Calcium Imaging and the Curse of Negativity

Gilles Vanwalleghem*, Lena Constantin and Ethan K. Scott

Neural Circuits and Behavior Laboratory, Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia

The imaging of neuronal activity using calcium indicators has become a staple of modern neuroscience. However, without ground truths, there is a real risk of missing a significant portion of the real responses. Here, we show that a common assumption, the non-negativity of the neuronal responses as detected by calcium indicators, biases all levels of the frequently used analytical methods for these data. From the extraction of meaningful fluorescence changes to spike inference and the analysis of inferred spikes, each step risks missing real responses because of the assumption of non-negativity. We first show that negative deviations from baseline can exist in calcium imaging of neuronal activity. Then, we use simulated data to test three popular algorithms for image analysis, CalmAn, suite2p, and CellSort, finding that suite2p may be the best suited to large datasets. We also tested the spike inference algorithms included in CalmAn, suite2p, and Cellsort, as well as the dedicated inference algorithms MLspike and CASCADE, and found each to have limitations in dealing with inhibited neurons. Among these spike inference algorithms, FOOPSI, from CalmAn, performed the best on inhibited neurons, but even this algorithm inferred spurious spikes upon the return of the fluorescence signal to baseline. As such, new approaches will be needed before spikes can be sensitively and accurately inferred from calcium data in inhibited neurons. We further suggest avoiding data analysis approaches that, by assuming non-negativity, ignore inhibited responses. Instead, we suggest a first exploratory step, using k-means or PCA for example, to detect whether meaningful negative deviations are present. Taking these steps will ensure that inhibition, as well as excitation, is detected in calcium imaging datasets.

Keywords: calcium imaging, zebrafish, GCaMP, baseline fluorescence, data analysis, cerebellar circuitry, segmentation, spike inference

INTRODUCTION

The advent of Genetically Encoded Calcium Indicators (GECI) has transformed the field of neuroscience by allowing the imaging of activity across large populations of neurons (Nakai et al., 2001; Pologruto et al., 2004; Tian et al., 2009), and these methods are now being integrated in other fields of biology (Balaji et al., 2017; Shannon et al., 2017; Stevenson et al., 2020). A concurrent boom in microscopy techniques has allowed the rapid volumetric imaging of these populations, *in vivo*, in models including larval zebrafish (Wyart et al., 2009; Ahrens et al., 2012; Constantin et al., 2020; Vanwalleghem et al., 2020); flies (Wang et al., 2003; Suh et al., 2004), and rodents (Chen et al., 2012; Cai et al., 2016; Klioutchnikov et al., 2020). The vast datasets produced by this approach have driven the development of computational tools designed to extract and process

activity information from populations of neurons (Mukamel et al., 2009; Freeman et al., 2014; Pachitariu et al., 2017; Giovannucci et al., 2019; Stringer and Pachitariu, 2019). A common assumption in most of these modern computational tools is the non-negativity of the GECI's signal.

However, negative deviations from the fluorescence baselines have been observed, and assumptions of non-negativity may cause the omission or misinterpretation of GECI data from populations with such negative deviations (Galizia et al., 2010; Munch and Galizia, 2017; Favre-Bulle et al., 2018; Marquez-Legorreta et al., 2019; Zimmerman et al., 2019). With the slow rise and decay of GECI probes, on the order of hundreds of milliseconds, a long-term average firing rate above 1 Hz would be convolved as a high fluorescence baseline. Such tonic activity can be found in vestibular neurons, even at rest (Shimazu and Precht, 1965; Cullen and McCrea, 1993), and in the primary visual cortex neurons (Baddeley et al., 1997) among a great many others. Notably, inhibition of tonically active neurons has been observed with electrophysiology in vestibular neurons (Shimazu and Precht, 1966), Purkinje cells (Tian et al., 2013), and distributed across the brain in response to stimulus-driven decisions (Steinmetz et al., 2019). Such inhibition of tonic neurons, convolved by the slow GECI kernels, translate to negative deviations from baseline as we and others have observed (Favre-Bulle et al., 2018; Zimmerman et al., 2019).

Many tools for GECI analysis include methods for inferring the spike train that generated the observed fluorescence signal, and again most of these spike deconvolution algorithms assume non-negativity (Vogelstein et al., 2010; Pachitariu et al., 2018). For example, the spikefinder online challenge had this implicit assumption in the datasets offered to the community (Theis et al., 2016), and their best performing algorithms were based on convolutional neural networks. This supervised approach, however, would miss inhibited response profiles as they have been trained on datasets with no negative deviation in the fluorescent traces.

Finally, this non-negative assumption is built into popular approaches for interpreting the patterns of activity across populations of neurons. For example, Non-negative Matrix Factorization (NMF), not to be confused with CNMF that is used to extract fluorescent traces from the videos (Pnevmatikakis et al., 2016), has been used as a dimensionality reduction or clustering tool on the fluorescent traces of individual neurons (Freeman et al., 2014; Mu et al., 2019; Torigoe et al., 2019). The NMF approach, when applied on extracted neuronal activity data normalized with z-scoring or $\Delta F/F_0$, discards negative deviations from the baseline fluorescent signal. Another approach that we and others have used, the binarization of the data based on a threshold of activity to generate "bar codes" of the brain activity, also has an intrinsic non-negative assumption (Kubo et al., 2014; Naumann et al., 2016; Heap et al., 2018; Daviu et al., 2020; Etter et al., 2020). Other threshold-based approaches, or even data cleaning steps, run the risk of discarding all negative deviations from baseline, biasing conclusions drawn from the dataset to exclude inhibition from the modeled system.

In summary, we find this non-negative assumption at all levels of calcium imaging analysis, from the extraction of fluorescence

TABLE 1 | Parameters used for the simulation of calcium datasets.

