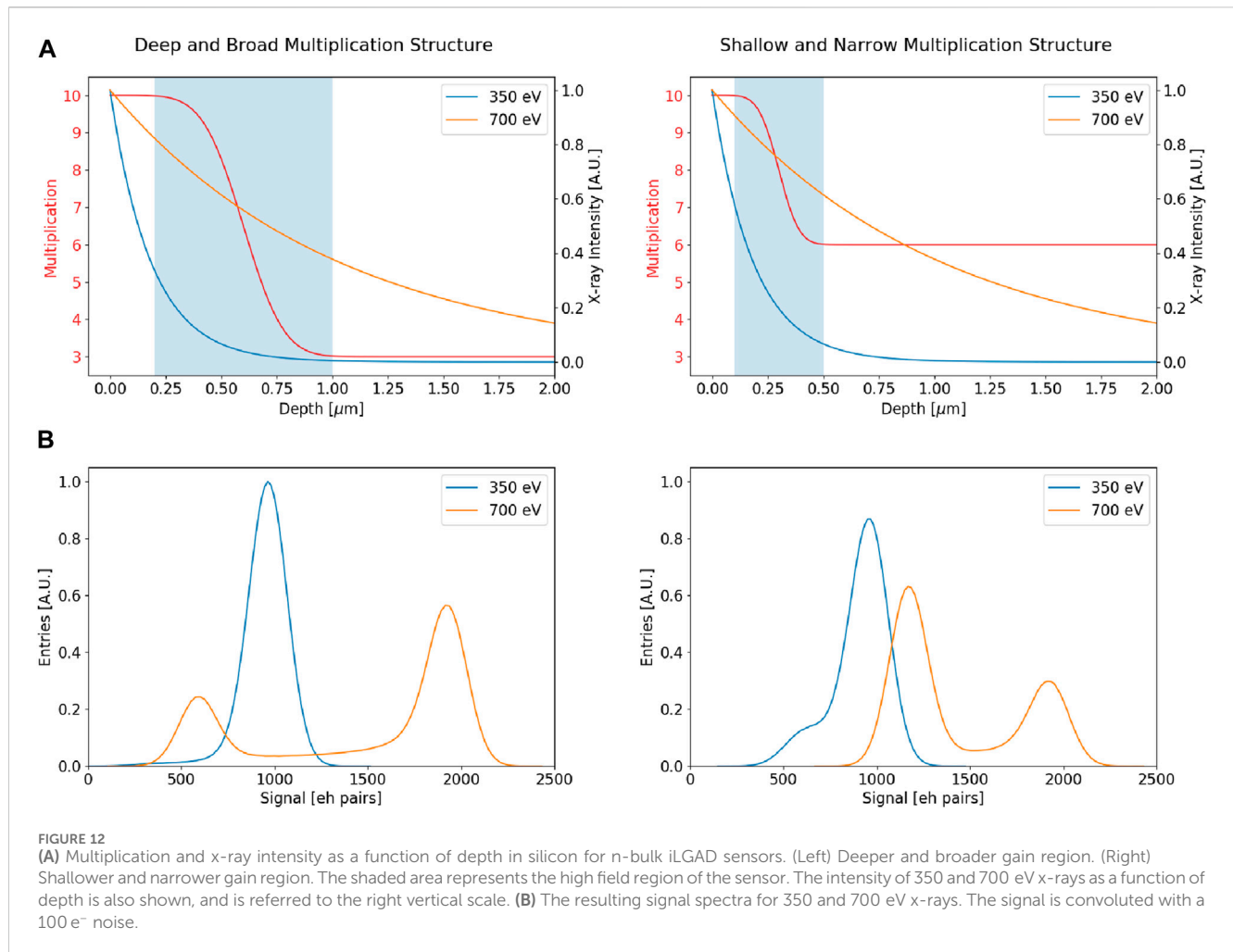
Placing the multiplication structure on the entrance window side of the sensor allows to achieve channel segmentation while avoiding to interrupt the multiplication structure. This avoids regions where the charge carriers from x-ray interactions do not reach the multiplication region. Inverted LGADs follow this design, and these are the only LGAD sensor type so far specifically developed for soft and tender x-ray detection [45]. With an absorption length spanning between $\sim 100 \text{ nm}$ and $\sim 10 \mu\text{m}$, x-rays in this energy region interact on both sides of the multiplication region when this is placed on the entrance window side of the sensor. This results in a spectrum of signals even for a single x-ray energy. These spectra are sensitive to the depth and width of the multiplication region, allowing for an optimization of these parameters. To look at a



possible optimization, two multiplication structures are considered in this subsection. A deep and broad multiplication structure with a high field region between 0.2 and 1 μm in the silicon, and a shallow and narrow structure presenting high field between 0.1 and 0.5 μm from the silicon surface. Both structures have a gain for electrons entering the multiplication region of $M_n = 10$. The gain for holes entering the multiplication region varies between the two designs, as a narrower multiplication structure requires a higher electric field to maintain the same gain for electrons. The gain for holes entering the multiplication region is $M_p = 3$ for the broad region and $M_p = 6$ for the narrow region. The gain and depth values chosen for these examples (as well as the $M(x)$ used later) do not come from an existing or proposed design and are arbitrary. They are however realistic as they fall in the region of possible values for an LGAD gain structure.

In the top row of Figure 11 the multiplication structures are applied to an iLGAD with p-type bulk (see Figure 5). The red line represents $M(x)$ and is referred to the left vertical scale, and the shaded area represents the high field region. Given the polarity of the bulk, the x-ray interactions taking place to the right of the multiplication structure result in electrons entering the high field region. The two colored lines represent the intensity of x-rays of two different energies (350 and 700 eV) and are referred to the vertical scale on the right. The entrance window is placed to the left of the multiplication structure, with the silicon starting at 0 and the bulk

extending for $\sim 200 \mu\text{m}$ to the right. The resulting signal spectra for the considered x-ray energies are shown in the bottom plots of Figure 11. The spectra are derived by considering the distribution of interaction depths for the different x-ray energies and the consequent gain value. The signal is convoluted with a $100 e^-$ noise, representing the total noise of the detector. This is again a value in the range of noises of readout electronics for synchrotron light applications [7, 8, 12]. The statistical fluctuations of the gain and creation of electrons hole pairs by the x-ray interaction are disregarded ($F = 1$, $F_{Si} = 0$). The area of the spectra is the same for all plots. Each spectra presents two peaks with different signal values for a given x-ray energy. These are the result of the x-ray interactions resulting in a gain of M_p or M_n for the charge carriers, with the peak at higher values being the one from electrons entering the gain region. The bottom left plot of Figure 11 shows the spectra for the deep and broad multiplication structure. The 350 eV spectrum for this structure is dominated by events where holes enter the multiplication region forming a peak at about $290 e^-$, while the 700 eV spectrum is dominated by events with electrons entering the gain structure with a peak at about $1940 e^-$. Looking at the spectra obtained from the shallow and narrow multiplication structure, bottom right plot of Figure 11, the effect of the different gain structure become evident. The higher value of M_p for the narrower structure results in a shift of the peak from events where the holes enter the multiplication region, making the



overall spectrum for each energy narrower. The narrower high field region also results in fewer interactions that receive values of gain between M_p and M_n . In addition, as the high field region is closer to the sensor surface, the fraction of events with M_n gain increases. The shallow and narrow multiplication structure results in improved spectra for a p-bulk iLGAD when compared to the deep and broad one. An element that is not evaluated in this example is the increased excess noise factor expected for a narrower gain structure that achieves the same M_n of a broader one. This affects both the fluctuation of the signal and shot noise, decreasing the sensor SNR. As the excess noise factor depends on the details of the electric field distribution in the gain layer no general conclusions can be drawn. The features of the spectra described in these examples are documented in presentations at conferences [69, 70].

It is interesting to see the effect of applying the multiplication structures described above to a n-bulk iLGAD sensor (see Figure 6). The results can be seen in Figure 12, where the same quantities as Figure 11 are shown, with the difference that the bulk to the right of the multiplication structure is n-type, resulting in the holes from the bulk entering the multiplication region. The only effect of the change in bulk polarity for this example is which side of the multiplication region reaches a gain M_n or M_p . The areas of the spectra in Figure 12 are the same as in Figure 11, allowing for a direct comparison. The spectrum of 350 eV x-rays for the deep and broad multiplication

structure (bottom left of Figure 12) is now dominated by events with M_n gain, while the 700 eV x-rays spectrum is similar to the one calculated for the same multiplication structure and p-bulk iLGAD. The shallow and narrow gain structure on a n-bulk iLGAD results in the spectra in the bottom right plot of Figure 12. For this structure the 700 eV x-ray spectrum is narrower, mainly as a result of the higher M_p of this structure. The 350 eV x-ray spectrum sees a reduction of the peak of events with M_n gain as the multiplication structure is closer to the silicon surface. In the case of the n-bulk iLGAD the deeper multiplication structure shows a better performance than the shallower one. Compared to the p-bulk iLGAD, the n-bulk iLGAD is a more suitable sensor for the lower energies of soft x-rays, while the p-bulk iLGAD performs better with higher energy x-rays. To better see the effect, Figure 13 shows the fraction of interactions resulting in M_p or M_n gain for a n-bulk iLGAD with a deep and broad multiplication structure and the same quantities for a p-bulk iLGAD with the narrow and shallow structure. The n-bulk iLGAD provides a gain of M_n predominantly to lower energy x-rays compared to the p-bulk iLGAD. However, if the gain for holes entering the multiplication structure is high enough, the SNR for the higher energy x-rays might be sufficient for detection as they create a larger number of charge carriers in their interaction with the silicon bulk. The n-bulk iLGAD, for low energy x-rays, can fulfill the conditions derived from the theory of

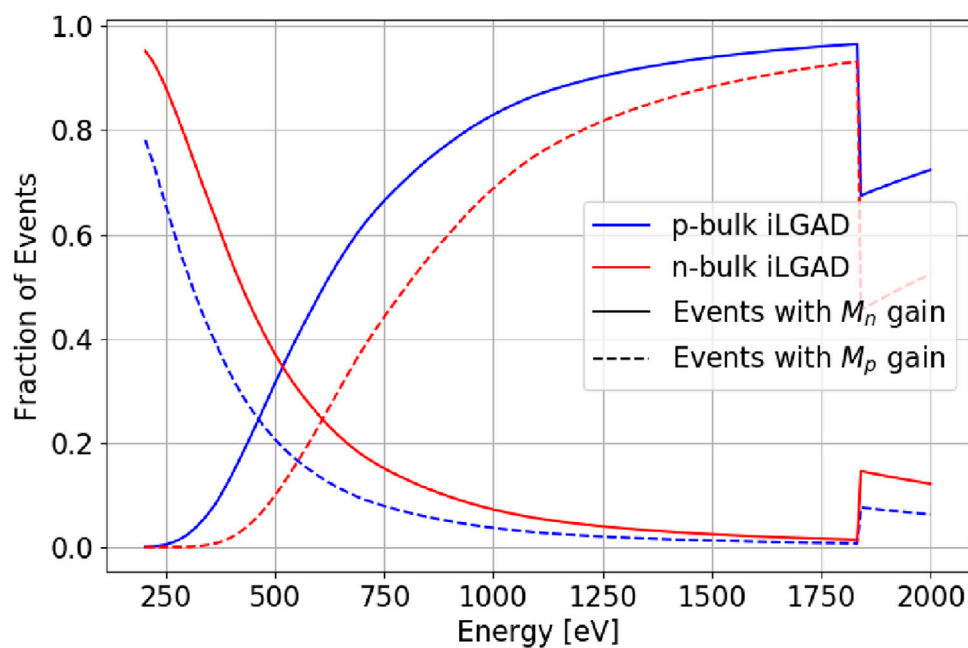


FIGURE 13
Fraction of events that results in the electrons (full lines) or holes (dashed lines) entering the multiplication region as a function of the x-ray energy. The iLGAD refers to the shallow and narrow multiplication structure shown in Figure 11 top right. The n-bulk iLGAD refers to the deep and broad multiplication structure shown in Figure 12 top left.

multiplication: a signal where the electrons enter the gain region and noise dominated by holes entering the gain region.

