Comparative Effects of Event Detection Methods on the Analysis and Interpretation of Ca²⁺ Imaging Data

Austin Neugornet[1], Bernadette O'Donovan[2] and Pavel Ivanovich Ortinski[1]*

[1] Department of Neuroscience, School of Medicine, University of Kentucky, Lexington, KY, United States, [2] Department of Pharmacology, Physiology and Neuroscience, University of South Carolina School of Medicine, Columbia, SC, United States

Calcium imaging has gained substantial popularity as a tool to profile the activity of multiple simultaneously active cells at high spatiotemporal resolution. Among the diverse approaches to processing of Ca²⁺ imaging data is an often subjective decision of how to quantify baseline fluorescence or F_0. We examine the effect of popular F_0 determination methods on the interpretation of neuronal and astrocyte activity in a single dataset of rats trained to self-administer intravenous infusions of cocaine and compare them with an F_0-independent wavelet ridgewalking event detection approach. We find that the choice of the processing method has a profound impact on the interpretation of widefield imaging results. All of the dF/F_0 thresholding methods tended to introduce spurious events and fragment individual transients, leading to smaller calculated event durations and larger event frequencies. Analysis of simulated datasets confirmed these observations and indicated substantial intermethod variability as to the events classified as significant. Additionally, most dF/F_0 methods on their own were unable to adequately account for bleaching of fluorescence, although the F_0 smooth approach and the wavelet ridgewalking algorithm both did so. In general, the choice of the processing method led to dramatically different quantitative and sometimes opposing qualitative interpretations of the effects of cocaine self-administration both at the level of individual cells and at the level of cell networks. Significantly different distributions of event duration, amplitude, frequency, and network measures were found across the majority of dF/F_0 approaches. The wavelet ridgewalking algorithm broadly outperformed dF/F_0-based methods for both neuron and astrocyte recordings. These results indicate the need for heightened awareness of the limitations and tendencies associated with decisions to use particular Ca²⁺ image processing pipelines. Both quantification and interpretation of the effects of experimental manipulations are strongly sensitive to such decisions.

Keywords: calcium imaging, event identification, network analysis, astrocytes, calcium transient detection

INTRODUCTION

Development of genetically encoded calcium indicators (GECIs) has encouraged a bloom of research to capture the activity of large cell populations at high spatiotemporal resolution. The activity of diverse cell types can be examined using Ca^{2+} imaging approaches. For example, amplitude and duration of neuronal Ca^{2+} transients are typically interpreted as proxies for action potentials (Ali and Kwan, 2020), and diverse features of astrocytic Ca^{2+} have been proposed to play functional roles in neural circuit regulation, behavior, and information processing (Guerra-Gomes et al., 2017). Accurate identification of Ca^{2+} transients is critical to facilitate understanding of dynamic behavior in individual cells as well as circuit relationships that may arise from cell–cell interactions. A variety of analytical approaches to process large amounts of imaging data are available, but the extent to which the choice of the analysis pipeline may impact interpretation of the underlying data remains unclear.

Most "traditional" event identification methods are based on transforming a fluorescence signal into dF/F_0 and then applying a threshold to find significant deviations from the presumed background noise. Definitions of baseline F_0 and thresholding methods vary widely across studies. In our analyses, we implement three common F_0 definitions: an initial segment of the fluorescence trace (F_0 initial), often used in recordings involving extracellular stimulations (e.g., Ellefsen et al., 2014; Lock et al., 2015; Rahmati et al., 2016); a minimally variable and dim segment (F_0 minimal); and as a fit to the background in a sliding window throughout the trace (F_0 smooth), such as implemented in a toolbox by Romano et al. (2017), SICT (Mancini et al., 2018), or FluoroSNNAP (Patel et al., 2015).

Wavelet ridgewalking algorithms, as originally developed for use in mass spectroscopy signal analysis (Du et al., 2006), have been recently employed to find events in neuronal imaging data (Prada et al., 2018). Wavelet ridgewalking algorithms make minimal assumptions about the characteristics of events they identify, requiring only that they are "peak-like" at some scale. Events across a wide range of durations and amplitude scales can be identified, even in the presence of high levels of noise. The wavelet method lacks precise event shape, duration, or amplitude requirements, increasing its versatility and potential to identify morphologically diverse Ca^{2+} waveforms. This is of particular importance for processing astrocytic data, as these cells are known to exhibit highly heterogeneous Ca^{2+} signals (Shigetomi et al., 2016; Verkhratsky and Nedergaard, 2018).

As Ca^{2+} imaging recordings capture activity within a broad local network, they can be used to explore network structure under a variety of paradigms. Graph theory has been developed to quantify and examine network characteristics, describing the organization of groups of objects of interest (nodes) through their connections to each other (edges) (Bullmore and Sporns, 2009). Graph theory is currently finding extensive use throughout neuroscience, with particularly prominent applications in neuroimaging and clinical psychiatry (Rubinov and Sporns, 2010; Bassett and Sporns, 2017) based on evaluation of pairwise connections between cells or brain areas (Sporns, 2018). Our analyses explore how detection of significant Ca^{2+} events at the level of single cells affects the analyses of network activity based on graph theory methods. To determine the impact of different analysis methodologies on interpretation of experimental outcomes, we utilize a dataset of nucleus accumbens shell imaging results from animals trained to self-administer cocaine. The nucleus accumbens is involved in motivation and drug-seeking behaviors. Multiple studies have shown that cocaine self-administration alters neuronal activity in the nucleus accumbens (Zhang et al., 2002; Pierce and Wolf, 2013; Calipari et al., 2016) and impacts astrocyte morphology (Scofield et al., 2016) with functional and behavioral consequences (Scofield, 2018; Kardos et al., 2019).

MATERIALS AND METHODS

Animal Training

Male Sprague–Dawley rats (*Rattus norvegicus*) weighing 225–250 g were obtained from Taconic Laboratories. Rats were individually housed with food and water available *ad libitum* in their home cages. A 12/12-h light/dark cycle was used with lights on 7:00 am–7:00 pm. All experimental procedures were performed during the light cycle. The experimental protocols were approved by the University of South Carolina and University of Kentucky Animal Care and Use Committees.

Two cohorts of eight rats each were used for the experiments. Both cohorts underwent jugular catheterization surgery: under isoflurane anesthesia (1.5–2.5% isoflurane in O$_2$), an indwelling silastic catheter (SAI Infusion technologies, Lake Villa, IL, United States) was placed into the right jugular vein and sutured in place. The catheter was routed subcutaneously to a mesh platform between the shoulder blades. Catheters were flushed daily with 0.3 ml of the antibiotic timentin (0.93 mg/ml) dissolved in heparinized saline. The catheters were sealed with plastic obturators when not in use.

Following catheter implantation, rats were placed in a stereotaxic apparatus under inhalational 1.5–2.5% isoflurane. Genetically encoded calcium indicator (GECI) GCaMP6f targeting either neurons (AAV9.Syn.GCaMP6f.WPRE.SV40, Addgene 100837) or astrocytes (pZac2.1.gfaABC1D.cytoGCamp6f.SV40, Addgene 52925) was injected bilaterally (2 µl/side) into the nucleus accumbens (NAc) shell *via* a Neuros syringe (Hamilton) targeting the following coordinates (relative to bregma): 1.0 mm anterior, ±1.0 mm lateral, and 7.0 mm ventral. Rats began behavioral training 14–20 days after viral microinjections.

In each cohort, four animals underwent cocaine self-administration training and four others served as their yoked saline counterparts. For self-administration training, rats were placed in operant chambers (Med Associates) for 2 h/day for 14–18 days during which time two levers (one "active," one "inactive") were extended. For rats undergoing self-administration training, an active lever press resulted in an infusion of cocaine solution (0.21 mg cocaine/50 µl saline over 2.8 s). Each infusion was followed by a 20-s timeout during which lever pressing had no scheduled consequences. All animals started on a fixed ratio (FR) 1 schedule that was

increased to FR5 following acquisition of stable responding (<10 variability in active lever presses across three consecutive days). The yoked saline animals received a saline infusion every time their yoke counterpart received cocaine. There were no scheduled consequences to lever pressing by the yoked saline animals. Behavior results are shown in **Supplementary Figure 1**.

Ca^{2+} Imaging

After self-administration training, the cohort of animals injected with the neuron-specific GCamp6f (hSyn animals) were sacrificed by guillotine. The brain was excised and 300-μm-thick slices were prepared using a vibratome (Leica VT1200S) in an ice-cold artificial cerebrospinal fluid (aCSF) cutting solution in which NaCl was replaced with an equiosmolar concentration of sucrose. A total of 19 hSyn slices were imaged from seven different rats (3 yoked saline and 4 cocaine) with each rat contributing 1–4 slices to the final dataset. The cohort of animals injected with the astrocyte-targeting GCamp6f (GFAP animals) produced 22 slices from eight different rats with each rat contributing 2–3 slices. The aCSF composition was as follows (in mM): 130 NaCl, 3 KCl, 1.25 NaH$_2$PO$_4$, 26 NaHCO$_3$, 10 glucose, 1 MgCl$_2$, and 2 CaCl$_2$, pH 7.2–7.4, when saturated with 95% O$_2$ and 5% CO$_2$; osmolarity was 305–315 mOsm. After cutting, the slices were kept in aCSF heated to 36°C for 30–40 min, following which they were maintained in aCSF at room temperature until transfer to the recording chamber.

