



OPEN ACCESS

EDITED BY

Humberto Rocha,
University of Coimbra, Portugal

REVIEWED BY

Sebastiaan Breedveld,
Erasmus Medical Center, Netherlands
James C. L. Chow,
University of Toronto, Canada

*CORRESPONDENCE

Florian Barkmann

✉ florian.barkmann@inf.ethz.ch

Yair Censor

✉ yair@math.haifa.ac.il

Niklas Wahl

✉ n.wahl@dkfz.de

RECEIVED 12 June 2023

ACCEPTED 18 September 2023

PUBLISHED 26 October 2023

CITATION

Barkmann F, Censor Y and Wahl N (2023)
Superiorization of projection algorithms for
linearly constrained inverse radiotherapy
treatment planning.
Front. Oncol. 13:1238824.
doi: 10.3389/fonc.2023.1238824

COPYRIGHT

© 2023 Barkmann, Censor and Wahl. 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.

Superiorization of projection algorithms for linearly constrained inverse radiotherapy treatment planning

Florian Barkmann^{1,2,3*}, Yair Censor^{4*} and Niklas Wahl^{2,3*}

¹Institute for Machine Learning, Department of Computer Science, ETH Zürich, Zurich, Switzerland,

²Department of Medical Physics in Radiation Oncology, German Cancer Research Center (DKFZ), Heidelberg, Baden-Württemberg, Germany, ³Heidelberg Institute for Radiation Oncology (HIRO) and National Center for Radiation Research in Oncology (NCRO), Heidelberg, Baden-Württemberg, Germany,

⁴Department of Mathematics, Faculty of Natural Sciences, University of Haifa, Haifa, Israel

Objective: We apply the superiorization methodology to the constrained intensity-modulated radiation therapy (IMRT) treatment planning problem. Superiorization combines a feasibility-seeking projection algorithm with objective function reduction: The underlying projection algorithm is perturbed with gradient descent steps to steer the algorithm towards a solution with a lower objective function value compared to one obtained solely through feasibility-seeking.

Approach: Within the open-source inverse planning toolkit matRad, we implement a prototypical algorithmic framework for superiorization using the well-established Agmon, Motzkin, and Schoenberg (AMS) feasibility-seeking projection algorithm and common nonlinear dose optimization objective functions. Based on this prototype, we apply superiorization to intensity-modulated radiation therapy treatment planning and compare it with (i) bare feasibility-seeking (i.e., without any objective function) and (ii) nonlinear constrained optimization using first-order derivatives. For these comparisons, we use the TG119 water phantom, the head-and-neck and the prostate patient of the CORT dataset.

Main results: Bare feasibility-seeking with AMS confirms previous studies, showing it can find solutions that are nearly equivalent to those found by the established piece-wise least-squares optimization approach. The superiorization prototype solved the linearly constrained planning problem with similar dosimetric performance to that of a general-purpose nonlinear constrained optimizer while showing smooth convergence in both constraint proximity and objective function reduction.

Significance: Superiorization is a useful alternative to constrained optimization in radiotherapy inverse treatment planning. Future extensions with other approaches to feasibility-seeking, e.g., with dose-volume constraints and more sophisticated perturbations, may unlock its full potential for high performant inverse treatment planning.

KEYWORDS

radiation therapy treatment planning, inverse planning, constrained treatment plan optimization, IMRT, superiorization method, feasibility-seeking algorithm

1 Introduction

Numerical optimization methods lie at the heart of state-of-the-art inverse treatment planning for intensity-modulated radiation therapy (IMRT) (1). Usually, a clinical prescription of the treatment goals forms the input to a nonlinear multi-criteria optimization (MCO) problem with or without additional constraints, depending on the desired patient dose distribution.

During the translation of the clinical goals into an MCO problem, one distinguishes between objectives, i.e., soft goals that compete with each other, and hard constraints designed to ensure, for example, maximal tolerance doses in an organ-at-risk (OAR) and minimal dosage of the target. This versatile approach enables the treatment planner to employ arbitrary combinations of suitable (convex) nonlinear objective functions along with any choice of constraints on the voxels' doses.

This mathematical modeling allows numerical optimization of the fluence of beam elements (beamlets) using a pre-computed normalized dose mapping (2). The resulting constrained nonlinear optimization problem is frequently solved by applying an extended (quasi-)Newton approach with sequential quadratic programming (SQP) and/or interior-point methods (1–7). Until now, the capabilities of inverse planning have been substantially extended through multi-criteria Pareto optimization with subsequent exploration of the Pareto surface (8, 9) or stochastic/robust optimization (10).

Computational difficulties may arise in the constrained nonlinear optimization approach. First, optimal convergence for problems of typical size in radiotherapy is tied to the availability of computationally efficient second-order derivatives. While, for example, van Haveren and Breedveld (11) showed that for many typical functions efficient formulations can be found, current research persistently adds new quantities, optimization strategies, and new types of problem formulations to inverse planning for photons and particles (see, e.g., 12–18) to which such strategies might not be directly applicable. Second, a common approach among successful optimizers for nonlinear constrained optimization is to transform the constrained problem into an unconstrained problem using, for example, barrier functions (in the case of interior point methods, e.g., 3, 19) and the method of Lagrange multipliers in combination with slack variables (3, 19, 20). This creates a computational burden when the number of constraints increases. Handling many constraints as, for example, linear inequalities for many or all individual voxel dose bounds, can inflate the computational effort because each constraint requires a Lagrange multiplier and an additional slack variable. Possible “workarounds” include minimax-optimization in combination with auxiliary variables or usage of continuous and differentiable maximum approximations like the LogSumExp and softmax functions (5).

Taking a step back, however, to the starting days of treatment planning research, shows that one does not necessarily need to use a mathematical *optimization* approach to solve the purely linearly constrained IMRT problem but could use feasibility-seeking projection algorithms instead (21, 22).

In the context of IMRT, such bare feasibility-seeking translates to seeking a feasible solution that will obey the prescribed lower and upper dose bounds on doses in voxels, without aiming to optimize

any objective function. While, in general, a bare feasibility-seeking task can be translated to a constrained optimization problem with a zero objective function, the literature demonstrates a wide spectrum of many efficient feasibility-seeking algorithms not derived from translation of the bare feasibility-seeking task to a constrained optimization problem (see, e.g., 23). If no feasible solution is found, these algorithms find a proximal solution, similar to the piece-wise least-squares approach. Even though they have seen further development over the last decades (24) and, more recently, also extension to dose-volume constraints (25–27), numerical optimizers have been the preferred choice in the field due to their abilities to handle the nonlinear objective functions, e.g., (generalized) equivalent uniform dose (EUD), which are often desired when prescribing treatment goals.

The work presented here now combines nonlinear objective functions as used in optimization with feasibility-seeking within linearly constraining dose bounds by applying the superiorization method (SM). To do so, the SM uses a *superiorized version of the basic algorithm*, the latter being a user-chosen iterative feasibility-seeking algorithm, which is perturbed by interlacing reduction steps of the chosen (nonlinear) objective function. This practically steers the iterates of the feasibility-seeking algorithm to a feasible solution point with a “superior”, i.e., smaller or equal objective function value, which is not necessarily a constrained minimization point.

The superiorization method thus works with the constraints data and the user's choice of objective function, much alike constrained optimization methods would. But it does not aim at an optimal point that minimizes the objective over all constraints like the latter do. In contrast, the SM aims at a point that will fulfill all constraints and have a reduced – not necessarily minimal – objective function value. Not finding the optimal solution, but instead aiming for a satisfactory or adequate result, is a reasonable decision strategy (“Satisficing”, see 28), particularly considering the degeneracy of the IMRT optimization problem (29). Hence, this aim suffices for the purpose of generating acceptable treatment plans. Combined with the simplicity of the gradient descent steps (i.e., not relying on second-order derivatives), superiorization can find a solution, in general, faster and with less investment of computing resources, and fewer conditions concerning design of the objective function.

Application of the SM to treatment planning is encouraged by the flexibility it has shown for applications in multiple fields:¹ It has demonstrated its effectiveness for image reconstruction in single-energy computed tomography (CT) (31, 32), dual-energy CT (33) and, more recently, in proton CT (34, 35), by reducing total variation (TV) during image reconstruction. The SM has also been successfully applied to diverse other fields of applications, such as tomographic imaging spectrometry (36) or signal recovery (37).

This work is – to the best of our knowledge – the first in-depth investigation of the SM as a potential alternative to constrained

1 Censor (<http://arxiv.org/abs/1506.04219>) provides a continuously updated bibliography of works using superiorization to present the state of research (30).

minimization algorithms for inverse radiotherapy treatment planning using common objective functions. To date, we could only identify an initial study of the applicability of SM in IMRT utilizing TV as objective function (38), which does not represent common choices in objective function design for treatment planning. Another work considering the use of SM in IMRT used superiorization to boost a specific lexicographic planning approach (39).

Expanding on those preliminary works, we develop, tune, and evaluate a prototypical superiorization solver for radiotherapy treatment planning problems. To show how this SM solver is able to replace a constrained minimization approach, and to maximize reproducibility and re-usability of our work, our superiorization approach is implemented into the validated open source radiation therapy treatment planning toolkit matRad (5) together with an instructive set of scripts to execute and reproduce the results of this work (see section 2.5). Within matRad and its included phantoms and patient cases, the SM is evaluated and tested on full-fledged IMRT and intensity-modulated proton therapy (IMPT) treatment planning problems. We compare to using non-linear constrained optimization using only first-order derivatives like the SM, that is, a quasi-Newton method construction a Hessian approximation.

This paper is structured as follows: In section 2, we describe the approaches and present the specific version of the SM that we use along with the feasibility-seeking algorithm embedded in it. Section 3 includes our computational results. Finally, in section 4, we discuss the potential of SM with possible future developments and conclude our work in section 5.

