A machine-learning tool to identify bistable states from calcium imaging data

Aalok Varma[1], Sathvik Udupa[2], Mohini Sengupta[1#], Prasanta Kumar Ghosh[2] and Vatsala Thirumalai[1*]

[1] National Centre for Biological Sciences,
Tata Institute of Fundamental Research,
Bellary Road, Bangalore 560065.

[2] Department of Electrical Engineering
Indian Institute of Science
Bangalore - 560 012.

[#]Current Address:
Washington University School of Medicine in St Louis
660 S Euclid Ave,
St. Louis, MO 63110.

*Correspondence: vatsala@ncbs.res.in

Acknowledgements:
The authors would like to thank the following sources of funding support: Wellcome Trust-DBT India Alliance Intermediate (VT; 500040/Z/09/Z) and Senior fellowships (VT; IA/S/17/2/503297), Department of Biotechnology (VT; BT/PR4983/MED/30/790/2012), Science and Engineering Research Board, Department of Science and Technology (VT; EMR/2015/000595), Department of Atomic Energy (VT) and Department of Science and Technology, Govt. of India (PKG). The authors would also like to thank members of the fishfolk lab at NCBS for their inputs at various stages of this project, and T.P. Jagadeesh for maintaining the fish facility. The authors would also like to thank Prof. Rohini Balakrishnan for initiating this collaboration.

Keywords:
Purkinje neuron, bistability, zebrafish, machine learning, convolutional neural network, recurrent neural networks, cerebellum.

Abstract

The use of calcium imaging to map the activation of neurons *in vivo* during behavioral tasks has resulted in advancing our understanding of how nervous systems encode sensory input and generate appropriate output. In almost all of these studies, calcium imaging is used to infer spike times or probabilities since action potentials activate voltage-gated calcium channels and increase intracellular calcium levels. However, it is well known that neurons not only fire action potentials but convey information via intrinsic dynamics such as by generating bistable membrane potential states. While a number of tools for spike inference have been developed and are currently being used, no tool exists for converting calcium imaging signals to maps of cellular state in bistable neurons.

Purkinje neurons (PNs), the GABA-ergic principal neurons of the cerebellum, exhibit membrane potential bistability, firing either tonically or in bursts. Several studies have implicated the role of a population code in cerebellar function, with bistability adding an extra layer of complexity to the code. In this manuscript we develop a tool, CaMLsort, which uses convolutional recurrent neural networks to classify calcium imaging traces as arising from either tonic or bursting cells. We validated the classifier using a number of different methods and we show that the tool performs well on simulated event rasters as well as real biological data that was previously not seen by the network. This tool offers a new way of analysing calcium imaging data from bistable neurons to understand how they participate in network computation and natural behaviors.

Introduction

Calcium imaging enables us to sample the activity of populations of neurons and even whole brains during animal behaviour (Ahrens et al., 2013; Grewe and Helmchen, 2009; Grienberger and Konnerth, 2012), providing us a window into the functioning of the nervous system. When a neuron fires an action potential, somatic calcium levels increase dramatically due to the opening of voltage-dependent calcium channels. Simultaneous electrophysiology and calcium imaging recordings of neuronal activity demonstrate the correspondence of somatic calcium imaging signals with action potential firing (Chen et al., 2013; Kerr et al., 2005; Kwan and Dan, 2012; Tian et al., 2009). Nevertheless, the calcium imaging signal is at best a delayed, low-pass filtered and non-linearly transformed proxy for neuronal activity. Non-linearities are introduced due to the dynamics of calcium release, and the kinetics of calcium binding and unbinding by the sensor molecules (Ali and Kwan, 2020). Yet, this technique is useful in cases where the neurons being studied have low baseline firing rate. However, this is not often the case and several neuronal classes, such as cerebellar Purkinje neurons and cortical interneurons are known to fire at very high spike rates spontaneously (Armstrong and Edgley, 1984; Markram et al., 2004). Further, many neuronal types have interesting intracellular dynamics apart from action potential firing that might strongly modulate calcium imaging signals (Lin et al., 2007; Moreaux and Laurent, 2007). Lastly, confounds may be introduced when more than one type of electrical event leads to somatic calcium increase (Lev-Ram et al., 1992; Ramirez and Stell, 2016).

Purkinje neurons (PNs) show all of the features mentioned above that make the interpretation of calcium signals in these neurons especially difficult. PNs are principal neurons of the cerebellum, integrate thousands of synaptic inputs and are indispensable for cerebellar function (Narayanan and Thirumalai, 2019). Therefore, studying PN population activity during motor behaviours is critical to deciphering cerebellar computation. PNs fire simple spikes and complex spikes - in mammals, climbing fiber (CF) input from the inferior olive leads to calcium entry within the dendrites and several spikelets that ride the depolarization wave (Kitamura and Häusser, 2011; Miyakawa et al., 1992); in fish, the olivary input causes an all-or-none AMPAR-mediated giant synaptic potential with concomitant calcium entry but without any spikelets (Sengupta and Thirumalai, 2015). While the CF input causes a relatively large calcium signal, bursts of simple spikes can also generate calcium signals via the activation of somatic voltage-dependent calcium channels, albeit smaller in amplitude (Knogler et al., 2019; Sengupta, 2015). This means that optically recorded calcium dF/F signals cannot be uniquely assigned to simple spikes or CF inputs.

In addition to the above, PNs exhibit membrane potential bistability - PNs can exist in one of two stable membrane potentials. At a depolarized state, they generate action potentials at a more or less regular rate and are said to be tonic. When

hyperpolarized, they generate bursts of action potentials (Sengupta and Thirumalai, 2015). Simple spikes in the bursting state are highly correlated with motor bouts whereas those in the tonic state are not (Sengupta, 2015). This implies that bistability has important ramifications for how motor bouts are represented across the entire PN population. We wondered if calcium imaging can be used to deduce the state of PNs. Since both CF inputs and simple spikes contribute to calcium signals, assigning optically recorded dF/F signals is not straight-forward and deduction of spike rate using existing spike inference algorithms is not possible.

Here, we first performed simultaneous electrophysiological recordings and calcium imaging to obtain ground truth data. Unsurprisingly, we found that none of the usually measured parameters such as peak amplitude, area under the curve etc., of the calcium signal uniquely correlated with cellular state. We therefore built a sorter using convolutional recurrent neural networks to identify cellular state from dF/F data. We have tested our sorter with data that was not part of the training set and validated that it performs at above 90% accuracy and with F1 scores greater than 0.8 in most cases. Our method demonstrates that calcium imaging datasets can be mined using such tools to glean greater insights into neuronal activity states.

Results

Zebrafish cerebellar Purkinje neurons are spontaneously active and exhibit membrane potential bistability

Purkinje neurons (PNs) are the principal neurons in the cerebellar circuit (Fig. 1A). In larval zebrafish, they are spontaneously active and exhibit two kinds of membrane potential events that can be distinguished solely based on their amplitude in both extracellular and intracellular recordings (Sengupta and Thirumalai, 2015) (Fig 1B and 1C). The smaller amplitude events are action potentials and are called simple spikes (Fig 1B-E, marked by cyan dots), while the larger amplitude events are all-or-none synaptic input from climbing fibers (Fig 1B-E, marked by magenta dots). Moreover, these neurons exhibit membrane potential bistability, i.e. they are found to be at one of two stable membrane potentials and fire either tonically or in bursts, depending on the membrane potential. Both these states can be distinguished visually from extracellular loose-patch recordings based on (i) the firing frequency of simple spikes and (ii) the presence of long intervals of inactivity in the burst mode (Fig 1B, left) which are not present in the tonic mode (Fig 1B, right). The same features can be used to distinguish between the two modes in whole-cell patch-clamp recordings, along with an extra feature - membrane potentials are more hyperpolarized in the bursting mode (Fig 1C, left) as opposed to the tonic mode (Fig 1C, right). Cells can spontaneously switch between these two states, but can also be taken from one state to the other by current injection - bursting cells become tonic

upon depolarizing current injection, whereas tonic cells fire in bursts when injected with hyperpolarizing currents.

Calcium imaging reports both simple spike bursts and individual CF inputs, independent of state

Given that PNs manifest two distinct activity patterns (tonic and bursting) and two types of electrical events (simple spike and CF events), we wondered if calcium imaging signals can reliably report these distinct patterns of activity. To do this, we performed simultaneous wide-field calcium imaging and targeted patch-clamp electrophysiology from zebrafish PNs. We used the carbonic anhydrase (Ca8) enhancer element and a minimal promoter (*cfos*) to drive expression of the genetically encoded calcium sensor GCaMP5G in PNs (Matsui et al., 2014). Embryos were microinjected at the 1-2 cell stage and those larvae that showed sparse labelling in PNs were selected for targeted patch-clamp electrophysiology and imaging (Fig 2A). We chose to perform extracellular, loose-patch recordings so as to be minimally invasive and to avoid the dialysis of cellular contents including GCaMP. We obtained such recordings from 15 cells in 12 larvae, out of which 9 cells were naturally in the tonic state, and 6 cells, in the bursting state.

Representative traces from this experiment for cells in the bursting and the tonic states are shown in Fig 2B and 2C, respectively. The first observation we made was that many of the calcium transients (marked with a hash symbol in Fig 2B and 2C) had an almost one-to-one correspondence with CF inputs (magenta lines in the rasters) in cells of both states. In tonic cells, most simple spikes are not reported as transients, as one would expect from cells that fire at high frequencies. However, in some cases, calcium transient peaks occurred in the absence of CF inputs, during times of simple spike firing (Fig 2C, see asterisk and cyan lines in raster). Similarly, in bursting cells, we found some calcium transients associated only with a burst of simple spikes (Fig 2B, asterisk). Naturally, then, one would expect to see calcium transients when CF inputs co-occurred with a burst of simple spikes, which is what we observed (marked with open circles in Fig 2B and 2C). Thus, regardless of the state of a cell, we found calcium transients that corresponded to individual CF inputs, bursts of simple spikes or a combination of both.

To quantify the correspondence between calcium imaging signals and simple spikes or CF events, we first measured the average calcium signal triggered with respect to individual simple spikes (Fig 2D(i)) or CF inputs (Fig 2D(ii)). We found that the calcium signal's decay is the dominant response when triggered to individual simple spikes. This may be attributed to the high frequency of simple spikes observed in these cells. When aligned to CF inputs, however, there is a modest increase in calcium response after the CF input, followed by a decay, which is the typical response of a calcium sensor. It is interesting to note here that the baseline calcium signal preceding either of the two events is roughly the same. These general

features of the event-triggered calcium signal remained the same when data was sorted by cellular state (Supp. Fig. 2.1)

While this method confirms that a calcium transient occurs after a CF input, suggesting a causal link between the CF input's arrival and the measured calcium influx into the cell, it is not useful in the case of simple spikes. Firstly, there are multiple simple spikes per frame - approximately 2-3 spikes per frame when imaged at 30Hz. This is further worsened by the timescales of GCaMP5G's rise and decay, which occurs on the order of 100s of milliseconds. Moreover, one cannot separate out the individual effects of simple spikes and CF inputs because the two events are interspersed and thus, when calculating the triggered average response for one event, there is bleed-through from the other.

As an alternate measure, we decided to look at the number of simple spikes and CF inputs relative to the peak of the calcium signal using peri-event time histograms (Fig 2E). GCaMP5G has a rise time constant of ~100ms, which means that the time it takes to reach a peak value from the baseline is ~267ms. This time point is marked in the histograms with a vertical dashed line. The y-axes of these histograms represent the z-score for each bin. We calculated the average z-score of bins 200-400ms before the GCaMP peak (which is the window in which the rise in the signal would have started, Fig 2E, yellow shaded area) and the average z-score of bins +/-100ms around the peak (which is roughly when the signal would start decaying, Fig 2E, grey shaded area). It is clear from these that there is a greater count of both simple spikes and CF inputs right before the GCaMP signal starts to rise, which suggests that the occurrence of both kinds of electrical events in a PN could lead to a calcium transient. We checked if this effect was state-dependent, and we find that it is not, since the histograms do not differ much between tonic and bursting cells (Supplementary Figure 2.2).

In summary, it is difficult to infer the underlying event train or even the cellular state of a Purkinje neuron directly from its calcium signal, as there is no one-to-one correspondence between calcium transients and their underlying electrophysiological events, as has also been reported previously (Knogler et al., 2019; Sengupta, 2015).

The calcium signal for a Purkinje neuron can be reconstructed from its electrophysiological recording

Though the calcium signal in PNs could arise from more than one electrical event, we hypothesised that there might be sufficient information in the calcium signal time series to allow us to determine whether it is tonically firing or bursting. One possible approach to this problem could be the application of machine learning algorithms to build a classifier that labels cells as tonic or bursting given a time series of calcium transients. For this, it was first necessary to generate a larger ground truth dataset comprising the electrical activity patterns and their corresponding calcium transients

for training the classifier. Mathematically, the calcium signal obtained by imaging can be thought of as the result of the convolution of spike trains with a calcium sensor kernel having a fast rise and a slow decay and described as a difference of exponentials. This transformation has been shown to work well in cells that are not spontaneously active, where every single action potential produces a large increase in intracellular calcium concentration (Rupprecht et al., 2021; Vogelstein et al., 2010; Yaksi and Friedrich, 2006). We hypothesised that a similar transformation could be used to reconstruct the calcium signal profiles of PNs from their electrophysiological recordings alone. Such reconstructed calcium imaging traces with their corresponding electrical recording traces could then be applied to build the classifier.

However, given that these neurons have not one, but two electrophysiological events, each contributing differently to the calcium signal obtained from imaging, we must assign each event a separate kernel (Fig 3A), which we assume to be independent of each other. This is a valid assumption, since simple spikes and CF events arise from distinct intrinsic mechanisms (Sengupta and Thirumalai, 2015). We used simultaneous imaging and electrophysiology recordings to find optimal kernels for simple spikes and CF events. For each PN recording, we obtained simple spike and CF event rasters. Each event raster was then independently convolved with its own exponential response kernel, and the resulting “event-specific signals” were added to produce the final reconstruction (Fig 3A). We then calculated the correlation coefficient and the mean squared error (MSE) between the reconstructed calcium signal time series and the experimentally obtained calcium imaging time series.