Placing the gain structure on the entrance window side of the sensor is the only strategy used so far for LGAD sensors dedicated to photon science applications, with the available sensors having a p-type bulk. N-bulk iLGADs were fabricated with the aim of detecting low-energy charged particles and have not yet been characterized using x-rays. The choice of the bulk polarity allows to choose which energy region receives the higher gain values. As a drawback, these sensors will produce a spectrum with two peaks for most x-ray energies, complicating the interpretation of measurements acquired with charge integrating or spectrally-discriminating ASICs.

Regarding the fabrication of these sensors, their complexity is increased with respect to sensors with the gain structure on the readout side. Both sides of the sensor require the realization of complex structures to allow for a fine-enough segmentation on the readout side and the gain structure coupled with the TEW on the opposite side. As the gain structure is close to the TEW, the optimization of these two components is entangled. Despite their complexity, common fabrication techniques can be used to realize these sensors as their components are placed close to the sensor surfaces.

6.3 Deep gain structure

Placing the gain structure a few μm or further from the readout side of the sensor allows for a continuous structure and a fine segmentation of the readout channels without areas where the charge carriers from the x-ray interactions do not receive

multiplication. If the gain structure is close to the readout side, almost all of the soft and tender x-ray interactions will result in electrons entering the multiplication region. A multiplication structure placed a few μm from entrance window of the sensor would result in part of the events where the holes enter the multiplication region receiving a lower gain. This is similar to the situation detailed in the previous sub section. The distance of the multiplication structure from the entrance window of the sensor determines which is the fraction of events resulting receiving M_n or M_p gain. This is shown in Figure 14 for a $1\ \mu\text{m}$ wide gain region placed at different depths in the sensor.

The dark current in a sensor with a deep gain structure results in both holes (from the n-type region) and electrons (from the p-type bulk) entering the multiplication region. Assuming the same generation rate for both the p- and n-type regions, placing the multiplication structure closer to the readout side increases the fraction of electrons from the dark current entering the multiplication region, and results in a higher noise compared to the same multiplication structure placed closer to the entrance window side of the sensor.

As the fraction of events receiving M_n gain favours a gain structure closer to the readout side and the noise favours a multiplication structure closer to the entrance window, an optimization of the position of the multiplication region is possible if a particular energy of x-rays is targeted for detection.

The realization of a deep gain structure requires methods not often employed in sensor fabrication. In case the structure is placed a few μm from one of the sensor surfaces, high energy ion implantation or epitaxial growth of silicon can be used. In the case of deeper structure, this can be realized through wafer-to-wafer bonding and eventual thinning of one of the starting wafers to

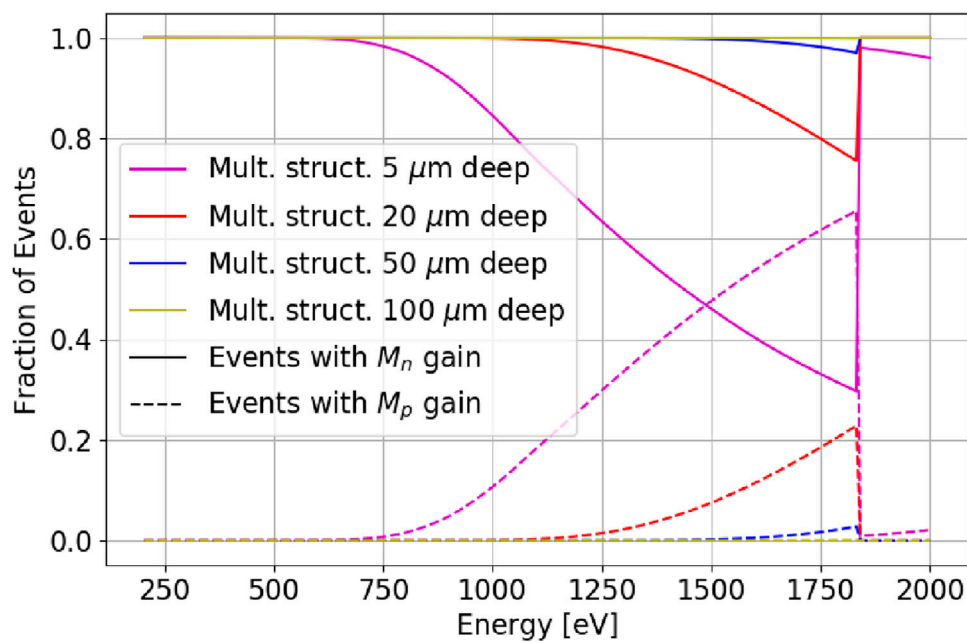


FIGURE 14
Fraction of events that results in the electrons (full lines) or holes (dashed lines) entering the multiplication region as a function of the x-ray energy. For DJ-LGADs with their multiplication structure at different depths (measured from the entrance window) in silicon.

achieve the desired position in the sensor. The wafer bonding and epitaxy procedures result in different silicon substrates being placed on either side of the multiplication structure, this can influence the optimization discussed above as the substrates can have different dark current generation rates.

7 Comparison of possible sensor structures

In this section different LGAD technologies are compared in their possible application in photon science. For this comparison it is assumed that all the sensor technologies are produced on thick-enough substrates and are equipped with a thin entrance window for soft x-rays. A few figures of merit are summarized in [Table 1](#) and described below.

7.1 Collected charge

The sign of collected charge might constraint the ASIC that can be connected to the sensor as some ASICs are optimized for either electrons or hole collection.

7.2 Signal

The sensor signal can either be uni- or bi-polar, the only sensor type with bipolar signals is the AC-LGAD. The AC-coupling of the channels renders impossible the use of these sensors with charge integrating electronics. The readout of AC-LGADs using photon counting ASICs is constraint by the time scale of the development of the bipolar signal: if the signals present both its positive and negative

lobe within the integration time constant of the ASIC amplifier, then the signal registered by the electronics might be too small to cross the set threshold. These considerations might render AC-LGADs incompatible with the available ASICs developed for photon science.

7.3 Gain structure side

The side where the gain structure is placed, together with the sensor polarity, has interesting implications for the signals produced by the sensors, as shown in [Section 6](#).

7.4 Fill factor

The fill factor of the sensors is qualitatively categorized. The standard LGAD technology does not allow for the channel segmentation necessary for photon science applications, while TI-LGADs might achieve a sufficient fill factor assuming that the performances obtained on $\sim 50 \mu\text{m}$ substrates hold true also for thicker substrates. The remaining sensor technologies are more promising in this regard.

7.5 Charge sharing

The charge sharing is classified in a qualitative manner. Standard LGADs and TI-LGADs sensors often do not offer this feature that is used to improve the position resolution of the sensors through interpolation. The other considered technologies feature charge sharing, with the AC-LGADs and DC-RSD having a resistive sharing mechanism instead of a charge sharing driven by charge carrier diffusion.

TABLE 1 Comparison of different LGAD technologies for their application in photon science.

	Standard	AC-coupled	Trench isolated	Inverted	N-bulk inverted	Deep junction	DC-RSD
Collected Charge	e ⁻	e ⁻	e ⁻	h ⁺	e ⁻	e ⁻	e ⁻
Signal	Unipolar	Bipolar	Unipolar	Unipolar	Unipolar	Unipolar	Unipolar
Gain Structure Side	Readout	Readout	Readout	TEW	TEW	Deep in the bulk	Readout
Fill Factor	Low	100%	Medium	100%	100%	100%	100%
Charge Sharing	No	Yes (resistive)	No	Yes	Yes	Yes	Yes (resistive)
Constant Gain	Yes	Yes	Yes	No	No	Yes	Yes
Fabrication Complexity	Low	Low	Medium	High	High	High	Medium
Demonstrated	Yes	Yes [†]	Yes [†]	Yes	Yes	Yes [†]	No
Integrated with TEW	No	No	No	Yes	Yes	No	No
“Large” Sensors*	Yes	Yes	Maybe	Yes	Yes	Yes	Maybe

* “Large” sensors are in the order of 4 × 4 cm² or larger.
† The sensor technology was either demonstrated on thin (45–60 μm) substrates or with a few prototypes.

7.6 Constant gain

The constant gain refers to the sensors response to monochromatic x-rays, as discussed in Section 6 the point of interaction of the x-ray photons can impact the gain for each event. All sensor technologies are expected to provide a constant gain for soft and tender x-rays with the exception of iLGADs and n-bulk iLGADs.

7.7 Fabrication complexity

The fabrication complexity qualitatively ranks the sensor technologies. All LGAD sensors are going to have a more complex fabrication process than standard silicon sensors due to the gain implant and the gain termination structure (junction termination extension). The fabrication complexity can have consequences on the cost of the sensors and their yield. Due to the presence of a thin entrance window all sensors require a double sided fabrication process with lithography operations performed on both sides of the silicon wafer. For sake of comparison, the standard and AC-coupled LGADs are assigned a “low” complexity as they present, in principle, the easiest production path. The TI-LGADs together with DC-RSD are assigned a “medium” complexity as the former require the etching and filling of trenches in the silicon, and the latter (yet to be produced) seem to require a precise control of the resistance between channels. The iLGADs of both bulk polarities and the DJ-LGADs are ranked with a “high” fabrication complexity. The iLGADs foresee complex structures on both sensor sides, with the realization of the TEW being entangled with the multiplication structure. The DJ-LGADs require fabrication processes that are not usually employed in standard silicon sensors in order to place the multiplication structure at a depth of more than a couple of μm in the sensor bulk.