Two-minute videos of spontaneous activity in the NAc shell were acquired with an ORCA-Flash 4.0 (V2) digital camera during LED excitation (X-Cite XLED1, Excelitas Technologies). Videos were taken at 512 × 512 pixels and collected with a 40× objective (0.065 μm/pixel) at 25 frames per second.
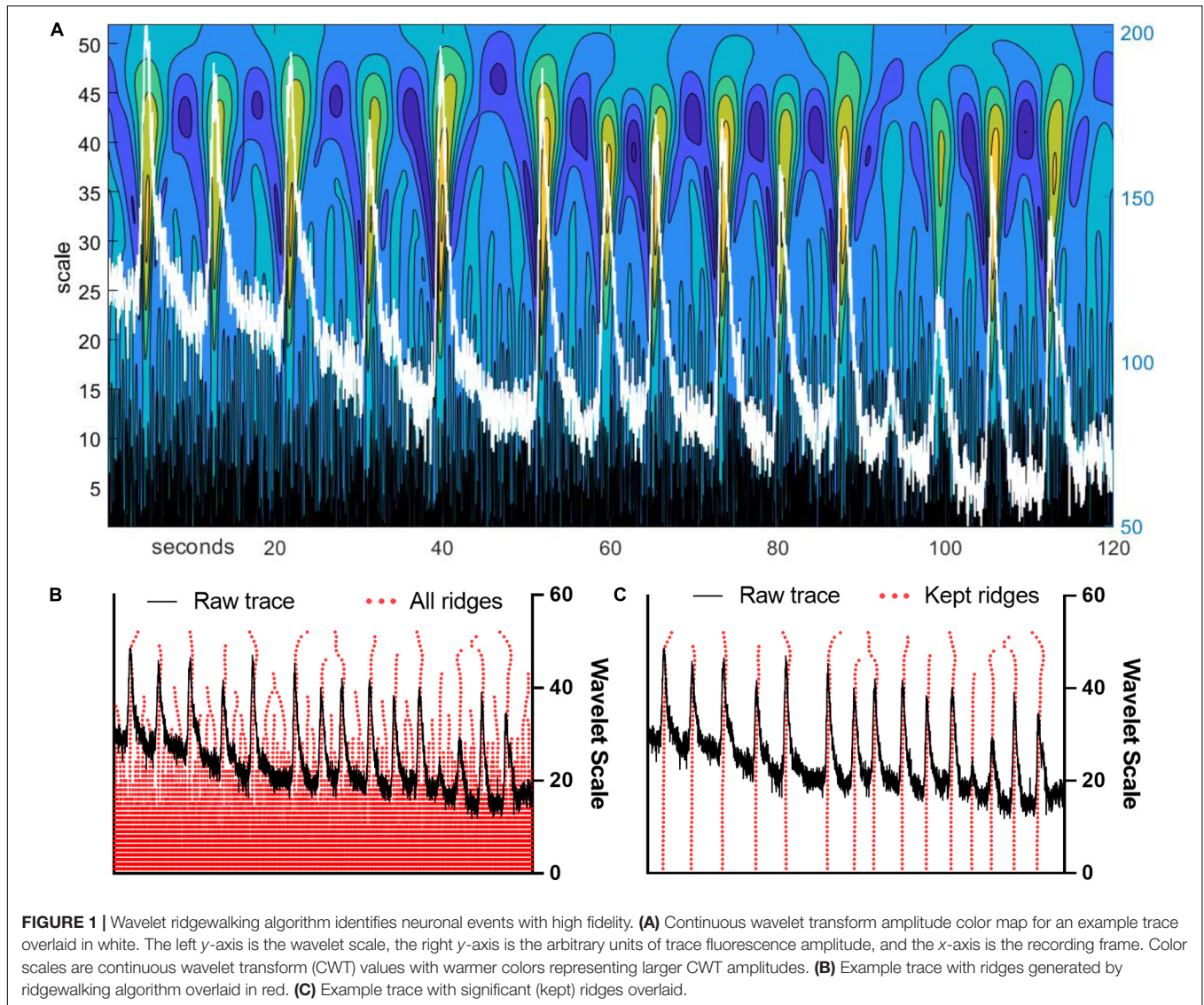
Image Analysis

We used a slightly modified version of ABLE (Reynolds et al., 2017) to segment our videos based on activity. The only changes we made were to remove the thresholds (caps) for maximum seed and final region of interest (ROI) size. For our dataset, we used ABLE parameters: alpha = 0.1, lambda = 150 (following Reynolds et al., 2017), and a radius of 20. Modified ABLE scripts are available upon request. The background and neuropil signal were identified for each ROI independently by finding the average (mean) trace in a 60-pixel-wide "halo" around a given ROI, excluding those points found to be in other ROIs. Individual neuropil signals were subtracted from the corresponding ROI traces. Once the ROIs were identified for a recording, those same ROIs and associated fluorescent signal traces were used for all subsequent method comparisons throughout the study. Examples of ABLE-detected ROIs for both neuron and astrocyte populations are presented in **Supplementary Figure 2**.

Event Detection

Events were identified from each ROI trace using either a dF/F_0 thresholding method or continuous wavelet transform of each trace followed by a custom-written ridgewalking algorithm. To construct dF/F_0 traces, we determined F_0 by three different methods: calculating the mean value of an initial window on the trace which we will refer to as "F_0 initial" (e.g.,

Ellefsen et al., 2014; Lock et al., 2015; Rahmati et al., 2016); finding the smoothed 8th percentile point in a sliding window producing a variable F_0 time series, referred to as "F_0 smooth" (e.g., Patel et al., 2015; Romano et al., 2017; Mancini et al., 2018); and calculating the mean value of a baseline segment found by minimizing the quantity (trace variance + trace mean^2)$^{1/2}$ in a sliding window on the trace, which we call "F_0 minimal." We then applied four different thresholding criteria to identify significant activity within a dF/F_0 trace: 2.5 standard deviations above the mean of either the baseline (baseline SD) or the total dF/F_0 trace (trace SD) and p-value of <0.05 as determined by z-test of either the baseline (baseline z-score) or the total dF/F_0 trace (trace z-score). The possible combinations of choice of F_0 and thresholding criteria yielded 12 different dF/F_0 event detection methods. These methods were compared to a wavelet ridgewalking algorithm that does not require definition of F_0 to identify significant events.

Before a wavelet ridgewalking algorithm is used for event detection, a continuous wavelet transform (CWT) is applied to a signal trace producing a two-dimensional sheet of coefficients revealing localized "ridges" of local maxima (**Figure 1A**). Significant events can be identified by the properties of these ridges (**Figures 1B,C**), particularly ridge length. The CWT was calculated using the jLab MATLAB toolbox (Lilly, 2019). We used a generalized Morse wavelet with (shape) parameter gamma = 3 and initial (scale) parameter beta = 2 (before scaling). Wavelets having gamma = 3, also called Airy wavelets, were chosen because they have the smallest Heisenberg area (Lilly and Olhede, 2012). We used a set of optimal scales/frequencies determined by jLab scripts for time series of length N = 3001, the number of frames in our videos. The completed CWT for a trace results in a two-dimensional sheet of CWT amplitudes, as transform amplitudes for wavelets of all scale sizes are calculated for each time point in the trace. We then generate ridges along the scale dimension using a custom-written MATLAB ridgewalking function. The algorithm is as follows: (a) All local CWT amplitude maxima for *each* scale are identified. These local maxima serve as new ridge initializations and candidates for ridge extension. (b) The local CWT amplitude maxima at the largest scale size are the first ridge initializations. Progressing to smaller scales, local maxima are linked into a ridge if they fall within the temporal range determined by the central window for the wavelet of the preceding scale size when the window is centered on the ridges and ridge initializing maxima at that (larger) scale. If multiple maxima at the smaller scale fall into the range determined for a ridge, only the largest of the new maxima is appended to the ridge. Others within the window initialize new ridges. If a local maximum does not fall into any of the previous ridges' or maxima's windows, it also initializes a new ridge. (c) The algorithm ends either after the maxima at the smallest scale have been processed, at which point all local maxima at all scales have been formed into ridges (or identified as single points), or once a user-specified scale has been reached. Gaps in scale along ridges are not allowed. Scale gaps terminate ridges. (d) Once all ridges for a trace are identified, ridge significance is determined by passing through at least 41 scales. Additionally, ridges whose largest CWT maxima occur at or below the 10th scale size

FIGURE 1 | Wavelet ridgewalking algorithm identifies neuronal events with high fidelity. **(A)** Continuous wavelet transform amplitude color map for an example trace overlaid in white. The left *y*-axis is the wavelet scale, the right *y*-axis is the arbitrary units of trace fluorescence amplitude, and the *x*-axis is the recording frame. Color scales are continuous wavelet transform (CWT) values with warmer colors representing larger CWT amplitudes. **(B)** Example trace with ridges generated by ridgewalking algorithm overlaid in red. **(C)** Example trace with significant (kept) ridges overlaid.

are excluded as short noise. Noise parameters were explicitly calculated by noise modeling. MATLAB scripts implementing this algorithm are available upon request.

Noise Modeling

We generated 10,000 white noise time series using pseudorandom number generators and subjected those time series to the ridgewalking algorithm. We constructed empirical cumulative density functions for ridge length and scale of maximum CWT amplitude along a ridge (event size). We repeated this analysis at several different combinations of noise parameters. From these cumulative density functions, we were able to determine thresholds yielding the desired exclusion rate. For event size, we determined the 98% exclusion threshold to be the 10th scale size. Thus, ridges whose maximum CWT amplitude occurred at a scale less than or equal to 10 were removed as "short noise." This threshold was found to be stable across parameters, including series length. For ridge length, we

chose a threshold length corresponding to 98% exclusion rate (41 scales). These length cutoffs were found to be robust across wide regimes of noise variance and series lengths. Ridges were only kept if they met both the ridge length and max CWT amplitude scale size criteria.

Rasterization, Calculation of Activity Features, and Network Measures

For dF/F_0-processed ROIs, rasters were generated based on one of four different thresholding methods as described above. To create rasters from the wavelet ridgewalking-processed ROI traces, we found the location in time and scale of the maximum CWT amplitude along each significant ridge. Maximum CWT amplitudes were used as event centers. Following rasterization, event frequency was calculated as the number of unique events found in the trace. The amplitude was taken as the maximum dF/F_0 value within an event window. For the wavelet ridgewalking algorithm, each event had an associated F_0 value

calculated from the segment of the trace immediately preceding the event. Each event trace itself was transformed to dF/F_0 using its unique F_0, and the amplitude of an event was defined as the maximum dF/F_0 value within the event window. Duration was calculated as the length of a rastered segment. For network analyses, connectivity matrices for the ROIs in each video were generated using the phi coefficient between ROI rasters. The phi coefficient is essentially a correlation coefficient for binarized data series, such as rasters. Phi coefficient values were thresholded at 0.2. Any value below 0.2 was set to zero. Connection matrices were binarized to unweighted and undirected connection matrices. Network measures and distance matrices were calculated using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). Specifically, we calculated the global clustering coefficient, assortativity, density, modularity, characteristic path length, and network efficiency for identified networks. The global clustering coefficient represents how common clustering is within a network where a "cluster" is formed around a node when many of the nodes connected to it are also connected to each other. Assortativity captures the tendency of network nodes to be connected to other nodes with similar numbers of connections. In an assortative network, nodes with similar numbers of connections preferentially connect to each other, while in a disassortative network, nodes preferentially connect to others with dissimilar numbers of connections. Density is a measure of the number of possible connections that are found to actually exist between nodes; thus, in a very dense network, most nodes may be mutually connected to each other. Modularity captures the tendency of a network to be formed of "modules" or subsets of densely interconnected nodes with more sparse connections between these subsets. Characteristic path length is the typical number of connections required to link one node to another in a network. Network efficiency is the inverse of the characteristic path length, reflecting the efficiency of information transfer throughout a network along the extant connections or paths between nodes.