With this description in place, the problem then becomes one of finding optimal coefficients for each response kernel. This optimization was done by gradient descent to minimise the MSE between the ground truth and the reconstruction. The process was performed independently for each cell in the dataset. Then, these optimal kernel combinations were used to make reconstructions for other cells in the dataset, and the kernel that generalised best across all cells, as quantified by least total MSE and highest total cross-correlation between the ground truth and reconstruction (Supplementary Figure 3), was taken to be the best kernel for the next steps. Kernel numbers 7 and 8 were very close to each other in both scores (Supplementary Figure 3), so kernel 7 was chosen for further analysis.

Although this is a simple forward model with calcium dynamics assumed to be linear and time-invariant, we found that this method to approximate the calcium signal works well, no matter the state of the neuron (Fig 3B). The quality of the reconstructions has been quantified by measuring the correlation coefficient between the ground truth and the reconstruction (Fig 3C). To benchmark the correlation coefficient values obtained against those obtained simply by chance, we also calculated the correlation coefficient between the ground truth and a scrambled version of the reconstruction, as well as that between the ground truth and a

reconstruction generated after shuffling all the event times in the recording (Fig 3C). In both the latter cases, the correlation disappears, suggesting that the reconstructions generated by this forward model are, indeed, faithful reconstructions.

Given that only groups/bursts of simple spikes produce a calcium transient amplitude comparable to that generated by an individual CF input, one would expect a much smaller response kernel for simple spikes than for the CF input. Indeed, the optimal kernels found by this method show that the CF input response kernel is approximately 7.5 times larger than the simple spike response kernel (Fig 3D).

A state-labelled calcium signal dataset was generated by reconstructing the calcium signal for a compilation of electrophysiological traces

Having a method of converting any given PN electrophysiology trace to its corresponding calcium imaging trace allowed us to generate simulated calcium imaging traces for a compilation of all electrophysiological recordings acquired in our lab (Fig 4A). To do so, we first pooled all the electrophysiological recordings, both extracellular and whole-cell, made during the period of 2011-2020, for a total of N=173 recordings. Recordings of poor quality, i.e., those with interruptions, or noisy baselines, unusually broad/filtered events, or indistinguishable simple spikes and CF inputs (N=37) were first excluded from the dataset. Retained recordings then had a state assigned to them, including state switches, if there were any. This collection was split into two sets, one with those recordings which had only one state throughout (N=131), and another with those which had a state switch (N=5). Events were detected in all the recordings, and then the calcium signal reconstructed, to finally produce two labelled datasets - a state-labelled reconstructed calcium signal dataset and an independent dataset with switches. The recordings used to generate this dataset were of different durations with a median duration of 60 seconds. In order to get multiple samples per recording, we decided to split the traces into 10-second-long non-overlapping chunks, representative samples from which are shown in Fig 4B. This produced a dataset of state-labelled calcium imaging traces (n=928 traces; n=580 bursting, n=348 tonic) that is much larger than could have been obtained by simultaneous imaging and electrophysiology of individual cells. The entire complement of non-overlapping samples is shown as a heatmap in Fig 4C.

To test whether there were properties of these traces that could be used to unambiguously classify their source state, we compared various properties of the traces between the two classes - the number of peaks, average peak amplitude, area under the curve, mean, standard deviation and coefficient of variation (Fig 4D). We found that most of the properties of the traces between the two classes were largely similar. While several of these differences are statistically significant (Table 2), they have small effect sizes, and our aim was not to just find properties that are statistically distinguishable from each other but to be able to classify a trace as having come from either a tonically firing cell or a bursting cell. To do so, we first

performed a PCA on the dataset of trace features and found that they had large overlaps (Fig 4E). Thus, linear methods of separating the two classes weren't successful and we needed to consider more complex models to solve this classification problem.

Machine learning methods have successfully been used to solve such problems when the data points are not linearly separable (He et al., 2016). Thus, we turned to developing a machine learning model to identify the state of a cell from a calcium signal time series.

Training a convolutional recurrent neural network to distinguish state from calcium imaging data alone

To find good machine learning models to solve this classification problem, we performed 5-fold cross validation, which is used to ensure that models generalised well and do not overfit the training data (Fig 5A). On each fold, only 60% of the training traces were shown to the network in the training phase (Fig 5A, white boxes), the network optimised on 20% of data used as a validation set (Fig 5A, yellow boxes), and the generalizability of the trained network assessed using the remaining 20% (Fig 5A, blue boxes). Each of the 5 folds had a different 60-20-20 split, such that the training and cross-validation datasets had a balance of traces from both classes.

We tested several different machine learning models on this classification problem - Support Vector Machines (SVMs), Deep Neural Networks (DNN) and Convolutional Neural Networks (CNNs). The prediction accuracies and F1 scores (a metric that adjusts for class imbalances as well as whether a prediction was a true/false positive/negative, defined in the methods section) for the various models tested are shown in Table 3. Of these models, the 1D CNN model performed best across all folds. However, these are only on snapshots of 10s-long snippets of the calcium trace, and the networks had more erroneous performance when predicting on the entire time series (data not shown). To overcome this, we decided to use a CNN-LSTM model. LSTM (Long short-term memory) networks (Hochreiter and Schmidhuber, 1997) are a kind of recurrent neural network (RNN), which has the ability to learn features across an entire sequence and not just across neighbouring positions in a trace sequence, like CNNs do.

The architecture of the network is shown in Fig 5B. The input to the network is a trace sampled at 30Hz, which is some multiple of 10 seconds in duration. It is first remapped to values between 0 and 1 using min-max normalisation and then processed by a CNN, which acts as a local feature extractor that also down samples the input data, such that it produces 4 vector outputs at a 1Hz sampling frequency. These outputs are passed to a bidirectional LSTM, one time step at a time. A forward LSTM cell looks at trends in the forward direction (Fig 5B, blue dotted box and

arrow), while a reverse LSTM cell goes through the data in the reverse direction (Fig 5B, red dotted box and arrow). Each of the two cells produces a “hidden state” as its output. Given that the LSTM is a recurrent neural network, this output is fed back into the network along with the next appropriate time step from the CNN’s output, a process which is repeated 9 more times (for every 10 second trace). In essence, the LSTM is detecting trends across the entire sequence. The final hidden states from the forward and reverse LSTMs, produced after passing through the entire trace, are retained and concatenated. These outputs are combined by passing them through a Linear layer, which performs a matrix multiplication to produce an output that is 10 samples long. At this step, each sample represents the likelihood that the trace in question comes from a cell in the bursting mode, predicted at 1 second intervals. The likelihood of each time step having come from the tonic state can then be inferred (Fig 5B, grayscale boxes on the right). The state with the larger likelihood score is taken to be the final classification at each time step.

Five networks were produced, one for each fold of cross-validation. The median accuracy was 80.15% for the cross-validation phase and 82.02% for the test phase (Fig 5C, top). The median F1-score was 0.7302 for the cross-validation phase and 0.8071 for the test phase (Fig 5C, bottom). We looked at the posterior scores of the predicted class in both the cross-validation and test stages (Fig 5D). In almost all the cases, the posterior scores of a network are higher in the true positive and true negative cases as opposed to the false positive and false negative cases. Of all the network folds, Fold 2 (orange) seemed to be particularly promising, because it consistently had high posterior scores, and had much higher scores for true positives and negatives when compared to false positives and negatives. Put together, this means that not only are these networks being trained well and making accurate predictions, but also confident about the predictions they made.

To be sure that networks weren’t performing well just because of peculiarities in the dataset used for training and testing, we presented an independent set of calcium signal reconstructions to the network that were not used in any phase of the training. These reconstructions also had state switches in them (Fig 6A). Networks performed fairly well even in these cases, with median F-1 scores higher than 0.8 (Fig. 6B). Fold 2 had a median F1-score of 1 and its predictions are shown in Fig 6A. The regions highlighted in red are those where the prediction didn’t match the ground truth. What is immediately clear from it is that most of the cells’ states were correctly identified, except in those cases where the cell switched state mid-recording, which is underrepresented in our training dataset as well.

As a further challenge to the trained networks, we decided to generate artificial spike trains that represent extreme cases of the tonic or bursting class (Supplementary Figure 6), reconstruct the calcium signal for these spike trains and ask the network to classify these completely synthetic data. We found that most folds of the network performed this task well (Fig 6D). Fold 2, in particular, predicted the state of 20/20

traces correctly (Fig 6C), with only one brief misclassified region, marked in the pink shaded area in Fig 6C.

Neural networks trained only on simulated calcium signals correctly identify Purkinje neuron state from real data

Thus far, all the inputs and challenges to the network were made using calcium signal reconstructions. To test the applicability of these networks on data from experiments, we went back to our original simultaneous electrophysiology-imaging data and asked if the network could call the state of those cells correctly based on the dF/F traces alone.

We found that the networks called the state of 7/15 cells correctly throughout the entire trace duration. In two other cells, more than 50% of the entire trace duration was called correctly (Fig 6E). It is important to note, though, that those traces which were incorrectly identified also had posterior scores that were not only low, but also similar for both classes (Fig 6E bottommost trace).

This outcome is interesting, because the networks have only been exposed to simulated calcium signals generated using a first-order generative model in the training and test phases, and yet are able to reasonably assess the state from real data, which have several challenging features - variable sensor expression levels, shot noise and signal bleaching, to name a few.

Among the five optimal networks identified from the 5-fold cross-validation, network #2 performed the best across all the challenges presented to it (Fig 5C, Fig 5D, and Fig 6), and thus is what we have decided to retain for future use, and call it CaMLsort_v1 (Calcium imaging and Machine Learning based tool to sort intracellular state). This will be made available as a Python package on GitHub.

CaMLsort not only makes calls about the state of a neuron but also produces class posterior scores, which essentially show how confident it was for each call. Thus, with some subjectivity on the part of the experimenter in combining the calls with the confidence scores, the error rate may be further reduced.

Discussion

In this paper, we describe the development of CaMLsort, a machine learning-based tool to identify the bistable state of Purkinje neurons from calcium imaging data. We develop this network and train it only on a simulated calcium signal dataset, but find that it can be used on experimental data, too.

Bistability in Purkinje neuron populations

The ability of neurons to exist in two stable states and exhibit distinct firing patterns/ excitabilities in those states has been documented in invertebrate (Dickinson and Nagy, 1983; Lechner et al., 1996; Marder et al., 1996; Russell and Hartline, 1982) and vertebrate nervous systems (Cowan and Wilson, 1994; Hounsgaard et al., 1984; Loewenstein et al., 2005; Misgeld et al., 1986; Schwindt and Crill, 1980).

Among the many systems in which neuronal bistability has been studied, Purkinje neurons of the cerebellum are particularly interesting because they are spontaneously active. Thus, they switch between firing either tonically or in bursts (Sengupta and Thirumalai, 2015), and such behavior is not only unique to Purkinje neurons but has also been reported in many other cell types such as thalamic (Jahnsen and Llinás, 1984) and cortical neurons (McCormick et al., 1985).

Though it is appreciated that bistability increases the information processing capacity of neurons, such as by allowing gating of synaptic inputs (Thirumalai and Jha, 2022), its implication for network computation or behavioural output is less well understood. This is true of Purkinje neurons as well - while it is established that these neurons are bistable and are important for motor behaviours (Engbers et al., 2013; Popa et al., 2018; Sengupta and Thirumalai, 2015), the role of bistability in Purkinje neuron computation or motor behaviour output has thus far not been clear.

We propose that that having more than one stable state allows a cell to toggle in and out of participating in circuit computations. In the cerebellum, specifically, bistability may have significant consequences by modulating PN neuronal synchrony with other PNs. PNs make weak inhibitory synapses onto deep cerebellar nuclear (DCN) cells in mammals and the eurydendroid cells in fish, with a high degree of convergence (Bae et al., 2009; De Zeeuw et al., 2011; Harmon et al., 2020; Ikenaga et al., 2006; Person and Raman, 2012). DCN cell spike times could be effectively controlled via inhibition from several PNs firing synchronously (Person and Raman, 2012). Since PNs fire tonically when synaptically-isolated and CF inputs can trigger bursts (Sengupta and Thirumalai, 2015), the bursting state is a network-embedded state while the tonic state could be thought of as an isolated state. By toggling between these two states PNs can increase or decrease the levels of population synchrony and thereby affect DCN firing differentially. Mapping state distribution across the PN population during various tasks will lead us to an understanding of how PNs synchronize or desynchronize during these tasks and how that in turn shapes cerebellar output and motor behaviors.

Inferring long-term dynamics from calcium imaging data

Calcium imaging has been one of the core experimental techniques in neuroscience for more than three decades and has seen many major developments in sensor quality, microscopy and analytical algorithms and software (Chen et al., 2013; Göbel

and Helmchen, 2007; Peron et al., 2015; Renninger and Orger, 2013; Tian et al., 2009; Yang and Yuste, 2017). Yet, the technique is largely used to identify regions of the brain that respond to presented stimuli or correlate with behavioral output under well-defined experimental conditions. This is resultant from the fact that the predominant mode of transient calcium elevation occurs due to the opening of voltage-gated calcium channels during action potential firing (Borst and Helmchen, 1998; Göbel and Helmchen, 2007). Thus, these calcium transients are quite brief, typically lasting only a few milliseconds and report spiking activity of neuronal populations during sensory stimulation and/or motor output. However, it is clear that neurons also exhibit longer time scale dynamics such as bursting, which carries information for downstream neurons (Zeldenrust et al., 2018). As the currently available calcium sensors themselves are much slower compared to the calcium transients, it is obvious that the recorded dF/F signals must manifest signatures of such long-term dynamics. CaMLsort addresses this issue by looking at longer term signatures in the calcium imaging time series data to identify cellular state.

Broader Applications of CaMLsort

The development of CaMLsort has opened up the possibility of a lot of future research in understanding PN bistability and its role in the function of the cerebellum as a whole. We could now experimentally address, for instance, what causes a switch in a PN's state by doing a high-throughput screen for chemical agents that affect bistability. We know that while bistability is a cell-intrinsic property, state switches could be a result of modulation of network activity. Neuromodulators are key candidates for this, as they can alter cell-intrinsic properties as well as network activity, and could thereby affect the bistability of a neuron, as has been demonstrated in multiple experimental systems (Abbinanti et al., 2012; Conway et al., 1988; Hounsgaard et al., 1988; Lechner et al., 1996; Williams et al., 2002).

