

Imperial College London
Department of Electrical and Electronic Engineering

**Detecting Cells and Cellular Activity from
Two-Photon Calcium Imaging Data**

Stephanie Reynolds

Supervised by Pier Luigi Dragotti and Simon Schultz

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy of Imperial College London

Declaration of Originality

I declare that this thesis and the research it contains are the product of my own work under the guidance of my thesis supervisors, Professors Pier Luigi Dragotti and Simon Schultz. All material that is not my own work has been properly acknowledged.

Stephanie Reynolds

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

To understand how networks of neurons process information, it is essential to monitor their activity in living tissue. Information is transmitted between neurons by electrochemical impulses called action potentials or spikes. Calcium-sensitive fluorescent probes, which emit a characteristic pulse of fluorescence in response to a spike, are used to visualise spiking activity. Combined with two-photon microscopy, they enable the spiking activity of thousands of neurons to be monitored simultaneously at single-cell and single-spike resolution. In this thesis, we develop signal processing tools for detecting cells and cellular activity from two-photon calcium imaging data.

Firstly, we present a method to detect the locations of cells within a video. In our framework, an active contour evolves guided by a model-based cost function to identify a cell boundary. We demonstrate that this method, which includes no assumptions about typical cell shape or temporal activity, is able to detect cells with varied properties from real imaging data.

Once the location of a cell has been identified, its spiking activity must be inferred from the fluorescence signal. We present a metric that quantifies the similarity between inferred spikes and the ground truth. The proposed metric assesses the similarity of pulse trains obtained from convolution of the spike trains with a smoothing pulse, whose width is derived from the statistics of the data. We demonstrate that the proposed metric is more sensitive than existing metrics to the temporal and rate precision of inferred spike trains.

Finally, we extend an existing framework for spike inference to accommodate a wider class of fluorescence signals. Our method, which is based on finite rate of innovation theory, exploits the known parametric structure of the signal to infer the unknown spike times. On *in vitro* imaging data, we demonstrate that the updated algorithm outperforms a state of the art approach.

Acknowledgements

First and foremost, I would like to thank my primary supervisor, Pier Luigi Dragotti. When I first decided to pursue a PhD, countless older and wiser people told me that the most important factor is finding the right supervisor and, four years later, I find that I agree. I am grateful for his clear guidance, the many opportunities he gave me to teach, learn and attend conferences, and his understanding. I would like to thank my second supervisor Simon Schultz, whose infectious enthusiasm made him great fun to work alongside. I am grateful for the all the support and advice he has given me during this time.

I would like to thank our collaborators Per Jesper Sjöström, Therese Abrahamsson and Renaud Schuck, who collected data that was essential to this thesis. I have learnt a lot from interactions with members of the Schultz lab, thanks in particular to Amanda Foust, Caroline Copeland and Peter Quicke. I am also thankful to my examiners, Timothy Constandinou and Dimitri Van De Ville for their valuable input regarding improvements to my thesis.

I am grateful to my friends in the C&SP group for brightening up the more difficult PhD days. I will miss the fiery lunchtime debates, office gossips, coffees and cocktails. Thanks to you all for being such great company. Special thanks go to Tricia, for being my organising partner and for keeping the more rowdy members of the group in check; to John, whose big smile and words of wisdom made him an awesome desk neighbour; to Alex and Michael, who shared the journey with me; and Charlie, Doris, Ilia, Sithan and Wilhelm, for the good times.

Finally, I would like to express my gratitude to my parents for their endless support. To my friends and family for reminding me what is important. And to Leo, for listening to every presentation, sharing both my difficulties and my successes, and for making me smile every day.

Contents

1	Introduction	23
1.1	Original contribution and outline of thesis	24
1.2	Publications	25
2	Background	27
2.1	Neuronal information processing	27
2.1.1	The action potential	28
2.2	Monitoring neuronal activity at cellular resolution	29
2.3	Calcium imaging	31
2.3.1	Signal model	33
2.4	Two-photon microscopy	34
2.4.1	Fluorescence excitation	34
2.4.2	Scattering and absorption in biological tissue	36
2.4.3	Fluorescence collection	36
2.4.4	Shot noise	38
2.4.5	Neuropil contamination	39
2.5	Outlook	40

3	ABLE: an activity-based level set segmentation algorithm	43
3.1	Introduction	44
3.2	Active contours	46
3.2.1	External energy	48
3.3	Level set method	49
3.3.1	External velocity	50
3.3.2	Regularisation	51
3.4	Proposed method	53
3.4.1	Dissimilarity metrics	55
3.4.2	Neighbouring cells	55
3.4.3	Implementation	57
3.5	Experimental methods	59
3.5.1	Simulations	59
3.5.2	Two-photon calcium imaging of hippocampal slices	60
3.5.3	In vivo imaging datasets	61
3.5.4	Performance metrics	61
3.6	Results	62
3.6.1	Robustness to heterogeneity in cell properties	62
3.6.2	Demixing overlapping cells	63
3.6.3	Results on hippocampal brain slices	65
3.6.4	Algorithm comparison on manually labelled dataset	65
3.7	Summary	67

4	CosMIC: A consistent metric for spike inference from calcium imaging	69
4.1	Introduction	69
4.2	Fuzzy set theory	72
4.2.1	Sørensen-Dice coefficient	74
4.3	Constructing the metric	76
4.3.1	Ancestor metrics	79
4.4	Temporal error tolerance	80
4.4.1	Parameter estimation	80
4.4.2	Cramér-Rao bound	81
4.4.3	Bound for spike inference	82
4.4.4	Pulse width	85
4.5	Experimental methods	85
4.5.1	Success rate	86
4.5.2	Spike train correlation	87
4.5.3	Two-photon calcium imaging of cortical slices	87
4.6	Results	88
4.6.1	CosMIC penalises overestimation	88
4.6.2	CosMIC rewards high temporal precision	90
4.6.3	Application to real imaging data	92
4.6.4	CosMIC discriminates precision and fall-out rate of spike trains	94
4.7	Summary	95

5	Spike inference using Finite Rate of Innovation theory	97
5.1	Introduction	98
5.2	Sampling theory	100
5.2.1	Signals with finite rate of innovation	101
5.2.2	Sampling kernels	102
5.2.3	Reconstruction of a stream of Diracs	104
5.3	Spike inference from calcium imaging data	106
5.3.1	Mapping samples to the FRI setting	108
5.3.2	Amplitude estimation	111
5.4	Spike inference from noisy data	112
5.4.1	Pre-whitening	113
5.4.2	Model order estimation	115
5.4.3	Spike inference algorithm	116
5.5	Experimental methods	117
5.5.1	Simulated data	117
5.6	Results	118
5.6.1	Comparison with instantaneous-rise algorithm	118
5.6.2	Model order estimation	119
5.6.3	Comparison with state of the art	121
5.7	Summary	122
6	Conclusion	123
6.1	Summary	123
6.2	Future research	124

A	ABLE: supplemental analytical results	127
A.1	External velocity	127
A.2	Euclidean dissimilarity metric	129
B	CosMIC: proof of analytical results	131
	Bibliography	135

List of Figures

2.1	Example of a whole-cell patch-clamp recording.	29
2.2	Model of the neuronal fluorescence signal.	32
2.3	Procedure of fluorescence excitation by a two-photon microscope.	35
2.4	Example of in vivo two-photon calcium imaging data.	37
2.5	Illustration of neuropil contamination of cellular signal.	40
3.1	Example of an active contour segmenting a piecewise smooth image.	47
3.2	The level set function.	49
3.3	The Chan-Vese active contours model.	52
3.4	Flow diagram of the proposed segmentation algorithm.	54
3.5	Evolution of multiple active contours.	57
3.6	Segmentation results on in vivo imaging data.	62
3.7	Segmentation results on simulated data containing overlapping cells.	63
3.8	Segmentation results on in vitro imaging data containing synchronously spiking cells.	64
3.9	Algorithm comparison on a manually labelled dataset.	66
4.1	Inconsistency in usage of spike inference metrics.	70
4.2	Membership functions of the union and intersection of fuzzy sets.	73

4.3	Flow diagram of the proposed metric.	77
4.4	Proposed metric example.	78
4.5	Metric pulse width.	84
4.6	Illustration of common metrics	86
4.7	Results on spike train estimates with varying overestimation ratio.	89
4.8	Results on spike train estimates with varying temporal precision.	91
4.9	Results on real imaging data	93
4.10	Results on spike train estimates with varying precision and fall-out rate.	96
5.1	Example of a slow-rise fluorescence signal.	99
5.2	Example of signals with finite rate of innovation.	102
5.3	Exponential reproduction with an E-spline of order 2.	103
5.4	Example of sampling and reconstruction of a stream of Diracs.	105
5.5	Example of spike inference with the FRI algorithm.	109
5.6	Model order estimation procedure.	114
5.7	Example of simulated data.	117
5.8	Comparison of the temporal error of the FRI algorithm with instantaneous-rise and slow-rise models.	119
5.9	Accuracy of the model order estimation procedure on simulated data.	120
5.10	Algorithm comparison on real data	122

List of Tables

2.1	Time constants of four commonly used calcium indicators.	34
3.1	Runtime of the segmentation algorithm (minutes) on synthetic data. . .	58
3.2	Algorithm success rate on a manually labelled dataset.	67
5.1	Amplitude of simulated calcium transients.	118

Nomenclature

Acronyms

2PCI	Two-photon calcium imaging
AP	Action potential
CRB	Cramér-Rao bound
FRI	Finite rate of innovation
GECI	Genetically encoded calcium indicator
i.i.d.	Independent identically distributed
ICC	Intracellular calcium concentration
LSF	Level set function
MSE	Mean square error
PDF	Probability density function
RMSE	Root mean square error
ROI	Region of interest
SDC	Sørensen-Dice coefficient
SNR	Signal-to-noise ratio
STC	Spike train correlation

Number Sets

\mathbb{R}	Real numbers
--------------	--------------

$\mathbb{R}_{>0}$	Positive real numbers
\mathbb{Z}	Integers
\mathbb{C}	Complex numbers: $x + iy$ for $x, y \in \mathbb{R}$

Variables

\mathbf{g}	Vector: $\mathbf{g} \in \mathbb{C}^m$ with $m > 1$
\mathbf{G}	Matrix: $\mathbf{G} \in \mathbb{C}^{m \times n}$ with $m, n > 1$
$g(t)$	Continuous-time signal
$g[n]$	Discrete-time signal

Functions

$g : A \rightarrow B$	A function, g , defined on domain A with image a subset of B
$\langle g_1(t), g_2(t) \rangle$	Inner product: $\int_{\mathbb{R}} g_1(t)g_2^*(t) dt$
$g_1(t) * g_2(t)$	Continuous-time convolution: $\int_{\mathbb{R}} g_1(\tau)g_2(t - \tau) d\tau$
$g_1[n] * g_2[n]$	Discrete-time convolution: $\sum_{m \in \mathbb{Z}} g_1[m]g_2[n - m]$
$\mathbb{1}_A(t)$	Indicator function, equal to one if $t \in A$ and 0 otherwise
$\mathbb{1}_{t>a}$	Indicator function of set $A = \{t \in \mathbb{R} : t > a\}$
$\delta(t)$	Dirac delta function
$\delta[n]$	Discrete Dirac delta function
\mathbf{G}^H	Hermitian (conjugate transpose) of matrix \mathbf{G}
$H(x)$	Heaviside function

Probability

$\mathbb{E}[U]$	Expectation of random variable U
$\mathbb{P}(A)$	Probability of event A
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean μ and standard deviation σ

Norms

$|a|$ Absolute value of a real or complex number a

$\|\mathbf{g}\|$ Euclidean norm of vector \mathbf{g} : $\sqrt{g[1]^2 + g[2]^2 + \dots + g[m]^2}$

$\|g(t)\|_1$ L1-norm of continuous-time function $g(t)$: $\int_{\mathbb{R}} |g(t)| dt$

$|A|_c$ Cardinality of a fuzzy set

Asymptotic notation

$O(h(x))$ Referred to as ‘Big-O’ notation, this indicates the set of functions, such that

$$\{g(x) : \exists c, x_0 \in \mathbb{R}_{>0} \text{ such that } 0 \leq g(x) < ch(x) \forall x \geq x_0\}$$

Chapter 1.