Simulated Dataset Generation

Pseudoneuronal and pseudoastrocytic events of randomized duration and frequency were embedded in a noise and data-derived bleaching model. Pseudoneuronal events were either peak values with exponential or linear decays. Peak amplitude was fixed for all events to recreate the mean signal-to-noise ratio (SNR) found in our biological datasets. Durations were varied by changing the decay constant of the exponential or increasing the base of the right triangle. Pseudoastrocytic events were isosceles triangles to simulate astrocytic transients with slower rise times and fixed peak values as for pseudoneuronal events. Event duration was varied by increasing the length of the triangle base. Noise was modeled as random white noise. To model fluorescence bleaching, a linear regression fit to each fluorescence trace was generated. Trace segments determined by the wavelet ridgewalking algorithm to be events were removed from the fit. We then found the mean slope of all fits to all traces which was used to model bleaching present in the simulated datasets. Experimental results also determined the SNR in simulated data. SNR was determined

based on events identified by our wavelet ridgewalking algorithm. Each trace background, defined as the portions of the trace not indicated as part of an event, had the bleaching removed by regression. The amplitude of each event was divided by the background noise standard deviation of its trace to determine an SNR for each event. The mean of these SNRs (SNR = 69.9 ± 4.7) was used for modeling. All synthetic time series were of identical length to our recordings (3,001 frames). We varied the frequency of events from 1 to 100 and the duration of events from 1 to 500 frames. All codes are available upon request.

Statistical Analysis

Analyses were done in Excel 2016 or GraphPad Prism 8. All data are reported as mean ± standard error of the mean and statistical significance threshold was set at $p < 0.05$.

RESULTS

Wavelet Analysis Outperforms Other Event Detection Methods

To quantify method performance, we calculated the probability that any ROI displayed significant events in a given image frame using each tested method. We reasoned that for spontaneous neuronal activity, this probability should be stable across time, and that deviations are likely caused by erroneous event identification. **Figure 2** illustrates the fraction of ROIs determined to be active at a given time point in the trace. Notice that the F_0 initial and F_0 minimal approaches produced activation probabilities that varied widely across time for the majority of applied thresholds (**Figures 2A,B**), while the F_0 smooth method produced more stable probabilities (**Figure 2C**). However, the performance of the F_0 smooth method suffered from a sharply defined degradation toward the end of the trace. Additionally, thresholding by the trace z-score failed to identify any significant events in F_0 smooth traces. End-of-time series degradation was likely due to declining number of points in the sliding window required by the F_0 smooth algorithm. The wavelet ridgewalking algorithm also found relatively stable event probabilities (**Figure 2D**) but did not show the end-of-time series degradation. Overall, the choice of thresholding criterion for the F_0 methods led to as much as a 7-fold difference in the fraction of signals considered to represent significant Ca^{2+} transients (**Figures 2A–C**). This caveat did not apply to the wavelet ridgewalking method where the event detection criterion has been empirically determined. The F_0 smooth method with a baseline SD threshold and the wavelet ridgewalking method produced event probabilities that converged for a large portion of the time series (**Figure 3A**). However, we found that while event probabilities were similar, the events captured within the rasters were not the same. For example, under conditions of high SNR (SNR = 27.6892), the wavelet ridgewalking algorithm and F_0 smooth found almost identical events, but the F_0 smooth method found those events to be of shorter duration and fragmented some events. In the example shown in **Figure 3**, the F_0 smooth method misses

FIGURE 2 | Different event identification methods find different event probabilities. **(A)** F_0 initial thresholding finds very different, temporally unstable firing probabilities for each thresholding scheme, as does F_0 minimal in panel **(B)**. **(C)** F_0 smooth thresholding identifies relatively stable firing probabilities, but the discovered rates vary widely with the thresholding scheme. **(D)** Wavelet ridgewalking algorithm also produces relatively stable firing probabilities.

a large event toward the end of the time series, beyond the range of optimal performance for this method. **Figure 3C** shows an example trace with a lower SNR (SNR = 8.9381) and events of shorter duration. In this example, wavelet analysis outperforms the F_0 smooth method, capturing all visible events, while the F_0 smooth method misses several. Under low SNR conditions, both the wavelet ridgewalking algorithm and F_0 smooth method identified what appeared to be false events which, however, were not the same for the two methods (**Figure 3B**, arrows). **Figure 3D** shows an extreme case: a trace that displayed substantial bleaching over the time series and without any recognizable events. For this trace, the wavelet ridgewalking algorithm correctly output an empty raster, while the F_0 smooth method identified events throughout the trace. This result demonstrated a particularly troubling quality for a method to be used in an automated unsupervised pipeline, as such a trace and the ROI associated with it would be kept in the dataset when it should clearly be removed. Therefore, the apparent similarity of performance between the wavelet ridgewalking algorithm and F_0 smooth method is misleading. It is likely that the tendency of the F_0 smooth method to falsely identify events is balancing its propensity for missing true events in such a way as to buoy its rasterization probability (under the baseline SD thresholding criterion) to approximately that of the wavelet ridgewalking algorithm.

Event Detection Methods Exhibit Differential Performance on Simulated Datasets

To more robustly quantify the fidelity with which different methods capture underlying Ca^{2+} transients, we generated model datasets with varying event numbers and widths (see section "Materials and Methods"). Since the number and timing of modeled events were predetermined, we could evaluate the exact rate with which the respective event detection methods were successful in capturing them. The results in **Figures 4**, **5** for right-triangle (pseudoneuronal) events and isosceles triangle (pseudoastrocytic) events, respectively, show two measures of performance across the four event detection methods: the plots on the left visualize the percent of the rasters that fall outside of the known event bounds (i.e., incorrectly identified events or "false positives"); the plots on the right show the percent of total event durations that are correctly captured by the respective method. The ideal method would exhibit low values on the left panels and high values on the right, indicating a low rate of both false positives as well as a high rate of event capture. In terms of false positives, we found relatively good performance for all methods. F_0 initial and F_0 minimal methods performed ideally (**Figures 4A$_i$,B$_i$**), while F_0 smooth and wavelet ridgewalking found some false events in the case of few actual

FIGURE 3 | Wavelet analysis outperforms F_0 smooth thresholding method for event identification. **(A)** Wavelet ridgewalking algorithm and F_0 smooth thresholding method produce similar raster probabilities. **(B)** In high signal-to-noise ratio (SNR) conditions, the two methods appear to identify almost identical events, but F_0 smooth finds those events to be shorter than wavelet ridgewalking. **(C)** In lower SNR conditions, the F_0 smooth thresholding method begins to fail, not identifying short spike events that wavelet ridgewalking readily detects. **(D)** F_0 smooth fails to identify "dead" ROIs whose traces contain no events, while wavelet analysis can be used to discard them. Red arrows indicate false events captured by wavelet ridgewalking, blue arrow indicates a false event captured by the F_0 smooth method.

events (**Figures 4C$_i$,D$_i$**). In terms of events captured, F_0 initial, F_0 minimal, and wavelet ridgewalking all outperformed F_0 smooth (**Figures 4A$_{ii}$–D$_{ii}$**). Interestingly, F_0 initial, F_0 minimal, and wavelet approaches were also noted to have well-defined ranges of parameters beyond which their performance sharply declined. For pseudoneuronal events, the F_0 initial method transition line between high and low performance follows a curve from 20 events/trace at the smallest event width to 5 events/trace at the largest event width (**Figure 4A$_{ii}$**). For F_0 minimal, which has a broader range of high performance, this transition follows a curve from 40 events/trace at the smallest event width to 7 events/trace at the largest (**Figure 4B$_{ii}$**). Wavelet ridgewalking began to fail when there were more than 40 events in a single trace of 3,001 frames. For right triangular events, wavelet ridgewalking exhibited high performance through the broadest ranges of event frequency and duration (**Figure 4D$_{ii}$**). We also modeled pseudoneuronal events with exponentially decaying amplitudes (**Figure 5**). The results followed the same patterns as observed for right triangular events (**Figure 4**), though all methods captured less of the simulated events (**Figures 5A$_{ii}$–D$_{ii}$**) and the wavelet ridgewalking algorithm in particular exhibited a marked reduction in performance for events of shorter duration (**Figure 5D$_{ii}$**). We observed generally similar patterns for pseudoastrocytic events modeled as isosceles

triangles (**Figure 6**). F_0 initial and F_0 minimal captured the smallest fraction of false positives (**Figures 6A$_i$,B$_i$**) relative to the F_0 smooth and wavelet approaches (**Figures 6C$_i$,D$_i$**). The methods varied in the degree to which they successfully captured events (**Figures 6A$_{ii}$–D$_{ii}$**), and the F_0 minimal method exhibited the highest overall performance for these data, with near-perfect event identification across the majority of the event frequency and duration space probed. These results support marked differences in fidelity of event detection across the examined methods.

Quantification Method Impacts the Interpretation of Neuronal Ca^{2+} Data

We noticed a general susceptibility of the dF/F_0-based event detection methods to two types of errors: event fragmentation and introduction of spurious events at time points that differed between methods (**Figures 7A–D**). It is tempting to ascribe the short fragments to noise and discard them; however, we found that many of these fragments occur within visible event boundaries (**Figures 7B,D**), indicating that F_0-based approaches would generally benefit from postprocessing to determine which events to exclude and which to merge together for a faithful reflection of neuronal activity. Additionally, the F_0 initial and

FIGURE 4 | Synthetic data for pseudoneuronal events with linear decay. The left-hand plots represent "false positives" or the fraction of rasters existing outside of known event bounds, and the right-hand plots represent the "events captured" or the fraction of total event bounds that are covered by the rasters. **(A$_{i,ii}$)** Results for F$_0$ initial thresholded at 2.5 baseline standard deviations. **(B$_{i,ii}$)** Results for F$_0$ minimal thresholded at 2.5 baseline standard deviations. **(C$_{i,ii}$)** Results for F$_0$ smooth thresholded at 2.5 baseline standard deviations. **(D$_{i,ii}$)** Results for wavelet ridgewalking algorithm.