Another question we can now directly address is how the state of a cell changes over development. We hypothesise that the tonic state could be a developmentally early state, with the burst-like phenotype developing as they get integrated into the developing cerebellar circuit. By birthdating PNs using a photoconvertible protein like Kaede and imaging their activity with GCaMP, we can use CaMLsort to ask whether younger cells tend to be more tonic or vice versa. While this question could potentially be addressed using electrophysiology, CaMLsort allows us to perform this experiment in a more high-throughput and minimally invasive fashion.

We focused on calcium imaging and electrophysiology data obtained from Purkinje neurons of larval zebrafish where we have shown that the tool predicts cellular state with very high accuracy and reliability (Fig. 6). A natural question is whether prediction of cellular state is useful in cell types other than the Purkinje neuron. A number of cell types exhibit bistability including cortical neurons (Cowan and Wilson, 1994; Egorov et al., 2002), striatal neurons (Misgeld et al., 1986), thalamic neurons (Fuentealba et al., 2005) and spinal motoneurons (Hounsgaard et al., 1984). In all of

these cases, cellular state dictates how synaptic inputs are integrated. Therefore, predicting cellular state is essential to understanding circuit computation and here CaMLsort will be immensely useful. It is natural to wonder whether CaMLsort will be equally effective in predicting cellular state in different cell types and in different model organisms. While CaMLsort needs to be tested in other cell types and organisms, our classifier is built using kernels derived from a small ground truth dataset and then used to build a larger database of calcium traces from electrophysiological recordings. A similar approach can be used without much more additional work to create parallel electrophysiology-calcium imaging databases of any of the above-mentioned cell types. Taken together, the development of CaMLsort opens up new avenues of research into the functioning of not just the cerebellar circuit, but also other circuits with bistable neurons.

Limitations of CaMLsort

Having said the above, CaMLsort is not free of limitations. For instance, the generative model we use to train the neural network is a linear, time-invariant model and results may improve with the use of more accurate models, such as ones that estimate the true increase in intracellular calcium concentration resulting from simple spikes and CF inputs.

Similarly, our model is trained and tested on GCaMP5G, for reasons of resource availability and experimental feasibility. However, newer generations of the GCaMP family of sensors continue to be developed, the latest being the 8th generation, which have much faster dynamics and a higher signal-to-noise ratio than the GCaMP5G we have used. Whether our results depend on the variant of calcium sensor being used is unclear, but we do not expect it to make a significant difference, because CaMLsort performs a window-wise min-max normalisation prior to identifying state. Since the relative ratio of simple spike to CF input calcium influx shouldn't depend on the sensor used, we predict that CaMLsort will work well regardless of the sensor being used. Moreover, the advantage of the method we have described is that a small ground truth dataset might be all that is necessary to retrain CaMLsort for a new sensor.

Lastly, it could be that CaMLsort may not work independent of the imaging method used. However, this is also easily fixed, either by the acquisition of a small ground-truth dataset, or by comparing the traces obtained from imaging to those in the reconstructions dataset, or by denoising the data prior to prediction using CaMLsort.

Materials and Methods

<u>Experimental Animals</u>

Experiments were performed on Indian wild type (Ind WT) zebrafish (*Danio rerio*) and were approved by the institutional animal ethics committee (IAEC) and institutional biosafety committee (IBSC). Embryos were obtained by setting up an incross between adult Ind WT zebrafish raised and housed in a ZebTec multi linking system (Tecniplast, Italy) with a pH setting of 7.8 and conductivity of 1200µS. Larvae were kept in a MIR-154 incubator (Sanyo, Japan) with 14:10h light-dark cycle maintained at 28°C in E3 medium (composition in mM: 5 NaCl, 0.17 KCl, 0.33 $CaCl_2$, and 0.33 $MgSO_4$, pH 7.8) in standard 90mm Petri dishes (Tarsons, Kolkata, India). The medium was routinely replaced, and starting at 5 dpf, larvae were fed Zeigler Larval diet AP100 (<100 microns) (Pentair AES, FL, United States). Larvae have not yet undergone sex specification by these stages, so experiments performed were agnostic to the sex of the animal.

Transient transgenesis by microinjection

Thin-walled borosilicate capillaries (1.0mm OD, 0.56mm ID; Warner Instruments, Hamden, CT, United States) were pulled using a Flaming-Brown P-97 pipette puller (Sutter Instruments, Novato, CA, United States) to produce long and thin pipettes for microinjection. Indian WT embryos at the 1-2 cell stage were injected with ~10nL of a mixture of tagRFP-T:PC:GCaMP5G plasmid (Matsui et al., 2014) (50 ng/µL) and Tol2 transposase mRNA (50 ng/µl) (Urasaki et al., 2006), using a Picospritzer III (Parker Hannifin, NH, United States) pressure injection apparatus, and raised as described above. Larvae were screened at 5 dpf at an epifluorescence dissection microscope (Olympus SZX-16) and those showing mosaic labelling of Purkinje neurons were selected for experiments, which were performed at 7 dpf at room temperature (25-28°C).

Electrophysiology

Larvae were anaesthetised in 0.01% MS-222 (Sigma-Aldrich; Missouri, USA) and transferred to a custom-made acrylic recording chamber. Fine tungsten wire (California Fine Wire, CA, USA) was used to pin the larva onto a piece of Sylgard (Dow Corning, Midland, MI, United States) glued to the recording chamber. Two pins were put through the notochord at the level of the swim bladder and the tail, respectively. A third pin was then used to position the larval head dorsal-up. The MS222 was replaced by external solution (composition in mM: 134 NaCl, 2.9 KCl, 1.2 $MgCl_2$, 10 HEPES, 10 glucose, 2.1 $CaCl_2$, 0.01 D-tubocurarine; pH 7.8; 290 mOsm) and skin on the dorsal surface of the larva was carefully removed using fine no. 5 forceps (Fine Science Tools, Foster City, USA) to expose the brain.

A Flaming-Brown P-97 pipette puller (Sutter Instruments, Novato, CA, United States) was used to pull patch pipettes made of standard-walled borosilicate capillaries with a filament (1.5 mm OD; 0.86 mm ID; Warner Instruments, Hamden, CT, United States), pulled to a final tip diameter of 1-1.5 µm. For loose patch recordings, pipettes were backfilled with external solution, while for whole-cell recordings, pipettes were backfilled with K-gluconate internal solution (composition in mM: 115 K

gluconate, 15 KCl, 2 $MgCl_2$, 10 HEPES, 10 EGTA, 4 Mg-ATP; pH 7.2; 290 mOsm). Solutions were filtered using a 0.2µm filter (Millipore, Merck, Germany) before backfilling pipettes, which typically had resistances between 8 and 10 MΩ. Sulforhodamine (Sigma-Aldrich, St.Louis, MO, United States) was added to the patch internal solution at a final concentration of 5µg/mL to facilitate visualisation of cellular morphology post-recording.

Recordings were acquired with a Multiclamp 700B amplifier, Digidata 1440A digitizer and pCLAMP 10 software (Molecular Devices, Sunnyvale, CA, United States). Data were low pass filtered at 2 kHz and sampled at 20-50 kHz with a gain of 1 (whole-cell recordings) or between 20-1000 (loose patch recordings), so as to optimally use the dynamic range of the digitizer.

Simultaneous calcium imaging and electrophysiology

Indian WT larvae microinjected with the Tol2-tagRFP-T:PC:GCaMP5G construct (gift from Dr. Hideaki Matsui, Niigata University, Japan) and screened for sparse labelling were used for this experiment. Purkinje neurons expressing GCaMP5G were targeted for the recordings. Widefield imaging was performed using a 60x water-immersion objective (1.0 NA) at an Olympus BX61WI microscope. Excitation light was provided using a pE300-ultra (CoolLED, Andover, England). Imaging was performed using an EVOLV EM-CCD camera (Photometrics, Tucson, AZ, United States) at the rig, with an exposure of 30 ms, and an on-chip gain of 170. An ROI was drawn around the targeted cell, so that only an isolated cell was imaged. Imaging was synchronised with electrophysiology by interfacing with the Digidata 1440A.

Image analysis was performed in Fiji/ImageJ (Schindelin et al., 2012). In order to identify cell boundaries, the averaged time series was used as a reference and an ROI drawn around the cell. Bleach correction was performed using the Exponential Fitting Method in the Bleach Correction plugin in Fiji. The average raw pixel intensity values within the ROI for each time point was extracted. The baseline intensity, F0, was taken to be the 5th percentile of these values. dF/F was then calculated for each frame using the formula (F - F0)/F0, where F represents the fluorescence in that frame and F0 the baseline fluorescence. Following this calculation, a Gaussian filter was applied to the time series in MATLAB R2019b (Mathworks, Natick, MA), using its function '*smoothdata'*, with a window size of 15.

Spike detection and analysis

All electrophysiological analyses were done using custom scripts written in MATLAB, using functions defined in a publicly available repository (https://github.com/wagenadl/mbl-nsb-toolbox) to read electrophysiology data. For extracellular recordings, spikes were identified as events that showed large fluctuations relative to a 25ms rolling window baseline and the peak timing and amplitude were extracted. Events were sorted based on amplitude, with small

amplitude events being classified as simple spikes and large amplitude events as climbing fiber (CF) inputs, as per (Sengupta and Thirumalai, 2015). For intracellular recordings, peak detection algorithms inbuilt in MATLAB were used to identify spikes. Once again, events were sorted as simple spikes or CF inputs based on their amplitude relative to baseline.

Calcium imaging reconstruction from electrophysiology data

Simultaneous calcium imaging and extracellular electrophysiology data was used to infer optimal GCaMP5G kernels, which are of the general form **A1*exp(-t/□$_{decay}$) - A2*exp(-t/□$_{rise}$)**, for simple spikes and CF inputs of zebrafish PNs. Simple spike and CF input timings were first extracted from the electrophysiology data for a single cell. Simple spike timings were convolved with a simple spike kernel, with amplitudes A1s and A2s and CF input timings were convolved with a similar CF input kernel, with amplitudes A1c and A2c. The sum of both these convolutions was taken to be the putative calcium imaging reconstruction, downsampled to match the sampling rate of calcium imaging. The mean squared error (MSE) between the reconstruction and the true calcium imaging trace was calculated. The values for A1s, A2s, A1c and A2c were obtained by minimising the MSE between the true calcium imaging trace and the reconstructed trace using gradient descent. To ensure that local minima were avoided, the gradient descent algorithm was repeated with randomised starting points. The kernels with these coefficients were taken to be the optimal simple spike and CF input kernels, respectively, for that cell.

The ‘optimal’ kernel would be one that generalised best. This was tested by using the kernels from one cell to reconstruct calcium imaging for other cells for which ground truth data was available. As before, the MSE and also the Pearson correlation coefficient between the reconstructed trace and the true trace were calculated. The final ‘optimal’ kernel was obtained by taking the kernel which generalised the best, i.e., had minimum MSE and maximum correlation across all cells. To test whether the correlation was real or not, the Pearson correlation coefficient between a scrambled reconstructed trace and the true trace was also calculated, for reference.

This process of convolving simple spike and CF timings with their optimal GCaMP5G kernels and summating the resultant traces was then used to generate a calcium imaging trace from both extracellular and intracellular Purkinje cell recordings.

Building a state-labelled calcium imaging dataset

Intracellular and extracellular recordings from larval zebrafish Purkinje neurons acquired during the period 2011-2020 by various members of the lab were compiled. Good quality recordings were those in which the resting membrane potential of the cell was not below -80mV or not above -20mV, were uninterrupted gap-free recordings with no current injections and had two clearly distinguishable event types. Only those recordings which met the aforementioned criteria were retained. Besides

poor quality recordings, those recordings with an ambiguous, indeterminate state were also excluded. These recordings were renumbered randomly, to remove grouping of data collected by the same individual. Recordings were then classified as tonic or bursting. If there was a switch in state within the recording, the recordings were labelled as “switch”, with the appropriate state switch timing and kind (tonic-to-bursting, or bursting-to-tonic) marked. A summary of the recording information is in Table 1.

Here, data was split into two groups - one that contained traces only with a single state throughout the recording, and another that contained traces with switches in state. Once split, the spike detection protocol described above was applied to each group of recordings to detect the simple spikes and CF inputs. These event timings were then used to generate simulated calcium imaging traces as described above. Combining the reconstruction with the assigned state resulted in the state-labelled calcium imaging dataset.

Principal Components Analysis

Principal Components Analysis was performed using the PCA module that is part of the sklearn package in Python (Pedregosa et al., 2011). The input to the model was a vector with the trace features for each 10 second long trace in the reconstructions dataset (number of peaks, peak amplitude, area under the curve, mean, standard deviation and coefficient of variation). A 2-component PCA was performed, and the resulting fit (shown in Figure 4E) explained 88% of the variance in the data.

Training the CNN-LSTM model for classification

The dataset of reconstructed calcium imaging time series was first split into 5 folds to perform cross validation (Fig 5A). The experiments are repeated by considering each fold as a test set, with one of the folds being the validation set and remaining 3 folds to be training sets such that the data split for each experiment is as follows - 60% for training, 20% for validation, 20% for testing. This splitting was done at the level of entire traces, such that the training, testing and cross-validation sets all had a balanced representation of traces from both the tonic and bursting classes. We chose to do a 5-fold cross-validation, i.e. the entire training dataset is split into 5 parts, each containing 20% of the data. Four of the folds are used for training and the remaining one is used for cross-validation.

The architecture of the network is shown in Fig 5B. The input to the network is the entire trace sequence at 30Hz, normalised using min-max normalisation for every non-overlapping 10 second period, thus setting the minimum to 0, and maximum to 1. The first layer of processing is a CNN with kernel size 30 and stride size 30. The CNN downsamples the input trace from 30Hz to 1Hz, and produces 4 vector outputs, which act as the input to the LSTM. The LSTM is bidirectional, so it looks at the sequence in both the forward and reverse order to produce an output. The LSTM has 24 hidden layers, so it produces 48 numbers per time step (Fig 5B). The feature

dimensions are collapsed by a Linear layer by taking their weighted average. This provides the classification logits at each position, which is passed through a softmax layer. The network is optimised using a cross entropy loss and the Adam optimizer (Kingma and Ba, 2017). The output of a trained neural network, then, is the probability of each timestep (at 1 Hz) belonging to the bursting class, from which the probability of it belonging to the tonic class can be derived. The class with the higher probability is assigned to each time step.