7.8 Demonstrated

All sensor technologies considered in the paper have been demonstrated with sensors either produced for high energy

physics or photon science, with the exception of DC-RSD that have not yet been produced. The AC-coupled and TI-LGADs were produced so far on thin (45–60 μm) substrates and their performance is yet to be characterized on thicker substrates. Only a few DJ-LGAD sensors have been produced, so this technology is not considered to be at the same level as others.

7.9 Integrated with TEW

For soft and tender x-ray detection the integration of a TEW with the sensor is necessary. So far only iLGADs of both bulk polarity have been produced with an integrated TEW.

7.10 “Large” sensors

The last figure of merit is the possibility to produce “large” sensors, with a dimension of about 4 × 4 cm² or larger. The possibility of covering a sufficient area is fundamental in order to obtain all the necessary data from measurements at synchrotron light sources, providing a large enough acceptance without gaps in the sensitive surface of the sensors. The distance between the sample being studied and the sensor is also important in order to exploit the “lever arm” so that the relevant features projected on the sensor appear far enough from each other to be resolved by the sensor. In this context a sensor with a too-small area will need to be placed either too close to the sample for this separation to be present or will not cover enough of the solid angle to collect all the needed information. In this section, the considered limit in realizing large sensors is given by the lithographic technology necessary to realize the sensors. TI-LGADs require the fabrication of trenches with a width of less than 1 μm, a feature size that is only possible with projection lithography. This lithography technique allows to expose only a part of the silicon wafer, limiting the size of a device obtained with a single exposure. A large device could be divided in smaller exposure areas that fit the exposure area and be composed in

multiple exposures (the so-called stitching method). This path could be taken to produce “large” sensors requiring projection lithography. These considerations hold true also for the foreseen DC-RSD where the contact between the metallization and the n^+ layer might require the use of projection lithography.

Looking at the sensor performances, the most promising sensor technologies for photon science applications seem to be iLGADs of both bulk polarities and DJ-LGADs. The favoured bulk polarity of iLGADs depends on the energy of the x-rays to be measured. DJ-LGADs would be more versatile, with a single technology covering a larger energy range, but this technology is less mature. Finally, DC-RSDs could be an interesting technology for photon science applications, however these sensors have yet to be produced.

8 Summary

Photon science measurement in the 200–2000 eV energy region can be applied to a wide variety of research topics. The currently available hybrid detectors find a limited use in this energy region due to their quantum efficiency and signal-to-noise ratio. LGAD sensors feature internal charge multiplication, providing a possible improvement of the signal-to-noise ratio of the detector system. Different LGAD technologies are available, and are described in this paper. A summary of the results obtained through with different LGAD technologies in x-ray detection is provided. The quantum efficiency of the sensors can be improved by the use of a thin entrance window, and the most relevant aspect of this component are described. A summary of the theory describing charge multiplication in LGADs is given and is used in the discussion of the features obtained by placing the LGAD multiplication structure in different parts of the sensor. Finally, the sensor technologies are compared using figures of merit relevant for their application in photon science measurements. The most promising technologies seem to be iLGADs of both bulk polarities, DJ-LGADs, and DC-RSD. These show different technological maturity, with DC-RSD not having been proven yet. The characterization work done on LGADs using x-rays and the fabrication of sensors dedicated to photon science is a sign of interest in this technology. The performances of the sensors and the level of detail reached in their characterization are steadily evolving. These sensors are close to being employed in measurements at synchrotron light

sources as opposed to being brought to the beamlines to be characterized.

Author contributions

MCV: Formal Analysis, Investigation, Visualization, Writing–original draft, Writing–review and editing. GP: Investigation, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to thank prof. Lucio Pancheri of University of Trento for the discussion regarding the theory of signal, gain, and noise. The development of LGAD sensors is seen by the authors as a result of the welcoming environment provided by the RD50 collaboration that allowed to easily share ideas among different groups working on silicon sensors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Wagner A, Duman R, Henderson K, Mykhaylyk V. In-vacuum long-wavelength macromolecular crystallography. *Acta Crystallogr Section D* (2016) 72:430–9. doi:10.1107/S2059798316001078
- Rose M, Senkbeil T, von Gundlach AR, Stühr S, Rumancev C, Dzhibaev D, et al. Quantitative ptychographic bio-imaging in the water window. *Opt Express* (2018) 26:1237–54. doi:10.1364/OE.26.001237
- Hitchcock AP. Soft X-ray spectromicroscopy and ptychography. *J Electron Spectrosc Relat Phenomena* (2015) 200:49–63. Special Anniversary Issue: Volume 200. doi:10.1016/j.elspec.2015.05.013
- Pfeiffer F. X-ray ptychography. *X-ray ptychography* (2018) 12:9–17. doi:10.1038/s41566-017-0072-5
- Scagnoli V, Staub U, Bodenthin Y, de Souza RA, García-Fernández M, Garganourakis M, et al. Observation of orbital currents in CuO. *Science* (2011) 332:696–8. doi:10.1126/science.1201061
- Rajendran R, Mochi I, Helfenstein P, Mohacs I, Redford S, Mozzanica A, et al. Towards a stand-alone high-throughput EUV actinic photomask inspection tool: RESCAN. In: Sanchez MI, editor. *Metrology, inspection, and process control for microlithography XXXI*, 10145. Bellingham, WA: International Society for Optics and Photonics SPIE (2017). p. 101450N. doi:10.1117/12.2258379
- Andr  M, Barten R, Bergamaschi A, Br ckner M, Casati N, Cervellino A, et al. Towards MYTHEN III - prototype characterisation of MYTHEN III.0.2. *J Instrumentation* (2019) 14:C11028. doi:10.1088/1748-0221/14/11/C11028
- Tinti G, Bergamaschi A, Cartier S, Dinapoli R, Greiffenberg D, Johnson I, et al. Performance of the EIGER single photon counting detector. *J Instrumentation* (2015) 10:C03011. doi:10.1088/1748-0221/10/03/C03011
- Mozzanica A, Bergamaschi A, Br ckner M, Cartier S, Dinapoli R, Greiffenberg D, et al. Characterization results of the JUNGFR U full scale readout ASIC. *J Instrumentation* (2016) 11:C02047. doi:10.1088/1748-0221/11/02/C02047