2 Materials and methods

This work compares three approaches to model the treatment planning problem in IMRT: (i) a *nonlinear constrained minimization approach* of minimizing an objective function subject to constraints with a quasi-Newton method relying on first-order derivatives, (ii) the *feasibility-seeking approach* searching for a feasible solution adhering to constraints without considering any objective functions to minimize, and finally, (iii) the *superiorization approach*, which perturbs the feasibility-seeking algorithm to reduce (not necessarily minimize) an objective function by gradient descent steps. Before introducing these approaches, we briefly recap the discretization of the inverse treatment planning problem.

2.1 Discretization of the inverse treatment planning problem

Computerized inverse treatment planning usually relies on a spatial discretization of the particle fluence, the patient anatomy, and, consequently, the radiation dose.

The patient is represented by a three-dimensional voxelized grid (image) with n voxels numbered $i = 1, 2, \dots, n$. Based on this image, Q volumes of interest (VOIs) S_q , $q = 1, 2, \dots, Q$ are segmented. This allows us to represent the dose as a vector $\mathbf{d} = (d_i)_{i=1}^n$, whose i -th

component is the radiation dose deposited within the i -th voxel. For each of the segmentations S_q , we can then easily identify its dosage by finding d_i for all $i \in S_q$.

The radiation fluence is represented as a vector intensities $\mathbf{x} = (x_j)_{j=1}^m$, whose j -th component is the intensity of the j -th beamlet. The dose deposition a_i^j for a unit intensity of beamlet j to voxel i can then be precomputed and stored in the *dose influence matrix* $A = (a_i^j)_{i=1, j=1}^{n, m}$, mapping \mathbf{x} to \mathbf{d} via $\mathbf{d} = \mathbf{A}\mathbf{x}$.

2.2 The constrained minimization approach

In the optimization approach to IMRT treatment planning, the clinically prescribed aims are represented by various (commonly differentiable) objective functions which map the vector of beamlet intensities to the positive real numbers (2).

For our purposes, we limit ourselves to objective functions $f_p : \mathbb{R}^n \rightarrow [0, \infty)$, $p = 1, 2, \dots, P$, operating on the radiation dose \mathbf{d} as surrogates for clinical, dose-based goals.

A comprehensive, exemplary list of such common objective functions can be found in Wieser et al. (5, Table 1) and, for the reader's convenience, also in [Supplementary Data Sheet](#) below. These objective functions, which depend on the dose, are related to the intensities \mathbf{x} via $\mathbf{d} = \mathbf{A}\mathbf{x}$, which is computed at each iterate/change of \mathbf{x} during optimization.

Wishing to fulfill or decide between multiple clinical goals, the resulting multi-objective optimization problem may be scalarized using a weighted sum of several different individual objective functions for the various VOIs S_q . This approach, first introduced for least-squares (as introduced by 40), can today explore a plethora of objective functions (2, 5) while also satisfying hard constraints (3, 5):

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{p=1}^P w_p f_p(\mathbf{d}(\mathbf{x}))$$

such that $c_t^L \leq c_t(\mathbf{d}(\mathbf{x})) \leq c_t^U$, $t = 1, 2, \dots, T$, (1)

$$\mathbf{x} \geq 0$$

Here $w_p \geq 0$, for all $p = 1, 2, \dots, P$, are user-specified weights reflecting relative importance, f_p are user-chosen individual objective functions, \mathbf{x} is the beamlet radiation intensities vector (which is physically bound to the nonnegative real orthant), and c_t are user-chosen individual constraints with lower and upper bounds c_t^L and c_t^U , respectively. While the constraints c_t can, in principle, be nonlinear constraints, we focus here on *linear inequality constraints* representing upper and lower dose prescription bounds.

The inverse planning problem from eq. (1), solved with numerical optimization techniques, is commonly used today

TABLE 1 Dose inequalities/prescriptions and penalty weights used for minimization and for AMS feasibility-seeking.

VOI	w_p	tolerance/inequality constraint
Target	1000	59 Gy < \mathbf{d} < 61 Gy
Core	100	\mathbf{d} < 20 Gy
Body	30	\mathbf{d} < 30 Gy

across treatment modalities (among others 2, 3, 5, 40, 41). SQP or interior point methods with a (quasi-)Newton approach are often used to solve the resulting constrained optimization problems (1–7, 42). In this work, we focus on a quasi-Newton approach using first-order derivatives only, since the superiorization approach (as described further below in section 2.4) has so far only been investigated using gradient descent steps itself.

2.3 The feasibility-seeking approach

Since the bare feasibility-seeking approach is the backbone of the SM, it will be outlined below using the notation from sections 2.1 and 2.2. Prior work has already suggested the feasibility-seeking approach to address the treatment planning problem (see, e.g., 43, and references therein).

To solve the treatment planning problem with feasibility-seeking, dose prescriptions are modeled as a system of linear inequalities: In general, the dose in every voxel is constrained with a lower and upper bound. Feasibility-seeking now seeks a solution, i.e., a beamlet intensity vector fulfilling these prescriptions.

With $d(x) = Ax$, the beamlet radiation intensities vector x now has to be recovered from a system of linear inequalities of the form

$$c_i^L \leq \sum_{j=1}^m a_j^i x_j \leq c_i^U, \quad i = 1, 2, \dots, n. \quad (2)$$

In principle, individual lower and upper bounds c_i^L and c_i^U can be chosen for each voxel i . Since prescriptions are usually grouped per VOI S_q , the system can be rewritten as:

$$\text{For all } q = 1, 2, \dots, Q: \quad \ell_q \leq \sum_{j=1}^m a_j^i x_j \leq u_q \quad \text{for all } i \in S_q, \quad (3)$$

with ℓ_q and u_q representing the lower and upper dose bounds per VOI S_q , respectively. Since it does not make sense to prescribe positive lower bounds to OARs, these are generally chosen to be equal to zero.

Geometrically, depending on which structure S_q a voxel i belongs to, each physical dose constraint set C_i in each voxel $i = 1, 2, \dots, n$, is a *hyper slab* (i.e., an intersection of two half-spaces) in the m -dimensional Euclidean vector space \mathbb{R}^m .

Aiming at satisfaction of all physical dose constraints along with the nonnegativity constraints is, thus, the following (which is a special case of the *convex feasibility problem* see, e.g., 23):

$$\text{Find an } x^* \in W := \left\{ x \in \mathbb{R}^m \mid \text{for all } q = 1, 2, \dots, Q, \ell_q \leq \sum_{j=1}^m a_j^i x_j \leq u_q, \text{ for all } i \in S_q, \text{ and } x \geq 0 \right\} \quad (4)$$

Such feasibility-seeking problems can typically be solved by a variety of efficient projection methods, whose main advantage, which makes them successful in real-world applications, is computational (see, e.g., 23, 44).

They commonly can handle very large-size problems of dimensions beyond which other, more sophisticated currently

available, methods start to stutter or cease to be efficient. This is because the building blocks of a projection algorithm are the projections onto the given individual sets. These projections are actually easy to perform, particularly in linear cases such as hyperplanes, half-spaces, or hyperslabs.

For the purpose of this paper, we define such an iterative feasibility-seeking algorithm via an algorithmic operator $\mathcal{A} : \mathbb{R}^m \rightarrow \mathbb{R}^m$,

$$x^0 \in \mathbb{R}^m, \quad x^{k+1} = \mathcal{A}(x^k), \quad k = 1, 2, \dots, \quad (5)$$

whose task is to (asymptotically) find a point in W .

The algorithmic structures of projection algorithms are sequential, simultaneous, or in-between, such as in the block-iterative projection (BIP) methods (see, e.g., 45, 46, and references therein) or in the more recent string-averaging projection (SAP) methods (see, e.g., 47, and references therein). An advantage of projection methods is that they work with the initial, raw data and do not require transformation of, or other operations on, the sets describing the problem.

For our prototype used here in conjunction with the SM, we rely on the well-established Agmon, Motzkin, and Schoenberg (AMS) relaxation method for linear inequalities (48, 49). Implemented sequentially and modified for handling the bounds $x \geq 0$, it is outlined in Algorithm 1. We denote $\ell := (\ell_q)_{q=1}^Q$ and $u := (u_q)_{q=1}^Q$.

```

1: function  $\mathcal{A}^{AMS}(x, A, u, \ell, \lambda, v)$ 
2:    $I = CS(n)$  *Select control sequence*
3:   for all  $i \in I$  do
4:     if  $a^i, x > u_q$  then * $u_q$  for containing  $i$ -th voxel*
5:        $x \leftarrow x - \lambda v_i \frac{\langle a^i, x \rangle - u_q}{\|a^i\|_2^2} a^i$ 
6:     else if  $\langle a^i, x \rangle < \ell_q$  then * $\ell_q$  for containing  $i$ -th voxel*
7:        $x \leftarrow x - \lambda v_i \frac{\ell_q - \langle a^i, x \rangle}{\|a^i\|_2^2} a^i$ 
8:     else *Do nothing*
9:   end if
10: end for
11: for  $j = 1, 2, \dots, m$  do *Ensure nonnegativity of  $x$ *
12:    $x_j \leftarrow \begin{cases} x_j, & \text{for } x_j \geq 0 \\ 0, & \text{for } x_j < 0 \end{cases}$ 
13: end for
14: return  $x$ 
15: end function

```

Algorithm 1. The AMS Sequential Relaxation Method's algorithmic operator \mathcal{A}^{AMS} .

During an iteration, Algorithm 1 iterates over all rows of the dose matrix A and handles sequentially the right-hand side and the left-hand side of individual constraints from eq. (3). The *control sequence* (CS) (50, Definition 5.1.1) determines the order of iterating through the matrix rows/constraints. When a corresponding voxel dose inequality is violated, the algorithm performs geometrically a projection of the current point x onto the violated half-space with a user-chosen relaxation parameter $0 < \lambda \leq 2$. The original AMS algorithm is modified in Algorithm 1 to allow the relaxation for each

voxel i to be weighted with v_i and by performing projections onto the nonnegative orthant of \mathbb{R}^m (in steps 11–13) to return only nonnegative intensities \mathbf{x} . The vector $\mathbf{a}^i = (a_i^j)_{j=1}^m$ is the i -th row of the dose matrix \mathcal{A} and is the normal vector to the half-space represented by that row and $\|\mathbf{a}^i\|_2^2$ is its square Euclidean norm.