Since the model acts on the entire trace sequence, the amount of independent data for training the model is significantly reduced. To overcome this issue, we used a data augmentation strategy, which allows the network to learn features on arbitrary length sequences. For a sequence of length N, we determine a random segment size n, with n between [k, N], where k is the minimum sequence length. We randomly determine a starting position x, with a constraint that $x < N - n$. In our case, k=300 samples, and n varies in multiples of k (i.e. n = 300, 600, 900…). In each iteration of the network pass during the training phase, this random sequential segment was taken as the input trace to the model and gradient descent was performed to update weights. This was done for all the traces in the training set, so that the network saw entire traces at least once and in a non-overlapping fashion during the 500 epochs of training. The training is stopped before the full 500 epochs, based on early stopping criteria - if the validation loss does not improve after few epochs. All models were trained using the PyTorch framework in Python 3 (Paszke et al., 2019).

For all evaluations - cross-validation, test and other independent validations, the network had to make predictions for the entire trace duration and not just the n-sample long chunks. Networks were evaluated using two metrics - classification accuracies and F1-scores. Classification accuracies are simply the fraction of predictions that were correct. F1-scores, on the other hand, are adjusted for class imbalances and account for the kind of errors made. By definition, an F1-score is the harmonic mean of the precision and recall metrics, where precision is the fraction of all predicted positives that were true positives, and recall is the fraction of all ground truth positives that were true positives.

Generating artificial Purkinje neuron spike trains

Synthetic spike trains of the tonic class were generated using a Poisson statistical model. The only parameter needed for this is an average event rate, λ, which we took to be a random number between 4-10 for simple spikes. Spike trains for bursting neurons do not follow Poisson statistics. Hence, we used an inverse sampling method to generate spike trains of the bursting class. For this, we used the compiled electrophysiology data to obtain the real cumulative distribution of interevent intervals for simple spikes across all bursting cells. A random number, y, is generated from the uniform distribution and it is used to find a value x from the reference cumulative interspike interval (ISI) distribution such that $P(ISI <= x) = y$.

Thus, by generating events sequentially with a given next interevent interval, we obtained artificial simple spike trains for the bursting class.

CF input event trains were generated using a Poisson statistics model with an average event rate between 0.2-1. The distributions of interevent intervals in the synthetic datasets were compared against the real distributions using histograms and QQ plots (Supplementary Figure 6C-F). Independent reviewers were asked to classify unlabelled spike trains as real or artificial and managed to do so with an accuracy of ~42% on average.

Figures, Tables and Figure Legends

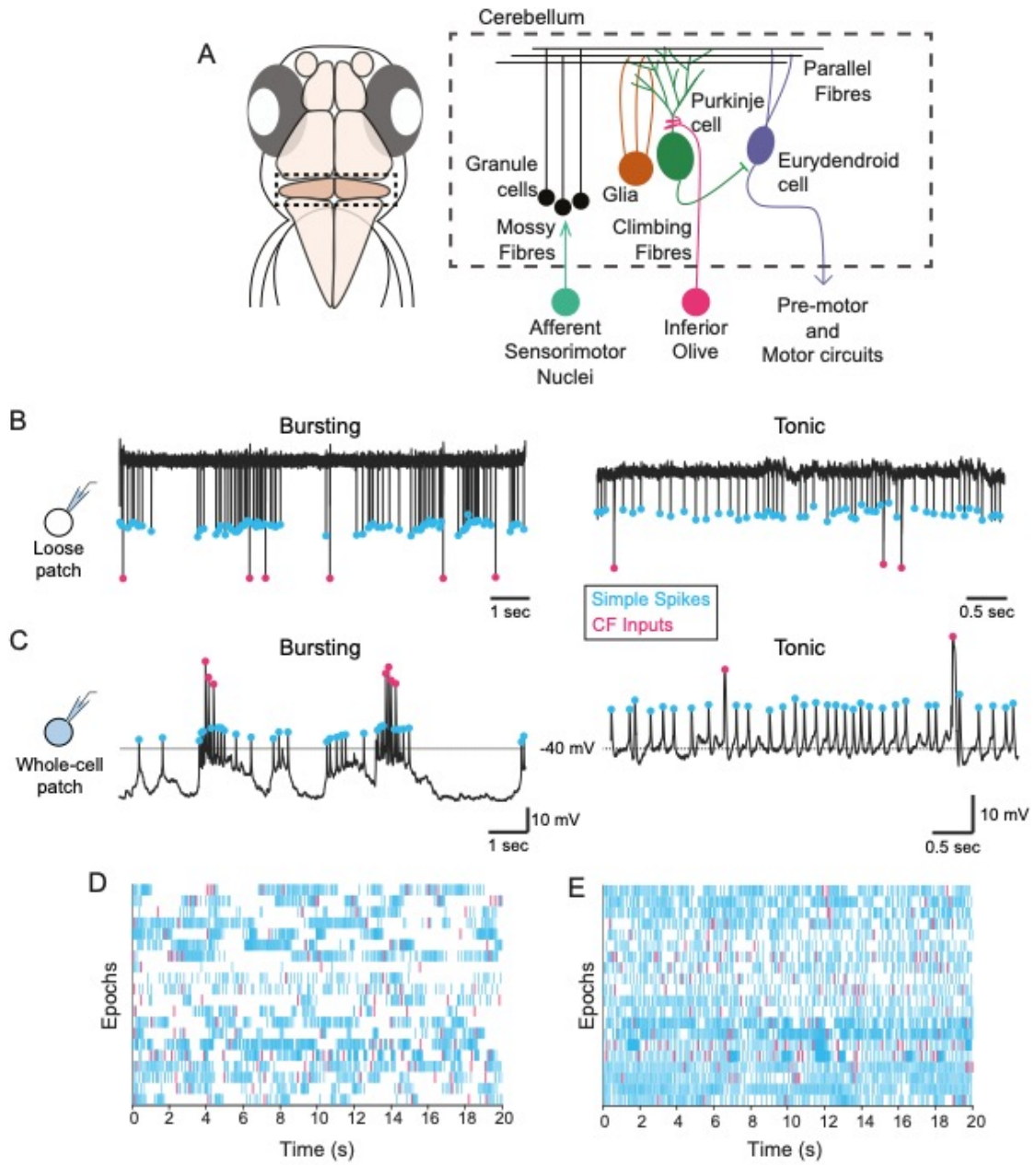


Figure 1 - Cerebellar Purkinje neurons (PNs) in larval zebrafish exhibit membrane potential bistability (A) Schematic of the cerebellar circuitry in larval zebrafish (*Danio rerio*), showing the location of the cerebellum in the brain (left) and the cell types and circuit architecture (right). Purkinje neurons (green) are the principal neurons in this circuit and receive multiple inputs from afferent sensorimotor nuclei outside the cerebellum via parallel fibers (black) and climbing fibers (magenta). **(B)** Representative loose patch recordings from larval zebrafish PNs showing the two modes of firing. The bursting mode recording is shown on the left, and the tonic mode is shown on the right. In each mode, electrical events can be

distinguished by their amplitude, with small-amplitude simple spikes marked in cyan and the large-amplitude CF inputs marked in magenta. Note that the time scales shown for each mode are different, with a longer duration shown for the bursting mode than the tonic mode. **(C)** Representative whole-cell current clamp recordings showing the two modes of firing in PNs, with the same colour scheme for events as in (B). These recordings were obtained from different cells from those shown in (B). The horizontal dotted line marks -40mV. **(D)** Raster plots of 20 second-long epochs of events from recordings from n=10 cells in the bursting mode. **(E)** Raster plots as in (D) for n=10 cells in the tonic mode.

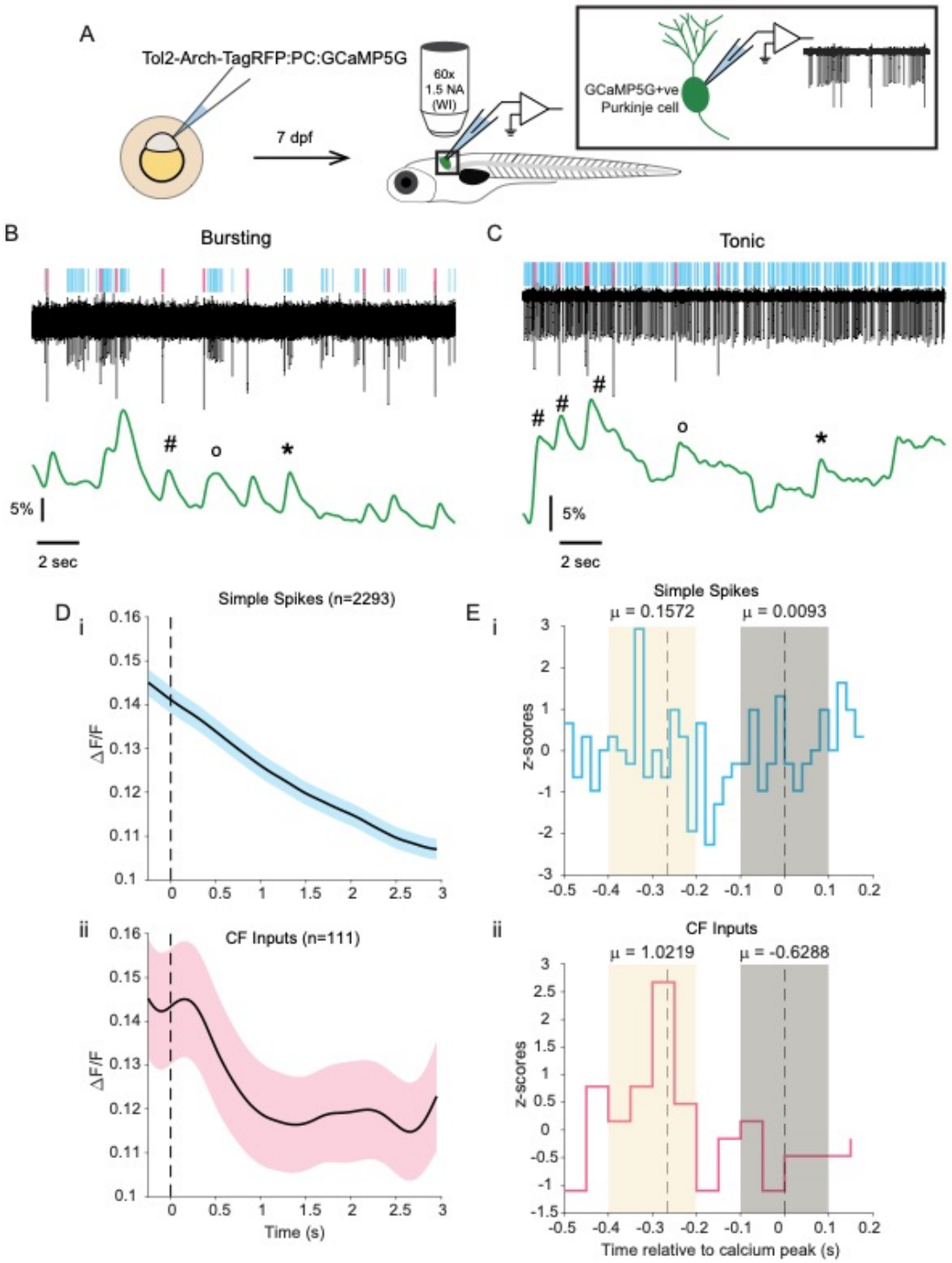
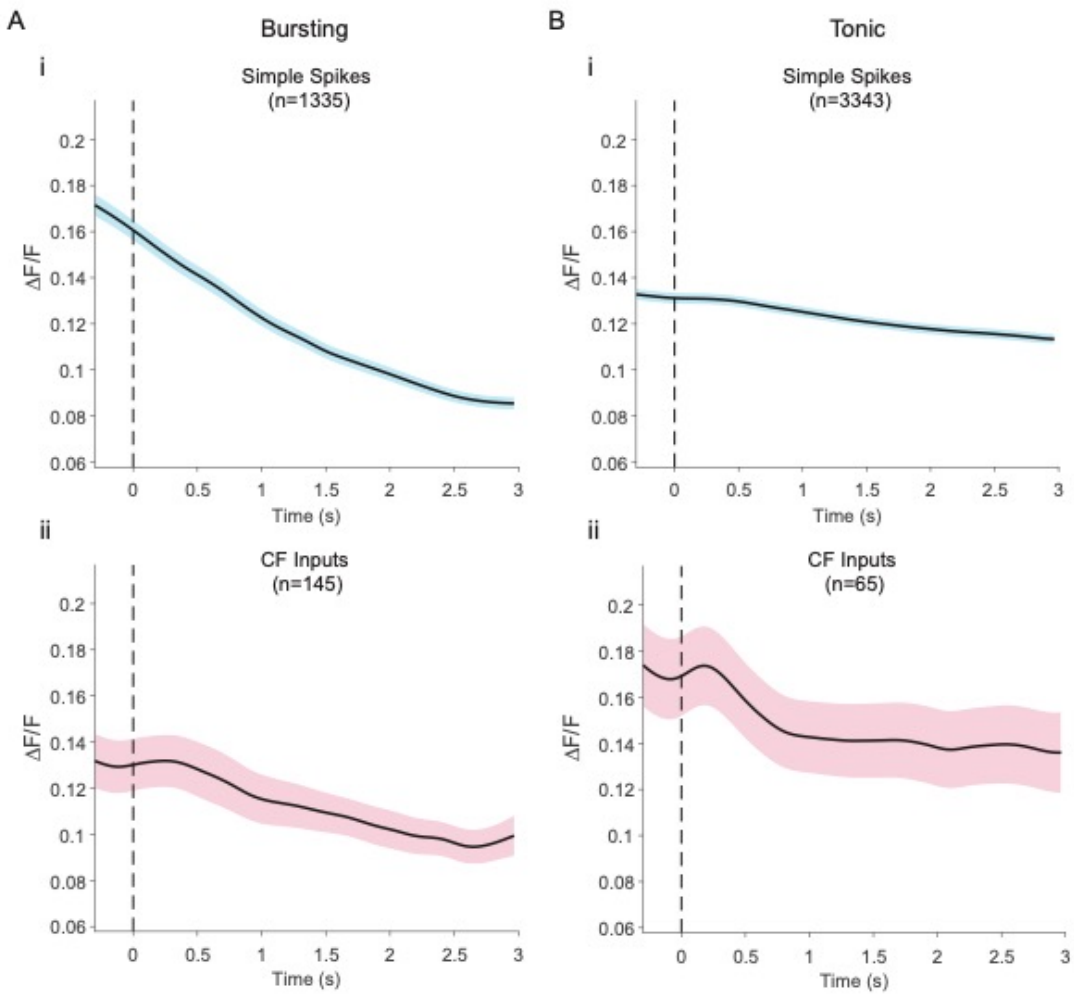
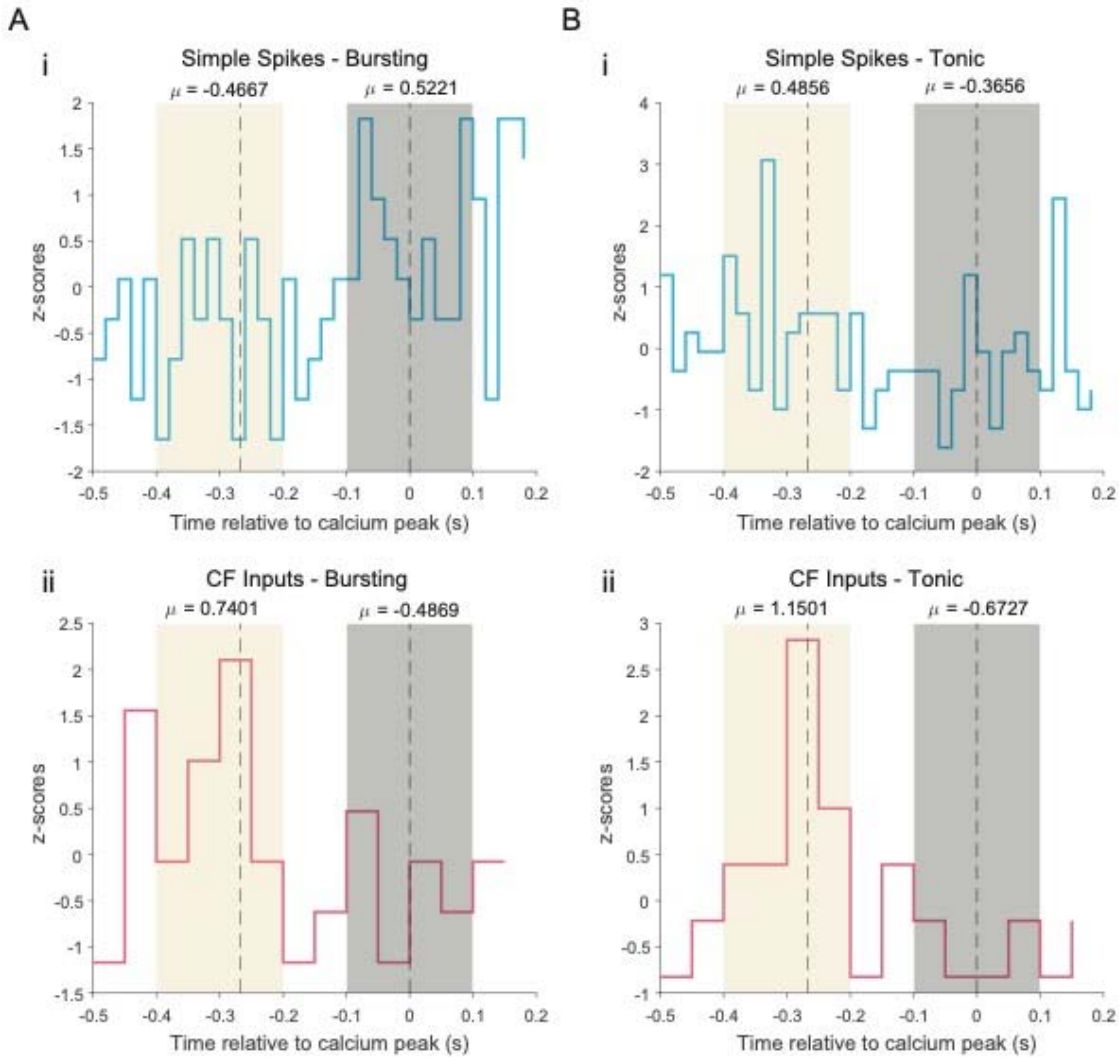