F_0 minimum baseline methods appeared to introduce large-scale time structure to rasters produced from our dataset (cf. **Figures 2A,B**). Where we expected relatively stable event probability for spontaneous transients, these methods exhibit pronounced changes in event probability as time progresses. These large-scale effects may be due to the selected methods' inabilities to account for fluorescence bleaching present during recording. The variability between methods that we observe

could introduce substantial uncertainty into the interpretation of data related to frequency and duration of Ca²⁺ transients. **Figures 7E,F** show the distributions of event durations and event frequencies found by each of the tested methods. Most methods found substantially different mean frequencies and distributions. Even when the same definition of F_0 is used, the choice of threshold alters the distribution mean and shape. We note that across the board, dF/F_0 thresholding methods

FIGURE 5 | Synthetic data for pseudoneuronal events with exponential decay. The left-hand plots represent "false positives" or the fraction of rasters existing outside of known event bounds, and the right-hand plots represent the "events captured" or the fraction of total event bounds that are covered by the rasters. **(A$_{i,ii}$)** Results for F$_0$ initial thresholded at 2.5 baseline standard deviations. **(B$_{i,ii}$)** Results for F$_0$ minimal thresholded at 2.5 baseline standard deviations. **(C$_{i,ii}$)** Results for F$_0$ smooth thresholded at 2.5 baseline standard deviations. **(D$_{i,ii}$)** Results for wavelet ridgewalking algorithm.

found shorter mean event durations and higher mean event frequencies, both of which can be traced back to those methods' tendencies to fragment events and misidentify spurious segments as significant transients.

To further evaluate the extent to which the choice of event detection algorithm impacts the interpretation of experimental data, we compared the characteristic features of neuronal Ca^{2+}

transients: amplitude, frequency, and duration of events. The data were split into imaging sessions from animals trained to self-administer cocaine and their yoked saline controls. Characteristic features were examined at three population scales: individual events across all ROIs (saline: N_{events} = 5,454–287,100; cocaine: N_{events} = 9,353–408,411), mean values for each ROI across all slices (saline N_{ROI} = 1,025; cocaine N$_{ROI}$ = 1,556), and

FIGURE 6 | Synthetic data for pseudoastrocytic events. The left-hand plots represent "false positives" or the fraction of rasters existing outside of known event bounds, and the right-hand plots represent the "events captured" or the fraction of total event bounds that are covered by the rasters. **(A$_{i,ii}$)** Results for F$_0$ initial thresholded at 2.5 baseline standard deviations. **(B$_{i,ii}$)** Results for F$_0$ minimal thresholded at 2.5 baseline standard deviations. **(C$_{i,ii}$)** Results for F$_0$ smooth thresholded at 2.5 baseline standard deviations. **(D$_{i,ii}$)** Results for wavelet ridgewalking algorithm.

mean values for each slice across all animals (saline N_{slice} = 7; cocaine N_{slice} = 12). The summary of our findings is represented in **Table 1**. Significant differences between saline and cocaine populations were determined by Mann–Whitney U tests. Most methods found significant differences in event duration between saline and cocaine self-administration conditions at the level of pooled events. The significant differences found between

conditions declined at the next tier of analysis (pooled ROI means) and vanished altogether at the level of pooled slice means. Considering that the direction of changes among statistically significant findings remained the same, this effect was likely driven by diminishing sample size at progressively higher levels of analysis. Encouragingly, all methods with the exception of F_0 smooth under 2.5 times total trace standard deviation threshold

FIGURE 7 | df/F$_0$ thresholding methods fragment events with effects on Ca^{2+} activity measures. **(A)** A 1-min sample of neuronal trace, with example rasters beneath. All F$_0$ methods were thresholded at 2.5 times the standard deviation of the baseline. Note that, where wavelet finds continuous events, F$_0$ methods have a tendency to introduce fragments. Furthermore, in the case of F$_0$ minimal, most of this trace is deemed active, with little to no segregation between events. **(B)** A zoomed short segment of a neuronal trace with high SNR. Notice that even in high SNR conditions, fragmentation still occurs. **(C)** A 1-min sample of a low SNR neuronal trace with example rasters beneath. All F$_0$ methods were thresholded as in **(A)**. With low SNR, there is increased fragmentation by F$_0$ smooth and F$_0$ minimal methods. F$_0$ initial thresholding fails utterly for this trace. **(D)** A zoomed short segment of a neuronal trace with low SNR. F$_0$ initial fails to identify this event. **(E)** Distributions of mean event durations for presumed neurons from saline and cocaine self-administration groups as calculated by the various methods. The number for each method indicates thresholding criteria as follows: 1, baseline sd*2.5; 2, trace sd*2.5; 3, z-score baseline p = 0.05; 4, trace z-score p = 0.05. Note that the wavelet method finds events to be of longer duration due to lack of fragmentation effects. **(F)** Neuronal event frequency distributions. All F$_0$ thresholding methods identify several times more events than the wavelet ridgewalking method for the same traces.
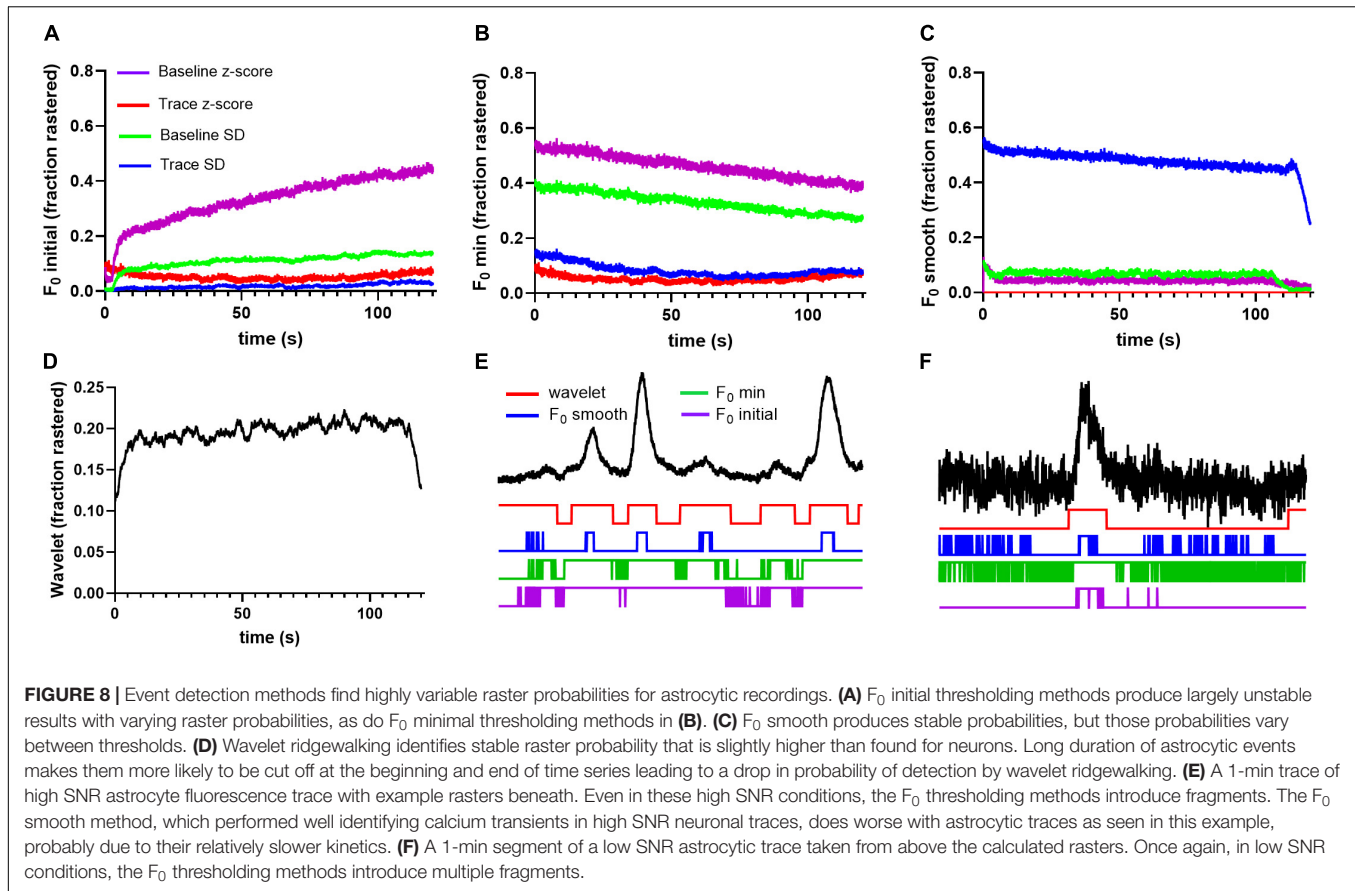
found the same qualitative effect of cocaine self-administration: increased event duration.

Some of the dF/F$_0$ thresholding methods identified significant decreases in event frequency associated with cocaine self-administration at the level of the ROI means (**Table 1**). In general, the dF/F$_0$ methods found dramatically more events than the wavelet method at both the cell population and slice mean levels of analysis, which, combined with the event duration results, supports our finding that dF/F$_0$ thresholding methods are capturing multiple short fragments of events compared with wavelet analysis' tendency to capture and represent events in a single segment.

Relative fluorescence amplitude is one of the most widely reported measures of Ca^{2+} activity. We found that most methods with the exception of the wavelet ridgewalking algorithm reported significant differences in event amplitude in the cocaine self-administration group at the level of individual events. However, the different methods disagreed on whether cocaine self-administration leads to increased or decreased event amplitude (**Table 1**). The results became more homogeneous at the level of the ROI means, with cocaine self-administration being found to produce higher-amplitude events, but the quantitative value of this difference varied substantially. For example, some negative mean amplitude values were reported

for the F$_0$ initial method with a baseline z-score threshold, which could be caused by failure to appropriately account for decreasing fluorescence due to bleaching. Additionally, the scale of analysis for the F$_0$ smooth method under a 2.5 baseline standard deviation method impacted the calculated amplitudes dramatically with the mean amplitude of all ROIs being much higher than the slice mean or events mean. This reflects a tendency to identify a subset of ROIs to have sparse but very high amplitude events compared to the general findings of this method. With few exceptions, the effects of cocaine self-administration appeared magnified when analyzed by the F$_0$ methods relative to the wavelet analysis at every scale.