Introduction

The human brain is a remarkably complex organ responsible for a vast array of tasks. It is believed to consist of approximately 100 billion neurons, whose activity is organised in hierarchical networks. Information is sent between neurons by electrochemical impulses called action potentials (APs) or spikes. Coordinated spiking activity across multiple spatial scales allows the brain to, for example, process incoming sensory information, encode memories and initiate motor activities.

To understand how networks of neurons process information, it is essential to observe brain activity in living tissue. Crucially, this can provide insight into the difference between healthy and unhealthy brains. For example, functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) have been successfully used to identify impairment in connectivity in patients with brain disorders [17]. Such findings can have clinical relevance both as a diagnostic tool and as a focus for new treatments.

While fMRI and EEG are effective for assessing larger-scale connectivity, the study of microcircuit activity has been limited by the challenges of recording at cellular resolution in living tissue. Traditionally, electrophysiological approaches have been used to monitor a cell's spiking activity. These approaches are, however, limited to sampling from a small number of unidentifiable cells. As an alternative, fluorescent sensors can be used to visualise cellular activity via changes in their fluorescence intensity. One such approach is calcium imaging, whereby calcium-sensitive probes monitor the levels of intracellular calcium, which, in turn, reflects a cell's spiking activity.

To perform calcium imaging, a light source is required that excites fluorescence from the probes. In lightsheet microscopy (LSM), for example, fluorescence is excited from

all neurons in a thin sheet of tissue simultaneously [1, 12]. This technology can be used to monitor neuronal activity at high speeds in transparent organisms, such as the larval zebrafish. When imaging through light scattering media, however, the spatial resolution of LSM suffers. In contrast, two-photon microscopy overcomes light scattering through the usage of a highly localised focal spot that is scanned across the tissue to generate an image. Due to its relatively high spatial resolution, this technology is unrivalled in its ability to image hundreds of microns deep in living tissue. State of the art two photon calcium imaging is able to monitor the activity of thousands of cells simultaneously in living animals [124, 100]. The prospects for this technology are thus vast; it has already been used, for example, to study Alzheimer’s disease in animal models [112] and to determine the spatial organisation of receptive fields [86, 10].

Before neuronal activity from two-photon calcium imaging (2PCI) videos can be analysed, two challenging signal processing problems must be tackled. Firstly, the locations of cells within the video must be identified. This task is complicated by heterogeneity in cellular properties, such as their shape and typical temporal activity, and the frequency of cellular overlap. Secondly, one must recover a cell’s spiking activity from its fluorescence signal. This problem, which we refer to as spike inference, is an inverse problem that can be solved by considering the known parametric structure of the fluorescence signal. In this thesis, we focus on these problems and outline our contributions below.

1.1 Original contribution and outline of thesis

The remainder of the thesis proceeds as follows.

In Chapter 2, we provide relevant background information on 2PCI data. We discuss how the data is generated and outline the sources of noise.

In Chapter 3, we present a novel segmentation algorithm for detecting the location of cells from 2PCI data. In our framework, multiple coupled active contours evolve, guided by a model based cost function, to identify cell boundaries. On in vivo imaging data, we demonstrate the versatility of the algorithm, which includes no priors on a cell’s stereotypical morphology or temporal activity. We also demonstrate the robustness of the algorithm on challenging datasets exhibiting highly correlated cellular activity and overlapping cells, respectively.

In Chapter 4, we consider the metrics that are used to compare the locations of inferred spikes with ground truth spiking activity. We highlight the limitations of existing metrics and propose an alternative. Rather than operating on the true and estimated spike trains directly, the proposed metric assesses the similarity of the pulse trains obtained from convolution of the spike trains with a smoothing pulse, whose width is derived from the statistics of the imaging data. On real and simulated data, we demonstrate the sensitivity of the proposed metric to the temporal and rate precision of inferred spike trains.

In previous work, it was established that the neuronal fluorescence signals that are induced by 2PCI belong to the class of signals with finite rate of innovation (FRI). Using FRI theory and a simplified model of the fluorescence signal, an algorithm was developed that inferred the locations of spikes with high accuracy [87]. In Chapter 5, we extend this framework to encompass a broader class of fluorescence signals, which are not well-matched by the original framework. We also introduce procedures to improve the algorithm's robustness to noise. Using the spike inference metric presented in Chapter 4, we demonstrate that the proposed algorithm outperforms a state of the art approach on in vitro imaging data.

Finally, in Chapter 6 we summarise the contributions of the thesis and discuss some possible directions for future research.

1.2 Publications

The materials presented in this thesis have led to the following publications:

Peer-reviewed journals

S. Reynolds, T. Abrahamsson, P. J. Sjöström, S. R. Schultz, and P. L. Dragotti. CosMIC: A consistent metric for spike inference from calcium imaging, *Neural Computation*, in press.

S. Reynolds, T. Abrahamsson, R. Schuck, P. J. Sjöström, S. R. Schultz, and P. L. Dragotti. ABLE: An activity-based level set segmentation algorithm for two-photon calcium imaging data, *eNeuro*, 4(5):ENEURO.0012-17.2017, 2017.

R. Schuck, M. A. Go, S. Garasto, S. Reynolds, P. L. Dragotti, and S. R. Schultz. Multiphoton minimal inertia scanning for fast acquisition of neural activity signals, *Journal of Neural Engineering*, 15(2):025003, 2018.

Peer-reviewed conferences

S. Reynolds, C. S. Copeland, S. R. Schultz, and P. L. Dragotti. An extension of the FRI framework for calcium transient detection. In *13th International Symposium on Biomedical Imaging (ISBI)*, pages 676 – 679, 2016. IEEE.

S. Reynolds, J. Oñativia, C. S. Copeland, S. R. Schultz, and P. L. Dragotti. Spike detection using FRI methods and protein calcium sensors: performance analysis and comparisons. In *11th international conference on Sampling Theory and Applications (SampTA)*, pages 533 – 537, 2015. IEEE.

Chapter 2.

Background

To understand function and dysfunction in the mammalian brain it is crucial to observe its components in action. In particular, long-term studies of brain activity in living animals ('in vivo') could be key to identifying how the brain encodes stimuli or orchestrates behaviours. Whilst there are many viable techniques for observing brain activity at larger spatial scales, two-photon calcium imaging is unparalleled in its ability to monitor cellular activity, deep in living tissue, at single-cell resolution. To emphasise the importance of this technology, in Section 2.1, we discuss some basic aspects of neuroscience. In Section 2.2, we discuss the relative merits of existing techniques to monitor neuronal activity at single-cell resolution. Finally, in Sections 2.3 and 2.4, we provide further detail on calcium imaging and two-photon microscopy, respectively.

2.1 Neuronal information processing

The mammalian cortex is responsible for a wide array of tasks, such as motor control, integration of sensory inputs and decision making [151]. It is believed that communication occurs across multiple spatial scales, with local networks of interconnecting neurons constituting microcircuits whose activity is aggregated in larger macrocircuits spread across multiple brain areas. Microcircuits, which consist of 1000 - 10,000 neurons, are hypothesised to be computational units that undertake specific information processing tasks [81]. For example, one 'canonical' microcircuit has been observed to repeat throughout the cortex and is presumed to be responsible for fundamental

computations [37]. Despite this apparent hierarchical organisation, local networks are heavily interconnected with other brain areas.

To understand information processing it is, therefore, necessary to monitor network activity at multiple spatial scales. At the macroscopic scale, technologies such as fMRI can be used to study coordinated activity across brain areas. At the microscopic scale, cellular resolution functional imaging has the potential to elucidate the dynamics of microcircuits. The action potential (AP), an electro-chemical impulse propagated between connected neurons, is the core method of information transmission between neurons [70]. It is not yet fully understood how neuronal networks encode information through coordinated spiking activity. Accordingly, many technologies are aimed at monitoring neuronal population activity at single-cell and single-spike resolution.

2.1.1 The action potential

Inputs from other neurons or sensory organs can excite a neuron and cause it to fire an electro-chemical impulse called an action potential. The resting potential of a neuron, which reflects the difference in ion concentrations inside and outside the cell, is approximately -70mV . Inputs from other neurons arrive in the form of inward current injections, causing the neuron's membrane potential to become less negative. This is referred to as depolarisation. If the stimulus is sufficiently large to depolarise the neuron above a (neuron-dependent) threshold, an AP is fired. Over a period of approximately 1ms , the membrane potential rapidly rises to $+30\text{mV}$, subsequently falling back to the resting potential. This rapid rise in membrane potential is referred to as an AP, or spike. An AP is thought to be an all-or-nothing event: sub-threshold inputs do not illicit an AP and all supra-threshold inputs illicit APs of the same amplitude. So, although sub-threshold inputs can be important in identifying connections between neurons, it is thought that APs are key to neural computation. A neuron's activity can therefore be characterised by the set of times at which that neuron fired a spike: $\{t_k\}_{k=1}^K$ with all $t_k \in \mathbb{R}$. This activity can be represented as a spike train

$$x(t) = \sum_{k=1}^K \delta(t - t_k), \quad (2.1)$$

where $\delta(\cdot)$ is the Dirac delta function.

Individual neurons encode information through the rate and timing of the spikes they

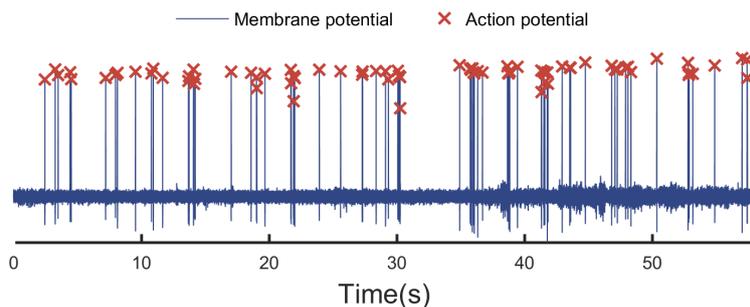


Figure 2.1: Example of a whole-cell patch-clamp recording. This recording, from an online dataset [132, 26], was sampled at 10kHz. Action potentials are visible as peaks in the membrane potential.

fire [70]. For example, the firing rate of motor neurons is related to the force of contraction of associated muscle fibres. In contrast, visual cortex neurons encode information about visual stimuli using precise timing of individual spikes [66]. In the cat visual system, temporal coding has been observed on the millisecond time-scale. To decipher the neural code, it is therefore necessary to record from the brain at sufficient temporal resolution to detect individual spikes.