Figure 2 - Both simple spikes and CF EPSPs contribute to the calcium signal in PNs irrespective of cellular state. (A) Schematic of the experimental design and setup of simultaneous calcium imaging and electrophysiology in larval zebrafish. Wild-type zebrafish embryos were injected at the 1-cell stage with the Tol2-Arch-TagRFP:PC:GCaMP5G construct (left). At 7 days post fertilisation (dpf) (right), isolated PNs expressing GCaMP (right, inset) were targeted for simultaneous electrophysiology and imaging (cell and pipette outlined in white in inset). **(B)** and **(C)**

Representative traces obtained during simultaneous imaging and electrophysiology for cells in the bursting (B) and tonic (C) mode. Simple spikes and CF inputs in the trace are identified and marked in the raster above the trace in cyan and magenta, respectively. The simultaneously-recorded change in GCaMP fluorescence (dF/F) is shown below in green. Calcium transients are marked with an asterisk (*) if they correspond to only simple spikes in the electrophysiology channel. Similarly, they are marked with a hash symbol (#) if they correspond only to CF inputs, and with an open circle (o) if they correspond to both simple spikes and CF inputs. **(D)** Event-triggered average calcium signals. **(i)** Simple spike-triggered average calcium signal. The black line represents the mean and the cyan-shaded region is the SEM. n=2293 events from N=15 cells. **(ii)** CF-input triggered average calcium signal, similar to (i). The black line represents the mean and the magenta region is the SEM. n=111 events from N=15 cells **(E)** Peri-event time histograms (PETHs) of **(i)** simple-spikes and **(ii)** CF inputs relative to peaks in the calcium signal. The vertical dashed lines mark t=-0.267s and t=0s, which is the interval during which the GCaMP5G increases from baseline to peak. The y-axis values represent z-scores of the event counts. The mean z-scores for the yellow and grey shaded regions are given above each shaded region.



Supplementary Figure 2.1 - Event-triggered average calcium signal for each state (A) Event-triggered average calcium signal in bursting cells. The simple spike-triggered average is on top, with a cyan shaded error bar and the CF input-triggered average is on the bottom, with a magenta shaded error bar. The black line is the mean and the error represented is the SEM. **(B)** Event-triggered average calcium signal in tonic cells, similar to **(A)**.



Supplementary Figure 2.2 - Peri-event time histograms of simple spikes and CF inputs around the peak calcium signal in each state (A) Peri-event time histogram (PETH) of **(i)** simple-spikes and **(ii)** CF inputs relative to peaks in the calcium signal for cells in the bursting mode. The PETHs are coloured cyan and magenta for simple spikes and CF inputs, respectively. The vertical dashed lines mark the period it takes for the GCaMP5G sensor to reach the peak from baseline. The mean z-score for the yellow (-400 to -200 ms relative to the calcium peak) and grey (-100 to +100 ms relative to the calcium peak) regions are given above the respective shaded bars. Data from n=6 cells. **(B)** PETHs for cells in the tonic mode, similar to **(A)**. Data from n=9 cells.

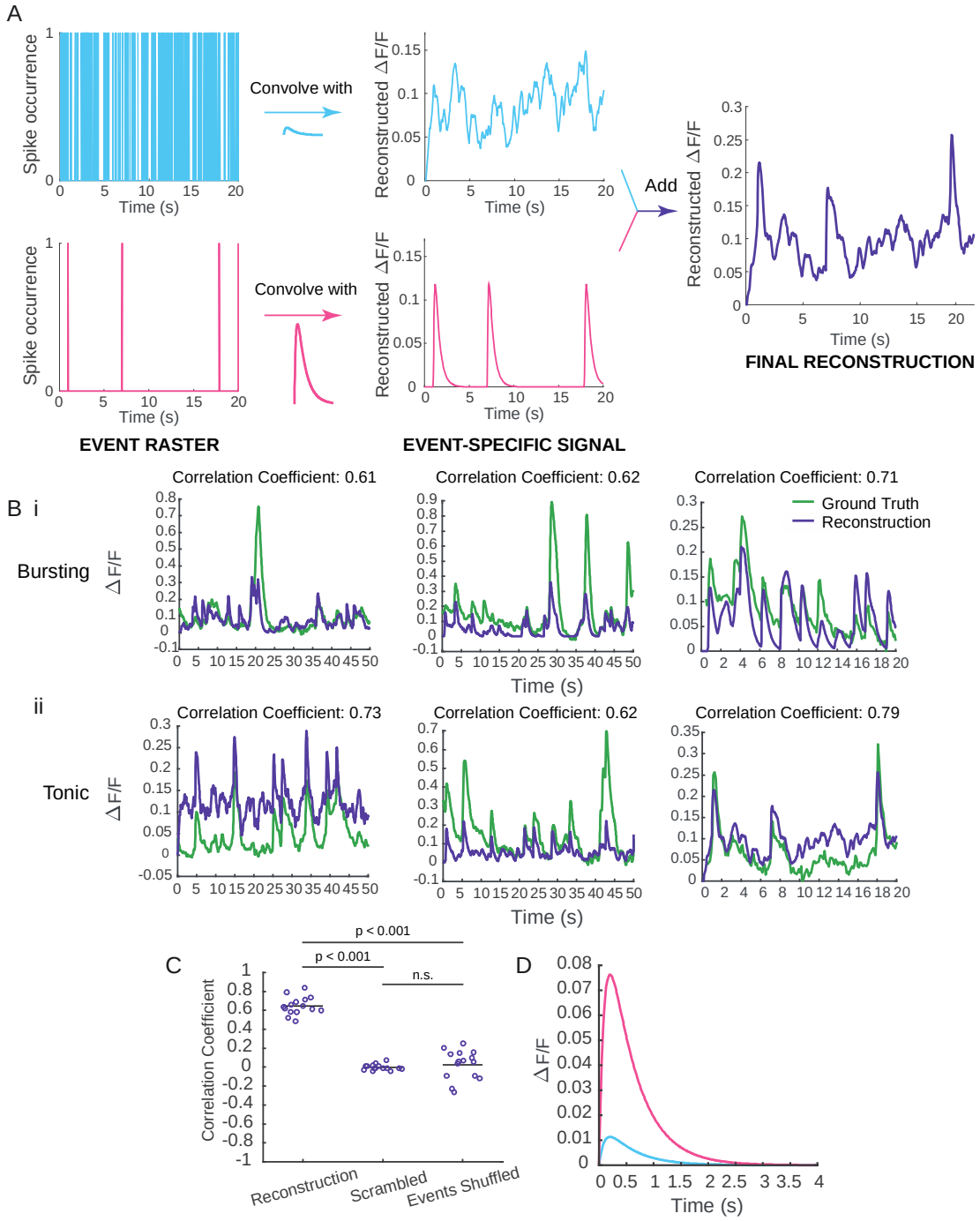


Figure 3 - Reconstructing the calcium signal from electrophysiology traces. (A) Representative example showing the method used to reconstruct the calcium signal from an electrophysiological trace. Simple spike (cyan) and CF input (magenta) rasters (left) are independently convolved with calcium sensor kernels (cyan for simple spikes, magenta for CF inputs) to generate an event-specific calcium signal time series (middle). These are then simply added together to get the final reconstruction (right), shown in dark purple. This example is the same cell as shown in the rightmost plot in B(ii). **(B)** Comparison of ground truth and reconstructions for 6

cells (3 bursting (i) and 3 tonic (ii)) chosen randomly from the dataset. The ground truth trace is shown in green and the reconstruction in purple, and the Pearson's correlation coefficient between the two traces is marked for each pair. **(C)** Quantification of the quality of reconstructions obtained using the optimal kernel combination. Correlation coefficients were calculated between the ground truth for each cell (n=15 cells; 6 bursting, 9 tonic) and either the true reconstruction, a scrambled version of the reconstruction, or the reconstruction for when simple spike and CF input event times were shuffled. A Kruskal-Wallis test yielded a p-value of 2.108 x 10^-7, which was followed by a post-hoc Dunn's test for individual comparisons, the latter of which is marked on the figure. **(D)** The optimal calcium sensor kernels for simple spikes (cyan) and CF inputs (magenta).