Frame rate	Simulated volume	Radius nuclei	τ of GECI	Time points
5 Hz	$90 \times 90 \times 50$	5.9 m	1.5	1,000

traces to spike inference and analyses of populations' dynamics. Our goal here was to assess how the most popular calcium imaging analyses responded to negative deviations from the baseline, including whether or not each approach was sensitive to traces that were typical of inhibitory signals in neural networks. We also hope to spark a discussion on how these assumptions may have biased past studies, and may continue to bias future work using GECIs.

MATERIALS AND METHODS

The imaging data came from Favre-Bulle et al. (2018). Briefly, experiments were carried on 6 day post-fertilization (dpf) *nacre* mutant zebrafish (Danio rerio) larvae of the Tüpfel long fin strain carrying the transgene *elavl3:H2B-GCaMP6s* (Chen et al., 2013). The larvae were immobilized in 2% low melting point agarose (Progen Biosciences, Australia) and imaged using a diffuse digitally scanned light-sheet microscope (Taylor et al., 2018) while an optical trap was applied to the otolith to simulate acceleration (Favre-Bulle et al., 2017, 2018, 2019, 2020). All procedures were performed with approval from the University of Queensland Animal Welfare Unit in accordance with approval SBMS/378/16/ARC.

Artificial datasets were generated using the Neural Anatomy and Optical Microscopy simulation toolbox (Charles et al., 2019). We used the parameters for nuclear simulation with GCaMP6f default (see **Table 1**). To simulate inhibited neuronal responses, we randomly attributed a spike number from a Poisson distribution (λ of 1, based on; Baddeley et al., 1997) to each 200 ms time window of 10 to 20 percent of all simulated neurons (since ∼20% of neurons were inhibited when observed by Steinmetz et al., 2019). We then set a time frame of 0.2 to 5 s of inhibition (0 spikes), which was used to simulate the neuronal activity and generate movies that were processed with the tools below. For the mixed activity, we used a similar approach on the second half from the time series of 20% of the neurons in the simulated dataset.

For fluorescence extraction and spike inference, we benchmarked the most cited calcium imaging toolboxes: suite2p (suite2p, version 0.8.0, RRID:SCR_016434) (Pachitariu et al., 2017), CaImAn version 1.8 (Giovannucci et al., 2019), and the PCA/ICA approach CellSort (Mukamel et al., 2009). We did not simulate motion, and as such did not use the registration algorithms included in either suite2p or CaImAn. The parameters used for each of these approaches can be found in the github repository. Briefly, for suite2p we used the sourcery roi extraction, with a τ of 2, frame rate of 5, diameter of neurons (4,6), threshold scaling of 0.5 and a high pass of 50. For CaImAn, we used the CNMFe implementation which shows a better accuracy for background estimation than CNMF and should

avoid the risks of spurious negative deviations due to background subtraction (Zhou et al., 2018). The parameters for CaImAn were a τ of 2, frame rate of 5, a gSig of 4 and autoregressive order of 2. For the deep-learning spike inference method CASCADE, we used the Universal_5Hz_smoothing200ms pretrained model to infer the spikes on our dataset (Rupprecht et al., 2020). For MLspike, we used $\tau = 2$, dt = 0.2, pnonlin = [0.55 0.03] as suggested for GCaMP6f in Deneux et al. (2016).

For the analysis of the responses, we used MATLAB (R2018b, RRID: SCR_001622). $\Delta F/F_0$ was computed as in Akerboom et al. (2012). We used the non-negative matrix factorization function nnmf with 15 factors to reanalyze the data from Favre-Bulle et al. (2018). We used the correlation coefficients tools from MATLAB to compute the 2-dimensional correlation between the regions of interest (ROIs) and the ideal components, as well as between the traces or spikes and the ideal traces or spikes.

Statistical tests and plotting were done in Graphpad Prism (8.4.3, RRID:SCR_002798), using ordinary ANOVA with Tukey's multiple comparison test.

All the code used to generate and analyze the data can be found on github.com/Scott-Lab-QBI/Negative CalciumResponses.

RESULTS

Real Data

First, we reanalyzed a zebrafish dataset from our previous study of vestibular processing in which we identified inhibited responses in hundreds of neurons across the thalamus and cerebellum (Favre-Bulle et al., 2018). For the analysis presented here, we focus on two representative neurons from the cerebellum and hindbrain of a larval zebrafish (**Figure 1A**) as larvae were subjected to vestibular stimuli (**Figure 1B**, shaded areas). As seen in the raw data (**Figure 1B**, arrows), we observe negative deviations from baseline during stimulation (**Figure 1B**, magenta traces), as well as positive responses (**Figure 1B**, green).

Our first observation was that the classical $\Delta F/F_0$ approach with a moving baseline window (Akerboom et al., 2012) creates positive artifacts following negative deviation from baseline as seen in **Figure 1B** (arrows). These positive artifacts could be construed as actual responses by some approaches, since they peak at the same level as the actual responses (magenta traces with arrows vs. adjacent green traces in **Figure 1B**). In the $\Delta F/F_0$ trace, the results do not correlate as well for the (magenta) inhibited neuron ($\rho = 0.599$) when compared to the (green) activated neuron ($\rho = 0.979$). However, the z-scored trace was perfectly correlated to the raw trace ($\rho = 1$) for both neurons. As such, we recommend the use of z-score as a normalization of calcium traces, and we will use this normalization in the following analysis.

Beyond these artifacts, there was the concern that popular data analysis methods could miss inhibited response profiles. NMF has been used to analyze larval zebrafish calcium imaging data (Mu et al., 2019; Torigoe et al., 2019), so we tested this method on the same vestibular dataset from our group (Favre-Bulle et al., 2018). As can be seen (**Supplementary Figure 1**),

the NMF approach failed to identify responses resembling the inhibited cluster identified by k-means while the other (non-negative) clusters were found with a high correlation ($\rho = 0.92$, $\rho = 0.94$, respectively, **Figure 1C**).

The major limitation of this analysis was that it lacked a ground truth, making it impossible to judge whether outputs from apparently successful approaches actually reflected physiology. To solve this problem, we turned to simulated data for which we control the ground truth.