10. Ramilli M, Bergamaschi A, Andrae M, Brückner M, Cartier S, Dinapoli R, et al. Measurements with MÖNCH, a 25 μm pixel pitch hybrid pixel detector. *J Instrumentation* (2017) 12:C01071. doi:10.1088/1748-0221/12/01/C01071
11. Allahgholi A, Becker J, Delfs A, Dinapoli R, Goettlicher P, Greiffenberg D, et al. The adaptive gain integrating pixel detector at the European XFEL. *J Synchrotron Radiat* (2019) 26:74–82. doi:10.1107/S1600577518016077
12. Llopart X, Alozy J, Ballabriga R, Campbell M, Casanova R, Gromov V, et al. Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning. *J Instrumentation* (2022) 17:C01044. doi:10.1088/1748-0221/17/01/C01044
13. Cartier S, Kagias M, Bergamaschi A, Wang Z, Dinapoli R, Mozzanica A, et al. Micrometer-resolution imaging using MÖNCH: towards G_2 -less grating interferometry. *J Synchrotron Radiat* (2016) 23:1462–73. doi:10.1107/S1600577516014788
14. Zhang J, Andrä M, Barten R, Bergamaschi A, Brückner M, Chirioti-Alvarez S, et al. Design and first tests of the gotthard-II readout ASIC for the European X-ray free-electron laser. *J Instrumentation* (2021) 16:P04015. doi:10.1088/1748-0221/16/04/P04015
15. Sadrozinski HF, Seiden A, Cartiglia N. 4D tracking with ultra-fast silicon detectors. *Rep Prog Phys* (2017) 81:026101. doi:10.1088/1361-6633/aa94d3
16. Pellegrini G, Fernández-Martínez P, Baselga M, Fleta C, Flores D, Greco V, et al. Technology developments and first measurements of Low Gain Avalanche Detectors (LGAD) for high energy physics applications. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2014) 765:12–6. doi:10.1016/j.nima.2014.06.008
17. Monaco V, Ali OH, Bersani D, Abujami M, Boscardin M, Cartiglia N, et al. Performance of lgad strip detectors for particle counting of therapeutic proton beams. *Phys Med Biol* (2023) 68:235009. doi:10.1088/1361-6560/ad02d5
18. Duranti M, Vagelli V, Ambrosi G, Barbanera M, Bertucci B, Catanzani E, et al. Advantages and requirements in time resolving tracking for astroparticle experiments in space. *Instruments* (2021) 5:20. doi:10.3390/instruments5020020
19. Moffat N, Bates R, Bullough M, Tartoni N. A novel detector for low-energy photon detection with fast response. In: *2018 IEEE nuclear science symposium and medical imaging conference proceedings (NSS/MIC)* (2018). p. 1–4. doi:10.1109/NSSMIC.2018.8824404
20. Galloway Z, Gee C, Mazza S, Ohldag H, Rodriguez R, Sadrozinski HW, et al. Use of “LGAD” ultra-fast silicon detectors for time-resolved low-keV X-ray science. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 923:5–7. doi:10.1016/j.nima.2019.01.050
21. Giacomini G. Fabrication of silicon sensors based on low-gain avalanche diodes. *Front Phys* (2021) 9:618621. doi:10.3389/fphy.2021.618621
22. Paternoster G, Arcidiacono R, Boscardin M, Cartiglia N, Cenna F, Dalla Betta GF, et al. Developments and first measurements of ultra-fast silicon detectors produced at FBK. *J Instrumentation* (2017) 12:C02077. doi:10.1088/1748-0221/12/02/C02077
23. Cartiglia N, Staiano A, Sola V, Arcidiacono R, Cirio R, Cenna F, et al. Beam test results of a 16 ps timing system based on ultra-fast silicon detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2017) 850:83–8. doi:10.1016/j.nima.2017.01.021
24. Sola V, Arcidiacono R, Boscardin M, Cartiglia N, Dalla Betta GF, Ficorella F, et al. First FBK production of 50 μm ultra-fast silicon detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 924:360–8. doi:10.1016/j.nima.2018.07.060
25. Galloway Z, Fadeyev V, Freeman P, Kougoukousis E, Gee C, Gruy B, et al. Properties of HPK UFSD after neutron irradiation up to $6 \times 10^{15} \text{ n/cm}^2$. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 940:19–29. doi:10.1016/j.nima.2019.05.017
26. Paternoster G, Borghi G, Boscardin M, Cartiglia N, Ferrero M, Ficorella F, et al. Trench-isolated low gain avalanche diodes (ti-lgads). *IEEE Electron Device Lett* (2020) 41:884–7. doi:10.1109/led.2020.2991351
27. Giacomini G, Chen W, D’Amen G, Tricoli A. Fabrication and performance of AC-coupled LGADs. *J Instrumentation* (2019) 14:P09004. doi:10.1088/1748-0221/14/09/p09004
28. Mandurrino M, Arcidiacono R, Boscardin M, Cartiglia N, Dalla Betta GF, Ferrero M, et al. Demonstration of 200-100- and 50- μm pitch resistive AC-coupled silicon detectors (RSD) with 100% fill-factor for 4D particle tracking. *IEEE Electron Device Lett* (2019) 40:1780–3. doi:10.1109/LED.2019.2943242
29. Cartiglia N, Arcidiacono R, Borghi G, Boscardin M, Costa M, Galloway Z, et al. LGAD designs for future particle trackers. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2020) 979:164383. doi:10.1016/j.nima.2020.164383
30. Tornago M, Arcidiacono R, Cartiglia N, Costa M, Ferrero M, Mandurrino M, et al. Resistive AC-Coupled Silicon Detectors: principles of operation and first results from a combined analysis of beam test and laser data. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2021) 1003:165319. doi:10.1016/j.nima.2021.165319
31. Menzio L, Arcidiacono R, Borghi G, Boscardin M, Cartiglia N, Centis VM, et al. DC-coupled resistive silicon detectors for 4D tracking. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1041:167374. doi:10.1016/j.nima.2022.167374
32. Pellegrini G, Baselga M, Carulla M, Fadeyev V, Fernández-Martínez P, García MF, et al. Recent technological developments on LGAD and iLGAD detectors for tracking and timing applications. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 831:24–8. doi:10.1016/j.nima.2016.05.066
33. Dalla BGF, Pancheri L, Boscardin M, Paternoster G, Piemonte C, Cartiglia N, et al. Design and TCAD simulation of double-sided pixelated low gain avalanche detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2015) 796:154–7. doi:10.1016/j.nima.2015.03.039
34. Bisht A, Borghi G, Boscardin M, Centis Vignali M, Ficorella F, Hammad Ali O, et al. Development of LGAD sensors at FBK. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022). 1039:167018. doi:10.1016/j.nima.2022.167018
35. Segal J, Kenney C. New thin-entrance window LGAD for soft X-ray detection at LCLS. In: *2020 IEEE nuclear science symposium and medical imaging conference (NSS/MIC)* (2020). p. 1–3. doi:10.1109/NSS/MIC42677.2020.9507940
36. Khalid W, Valantan M, Doblas A, Flores D, Hidalgo S, Konrad G, et al. First results for the pLGAD sensor for low-penetrating particles. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1040:167220. doi:10.1016/j.nima.2022.167220
37. Koybasi O, Kok A, Povoli M, Segal J, Kenney C. Design, fabrication, and preliminary test results of a new inverse-LGAD for soft X-ray detection. In: *Presentation at the 13th international “hiroshima” symposium on the development and application of semiconductor tracking detectors (HSTD13)*. Vancouver, Canada: The next “Hiroshima” Symposium (HSTD13) will be held at Simon Fraser University SFU (2023). Available from: <https://indico.cern.ch/event/1184921/contributions/5574849/> (Accessed June 26, 2024).
38. Senger M, Macchiolo A, Kilminster B, Paternoster G, Centis Vignali M, Borghi G. A comprehensive characterization of the TI-LGAD technology. *Sensors* (2023) 23:6225. doi:10.3390/s23136225
39. Zhao Y, Ayyoub S, Chen W, Gee C, Islam R, Mazza SM, et al. A new approach to achieving high granularity for silicon diode detectors with impact ionization gain. *J Phys Conf Ser* (2022) 2374:012171. doi:10.1088/1742-6596/2374/1/012171
40. Seiden A, Schumm BA, Gee C, Sadrozinski H, Ott J, Mazza SM, et al. Deep Junction LGAD: a new approach to high granularity LGAD. In: *Presentation at the 41st RD50 workshop 2022* (2022). Available from: <https://indico.cern.ch/event/1132520/contributions/5140036/> (Accessed June 26, 2024).
41. Zhang J, Andrä M, Barten R, Bergamaschi A, Brückner M, Dinapoli R, et al. Performance evaluation of the analogue front-end and ADC prototypes for the Gotthard-II development. *J Instrumentation* (2017) 12:C12052. doi:10.1088/1748-0221/12/12/C12052
42. Zhang J, Andrä M, Barten R, Bergamaschi A, Brückner M, Dinapoli R, et al. Towards Gotthard-II: development of a silicon microstrip detector for the European X-ray Free-Electron Laser. *J Instrumentation* (2018) 13:P01025. doi:10.1088/1748-0221/13/01/P01025
43. Mozzanica A, Bergamaschi A, Dinapoli R, Gozzo F, Henrich B, Kraft P, et al. MythenII: a 128 channel single photon counting readout chip. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2009) 607:250–2. Radiation Imaging Detectors 2008. doi:10.1016/j.nima.2009.03.166
44. Andrä M, Zhang J, Bergamaschi A, Barten R, Borca C, Borghi G, et al. Development of low-energy X-ray detectors using LGAD sensors. *J Synchrotron Radiat* (2019) 26:1226–37. doi:10.1107/S1600577519005393
45. Zhang J, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Development of LGAD sensors with a thin entrance window for soft X-ray detection. *J Instrumentation* (2022) 17:C11011. doi:10.1088/1748-0221/17/11/C11011
46. Doblas A, Flores D, Hidalgo S, Moffat N, Pellegrini G, Quirion D, et al. Inverse LGAD (iLGAD) periphery optimization for surface damage irradiation. *Sensors* (2023) 23:3450. doi:10.3390/s23073450
47. Mazza S, Saito G, Zhao Y, Kirkes T, Yoho N, Yerdea D, et al. Synchrotron light source X-ray detection with low-gain avalanche diodes. *J Instrumentation* (2023) 18: P10006. doi:10.1088/1748-0221/18/10/P10006
48. Currás E, Fernández M, Moll M. Gain reduction mechanism observed in low gain avalanche diodes. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1031:166530. doi:10.1016/j.nima.2022.166530
49. Liguori A, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, Boscardin M, et al. Characterization of iLGADs using soft X-rays. *J Instrumentation* (2023) 18:P12006. doi:10.1088/1748-0221/18/12/P12006
50. Mozzanica A, Bergamaschi A, Schmitt B, mezza D, Greiffenberg D, Baruffaldi F, et al. Photon science detector development at PSI (2023). *Presentation at High Precision X-Ray Measurements* (2023). Available from: <https://agenda.infn.it/event/34523/contributions/200941/> (Accessed June 26, 2024).
51. Butcher TA, Phillips NW, Chiu CC, Wei CC, Ho SZ, Chen YC, et al. *Ptychographic nanoscale imaging of the magnetoelectric coupling in freestanding BiFeO₃* (2023). doi:10.48550/arXiv.2308.13465
52. Hegedüs N, Balázs K, Balázs C. Silicon nitride and hydrogenated silicon nitride thin films: a review of fabrication methods and applications. *Materials* (2021) 14:5658. doi:10.3390/ma14195658

53. Ott J, Pasanen TP, Repo P, Seppänen H, Vähäniemi V, Savin H. Passivation of detector-grade float zone silicon with atomic layer deposited aluminum oxide. *physica status solidi (a)* (2019) 216:1900309. doi:10.1002/pssa.201900309
54. X-Ray Interactions With Matter. X-ray interactions with matter (1995). Available from: https://henke.lbl.gov/optical_constants/ (Accessed June 26, 2024).
55. Henke B, Gullikson E, Davis J. X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E = 50\text{--}30,000\text{ eV}$, $Z = 1\text{--}92$. *At Data Nucl Data Tables* (1993) 54:181–342. doi:10.1006/adnd.1993.1013
56. Segal J, Kenney C, Kowalski JM, Kowalski JE, Blaj G, Rozario L, et al. Thin-entrance window process for soft x-ray sensors. *Front Phys* (2021) 9:618390. doi:10.3389/fphy.2021.618390
57. Carulla M, Vignali MC, Barten R, Baruffaldi F, Bergamaschi A, Borghi G, et al. Study of the internal quantum efficiency of FBK sensors with optimized entrance windows. *J Instrumentation* (2023) 18:C01073. doi:10.1088/1748-0221/18/01/C01073
58. Harada T, Teranishi N, Watanabe T, Zhou Q, Bogaerts J, Wang X. High-exposure-durability, high-quantum-efficiency (>90%) backside-illuminated soft-X-ray CMOS sensor. *Appl Phys Express* (2019) 13:016502. doi:10.7567/1882-0786/ab5b5e
59. Torregrosa F, Laviron C, Milesi F, Hernandez M, Faik H, Venturini J. Ultra shallow p+/n junctions using plasma immersion ion implantation and laser annealing for sub $0.1\text{ }\mu\text{m}$ CMOS devices. *Nucl Instr Methods Phys Res Section B: Beam Interactions Mater Atoms* (2005) 237:18–24. doi:10.1016/j.nimb.2005.04.108
60. Shi L, Nihtianov S, Scholze F, Gottwald A, Nanver L. High-sensitivity high-stability silicon photodiodes for duv, vuv and euv spectral ranges. In: *UV, X-ray, and gamma-ray space instrumentation for astronomy XVII (SPIE)*, 8145 (2011). p. 219–27. doi:10.1117/12.891865
61. Hoenk ME, Nikzad S, Carver AG, Jones TJ, Hennessy J, Jewell AD, et al. Superlattice-doped silicon detectors: progress and prospects. *High Energy Opt Infrared Detectors Astron VI* (2014) 9154:324–36. doi:10.1117/12.2057678
62. McIntyre R. Multiplication noise in uniform avalanche diodes. *IEEE Trans Electron Devices* (1966) ED-13:164–8. doi:10.1109/T-ED.1966.15651
63. Teich M, Matsuo K, Saleh B. Excess noise factors for conventional and superlattice avalanche photodiodes and photomultiplier tubes. *IEEE J Quan Elect* (1986) 22:1184–93. doi:10.1109/JQE.1986.1073137
64. McIntyre R. A new look at impact ionization-Part I: a theory of gain, noise, breakdown probability, and frequency response. *IEEE Trans Electron Devices* (1999) 46:1623–31. doi:10.1109/16.777150
65. Yuan P, Anselm K, Hu C, Nie H, Lenox C, Holmes A, et al. A new look at impact ionization-Part II: gain and noise in short avalanche photodiodes. *IEEE Trans Electron Devices* (1999) 46:1632–9. doi:10.1109/16.777151
66. Redus RH, Farrell R. Gain and noise in very high-gain avalanche photodiodes: theory and experiment. In: Hoover RB, Doty FP, editors. *Hard X-ray/gamma-ray and neutron optics. Sensors, and applications*, 2859. Bellingham, WA: International Society for Optics and Photonics SPIE (1996). p. 288–97. doi:10.1117/12.245118
67. Currás Rivera E, Moll M. Study of impact ionization coefficients in silicon with low gain avalanche diodes. *IEEE Trans Electron Devices* (2023) 70:2919–26. doi:10.1109/TED.2023.3267058
68. Gatti E, Manfredi P, Sampietro M, Speziali V. Suboptimal filtering of 1/-noise in detector charge measurements. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1990) 297:467–78. doi:10.1016/0168-9002(90)91331-5
69. Liguori A, Mozzanica A, Bergamaschi A, Schmitt B, Cuenca CL, Ruder C, et al. Investigation of soft X-ray detection using iLGAD sensors. In: *Presentation at the 18th "Trento" workshop 2023* (2023). Available from: <https://indico.cern.ch/event/1223972/contributions/5262043/> (Accessed June 26, 2024).
70. Carulla M, Baruffaldi F, Bergamaschi A, Boscardin M, Vignali MC, Dinapoli R, et al. Soft X-ray detection with single photon resolution using LGAD sensors. In: *Presentation at high precision X-ray measurements 2023* (2023). Available from: <https://agenda.infn.it/event/34523/contributions/200935/> (Accessed June 26, 2024).