2.2 Monitoring neuronal activity at cellular resolution

To enable analysis of neuronal microcircuit activity, recording techniques aim to maximise the number of neurons that can be monitored at single-cell and single-spike resolution. The patch-clamp technique is the current gold standard for monitoring the voltage of a single cell [50, 115] and is commonly used as a benchmark for new technologies. To obtain a patch-clamp recording, an electrode is inserted into brain tissue and attached to a cell membrane. This configuration allows the cell’s voltage to be recorded at a sampling rate on the order of 10kHz, which is sufficient to resolve single APs (Fig. 2.1). As one electrode is required per cell, the patch-clamp technique is limited by the number of neurons that can be monitored simultaneously. Although state of the art implementations can record from up to 8 neurons at the same time [16, 150], this is not sufficient for analyses of population activity.

In contrast to the patch-clamp technique, which monitors the intracellular voltage signal, extracellular approaches take advantage of the fact that an AP can be detected from voltage fluctuations outside of the cell body. Instead of pairing one electrode

with one cell, these techniques insert electrodes with multiple recording sites into brain tissue [18]. Each recording site detects transmembrane potentials, which are indicative of APs, from several cells. Activity must then be assigned to the cell from which it originated by a process known as spike sorting. At each site, spikes may stem from neurons as far as $150\text{-}200\mu\text{m}$ away (a neuron's cell body has diameter $10\mu\text{m}$), although typically only activity from cells within $50\mu\text{m}$ will have sufficient amplitude to be detected [51]. Despite their relatively high temporal resolution and signal-to-noise ratio (SNR), extracellular electrophysiological approaches are not well-suited to long-term studies, which are necessary to understand learning, memory and plasticity in the living brain. This limitation is due to the instability in the monitored neuronal populations. In one study, statistical analysis of spike waveforms and distribution was used to show that the number of units recording consistently from the same neuronal populations fell to 57% over one week [33]. In contrast, fluorescence imaging, which can provide both functional and anatomical information, can be used to monitor the same neuronal populations over weeks and months [98].

Whereas an electrode must be in contact with the tissue that it monitors, fluorescence imaging can be performed non-invasively to monitor cellular activity *in vivo*. In particular, functional imaging utilises sensors in the brain whose fluorescence intensity varies in response to cellular activity. One option is to use voltage-sensitive fluorescent indicators [101], which visualise changes in a cell's membrane potential. Due to the fast time scale of APs (on the order of 1ms), to visualise cellular activity with a voltage sensor it is necessary to have sub-millisecond temporal resolution, which, in turn, limits the signal-to-noise ratio of the acquired signals. This limitation means that, in practice, voltage imaging is mostly used to perform widefield imaging of cellular aggregates [71]. Alternatively, a secondary variable can be imaged, from which spiking activity can be inferred. In this thesis, we focus on calcium imaging data, which uses fluorescent sensors to visualise a cell's intracellular calcium concentration (ICC). An AP causes the ICC to rise to levels that are 10 to 100 times the levels at baseline [9], this high dynamic range makes the ICC a favourable signal to monitor. The duration of the ICC elevation after an AP is an order of magnitude slower than the duration of an AP. As a consequence, cells can be monitored with calcium imaging at lower sampling rates whilst still capturing high-frequency spiking activity.

Fluorescence imaging approaches require a light source, with which to excite fluorescence from the sample (the biological tissue), and a detector, with which to collect the fluorescence. Approaches that excite fluorescence from a large volume simultaneously,

such as confocal [97] and lightsheet [1] microscopy, treat excited photons differently depending on how they arrive at the detector. Confocal microscopes reject all photons outside of a detection pinhole, thereby ideally only collecting photons from a localised region. In contrast, lightsheet microscopes use detector arrays, such as CMOS cameras. Each detector represents a pixel in the output image and, in theory, collects excited photons from a localised region of the sample. While appropriate for transparent specimens or superficial layers of tissue, these approaches are not well-suited to imaging through light-scattering media. Changes in the refractive index, which can occur every nanometre in the mammalian brain [151], divert light from its path. This phenomenon is referred to as scattering (see Section 2.4.2). It is hard to determine the source in the sample from which scattered photons originate, which complicates the use of detection pinholes and camera arrays. Two-photon microscopy, on the other hand, uses a highly localised laser to excite fluorescence [31]. This allows all photons emitted from the sample to be collected by a single detector, regardless of the angle at which they arrive at the detector, as they reliably stem from a small volume.

2.3 Calcium imaging

Calcium ions are a crucial component in many aspects of cellular function [9]. For example, they trigger the contraction of heart muscle cells and regulate both cell death and proliferation. In neurons, the ICC is tightly linked to spiking activity. When an AP is fired in a neuron, the depolarisation causes voltage-gated calcium ion channels to open. As a consequence, there is a brief influx and a subsequent slower efflux of calcium ions. This transient elevation in the ICC is referred to as a calcium transient. Compared to the rapid timescale of an AP, a calcium transient is relatively slow; the calcium ion influx lasts approximately 1ms [11] and the efflux resembles an exponential decay with time constant 0.5s - 1s [46].

The development of calcium-sensitive fluorescent indicators ('calcium indicators') has enabled the visualisation of the ICC and, indirectly, neuronal spiking activity. The indicators consist of a fluorophore and a domain to which calcium ions can bind. Calcium binding alters the properties of the indicator and, in turn, elicits a stereotypical change in the fluorescence, the nature of which varies depending on the indicator used [14]. Calcium indicators can be categorised into two groups: synthetic indicators and genetically encoded indicators (GECIs). The former group, which has been in devel-

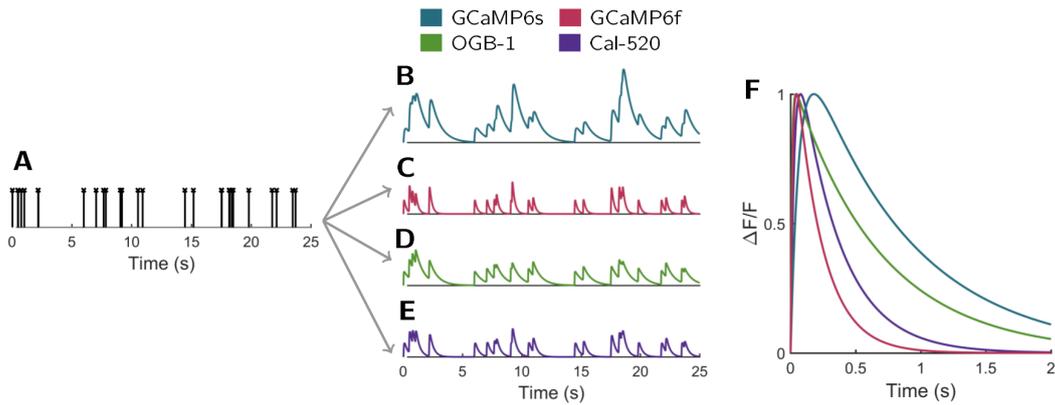


Figure 2.2: Model of the neuronal fluorescence signal. The fluorescence emitted from a neuron can be modelled as that neuron’s spike train convolved with a characteristic pulse. The rise and decay rate of the pulse varies with the fluorescent indicator used to generate the imaging data (**F**). We display simulated fluorescence traces generated from the spiking activity in **A** and the characteristic pulse shapes of the indicators GCaMP6s (**B**), GCaMP6f (**C**), OGB-1 (**D**) and Cal-520 (**E**).

opment since the early 1980s [138], can be used to label large cell populations in intact tissue through bolus injection [130]. They are, however, limited to acute imaging experiments (on the order of hours) [71]. GECIs, on the other hand, can be encoded into the DNA of transgenic animals [80, 2, 26]. In that case, they are naturally expressed in neurons and can be targeted to specific cell types and subcellular compartments. This makes GECIs ideal for chronic imaging experiments over multiple weeks and months [55]. Moreover, recent advances have produced GECIs whose sensitivity and signal quality surpasses those of synthetic indicators [26], which was not the case as recently as 2009 [136].

In the following, we introduce a mathematical model relating the fluorescence intensity of somatic calcium indicators to spiking activity. The mathematical model introduced below is prevalent in the literature, in particular it is commonly used by algorithms that infer spike timing from calcium imaging data [148, 147, 30, 73, 87, 109, 105]. For a more comprehensive coverage of the biophysical processes underlying the mathematical model, the reader is referred to [14].

2.3.1 Signal model

When an AP is fired from a neuron, there is a rapid influx and subsequent slower efflux of calcium ions, which is referred to as a calcium transient. This calcium response is modelled with a rapid rise and an exponential decay. A corresponding fluorescence pulse is emitted from the calcium-sensitive fluorescent indicators in the neuron. The pulse emitted as a result of an AP at time $t = 0$ is proportional to

$$p(t) = c \left(1 - e^{-t/\tau_{\text{on}}}\right) e^{-t/\tau_{\text{off}}} \mathbb{1}_{t>0}, \quad (2.2)$$

where τ_{on} and τ_{off} are parameters which define the rate of the rise and decay, respectively, and c is a normalisation constant that ensures that $p(t)$ has peak value equal to 1.

The rise time is predominantly influenced by the rate at which ions bind to the fluorescent indicator (the association rate). The rate at which the fluorescence pulse decays is affected by cellular properties, such as the extrusion rate of calcium channels, and properties of the indicator, including the association and disassociation rates. In Table 2.1, we indicate values of τ_{on} and τ_{off} obtained from the literature for the following indicators: Cal-520 AM [133], OGB-1 AM [73], GCaMP6f and GCaMP6s [26]. On new data generated by those indicators, these values can be used as a guideline. The corresponding characteristic pulses are plotted in Fig. 2.2F.

Each time a neuron fires an AP, a stereotypical pulse of fluorescence is emitted. Writing a_k as the amplitude of the pulse emitted in response to an AP at time t_k , the fluorescence signal, $f : \mathbb{R} \rightarrow \mathbb{R}$, is

$$f(t) = b(t) + \left(\sum_{k=1}^K a_k \delta(t - t_k) \right) * p(t) \quad (2.3)$$

$$= b(t) + c \sum_{k=1}^K a_k \left(1 - e^{-(t-t_k)/\tau_{\text{on}}}\right) e^{-(t-t_k)/\tau_{\text{off}}} \mathbb{1}_{t>t_k}, \quad (2.4)$$

where $b(t)$ reflects the baseline fluorescence. The baseline component is typically modelled as either a low-frequency drift [30, 106] or a piecewise constant function [44, 104]. This can reflect ICC changes in response to other cellular processes [48] or a steady decline in baseline fluorescence due to bleaching of the fluorescent indicator.