Simulated Data

We used the Neural Anatomy and Optical Microscopy (NAOMi) Simulation toolbox (Charles et al., 2019) to generate 10 datasets of simulated nuclear-targeted GCaMP6f data, as described in the Materials and Methods. Briefly, each dataset contained about 90 neurons, and for each, we randomly selected either 10 or 20% of the neurons to be inhibited. For each inhibited neuron, we simulated tonic firing, based on an observed Poisson distribution (Baddeley et al., 1997), which was randomly interrupted for 0.2 to 5 s to simulate inhibition (**Figure 2A**). We chose a random inhibition pattern as both suite2p and CaImAn depend on the correlation between pixels to generate the ROIs, and we wanted to make the inhibited neurons as easy to identify as possible, since most methods depend on local correlations to identify the neurons. The simulated spiking (**Figure 2A**) was then convolved with a GCaMP6f kernel to simulate neural activity (**Figure 2B**), which was then used to generate movies using NAOMi (**Figure 2C**). As most simulated neurons would be below the detection threshold, we used NAOMi to output the ideal responses corresponding to what would be detected with a microscope. While other algorithms occasionally identified additional neurons, the effect was marginal (<1%), so we decided to use the ideal responses as ground truth for the sake of simplicity (Charles et al., 2019).

Each fluorescence dataset was processed through suite2p (Pachitariu et al., 2017), CaImAn (Giovannucci et al., 2019), or CellSort (Mukamel et al., 2009), and the outputs for each approach were then analyzed in the same manner. We did not investigate whether the suite2p default classifier or the CaImAn components evaluation would exclude inhibited neurons, and as such, we kept all the ROIs either algorithm identified. The raster plots of the ten datasets (**Figure 3A**) show that CaImAn identifies the highest number of ROIs, with CellSort and suite2p identifying a similar number of ROIs (Ideal = 94.3 ± 4.7, CaImAn = 84.7 ± 14.4, CellSort = 55.5 ± 3.8, suite2p = 56.1 ± 4.7).

The segmentation of the simulated fluorescent movies gave good results for all three algorithms, with well-defined regions of interest that correlated well with the ideal ROIs (**Figures 3B,C**, $\rho_{CaImAn} = 0.74 ± 0.06$, $\rho_{CellSort} = 0.80 ± 0.02$, $\rho_{suite2p} = 0.79 ± 0.03$). We then correlated the ideal traces of activated or inhibited simulated neurons to the traces extracted by each algorithm, and for each dataset, we averaged the maximum correlations to each ideal trace (**Figure 3D**). All three algorithms succeeded in extracting the relevant traces for the activated neurons (**Figure 3D**, left, indicated by green bar), but CellSort and suite2p outperformed CaImAn for the inhibited traces
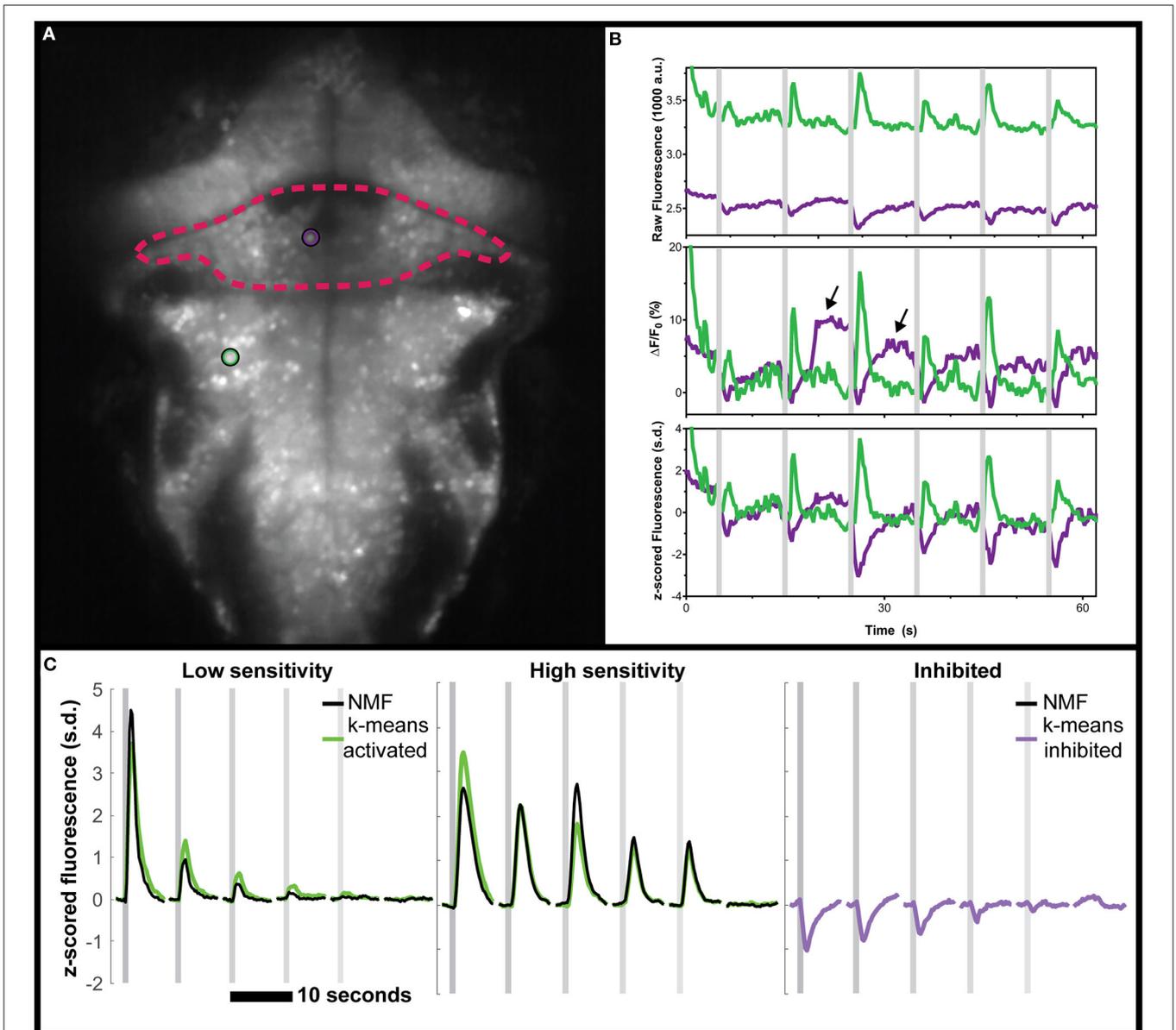
FIGURE 1 | Negative deviations from baseline in real data from the cerebellum of zebrafish, and performance of various analysis tools. **(A)** Mean fluorescence image of a 6 dpf zebrafish expressing nuclear-targeted GCaMP6s (Chen et al., 2013). The cerebellum is outlined in red, and an inhibited neuron is indicated with a magenta circle. The green circle indicates an activated neuron in the hindbrain. **(B)** Time traces of the raw (top), $\Delta F/F_0$ (middle), or z-scored (bottom) fluorescence for these two neurons, in their respective colors. Arrows indicate positive deviation artifacts resulting from the cessation of inhibition on the inhibited neuron. **(C)** Comparisons between the clusters identified using k-means (green for activated, magenta for inhibited) and those identified with NMF (black). No inhibited cluster was identified by NMF. Gray shaded areas indicate the time of vestibular stimulation (Favre-Bulle et al., 2018), with a progression from strong to weak stimuli across the stimulus train.