In summary, the algorithmic operator in Algorithm 1 describes a single complete sweep of projections sequentially over all constraints (half-spaces) followed by a projection onto the nonnegative orthant thus ensuring the nonnegativity constraint. Such sweeps will be executed iteratively.

The theory behind this algorithm guarantees that, under reasonable conditions, if the feasibility-seeking sweeps are performed endlessly then any sequence of iteration vectors $\{\mathbf{x}^k\}_{k=0}^\infty$ converges to a point that satisfies all constraints.

Choosing to define an algorithmic operator \mathcal{A} in Algorithm 1, allows us to concisely display the superiorization approach independent from the chosen projection algorithm below (see step 21 inside Algorithm 2).

2.4 The superiorization method and algorithm

The SM is built upon application of a feasibility-seeking approach (section 2.3) to the constraints in the second and third lines of eq. (1). But instead of handling the constrained minimization problem of eq. (1) with a full-fledged algorithm for constrained minimization, the SM interlaces into the feasibility-seeking iterative process (i.e., “the basic algorithm”) steps that reduce locally in each iteration the objective function value.

Accordingly, the SM does not aim at finding a constrained minimum of the combined objective function $f(\mathbf{x}) = \sum_{p=1}^P w_p f_p(\mathbf{x})$ of eq. (1) over the constraints. It rather strives to find a feasible point that satisfies the constraints and has a reduced – not necessarily minimal – value of f .

In the following, we give a brief and focused introduction to SM. A more detailed explanation and review can be found in, e.g., Censor et al. (51, Section II) and references therein (see also 31, 35, 45, 52–55).

In general, the SM is intended for *constrained function reduction problems* of the following form (55, Problem 1):

Problem 1. The constrained function reduction problem of the SM

Let W be a given set (such as in eq. (4)) and let $f: \mathbb{R}^m \rightarrow \mathbb{R}$ be an objective function (such as in eq. (1)). Let \mathcal{A} from eq. (5) be an algorithmic operator that defines an iterative basic algorithm for feasibility-seeking of a point in W . Find a vector $\mathbf{x}^* \in W$ whose function value is smaller or equal (but not necessarily minimal) than that of a point in W that would have been reached by applying the basic algorithm alone.

The SM approaches this question by investigating the *perturbation resilience* (52, Definitions 4 and 9) of \mathcal{A} , and then proactively using such perturbations, to locally reduce the values f of the iterates, in order to steer the iterative sequence generated by algorithm \mathcal{A} to a solution with smaller or equal objective function value. The structure

of the superiorization algorithm implemented here is given by Algorithm 2 with explanations here and in section 2.4.1.

Except for the initialization in steps 1–3, Algorithm 2 consists of the perturbations phase (steps 5–19) and the feasibility-seeking phase (steps 20–23).

In the perturbation phase, the objective function f is reduced using negative gradient descent steps. The step-size β of these gradient updates is calculated by α^s where α is a fixed user-chosen constant, called *kernel*, $0 < \alpha < 1$ so that the resulting step-sizes are nonnegative and form a summable series. The power s is incremented by one until the objective function value of the newly acquired point is smaller or equal to the objective function value of the point with which the current perturbations phase was started.

The parameter N determines how many perturbations are executed before applying the next full sweep of the feasibility-seeking phase. The basic Algorithm 1 with algorithmic operator \mathcal{A}^{AMS} , used throughout this work, is indeed perturbation resilient (56).

The superiorization approach has the advantage of letting the user choose any task-specific algorithmic operator \mathcal{A} that will be computationally efficient, independently of the perturbation phase, as long as perturbation resilience is preserved.

```

1:  $k \leftarrow 0$ 
2:  $\mathbf{x}^k \leftarrow \mathbf{x}^0$ 
3:  $s \leftarrow -1$ 
4: while stopping rule not met do
5:    $t \leftarrow 0$  *start of perturbation phase*
6:    $\mathbf{x}^{k,t} \leftarrow \mathbf{x}^k$ 
7:   while  $t < N$  do *apply N function reductions*
8:      $\text{loop} \leftarrow \text{true}$ 
9:     while  $\text{loop}$  do
10:       $s \leftarrow s + 1$ 
11:       $\beta \leftarrow \alpha^s$  *Step size adaptation*
12:       $\mathbf{z} \leftarrow \mathbf{x}^{k,t} - \beta \nabla f(\mathbf{x}^{k,t})$  *Function reduction step*
13:      if  $f(\mathbf{z}) \leq f(\mathbf{x}^{k,t})$  then *Function reduction check*
14:         $t \leftarrow t + 1$ 
15:         $\mathbf{x}^{k,t} \leftarrow \mathbf{z}$ 
16:         $\text{loop} \leftarrow \text{false}$ 
17:      end if
18:    end while
19:  end while
20:  $\mathbf{v}_k \leftarrow \eta^k \mathbf{v}$  *start of feasibility-seeking phase*
21:  $\mathbf{x}^{k+1} \leftarrow \mathcal{A}^{\text{AMS}}(\mathbf{x}^{k,t}, \mathcal{A}, \mathbf{u}, \ell, \lambda, \mathbf{v}_k)$ 
22:  $k \leftarrow k + 1$ 
23: end while
24: return  $\mathbf{x}^k$ 

```

Algorithm 2. Superiorization of the feasibility-seeking basic algorithm described by the operator $\mathcal{A} = \mathcal{A}^{\text{AMS}}$.

For our IMRT treatment planning problem using voxel dose constraints as introduced in eqs. (2) – (4), \mathcal{A} can be – besides the chosen AMS algorithm – any of the wide variety of feasibility-seeking algorithms (see, e.g., 23, 44, 50, 57).

The principles of the SM have been presented and studied in previous publications (consult, e.g. 31, 52, 54), but, to the best of our knowledge, this is the first work applying the SM to a treatment planning problem with an objective function of the general form $f(\mathbf{x}) := \sum_{p=1}^P w_p f_p(\mathbf{x})$ from eq. (1).

2.4.1 Modifications of the prototypical superiorization algorithm

To control the initial step-size, we warm start the algorithm with larger kernel powers s within the first iteration, which substantially improves the algorithm's runtime. For our purposes, we chose an initial increment of $s \leftarrow s + 25$.

In the feasibility-seeking phase, instead of weighting all projections onto the half-spaces equally via the relaxation parameters, each projection can also be given an individual weight $0 < v_i < 1$ representing the importance of the i -th inequality constraint (i. e., voxel).

Further, as shown in step 20 of Algorithm 2, weights can be reduced after each iteration to improve stability. Similar to how the step-sizes are reduced in the perturbation phase, we utilize another kernel $0 < \eta < 1$ and use its powers η^k to reduce the weights in step 20 by incrementing k after each feasibility-seeking sweep. The new weights are then calculated by $\eta^k \cdot v$, where v are the initial weights.

Finally, we integrate four different control sequences to iterate through the rows of A . Apart from following the cyclic order according to voxel indices, we experimented with a random order and with sequences choosing rows with increasing or decreasing weights v_i .

2.4.2 Stopping criteria

The algorithm was terminated after a given maximal number of iterations was reached or after a certain time limit was exceeded, or when the stopping criterion formulated below was met. The default number of maximum iterations was 500 and the default wall-clock duration was set to 50min.

The stopping criterion that we used consists of two parts, both of which must be met for three consecutive iterations for the algorithm to stop. The first part of the stopping criterion is that the relative change of the objective function f defined by

$$\left| \frac{f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k)}{\max\{1, f(\mathbf{x}^k)\}} \right| \quad (6)$$

becomes smaller than 10^{-4} .

For the second part of the stopping criterion, we define the square of the weighted L_2 -norm of the constraints violations by²

$$V(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^n \frac{(\ell_q - \langle \mathbf{a}^i, \mathbf{x} \rangle)_+^2 + (\langle \mathbf{a}^i, \mathbf{x} \rangle - u_q)_+^2}{\|\mathbf{a}^i\|_2^2} \quad (7)$$

where ℓ_q and u_q depend on which structure the i -th voxel belongs to. This second part of the stopping rule is met if the relative change of V defined by

$$\left| \frac{V(\mathbf{x}^{k+1}) - V(\mathbf{x}^k)}{\max\{1, V(\mathbf{x}^k)\}} \right| \quad (8)$$

is smaller than 10^{-3} .

All tolerances of the stopping criteria can be customized and also set to a negative number to turn off single stopping criteria or early stopping altogether.

2.5 Implementation

The superiorization prototype described above was implemented in the open-source crossplatform software "matRad" (5, 58, 59), which is a multi-modality radiation dose calculation and treatment planning toolkit written in Matlab. The implementation is publicly available on the matRad GitHub repository on a research branch.³

The superiorization solver is implemented as the class `matRad_OptimizerSuperiorization.m` within matRad's optimization framework. The class defines various user-configurable properties such as the maximum number of iterations, maximum wall time, different warm-start settings, two different feasibility-seeking algorithms, and various control sequences. Once the optimizer has been initialized, the `optimize` method can be called to generate a solution to the plan. The `optimize` method requires the following inputs: a starting point, the objective function with its gradient, the linear constraints, and the dose projection matrix. The perturbation phase, as well as the two provided feasibility-seeking algorithms, are implemented as additional methods. Furthermore, within the class, an additional method `PlotFunction` is available. This method facilitates the visualization of key metrics, such as the objective function value, the maximum constraint violation, and the proximity of the solution to the set of feasible solutions. Multiple scripts to reproduce the results presented herein are provided in an additional GitHub repository.⁴

The implementation in matRad facilitates comparison against plans generated on the same datasets with a nonlinear optimizer, as matRad implements a number of common objective functions used in treatment planning (compare to Supplementary Data Sheet and Wieser et al. (5, Table 1)). While matRad provides interfaces to both the open-source Interior Point OPTimizer (IPOPT) (19) as well as to Matlab's built-in interior-point algorithm from `fmincon`, only the first was used for our comparisons.