The amplitude of calcium transients varies with the local spike rate of neurons. The

Fluorescent indicator	τ_{on} (ms)	τ_{off} (ms)
GCaMP6f	18	205
GCaMP6s	72	794
OGB-1	10	667
Cal-520	32	314

Table 2.1: Time constants of four commonly used calcium indicators. The pulse shape is defined by two parameters, τ_{on} and τ_{off} , which represent the characteristic delays. These values can be used as a guideline for each calcium indicator; in practice, they will vary with the indicator expression level as well as the cell type.

co-operative calcium binding of some GECIs [14] results in increasing amplitudes as the spike rate increases [2, 26]. In contrast, synthetic indicators are known to saturate at high spike rates [64], producing spikes of decreasing amplitude. The amplitude of spikes can also steadily decrease as a result of indicator bleaching. In Fig. 2.2, we display simulated fluorescence signals generated by the same spiking activity for four common calcium indicators.

2.4 Two-photon microscopy

Two-photon microscopy is a powerful technique for imaging deep in light-scattering media. A beam of light is used to excite fluorescence from a concentrated focal spot in the sample, this process is discussed in Section 2.4.1. The effectiveness of the excitation of fluorescence in biological tissues is affected by scattering and absorption, which is discussed in Section 2.4.2. Finally, in Section 2.4.3, the acquisition process of a two-photon microscope image is described and sources of noise are discussed in Sections 2.4.4 and 2.4.5.

2.4.1 Fluorescence excitation

Fluorescence microscopy relies on the use of fluorophores (fluorescent molecules) that can be excited by the absorption of photons. When a fluorophore absorbs a photon of sufficient energy, it transitions from a ground state to an excited state. The fluorophore subsequently relaxes to the ground state by either a radiative or non-radiative process. In the latter case, light energy is absorbed as thermal energy. In the former, a photon

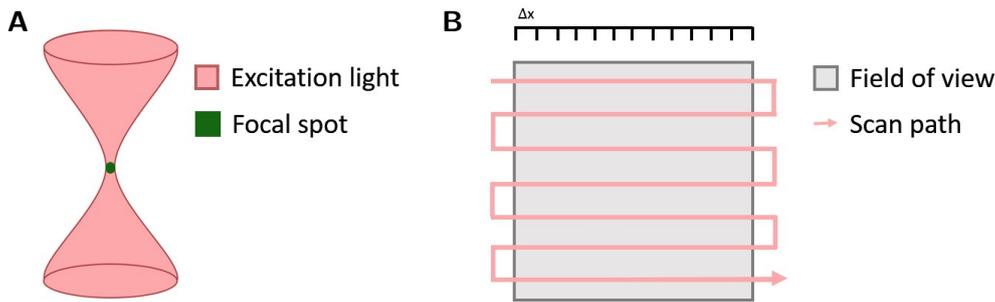


Figure 2.3: Procedure of fluorescence excitation by a two-photon microscope. A beam of light is focused through the microscope objective such that fluorescence is only excited in the vicinity of the focal spot (**A**). The focal spot is scanned sequentially through the field of view to generate one frame of an image (**B**). The beam moves discretely along the scan path with increment Δx , spending a fixed time at each point.

is emitted by the fluorophore. This process, which occurs on the nanosecond timescale, is known as fluorescence.

In standard fluorescence microscopy, a single photon is absorbed by the fluorophore, some energy is transformed into thermal energy and a lower-energy photon is subsequently emitted. In two-photon microscopy, excitation only occurs when two photons have been absorbed essentially simultaneously. With the combined energies of the two absorbed photons, the fluorophore transitions into an excited state. To relax from this state, some energy is absorbed as thermal energy and a higher-energy photon is emitted. Typically, the excitation beam has wavelength in the near-infrared range and the light emitted from the fluorophore is in the visible spectral range [151].

A two-photon excitation event can only occur when two photons arrive at the same molecule within approximately 0.5 femtoseconds of each other [52]. To enable this event, the photons in the excitation light beam must be highly concentrated both spatially and temporally. To concentrate photons temporally, the excitation light beams consist of ultrashort pulses (less than a picosecond in length) with high peak intensities [52]. Spatial density is achieved by focusing the excitation light through a microscope objective with high numerical aperture. As a result, fluorescence is only excited in a highly localised focal spot (Fig. 2.3A). Indeed, the probability of excitation scales with the square of the light intensity, meaning that the excitation decays sharply both axially and laterally.

2.4.2 Scattering and absorption in biological tissue

The scattering and absorption of light in biological tissue have a large impact on the signal generated in fluorescence microscopy. Biological tissues contain numerous endogenous molecules that absorb light of different wavelengths [151]. Before photons can reach the focal spot, they may be absorbed by these molecules. Although this absorption does attenuate the light at the focal spot, the greater impact stems from the associated tissue damage. This damage, which limits the viability of confocal microscopy for long-term imaging of biological tissue, is mitigated in two-photon microscopy [128]. In the latter case, the localised focal spot spares the majority of the tissue from high-intensity light. Moreover, the excitation beam uses near-infrared light, which is absorbed by few endogenous molecules [131].

A greater obstacle is presented by the scattering of light by biological tissue. Light is deflected from its original path as the refractive index of a medium changes. This phenomenon is referred to as scattering. Before light reaches its focal spot, it passes through multiple components, such as cells and blood vessels, each of which has a different refractive index. Even within these components, the refractive index may change on the nanometer scale [151]. Consequently, the intensity of non-scattered light decreases exponentially with penetration depth [52]. This, in turn, reduces the intensity of fluorescence excited at the focal spot, with greater effect as the imaging depth increases. In vivo fluorescence microscopy is thus fundamentally limited by scattering. As longer wavelength light is affected less in this regard, the near-infrared beam used in two-photon microscopy mitigates some of this effect. The maximum imaging depth, which depends on the sparseness and brightness of fluorophores, is hard to quantify. Two-photon microscopy is, however, sufficiently robust to scattering that it is routinely used to image hundreds of microns deep in tissue [59].

2.4.3 Fluorescence collection

As the spatial density of scattered photons is generally low, it can be assumed that they do not excite significant fluorescence. The majority of emitted photons, therefore, can be assumed to originate from the vicinity of the focal spot. Consequently, a two-photon microscope collects all emitted photons, regardless of the angle from which they appear to originate [52].

To generate an image, the focal spot is typically scanned in a raster pattern across

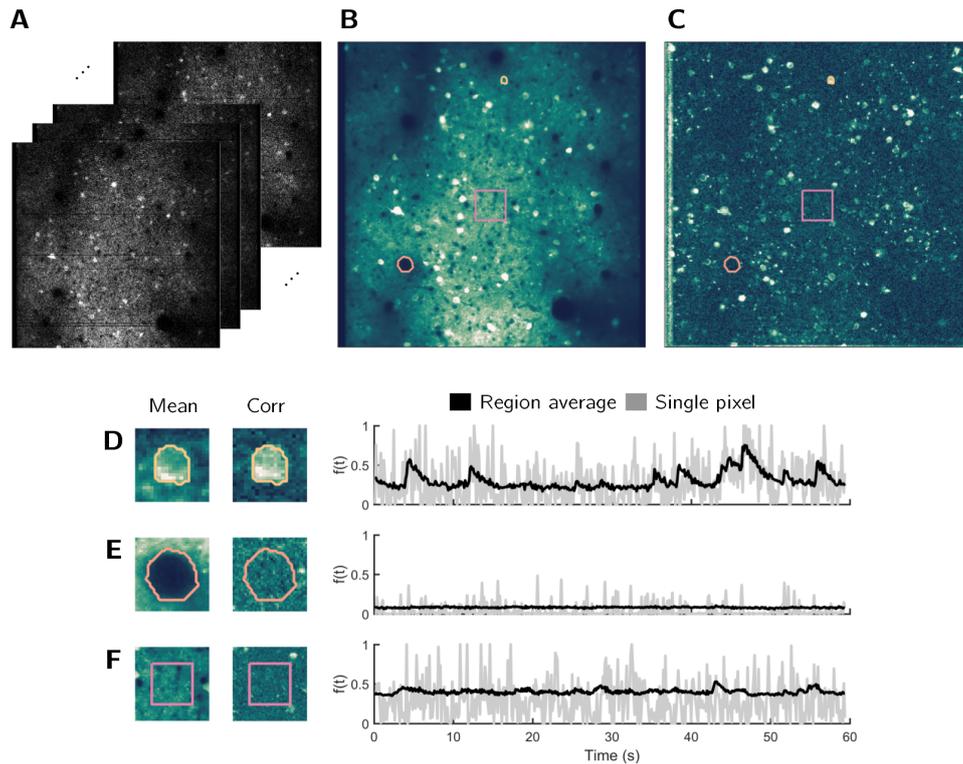


Figure 2.4: Example of in-vivo two-photon calcium imaging data. The excitation laser scans the field of view once to produce a single frame of the video (A). In the mean image, which displays the mean of each pixel’s activity, bright spots correspond to cells with high baseline fluorescence (B). In C, we display the correlation image, which, at each pixel, is computed as the mean of the correlation coefficients between that pixel’s temporal activity and the activity of neighbouring pixels. Bright spots in the correlation image correspond to active cells. In D, E and F, we focus on the three regions highlighted in the mean and correlation images. The regions visible in D, E and F correspond to a cell, blood vessel and background region, respectively. In each subfigure, we display the zoomed in mean and correlation images. We also plot the average temporal activity from all pixels in the interior of the object and the activity from a single pixel within the object. The data presented in this figure was collected by Svoboda et al. [99].

the field of view (Fig. 2.3B). The scanning is performed discretely; after dwelling Δt seconds in one location, the laser either moves Δx microns in the x-direction or starts a new row. Generally the field of view is square; the between-row spacing is equal to the within-row (column) spacing and there are the same number of rows as columns. A single image is generated by completing the scan path once. A video of N frames is

generated by completing the scan path N times, each completed scan path generates the data for one frame of the video (Fig. 2.4A).

The spatial increment, Δx , must be sufficiently small that single cells can be resolved. Meanwhile, the area covered — $(M\Delta x)^2$ with M the number of rows/columns — must be sufficiently large to contain enough cells for a meaningful analysis. To maximise the number of monitored cells, therefore, one would wish to maximise M . Similarly, the dwell-time at each spatial location, Δt , must be long enough to obtain a usable SNR (see Section 2.4.4). However, the sampling period (the temporal spacing between samples at the same spatial location), which is approximately equal to $M^2\Delta t$, must be sufficiently small that individual APs can be observed. To minimise the sampling period, therefore, the dwell-time or number of monitored cells must be compromised. Typically, Δx and Δt are in the range $1\text{-}5\mu\text{m}$ and $10\text{-}100\text{ms}$, respectively.

The trade-off between temporal resolution and other factors is lessened by the use of calcium-sensitive fluorescent probes, as calcium transients are significantly slower than the APs that induce them. To further loosen the constraints on the data, signal processing methods are required that are robust in low-noise settings. This will reduce the necessary dwell-time and further increase the number of cells that can be monitored at single-cell and single-spike resolution.

2.4.4 Shot noise

In the following, we introduce a mathematical model for the noise introduced during the acquisition process. This model relates to the statistical fluctuations in photon count and electron emission and is standard in the literature [97, 43, 27]. The photons emitted from the sample are detected by a photomultiplier tube, which is extremely sensitive to low levels of light. On detecting photons, a photocathode emits electrons, whose number is subsequently multiplied by a sequence of dynodes [97]. The pixel intensity value is obtained by analogue-to-digital conversion of current emitted from the final dynode.