($\rho_{CaImAn} = 0.43 \pm 0.09$, $\rho_{CellSort} = 0.82 \pm 0.07$, $\rho_{suite2p} = 0.83 \pm 0.08$, **Figure 3D**, right, magenta).

To assess the proportion of true positives, we identified the ideal fluorescent trace to which each ROI's fluorescent trace best correlated. We only counted the unique ROIs that passed a 0.5 correlation cut-off, as all algorithms over-segment some of the sources in duplicated fluorescent traces (Charles et al., 2019). When comparing the proportions of identified ideal activated neurons, CellSort outperformed suite2p

slightly, followed by CaImAn (proportions of $0.38 \pm 0.05_{CaImAn}$, $0.58 \pm 0.06_{CellSort}$, and $0.54 \pm 0.04_{suite2p}$, **Figure 3E** left). For inhibited neurons, CellSort outperformed suite2p slightly again, but the divide with CaImAn grew (proportions of 0.34 ± 0.19 $_{CaImAn}$, 0.86 ± 0.10 $_{CellSort}$ and 0.82 ± 0.10 $_{suite2p}$, **Figure 3E**, right). All algorithms seemed insensitive to the ratio of inhibited neurons presented, as we saw no difference in those metrics between datasets with 10 vs. 20% inhibited neurons.

FIGURE 2 | Creating simulated calcium imaging datasets. **(A)** An example dataset of simulated activity, showing spike numbers for one neuron (green) activated and one (magenta) inhibited by a hypothetical stimulus (gray rectangles). **(B)** The spike trains are convolved with a GCaMP6f kernel and noise to generate fluorescence traces. **(C)** The simulated neuronal activity was used to create an artificial movie as captured by a microscope.

These results are lower than the results from Charles et al. (2019), who found that both CaImAn and suite2p outperformed CellSort (proportions of 0.71, 0.69 and 0.33, respectively). One possible explanation for the difference is that our use of nuclear-targeted GCaMP simulations, like our real datasets, may favor CellSort.

Spike Inference From Simulated Calcium Traces

In theory, inferring the spike trains responsible for calcium traces is one way to improve the temporal resolution, as you get rid of the convolved GCaMP kernel, but the frame rate of acquisition often makes such deconvolution impractical and unreliable. Each of the above algorithms offers some form of spike inference (**Figure 4A**), and multiple other approaches have been proposed during an online challenge (Berens et al., 2018). CaImAn offers multiple options for spike inference, among which we selected their fast non-negative deconvolution (FOOPSI) method (Vogelstein et al., 2010). For suite2p, we used the Online Active Set method to Infer Spikes (OASIS) (Friedrich et al., 2017). We also tested a recent spike inference method based on deep learning, CASCADE, which offers universal pre-trained models (Rupprecht et al., 2020) and a maximum likelihood approach to the probable spiked train, MLspike (Deneux et al., 2016).

Using this approach, we tested how accurate each spike detection algorithm was on our datasets. To avoid any confounding issues from the detection algorithm, we used the ideal calcium responses as the basis for the spike detection. Based on our results with the moving baseline of $\Delta F/F_0$ (**Figure 1B**), we also did not pre-process the data for the spike inference with suite2p, and used a global minimum to normalize for CASCADE

and MLspike. The CellSort deconvolution approach had limited success with both activated and inhibited neurons ($\rho_{CellSort} = 0.08 \pm 0.08$, and $\rho_{CellSort} = 0.003 \pm 0.013$ respectively, **Figure 4B**, **Supplementary Figure 2**). The more recent CaImAn and suite2p did well for the activated neurons, ($\rho_{CaImAn} = 0.53 \pm 0.04$, $\rho_{suite2p} = 0.55 \pm 0.03$), but CaImAn outperformed suite2p for inhibited neurons ($\rho_{CaImAn} = 0.60 \pm 0.05$, $\rho_{suite2p} = 0.37 \pm 0.03$). The universal model of CASCADE performed better than the rest on the activated neurons ($\rho_{CASCADE} = 0.66 \pm 0.03$), but worse on the inhibited neurons ($\rho_{CASCADE} = 0.03 \pm 0.03$). MLspike performed slightly worse than the above algorithms on the activated neurons ($\rho_{MLspike} = 0.53 \pm 0.03$), but was intermediate on the inhibited neurons ($\rho_{MLspike} = 0.28 \pm 0.02$).

Those performances were also tested in a scenario containing neurons with a mixture of activation and inhibition (**Supplementary Figure 3**). In that scenario there was little difference between the activated neurons and the mixed activity neurons with the different algorithms for ROI detections. The spike inference results were slightly worse across the board, but not as strongly as in **Figure 4**.