We chose to use matRad's optimization implementation as a benchmark for mainly two reasons: First, matRad has been used in numerous research works demonstrating its ability to create acceptable treatment plans. Second, as an open-source tool, matRad does allow direct modifications of the algorithms and respective parameters and stopping criteria, running them under truly similar conditions. This means that the evaluation of the objective function and its gradient itself use *exactly* the same code. Benchmarking against other closed-source treatment planning systems would be inconsequential due to hidden computational optimizations, simplifications, and unknown mathematical formulations of objectives and constraints.

2 For any real number r we use $(r)_+ := \max\{0, r\}$

3 <https://github.com/e0404/matRad/tree/research/superiorization>. Commit hash at the time of this work: ec04f5a5

4 <https://github.com/e0404/paper-superiorization-imrt/>. Commit hash at the time of this work: a787e94.

As motivated in section 2.2, no second-order derivatives were used in the nonlinear optimization approach, but instead a limited-memory Hessian approximation using first-order derivatives was chosen. While second-order derivatives can be used within matRad, it does not make use of fast exact Hessian computation strategies (11), reducing the value of a runtime comparison.

matRad performs all computations in a fully-discretized model with a voxel grid. The “dose matrix” A is stored as a compressed sparse column matrix computed for all analyses using matRad’s singular value decomposed pencil-beam algorithm (60) for photons and a singleGaussian pencil-beam algorithm for protons, both validated against clinical implementations (5).

3 Results

3.1 Proof-of-work: Phantom plan

To demonstrate the applicability of superiorization to the IMRT treatment planning problem, we first evaluate a small example using the horseshoe phantom defined in the AAPM TG119 Report (61). The phantom is part of the CORT dataset (62) and consequently available with matRad.

We created an equidistantly spaced 5-field IMRT photon plan with $5\text{mm} \times 5\text{mm}$ beamlet doses (resulting in 1918 pencil-beams and a corresponding sparse dose influence matrix with 9.3×10^7 non-zero entries in 3.5×10^6 voxels).

With this setup, we generated treatment plans using three different approaches: (i) constrained minimization with IPOPT, (ii) the AMS algorithm for feasibility-seeking only, and (iii) the SM

with the AMS algorithm. Different combinations of nonlinear objective functions and linear inequality constraints on dose were evaluated and compared across these approaches.

For analysis, we use dose-volume histograms (DVHs) and axial dose (difference) slices, as well as the evolution plots of the objective function values and the constraint violations.

3.1.1 General usability of the AMS feasibility-seeking projection algorithm

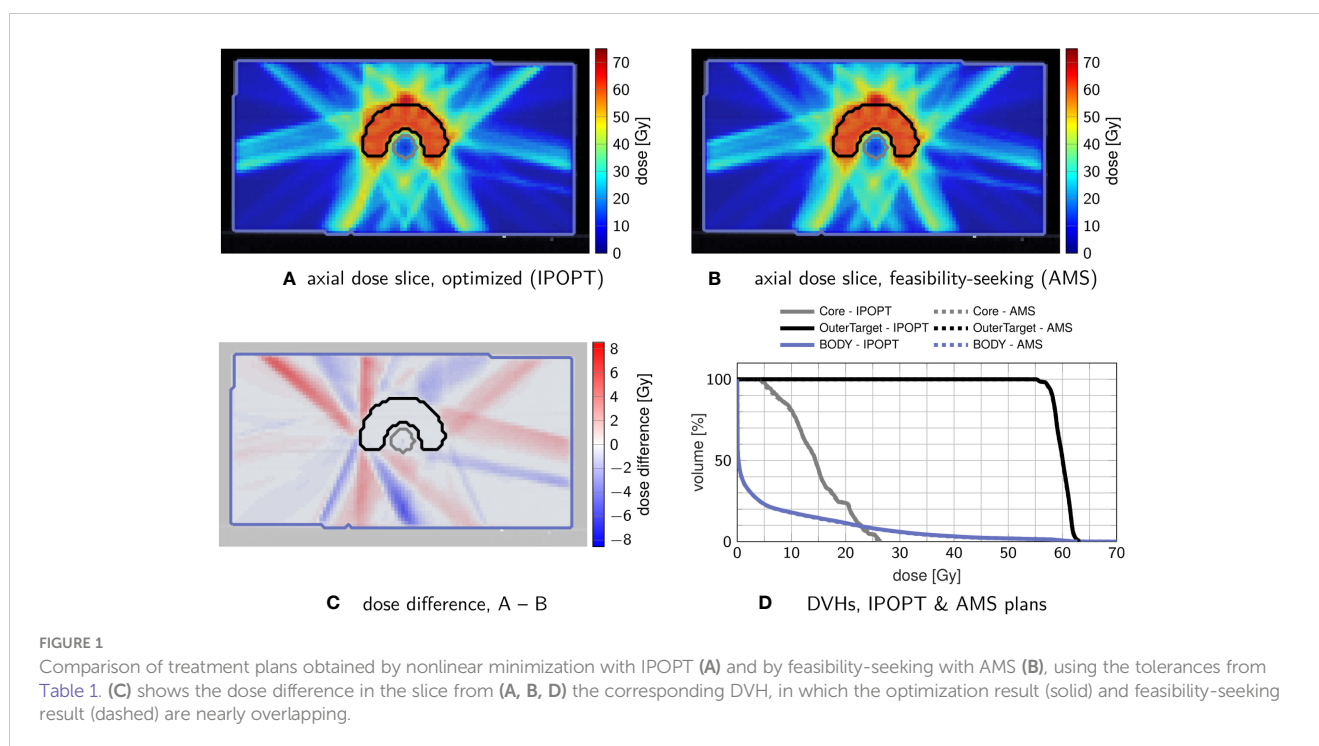
We first validate that our implemented projection algorithm AMS is capable of finding comparable treatment plans to those found by established optimization algorithms when applied to a straightforward piece-wise least-squares objective function for the unconstrained minimization of residuals.

The setup prescribes 60 Gy to the C-shaped target. To achieve this prescription, we bound the dose in the target by (60 ± 1) Gy. To the two OARs, “Core” and “Body”, upper bounds (a.k.a. tolerance doses) are prescribed, resulting in the parameters given in Table 1.

For nonlinear minimization with IPOPT, the tolerance doses serve as parameters for respective penalized piece-wise least-squares objective functions while for AMS the tolerances directly translate into linear inequalities and the weights proportionally increase the relaxation parameters.

Figure 1 confirms that feasibility-seeking with weighted AMS is able to find dose distributions of similar quality as conventional nonlinear unconstrained minimization of a piece-wise leastsquares objective function. While resulting in different intensity-modulation patterns, nearly congruent DVHs are observed.

A crude performance analysis though measures substantially longer runtimes for the AMS approach (about five times slower



than unconstrained minimization). This difference is mainly driven by the fact that AMS does sequential iteration by iterations through the matrix rows in each sweep.

This investigated scenario is, however, not intended to display any performance advantages of the AMS algorithm, but only to validate its behavior and confirm the long-known ability of such feasibility-seeking algorithms to yield acceptable treatment plans (21, 22).

3.1.2 Inverse planning with superiorization

Using the same phantom and irradiation geometry as in section 3.1.1, the feasibility problem used in 3.1.1 was modified to enforce some hard linear inequality constraints while minimizing an objective function. When the constraints are feasible, superiorization using AMS as the basic algorithm will find a feasible point while perturbing the iterates of the feasibility-seeking algorithm towards smaller or equal (not necessarily minimal) function values with objective function reduction steps.

TABLE 2 Dose inequality constraints, objective functions, and penalty weights used separately for constrained minimization and for superiorization.

VOI	w_p	$c(d)$	$f(d)$
Target	1000	$59 \text{ Gy} < d < 61 \text{ Gy}$ (I & II)	$f_{sqdev}(d; 60 \text{ Gy})$
Core	100	$d < 30 \text{ Gy}$ (II & III)	$f_{sqdev+}(d; 20 \text{ Gy})$
Body	30	–	$f_{sqdev+}(d; 30 \text{ Gy})$

The Roman numerals in parentheses for the inequality constraints describe their usage in the plans, respectively. The functions in the right-hand column stem from Wieser et al. (5, Table 1) and are identified here in Supplementary Data Sheet below.