As photons arrive independently at the photocathode, the number arriving during the acquisition period is modelled as a Poisson random variable. The number of electrons emitted in response to these photons, E_0 , can be shown to be Poisson distributed [27] with mean dependent on the quantum efficiency of the photocathode and the rate of photon arrival. At each subsequent dynode, the number of electrons is multiplied by

an amount dependent on an amplification factor, δ . The emission processes instigated by different incident electrons are physically independent [27]. As such, the number of electrons emitted from the i^{th} dynode, E_i , is a sum of E_{i-1} independent random variables, which model the output in response to the electrons emitted by the previous dynode. Given N_D dynodes in total, we have

$$E_i = \sum_{m=1}^{E_{i-1}} \eta_{i,m}, \quad (2.5)$$

for $i \in \{1, \dots, N_D\}$ and $\eta_{i,m} \sim \text{Poisson}(\delta)$.

It has been shown empirically that, when the rate of photon influx is not too low, the distribution of E_{N_D} can be approximated by a Gaussian, whose parameters are dependent upon the rate of photon influx, the amplification factor of the dynodes and the quantum efficiency of the photocathode. As the rate of photon influx varies at each pixel depending on the level of activity at the source, the statistics of the noise are not constant spatially. Rather, the noise is approximated as a heteroskedastic Gaussian. The signal at a pixel \mathbf{x} is written

$$\tilde{I}(\mathbf{x}) = I(\mathbf{x}) + \sigma(I(\mathbf{x}))\xi, \quad (2.6)$$

where $I(\mathbf{x})$ is the underlying, deterministic signal and $\xi \sim \mathcal{N}(0, 1)$. Assuming that the analogue-to-digital conversion process is linear, we have [27]

$$\sigma(I(\mathbf{x})) = aI(\mathbf{x}) + b, \quad (2.7)$$

i.e. the noise level at a pixel increases linearly with the level of activity at that pixel. In practice, this model is simplified by assuming that σ is constant at each pixel [147, 105, 106, 44, 127].

2.4.5 Neuropil contamination

As discussed in Section 2.4.3, in two-photon microscopy, fluorescence is excited from a localised focal spot within the sample. Photons collected at one point on the scan path stem from a small volume centred around this focal spot. The size and shape of this volume are characterised by the point spread function of the microscope [52]. Particularly when imaging at depth, as scattering of light becomes an increasing problem, the resolution of this focal spot suffers. At best, in-plane (lateral) resolution is

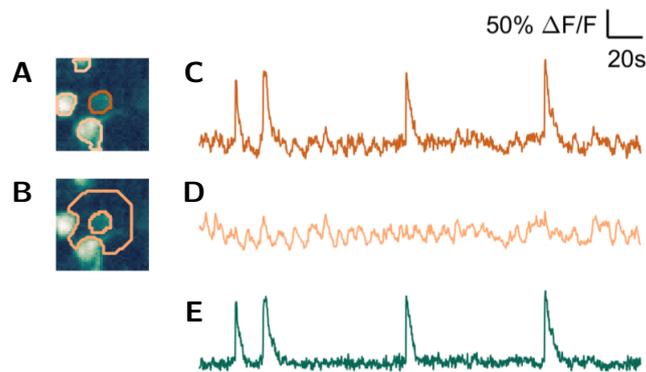


Figure 2.5: Illustration of neuropil contamination of cellular signal. The average temporal signal from pixels inside the cell delineated by the dark orange contour in **A** are plotted in **C**. The neighbouring pixels that do not belong to another cell (highlighted in **B**) exhibit neuropil activity (**D**), which originates from densely packed cellular processes. Due to the relatively low axial resolution of the focal volume, this activity is present in the cellular signal. Subtracting the weighted neuropil signal from the raw cellular signal removes this artefact (**E**). The data presented in this figure was collected by Svoboda et al. [99].

sub-micron and axial (perpendicular to the imaging plane) resolution is on the order of several microns [51].

The effect of the relatively low axial resolution is that cellular signals are contaminated with signals from neighbouring objects. In the cortex, where cells are relatively sparsely distributed, the main confounding factor is neuropil contamination [46]. Neuropil signal, which corresponds to the average activity of densely packed cellular processes [46], is present in ‘background’ pixels that do not contain cell bodies. It has been found to correlate strongly with the electrocortigram [64] and be tuned to behavioural activity [100]. Pixels in cell somata are often contaminated with neuropil activity (Fig. 2.5). To decontaminate cellular signals, it is common to subtract the weighted local neuropil signal from the cellular signal. The value of the weight parameter chosen varies depending on the imaging set-up [100, 26, 63].

2.5 Outlook

Two-photon microscopy is a powerful tool for imaging with cellular resolution several hundred microns deep in living tissue. New generations of calcium indicators have

sufficient dynamic range to visualise single APs. Moreover, genetically encoded indicators can be expressed chronically, through non-invasive methods, in specific cell populations. Paired together, these technologies allow the same neuronal populations to be monitored at single-cell and single-spike resolution over periods of weeks and months [98]. They have been used to study the spatial organisation of receptive fields [86, 10] and to monitor changes in neural activity during learning [55, 100].

At present, cellular resolution imaging tends to occur over areas less than 1mm^2 on a single plane; even state of the art microscopes are limited to imaging thousands of cells simultaneously [124]. To allow more complex behaviours to be studied, the expansion of a microscope's field of view is the focus of constant development [59]. The trade-off between SNR and cell number is a limiting factor in this regard — a larger field of view necessitates a compromise on dwell-time and, therefore, noise level at each pixel. To further push the boundary of this technology, robust signal processing methods are required, which can maximise the information gleaned from noisy data and, in turn, relax the constraints on the imaging systems.

Chapter 3.

ABLE: an activity-based level set segmentation algorithm

A state of the art two-photon calcium imaging video may contain activity from thousands of neurons. Before neuronal activity can be analysed, the location of each cell within the video must be identified. This task is complicated by the heterogeneity in cellular properties, both spatial and temporal. For example, the calcium indicator used to generate a video affects a cell’s stereotypical temporal activity and its apparent shape — some genetically encoded indicators are excluded from the nucleus and therefore produce fluorescent ‘doughnuts’. Segmentation algorithms are further impeded by the frequency of cellular overlap, which occurs due to the relatively low axial resolution of the two-photon microscope. These challenges necessitate flexible segmentation algorithms with minimal assumptions on the properties of regions of interest (ROIs).

In this chapter, we present an algorithm for detecting the location of cells from two-photon calcium imaging data. In our framework, multiple coupled active contours evolve, guided by a model-based cost function, to identify cell boundaries. An active contour seeks to partition a local region into two subregions, a cell interior and exterior, in which all pixels have maximally ‘similar’ time courses. This simple, local model allows contours to be evolved predominantly independently. When contours are sufficiently close, their evolution is coupled, in a manner that permits overlap. We illustrate the ability of the proposed method to identify the boundaries of overlapping cells. We also demonstrate the versatility of our framework, which includes no priors on a cell’s morphology or stereotypical temporal activity, on mouse in vivo imaging

The work presented in this chapter led to the following publication [107].

data.

3.1 Introduction

A broad spectrum of methods have been proposed to overcome the challenges inherent in segmenting calcium imaging data. One confounding factor in the development of such algorithms is the heterogeneity in neuroscientists' selection criteria for ROIs. While some are only interested in somata [86], others also require the detection of dendrites [120]. While some require only active cells [78], others analyse the proportion of cells that are active [100]. A set of such selection criteria, which can be hard to define mathematically, can instead be implicitly incorporated through the use of supervised learning from manually labelled data [143, 4, 65]. Whereas an earlier approach used a non-binary decision tree on pixel features, such as mean, variance and covariance [143], later approaches have utilised convolutional neural networks (CNNs) [4, 65]. One CNN has been found to outperform most state of the art unsupervised methods on a benchmark dataset [65]. However, an obstacle for all supervised approaches is that the statistics of imaging videos vary significantly with factors such as the brain area, the microscope zoom and objective. In this case, the benchmark test dataset contained no 'unseen' data — all test samples had a corresponding training sample acquired under the same imaging conditions. It is, therefore, not yet clear how well supervised learning approaches generalise to new data from different imaging conditions.

A second subset of approaches bypass the (2+1)-D imaging video, utilising instead a 2-D summary image [122, 92, 60, 65]. This approach is practical; by significantly reducing the dimension of the input data, it allows videos to be processed that cannot be stored in the RAM of a typical laboratory workstation. Typically, one of two summary images is used: the mean image [92], which displays the mean of each pixel's activity, and the correlation image [122, 60, 65], which displays the average correlation between the activity of a pixel and that of its neighbours. While efficient, the usage of a 2D summary image limits the types of cells that can be detected. Only active cells are visible in the correlation image, whereas sparsely firing cells have low visibility in the mean image. Indeed, when a CNN was trained on manually labelled (2+1)-D imaging videos rather than the corresponding mean images, it was found to detect more transiently active neurons [4]. A further impediment to such approaches is the prevalence of cells in 2D summary images that appear to overlap. To identify the true

boundaries of overlapping cells it is necessary to consider the (2+1)-D imaging data.

The (2+1)-D imaging video, which consists of two spatial dimensions and one temporal dimension, is often prohibitively large to work on directly. One family of approaches therefore reshapes the (2+1)-D imaging video into a 2-D matrix. The resulting matrix admits a decomposition, derived from a generative model of the imaging video, into two matrices, each encoding spatial and temporal information. Early approaches applied independent component analysis to find the spatial components that have maximally statistically independent temporal activity [118, 82]. The constraint of temporal independence, however, causes algorithm performance to deteriorate in the case of overlapping cells [105, 154]. By instead including a constraint on non-negativity of the matrix entries, Maruyama et al. estimated the component matrices using non-negative matrix factorization (NMF) [76]. Pnevmatikakis et al. further developed this approach into a widely-used constrained NMF (CNMF) algorithm [105, 45]. The additional constraints on neuronal temporal activity, which is assumed to be an auto-regressive process derived from calcium dynamics, and sparsity constraints on the size of spatial components, have been shown to improve the performance of CNMF over plain NMF [105]. Related methods suggest performance gains can be achieved by including a more accurate model of the background activity [94, 56] or learning a dictionary of cellular features from the data [34]. By expressing the (2+1)-D imaging video as a 2-D matrix, this type of approach can achieve relatively high processing speeds. This does, however, come at the cost of discarded spatial information, which occurs when the two spatial dimensions are reduced to one. To mitigate this, either morphological filters [82, 94, 105, 45] or manual curation [118, 76] are required in post-processing steps.

In this chapter, we present a segmentation method in which cell boundaries are detected by multiple coupled active contours. To evolve an active contour we use the level set method, which is a popular tool in bioimaging due to its topological flexibility [28]. To each active contour, we associate a higher-dimensional function, referred to as a level set function, whose zero level set encodes the contour location. We implicitly evolve an active contour via the level set function, whose evolution is driven by a local model of the imaging data temporal activity. The model includes no assumptions on a cell's morphology or stereotypical temporal activity. We demonstrate that this versatility allows us to segment a variety of cell types and shapes and both active and inactive cells. For convenience, we refer to our method as ABLE (an Activity-Based LLevel set method). In Sections 3.2 and 3.3, we discuss relevant background on active contours and the level set method, respectively. Then, in Section 3.4, we present the proposed

segmentation algorithm. Finally, in Sections 3.5 and 3.6, we describe the data on which we evaluate ABLE and assess ABLE’s segmentation performance, respectively.