We performed Kruskal–Wallis tests to determine which methods produced the most disparate results. These analyses were done only at the level of ROI means. For event durations, we found that out of 66 pairwise comparisons, 59 differed significantly from each other in both saline and cocaine self-administration conditions, and those methods that were not statistically different from each other were identical for the saline and cocaine datasets (**Supplementary Table 1**). Similar to event durations, event frequency data varied significantly across most methods. Out of eight methods that produced similar results in the saline group and seven methods that produced similar results in the cocaine group, five were shared between

TABLE 1 | Mean values of neuronal calcium event measures found using each of the tested methods.

	Wavelet		F_0 Smooth						F_0 Minimal								F_0 Initial							
			1		2		3		1		2		3		4		1		2		3		4	
	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c
Duration$_{All}$	61.23	77***	1.73	1.9****	10.16	9.8***	1.17	1.2	3.33	3.62	2.54	2.7***	3.29	3.4****	1.41	1.5****	3.27	3.7	2.12	2.3***	3.47	3.62	1.41	1.5***
Duration$_{ROI}$	42.41	54.6***	2.82	2.88	16.32	16.97	1.43	1.5***	4.72	7.06	5.57	6.82	5.26	7.07	3.29	4.26	3.13	5.49***	5.22	5.4**	5.19	6.78	3.29	4.26
Duration$_{Slice}$	60.55	77.13	1.75	2	10.07	9.95	1.17	1.22	3.3	3.81	2.54	2.73	3.28	3.74	1.41	1.54	3.26	3.75	2.16	2.45	3.49	3.76	1.41	1.54
Frequency$_{ROI}$	5.32	6.01	203.62	174****	175.59	167***	113	107****	203	186***	124.25	120.82	280.1	262.5*	79.11	75.03	28.32	35.72	14.69	19.44	226.76	220.03	79.11	75.03
Frequency$_{Slice}$	5.27	6.34	198.97	168.84	175.85	165.67	113.03	105.74	202.24	181.25	123.71	120.82	280.91	256.83	80.26	73.78	29.83	36.63	16.03	20.46	229.59	217.61	80.26	73.78
Amplitude$_{All}$	3.42	3.88	3.98	10****	4.83	11****	1.55	2.1****	3.15	10.6***	3.14	9.77****	0.42	4.72****	1.58	0.16	5.35	2.25****	6.1	3.91****	0.31	-0.4***	0.11	0.3*
Amplitude$_{ROI}$	2.22	2.41***	96.24	390****	4.29	9.79*	1.38	8.23	3.96	9.34**	5.71	13.6***	0.54	3.1	1.05	0.64	1.15	1.38*	2.89	2.81	0.12	-1.13	0.17	0.66
Amplitude$_{Slice}$	3.2	3.69	4.65	13.73	5.05	11.67	2.24	3.13	3.23	9.88**	3.23	10.2****	0.51	3.64	1.61	0.12	5.59	2.36	6.39	4.06	0.39	-0.45	0.13	0.34

*p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. Significant differences were determined by Mann–Whitney U test. Shaded cells with bolded values indicate significantly different saline/cocaine group pairs.

groups (**Supplementary Table 2**). The differences in amplitude of detected events were also significant across most of the methods and 12 comparisons that did not reach significance were shared between the saline and cocaine groups (**Supplementary Table 3**).

Quantification Method Impacts the Interpretation of Astrocytic Ca^{2+} Data

Our synthetic data (**Figures 4–6**) indicated that differences in kinetics of Ca^{2+} events may have an impact on the relative performance of event detection algorithms. Therefore, we compared the method performance when applied to the analysis of astrocytic Ca^{2+} transients which are substantially more heterogeneous than neuronal events and exhibit generally slower rise and decay times. As for the neuronal recordings, we began by analyzing the probability of astrocytic ROI activation found by each method based on events pooled across saline and cocaine conditions. The probability to detect an event varied as a function of time for F_0 initial and F_0 minimal approaches, although the magnitude of variation was less than exhibited for neurons (**Figures 8A,B**). Once again, the F_0 smooth and wavelet ridgewalking algorithm displayed the most stable performance across time. However, in the case of astrocytes, activation probabilities were dissimilar between these two methods, with the wavelet ridgewalking algorithm finding a slightly increased probability of activation for astrocytes relative to the F_0 smooth algorithm (**Figures 8C,D**). Event fragmentation and foreshortening of event duration caused by raster fragments were also obvious with the astrocyte data (**Figures 8E,F**).

To examine how these discrepancies would impact the interpretation of experimental manipulations, we compared frequency, amplitude, and duration of astrocytic Ca^{2+} transients in brain slices of animals trained to self-administer cocaine and saline controls. The wavelet ridgewalking method identified astrocytic Ca^{2+} transients as tending to be significantly longer than those observed in neuronal recordings. The dF/F_0 methods found mean astrocytic calcium transients to have durations that are of similar, larger, or smaller magnitude than those found for neuronal transients, depending on the method (**Table 2**). All methods identified significant differences in event duration between the saline and cocaine self-administration populations when all events from each condition were pooled, and most methods converged on the same qualitative conclusion: cocaine self-administration leads to shorter calcium transients in astrocytes. The single exception was the F_0 smooth method used with a trace SD threshold which found that cocaine exposure was associated with an increase in duration of astrocytic transients. All significant differences for event duration between the saline and cocaine groups vanished across all methods when the populations of ROI mean durations or slice mean durations were compared.

Three dF/F_0 thresholding methods identified differences in frequency between the total population of ROIs converging on an increase in event frequency following cocaine self-administration (**Table 2**). The other nine methods failed to detect any significant frequency differences. With regard to event amplitude, most dF/F_0 thresholding methods found significant differences

FIGURE 8 | Event detection methods find highly variable raster probabilities for astrocytic recordings. **(A)** F_0 initial thresholding methods produce largely unstable results with varying raster probabilities, as do F_0 minimal thresholding methods in **(B)**. **(C)** F_0 smooth produces stable probabilities, but those probabilities vary between thresholds. **(D)** Wavelet ridgewalking identifies stable raster probability that is slightly higher than found for neurons. Long duration of astrocytic events makes them more likely to be cut off at the beginning and end of time series leading to a drop in probability of detection by wavelet ridgewalking. **(E)** A 1-min trace of high SNR astrocyte fluorescence trace with example rasters beneath. Even in these high SNR conditions, the F_0 thresholding methods introduce fragments. The F_0 smooth method, which performed well identifying calcium transients in high SNR neuronal traces, does worse with astrocytic traces as seen in this example, probably due to their relatively slower kinetics. **(F)** A 1-min segment of a low SNR astrocytic trace taken from above the calculated rasters. Once again, in low SNR conditions, the F_0 thresholding methods introduce multiple fragments.

between saline and cocaine self-administration condition at the level of all events with some of the identified differences contradicting qualitatively (**Table 2**). Only two methods found significant differences in amplitude at the level of ROI mean values: wavelet ridgewalking, which reported a significant decrease in event amplitude in the cocaine group, and F_0 initial thresholded by baseline z-score, which reported a significant increase in amplitude.

Implications for the Interpretation of Network Structure

One of the advantages of Ca^{2+} imaging is that it allows researchers to probe relationships between multiple cells comprising a local network. We analyzed the impact of event detection methods on neuronal network structure by calculating the global clustering coefficient, assortativity, density, modularity, characteristic path length, and network efficiency for networks found in our slice recordings. **Table 3** contains the summary results for neuronal networks. While the majority of methods found no significant differences for any of these measures, the mean values varied widely across methods. For example, network modularity, a measure of connection density within versus outside of detected modules, was qualitatively different between F_0 minimal (baseline SD) and F_0 initial (baseline SD) with the former detecting a significant increase

and the latter reporting a decrease in modularity associated with cocaine self-administration.

We also carried out the analysis of network measures for astrocytic networks. The results are summarized in **Table 4**. As with neurons, the mean values of the network measures varied by method, and most methods did not find any significant differences between the two populations. Looking at the characteristic path length and global efficiency, network measures with the strongest trends toward significance, we found that the directionality of the relationship between the two populations is opposite for the wavelet ridgewalking algorithm and dF/F_0 thresholding methods: wavelet ridgewalking reported increased characteristic path length and thus decreased global network efficiency associated with cocaine self-administration, while all the F_0 minimal methods and the F_0 initial (z-test thresholds) methods found the opposite.

Comparing across methods, we found significant differences between dF/F_0 methods and wavelet ridgewalking as well as among the dF/F_0 methods themselves for characteristic path length. In the cocaine condition, all significant differences were found to be between the wavelet ridgewalking algorithm and dF/F_0 thresholding methods. For network efficiency, there were very few significant differences found in the saline condition, but the wavelet ridgewalking results for the cocaine self-administration population were found to be significantly different from all but one of the dF/F_0 methods analyzed

TABLE 2 | Mean values of astrocytic calcium event measures found using each of the tested methods.