OPEN ACCESS

EDITED BY

Cornelia B. Wunderer,
Helmholtz Association of German Research
Centres (HZ), Germany

REVIEWED BY

Pietro Zambon,
Dectris AG, Switzerland
Hayden Salway,
University of Cambridge, United Kingdom

*CORRESPONDENCE

Michael Fiederle,
✉ michael.fiederle@fmf.uni-freiburg.de

RECEIVED 04 March 2024

ACCEPTED 21 June 2024

PUBLISHED 30 July 2024

CITATION

Fiederle M and Baumbach T (2024), Perspective
of perovskite-based X-ray hybrid pixel
array detectors.
Front. Phys. 12:1395690.
doi: 10.3389/fphy.2024.1395690

COPYRIGHT

© 2024 Fiederle and Baumbach. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Perspective of perovskite-based X-ray hybrid pixel array detectors

Michael Fiederle^{1,2*} and Tilo Baumbach²

¹Materials Research Center FMF, Albert-Ludwigs-University Freiburg, Freiburg, Germany, ²Institute for Photon Science and Synchrotron Radiation IPS, Karlsruhe Institute of Technology, Karlsruhe, Germany

Compound semiconductors are playing a major role in the production of X-ray pixel detectors for the application in laboratories and beamlines at photon sources. The performance of these detectors has constantly been improved for the last decades but experiments are still limited by the properties of the detector material, especially under high flux illumination. The fast development of perovskite crystals opens the possibility for new materials to be used as highly efficient X-ray pixel detectors. The published data until now, of the transport properties, demonstrate the large potential of perovskite semiconductors. The achieved values are comparable with the ones of CdTe-based detectors. This paper presents potential perovskite-based detector materials and compares their performance with the state-of-the-art CdTe-based detectors. The perspectives of perovskite semiconductors are promising for the production of large area X-ray detectors but still some challenges remain.

KEYWORDS

X-ray detectors, high-Z semiconductors, perovskites, pixel detectors, synchrotron detectors

1 Introduction

During recent years the use of Hybrid Pixel Array Detectors (HPAD) has become a standard for photon detection at synchrotron storage ring and free-electron sources for certain applications like X-ray diffraction of crystalline materials and Non-Destructive Testing by Computed Tomography. It is generally accepted that photon counting detectors offer a large improvement over current detectors based on CCD cameras. Examples are the Pilatus [1], the Medipix [2] and the XPAD [3] detectors, which have been developed in various places in Europe.

Several research projects have investigated the possibilities of achieving highly efficient detector materials and thus successfully developed the technology for their use at photon sources. In this context, a large number of semiconductors have been investigated as possible sensor materials in the last 30 years, such as PbO [4], HgI₂ [5], TlBr [6], GaAs [7] and Cd-based semiconductors like CdTe, Cd (Te,Se) and (Cd,Zn) Te (CZT) [8].

The common feature of these detector-materials is the dependence of the detection efficiency on the thickness of the sensor. X- and gamma-ray detectors are bulk devices in contrast to most electronic semiconductor devices. The need for a high detection efficiency defines the requirements for the material properties.

- High-Z for efficient absorption of X- and gamma-rays, especially at high photon energies >20 keV.
- High resistivity ρ for low leakage current ($\rho > 10^8 \Omega \text{ cm}$)
- High mobility-lifetime product $\mu\tau$ of both charge carrier types for an optimum mean free path length ($\mu\tau > 10^{-3} \text{ cm}^2 \text{ V}^{-1}$)
- High mobility at least of one charge carrier for prompt signal formation

- Homogeneous distribution of electronic material properties on a micrometer scale for active detector areas with several cm².
- High structural quality given by single crystalline material.
- Low defect concentration.
- Reliable contact and interconnection technology to produce hybrid pixel detectors on a micrometer scale.
- Reliable production sources and availability at a reasonable cost

So far only the Cd-based semiconductor materials fit all requirements and yield the correct performance. Thus, the detector systems using CdTe and CZT currently define the standards regarding noise (leakage current), efficiency (mobility lifetime products $\mu\tau_e$ and $\mu\tau_h$), spatial resolution (pixel pitch) and homogeneity (defect concentrations). Since 2021 Computed Tomography systems using CdTe detectors have been installed in clinics for the application in human medicine [9].

An overview about CdTe-based pixel detectors is shown in [10] and by Gruner et al. for general perspectives for hard X-ray pixel detectors [11].

There is a growing need for large area, highly effective X-ray pixel detectors. The application of X-ray Diffraction at higher X-ray energies at light sources as well as Computed Tomography need effective pixel detectors with several cm² active area. For instance, reducing radiation dose for operators and patients is only possible if high-Z detectors are available. This leads to enhanced diagnostics information in the medical field and applications in inspection and analysis. The commercial use of highly effective X-ray pixel detectors is still limited by the costs and availability of the semiconductor materials. Novel, radical production techniques have to be considered to use the advantages of semiconductor X-ray sensors. Fast deposition techniques used in the development of flat panel detectors could be one possibility if the required material quality can be obtained for a new semiconductor material.

Yet, up to now, the production of large area sensors based on highly efficient semiconductors has been limited. These limitations are related to the growth and fabrication processes. The production of CdTe-based semiconductors is very complex. The materials are usually produced by melt growth techniques with temperatures above 1,100°C including toxic gases and solutions. This complexity generates a limitation in the production of the material and of the devices. Therefore, alternative materials are required.

Beyond these semiconductors a new class of materials are now in the focus of scientific research, namely, the perovskite materials. The term perovskite refers to a type of mineral structure that has a specific crystallographic arrangement. The crystal structure following the formula ABX₃, where A and B are positively charged ions and X is a negatively charged ion. The name “perovskite” originates from the mineral calcium titanate (CaTiO₃), which was first discovered in the Ural Mountains of Russia and named after the Russian mineralogist Lev Perovski.

This material class already yields a breakthrough for the application in photovoltaic cells. The perovskite films have pushed the photovoltaics market level up by combining low-cost production with high efficiency, thus exceeding the performance of silicon-based cells and achieving a similar performance within a few

years instead of decades [12]. Perovskites represent a class of materials offering tailored performance in different fields of application.

Perovskite films can be simply produced through wet chemistry and show excellent charge transport properties with mobility-lifetime products ($\mu\tau$) better than 10⁻⁴ cm²/V. The requirements for X-ray detectors, for example, the high $\mu\tau$ and homogeneous material properties are similar to the photovoltaic cells, adding high absorption and required thickness. Recent publications already demonstrated the possibility to use perovskite for the development of X-ray detectors [13–15].

The published data of the perovskite materials [16] showed remarkable results similar to the performance of state-of-the-art of detector material such as CdTe. Using the right class of perovskites can lead to non-toxic detectors with high absorption efficiency for X-ray energies up to 120 keV. This will open the opportunity to produce large area detector systems with high efficiency and low costs with increased sustainability.

An overview of a selection of available perovskite detectors is given in the next paragraph. This will be followed by the detailed presentation of the most promising materials. The discussion will compare the performance of the selected perovskite crystals with the state-of-the-art of CdTe and CZT based pixel X-ray detectors. A summary and perspective of the new class of materials in the field of X- and gamma-ray detectors shall finalize the paper.

2 Overview of X-ray detectors based on perovskites

Perovskites represent a class of materials containing insulators and semiconductors. This paper concentrates only on the perovskite semiconductors due to their capability of direct detection of X-rays. Subdivisions are organic, e.g., Methylammonium Lead Iodide MAPbI₃ and Methylammonium Lead Bromide MAPbBr₃, as well as inorganic materials, like CsPbBr₃ or Cs₂AgBiBr₆. Latest review articles [17]; [15, 18]; provide a nearly complete list of perovskite materials working either as scintillator materials or semiconductors.

The sensitivity is given by Kasap (2000) [19]. The X-ray sensitivity *S* of the detector is defined as the charge *Q* released per unit area *A* per unit exposure of radiation *X*: *S* = *Q*/*AX*. The Schubweg limited X-ray sensitivity *S* is proportional to Φ given by:

$$\Phi = \frac{5.45 \cdot 10^{13} \cdot q}{\alpha_{air} / \rho_{air} \cdot W_{pair}} \cdot \frac{\alpha_{en}}{\alpha}$$

Where *W*_{pair} is the pair creation energy, *q* the elementary charge, α_{air}/ρ_{air} is the mass energy-absorption coefficient of air. α , α_{en} are the mass energy-absorption coefficient and linear attenuation coefficient of the semiconductor material, respectively.

References of the table [20–31]:

In Table 1 a selection of perovskite semiconductors is presented showing important material properties for X-ray detectors. Part of the data are taken from Lin, C.-F. et al. [15]. From Table 1 three materials have been selected representing the best compromise between material properties and available size and thickness of detector materials. These materials are the organic based perovskites MAPbI₃ or MAPbBr₃, the single crystalline material

TABLE 1 Comparison of material properties of selected perovskite crystals.