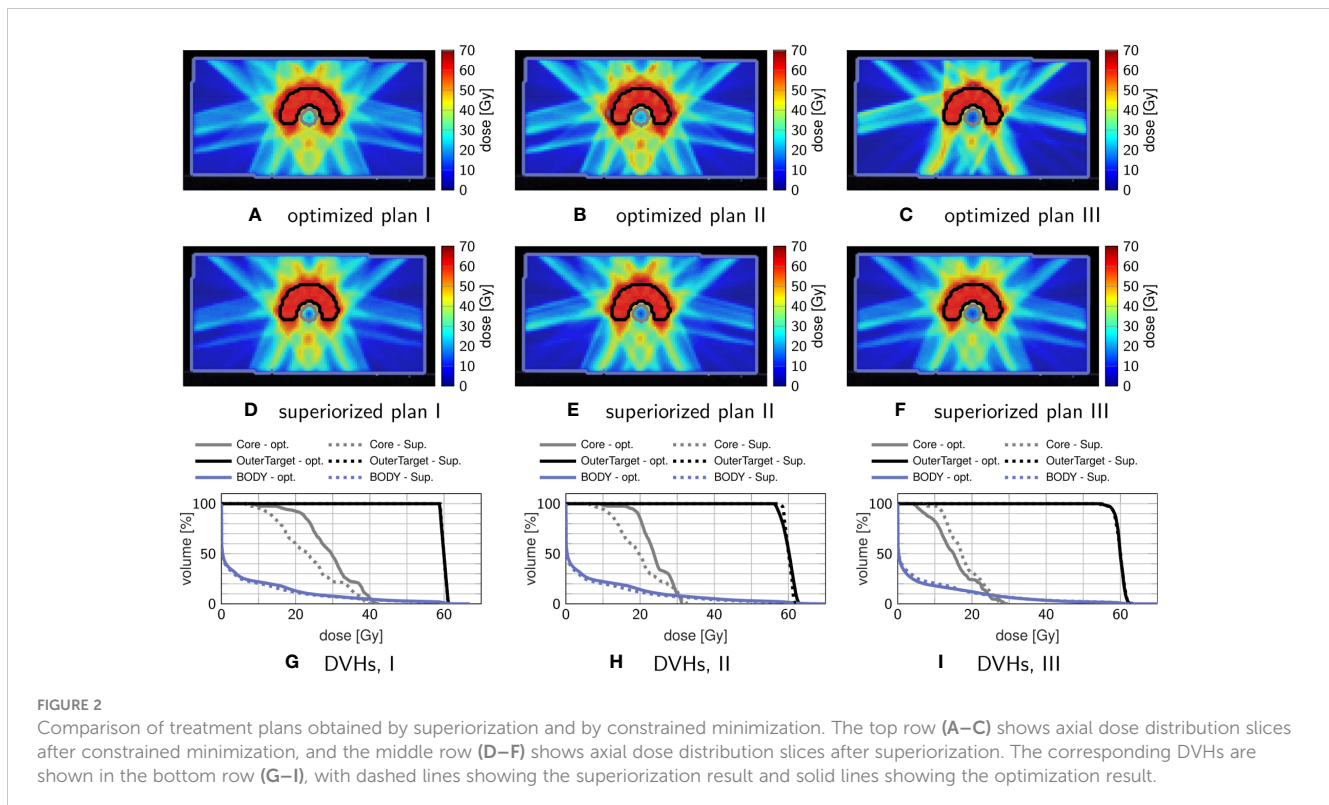
As reference, nonlinear constrained minimization with IPOPT with a logistic maximum approximation for minimum/maximum (compare (5), Table 1), was used. Three prescription scenarios were investigated: (I) linear inequalities on the target ($59 \text{ Gy} < d < 61 \text{ Gy}$), (II) additional linear inequalities on the “Core” structure ($d < 30 \text{ Gy}$), and (III) only linear inequalities on the “Core” ($d < 30 \text{ Gy}$). The parameters are detailed in Table 2.

Figure 2 compares dose distributions and DVHs after superiorization and after constrained minimization. The respective evolution of the objective function values and the constraint violations (calculated by the infinity norm over all inequality constraint functions, corresponding to the maximum residual) is exemplarily shown in Figure 3 for plan I.

Comparing plan quality, both plans adhere to the linear inequality constraints when the problem is feasible (which is the case for plans I & III) as seen in the DVHs. In plan I, superiorization appears to reach better OAR sparing with reduced mean and maximum dose, while in plan III constrained minimization achieves better OAR sparing. For plan II, which poses an infeasible problem, both target coverage and mean OAR sparing are improved for superiorization, yet at higher OAR maximum dose than obtained through constrained minimization.

The evolution of the objective function and constraint violation for plan I in Figure 3 exhibits a “typical” behavior of superiorization, seeing a strong decrease in the objective function values within the first iterations, followed by a slower slight increase as the perturbations’ step-sizes diminish. Both approaches were stopped after the maximum number of iterations (1000) was reached.

Nearly similar constraint violation is achieved by both methods, while constrained minimization resulted in higher objective



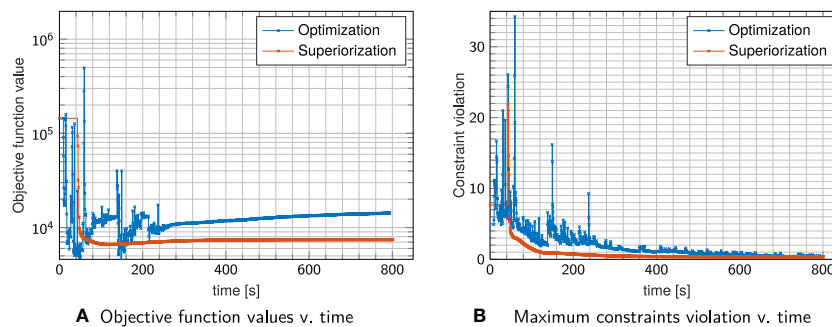


FIGURE 3 Objective Function values (A) and maximum constraint violation (B) over time for plan I shown in Figure 2. Each cross indicates a full iteration.

function values than superiorization, which can be attributed to the difference in OAR sparing. For all investigated plans I–III, superiorization showed a much “smoother” evolution of objective function and constraint violation than observed in the constrained minimization approach.

3.2 Head-and-neck case

To prove the usability of superiorization in a conventional planning setting, we applied the SM to a head-and-neck case with a wider range of available objective functions, i.e., including common DVH-based objectives.

Coverage of the planning target volumes (PTVs) was enforced using voxel inequality constraints. Again, the results of superiorization were compared to those obtained by solving the constrained minimization problem. All objectives and constraints are given in Table 3.

Both solvers use the same stopping criteria for the maximum constraint violation (smaller than 0.01 Gy is acceptable) and

objective function change of value (smaller than 0.1% in three consecutive iterations/sweeps).

Figure 4 shows exemplary axial dose slices and the DVHs for the plans generated with constraint minimization and with the SM. Quantitative runtime information and evolution of objective function and constraint violation are provided in Figure 5.

Both techniques were able to generate a plan that satisfies the linear inequalities up to the allowed violation threshold. Considering absolute runtime, the plan generated with the SM satisfied the stopping criteria after 400s, with constrained minimization failing to converge until the maximum number of iterations was reached.

SM spent most of the time in the first sweep/iteration, where it focuses on multiple objective function evaluations to generate a large initial decrease (as already observed above). It continuously decreases the objective function values together with decreasing constraints violation, reaching acceptable constraints violation more slowly than the run with constrained minimization.

However, using the same stopping criteria, the SM reached a solution with a much lower objective function value (approximately one-third of the value achieved by the constrained minimization plan). This is also visible in the dose slices and DVH, which show more normal tissue/OAR sparing for the SM plan. All results are, naturally, only valid for the experiments we performed. Further work, with varying algorithmic parameters, initialization points, and stopping criteria, is necessary to make more general statements.

TABLE 3 Dose inequality constraints, objective functions and penalty weights used for optimization and for superiorization on the head-and-neck case.

VOI	w_p	$c(d)$	$f(d)$
PTV70	1000	$66.5 \text{ Gy} < d < 77 \text{ Gy}$	$f_{sqdev}(d; 70 \text{ Gy})$
PTV63	1000		$f_{sqdev}(d; 63 \text{ Gy})$
PTV63	1000		$f_{minDVH}(d; 60 \text{ Gy}, 95\%)$
Spinal Cord PRV	100	$d < 50 \text{ Gy}$	$f_{sqdev+}(d; 15 \text{ Gy})$
Parotid L & R	100		$f_{sqdev+}(d; 10 \text{ Gy})$
Optic Nerve L & R	100		$f_{maxDVH}(d; 50 \text{ Gy}, 10\%)$
Larynx	300		$f_{sqdev+}(d; 15 \text{ Gy})$
Chiasm	100		$f_{maxDVH}(d; 50 \text{ Gy}, 10\%)$
Cerebellum	100		$f_{sqdev+}(d; 15 \text{ Gy})$
Brainstem PRV	100	$d < 30 \text{ Gy}$	$f_{sqdev+}(d; 15 \text{ Gy})$
NT/Body	100		$f_{mean}(d)$

The functions in the right-hand side column are identified here in Supplementary Data Sheet.

3.3 Prostate case

To demonstrate how the superiorization approach translates to a second patient, using a different irradiation modality, we create prostate IMPT plans with opposing fields on a 5mm spot grid using both superiorization and constrained minimization.

Figure 6 shows exemplary axial dose slices and the DVHs for the plans generated with constraint minimization and with the SM for the objective and constraint functions stated in Table 4.

The superiorized plan matches the dosimetric performance of the constrained minimization approach. Little increased dose in the rectum and bladder are traded against a slightly more homogeneous target coverage and reduced dose in the femoral heads.

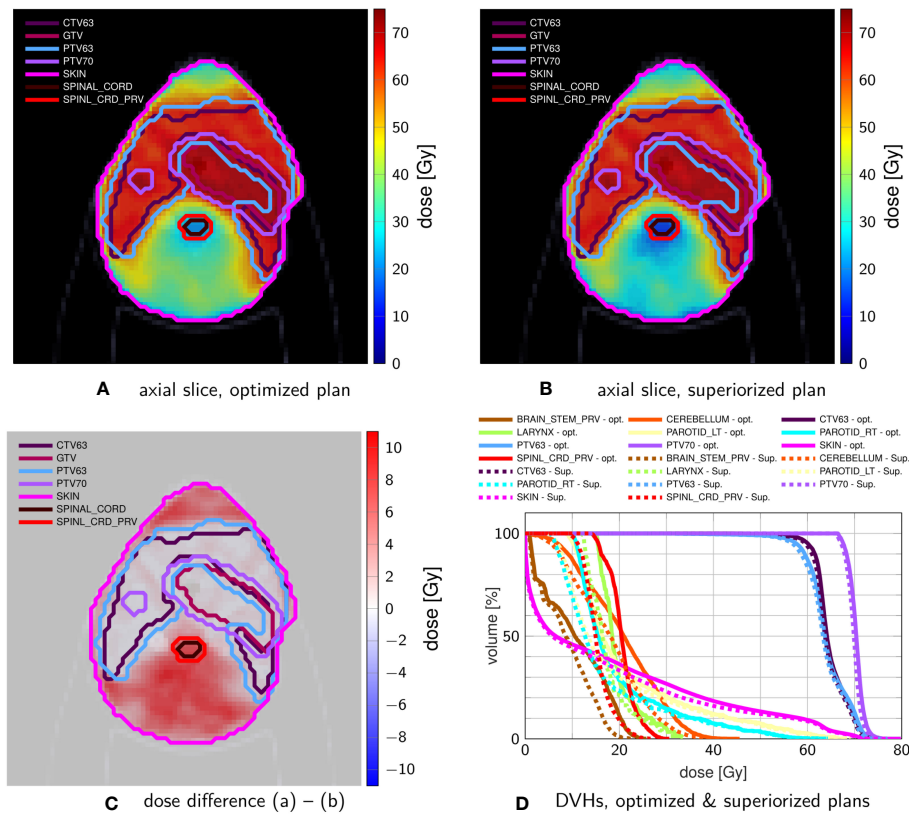


FIGURE 4 Comparison of head-and-neck treatment plans after (A) constrained minimization and after (B) superiorization (with AMS as the basic algorithm) using the tolerances from Table 3. (C) shows the dose difference in the same slice displayed in (A, B). (D) compares the resulting DVHs after optimization (solid) and superiorization (dashed).