DISCUSSION

In this study, we show that the often implicit assumption of non-negativity for calcium imaging data can lead to missing real responses from inhibited neurons. Current approaches run the risk of missing a significant fraction of responses at every step of the analysis pipeline, including cleaning the data, processing, feature extraction, dimensionality reduction, and clustering.

We have shown these negative deviations exist in real data from zebrafish, as we previously observed (**Figure 1**; Favre-Bulle

FIGURE 3 | Various analyses' performances on simulated data. **(A)** Raster plots of ideal responses from NAOMi, and extracted fluorescence traces from CalmAn, CellSort, and suite2p. All the fluorescent traces were z-scored from−3 to 6 s.d. White horizontal lines separate the individual datasets. **(B)** Segmentation of the regions of interest (ROIs) by each algorithm, as for the raster plots in **(A)**, for one representative dataset. **(C)** Quantification of the correlation between the ROIs identified by each of the three algorithms and the ideal ROIs. Symbol color indicate the percentage of inhibited neurons ($n = 5$ datasets with 10% inhibited neurons in blue, and $n = 5$ datasets with 20% inhibited neurons in yellow). **(D)** Average maximum correlations between the traces identified by each algorithm and the ideal responses for the activated neurons (left, green rectangle) and the inhibited neurons (right, magenta rectangle). **(E)** Fraction of the ideal responses identified with a correlation above 0.5 by the three algorithms for the activated neurons (left) and the inhibited neurons (right).

et al., 2018), and as observed in mice (Steinmetz et al., 2019) and flies (Galizia et al., 2010; Munch and Galizia, 2017).

Hyperpolarization is well-known to decrease GCaMP signals in cells that are partially active at resting potential and that can be further inactivated by hyperpolarization (Zhao et al., 2018). We speculate that the negative deflections that we observed in

GCaMP6 signals of our real dataset (**Figure 1**) are, based on the spatial location and activity of the ROIs, Purkinje cells. In zebrafish, Purkinje cells receive excitatory inputs from granule cells and climbing fibers, and inhibitory inputs from stellate interneurons. Their cell bodies are located within one of the most superficial layers of the cerebellum, and hence would

FIGURE 4 | Spike inference from simulated calcium traces. **(A)** We applied each of the five spike inference algorithm on simulated GECI fluorescence data generated from a synthetic ground truth. The inferred spikes were then compared to the synthetic ground truth using a correlation. We show example inferred spikes from each algorithm, as well as the synthetic ground truth and simulated fluorescence. **(B)** Correlation between the inferred spikes from the simulated calcium traces and the actual spikes for the activated neurons (left, green rectangle) and the inhibited neurons (right, magenta rectangle). Each datapoint represents the performance on one simulated dataset ($n = 5$ datasets with 10% inhibited neurons in blue, and $n = 5$ datasets with 20% inhibited neurons in yellow).

be the first cells to be optically sectioned from the dorsal orientation. Purkinje cells display bistable spontaneous activity, where they switch between the steady production of tonic or depolarising "up" spikes and short bursts of intermittent activity or hyperpolarizing "down" states (Sengupta and Thirumalai, 2015). Therefore, the negative defections in GCaMP6 signals that we observed in our real dataset in **Figure 1** could be tonically-active Purkinje cells in the superficial layers of cerebellum that have toggled to their climbing-fiber induced bursting or "down" state (Engbers et al., 2013), which would produce negative voltages.

We have demonstrated that a moving baseline, such as for $\Delta F/F_0$, may create artifacts in inhibited neurons, which may lead to the generation of spurious positive signals. Finally, inhibited responses, when normalized, can also be lost when using NMF or thresholding approaches to analyze and visualize the data (**Figure 1C**). Even pixel-wise NMF approaches such as Thunder could miss inhibited responses, if they include a preprocessing step such as $\Delta F/F_0$ (Freeman et al., 2014). It would be interesting to revisit the data from studies that used these approaches (Mu et al., 2019; Torigoe et al., 2019) to see whether inhibited neurons are present in the datasets. We suggest that an initial unbiased step of data exploration of the dataset should be performed to ensure that no inhibited responses are present before pursuing steps including the above methods that assume non-negativity. Principal component analysis, or other dimensionality reduction tools, could be used to explore the data in the case of spontaneous activity or complex stimuli. Alternatively, for stimulus-driven activity, a correlation or linear regression should reveal any neuronal activity that deviates negatively from baseline.

By using simulated data (**Figure 2**), we tested how reliably CellSort, suite2p, and CaImAn could detect inhibited neurons in

a calcium imaging dataset. CellSort was the best algorithm in our specific analysis of nuclear-targeted GCaMP (**Figure 3**), which is at odds with other comparisons (Charles et al., 2019). However, both CaImAn and suite2p are better suited to larger datasets of thousands of neurons. Between these two approaches, suite2p outperformed CaImAn for the detection of activated responses both in terms of the fidelity of the extracted response (**Figure 3D**, mean difference of 0.056 and $p = 0.0006$) and the fraction of responses identified (**Figure 3E**, mean difference of 0.16 and $p < 0.0001$). For inhibited responses, suite2p largely outperformed CaImAn with more than twice the fraction of ideal inhibited responses recovered (mean difference = 0.47 and $p < 0.0001$). CellSort is a good option for smaller datasets as it requires an *a priori* estimate of the number of components (neurons) and does not perform as well at low SNR typical of endoscopic recordings (Resendez et al., 2016; Zhou et al., 2018). Among the currently available approaches, we therefore favor suite2p, or CellSort for smaller datasets, in order to recover the most inhibited responses from calcium imaging of neuronal activity.