3.2 Active contours

The goal of image segmentation is to partition the domain of an image, Ω , into multiple subregions. When the boundary of a single object is sought, the domain is partitioned into two disjoint subregions: Ω^{in} , the interior, and Ω^{out} , the background. The object boundary, $\delta\Omega^{\text{in}}$, is then obtained from the closed curve delineating the border of Ω^{in} (Fig. 3.1A). Although in our application the video has three dimensions, we only wish to partition the spatial domain of the video, which is two-dimensional. This is due to the assumption that the objects are static during an imaging experiment. Otherwise, it would be necessary to partition the whole spatiotemporal domain [79, 110]. The image is denoted with a function, $\mathbf{I} : \Omega \rightarrow \mathbb{R}^m$, that associates a vector of length m with every pixel, \mathbf{x} , in the image domain, $\Omega \subset \mathbb{R}^2$. When $m = 1$, we have a single-frame image, when $m > 1$, we have a video consisting of m frames.

An active contour is a commonly-used tool for identifying an object’s boundary. Starting from a seed location, the contour evolves towards the object boundary, see Fig. 3.1. Its evolution is guided by a cost function that is designed to be minimised when the contour is located at the true boundary. We denote the active contour, which evolves in artificial time, $\tau \in \mathbb{R}$, with $\Gamma(\tau)$. At each timestep, $\Gamma(\tau)$ represents the most recent estimate of $\delta\Omega^{\text{in}}$. Active contours are versatile; they can accommodate a broad range of cost functions and evolution frameworks (for a review of the bioimaging literature see [28]). Consequently, since their conception in 1988 [61], they have found widespread usage in image processing, both in segmentation and, for example, in motion detection [58] and multi-view scene modelling [8].

There are two main approaches to modelling an active contour: parametric and implicit. In the former approach, a curve is represented by a continuous parameter, a discrete set of coefficients and a method of interpolation from the discrete coefficients to the continuous curve [61, 57]. For example, curves can efficiently be constructed as linear combinations of B-spline basis functions¹ [42, 15]. While effective for objects that have stereotypical shapes, this approach is not ideal for irregularly-shaped objects

¹A B-spline is a type of piecewise polynomial function (a ‘spline’) with minimal support for a given degree of smoothness.

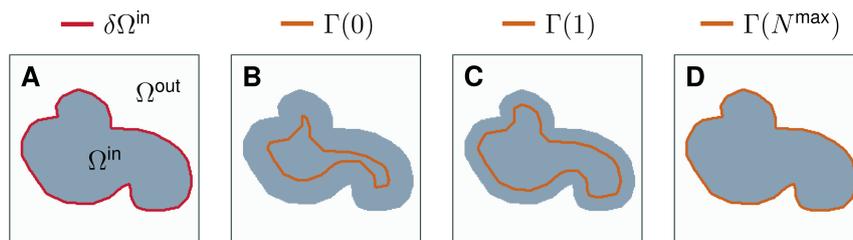


Figure 3.1: Example of an active contour segmenting a piecewise smooth image. The image consists of an object, Ω^{in} , and a background, Ω^{out} (A). An active contour, Γ , evolves guided by a cost function that is minimised when it is located at the true object boundary, $\delta\Omega^{\text{in}}$ (B - D). After N^{max} iterations, the contour converges at the true object boundary.

whose shape cannot easily be represented by a combination of basis functions.

The second class of approaches evolves the contour implicitly via a higher-dimensional function [21, 90]. In this framework, the active contour is defined as the zero level set of a manifold, $\phi : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$, that also evolves in time, such that

$$\Gamma(\tau) = \{\mathbf{x} \in \Omega : \phi(\mathbf{x}, \tau) = 0\}. \quad (3.1)$$

Unless shape constraints are explicitly included in the cost function, this class of approaches allows a greater degree of topological flexibility than parametric active contours. A further advantage of this framework is that it can easily be extended to images with more than two spatial dimensions. This is fitting for calcium imaging data, which is starting to be collected in three spatial dimensions [1]. This is the framework we adopt for our segmentation algorithm, which we introduce in more detail in Section 3.3.

The evolution of an active contour is guided by a cost function, which is designed to be minimised when the contour is located at the true object boundary. The cost function is typically formed of two components: one relating to the image, which is referred to as the external energy, and another relating to the properties of the contour itself, which we refer to as the regulariser. Both components are weighted by scalar parameters, $\mu, \lambda \in \mathbb{R}$, to form the cost function:

$$\mathcal{E}(\Gamma) = \lambda \mathcal{E}_{\text{ext}}(\Gamma) + \mu \mathcal{E}_{\text{reg}}(\Gamma), \quad (3.2)$$

where \mathcal{E}_{ext} is the external energy and \mathcal{E}_{reg} is the regulariser. The external energy, which we discuss in the following section, incorporates assumptions about the properties of the object to be detected. The regulariser, which we discuss in Section 3.3.2, encourages the object boundary to conform to a set of pre-defined conditions. For example, jagged edges or excessively long boundaries may be penalised.

3.2.1 External energy

To construct the external energy, quantitative image features are computed. The features are typically one of two types: boundary-based and region-based features. As objects in calcium imaging videos tend not to possess sharp edges (due to the blurring induced by the point spread function), we focus on the latter approach. Two features $\mathbf{f}^{\text{in}} \in \mathbb{R}^m$ and $\mathbf{f}^{\text{out}} \in \mathbb{R}^m$ are derived from values of the image in the interior and exterior of Γ , respectively. These features, which are re-computed each time the contour is updated, are then used to classify pixels into the interior or exterior regions. Using a function, $D : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$, that computes the discrepancy between the image at pixel \mathbf{x} and a feature, we have

$$\mathcal{E}_{\text{ext}}(\Gamma) = \lambda^{\text{in}} \int_{\text{inside } \Gamma} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{in}}) d\mathbf{x} + \lambda^{\text{out}} \int_{\text{outside } \Gamma} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{out}}) d\mathbf{x}, \quad (3.3)$$

where λ^{in} and λ^{out} are real-valued weight parameters. The external energy thus penalises discrepancies between the value of an image at pixel \mathbf{x} and the feature of the subregion to which it belongs. The relative impact of discrepancies in the interior and exterior is controlled by the weights, λ^{in} and λ^{out} .

Chan and Vese proposed a region-based active contours approach for piecewise smooth, single-frame images containing a single object and a background [22, 24]. The external energy, which is inspired by the Mumford-Shah functional for the segmentation of piecewise smooth images [83], is equal to

$$\mathcal{E}_{\text{ext}}(\Gamma) = \lambda^{\text{in}} \int_{\text{inside } \Gamma} \|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{in}}\|^2 d\mathbf{x} + \lambda^{\text{out}} \int_{\text{outside } \Gamma} \|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{out}}\|^2 d\mathbf{x}, \quad (3.4)$$

where \mathbf{f}^{in} and \mathbf{f}^{out} are the average intensity of pixels in the interior and exterior, respectively. This energy penalises discrepancies between a pixel's intensity and the average intensity of the subregion to which it belongs. The initial formulation was designed for single-frame images, where $\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{in}}, \mathbf{f}^{\text{out}} \in \mathbb{R}$, and later extended to vector-valued images [23] and images of more than two pieces [25].

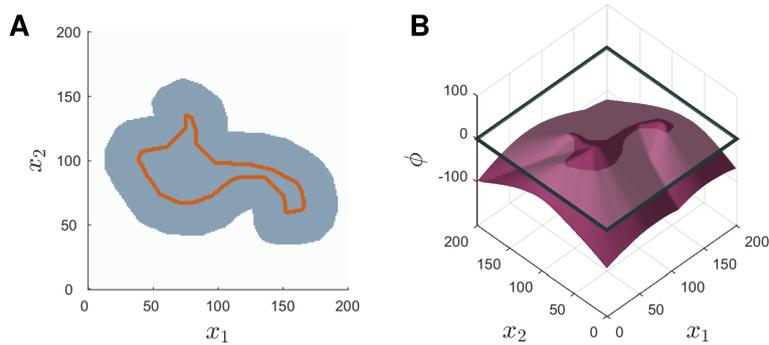


Figure 3.2: The level set function (LSF). The active contour (**A**) is encoded as the zero level set of the LSF (**B**). The LSF is positive inside the contour, negative outside and zero on the contour.

3.3 Level set method

To update the active contour we use the level set method of Osher and Sethian [91]. This method was first introduced to image processing by [20] and [74]; it has since found widespread use in the field. We represent the active contour by a function ϕ , which is positive for all \mathbf{x} inside the contour, negative for \mathbf{x} outside the contour and equal to zero for \mathbf{x} on the boundary (see Fig. 3.2). We refer to ϕ as the level set function (LSF), as its zero level set encodes the contour of interest. We note that we have reversed the convention of negative ϕ inside the contour and positive ϕ outside in order to improve the clarity of figures representing the LSF.

We aim to find the active contour that minimises the energy in Eq. (3.2). We thus seek the LSF that minimises the analogous cost function:

$$\mathcal{E}(\phi) = \lambda \mathcal{E}_{\text{ext}}(\phi) + \mu \mathcal{E}_{\text{reg}}(\phi), \quad (3.5)$$

where we have replaced Γ with ϕ . As it is in general not possible to minimise this cost function directly, a standard way to obtain the LSF that minimises Eq. (3.5) is to find the steady-state solution to the gradient flow equation [5]:

$$\frac{\partial \phi}{\partial \tau} = -\frac{\partial \mathcal{E}}{\partial \phi}. \quad (3.6)$$

As artificial time, denoted by τ , advances, the LSF is thus evolved in the image plane in a manner that yields the sharpest decrease in the cost function. From Eq. (3.5),

this partial differential equation (PDE) becomes

$$\frac{\partial \phi}{\partial \tau} = - \left(\lambda \frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi} + \mu \frac{\partial \mathcal{E}_{\text{reg}}}{\partial \phi} \right). \quad (3.7)$$

The evolution of the LSF is thus driven by two components: the former provides the impetus from the image data, and is referred to as the external velocity, and the latter the impetus from the regulariser. In Sections 3.3.1 and 3.3.2, we discuss the computation of each component.

To solve the PDE numerically, it is necessary to discretise Eq. (3.7). We approximate ϕ by storing and updating its value on the uniformly-spaced image grid. The evolution parameter τ is also discretised, such that

$$\phi(\tau + \Delta\tau) = \phi(\tau) - \Delta\tau \left(\lambda \frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi} + \mu \frac{\partial \mathcal{E}_{\text{reg}}}{\partial \phi} \right). \quad (3.8)$$

At every timestep τ , the LSF is updated until a maximum number of iterations, N^{max} , has been reached or the active contour has converged. Using one or both of these conditions is common in the active contour literature, see, for example, [29, 69]. We retain $\mu\Delta\tau < 0.25$ in order to satisfy the Courant-Friedrichs-Lewy condition [69] — a necessary condition for the convergence of a numerically-solved PDE. This condition requires that the numerical waves propagate at least as fast as the physical waves [89].