4 Discussion

In this work, we applied the novel superiorization method, which solves a system of linear inequalities while reducing a nonlinear objective function, to inverse radiotherapy treatment planning. On a phantom and on a head-and-neck case, we

demonstrated that superiorization can produce treatment plans of similar quality to plans generated with constrained minimization.

Superiorization showed a smooth convergence behavior for both objective function reduction and constraint violation decrease, including the “typical” behavior of strong initial objective function reduction with subsequent diminishing objective function reduction

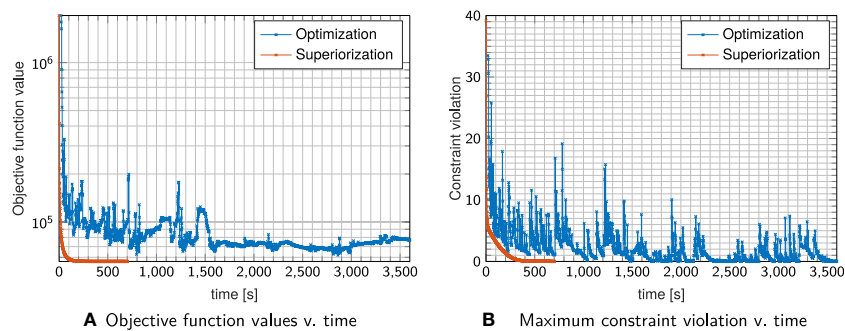


FIGURE 5 Evolution of objective Function values (A) and constraint violation (B) with runtime for the plan shown in Figure 4.

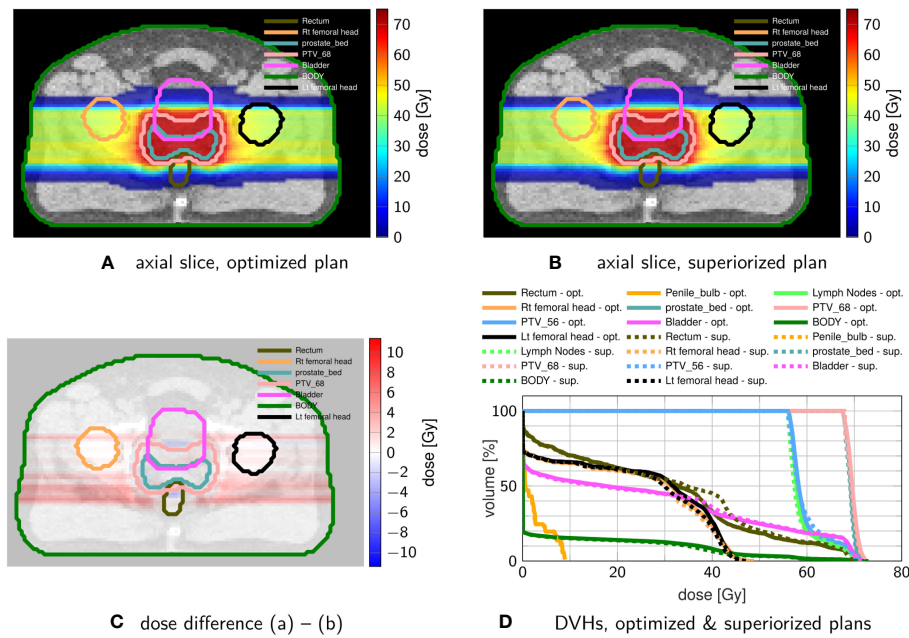


FIGURE 6 Comparison of prostate proton treatment plans after (A) constrained minimization with IPOPT and after (B) superiorization (with AMS as the basic algorithm) using the tolerances from Table 4. (C) shows the dose difference in the same slice displayed in (A, B). (D) compares the resulting DVHs after optimization (solid) and superiorization (dashed).

TABLE 4 Dose inequality constraints, objective functions and penalty weights used for optimization and for superiorization on the prostate proton case.

VOI	w_p	$c(d)$	$f(d)$
PTV68	5000	$68 \text{ Gy} < d < 72 \text{ Gy}$	$f_{sqdev}(d; 68 \text{ Gy})$
PTV56	3000	$56 \text{ Gy} < d < 72 \text{ Gy}$	$f_{sqdev}(d; 56 \text{ Gy})$
Rectum	100		$f_{sqdev+}(d; 30 \text{ Gy})$
Rectum	300		$f_{maxDVH}(d; 50 \text{ Gy}, 20\%)$
Bladder	100		$f_{sqdev+}(d; 30 \text{ Gy})$
Bladder	300		$f_{maxDVH}(d; 50 \text{ Gy}, 20\%)$
Femoral Heads	100		$f_{sqdev+}(d; 30 \text{ Gy})$
NT/Body	100		$f_{mean}(d)$

The functions in the right-hand side column are identified here in [Supplementary Data Sheet](#).

– including potential slight increase – while proximity to the feasible set within the dose inequality constraints is achieved.

4.1 The mathematical framework of constrained minimization and of superiorization for treatment planning

At the heart of the superiorization algorithm lies a feasibility-seeking algorithm (in this work, the AMS relaxation method for linear inequalities). This means that superiorization handles the treatment

planning problem as a feasibility-seeking problem for linear inequality dose constraints that should be fulfilled while reducing (not necessarily minimizing) an objective function along the way.

Constrained optimization algorithms, on the other hand, tackle the same data, i.e., constraints and objective function, as a full-fledged optimization problem. With the IPOPT package, for example, inequality constraints become logarithmic barrier functions and are incorporated as a linear combination into the Lagrangian function, whose minimization then enforces the constraints (19).

When the problem is hardly feasible, finding the right Lagrange multipliers may then dominate the optimization problem in its initial stages. Superiorization with a feasibility-seeking projection algorithm will smoothly reduce the proximity to the constraints, even for infeasible constrained problems, while the perturbations in the objective function reduction phase reduce the objective function value.

Our current implementation is, however, specifically geared for linear constraints. Yet other works on feasibility-seeking have shown that other relevant constraints, like, e.g., DVH constraints, can be incorporated into the feasibility-seeking framework, since they can still be interpreted as linear inequalities on a subset (relative volume) of voxels (25–27).

4.2 Comparability of runtime, convergence and stopping criteria

We demonstrated that feasibility-seeking for inverse IMRT treatment planning is practically equivalent to the least-squares

approach if similar prescriptions are set. However, obtaining the final solution with feasibility-seeking took more time than with unconstrained minimization with our prototype implementation in Matlab.

Stopping criteria, convergence and runtimes are more comparable when considering the constrained minimization vis-à-vis superiorization. Our prototype superiorization algorithm “converged” as fast as the used constrained nonlinear minimization algorithm when using the same objective functions and linear inequalities, exhibiting smoother progress during the iterations. It is interesting to note that even so SM is not guaranteed to find an optimal solution it sometimes exhibits better initial behavior than the constrained minimization algorithm. A similar phenomenon has been observed in the past by Censor et al. (53), wherein the SM was compared with a projected subgradient method (PSM) on a CT image reconstruction problem from projections in computerized tomography.

Recognizing the limited scope of the experiments presented here, our results about the superiorization method need further work to become well established. For example, the stopping criteria play a substantial role in both optimization and superiorization. Further modification of the respective parameters may lead to earlier or later stopping of either of the algorithms. Particularly the quasi-Newton algorithm will likely improve on its solution when allowing more iterations/longer runtimes. However, we suspect that the Lagrangian is particularly difficult to navigate when using a Hessian approximation over exact Hessian computations in these heavily constrained examples. This suspicion is supported by a solver benchmark performed by ten Eikelder et al. (63).

Consequently, runtime and convergence of a constrained nonlinear optimization algorithm would expectedly improve when incorporating second derivatives, such as proposed by van Haveren and Breedveld (11), instead of relying on a low-memory approximation to the quasi-Newton approach. In addition, alternative nonlinear minimum/maximum dose constraint implementations are possible. An advantage of the SM is that such “workarounds” are not necessary.

For superiorization, computational complexity and convergence are heavily dependent on the chosen feasibility-seeking algorithm. While the function reduction in superiorization has the computational complexity of gradient descent steps, the basic AMS algorithm used as a starting point performs sequential projections over all constraints. The complexity is thus principally comparable to the corresponding submatrix-vector products, however, the algorithm’s sequential structure complicates parallelization and other computational optimizations. Thus, modifications of the AMS algorithm are still actively researched (e.g., 64). Computational complexity and convergence properties of projection algorithms are a topic of ongoing research (see, e.g., 65, where it is discussed in a more general setting).

Despite these limitations, we demonstrated that a straightforward superiorization implementation was able to solve the given treatment planning problem arriving at dosimetrically comparable treatment plans.