As for the spike inference, the algorithm included with CellSort did poorly on both activated and inhibited neurons. MLspike was outperformed by suite2p and CaImAn performed similarly to one another with activated neurons, in line with published results (Pachitariu et al., 2018). However, for inhibited responses, suite2p's performance collapsed when using OASIS. CASCADE performed well on the activated neurons, but the lack of inhibited neurons in the training datasets mean it performed poorly when detecting our inhibited responses, as such the use of a more varied training dataset could improve its performance. Overall CaImAn, using FOOPSI, presents the best approach to infer spikes from inhibited neurons (Vogelstein et al., 2010). Several other methods of spike inference have been benchmarked

(Berens et al., 2018), and it would be interesting to benchmark these with simulated inhibited neurons. Finally, we want to point out that all the spike inference algorithms mistakenly inferred a strong spiking probably/rate when the inhibition ended (**Figure 4A**, bottom), this would need to be accounted for in any downstream analysis of these inferences.

We saw no significant differences between simulated datasets with 10 or 20% inhibited neurons in any of the above metrics, showing that the proportion of inhibited neurons should not affect the detection of the activated neurons.

Overall, we suggest that the PCA/ICA approach, such as implemented in CellSort should be favored when dealing with smaller datasets, where the number of ROIs can be estimated before processing, and nuclear-targeted GECIs. For larger datasets however, we suggest using suite2p, which has worked well both with nuclear-targeted simulations in this study, and with a cytoplasmic GECI simulation (Charles et al., 2019). With regard to spike inference, the FOOPSI approach gave the best results, so we would favor this method when inferring spikes. In terms of data analysis, NMF or thresholding based on activity should be avoided before an unbiased analysis such as PCA, or k-means can be used to ensure the absence of relevant inhibited neurons.

Another way to address the issue of negative deviations is to change the calcium indicator used. For example, an inverse-response GECI (Zhao et al., 2018) has been specifically designed to easily visualize neuronal inhibition in flies, but then the activated neurons would be the ones deviating negatively from the baseline. A powerful alternative is the use of Genetically Encoded fluorescent Voltage Indicators (GEVIs), which directly report membrane potential (Akemann et al., 2010; Gong et al., 2015; Bando et al., 2019). Those GEVIs can be used to visualize action potentials with millisecond time resolution, their signal to noise ratio are constantly improving and they would offer an accurate measure of the actual spikes of the imaged neurons. However, the requirement of ∼kHz imaging speed precludes their use for volumetric or whole-brain imaging with the current technologies.

Finally, GECIs allow the genetic targeting of the calcium indicator to subtypes of neuronal cells, providing information on the expected firing rate and behavior of the neurons (Scott et al., 2007; Forster et al., 2017). Alternatively, imaged neurons can be identified *post-hoc* using fixation and labeling (Lovett-Barron et al., 2017), which can be used to choose an analysis method that would be more appropriate if one expects negative deviations.

In summary, we have shown that assumptions of non-negativity can lead to the omission of real and simulated inhibited responses, and can produce spurious positive signals during the analysis of neural calcium imaging datasets. We have tested three popular and readily available approaches for analyzing such data, and provide recommendations for the best approaches to use when analyzing calcium imaging data that may contain inhibited signals.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncir.2020.607391/full#supplementary-material

REFERENCES

Ahrens, M. B., Li, J. M., Orger, M. B., and Robson, D. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* 485, 471–477. doi: 10.1038/nature11057

Akemann, W., Mutoh, H., Perron, A., Rossier, J., and Knopfel, T. (2010). Imaging brain electric signals with genetically targeted voltage-sensitive fluorescent proteins. *Nat. Methods* 7, 643–649. doi: 10.1038/nmeth.1479

Akerboom, J., Chen, T. W., Wardill, T. J., and Tian, L. (2012). Optimization of a GCaMP calcium indicator for neural activity

imaging. *J. Neurosci.* 32, 13819–13840. doi: 10.1523/JNEUROSCI.2601-12.2012

Baddeley, R., Abbott, L. F., Booth, M. C., and Sengpiel, F. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. Biol. Sci.* 264, 1775–1783. doi: 10.1098/rspb.1997.0246

Balaji, R., Bielmeier, C., Harz, H., Bates, J., Stadler, C., Hildebrand, A., et al. (2017). Calcium spikes, waves and oscillations in a large, patterned epithelial tissue. *Sci. Rep.* 7:42786. doi: 10.1038/srep42786

Bando, Y., Sakamoto, M., Kim, S., Ayzenshtat, I., and Yuste, R. (2019). Comparative evaluation of genetically encoded voltage indicators. *Cell Rep.* 26, 802–813.e804. doi: 10.1016/j.celrep.2018.12.088

Berens, P., Freeman, J., Deneux, T., Chenkov, N., McColgan, T., Speiser, A., et al. (2018). Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *PLoS Comput. Biol.* 14:e1006157. doi: 10.1371/journal.pcbi.1006157

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534, 115–118. doi: 10.1038/nature17955

Charles, A. S., Song, A., Gauthier, J. L., and Pillow, J. W. (2019). Neural anatomy and optical microscopy (NAOMi) simulation for evaluating calcium imaging methods. *bioRxiv [Preprint]*. doi: 10.1101/726174

Chen, Q., Cichon, J., Wang, W., Qiu, L., and Lee, S. J. (2012). Imaging neural activity using Thy1-GCaMP transgenic mice. *Neuron* 76, 297–308. doi: 10.1016/j.neuron.2012.07.011

Chen, T. W., Wardill, T. J., Sun, Y., and Pulver, S. R. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300. doi: 10.1038/nature12354

Constantin, L., Poulsen, R. E., Scholz, L. A., and Favre-Bulle, I. (2020). Altered brain-wide auditory networks in a zebrafish model of fragile X syndrome. *BMC Biol.* 18:125. doi: 10.1186/s12915-020-00857-6

Cullen, K. E., and McCrea, R. A. (1993). Firing behavior of brain stem neurons during voluntary cancellation of the horizontal vestibuloocular reflex. I. Secondary vestibular neurons. *J. Neurophysiol.* 70, 828–843. doi: 10.1152/jn.1993.70.2.828