3.3.1 External velocity

To compute the external velocity, we rewrite the external energy from Eq. (3.3) in terms of the Heaviside function, such that

$$\mathcal{E}_{\text{ext}}(\phi) = \int_{\Omega} A^{\text{in}}(\mathbf{x})H(\phi(\mathbf{x})) + A^{\text{out}}(\mathbf{x})(1 - H(\phi(\mathbf{x})))d\mathbf{x}, \quad (3.9)$$

where $A^{\text{in}}(\mathbf{x}) = \lambda^{\text{in}}D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{in}})$ and $A^{\text{out}}(\mathbf{x}) = \lambda^{\text{out}}D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{out}})$. Here, we are implicitly using $H(\phi)$ and $(1 - H(\phi))$ as indicator functions of the subregions inside and outside the active contour, respectively. To calculate the derivative with respect to ϕ , we use the standard variational formulation [23, 90, 5]. The Heaviside function is

approximated by a smooth function, $H_\epsilon(x)$, where

$$H_\epsilon(x) = \begin{cases} 1 & x > \epsilon \\ \frac{1}{2} \left(1 + \frac{x}{\epsilon} + \frac{1}{\pi} \sin\left(\frac{\pi x}{\epsilon}\right) \right) & |x| \leq \epsilon \\ 0 & x < -\epsilon. \end{cases} \quad (3.10)$$

The derivative of the Heaviside function, the delta function, is approximated by $\delta_\epsilon(x)$, where

$$\delta_\epsilon(x) = \begin{cases} \frac{1}{2\epsilon} \left(1 + \cos\left(\frac{\pi x}{\epsilon}\right) \right) & |x| \leq \epsilon \\ 0 & |x| > \epsilon. \end{cases} \quad (3.11)$$

We note that $\delta_\epsilon(x) = \frac{d}{dx}H_\epsilon(x)$. Replacing the Heaviside function in Eq. (3.9) with its approximation and computing the derivative, the external velocity at $\mathbf{x} \in \Omega$ is

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi}(\mathbf{x}) = \delta_\epsilon(\phi(\mathbf{x})) (A^{\text{in}}(\mathbf{x}) - A^{\text{out}}(\mathbf{x})) \quad (3.12)$$

$$= \delta_\epsilon(\phi(\mathbf{x})) (\lambda^{\text{in}} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{in}}) - \lambda^{\text{out}} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{out}})). \quad (3.13)$$

The term $\delta_\epsilon(\phi(\mathbf{x}))$, which is only non-zero at pixels on or near the active contour, acts as a localization operator, ensuring that the velocity only impacts the LSF in the vicinity of the active contour. The parameter ϵ determines the radius of the non-zero band.

An example of the external velocity for the single-image Chan-Vese active contour model is shown in Fig. 3.3. From Eq. (3.4), setting $\lambda^{\text{in}} = \lambda^{\text{out}} = 1$, we have

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi}(\mathbf{x}) = \delta_\epsilon(\phi(\mathbf{x})) (\|I(\mathbf{x}) - \mathbf{f}^{\text{in}}\|^2 - \|I(\mathbf{x}) - \mathbf{f}^{\text{out}}\|^2), \quad (3.14)$$

where, in this case, \mathbf{f}^{in} and \mathbf{f}^{out} are the average intensity values within each subregion. The contour is thus moved towards the pixels whose values are more similar to \mathbf{f}^{in} than \mathbf{f}^{out} , and repelled by other pixels.

3.3.2 Regularisation

The external velocity updates the LSF only in the vicinity of the zero level set. As the LSF is updated, the position of the zero level set is altered and the LSF is subsequently updated at new locations. Consequently, it is beneficial to maintain an LSF that is

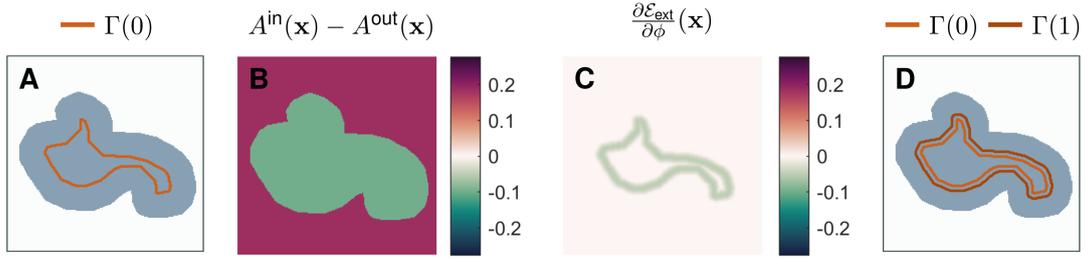


Figure 3.3: The Chan-Vese active contours model. In **A**, we display the initial estimate of the object boundary, $\Gamma(0)$. The external velocity at each pixel is equal to a data-based term, $A^{\text{in}}(\mathbf{x}) - A^{\text{out}}(\mathbf{x})$, multiplied by a localisation operator. In **B**, we display the former, which is negative at pixels that are more similar to the interior than the exterior of the active contour, and positive otherwise. The localisation operator suppresses the contribution of the data-based term at pixels distant from the active contour (**C**). In **D**, we plot the updated active contour, whose evolution is driven by the external velocity in **C**.

smoothly varying in that region. We use the regulariser of Li et al. [69], which has been implemented in numerous medical imaging applications [68, 72, 146]. The goal of the regulariser is to ensure that ϕ approximates a signed distance function in the vicinity of the contour and is flat elsewhere. In this context, a distance function satisfies:

$$d(\mathbf{x}) = \min_{\mathbf{x}_c \in \Gamma} \|\mathbf{x} - \mathbf{x}_c\|, \quad (3.15)$$

i.e. the value of $d(\mathbf{x})$ is the distance from \mathbf{x} to the closest point on the contour. It follows that, for all \mathbf{x} on the contour, the distance is equal to 0. A signed distance function encodes both distance to the contour and the subregion that a point belongs to; it is equal to $d(\mathbf{x})$ for pixels in the curve interior and $-d(\mathbf{x})$ for those in the exterior, see Fig. 3.2B. A signed distance function, therefore, satisfies the properties of an LSF and is sufficiently smoothly varying. The regularisation energy is

$$\mathcal{E}_{\text{reg}}(\phi) = \int_{\Omega} p(\|\nabla\phi\|) \, d\mathbf{x}, \quad (3.16)$$

where

$$p(s) = \begin{cases} \frac{1}{(2\pi)^2} (1 - \cos(2\pi s)) & \text{if } s \leq 1, \\ \frac{1}{2} (s - 1)^2 & \text{otherwise.} \end{cases} \quad (3.17)$$

The partial derivative is then [69]

$$\frac{\partial \mathcal{E}_{\text{reg}}}{\partial \phi} = \text{div} \left(\nabla \phi \frac{p'(\|\nabla \phi\|)}{\|\nabla \phi\|} \right), \quad (3.18)$$

where $\text{div}(\cdot)$ is the divergence operator.

3.4 Proposed method

In the approaches presented thus far, an active contour partitions the whole image domain into two distinct subregions: the object and the background. In the case of calcium imaging data, the objects to be detected (cells or cellular components) are multiple orders of magnitude smaller than the spatial domain. We therefore adopt a more local approach to the segmentation problem. In the following, we start by considering the problem of detecting the boundary of a single, isolated cell.

In our framework, an active contour seeks to partition a local region of the video into a cell interior and exterior. At each iteration, the cell exterior is defined as the set of pixels exterior to the contour within a fixed distance, see Fig. 3.4B. The default distance is taken to be two times the expected radius of a cell. We refer to this exterior region as the narrowband to emphasise its proximity to the contour of interest. As an active contour is updated, so is the corresponding narrowband (Fig. 3.4F). The region of the video for which the optimal partition is sought is therefore not static; rather, it evolves as an active contour evolves.

Given an instance of the active contour, $\Gamma(\tau)$, the corresponding LSF is defined on the contour interior and narrowband (see Fig. 3.4C). We compute $\mathbf{f}^{\text{in}} \in \mathbb{R}^m$ and $\mathbf{f}^{\text{out}} \in \mathbb{R}^m$ as the average cell interior and narrowband time courses, respectively, where m is the number of frames in the video. As before, the external energy is

$$\mathcal{E}_{\text{ext}}(\phi) = \int_{\phi > 0} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{in}}) \text{d}\mathbf{x} + \int_{\phi \leq 0} D(\mathbf{I}(\mathbf{x}), \mathbf{f}^{\text{out}}) \text{d}\mathbf{x}. \quad (3.19)$$

where we take $\lambda^{\text{out}} = \lambda^{\text{in}} = 1$.

In the following section, we introduce the dissimilarity metrics used. Then, in Section 3.4.2, we extend the framework to multiple active contours.

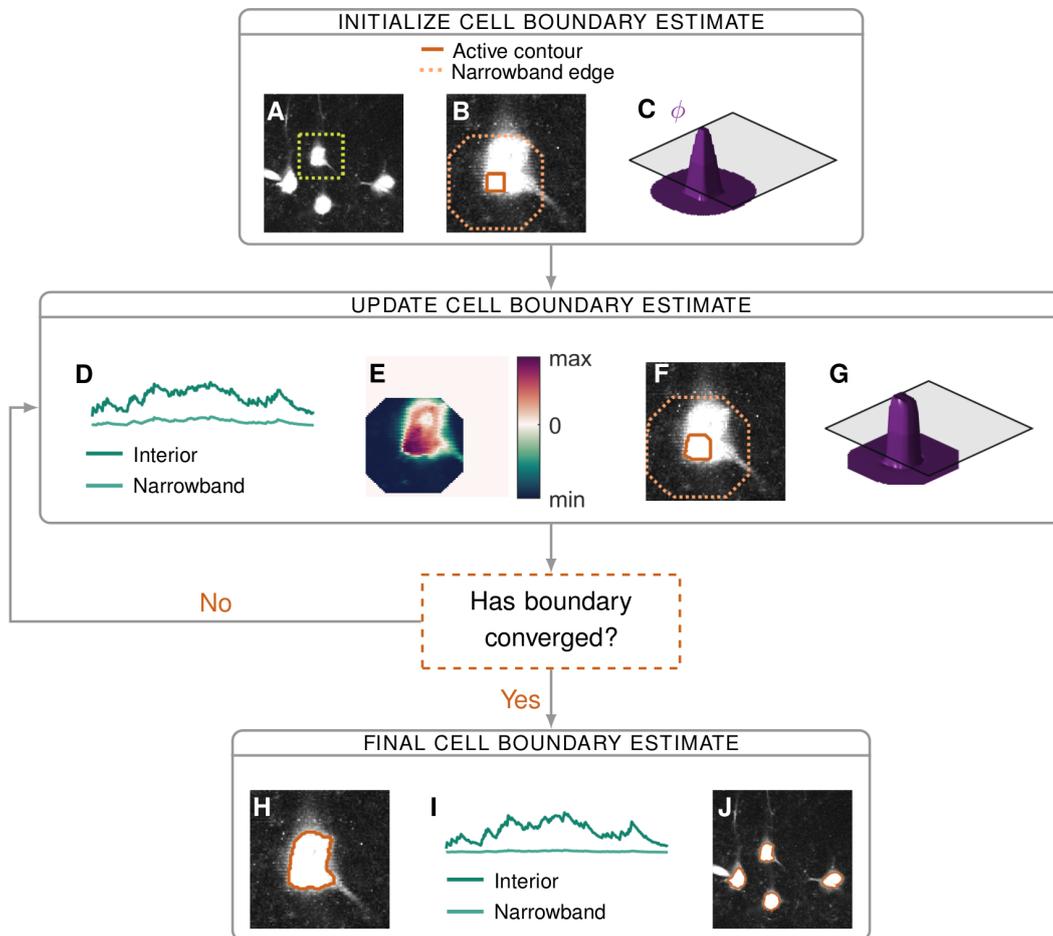


Figure 3.4: Flow diagram of the proposed segmentation algorithm. The three stages are: initialization (**A-C**), iterative updates of the estimate (**D-G**) and convergence (**H-J**). When cells are sufficiently far apart we can segment them independently — in this example we focus on the isolated cell in the dashed box on the maximum intensity image in **A**. We make an initial estimate of the cell interior, from which we form the corresponding narrowband (**B**) and level set function ϕ (**C**). Based on the discrepancy between a pixel’s time course and the time courses of the interior and narrowband regions (**D**), we calculate the velocity of ϕ at each pixel (**E**). ϕ evolves according to this velocity (**G**), which updates the location of the interior and narrowband (**F**). Final results for: one cell (**H**), the average signals from the corresponding interior and narrowband (**I**) and segmentation of all four cells (**J**).