A

How well does each MSE-optimized kernel generalize? MSE

Cell No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	"Average"
1	0.009	0.161	0.030	0.054	0.156	0.097	0.022	0.030	0.021	0.364	0.470	0.017	0.039	0.021	0.020	0.048
2	0.060	0.018	0.035	0.026	0.018	0.020	0.040	0.035	0.040	0.030	0.039	0.080	0.031	0.082	0.082	0.028
3	0.038	0.080	0.035	0.036	0.078	0.052	0.033	0.032	0.034	0.182	0.247	0.057	0.037	0.065	0.064	0.035
4	0.028	0.023	0.025	0.019	0.026	0.020	0.024	0.021	0.025	0.043	0.068	0.038	0.024	0.044	0.042	0.019
5	0.338	0.101	0.208	0.150	0.103	0.115	0.233	0.201	0.238	0.167	0.242	0.457	0.186	0.477	0.476	0.162
6	0.059	0.038	0.045	0.042	0.043	0.040	0.048	0.045	0.048	0.049	0.072	0.072	0.043	0.074	0.074	0.042
7	0.125	0.265	0.116	0.136	0.267	0.186	0.113	0.115	0.114	0.559	0.747	0.170	0.122	0.184	0.182	0.128
8	0.094	0.124	0.096	0.092	0.124	0.104	0.093	0.091	0.095	0.198	0.251	0.110	0.097	0.119	0.117	0.092
9	0.052	0.114	0.030	0.064	0.145	0.094	0.033	0.038	0.031	0.249	0.403	0.067	0.030	0.067	0.067	0.048
10	0.215	0.140	0.195	0.161	0.132	0.145	0.198	0.185	0.202	0.132	0.128	0.249	0.190	0.263	0.260	0.170
11	0.339	0.214	0.294	0.257	0.211	0.232	0.302	0.286	0.306	0.188	0.181	0.383	0.284	0.394	0.393	0.266
12	0.021	0.283	0.070	0.120	0.280	0.189	0.055	0.073	0.053	0.571	0.721	0.012	0.087	0.012	0.012	0.106
13	0.050	0.067	0.038	0.038	0.066	0.047	0.038	0.037	0.039	0.146	0.195	0.070	0.037	0.074	0.074	0.037
14	0.024	0.380	0.094	0.160	0.373	0.253	0.073	0.098	0.070	0.763	0.955	0.009	0.118	0.010	0.010	0.143
15	0.147	0.154	0.074	0.170	0.251	0.197	0.091	0.107	0.082	0.192	0.411	0.128	0.069	0.097	0.104	0.125
SUM	0.160	0.216	0.138	0.152	0.227	0.179	**0.140**	0.139	0.140	0.383	0.513	0.192	0.139	0.198	0.198	0.145

Mean Squared-Error ($\times 10^{-1}$): 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

Total MSE: 0.2, 0.4

Kernel Used

B

How well does each MSE-optimized kernel generalize? Correlations

Cell No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	"Average"
1	0.84	0.55	0.62	0.68	0.62	0.64	0.69	0.68	0.67	0.46	0.53	0.88	0.58	0.88	0.88	0.66
2	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
3	0.69	0.48	0.54	0.60	0.54	0.56	0.60	0.59	0.59	0.39	0.46	0.70	0.51	0.70	0.70	0.57
4	0.69	0.72	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.69	0.72	0.64	0.72	0.65	0.64	0.73
5	0.63	0.82	0.81	0.79	0.81	0.80	0.79	0.79	0.79	0.82	0.82	0.48	0.82	0.52	0.50	0.80
6	0.52	0.66	0.64	0.61	0.64	0.63	0.61	0.61	0.62	0.67	0.66	0.46	0.65	0.48	0.47	0.62
7	0.57	0.56	0.57	0.58	0.57	0.57	0.58	0.58	0.58	0.53	0.55	0.55	0.56	0.56	0.55	0.58
8	0.62	0.52	0.55	0.58	0.55	0.56	0.58	0.58	0.57	0.47	0.51	0.62	0.54	0.62	0.62	0.57
9	0.72	0.86	0.86	0.84	0.85	0.85	0.84	0.84	0.84	0.86	0.86	0.62	0.86	0.64	0.63	0.85
10	0.61	0.58	0.60	0.62	0.60	0.61	0.62	0.62	0.61	0.54	0.58	0.58	0.59	0.59	0.59	0.61
11	0.59	0.65	0.65	0.64	0.65	0.65	0.64	0.65	0.65	0.64	0.64	0.53	0.65	0.54	0.54	0.65
12	0.70	0.58	0.61	0.63	0.61	0.62	0.63	0.63	0.63	0.55	0.58	0.71	0.59	0.72	0.71	0.62
13	0.50	0.48	0.51	0.52	0.51	0.51	0.52	0.52	0.52	0.45	0.48	0.47	0.49	0.48	0.47	0.51
14	0.76	0.55	0.61	0.66	0.61	0.63	0.66	0.65	0.65	0.49	0.55	0.78	0.58	0.78	0.78	0.64
15	0.34	0.59	0.54	0.48	0.54	0.52	0.48	0.49	0.49	0.64	0.60	0.27	0.57	0.29	0.28	0.51
SUM	9.498	9.306	9.540	9.673	9.541	9.598	**9.674**	9.665	9.655	8.921	9.263	9.001	9.421	9.138	9.066	9.624

Pearson Correlation Coefficient: 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1

Total Correlation Coefficients: 9, 9.2, 9.4, 9.6

Kernel Used

Supplementary Figure 3 - Finding the most generalizable GCaMP5G kernel (A) Mean-squared errors (MSEs) between the ground truth and reconstructed calcium signal for each cell and kernel combination used. Each row represents a single cell and each column, the kernels used. The “average” kernel’s coefficients were the mean of the coefficients of the kernels optimised for each cell. The last row, with a different colormap, represents the total MSE across all reconstructions for a kernel. The values corresponding to each colour are shown in the colour bars on the right. Bright colours represent lower MSEs. **(B)** Similar representation as above, but for Pearson’s correlation coefficients between the ground truth and reconstructions. Note that the colormaps are reversed as compared to (A), so that brighter colours represent higher correlations. In both (A) and (B), the kernel chosen for all future work (i.e. no 7) is highlighted in bold.

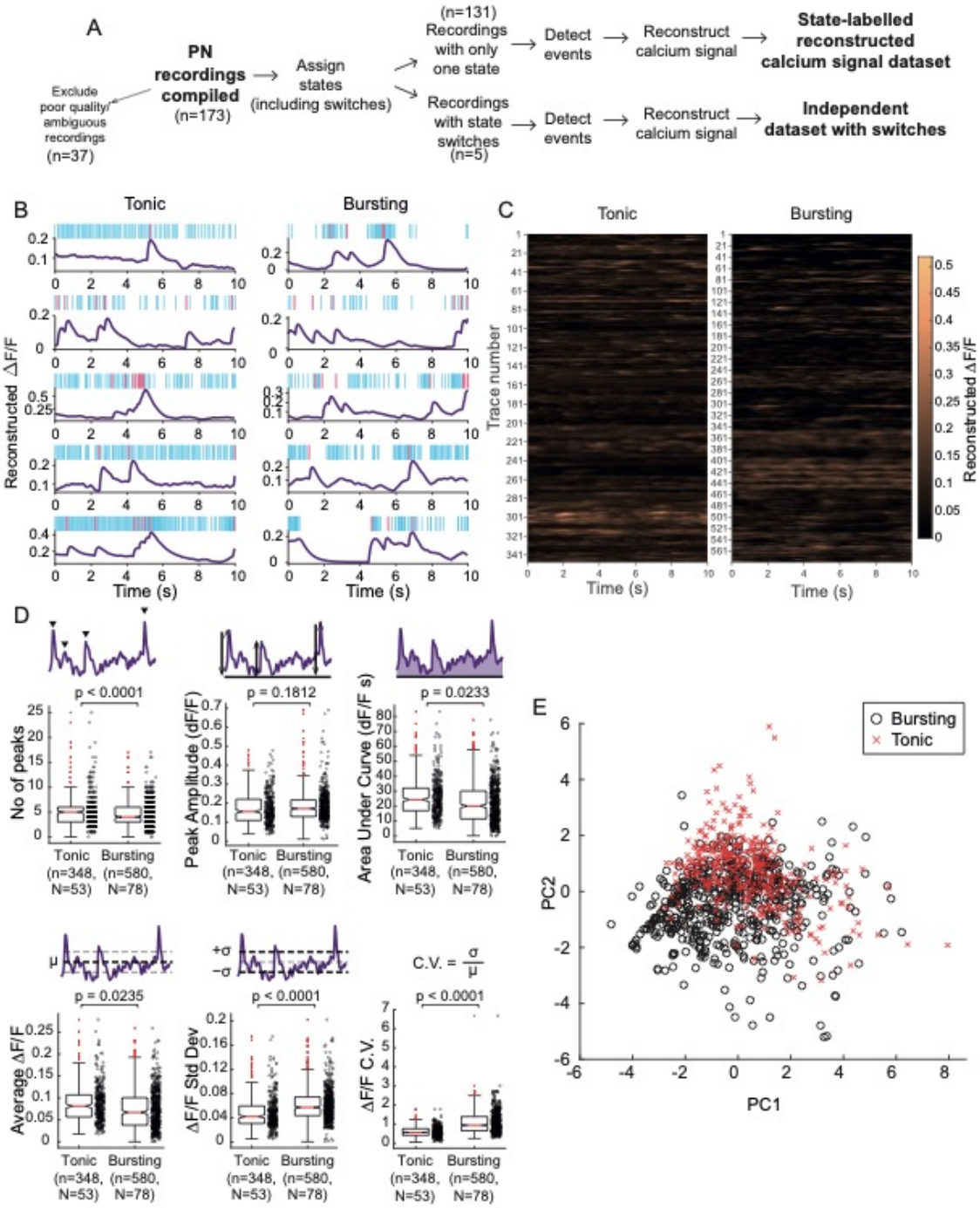


Figure 4 - Generating a labelled dataset of reconstructed calcium signal traces **(A)** Flowchart showing how electrophysiological recordings from PNs were processed to generate the state-labelled calcium signal database. **(B)** Representative randomly sampled traces from the state-labelled reconstructed calcium signal dataset. The left column has reconstructions from tonic cells and the right column has reconstructions from bursting cells. Above each trace is the raster showing the events in the source recording, with simple spikes in cyan, and CF

inputs in magenta. **(C)** Heatmaps representing all the reconstructions in the state-labelled calcium signal dataset, for each state, after splitting all traces into 10 second-long non-overlapping chunks. (n=348 tonic traces; n=580 bursting traces). **(D)** Comparing the distribution of trace properties for both states. The top three plots show (from left to right) the number of peaks, peak amplitude and area under the curve of the reconstructions in the dataset. The bottom three plots show (from left to right) the mean, standard deviation and the coefficient of variation of the traces. n's are the number of 10s-long samples. N's are the number of recordings used to generate the samples. p-values were computed using linear mixed-effects models (see Table 2). **(E)** Principal Components Analysis of trace properties, with data points marked by state - tonic (red crosses) and bursting (black open circles).

Table 1 - Number of recordings compiled for the state-labelled imaging dataset

	Extracellular	Intracellular	**Total**
Bursting	39	48	**87**
Tonic	11	43	**54**
Switch	0	7	**7**
Ambiguous	0	25	**25**
Total	**50**	**123**	**173**

Table 2 - Linear mixed-effects model statistics for comparing the various trace properties (related to Fig 4D)

Parameter	Model #	Model	DF	AIC	Test	p-value	Effect Size (*b*) [Tonic relative to Bursting]
No of peaks (NPeaks)	1	NPeaks ~ 1	2	5066.33	-	-	-
	2	NPeaks ~ 1|Cell	3	4550.27	1 vs 2	<0.0001	-
	3	NPeaks ~ State|Cell	4	4532.48	2 vs 3	**<0.0001**	**2.3575**
Average peak amplitude (Amp)	1	Amp ~ 1	2	-2549.08	-	-	-
	2	Amp ~ 1|Cell	3	-3000.83	1 vs 2	<0.0001	-
	3	Amp ~ State|Cell	4	-3000.62	2 vs 3	**0.1812**	**-0.0122**
Area Under the Curve (AUC)	1	AUC ~ 1	2	7192.146	-	-	-
	2	AUC ~ 1|Cell	3	6315.809	1 vs 2	<.0001	-
	3	AUC ~ State|Cell	4	6312.661	2 vs 3	**0.0233**	**4.4053**
Mean dF/F (Mean)	1	Mean ~ 1	2	-3365.760	-	-	-
	2	Mean ~ 1|Cell	3	-4244.373	1 vs 2	<.0001	-
	3	Mean ~ State|Cell	4	-4247.505	2 vs 3	**0.0235**	**0.0147**
Standard deviation of dF/F (Stdev)	1	Stdev ~ 1	2	-4566.303	-	-	-
	2	Stdev ~ 1|Cell	3	-5034.750	1 vs 2	<.0001	-
	3	Stdev ~ State|Cell	4	-5052.501	2 vs 3	**<.0001**	**-0.0135**
Coefficient of Variation of dF/F (CV)	1	CV ~ 1	2	1378.348	-	-	-
	2	CV ~ 1|Cell	3	558.4086	1 vs 2	<.0001	-
	3	CV ~ State|Cell	4	511.7434	2 vs 3	**<.0001**	**-0.4744**

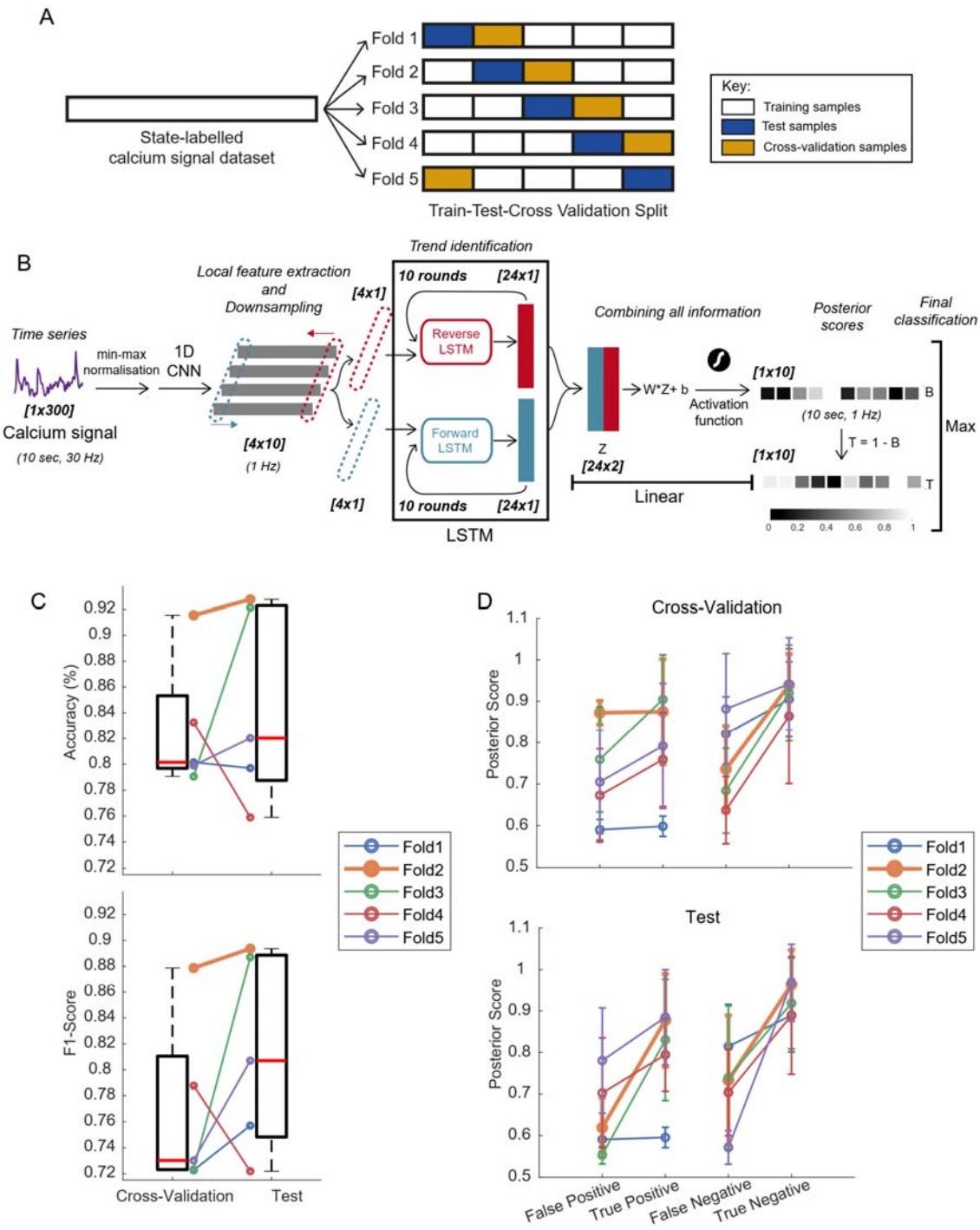


Figure 5 - Design and training of CaMLsort (A) Flowchart showing how the state-labelled calcium signal dataset was split into training and test datasets using a 5-fold 60-20-20 split into training (white boxes), test (blue boxes) and cross-validation (yellow boxes) samples, respectively. The networks that performed well in the training phase, as assessed by performance on the cross-validation samples, were then challenged with the unseen test data (blue boxes) in each fold. An independent