4.3 Dosimetric performance

The treatment plans obtained with constrained minimization and with superiorization show some dosimetric differences. For the

three different linearly constrained setups on the TG199 phantom, these differences were most pronounced on the OAR, and less pronounced for the target dosage.

In the setups with target dose inequality constraints, superiorization reached better OAR sparing. This may be a result of multiple interacting factors: the strong initial objective function decrease in superiorization pulling down the dose in the OAR, and potential too early stopping of the constrained minimizer.

Further, in the infeasible setting with linear inequality constraints on both target and OAR, superiorization has the advantage that the feasibility-seeking algorithm will still smoothly converge to a proximal point.

The improved OAR sparing did not occur when only using dose inequality constraints on the OAR. However, in this case, the differences in DVHs of the OAR are only substantial below a dose of 20Gy and, thus, of limited significance, since a piece-wise least-squares objective was used that does not contribute to the objective function at dose values below 20Gy.

The head-and-neck case also reproduces the better OAR sparing for all evaluated OARs, at slightly reduced target coverage for the non-constrained CTV63 and PTV63. Here, the difference in convergence speed was most significant. Through all cases, the superiorization exhibited the smooth evolution of both objective function value and constraint violation, which in turn suggests robustness against changes in the stopping criteria as well. This behavior of superiorization could be underlined by translating it to IMPT on a prostate case.

These encouraging results show that superiorization can create acceptable and apparently “better” treatment plans. Additional work on more cases or planning benchmarks, with varying tuning parameters of both constrained minimization and superiorization approaches is needed to assess the convergence, runtime, and dosimetric quality of the solutions.

4.4 Outlook

With the proof-of-concept put forward in this work, there are many possible directions to further investigate the application of superiorization algorithms to the radiotherapy inverse treatment planning problem. From the perspective of a treatment planner, one may focus on enabling further constraints, e.g., DVH-based constraints, that are often used in treatment planning.

Some of these constraints are also representable as modified linear inequalities or convex and non-convex sets and, thus, can efficiently be solved using a feasibility-seeking algorithm. Even nonlinear constraints that are based, for example, on normal-tissue complication probability or equivalent uniform dose could be incorporated in the current definition of the superiorization algorithm if the “basic algorithm” in the feasibility-seeking phase of the SM is replaced by any other perturbation resilient projection method that can handle nonlinear constraints. Such algorithms exist in the literature.

Moreover, superiorization might also be extended to use more complex function reduction steps and inherent criteria. For example, a “true” backtracking line search could be performed,

similar to approaches in optimization, since a perturbation resilient “basic algorithm” might be able to handle much more complex function reduction steps.

Considering these algorithmic and application-focused improvements, the SM should also be rigorously tested on radiotherapy optimization/inverse planning benchmark problems, like the TROTS dataset (66), as soon as it is able to handle the respective problem formulations. With this, transferability to other modalities like ion therapy or volumetric modulated arc therapy (VMAT) is also within reach.

5 Conclusions

We introduced superiorization as a novel inverse planning technique, merging feasibility-seeking for linear inequality dose constraints with objective function reduction. Our initial comparison of superiorization with constrained minimization using linear dose-inequalities suggests possible dosimetric benefits and smoother convergence. Superiorization is thus a valuable addition to the algorithmic inverse treatment planning toolbox.

Data availability statement

All data used in this study is publicly available through matRad. The code to reproduce our results is available on GitHub: <https://github.com/e0404/paper-superiorization-imrt>. Further inquiries can be directed to the corresponding authors.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

NW and YC conceived the original idea for the research. YC developed the theoretical framework for the study, while FB implemented the method. NW conducted the analyses of the

algorithm. All authors contributed to the writing and editing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Open access funding by ETH Zurich. This work is supported by the ISF-NSFC joint research program grant No. 2874/19 (YC), by the German Research Foundation (DFG), grant No. 443188743 (NW) and by the U.S. National Institutes of Health, grant No. R01CA266467 (YC and NW). Parts of this work pertain to project support (YC and NW) from the Cooperation Program in Cancer Research of the German Cancer Research Center (DKFZ) and Israel's Ministry of Innovation, Science and Technology (MOST).

Acknowledgments

We thank Mark Bangert for taking part in the first discussion rounds leading up to this work. Finally, the authors would like to thank the editor and two reviewers for their constructive comments and suggestions that have helped to improve the quality 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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1238824/full#supplementary-material>

References

1. Bortfeld T, Thieke C. Optimization of Treatment Plans, Inverse Planning. In: Schlegel W, Bortfeld T, Grosu AL, editors. *New Technologies in Radiation Oncology*. Berlin/Heidelberg: Springer-Verlag (2006). p. 207–20. doi: 10.1007/3-540-29999-8_17
2. Wu Q, Mohan R. Algorithms and functionality of an intensity modulated radiotherapy optimization system. *Med Phys* (2000) 27:701–11. doi: 10.1118/1.598932