	Wavelet		F_0 Smooth						F_0 Minimal								F_0 Initial							
			1		2		3		1		2		3		4		1		2		3		4	
	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C
Duration$_{All}$	**111.14**	**105.3****	**2.16**	**1.88******	**6.4**	**6.6******	**1.18**	**1.18******	**3.82**	**3.48******	**2.67**	**2.45*****	**2.96**	**2.87******	**1.48**	**1.45****	**6.36**	**6.07******	**3.07**	**2.98****	**3.33**	**3.12******	**1.48**	**1.45****
Duration$_{ROI}$	62.11	57.61	3.31	2.8	9.65	9.03	1.53	1.46	10.18	8.8	8.3	7.31	9.82	6.96	8.74	7.61	7.5	6	8.46	6.98	5.94	5.34	8.74	7.61
Duration$_{Slice}$	101.74	77.62	2.56	1.97	8.11	6.64	1.31	1.18	4.88	3.34	3.19	2.31	4.15	2.88	2.13	1.48	7.06	5.36	4.02	2.76	4.19	3.05	2.13	1.48
Frequency$_{ROI}$	5.85	5.66	75.77	86.84	183.41	186.13	104.77	107.51	**151.81**	**175******	79.56	83.55	**291**	**316******	102.01	101.19	**35.27**	**38.32****	16.3	17.1	212.92	224.8	102.01	101.19
Frequency$_{Slice}$	6.29	5.28	86.48	89.53	178.49	190.48	99.12	109.91	162.23	179.67	84.19	85.32	291.46	328.72	94.84	104.69	35.46	33.99	14.48	16.28	217.44	232.08	94.84	104.69
Amplitude$_{All}$	1.83	1.47	**37.79**	**38.9******	**17.21**	**20.7******	**5.13**	**-17******	**18.77**	**33.4******	**15.76**	**27.7******	**1.67**	**16.2******	**1.21**	**6.9******	**3.86**	**4.1******	**5.97**	**7.8******	**-1.59**	**-0.82******	-0.35	-0.13
Amplitude$_{ROI}$	**1.5**	**1.29***	526.96	884	17.5	20.55	19.03	-48.73	30.55	50.58	67.44	66.96	2.63	14.28	-1.1	7.59	4.91	4.47	6.1	5.52	**-0.89**	**-0.53***	-0.07	0.01
Amplitude$_{Slice}$	1.62	1.31	34.34	39.76	16.46	20.92	8.67	-16.33	18.46	41.24	15.47	34.2	3.24	17.73	1.14	7.2	3.87	4.98	5.76	7.4	**-1.67**	**-1.26**	-0.41	-0.25

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Significant differences were determined by Mann–Whitney U test. Shaded cells with bolded values indicate significantly different saline/cocaine group pairs.*

TABLE 3 | Neuron network measures found using each of the tested methods.

	Wavelet		F_0 Smooth						F_0 Minimal								F_0 Initial							
			1		2		3		1		2		3		4		1		2		3		4	
	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C
Clustering	0.263	0.281	0.552	0.43	0.716	0.718	0.22	0.291	0.833	0.823	0.839	0.832	0.834	0.806	0.586	0.638	0.276	0.397	0.303	0.417	0.882	0.856	0.586	0.638
Assortativity	0.48	0.36	−0.09	0.151	0.364	0.471	−0.124	−0.006	0.048	0.196	−0.039	0.153	0.076	0.18	−0.211	−0.146	0.379	0.432	−0.134	0.173	−0.103	−0.046	−0.211	−0.146
Density	0.036	0.06	0.265	0.174	0.244	0.255	0.359	0.308	0.541	0.455	0.582	0.488	0.499	0.417	0.164	0.195	0.051	0.094	0.331	0.306	0.684	0.639	0.164	0.195
Modularity	0.599	0.523	0.355	0.416	0.383	0.347	0.446	0.369	**0.09**	**0.16******	0.047	0.106	0.103	0.185	0.306	0.318	**0.739**	**0.58***	0.475	0.424	0.067	0.076	0.306	0.318
Path length	4.283	3.925	1.579	1.911	1.983	1.954	1.208	1.333	1.637	1.86	1.777	1.64	1.989	1.875	2.169	2.058	2.448	2.281	1.348	1.59	1.356	1.439	2.169	2.058
Efficiency	0.295	0.333	0.742	0.68	0.674	0.669	0.896	0.839	0.76	0.713	0.752	0.761	0.7	0.699	0.566	0.601	0.586	0.63	0.842	0.786	0.839	0.813	0.566	0.601

*$p < 0.05$; ****$p < 0.0001$. Significant differences were determined by Mann–Whitney U test. Shaded cells with bolded values indicate significantly different saline/cocaine group pairs.*

TABLE 4 | Astrocyte network measures found using each of the tested methods.

	Wavelet		F0 Smooth						F0 Minimal								F0 Initial							
			1		2		3		1		2		3		4		1		2		3		4	
	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C
Clustering	0.18	0.183	0.136	0.176	0.567	0.731**	0.154	0.12	0.541	0.591	0.565	0.566	0.527	0.614	0.321	0.377	0.323	0.311	0.308	0.239	0.585	0.727	0.321	0.377
Assortativity	0.2	0.265	0.151	0.44	0.41	0.541*	0.047	0.288	0.007	−0.056	0.004	−0.081	0.116	−0.024	0.208	0.156	0.143	0.315	0.378	0.112	−0.11	−0.201	0.207	0.156
Density	0.051	0.034	0.257	0.346	0.232	0.326	0.149	0.307	0.225	0.288	0.206	0.221	0.211	0.27	0.216	0.217	0.239	0.122*	0.114	0.1	0.324	0.468	0.216	0.217
Modularity	0.639	0.67	0.463	0.373	0.377	0.327	0.322	0.262	0.358	0.252	0.385	0.313	0.37	0.275	0.459	0.415	0.408	0.561	0.597	0.669	0.184	0.11	0.459	0.415
Path length	3.753	4.715	1.83	1.875	2.522	2.186	2.162	1.606	2.795	2.073	2.558	2.409	2.669	2.122	2.872	2.456	1.855	2.111	2.0108	1.996	1.999	1.644	2.872	2.456
Efficiency	0.405	0.302	0.747	0.783	0.552	0.624	0.805	0.821	0.52	0.62	0.544	0.567	0.543	0.624	0.559	0.581	0.693	0.638	0.708	0.68	0.617	0.722*	0.559	0.581

*p < 0.05; **p < 0.01. Significant differences were determined by Mann–Whitney U test. Shaded cells with bolded values indicate significantly different saline/cocaine group pairs.

(data not shown). Thus, we found that predominant differences for characteristic path length and global network efficiency were between the wavelet ridgewalking method and the dF/F_0 thresholding methods, which was the case for most of the network measures analyzed. For both neurons and astrocytes, the differences between wavelet ridgewalking and dF/F_0 thresholding method network measures were exacerbated in the cocaine self-administration population (data not shown).

DISCUSSION

Imaging Method

Widefield fluorescence microscopy allows the simultaneous visualization of fluorescence from multiple sources across a large area. Superior temporal resolution of widefield microscopy has benefits for the interpretation of relationships between dynamic events, such as intracellular Ca²⁺ transients. However, widefield imaging is vulnerable to contamination by signals from neighboring areas within the field of view as well as by signals originating both below and above the optical imaging plane. While the development of robust automatic image segmentation was not the goal of this study, we attempted to minimize signal contamination by using a modified version of the ABLE algorithm (Reynolds et al., 2017). In ABLE, the initial "seed" ROIs are evolved through algorithmic iterations based on correlated activity between neighboring pixels to produce final ROIs each of which contains all the pixels whose fluorescence activity was significantly similar. Fluorescence from a single pixel is allowed to contribute to any number of ROIs such that, ultimately, uncorrelated signal is excluded while distinct cells remain detectable even in the presence of structural overlap. However, it must also be borne in mind that when discrete domains are capable of generating isolated signals within the soma of a single cell, as is the case for astrocytic Ca²⁺ (Semyanov et al., 2020), the ability of the ABLE algorithm to identify signal overlaps is offset by failure to predict which of those overlaps belong to discrete cells and whether fluorescence spreading through the cell soma is identified as a single or multiple ROIs. This caveat, however, was not explicitly tested in our study. In the context of an automated analysis pipeline, this activity-based approach lends some confidence that only active pixels are included within the ROI outlines and is distinct from approaches that identify ROIs based on thresholds or watershed transforms applied to an average or maximal imaged field (Wong et al., 2010; Shen et al., 2018) or based on secondary fluorescent markers (Wardill et al., 2013). To further reduce potential signal contamination, we removed the local, rather than global, neuropil signal from each individual ROI as has been done in previous studies (Keemink et al., 2018; Pnevmatikakis, 2019; Soltanian-Zadeh et al., 2019). We did not seek to remove the commonly observed bleaching effects from the recordings, although various methods for doing so have been implemented by other groups (Balkenius et al., 2015; Shkryl, 2020). Our results indicate that wavelet ridgewalking outperformed all tested dF/F_0 thresholding methods in terms of faithfully identifying events within our *ex vivo* recording data, even in the presence of bleaching.

We observed considerable variability in SNR in our recordings, both within and between the imaged fields. The SNRs of our data were in many cases below SNRs achieved with single-photon or multiphoton confocal systems (Weisenburger et al., 2017; Song, 2019). Our results highlight the sensitivity of the various event detection methods to variability in SNR (e.g., **Figures 3**, **7**). In general, we found that the performance of the wavelet ridgewalking algorithm was on par with or better than the dF/F_0 approaches at high SNRs. At lower SNRs, the dF/F_0 thresholding methods tended to identify an increasing number of spurious events, while the performance of the wavelet ridgewalking algorithm remained stable.

dF/F$_0$ Thresholding

Our results indicate that the definition of F_0 as well as the choice of event detection threshold has major impacts on dF/F_0-based analyses. For example, event frequency differed by as much as 50-fold across detection methods and by as much as 15-fold between different choices of threshold applied with identical definitions of F_0 (**Figure 7** and **Table 1**). Similarly, a broad range of values were found across the analyses of event duration and amplitude. Amplitude-based approaches such as F_0 must take special care to ensure faithful detection of fluorescence fluctuations in part because all other measures are linked to amplitude by definition. In our analyses, we applied the tested methods to traces with unstable fluorescence baselines common in widefield imaging, and our recordings were taken under spontaneous firing conditions. It is possible that further pre- or postprocessing may cause the event identification methods tested to converge on a common set of events and thus produce more similar activity feature results. We did not undertake any preprocessing beyond local neuropil removal and undertook no method-specific postprocessing.

In studies that monitor Ca^{2+} activity triggered by external stimuli (e.g., electrical stimulation, drug application, and behavioral events), normalization of a recorded trace to the fluorescence signal immediately prior to stimulation (e.g., Ellefsen et al., 2014; Lock et al., 2015; Rahmati et al., 2016) presents a reasonable solution for amplitude-based detection, but this cannot be applied to spontaneous fluorescence since imaging starts at an undetermined point with regard to cell activity profile. Moreover, quantification of frequency and duration is less likely to benefit from stimulus-based approaches because they do not solve the problem of event fragmentation that our results indicate are common to F_0-based approaches.