Crystal structure	Materials	Crystal type	Growth method	Thickness (mm)	Electric field E (V mm ⁻¹)	Product of mobility-lifetime μT (cm ² ·V ⁻¹)	Sensitivity S (μC·Gy _{air} ⁻¹ ·cm ⁻²)	Ref.
ABX ₃ (Organic)	DMAPbI ₃	Single crystals	ITC	1.2 ± 0.04	NA	7.2 × 10 ⁻³	1.18 × 10 ⁴	[20]
	GAMAPbI ₃	Single crystals	ITC	1.2 ± 0.04	NA	1.3 × 10 ⁻²	2.31 × 10 ⁴	[20]
	MAPbBr ₃	Single crystals	ITC	NA	0.83	4.1 × 10 ⁻²	259.9	[21]
	MA _{0.6} Cs _{0.4} PbBr ₃	Single crystals	Solution	2	NA	4.64 × 10 ²	2017	[22]
	FAPbBr ₃	Single crystals	Solution	1	225	1.56 × 10 ⁻⁴	21,386	[23]
ABX ₃ (Inorganic)	CsPbBr ₃	Single crystals	Solution	1	20	(2.5 ± 0.2) × 10 ⁻³	1,256	[23]
	CsPbBr ₃	Quasi-monocrystal	Hot pressing	0.24	4.2	1.32 × 10 ⁻²	55,684	[26]
	CsPbBr ₃	Single crystals	Melt (Bridgman)	2	100	8.2 × 10 ⁻³	n.a	[23]
1D	CsPbI ₃	Polycrystals	Solution	NA	NA	3.63 × 10 ⁻³	2,370	[27]
A ₂ B ₂ X ₆	Cs ₂ AgBiBr ₆	Single crystals	Solution	2	NA	NA	105	[28]
	Cs ₂ AgBiBr ₆	Single crystals	Solution	2	25	NA	105	[29]
Large area process	MAPbI ₃	Polycrystals	Doctor blade coating	0.83	10–240	1.0 × 10 ⁻⁴	3.8 × 10 ³	[30]
	CsPbI ₂ Br	Polycrystals	ALS process	0.04	125	NA (mobility = 1.14 cm ² /Vs)	148,000	[31]

TABLE 2 Compares the material properties of CdTe and CdZnTe with the most promising perovskite material CsPbBr₃.

Crystal	Growth method	Density (g/cm)	Band gap (eV)	Resistivity (Ωcm)	Pair creation energy (eV)	μ _e τ _e (cm ² /V)	μ _e (cm ² /Vs)	μ _h τ _h (cm ² /V)	μ _h (cm ² /Vs)	Energy resolution @662 keV (percent)
CdTe [65]	THM	5.85	1.44	10 ⁹	4.43	10 ⁻³	800	10 ⁻⁴	80	0.5
CdZnTe [65]	THM	5.78	1.57	10 ¹⁰	4.64	10 ⁻¹	1,000	10 ⁻⁵	60	0.5
CsPbBr ₃ [66]	Bridgman	4.73	2.23	10 ¹⁰	6.61	10 ⁻⁴	22	10 ⁻³	22	5.5

CsPbBr₃ and the lead-free semiconductors CsAgBiBr₆. The presented perovskite crystals have two different structural appearances either thin films or bulk crystals.

Methylammonium lead perovskites, often referred to as MAPbX₃ (X stands for a halide such as iodine, bromine, or chlorine), are a class of hybrid organic-inorganic materials that have gained significant attention in the field of photovoltaics, particularly in the development of solar cells.

The ABX₃ perovskite crystals can be purely inorganic or hybrid-organic materials. The melting point are higher for the inorganic crystals, e.g., CsPbBr₃ has a melting point of 567°C [32].

These perovskite materials have the chemical formula CH₃NH₃PbX₃, where the methylammonium (MA) cation occupies the A-site, lead (Pb) occupies the B-site, and the halide (X) ions occupy the X-site in the perovskite crystal structure. The MAPbI₃ is the most investigated perovskite semiconductor with applications in photovoltaic cells as well as in X-ray detectors. The production of areas up to 20 × 20 mm² and with thickness up to 1 mm of thin films and bulk crystals is quite simple and the achieved transport properties fit the defined requirements for X-ray detectors.

The material properties of the crystals given in Table 1 demonstrate the high potential for successful X-ray detection in

terms of the requirements listed above. The typical thickness is around 1 mm. The absorption efficiency can be calculated using the average Z number of the material and thickness [33]. The efficiency is close to 90 percent for absorbing 50 keV photons with a thickness of 1 mm for CsPbBr₃ [34]. The MA-based perovskites yield an almost identical efficiency at 90 percent for 50 keV photons [18].

The resistivity of the intentionally undoped semiconductors is higher than $10^9 \Omega\text{cm}$, resulting in a low leakage current for applied electrical voltages up to 200 V. The most important properties are the values for the product of lifetime and mobility of the charge carriers. The multiplication of the $\mu\tau$ with the applied electrical field yields the mean free path length and is directly correlated with the charge collection efficiency of the detector material. The perovskite reach $\mu\tau$ -products of $6 \times 10^{-2} \text{ cm}^2\text{V}^{-1}$, which corresponds to a mean free path length of 12 cm for an electric field of 200 V/mm. The importance of this material property becomes apparent in comparison with conventional X-ray detector semiconductors like GaAs and CdTe. In the review article [35] the transport properties are compared to results from CdTe crystals. This comparison shows a value of $10^{-3} \text{ cm}^2/\text{V}$ for CdTe and several data points higher than $10^{-3} \text{ cm}^2/\text{V}$ for MAPI and Cs₂AgBiBr₆.

The crystals of these selected perovskites were produced by deposition techniques, solution growth processes as well as by melt growth method known as the Bridgman method [36]. The inverse temperature crystallization (ITC) is a solution-based process working at moderate temperature below 350°C.

In semiconductor growth, the Bridgman method is a technique used to produce high-quality single-crystal semiconductor materials. The drawback of the Bridgman method is the high growth temperature above the melting point of the material. This generates additional defects in the semiconductor.

The next paragraphs will compare perovskite crystals CsPbBr₃ with state-of-the-art pixel detectors based on CdTe and CZT. In the last years remarkable results have been presented for CsPbBr₃ regarding efficiency, size of detectors, and energy resolution - and therefore CsPbBr₃ appears to hold the most promise for application in HPADs.

3 Comparison of perovskite-based X-ray detectors with state-of-the-art pixel detector materials

Several crystals of the perovskites fulfil the requirements for a semiconductor X-ray detector given in the first section of this paper. Several classical semiconductors fulfil these requirements as well, however, only CdTe-based materials CdTe and (Cd,Zn)Te CZT demonstrated the capability for the use as a HPAD [10]. The data are presented in Table 2.

The density, the bandgap, the resistivity, and the $\mu\tau$ -products of CsPbBr₃ are comparable to the ones of CdTe and CZT. The mobility values in the table are lower and this has a strong influence on the signal formation. Consequently, the signal formation leads to a poor energy resolution. The mobilities are strongly correlated with the quality and purity of the material. Less concentration of structural defects and impurities will improve the mobilities. Z. Zhang and B. Saparov [37] collected the data of the mobilities of different perovskite crystals. Values above $100 \text{ cm}^2/\text{Vs}$ have been achieved.

However, the performance of CsPbBr₃ is quite promising taking into account the few years of research applied in the field of perovskite materials. Purification processes and advanced preparation of precursor materials have just been started.

The first published paper about CdTe was in 1955 [38]. It has taken over 60 years for this material to be successfully applied as a HPAD in Human Computed Tomography scanners, in laboratories and beamlines [9–11]. These applications usually use pixel pitches between 50 μm and 250 μm . This implies homogeneous properties on the micrometer scale for active detector areas with several cm^2 , regarding the specified requirements from the introduction.

The success of CdTe-based detectors is controlling the concentration of structural defects and impurities. In CdTe and CZT the concentration of deep levels is relatively high. It is a very complex balance of shallow and deep levels required to obtain the high values for the $\mu\tau$ -products for both carrier polarities [39]. In the last decades the purification of the material and the reduction of defect concentrations are continuously improving. Nonetheless, they are still high compared to silicon but in orders of magnitude, lower than in other compound semiconductors. The concentration of the defects has a strong influence on the $\mu\tau$ -products for electrons and holes, respectively. Defects are electronically active and act as traps for charge carriers. They can be classified by cross sections and relaxation times.

The application for X-ray detection in modern storage rings and XFELs [40] or in Computed Tomography [41] place significantly higher demands on the material and the available technologies. One common feature is the high photon flux. In Human CT the dose rate is $6 \times 10^8 \text{ photons/mm}^2\text{s}$. Latest publications identified a balance between the $\mu\tau$ -products of electrons and holes as a mandatory requirement for the applications under high photon fluxes [42]. A comparison with photon fluxes from $8 \times 10^7 \text{ photons/mm}^2\text{s}$ up to $1 \times 10^{10} \text{ photons/mm}^2\text{s}$ given in [43] discussed the effect of afterglow with different types of materials and configurations with CdTe ohmic and Schottky contacts, GaAs, latest “high flux CZT” by Redlen and silicon p-i-n structure. In the latter device no polarization effect is visible due to very small concentration of deep levels. The edge-on geometry has been applied to increase the absorption length up to 11.4 mm and maximize the efficiency [44].

High flux applications [16] and the influence of deep levels are part of the actual research of perovskite semiconductors. At the moment the database of deep levels is small but further improvements of the electronic properties are the next logical steps. This leads to more research in the involved defects. Several papers have been published analysing the nature and influence of deep levels in perovskites:

- Thin film and bulk material are characterized by Deep Level Transient Spectroscopy DLTS as well as by Photo Induced Current Transient Spectroscopy PICTS to determine the influence on the passivation of solar cells [46] and for optoelectronic devices [47,48].
- The role of impurities is the subject of the analysis as well as the identification of native defects in publications [45,49,50].
- In Figure 1 native defects and hydrogen related deep levels are summarized with their possible levels of ionization. Figure 1 was originally published in [49].