3. Breedveld S, van den Berg B, Heijmen B. An interior-point implementation developed and tuned for radiation therapy treatment planning. *Comput Optimization Appl* (2017) 68:209–42. doi: 10.1007/s10589-017-9919-4
4. Fogliata A, Nicolini G, Alber M, Åsell M, Clivio A, Dobler B, et al. On the performances of different IMRT treatment planning systems for selected paediatric cases. *Radiat Oncol* (2007) 2:7. doi: 10.1186/1748-717X-2-7
5. Wieser HP, Cisternas E, Wahl N, Ulrich S, Stadler A, Mescher H, et al. Development of the open-source dose calculation and optimization toolkit matRad. *Med Phys* (2017) 44:2556–68. doi: 10.1002/mp.12251
6. Luenberger DG, Ye Y. Linear and Nonlinear Programming. In: *International Series in Operations Research and Management Science, 3rd ed.* New York, NY: Springer (2008). p. 546.
7. Bazaraa MS, Sherali HD, Shetty CM. *Nonlinear Programming: Theory and Algorithms, 3rd ed.* Hoboken, NJ: Wiley-Interscience (2006). p. 872.
8. Küfer KH, Monz M, Scherrer A, Süß P, Alonso F, Sultan ASA, et al. *Multicriteria Optimization in Intensity Modulated Radiotherapy Planning*. Kaiserslautern: ITWM Report 77. Fraunhofer (ITWM) (2005).
9. Thieke C, Küfer KH, Monz M, Scherrer A, Alonso F, Oelfke U, et al. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. *Radiotherapy Oncol* (2007) 85:292–8. doi: 10.1016/j.radonc.2007.06.020
10. Unkelbach J, Alber M, Bangert M, Bokrantz R, Chan TCY, Deasy JO, et al. Robust radiotherapy planning. *Phys Med Biol* (2018) 63:22TR02. doi: 10.1088/1361-6560/aae659
11. Van Haveren R, Breedveld S. Fast and exact Hessian computation for a class of nonlinear functions used in radiation therapy treatment planning. *Phys Med Biol* (2019) 64:16NT01. doi: 10.1088/1361-6560/ab1e17
12. Kierkels RG, Korevaar EW, Steenbakkers RJ, Janssen T, van't Veld AA, Lan-653 gendijk JA, et al. Direct use of multivariable normal tissue complication probability models in treatment plan optimisation for individualised head and neck cancer radiotherapy produces clinically acceptable treatment plans. *Radiotherapy Oncol* (2014) 112:430–6. doi: 10.1016/j.radonc.2014.08.020
13. Christiansen E, Heath E, Xu T. Continuous aperture dose calculation and optimization for volumetric modulated arc therapy. *Phys Med Biol* (2018) 63:21NT01. doi: 10.1088/1361-6560/aae65e
14. Gao H, Lin B, Lin Y, Fu S, Langen K, Liu T, et al. Simultaneous dose and dose rate optimization (SDDRO) for FLASH proton therapy. *Med Phys* (2020) 47:6388–95. doi: 10.1002/mp.14531
15. Hahn C, Heuchel L, Öden J, Traneus E, Wulff J, Plaude S, et al. Comparing biological effectiveness guided plan optimization strategies for cranial proton therapy: potential and challenges. *Radiat Oncol (London England)* (2022) 17:169. doi: 10.1186/s13014-022-02143-x
16. Ten Eikelder SCM, Ajdari A, Bortfeld T, den Hertog D. Adjustable robust treatment-length optimization in radiation therapy. *Optimization Eng* (2022) 23:1949–86. doi: 10.1007/s11081-021-09709-w
17. Faddegon BA, Blakely EA, Burigo LN, Censor Y, Dokic I, D-Kondo JN, et al. Ionization detail parameters and cluster dose: a mathematical model for selection of nanodosimetric quantities for use in treatment planning in charged particle radiotherapy. *Phys Med Biol* (2023) 68:175013. doi: 10.1088/1361-6560/acea16
18. Liu R, Charyyev S, Wahl N, Liu W, Kang M, Zhou J, et al. An Integrated Physical Optimization framework for proton SBRT FLASH treatment planning allows dose, dose rate, and LET optimization using patient-specific ridge filters. *Int J Radiat Oncology Biology Phys* (2023) 116(4):949–59. doi: 10.1016/j.ijrobp.2023.01.048
19. Wächter A, Biegler LT. On the implementation of an interior-point filter linesearch algorithm for large-scale nonlinear programming. *Math Programming* (2006) 106:25–57. doi: 10.1007/s10107-004-0559-y
20. Nocedal J, Wright SJ. *Numerical Optimization*. New York: Springer (1999). p. 636.
21. Censor Y, Altschuler MD, Powlis WD. A computational solution of the inverse problem in radiation-therapy treatment planning. *Appl Mathematics Comput* (1988) 25:57–87. doi: 10.1016/0096-3003(88)90064-1
22. Powlis WD, Altschuler MD, Censor Y, Buhle EL. Semi-automated radiotherapy treatment planning with a mathematical model to satisfy treatment goals. *Int J Radiat OncologyBiologyPhysics* (1989) 16:271–6. doi: 10.1016/0360-3016(89)90042-4
23. Bauschke HH, Borwein JM. On projection algorithms for solving convex feasibility problems. *SIAM Rev* (1996) 38:367–426. doi: 10.1137/S0036144593251710
24. Cho PS, Marks RJ. Hardware-sensitive optimization for intensity modulated radiotherapy. *Phys Med Biol* (2000) 45:429–40. doi: 10.1088/0031-9155/45/2/312
25. Penfold S, Zalas R, Casiraghi M, Brooke M, Censor Y, Schulte R. Sparsity constrained split feasibility for dose-volume constraints in inverse planning of intensity-modulated photon or proton therapy. *Phys Med Biol* (2017) 62:3599–618. doi: 10.1088/1361-6560/aa602b
26. Brooke M, Censor Y, Gibali A. Dynamic string-averaging CQ-methods for the split feasibility problem with percentage violation constraints arising in radiation therapy treatment planning. *Int Trans Operational Res* (2020) 30:181–205. doi: 10.1111/itor.12929
27. Gadoue SM, Toomeh D, Schultze BE, Schulte RW. A dose-volume constraint (DVC) projection-based algorithm for IMPT inverse planning optimization. *Med Phys* (2022) 49(4):2699–708. doi: 10.1002/mp.15504
28. Simon HA. Rational choice and the structure of the environment. *psychol Rev* (1956) 63:129–38. doi: 10.1037/h0042769
29. Alber M, Meedt G, Nüsslin F, Reemtsen R. On the degeneracy of the IMRT optimization problem. *Med Phys* (2002) 29:2584. doi: 10.1118/1.1500402
30. Censor Y. Superiorization and perturbation resilience of algorithms: A continuously updated bibliography. (2022). doi: 10.48550/arXiv.1506.04219
31. Herman GT, Garduño E, Davidi R, Censor Y. Superiorization: An optimization heuristic for medical physics. *Med Phys* (2012) 39:5532–46. doi: 10.1118/1.4745566
32. Guenter M, Collins S, Ogilvy A, Hare W, Jirasek A. Superiorization versus regularization: A comparison of algorithms for solving image reconstruction problems with applications in computed tomography. *Med Phys* (2022) 49:1065–82. doi: 10.1002/mp.15373
33. Yang Q, Cong W, Wang G. Superiorization-based multi-energy CT image reconstruction. *Inverse Problems* (2017) 33:044014. doi: 10.1088/1361-6420/aa5e0a
34. Penfold S, Censor Y. Techniques in iterative proton CT image reconstruction. *Sens Imaging* (2015) 16:19. doi: 10.1007/s11220-015-0122-3
35. Schultze B, Censor Y, Karbasi P, Schubert KE, Schulte RW. An improved method of total variation superiorization applied to reconstruction in proton computed tomography. *IEEE Trans Med Imaging* (2020) 39:294–307. doi: 10.1109/TMI.2019.2911482
36. Han W, Wang Q, Cai W. Computed tomography imaging spectrometry based on superiorization and guided image filtering. *Optics Lett* (2021) 46:2208–11. doi: 10.1364/OL.418355
37. Pakkaranang N, Kumam P, Berinde V, Suleiman YI. Superiorization methodology and perturbation resilience of inertial proximal gradient algorithm with application to signal recovery. *J Supercomputing* (2020) 76:9456–77. doi: 10.1007/s11227-020-03215-z
38. Davidi R, Censor Y, Schulte R, Geneser S, Xing L. Feasibility-Seeking and Superiorization Algorithms Applied to Inverse Treatment Planning in Radiation Therapy. In: Reich S, Zaslavski A, editors. *Contemporary Mathematics*, vol. 636. Providence, Rhode Island: American Mathematical Society (2015). p. 83–92. doi: 10.1090/conm/636/12729
39. Bonacker E, Gibali A, Küfer KH, Süß P. Speedup of lexicographic optimization by superiorization and its applications to cancer radiotherapy treatment. *Inverse Problems* (2017) 33:044012. doi: 10.1088/1361-6420/33/4/044012
40. Alber M, Reemtsen R. Intensity modulated radiotherapy treatment planning by use of a barrier-penalty multiplier method. *Optimization Methods Software* (2007) 22:391–411. doi: 10.1080/10556780600604940
41. Bortfeld T, Bürkelbach J, Boesecke R, Schlegel W. Methods of image reconstruction from projections applied to conformation radiotherapy. *Phys Med Biol* (1990) 35:1423–34. doi: 10.1088/0031-9155/35/10/007
42. Carlsson F, Forsgren A. Iterative regularization in intensity-modulated radiation therapy optimization. *Med Phys* (2006) 33:225–34. doi: 10.1118/1.2148918
43. Censor Y, Bortfeld T, Martin B, Trofimov A. A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys Med Biol* (2006) 51:2353–65. doi: 10.1088/0031-9155/51/10/001
44. Bauschke HH, Koch VR. Projection Methods: Swiss Army Knives for Solving Feasibility and Best Approximation Problems with Halbspaces. In: Reich S, Zaslavski AJ, editors. *Contemporary Mathematics*, vol. 636 Providence, RI: American Mathematical Society (AMS) (2015). p. 1–40. doi: 10.1090/conm/636
45. Davidi R, Herman G, Censor Y. Perturbation-resilient block-iterative projection methods with application to image reconstruction from projections. *Int Trans Operational Res* (2009) 16:505–24. doi: 10.1111/j.1475-3995.2009.00695.x
46. Gordon D, Gordon R. Component-averaged row projections: A robust, blockParallel scheme for sparse linear systems. *SIAM J Sci Computing* (2005) 27:1092–117. doi: 10.1137/040609458
47. Bargetz C, Reich S, Zalas R. Convergence properties of dynamic string-averaging projection methods in the presence of perturbations. *Numerical Algorithms* (2018) 77:185–209. doi: 10.1007/s11075-017-0310-4
48. Agmon S. The relaxation method for linear inequalities. *Can J Mathematics* (1954) 6:382–92. doi: 10.4153/CJM-1954-037-2
49. Motzkin TS, Schoenberg IJ. The relaxation method for linear inequalities. *Can J Mathematics* (1954) 6:393–404. doi: 10.4153/CJM-1954-038-x
50. Censor Y, Zenios SA. *Parallel optimization: Theory, Algorithms, and Applications*. New York, NY, USA: Oxford University Press (1998).
51. Censor Y, Schubert KE, Schulte RW. Developments in mathematical algorithms and computational tools for proton CT and particle therapy treatment planning. *IEEE Trans Radiat Plasma Med Sci* (2022) 6:313–24. doi: 10.1109/TRPMS.2021.3107322
52. Censor Y. Weak and strong superiorization: between feasibility-seeking and minimization. *Analele Universitatii "Ovidius" Constanta - Seria Matematica* (2015) 23:41–54. doi: 10.1515/auom-2015-0046
53. Censor Y, Davidi R, Herman GT, Schulte RW, Tetrushvili L. Projected subgradient minimization versus superiorization. *J Optimization Theory Appl* (2014) 160:730–47. doi: 10.1007/s10957-013-0408-3
54. Herman GT. Superiorization for Image Analysis. In: *Proceedings of the 16th International Workshop on Combinatorial Image Analysis*, vol. 8466. Berlin, Heidelberg: Springer-Verlag, IWCIA (2014). p. 1–7. doi: 10.1007/978-3-319-07148-0_1

55. Censor Y. Can linear superiorization be useful for linear optimization problems? *Inverse Problems* (2017) 33:044006. doi: 10.1088/1361-6420/33/4/044006
56. Censor Y, Zaslavski AJ. Convergence and perturbation resilience of dynamic string-averaging projection methods. *Comput Optimization Appl* (2013) 54:65–76. doi: 10.1007/s10589-012-9491-x
57. Cegielski A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*. New York: Springer (2012). p. 316.
58. Ackermann B, Bangert M, Bennan ABA, Burigo L, Cabal G, Cisternas E, et al. *matRad. Version 2.10.1*. Heidelberg: Deutsches Krebsforschungszentrum (2020). doi: 10.5281/zenodo.7107719
59. Cisternas E, Mairani A, Ziegenhein P, Jäkel O, Bangert M. matRad - a multi-modality open source 3D treatment planning toolkit. In: *IFMBE Proceedings*, vol. 51. Cham: Springer International Publishing (2015). p. 1608–11. doi: 10.1007/978-3-319-19387-8_391
60. Bortfeld T, Schlegel W, Rhein B. Decomposition of pencil beam kernels for fast dose calculations in three-dimensional treatment planning. *Med Phys* (1993) 20:311–8. doi: 10.1118/1.597070
61. Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys* (2009) 36:5359–73. doi: 10.1118/1.3238104
62. Craft D, Bangert M, Long T, Papp D, Unkelbach J. Shared data for intensity modulated radiation therapy (IMRT) optimization research: the CORT dataset. *GigaScience* (2014) 3:2047–217X–3–37. doi: 10.1186/2047-217X-3-37
63. Ten Eikelder SCM, Ajdari A, Bortfeld T, den Hertog D. Conic formulation of fluence map optimization problems. *Phys Med Biol* (2021) 66:225016. doi: 10.1088/1361-6560/ac2b82
64. De Loera JA, Haddock J, Needell D. A sampling kaczmarz–motzkin algorithm for linear feasibility. *SIAM J Sci Computing* (2017) 39:S66–S87. doi: 10.1137/16M1073807
65. Chen C, Fu X, He B, Yuan X. On the iteration complexity of some projection methods for monotone linear variational inequalities. *J Optimization Theory Appl* (2017) 172:914–28. doi: 10.1007/s10957-016-1051-6
66. Breedveld S, Heijmen B. Data for TROTS – the radiotherapy optimisation test set. *Data Brief* (2017) 12:143–9. doi: 10.1016/j.dib.2017.03.037