In the presence of fluorophore bleaching, identifying a single section of trace to serve as a baseline F_0 is particularly problematic. Among the event detection methods tested, we found that F_0 minimal and F_0 initial methods had the greatest tendency to introduce large fluctuations in the probability of identifying an event at a given time throughout a trace. These results are unsurprising since both methods assume that fluorescence in a fixed segment of time represents the baseline activity level of the entire trace. Theoretically, F_0 minimal should be superior to F_0 initial as it aims to minimize the likelihood of categorizing active sections of the trace and high-frequency noise

as baseline. This was confirmed by simulated data (**Figures 4–6**). However, in practice, we found that F_0 minimal also tended to overidentify events and did not adequately account for bleaching or low-frequency oscillations (e.g., **Figure 7**). Wavelet ridgewalking analysis avoids these problems by identifying events independent of trace baselines. F_0 can thus be determined for each transient individually subsequent to event identification. The section of trace associated with each particular transient can then be transformed into dF/F_0 to provide a relative amplitude as we have done for our dataset.

In the context of an automated algorithm, there is in general no reason that an event amplitude threshold should be the sole determinant of significant events. The disparity in the performance of different detection methods on biological and simulated traces exemplify this issue. There is a clear distinction between high and low performance for these methods across the spectrum of event widths and frequencies in our simulated data. For example, F_0 initial and F_0 minimal both performed well with simulated data. Within their respective high-performance domains, both methods captured nearly all of the pseudoneuronal events successfully with minimal introduction of falsely identified events. However, despite the encouraging results with the simulated datasets (within their high-performance regimes), both the F_0 initial and F_0 minimal performed poorly with traces extracted from biological data. We believe this to be due to a combination of signal-to-noise variability in biological but not simulated data as well as the nature of "noise" in the two types of data. Thresholds that rely on standard deviation from the mean are predicted to perform well when the noise has a regular/stationary distribution such as white or Gaussian distributed noise. When the background noise is non-stationary and cannot be approximated by such a distribution, the methods may fail for all traces, as shown in our results based on biological data (e.g., **Figures 2**, **7**, **8**).

Wavelet Ridgewalking

Wavelet ridgewalking avoids both the problem of defining F_0 and the problem of threshold selection. In addition, we found that it was resilient to fluctuations in signal-to-noise ratio and bleaching effects. However, our results suggest that the performance of wavelet ridgewalking algorithms may be limited by the frequency of Ca^{2+} transients. For example, in our simulation of pseudoneuronal events, performance of the wavelet ridgewalking algorithm degraded abruptly when there were more than 40 events (of any duration) in the modeled time series (**Figure 4D$_{ii}$**). We have not explored the extent to which the problem persists across a range of time series lengths, as all simulated time series were constructed to be the same length as our recordings (3,001 frames). The possibility that wavelet ridgewalking algorithm-based event detection methods may fail when trying to image cells with very high frequencies of calcium transients should be kept in mind.

Technically, the wavelet ridgewalking algorithm employs thresholds as well. In our analyses, we removed ridges below a 98% confidence threshold as determined by white noise modeling. However, this determination is not simply based on amplitude but rather represents a 98% confidence that a

particular transient is not the product of white noise fluctuations regardless of transient shape, amplitude, or duration. Better or more thorough noise modeling may improve the performance of wavelet analysis further, as it is possible that noise artifacts are identified as significant events because they do not conform to the white noise assumptions.

Other Event Identification Methods

There are, of course, many other event detection tools available for Ca^{2+} imaging data, some of them also incorporating automatic ROI detection (e.g., CALIMA: Radstake et al., 2019; CaImAn: Giovannucci et al., 2019). Our intent was not to provide an exhaustive comparison between all available tools, but to highlight limitations associated with the use of classic dF/F$_0$ approaches, particularly relative to the wavelet ridgewalking algorithm as an alternative. We avoided library look-up methods (e.g., FluoroSNNAP, Patel et al., 2015) in our comparisons because they require foreknowledge of event shape, which may limit method versatility when applied to morphologically heterogeneous Ca^{2+} transients such as those displayed by astrocytes. Other nuanced methods, such as CNMF (constrained non-negative matrix factorization), are gaining popularity. CNMF (Pnevmatikakis et al., 2016) identifies ROIs, demixes overlapping ROI components, and denoises and deconvolves neuronal spiking activity from fluorescent traces in one fell swoop. CNMF is optimized for two-photon or light sheet microscopy and often has poor performance when used with more complicated background signals (Barbera et al., 2016), such as present in our recordings. CNMF-E is designed for microendoscopic data, which displays a more complicated profile of background signals (Zhou et al., 2018). Both CNMF and CNMF-E are optimized for the detection of neuronal signals, and it is unclear how well these methods could be adapted for analysis of astrocytic calcium transients. Nevertheless, CNMF-E should be appropriate for the analysis of widefield imaging data, and it may be interesting to evaluate its performance relative to wavelet ridgewalking in the future. Even in the absence of testing the breadth of extant event detection methods, our findings indicate that approaches relying on dF/F$_0$ thresholding must be subject to rigorous testing to justify the choice of baseline and event threshold.

Consequences for the Interpretation of Experimental Data

Our data indicate that the choice of an event detection algorithm has a substantial impact on the interpretation of experimental data both quantitatively and qualitatively. Quantitatively, the measured values of amplitude, frequency, and duration calculated from the same set of ROIs and traces varied significantly between methods even when similar baselines or detection thresholds were used (**Supplementary Tables 1–6**). These differences could lead to conflicting or even opposite conclusions. For example, in our dataset, cocaine self-administration was found to either suppress or increase the amplitude of neuronal Ca^{2+} transients depending on which approach was chosen to analyze the data. Examples of values trending in opposite direction between saline and cocaine groups could also be found in the analyses

of event duration and frequency, although among significant findings, all methods agreed that event duration increased and event frequency decreased in neurons following cocaine self-administration. The convergent findings of decreased frequency are consistent with the reports that action potential firing is reduced after cocaine self-administration (Peoples et al., 2007; Mu et al., 2010). Combined with an increased Ca^{2+} event amplitude, this finding could be interpreted to indicate an elevated number of action potential bursts in the NAc MSNs on the background of general neuronal hypoactivity. This interpretation would have to be modified if cocaine-induced decrease in amplitude was concluded based, for example, on the data from several F$_0$ initial methods (**Table 1**).

The results from imaging of astrocytes support our main conclusion: The effect of any experimental manipulation on Ca^{2+} fluorescence can only be ascertained with confidence if the analytical method used faithfully interprets the underlying signal. Automated analysis of astrocytic Ca^{2+} transients is further complicated not only by substantial variability in event shape but also by ambiguities with respect to spatial distribution within the soma of a single cell. Astrocytic Ca^{2+} transients are commonly observed in multiple discrete or overlapping microdomains within a single cell (Khakh and McCarthy, 2015; Agarwal et al., 2017; Verkhratsky and Nedergaard, 2018). Combined with the possibility that astrocytic Ca^{2+} may propagate across cellular subdomains, this poses significant analytic challenges especially since the mechanisms regulating spatial constraints on microdomain signals remain unknown. These considerations additionally pose the problem of image segmentation in stark relief. Activity-based methods of ROI identification, such as the one utilized in our study, may be effective in isolating microdomain signals but might not attribute them to single cells. The problem is amplified by extensive branching of astrocytic processes which may result in small "specks" of activity dispersed throughout the imaged field, a particular challenge in widefield microscopy experiments.

The ability to evaluate activity across multiple cells simultaneously afforded by Ca^{2+} imaging presents an opportunity to examine interactions within cellular networks. We analyzed functional connectivity between the ROIs in our recordings, meaning that we identified connections between ROIs whose rasters were significantly correlated in time. This generates networks of ROIs that are linked by similarities in their activity patterns. However, our approach does not establish effective (causal) or structural (physical) connections between the imaged cells though these are also important for network function. These could be calculated using a lagged phi coefficient, mutual information, or other measures of connectivity and further enhanced by differential labeling and multichannel fluorescence recording (Garofalo et al., 2009; Zhou et al., 2009; Rubinov and Sporns, 2010; Timme and Lapish, 2018). We acknowledge that the power of our slice network analyses may be limited by a relatively small number of slices represented in the saline and cocaine datasets. Ultimately, the disparities in network analysis results that we found here could be traced to the disparities in Ca^{2+} transient identification between the methods used, as

all networks were formed based on pairwise comparison of the rasters *via* the phi coefficient. Beyond discrepancies in the identification of specific events, the introduction of any non-random structure to rasters by a particular method such as exhibited in **Figures 2**, **8** will artificially increase correlation values between rasters, leading to differences in network measures. Effective Ca^{2+} transient identification is, therefore, the crux not only of fluorescent signal activity analyses but for any functional network analysis as well.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The animal study was reviewed and approved by the University of South Carolina and University of Kentucky Animal Care and Use Committees.