dataset with state switches was also used to further validate the networks obtained (see Fig 6). **(B)** The architecture of the convolutional recurrent neural networks used to solve the classification problem. The network takes an input time series trace sampled at 30Hz (purple) and that is a multiple of 10 seconds long in duration (here just 10 seconds long). This trace is first normalized using min-max normalization, following which it is passed through a 1-dimensional convolutional neural network (1D CNN), which has 4 kernels with a step size and stride length of 30 each. The resulting 4 traces (grey boxes) are downsampled, local feature-extracted versions of the original trace. All 4 outputs are passed to the LSTM module (central black box) for trend identification. The LSTM is a bidirectional one, which processes traces in both the forward (blue arrow) and reverse (red arrow) directions. A single time step is taken from all the CNN outputs (red and blue dotted boxes) and passed to the respective LSTM cell, which then processes the input to produce a resultant “hidden state” output (red and blue filled boxes). This output is passed back into the LSTM cell along with the next sample from the CNN’s output traces, until every time step from it has been processed (here 10 rounds). Once this is done, the final “hidden state” from each LSTM cell is retained and concatenated before their weighted average is taken by a Linear layer. The resulting average is converted to a “posterior score” using an activation function. These posterior scores represent the likelihood of the cell being in the bursting state at each time step. The likelihood of it being in the tonic state can be calculated from this. The final call is taken to be the state which has a higher posterior score. The text in italics represents the net effect of each phase of the neural network. The numbers in parentheses indicate the sampling frequency and/or the duration of the resultant vectors at various stages of processing by the neural network. Similarly, the numbers in bold and italics within square brackets indicate the size of the vectors/matrices at various stages of passing through the network. **(C)** Average classification accuracy (top) and F1-scores (bottom) at the cross-validation and the test phases, for trained networks. Each fold is represented in a different colour, marked in the legend to the right of panels (C) and (D). **(D)** Posterior scores for the predicted class in the cross-validation (top) and test (bottom) phases. Each dot represents the mean posterior score for a trained network, with error bars representing standard deviation.

Table 3 - Prediction accuracies and F1-scores for various models trained to solve the classification problem.

Model	Fold	Validation		Test	
		Accuracy (%)	F1-score	Accuracy(%)	F1-score
Support Vector Machine (SVM)	1	81.41	0.7206	70.19	0.6077
	2	72.68	0.593	79.66	0.6842
	3	75.84	0.703	72.55	0.6055
	4	73.10	0.6422	73.65	0.671
	5	70.67	0.629	74.89	0.6755
	Mean ± Stdev	74.74 ± 4.16	0.6576 ± 0.0531	74.19 ± 3.51	0.6488 ± 0.0388
Deep Neural Network	1	83.91	0.7893	72.99	0.6649
	2	77.71	0.7205	82.89	0.7849
	3	74.85	0.7099	76.10	0.6788
	4	73.31	0.6535	75.12	0.7143
	5	77.82	0.7060	75.24	0.6629
	Mean ± Stdev	77.52 ± 4.06	0.7158 ± 0.0486	76.47 ± 3.77	0.7012 ± 0.0511
1D CNN	1	87.20	0.8163	84.00	0.7933
	2	80.67	0.7541	88.33	0.8336
	3	82.66	0.7880	80.09	0.7263
	4	78.28	0.7185	81.07	0.7719
	5	82.60	0.7604	78.34	0.7004

	Mean ± Stdev	82.28 ± 3.28	0.7675 ± 0.0369	82.37 ± 3.91	0.7651 ± 0.0530
1D CNN (3x)	1	87.54	0.8226	84.00	0.7893
	2	80.61	0.7535	86.97	0.8142
	3	80.61	0.7650	78.41	0.7089
	4	78.62	0.73	73.66	0.6592
	5	82.49	0.7592	76.83	0.6775
	Mean ± Stdev	81.97 ± 3.40	0.7661 ± 0.0343	79.97 ± 5.42	0.7298 ± 0.0686
CNN-LSTM	1	86.74	0.8246	82.87	0.7831
	2	**79.32**	**0.7358**	**87.03**	**0.8259**
	3	87.03	0.8259	81.25	0.7288
	4	84.83	0.8036	80.54	0.7542
	5	82.60	0.7596	71.93	0.7583
	Mean ± Stdev	84.10 ± 3.21	0.7899 ± 0.0404	80.72 ± 5.52	0.7701 ± 0.0367

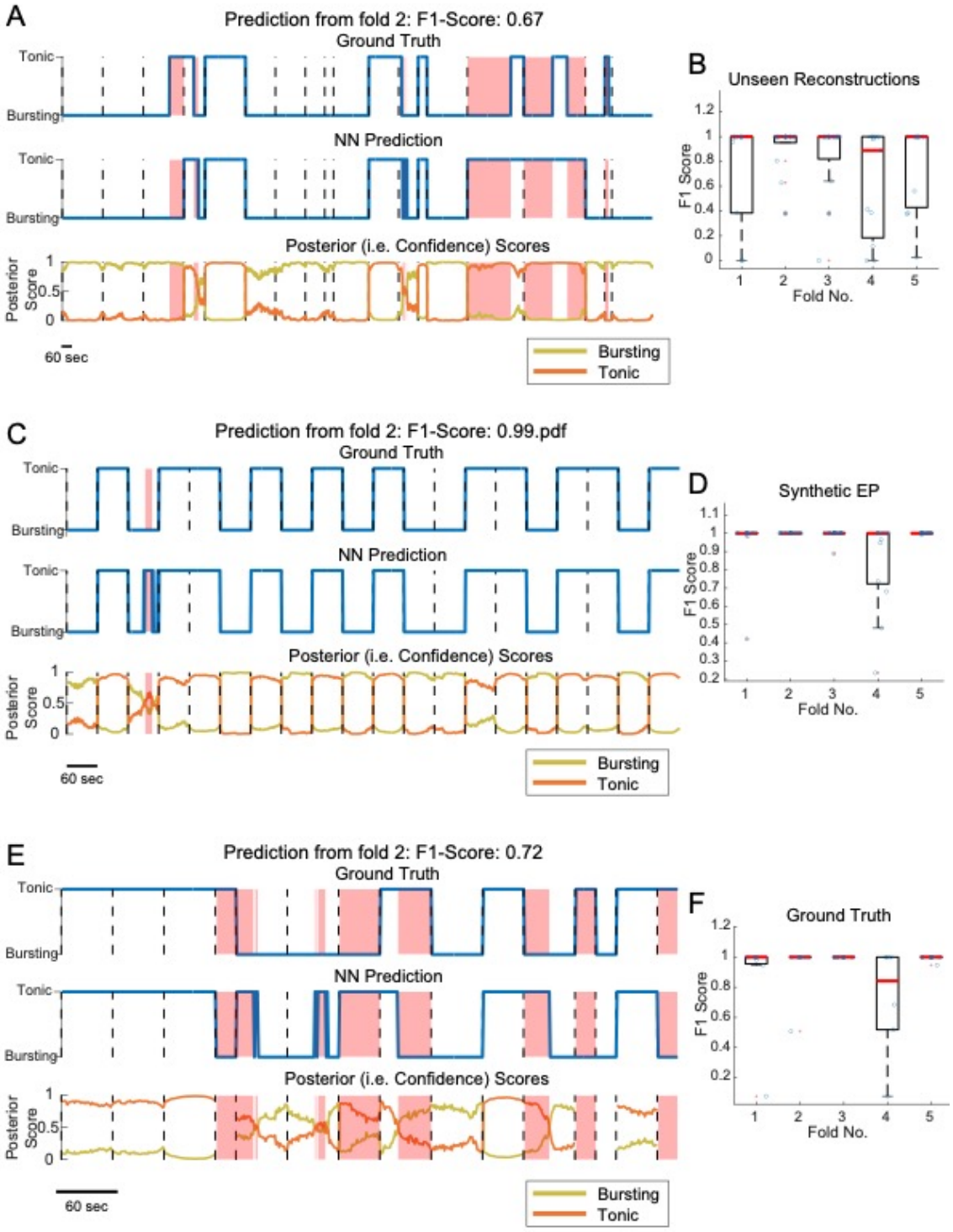
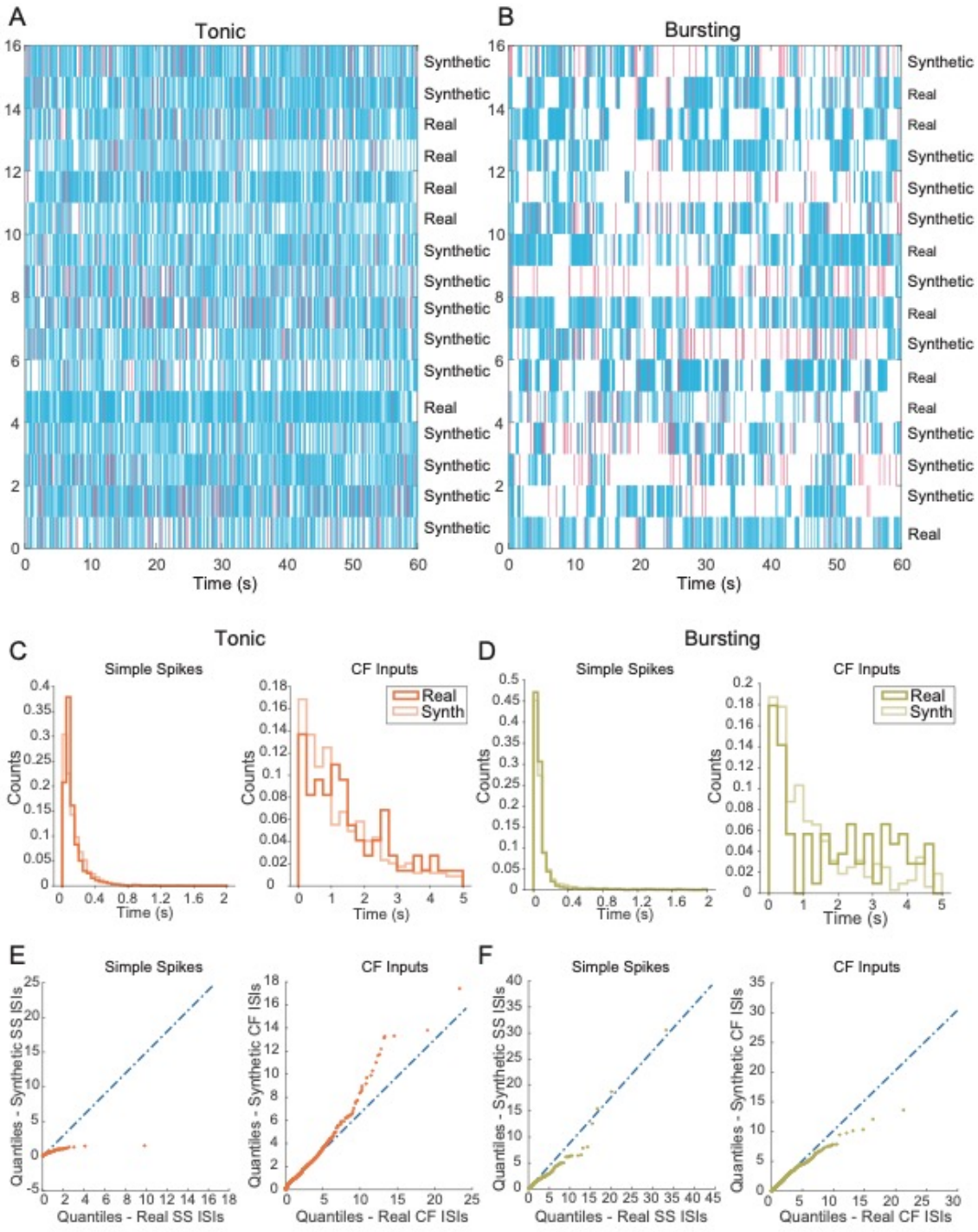


Figure 6 - Validation of CaMLsort on independent datasets (A) CaML sort predictions on calcium signal reconstructions from previously unseen data (n=19 recordings). Predictions from CaMLsort (middle) against ground truth state (top). Predictions for each recording were made independently, but individual recordings have been stitched here for the purposes of representation. Each recording is

separated by a vertical dashed line. The areas shaded in pink represent those regions where the prediction and ground truth do not match. The bottommost plot has posterior scores for each time step. The orange trace is the confidence score for the tonic class, while the green one shows the confidence scores for the bursting class. The F1 score in the heading shows the total F1-score obtained from comparing the two stitched time series. **(B)** Individual F1-scores for each cell and each fold of networks trained. The data represented in (A) is from network fold 2. **(C)** Predictions as in (A), but on calcium signal reconstructions from artificially generated PN spike trains of each class (n=20 recordings). **(D)** Individual F1-scores for data in (C) for all trained networks. **(E)** CaMLsort predictions on experimental data obtained from simultaneous imaging and electrophysiology (n=15 recordings). **(F)** Individual F1-scores for data in (E) across all folds of trained networks, similar to (B) and (D).