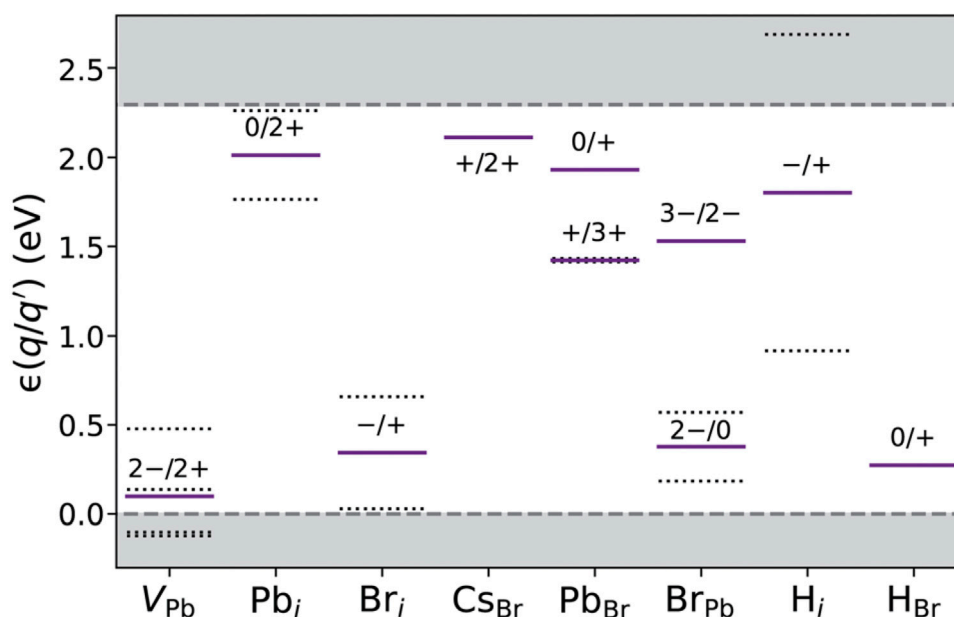


FIGURE 1
Thermodynamic charge state transition levels [45].

4 Discussion of the perspective of perovskite-based X-ray detectors

The improvement of solar cells is the main goal in the research of photo voltaic. Perovskite crystals have minimalised this breakthrough down to a few years instead of decades compared to the research of silicon-based cells in the last 50 years [12]. It can be expected that the development of X-ray detectors using perovskite semiconductors will follow a similar fast track:

The perovskite crystals are not like conventional semiconductors [51]. The remarkable data concerning the transport properties of perovskite semiconductors (see Table 1) were obtained without any additional modifications or purification procedures. The defect densities in all of these crystals are above 10^{13} cm^{-3} , not counting dislocation densities [52]. Theoretical studies like in [53] and experimental data provided in [54] describe a “self-regulation” phenomenon. Siekmann et al. [55], as well as Brandt et al. [56], use the term “defect tolerance” to explain the excellent transport properties. The phenomenon of this “defect tolerance” is described as the higher energetically probability of the formation of structural defects into shallow levels instead of deep levels [57, 58]. This notion is mainly used to explain the excellent performance of halide perovskites in optoelectronic applications. One paper [59] identified the formation of antisite defects in CsPbI_3 . These deep levels were not effective nonradiative recombination centers. The negative influence on the transport properties were neglectable. A recent view by X. Wang et al. [60] summarizes the formation of intraband levels by extrinsic interstitials and vacancies and an electrical behavior of shallow levels.

In the perovskite semiconductors the influence of structural defects and impurities is small. It should be easier to improve the transport properties. The concentration of defects must not be reduced like in CdTe or GaAs semiconductors.

A similar situation is known by the performance of GaN-based LEDs. In GaN the dislocation density is typically above 10^7 1/cm^2 and the material can be used for application as a LED [61]. In GaAs the dislocation density has to be lower than 10^4 1/cm^2 for a working LED.

The measurement of dislocation density is a suitable analysis of defect concentration in a semiconductor. Dislocations are structural defects and can mostly be correlated with grain boundaries and impurities. The dislocation density is a measure of the number of dislocations in a unit volume of a crystalline material. It means the total length of all dislocation lines per unit of volume in a crystalline solid. The dislocations can be visualized by chemical etching. A count of the number of etch pits per unit area on the etched surface defines the dislocation density.

The actual research community is still focused on the development of materials. Several presentations of pixel detectors are quite convincing but the level of the device technology is still in development [62].

The experience in the research of CdTe-based detectors will be very useful to shorten the duration to develop the perovskite crystals. The process of crystal growth and material improvement have to be finalized before the development of technology. The concept of doping and controlling of shallow and deep levels can be applied to perovskite materials to face the challenges like the ion migration and stability of contacts. These issues are strongly correlated to the quality of the material. For example, the ion migration was successfully inhibited by co-doping of extrinsic interstitials [63] and by using chenodeoxycholic acid additive [23].

Future research of perovskite-based X-ray pixel detectors should be focused on the development of the material. Reduction of impurities and defects will increase the mobilities. Higher mobilities will yield increased $\mu\text{-}\tau$ -products, and thus we expect significantly improved energy resolution to be achievable with better quality crystals. This will solve most of the described challenges and will open new technologies for the production of large area pixel

detectors. It would be possible to use the advantages of high-Z semiconductors with reasonable costs.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

MF: Writing–review and editing, Writing–original draft. TB: Writing–review and editing, Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The

research was supported by the German Federal Ministry of Education and Research (BMBF) under the funding code (Förderkennzeichen) 05K22VFA.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Dectris Home. *Dectris* (2024). Available from: <http://www.dectris.com/> (Accessed February 28, 2024).
- CERN Accelerating Science. *Medipix* (2024). Available from: <https://medipix.web.cern.ch/home> (Accessed February 28, 2024).
- Pangaud P, Basolo S, Boudet N, Berar JF, Chantepie B, Delpierre P, et al. XPAD3: a new photon counting chip for X-ray CT-scanner. *Nucl Instrum Methods Phys Res A* (2007) 571(1–2):321–4. doi:10.1016/j.nima.2006.10.092
- Simon M, Ford RA, Franklin AR, Grabowski SP, Menser B, Much G, et al. PBO as direct conversion X-ray detector material. In: SPIE Proceedings (2004). San Diego, CA, USA.
- Schieber M. Fabrication of HgI₂ nuclear detectors. *Nucl Instr Methods* (1977) 144(3):469–77. doi:10.1016/0029-554x(77)90011-8
- Hitomi K, Tada T, Kim S, Wu Y, Tanaka T, Shoji T, et al. Recent development of tBr gamma-ray detectors. *IEEE Trans Nucl Sci* (2011) 58(4):1987–91. doi:10.1109/tns.2011.2123115
- Tyazhev AV, Budnitsky D, Koretskay O, Novikov V, Okaevich L, Potapov A, et al. GaAs radiation imaging detectors with an active layer thickness up to 1mm. *Nucl Instrum Methods Phys Res A* (2003) 509(1–3):34–9. doi:10.1016/s0168-9002(03)01545-6
- Fiederle M, Ebling D, Eiche C, Hofmann D, Salk M, Stadler W, et al. Comparison of CdTe, Cd_{0.9}Zn_{0.1}Te and CdTe_{0.9}Se_{0.1} crystals: application for γ- and X-ray detectors. *J Cryst Growth* (1994) 138(1–4):529–33. doi:10.1016/0022-0248(94)90863-x
- Lell M, Kachelrieß M. Computed tomography 2.0. *Invest Radiol* (2023) 58(8):587–601. doi:10.1097/rli.0000000000000995
- Fiederle M, Procz S, Hamann E, Fauler A, Fröjd C. Overview of GaAs und CdTe pixel detectors using Medipix Electronics. *Cryst Res Technol* (2020) 55(9). doi:10.1002/crat.202000021
- Gruner SM, Carini G, Miceli A. Considerations about future hard X-ray area detectors. *Front Phys* (2023) 11:1285821. doi:10.3389/fphy.2023.1285821
- National Renewable Energy Laboratory (NREL). *Best research-cell efficiency chart* (2024). Available from: <https://www.nrel.gov/pv/cell-efficiency.html>.
- Li Z, Zhou F, Yao H, Ci Z, Yang Z, Jin Z. Halide perovskites for high-performance X-ray detector. *Mater Today* (2021) 48:155–75. doi:10.1016/j.mattod.2021.01.028
- He Y, Wang Z, Wang Z, Dun GH, Geng X, Wang C, et al. Perovskite band engineering for high-performance X-ray detection. *Front Phys* (2023) 11:1114242. doi:10.3389/fphy.2023.1114242
- Lin C-F, Huang K-W, Chen Y-T, Hsueh S L, Li M-H, Chen P, et al. Perovskite-based X-ray detectors. *Nanomaterials* (2023) 13:2024. doi:10.3390/nano13132024
- Pan L, He Y, Klepov VV, De Siena MC, Kanatzidis MG. Perovskite CsPbBr₃ single crystal detector for high flux X-ray photon counting. *IEEE Trans Med Imaging* (2022) 41(11):3053–61. doi:10.1109/TMI.2022.3176801
- Basiricò L, Ciavatti A, Fraboni B. Solution-Grown organic and perovskite x-ray detectors: a new paradigm for the direct detection of ionizing radiation. *Adv Mater Tech* (2020) 6(1). doi:10.1002/admt.202000475
- Wang Y, Lou H, Yue CY, Lei XW. Applications of halide perovskites in X-ray detection and imaging. *CrystEngComm* (2022) 24:2201–12. doi:10.1039/d1ce01575c
- Kasap SO. X-ray sensitivity of photoconductors: application to stabilized A-SE. *J Phys D: Appl Phys* (2000) 33(21):2853–65. doi:10.1088/0022-3727/33/21/326
- Huang Y, Qiao L, Jiang Y, He T, Long R, Yang F, et al. A-site cation engineering for highly efficient MAPbI₃ single-crystal x-ray detector. *Angew Chem Int Edition* (2019) 58(49):17834–42. doi:10.1002/anie.201911281
- Geng X, Feng Q, Zhao R, Hirtz T, Dun G, Yan Z, et al. High-quality single crystal perovskite for highly sensitive X-ray detector. *IEEE Electron Device Lett* (2020) 41(2):256–9. doi:10.1109/led.2019.2960384
- Fan Z, Liu J, Zuo W, Liu G, He X, Luo K, et al. Mixed-cation MA_xCs_{1-x}PbBr₃ perovskite single crystals with composition management for high-Sensitivity X-Ray Detection. *physica status solidi (RRL) – Rapid Res Lett* (2020) 14(10). doi:10.1002/pssr.202000226
- Jiang W, Ren J, Li H, Liu D, Yang L, Xiong Y, et al. Improving the performance and high-field stability of FAPbBr₃ single crystals in x-ray detection with Chenodeoxycholic acid additive. *Small Methods* (2023) 7(4):e2201636. doi:10.1002/smt.202201636
- He Y, Petryk M, Liu Z, Chica DG, Hadar I, Leak C, et al. CsPbBr₃ perovskite detectors with 1.4% energy resolution for high-energy γ-Rays. *Nat Photon* (2020) 15(1):36–42. doi:10.1038/s41566-020-00727-1
- Zhang H, Wang F, Lu Y, Sun Q, Xu Y, Zhang BB, et al. High-sensitivity X-ray detectors based on solution-grown caesium lead bromide single crystals. *J Mater Chem C* (2020) 8(4):1248–56. doi:10.1039/c9tc05490a
- Pan W, Yang B, Niu G, Xue K, Du X, Yin L, et al. Hot-pressed CsPbBr₃ quasi-monocrystalline film for sensitive direct X-ray detection. *Adv Mater* (2019) 31(44):e1904405. doi:10.1002/adma.201904405
- Zhang B-B, Liu X, Xiao B, Hafsia AB, Gao K, Xu Y, et al. High-performance X-ray detection based on one-dimensional inorganic halide perovskite CsPbI₃. *The J Phys Chem Lett* (2019) 11(2):432–7. doi:10.1021/acs.jpclett.9b03523
- Steele JA, Pan W, Martin C, Keshavarz M, Debroye E, Yuan H, et al. Photophysical pathways in highly sensitive Cs₂AgBiBr₆ double-perovskite single-crystal x-ray detectors. *Adv Mater* (2018) 30(46):e1804450. doi:10.1002/adma.201804450
- Pan W, Wu H, Luo J, Deng Z, Ge C, Chen C, et al. Cs₂AgBiBr₆ single-crystal X-ray detectors with a low detection limit. *Nat Photon* (2017) 11(11):726–32. doi:10.1038/s41566-017-0012-4
- Kim YC, Kim KH, Son DY, Jeong DN, Seo JY, Choi YS, et al. Printable organometallic perovskite enables large-area, low-dose X-ray imaging. *Nature* (2017) 550(7674):87–91. doi:10.1038/nature24032