Daviu, N., Fuzesi, T., Rosenegger, D. G., and Rasiah, N. P. (2020). Paraventricular nucleus CRH neurons encode stress controllability and regulate defensive behavior selection. *Nat. Neurosci.* 23, 398–410. doi: 10.1038/s41593-020-0591-0

Deneux, T., Kaszas, A., Szalay, G., Katona, G., Lakner, T., Grinvald, A., et al. (2016). Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations *in vivo*. *Nat. Commun.* 7:12190. doi: 10.1038/ncomms12190

Engbers, J. D., Fernandez, F. R., and Turner, R. W. (2013). Bistability in purkinje neurons: ups and downs in cerebellar research. *Neural. Netw.* 47, 18–31. doi: 10.1016/j.neunet.2012.09.006

Etter, G., Manseau, F., and Williams, S. (2020). A probabilistic framework for decoding behavior from *in vivo* calcium imaging data. *Front. Neural. Circuits* 14:19. doi: 10.3389/fncir.2020.00019

Favre-Bulle, I. A., Stilgoe, A. B., Rubinsztein-Dunlop, H., and Scott, E. K. (2017). Optical trapping of otoliths drives vestibular behaviours in larval zebrafish. *Nat. Commun.* 8:630. doi: 10.1038/s41467-017-00713-2

Favre-Bulle, I. A., Stilgoe, A. B., Scott, E. K., and Rubinsztein-Dunlop, H. (2019). Optical trapping *in vivo*: theory, practice, and applications. *Nanophotonics* 8, 1023–1040. doi: 10.1515/nanoph-2019-0055

Favre-Bulle, I. A., Taylor, M. A., Marquez-Legorreta, E., Vanwalleghem, G., Poulsen, R., Rubinsztein-Dunlop, H., et al. (2020). Sound generation in zebrafish with bio-opto-acoustics (BOA). *bioRxiv [Preprint]*. doi: 10.1101/2020.06.09.143362

Favre-Bulle, I. A., Vanwalleghem, G., Taylor, M. A., Rubinsztein-Dunlop, H., and Scott, E. (2018). Cellular-resolution imaging of vestibular processing across the larval zebrafish brain. *Curr. Biol.* 28, 3711–3722.e3713. doi: 10.1016/j.cub.2018.09.060

Forster, D., Arnold-Ammer, I., Laurell, E., Barker, A. J., and Fernandes, A. (2017). Genetic targeting and anatomical registration of neuronal populations in the zebrafish brain with a new set of BAC transgenic tools. *Sci. Rep.* 7:5230. doi: 10.1038/s41598-017-04657-x

Freeman, J., Vladimirov, N., Kawashima, T., Mu, Y., and Sofroniew, N. J. (2014). Mapping brain activity at scale with cluster computing. *Nat. Methods* 11, 941–950. doi: 10.1038/nmeth.3041

Friedrich, J., Zhou, P., and Paninski, L. (2017). Fast online deconvolution of calcium imaging data. *PLoS Comput. Biol.* 13:e1005423. doi: 10.1371/journal.pcbi.1005423

Galizia, C. G., Munch, D., Strauch, M., Nissler, A., and Ma, S. (2010). Integrating heterogeneous odor response data into a common response model: a DoOR to the complete olfactome. *Chem. Senses.* 35, 551–563. doi: 10.1093/chemse/bjq042

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

Gong, Y., Huang, C., Li, J. Z., and Grewe, B. F. (2015). High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor. *Science* 350, 1361–1366. doi: 10.1126/science.aab0810

Heap, L. A. L., Vanwalleghem, G., Thompson, A. W., and Favre-Bulle, I. A. (2018). Luminance changes drive directional startle through a thalamic pathway. *Neuron* 99, 293–301.e294. doi: 10.1016/j.neuron.2018.06.013

Klioutchnikov, A., Wallace, D. J., Frosz, M. H., and Zeltner, R. (2020). Three-photon head-mounted microscope for imaging deep cortical layers in freely moving rats. *Nat. Methods* 17, 509–513. doi: 10.1038/s41592-020-0817-9

Kubo, F., Hablitzel, B., Dal Maschio, M., Driever, W., Baier, H., Arrenberg, A., et al. (2014). Functional architecture of an optic flow-responsive area that drives horizontal eye movements in zebrafish. *Neuron* 81, 1344–1359. doi: 10.1016/j.neuron.2014.02.043

Lovett-Barron, M., Andalman, A. S., Allen, W. E., and Vesuna, S. (2017). Ancestral circuits for the coordinated modulation of brain state. *Cell* 171, 1411–1423.e1417. doi: 10.1016/j.cell.2017.10.021

Marquez-Legorreta, E., Constantin, L., Piber, M., Favre-Bulle, I. A., and Taylor, M. (2019). Brain-wide visual habituation networks in wild type and & zebrafish. *bioRxiv [Preprint]*. doi: 10.1101/722074

Mu, Y., Bennett, D. V., Rubinov, M., Narayan, S., and Yang, C. (2019). Glia accumulate evidence that actions are futile and suppress unsuccessful behavior. *Cell* 178, 27–43.e19. doi: 10.1016/j.cell.2019.05.050

Mukamel, E. A., Nimmerjahn, A., and Schnitzer, M. J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* 63, 747–760. doi: 10.1016/j.neuron.2009.08.009

Munch, D., and Galizia, C. G. (2017). Take time: odor coding capacity across sensory neurons increases over time in Drosophila. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* 203, 959–972. doi: 10.1007/s00359-017-1209-1

Nakai, J., Ohkura, M., and Imoto, K. (2001). A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. *Nat. Biotechnol.* 19, 137–141. doi: 10.1038/84397

Naumann, E. A., Fitzgerald, J. E., Dunn, T. W., and Rihel, J. (2016). From whole-brain data to functional circuit models: the zebrafish optomotor response. *Cell* 167, 947–960.e920. doi: 10.1016/j.cell.2016.10.019

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., and Rossi, L. F. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv [Preprint]*. doi: 10.1101/061507