AUTHOR CONTRIBUTIONS

AN and PO designed the study. AN performed the data analysis, wrote all the codes, and wrote the manuscript. BO'D collected the data used in the analyses. PO wrote and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.620869/full#supplementary-material

REFERENCES

Agarwal, A., Wu, P. H., Hughes, E. G., Fukaya, M., Tischfield, M. A., Langseth, A. J., et al. (2017). Transient opening of the mitochondrial permeability transition pore induces microdomain calcium transients in astrocyte processes. *Neuron* 93, 587.e587–605.e587. doi: 10.1016/j.neuron.2016.12.034

Ali, F., and Kwan, A. C. (2020). Interpreting in vivo calcium signals from neuronal cell bodies, axons, and dendrites: a review. *Neurophotonics* 7:011402. doi: 10.1117/1.NPh.7.1.011402

Balkenius, A., Johansson, A. J., and Balkenius, C. (2015). Comparing analysis methods in functional calcium imaging of the insect brain. *PLoS One* 10:e0129614. doi: 10.1371/journal.pone.0129614

Barbera, G., Liang, B., Zhang, L., Gerfen, C. R., Culurciello, E., Chen, R., et al. (2016). Spatially compact neural clusters in the dorsal striatum encode locomotion relevant information. *Neuron* 92, 202–213. doi: 10.1016/j.neuron.2016.08.037

Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* 20, 353–364. doi: 10.1038/nn.4502

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575

Calipari, E. S., Bagot, R. C., Purushothaman, I., Davidson, T. J., Yorgason, J. T., Pena, C. J., et al. (2016). In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2726–2731. doi: 10.1073/pnas.1521238113

Du, P., Kibbe, W. A., and Lin, S. M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics* 22, 2059–2065. doi: 10.1093/bioinformatics/btl355

Ellefsen, K. L., Settle, B., Parker, I., and Smith, I. F. (2014). An algorithm for automated detection, localization and measurement of local calcium signals from camera-based imaging. *Cell Calcium* 56, 147–156. doi: 10.1016/j.ceca.2014.06.003

Garofalo, M., Nieus, T., Massobrio, P., and Martinoia, S. (2009). Evaluation of the performance of information theory-based methods and cross-correlation to estimate the functional connectivity in cortical networks. *PLoS One* 4:e6482. doi: 10.1371/journal.pone.0006482

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., et al. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. *Elife* 8:e38173. doi: 10.7554/eLife.38173

Guerra-Gomes, S., Sousa, N., Pinto, L., and Oliveira, J. F. (2017). Functional roles of astrocyte calcium elevations: from synapses to behavior. *Front. Cell Neurosci.* 11:427. doi: 10.3389/fncel.2017.00427

Kardos, J., Dobolyi, A., Szabo, Z., Simon, A., Lourmet, G., Palkovits, M., et al. (2019). Molecular plasticity of the nucleus accumbens revisited-astrocytic waves shall rise. *Mol. Neurobiol.* 56, 7950–7965. doi: 10.1007/s12035-019-1641-z

Keemink, S. W., Lowe, S. C., Pakan, J. M. P., Dylda, E., van Rossum, M. C. W., and Rochefort, N. L. (2018). FISSA: a neuropil decontamination toolbox for calcium imaging signals. *Sci. Rep.* 8:3493. doi: 10.1038/s41598-018-21640-2

Khakh, B. S., and McCarthy, K. D. (2015). Astrocyte calcium signaling: from observations to functions and the challenges therein. *Cold Spring Harb. Perspect. Biol.* 7:a020404. doi: 10.1101/cshperspect.a020404

Lilly, J. M. (2019). *jLab: A Data Analysis Package for Matlab. 1.6.6.* Available online at: http://www.jmlilly.net/jmlsoft.html (accessed June 14, 2019).

Lilly, J. M., and Olhede, S. C. (2012). Generalized morse wavelets as a superfamily of analytic wavelets. *IEEE Trans. Signal Process.* 60, 6036–6041.

Lock, J. T., Parker, I., and Smith, I. F. (2015). A comparison of fluorescent Ca(2)(+) indicators for imaging local Ca(2)(+) signals in cultured cells. *Cell Calcium* 58, 638–648. doi: 10.1016/j.ceca.2015.10.003

Mancini, R., van der Bijl, T., Bourgeois-Jaarsma, Q., Lasabuda, R., and Groffen, A. J. (2018). SICT: automated detection and supervised inspection of fast Ca(2)+ transients. *Sci. Rep.* 8:15523. doi: 10.1038/s41598-018-33847-4

Mu, P., Moyer, J. T., Ishikawa, M., Zhang, Y., Panksepp, J., Sorg, B. A., et al. (2010). Exposure to cocaine dynamically regulates the intrinsic membrane excitability of nucleus accumbens neurons. *J. Neurosci.* 30, 3689–3699. doi: 10.1523/JNEUROSCI.4063-09.2010

Patel, T. P., Man, K., Firestein, B. L., and Meaney, D. F. (2015). Automated quantification of neuronal networks and single-cell calcium dynamics using calcium imaging. *J. Neurosci. Methods* 243, 26–38. doi: 10.1016/j.jneumeth.2015.01.020

Peoples, L. L., Kravitz, A. V., Lynch, K. G., and Cavanaugh, D. J. (2007). Accumbal neurons that are activated during cocaine self-administration are spared from inhibitory effects of repeated cocaine self-administration. *Neuropsychopharmacology* 32, 1141–1158. doi: 10.1038/sj.npp.1301203

Pierce, R. C., and Wolf, M. E. (2013). Psychostimulant-induced neuroadaptations in nucleus accumbens AMPA receptor transmission. *Cold Spring Harb. Perspect. Med.* 3:a012021. doi: 10.1101/cshperspect.a012021

Pnevmatikakis, E. A. (2019). Analysis pipelines for calcium imaging data. *Curr. Opin. Neurobiol.* 55, 15–21. doi: 10.1016/j.conb.2018.11.004

Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* 89, 285–299. doi: 10.1016/j.neuron.2015.11.037

Prada, J., Sasi, M., Martin, C., Jablonka, S., Dandekar, T., and Blum, R. (2018). An open source tool for automatic spatiotemporal assessment of calcium transients and local 'signal-close-to-noise' activity in calcium imaging data. *PLoS Comput. Biol.* 14:e1006054. doi: 10.1371/journal.pcbi.1006054

Radstake, F. D. W., Raaijmakers, E. A. L., Luttge, R., Zinger, S., and Frimat, J. P. (2019). CALIMA: the semi-automated open-source calcium imaging analyzer. *Comput. Methods Programs Biomed.* 179:104991. doi: 10.1016/j.cmpb.2019.104991

Rahmati, V., Kirmse, K., Markovic, D., Holthoff, K., and Kiebel, S. J. (2016). Inferring neuronal dynamics from calcium imaging data using biophysical models and bayesian inference. *PLoS Comput. Biol.* 12:e1004736. doi: 10.1371/journal.pcbi.1004736

Reynolds, S., Abrahamsson, T., Schuck, R., Sjostrom, P. J., Schultz, S. R., and Dragotti, P. L. (2017). ABLE: an activity-based level set segmentation algorithm for two-photon calcium imaging data. *eNeuro* 4:ENEURO.0012-17.2017. doi: 10.1523/ENEURO.0012-17.2017

Romano, S. A., Perez-Schuster, V., Jouary, A., Boulanger-Weill, J., Candeo, A., Pietri, T., et al. (2017). An integrated calcium imaging processing toolbox for the analysis of neuronal population dynamics. *PLoS Comput. Biol.* 13:e1005526. doi: 10.1371/journal.pcbi.1005526

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Scofield, M. D. (2018). Exploring the role of astroglial glutamate release and association with synapses in neuronal function and behavior. *Biol. Psychiatry* 84, 778–786. doi: 10.1016/j.biopsych.2017.10.029

Scofield, M. D., Li, H., Siemsen, B. M., Healey, K. L., Tran, P. K., Woronoff, N., et al. (2016). Cocaine self-administration and extinction leads to reduced glial fibrillary acidic protein expression and morphometric features of astrocytes in the nucleus accumbens core. *Biol. Psychiatry* 80, 207–215. doi: 10.1016/j.biopsych.2015.12.022

Semyanov, A., Henneberger, C., and Agarwal, A. (2020). Making sense of astrocytic calcium signals - from acquisition to interpretation. *Nat. Rev. Neurosci.* 21, 551–564. doi: 10.1038/s41583-020-0361-8

Shen, S. P., Tseng, H. A., Hansen, K. R., Wu, R., Gritton, H. J., Si, J., et al. (2018). Automatic cell segmentation by adaptive thresholding (ACSAT) for large-scale calcium imaging datasets. *eNeuro* 5:ENEURO.0056-18.2018. doi: 10.1523/ENEURO.0056-18.2018

Shigetomi, E., Patel, S., and Khakh, B. S. (2016). Probing the complexities of astrocyte calcium signaling. *Trends Cell Biol.* 26, 300–312. doi: 10.1016/j.tcb.2016.01.003

Shkryl, V. M. (2020). Error correction due to background subtraction in ratiometric calcium measurements with CCD camera. *Heliyon* 6:e04180. doi: 10.1016/j.heliyon.2020.e04180

Soltanian-Zadeh, S., Sahingur, K., Blau, S., Gong, Y., and Farsiu, S. (2019). Fast and robust active neuron segmentation in two-photon calcium imaging using spatiotemporal deep learning. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8554–8563. doi: 10.1073/pnas.1812995116

Song, A. (2019). *Large-Scale Volumetric In-Vivo Two-Photon Calcium Imaging*. Ph. D. thesis, Princeton University, Princeton.

Sporns, O. (2018). Graph theory methods: applications in brain networks. *Dialogues Clin. Neurosci.* 20, 111–121.

Timme, N. M., and Lapish, C. (2018). A tutorial for information theory in neuroscience. *eNeuro* 5:ENEURO.0052-18.2018. doi: 10.1523/ENEURO.0052-18.2018

Verkhratsky, A., and Nedergaard, M. (2018). Physiology of astroglia. *Physiol. Rev.* 98, 239–389. doi: 10.1152/physrev.00042.2016

Wardill, T. J., Chen, T. W., Schreiter, E. R., Hasseman, J. P., Tsegaye, G., Fosque, B. F., et al. (2013). A neuron-based screening platform for optimizing genetically-encoded calcium indicators. *PLoS One* 8:e77728. doi: 10.1371/journal.pone.0077728

Weisenburger, S., Prevedel, R., and Vaziri, A. (2017). Quantitative evaluation of two-photon calcium imaging modalities for high-speed volumetric calcium imaging in scattering brain tissue. *bioRxiv [Preprint]*. doi: 10.1101/115659

Wong, L. C., Lu, B., Tan, K. W., and Fivaz, M. (2010). Fully-automated image processing software to analyze calcium traces in populations of single cells. *Cell Calcium* 48, 270–274. doi: 10.1016/j.ceca.2010.09.008

Zhang, X. F., Cooper, D. C., and White, F. J. (2002). Repeated cocaine treatment decreases whole-cell calcium current in rat nucleus accumbens neurons. *J. Pharmacol. Exp. Ther.* 301, 1119–1125. doi: 10.1124/jpet.301.3.1119

Zhou, D., Thompson, W. K., and Siegle, G. (2009). MATLAB toolbox for functional connectivity. *Neuroimage* 47, 1590–1607. doi: 10.1016/j.neuroimage.2009.05.089

Zhou, P., Resendez, S. L., Rodriguez-Romaguera, J., Jimenez, J. C., Neufeld, S. Q., Giovannucci, A., et al. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *Elife* 7:e28728. doi: 10.7554/eLife.28728