31. Qian W, Xu X, Wang J, Xu Y, Chen J, Ge Y, et al. An aerosol-liquid-solid process for the general synthesis of halide perovskite thick films for direct-conversion X-ray detectors. *Matter* (2021) 4(3):942–54. doi:10.1016/j.matt.2021.01.020
32. Kanak A, Kopach O, Kanak L, Levchuk I, Isaiev M, Brabec CJ, et al. Melting and crystallization features of CsPbBr₃ perovskite. *Cryst Growth Des* (2022) 22(7):4115–21. doi:10.1021/acs.cgd.1c01530
33. NIST Standard Reference Database (XGAM). *XCOM: photon cross sections database* (2024). Available from: <https://www.nist.gov/pml/xcom-photon-cross-sections-database>.
34. He Y, Hadar I, De Siena MC, Klepov VV, Pan L, Chung DY, et al. Sensitivity and detection limit of spectroscopic-grade perovskite CsPbBr₃ crystal for hard x-ray detection. *Adv Funct Mater* (2022) 32(24). doi:10.1002/adfm.202112925
35. Wu H, Ge Y, Niu G, Tang J. Metal halide perovskites for X-ray detection and imaging. *Matter* (2021) 4(1):144–63. doi:10.1016/j.matt.2020.11.015
36. Wang X, Zhao D, Qiu Y, Huang Y, Wu Y, Li G, et al. PIN diodes array made of perovskite single crystal for X-Ray Imaging. *physica status solidi (RRL) – Rapid Res Lett* (2018) 12(10). doi:10.1002/pssr.201800380
37. Zhang Z, Saparov B. Charge carrier mobility of metal halide perovskites: from Fundamentals to Ionizing Radiation Detection. In: *Metal-halide perovskite semiconductors* (2023). p. 89–109. doi:10.1007/978-3-031-26892-2_6
38. Kröger FA, de Nobel D. XXIV. preparation and electrical properties of CdTe single crystals. *J Electronics Control* (1955) 1(2):190–202. doi:10.1080/00207215508961407
39. Fiederle M, Eiche C, Salk M, Schwarz R, Benz KW, Stadler W, et al. Modified compensation model of CdTe. *J Appl Phys* (1998) 84(12):6689–92. doi:10.1063/1.368874
40. Graafsma H, Becker J, Gruner SM. Integrating hybrid area detectors for storage ring and free-electron laser applications. In: *Synchrotron light sources and free-electron lasers* (2020). p. 1225–55. doi:10.1007/978-3-030-23201-6_37
41. Bhattarai M, Bache S, Abadi E, Samei E. Exploration of the pulse pileup effects in a clinical CdTe-based photon-counting computed tomography. *Med Phys* (2023) 50(11):6693–703. doi:10.1002/mp.16671
42. Veale MC, Angelsen C, Booker P, Coughlan J, French MJ, Hardie A, et al. Cadmium zinc telluride pixel detectors for high-intensity X-ray imaging at Free Electron Lasers. *J Phys D: Appl Phys* (2018) 52(8):085106. doi:10.1088/1361-6463/aaf556
43. Greiffenberg. High-Z sensors for synchrotron sources and FELs; ‘room temperature semiconductor detectors conference 2023’. In: 2023 IEEE Nuclear Science Symposium, medical imaging conference and international symposium on room-temperature semiconductor detectors (NSS MIC RTSD); November 4–11, 2023; Vancouver, Canada. IEEE (2023).
44. Yan T, Yang C, Cui X. An edge-on energy-resolved X-ray semiconductor detector. *Solid State Commun* (2021) 332:114339. doi:10.1016/j.ssc.2021.114339
45. Bao C, Gao F. Physics of defects in metal halide perovskites. *Rep Prog Phys* (2022) 85(9):096501. doi:10.1088/1361-6633/ac7c7a
46. Ren X, Zhang B, Zhang L, Wen J, Che B, Bai D, et al. Deep-level transient spectroscopy for effective Passivator selection in perovskite solar cells to attain high efficiency over 23%. *ChemSusChem* (2021) 14(15):3182–9. doi:10.1002/cssc.202100980
47. Luo D, Li X, Dumont A, Yu H, Lu Z. Recent progress on perovskite surfaces and interfaces in optoelectronic devices. *Adv Mater* (2021) 33(30):e2006004. doi:10.1002/adma.202006004
48. Armaroli G, Maserati L, Ciavatti A, Vecchi P, Piccioni A, Foschi M, et al. Photoinduced current transient spectroscopy on metal halide perovskites: electron trapping and Ion Drift. *ACS Energ Lett* (2023) 8(10):4371–9. doi:10.1021/acsenrgylett.3c01429
49. Swift MW, Lyons JL. Deep levels in cesium lead bromide from native defects and hydrogen. *J Mater Chem A* (2021) 9(12):7491–5. doi:10.1039/d0ta11742k
50. Levine I, Vera OG, Kulbak M, Ceratti DR, Rehmann C, Márquez JA, et al. Deep defect states in wide-band-gap ABX₃ halide perovskites. *ACS Energ Lett* (2019) 4(5):1150–7. doi:10.1021/acsenrgylett.9b00709
51. Yin W-J, Shi T, Yan Y. Unusual defect physics in CH₃NH₃PBI₃ perovskite solar cell absorber. *Appl Phys Lett* (2014) 104(6). doi:10.1063/1.4864778
52. Le Corre VM, Duijnste EA, El Tambouli O, Ball JM, Snaith HJ, Lim J, et al. Revealing charge carrier mobility and defect densities in metal halide perovskites via space-charge-limited current measurements. *ACS Energ Lett* (2021) 6(3):1087–94. doi:10.1021/acsenrgylett.0c02599
53. Walsh A, Scanlon DO, Chen S, Gong XG, Wei SH. Self-regulation mechanism for charged point defects in hybrid halide perovskites. *Angew Chem* (2014) 127(6):1811–4. doi:10.1002/ange.201409740
54. Yao Z, Xu Z, Zhao W, Zhang J, Bian H, Fang Y, et al. Enhanced efficiency of inorganic CsPbI_{3-x}Br_x perovskite solar cell via self-regulation of antisite defects. *Adv Energ Mater* (2021) 11(23). doi:10.1002/aenm.202100403
55. Siekmann J, Ravishanker S, Kirchartz T. Apparent defect densities in halide perovskite thin films and single crystals. *ACS Energ Lett* (2021) 6(9):3244–51. doi:10.1021/acsenrgylett.1c01449
56. Brandt RE, Poindexter JR, Gorai P, Kurchin RC, Hoye RLZ, Nienhaus L, et al. Searching for “defect-tolerant” photovoltaic materials: combined theoretical and experimental screening. *Chem Mater* (2017) 29(11):4667–74. doi:10.1021/acs.chemmater.6b05496
57. Kim G, Petrozza A. Defect tolerance and intolerance in metal-halide perovskites. *Adv Energ Mater* (2020) 10(37). doi:10.1002/aenm.202001959
58. Kim J, Chung C-H, Hong K-H. Understanding of the formation of shallow level defects from the intrinsic defects of lead tri-halide perovskites. *Phys Chem Chem Phys* (2016) 18(39):27143–7. doi:10.1039/c6cp02886a
59. Zhang X, Turiansky ME, Van de Walle CG. Correctly assessing defect tolerance in halide perovskites. *The J Phys Chem C* (2020) 124(11):6022–7. doi:10.1021/acs.jpcc.0c01324
60. Wang X, Zhang M, Hou T, Sun X, Hao X. Extrinsic interstitial ions in metal halide perovskites: a Review. *Small* (2023) 19(46):e2303060. doi:10.1002/sml.202303060
61. Li C, Xu L. The dislocation density of gan led epitaxial films grown on si substrates characterized by x - ray diffraction. In: 2020 7th International Conference on Information Science and Control Engineering (ICISCE) (2020).
62. He Z. Pixelated CsPbBr₃ semiconductor gamma spectrometers, ‘room temperature semiconductor detectors conference 2023’. In: 2023 IEEE Nuclear Science Symposium, medical imaging conference and international symposium on room-temperature semiconductor detectors (NSS MIC RTSD), Vancouver, BC, November 4–11, 2023 (IEEE) (2023).
63. Zhao Y, Yavuz I, Wang M, Weber MH, Xu M, Lee JH, et al. Suppressing ion migration in metal halide perovskite via interstitial doping with a trace amount of multivalent cations. *Nat Mater* (2022) 21(12):1396–402. doi:10.1038/s41563-022-01390-3
64. Suzuki K, Shiro A, Toyokawa H, Saji C, Shobu T. Double-exposure method with synchrotron white X-ray for stress evaluation of coarse-grain materials. *Quan Beam Sci* (2020) 4(3):25. doi:10.3390/qubs4030025
65. Zappettini A. Cadmium telluride and cadmium zinc telluride. In: *Single crystals of electronic materials* (2019). p. 273–301. doi:10.1016/b978-0-08-102096-8.00008-2
66. Pan L, Feng Y, Kandlakunta P, Huang J, Cao LR. Performance of perovskite CSPBBR₃ single crystal detector for gamma-ray detection. *IEEE Trans Nucl Sci* (2020) 67(2):443–9. doi:10.1109/tns.2020.2964306

Frontiers in Physics

Investigates complex questions in physics to understand the nature of the physical world

Addresses the biggest questions in physics, from macro to micro, and from theoretical to experimental and applied physics.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