Pachitariu, M., Stringer, C., and Harris, K. D. (2018). Robustness of spike deconvolution for neuronal calcium imaging. *J. Neurosci.* 38, 7976–7985. doi: 10.1523/JNEUROSCI.3339-17.2018

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

Pologruto, T. A., Yasuda, R., and Svoboda, K. (2004). Monitoring neural activity and [Ca2+] with genetically encoded Ca2+ indicators. *J. Neurosci.* 24, 9572–9579. doi: 10.1523/JNEUROSCI.2854-04.2004

Resendez, S. L., Jennings, J. H., Ung, R. L., Namboodiri, V. M., Zhou, Z. C., Otis, J. M., et al. (2016). Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nat. Protoc.* 11, 566–597. doi: 10.1038/nprot.2016.021

Rupprecht, P., Carta, S., Hoffmann, A., Echizen, M., Kitamura, K., Helmchen, F., et al. (2020). A deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging data. *bioRxiv [Preprint]*. doi: 10.1101/2020.08.31.272450

Scott, E. K., Mason, L., Arrenberg, A. B., Ziv, L., Gosse, N. J., Xiao, T., et al. (2007). Targeting neural circuitry in zebrafish using GAL4 enhancer trapping. *Nat. Methods* 4, 323–326. doi: 10.1038/nmeth1033

Sengupta, M., and Thirumalai, V. (2015). AMPA receptor mediated synaptic excitation drives state-dependent bursting in purkinje neurons of zebrafish larvae. *Elife* 4:e09158. doi: 10.7554/eLife.09158.020

Shannon, E. K., Stevens, A., Edrington, W., Zhao, Y., Jayasinghe, A. K., Page-McCaw, A., et al. (2017). Multiple mechanisms drive calcium signal dynamics around laser-induced epithelial wounds. *Biophys. J.* 113, 1623–1635. doi: 10.1016/j.bpj.2017.07.022

Shimazu, H., and Precht, W. (1965). Tonic and kinetic responses of cat's vestibular neurons to horizontal angular acceleration. *J. Neurophysiol.* 28, 991–1013. doi: 10.1152/jn.1965.28.6.991

Shimazu, H., and Precht, W. (1966). Inhibition of central vestibular neurons from the contralateral labyrinth and its mediating pathway. *J. Neurophysiol.* 29, 467–492. doi: 10.1152/jn.1966.29.3.467

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., and Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576, 266–273. doi: 10.1038/s41586-019-1787-x

Stevenson, A. J., Vanwalleghem, G., Stewart, T. A., Condon, N. D., Lloyd-Lewis, B., Marino, N., et al. (2020). Multiscale activity imaging in the mammary gland reveals how oxytocin enables lactation. *Proc Natl Acad Sci USA* 117, 26822–26832. doi: 10.1101/657510

Stringer, C., and Pachitariu, M. (2019). Computational processing of neural recordings from calcium imaging data. *Curr. Opin. Neurobiol.* 55, 22–31. doi: 10.1016/j.conb.2018.11.005

Suh, G. S., Wong, A. M., Hergarden, A. C., Wang, J. W., Simon, A. F., Benzer, S., et al. (2004). A single population of olfactory sensory neurons mediates an innate avoidance behaviour in Drosophila. *Nature* 431, 854–859. doi: 10.1038/nature02980

Taylor, M. A., Vanwalleghem, G. C., Favre-Bulle, I. A., and Scott, E. K. (2018). Diffuse light-sheet microscopy for stripe-free calcium imaging of neural populations. *J. Biophoton.* 11:e201800088. doi: 10.1002/jbio.201800088

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Roman Roson, M., Baden, T., et al. (2016). Benchmarking spike rate inference in population calcium imaging. *Neuron* 90, 471–482. doi: 10.1016/j.neuron.2016.04.014

Tian, J., Tep, C., Zhu, M. X., and Yoon, S. O. (2013). Changes in Spontaneous firing patterns of cerebellar purkinje cells in p75 knockout mice. *Cerebellum* 12, 300–303. doi: 10.1007/s12311-012-0439-6

Tian, L., Hires, S. A., Mao, T., Huber, D., Chiappe, M. E., Chalasani, S. H., et al. (2009). Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* 6, 875–881. doi: 10.1038/nmeth.1398

Torigoe, M., Islam, T., Kakinuma, H., Fung, C. C. A., Isomura, T., Shimazaki, H., et al. (2019). Future state prediction errors guide active avoidance behavior by adult zebrafish. *bioRxiv [Preprint]*. doi: 10.2139/ssrn.3345551

Vanwalleghem, G., Schuster, K., Taylor, M. A., Favre-Bulle, I. A., and Scott, E. K. (2020). Brain-wide mapping of water flow perception in zebrafish. *J. Neurosci.* 40, 4130–4144. doi: 10.1523/JNEUROSCI.0049-20.2020

Vogelstein, J. T., Packer, A. M., Machado, T. A., Sippy, T., Babadi, B., Yuste, R., et al. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* 104, 3691–3704. doi: 10.1152/jn.01073.2009

Wang, J. W., Wong, A. M., Flores, J., Vosshall, L. B., and Axel, R. (2003). Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* 112, 271–282. doi: 10.1016/S0092-8674(03)00004-7

Wyart, C., Del Bene, F., Warp, E., Scott, E. K., Trauner, D., Baier, H., et al. (2009). Optogenetic dissection of a behavioural module in the vertebrate spinal cord. *Nature* 461, 407–410. doi: 10.1038/nature08323

Zhao, Y., Bushey, D., Zhao, Y., Schreiter, E. R., Harrison, D. J., Wong, A. M., et al. (2018). Inverse-response Ca(2+) indicators for optogenetic visualization of neuronal inhibition. *Sci. Rep.* 8:11758. doi: 10.1038/s41598-018-30080-x

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

Zimmerman, C. A., Huey, E. L., Ahn, J. S., Beutler, L. R., Tan, C. L., Kosar, S., et al. (2019). A gut-to-brain signal of fluid osmolarity controls thirst satiation. *Nature* 568, 98–102. doi: 10.1038/s41586-019-1066-x