Supplementary Figure 6 - Generating synthetic Purkinje neuron spike trains (A) Rasters of simple spikes (cyan) and CF inputs (magenta) for 'tonic' traces, from both real cells and artificially generated (synthetic) spike trains, as indicated by the label on the right. **(B)** Rasters of simple spikes and CF inputs for a mix of 'bursting' traces from real cells and artificially generated (synthetic) spike trains, as in (A). **(C)** Histograms of the distributions of interevent intervals of simple spikes (left) and CF inputs (right) in real and synthetic traces in tonic cells. **(D)** Histograms of the

distributions of interevent intervals of simple spikes (left) and CF inputs (right) in real and synthetic traces in bursting cells. **(E)** QQ plots of the interevent interval distributions between real and synthetic traces for tonic cells. The left plot is for simple spikes and the right one for CF inputs. **(F)** QQ plots of the interevent interval distributions between real and synthetic traces for bursting cells. The left plot is for simple spikes and the right one for CF inputs.

Reference:

Abbinanti MD, Zhong G, Harris-Warrick RM. 2012. Postnatal emergence of serotonin-induced plateau potentials in commissural interneurons of the mouse spinal cord. *Journal of Neurophysiology* **108**:2191–2202. doi:10.1152/jn.00336.2012

Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ. 2013. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat Methods* **10**:413–420. doi:10.1038/nmeth.2434

Ali F, Kwan AC. 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

Armstrong DM, Edgley SA. 1984. Discharges of Purkinje cells in the paravermal part of the cerebellar anterior lobe during locomotion in the cat. *J Physiol (Lond)* **352**:403–424.

Bae Y-K, Kani S, Shimizu T, Tanabe K, Nojima H, Kimura Y, Higashijima S, Hibi M. 2009. Anatomy of zebrafish cerebellum and screen for mutations affecting its development. *Dev Biol* **330**:406–426. doi:10.1016/j.ydbio.2009.04.013

Borst JG, Helmchen F. 1998. Calcium influx during an action potential. *Methods Enzymol* **293**:352–371. doi:10.1016/s0076-6879(98)93023-3

Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS. 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**:295–300. doi:10.1038/nature12354

Conway BA, Hultborn H, Kiehn O, Mintz I. 1988. Plateau potentials in alpha-motoneurones induced by intravenous injection of L-dopa and clonidine in the spinal cat. *J Physiol (Lond)* **405**:369–384. doi:10.1113/jphysiol.1988.sp017337

Cowan RL, Wilson CJ. 1994. Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology* **71**:17–32. doi:10.1152/jn.1994.71.1.17

De Zeeuw CI, Hoebeek FE, Bosman LWJ, Schonewille M, Witter L, Koekkoek SK. 2011. Spatiotemporal firing patterns in the cerebellum. *Nature Reviews Neuroscience* **12**:327–344. doi:10.1038/nrn3011

Dickinson PS, Nagy F. 1983. Control of a central pattern generator by an identified modulatory interneurone in crustacea. II. Induction and modification of plateau properties in pyloric neurones. *Journal of Experimental Biology* **105**:59–82. doi:10.1242/jeb.105.1.59

Egorov AV, Hamam BN, Fransén E, Hasselmo ME, Alonso AA. 2002. Graded persistent activity in entorhinal cortex neurons. *Nature* **420**:173–178. doi:10.1038/nature01171

Engbers JDT, Fernandez FR, Turner RW. 2013. Bistability in Purkinje neurons: Ups and downs in cerebellar research. *Neural Networks*, Computation in the Cerebellum **47**:18–31. doi:10.1016/j.neunet.2012.09.006

Fuentealba P, Timofeev I, Bazhenov M, Sejnowski TJ, Steriade M. 2005. Membrane Bistability in Thalamic Reticular Neurons During Spindle Oscillations. *Journal of Neurophysiology* **93**:294–304. doi:10.1152/jn.00552.2004

Göbel W, Helmchen F. 2007. In Vivo Calcium Imaging of Neural Network Function. *Physiology* **22**:358–365. doi:10.1152/physiol.00032.2007

Grewe BF, Helmchen F. 2009. Optical probing of neuronal ensemble activity. *Current Opinion in Neurobiology*, Neuronal and glial cell biology ● New technologies **19**:520–529. doi:10.1016/j.conb.2009.09.003

Grienberger C, Konnerth A. 2012. Imaging Calcium in Neurons. *Neuron* **73**:862–885. doi:10.1016/j.neuron.2012.02.011

Harmon TC, McLean DL, Raman IM. 2020. Integration of Swimming-Related Synaptic Excitation and Inhibition by olig2+ Eurydendroid Neurons in Larval Zebrafish Cerebellum. *J Neurosci* **40**:3063–3074. doi:10.1523/JNEUROSCI.2322-19.2020

He K, Zhang X, Ren S, Sun J. 2016. Deep Residual Learning for Image Recognition2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778. doi:10.1109/CVPR.2016.90

Hochreiter S, Schmidhuber J. 1997. Long Short-Term Memory. *Neural Computation* **9**:1735–1780. doi:10.1162/neco.1997.9.8.1735

Hounsgaard J, Hultborn H, Jespersen B, Kiehn O. 1988. Bistability of alpha-motoneurones in the decerebrate cat and in the acute spinal cat after intravenous 5-hydroxytryptophan. *The Journal of Physiology* **405**:345–367. doi:10.1113/jphysiol.1988.sp017336

Hounsgaard J, Hultborn H, Jespersen B, Kiehn O. 1984. Intrinsic membrane properties causing a bistable behaviour of alpha-motoneurones. *Exp Brain Res* **55**:391–394. doi:10.1007/BF00237290

Ikenaga T, Yoshida M, Uematsu K. 2006. Cerebellar efferent neurons in teleost fish. *Cerebellum* **5**:268–274. doi:10.1080/14734220600930588

Jahnsen H, Llinás R. 1984. Electrophysiological properties of guinea-pig thalamic neurones: an in vitro study. *The Journal of Physiology* **349**:205–226. doi:10.1113/jphysiol.1984.sp015153

Kerr JND, Greenberg D, Helmchen F. 2005. Imaging input and output of neocortical networks in vivo. *Proceedings of the National Academy of Sciences* **102**:14063–14068. doi:10.1073/pnas.0506029102

Kingma DP, Ba J. 2017. Adam: A Method for Stochastic Optimization. doi:10.48550/arXiv.1412.6980

Kitamura K, Häusser M. 2011. Dendritic calcium signaling triggered by spontaneous and sensory-evoked climbing fiber input to cerebellar Purkinje cells in vivo. *J Neurosci* **31**:10847–10858. doi:10.1523/JNEUROSCI.2525-10.2011

Knogler LD, Kist AM, Portugues R. 2019. Motor context dominates output from purkinje cell functional regions during reflexive visuomotor behaviours. *eLife* **8**:e42138. doi:10.7554/eLife.42138

Kwan AC, Dan Y. 2012. Dissection of Cortical Microcircuits by Single-Neuron Stimulation In Vivo. *Current Biology* **22**:1459–1467. doi:10.1016/j.cub.2012.06.007

Lechner HA, Baxter DA, Clark JW, Byrne JH. 1996. Bistability and its regulation by serotonin in the endogenously bursting neuron R15 in Aplysia. *J Neurophysiol* **75**:957–962.

Lev-Ram V, Miyakawa H, Lasser-Ross N, Ross WN. 1992. Calcium transients in cerebellar Purkinje neurons evoked by intracellular stimulation. *Journal of Neurophysiology* **68**:1167–1177. doi:10.1152/jn.1992.68.4.1167

Lin B-J, Chen T-W, Schild D. 2007. Cell type-specific relationships between spiking and [Ca2+]i in neurons of the Xenopus tadpole olfactory bulb. *The Journal of Physiology* **582**:163–175. doi:10.1113/jphysiol.2006.125963

Loewenstein Y, Mahon S, Chadderton P, Kitamura K, Sompolinsky H, Yarom Y, Häusser M. 2005. Bistability of cerebellar Purkinje cells modulated by sensory stimulation. *Nat Neurosci* **8**:202–211. doi:10.1038/nn1393

Marder E, Abbott LF, Turrigiano GG, Liu Z, Golowasch J. 1996. Memory from the dynamics of intrinsic membrane currents. *Proceedings of the National Academy of Sciences* **93**:13481–13486. doi:10.1073/pnas.93.24.13481

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. 2004. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* **5**:793–807. doi:10.1038/nrn1519

Matsui H, Namikawa K, Babaryka A, Köster RW. 2014. Functional regionalization of the teleost cerebellum analyzed in vivo. *Proc Natl Acad Sci USA* **111**:11846–11851. doi:10.1073/pnas.1403105111

McCormick DA, Connors BW, Lighthall JW, Prince DA. 1985. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of Neurophysiology* **54**:782–806. doi:10.1152/jn.1985.54.4.782

Misgeld U, Calabresi P, Dodt HU. 1986. Muscarinic modulation of calcium dependent plateau potentials in rat neostriatal neurons. *Pflugers Arch* **407**:482–487. doi:10.1007/BF00657504

Miyakawa H, Lev-Ram V, Lasser-Ross N, Ross WN. 1992. Calcium transients evoked by climbing fiber and parallel fiber synaptic inputs in guinea pig cerebellar Purkinje neurons. *J Neurophysiol* **68**:1178–1189. doi:10.1152/jn.1992.68.4.1178

Moreaux L, Laurent G. 2007. Estimating firing rates from calcium signals in locust projection neurons in vivo. *Frontiers in Neural Circuits* **1**.

Narayanan S, Thirumalai V. 2019. Contributions of the Cerebellum for Predictive and Instructional Control of Movement. *Curr Opin Physiol* **8**:146–151. doi:10.1016/j.cophys.2019.01.011

Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning LibraryAdvances in Neural Information Processing Systems. Curran Associates, Inc.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**:2825–2830.

Peron S, Chen T-W, Svoboda K. 2015. Comprehensive imaging of cortical networks. *Current Opinion in Neurobiology*, Large-Scale Recording Technology (32) **32**:115–123. doi:10.1016/j.conb.2015.03.016

Person AL, Raman IM. 2012. Purkinje neuron synchrony elicits time-locked spiking in the cerebellar nuclei. *Nature* **481**:502–505. doi:10.1038/nature10732

Popa LS, Streng ML, Ebner TJ. 2018. Purkinje Cell Representations of Behavior: Diary of a Busy Neuron. *Neuroscientist* 1073858418785628. doi:10.1177/1073858418785628

Ramirez JE, Stell BM. 2016. Calcium Imaging Reveals Coordinated Simple Spike Pauses in Populations of Cerebellar Purkinje Cells. *Cell Reports* **17**:3125–3132. doi:10.1016/j.celrep.2016.11.075

Renninger SL, Orger MB. 2013. Two-photon imaging of neural population activity in zebrafish. *Methods*, Zebrafish Methods **62**:255–267. doi:10.1016/j.ymeth.2013.05.016

Rupprecht P, Carta S, Hoffmann A, Echizen M, Blot A, Kwan AC, Dan Y, Hofer SB, Kitamura K, Helmchen F, Friedrich RW. 2021. A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. *Nat Neurosci* **24**:1324–1337. doi:10.1038/s41593-021-00895-5

Russell DF, Hartline DK. 1982. Slow active potentials and bursting motor patterns in pyloric network of the lobster, Panulirus interruptus. *Journal of Neurophysiology* **48**:914–937. doi:10.1152/jn.1982.48.4.914

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682. doi:10.1038/nmeth.2019

Schwindt P, Crill W. 1980. Role of a persistent inward current in motoneuron bursting during spinal seizures. *Journal of Neurophysiology* **43**:1296–1318. doi:10.1152/jn.1980.43.5.1296

Sengupta M. 2015. Purkinje neurons in action: From single cells to ensembles. Mumbai: Tata Institute of Fundamental Research.

Sengupta M, Thirumalai V. 2015. AMPA receptor mediated synaptic excitation drives state-dependent bursting in Purkinje neurons of zebrafish larvae. *Elife* **4**. doi:10.7554/eLife.09158

Thirumalai V, Jha U. 2022. RECRUITMENT OF MOTONEURONSVertebrate Motoneurons, In Press. Springer-Verlag.

Tian L, Hires SA, Mao T, Huber D, Chiappe ME, Chalasani SH, Petreanu L, Akerboom J, McKinney SA, Schreiter ER, Bargmann CI, Jayaraman V, Svoboda K, Looger LL. 2009. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat Methods* **6**:875–881. doi:10.1038/nmeth.1398

Urasaki A, Morvan G, Kawakami K. 2006. Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition. *Genetics* **174**:639–649. doi:10.1534/genetics.106.060244

Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R, Paninski L. 2010. Fast Nonnegative Deconvolution for Spike Train Inference From Population Calcium Imaging. *Journal of Neurophysiology* **104**:3691–3704. doi:10.1152/jn.01073.2009

Williams SR, Christensen SR, Stuart GJ, Häusser M. 2002. Membrane potential bistability is controlled by the hyperpolarization-activated current I(H) in rat cerebellar Purkinje neurons in vitro. *J Physiol (Lond)* **539**:469–483.

Yaksi E, Friedrich RW. 2006. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. *Nat Methods* **3**:377–383. doi:10.1038/nmeth874

Yang W, Yuste R. 2017. In vivo imaging of neural activity. *Nat Methods* **14**:349–359. doi:10.1038/nmeth.4230

Zeldenrust F, Wadman WJ, Englitz B. 2018. Neural Coding With Bursts—Current State and Future Perspectives. *Frontiers in Computational Neuroscience* **12**.