3.4.1 Dissimilarity metrics

Due to the heterogeneity of calcium imaging data, we do not use a universal dissimilarity metric. When both the pattern and magnitude of a pixel’s temporal activity are informative, as is typically the case for synthetic dyes, we use

$$D^E(I(\mathbf{x}), \mathbf{f}) = \frac{1}{m} \|I(\mathbf{x}) - \mathbf{f}\|^2, \quad (3.20)$$

for $\mathbf{f} \in \mathbb{R}^m$. When we have an image not a video (i.e. $\mathbf{I}(\mathbf{x})$ and \mathbf{f} are one-dimensional) this dissimilarity metric reduces to the fitting term introduced by Chan and Vese [24].

For datasets in which the fluorescence expression level varies throughout cells and, as a consequence, pixels in the same cell exhibit the same pattern of activity at different magnitudes, we use a metric based on the correlation, such that

$$D^C(I(\mathbf{x}), \mathbf{f}) = 1 - \text{corr}(I(\mathbf{x}), \mathbf{f}), \quad (3.21)$$

where corr represents the Pearson correlation coefficient.

3.4.2 Neighbouring cells

We now extend the cost function presented in Eq. (3.19) to one suitable for partitioning a region into multiple cell interiors and a global exterior, which encompasses the narrowbands of all the cells. We denote with ϕ_m the LSF that corresponds to the m^{th} cell and $\mathbf{f}^{\text{in},m}$ the corresponding interior time course. Due to the relatively low axial resolution of a two-photon microscope, fluorescence intensity at one pixel can originate from multiple cells in neighbouring z-planes. Therefore, in contrast to previous approaches to coupling active contours [156, 40], we do not penalise overlap of contour interiors. Instead, we allow them to overlap when this best fits the data. In particular, we assume that a pixel in multiple cells would have a time course well fit by the sum of the interior time courses for each cell. The external energy in the case of $M > 0$ cells is thus

$$\mathcal{E}_{\text{ext}}(\phi_1, \phi_2, \dots, \phi_M) = \int_{\text{outside}} D(I(\mathbf{x}), \mathbf{f}^{\text{out}}) d\mathbf{x} + \int_{\text{inside}} D(I(\mathbf{x}), \sum_{m \in \mathcal{C}(\mathbf{x})} \mathbf{f}^{\text{in},m}) d\mathbf{x} \quad (3.22)$$

where ‘outside’ denotes the global exterior, the union of the narrowbands excluding any cell interior, and ‘inside’ denotes the union of the contour interiors. The function

$C(\mathbf{x})$ identifies all cells whose interior contains $\mathbf{x} \in \Omega$. We denote the local spatial domain, which incorporates the global exterior and the union of contour interiors, with $\bar{\Omega}$. When the region to be partitioned contains only one cell, the external energy in Eq. (3.22) reduces to that in Eq. (3.19).

We can rewrite the RHS of Eq. (3.22) in terms of the LSFs using the Heaviside function implicitly as an indicator function, such that

$$\mathcal{E}_{\text{ext}}(\phi_1, \phi_2, \dots, \phi_M) = \int_{\bar{\Omega}} A(\mathbf{x}) \prod_{k \in C(\mathbf{x})} H(\phi_k(\mathbf{x})) \prod_{n \notin C(\mathbf{x})} (1 - H(\phi_n(\mathbf{x}))) \, d\mathbf{x}, \quad (3.23)$$

where we simplify the integrand by writing

$$A(\mathbf{x}) = \begin{cases} D(I(\mathbf{x}), \sum_{m \in C(\mathbf{x})} \mathbf{f}^{\text{in},m}) & \text{if } C(\mathbf{x}) \neq \emptyset, \\ D(I(\mathbf{x}), \mathbf{f}^{\text{out}}) & \text{otherwise.} \end{cases} \quad (3.24)$$

To obtain the external velocity of the m^{th} active contour we compute the derivative of Eq. (3.23) with respect to ϕ_m . We state the result here and provide the derivation in Appendix A.1. As in Eq. (3.12), the velocity is the product of a data-based term and a localisation operator, such that

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m}(\mathbf{x}) = \delta_\epsilon(\phi_m(\mathbf{x})) (D(I(\mathbf{x}), \mathbf{f}^{\text{in},m}) - D(I(\mathbf{x}), \mathbf{f}^{\text{out}})), \quad (3.25)$$

if \mathbf{x} is not in a neighbouring cell, i.e. $H(\phi_k(\mathbf{x})) = 0$ for all $k \neq m$. Otherwise

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m}(\mathbf{x}) = \delta_\epsilon(\phi_m(\mathbf{x})) \left(D\left(I(\mathbf{x}), \mathbf{f}^{\text{in},m} + \sum_{\substack{j \in C(\mathbf{x}) \\ j \neq m}} \mathbf{f}^{\text{in},j}\right) - D\left(I(\mathbf{x}), \sum_{\substack{j \in C(\mathbf{x}) \\ j \neq m}} \mathbf{f}^{\text{in},j}\right) \right). \quad (3.26)$$

The external velocity can be interpreted as follows: if a pixel, not in another cell, has time course more similar to that of the contour interior than the narrowband, then the contour moves to incorporate that pixel. If a pixel in another cell has time course better matched by the sum of the interior time courses of cells containing that pixel plus the interior time course of the evolving active contour, then the contour moves to incorporate it. Otherwise, the contour is repelled from that pixel.

Although the external velocity contains contributions from all cells in the video, the localisation operator, δ_ϵ , ensures that the problem remains local. We only evaluate the image-based terms in Eq. (3.25) and (3.26) at pixels where the localisation operator is

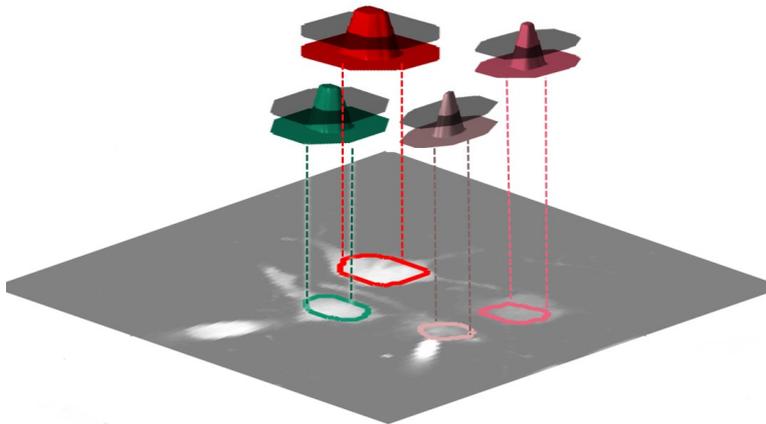


Figure 3.5: Evolution of multiple active contours. Each active contour seeks to partition its local region into a cell interior and exterior. When contours become sufficiently close, their evolution is coupled.

non-zero. As noted in Section 3.3.1, this is a band of pixels around the active contour, with width on the order of a single pixel. Consequently, many active contours can be evolved independently. Mathematically, \mathbf{f}^{out} is the average time course corresponding to the global exterior, which is the union of all contours' narrowbands. In practice, when we compute the velocity of each active contour we utilise a local \mathbf{f}^{out} , which is computed from pixels in that contour's narrowband.

The active contours are evolved in parallel, so that, at each time step, τ , we compute

$$\phi_m(\tau + \Delta\tau) = \phi_m(\tau) - \Delta\tau \left(\lambda \frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m} + \mu \frac{\partial \mathcal{E}_{\text{reg}}}{\partial \phi_m} \right). \quad (3.27)$$

for all $m \in \{1, \dots, M\}$, where the regulariser is the same as in Section 3.3.2. Then, the features, $\mathbf{f}^{\text{in},m}$ and \mathbf{f}^{out} , are updated and each contour is evolved again. This process is repeated until a maximum number of iterations has been reached or the contours have converged. In the next section, we discuss details of the implementation, such as the conditions for convergence and the initialisation procedure.

3.4.3 Implementation

Initialisation

We devised an automatic initialization algorithm which selects connected areas of peak local correlation and/or peak mean intensity as initial ROI estimates. The correlation

		Num. cells		
		25	125	225
Num. frames (m)	100	1.1	6.5	11.2
	1000	1.3	6.5	12.7

Table 3.1: Runtime of the segmentation algorithm (minutes) on synthetic data. On data with dimensions $512 \times 512 \times m$, the runtime of ABLE (minutes) increases linearly with the number of cells and is not significantly affected by increasing number of frames, m . Runtime was measured on a PC with 3.4GHz Intel Core i7 CPU.

image displays, at each pixel, the average correlation between that pixel’s time course and those of neighbouring pixels. Local peaks in this image and the mean intensity image are identified (by a built-in MATLAB function, ‘`imextendedmax`’) as candidate ROIs. The selectivity of the initialization is set by a tuning parameter, which defines the relative height with respect to neighbouring pixels (in units of standard deviation of the input image) of the peaks that are suppressed. The higher the value of this parameter, the more conservative the initialization. Empirically, we have found values in the range 0.2 - 0.8 to be suitable. This typically overestimates the number of ROIs; redundant estimates are automatically pruned during the update phase of the algorithm.

Convergence

We stop updating a contour estimate when a maximum number of iterations, N_{\max} , has been reached or the active contour has converged — using one or both of these conditions is common in the active contour literature, see, for example, [29, 69]. A contour is deemed to have converged if, in N_{con} consecutive iterations, the number of pixels that are added to or removed from the interior is less than ρ . As default, we take $N_{\max} = 100$, $N_{\text{con}} = 40$ and $\rho = 2$.

The complexity of the level set method is intrinsically related to the dimensionality of the active contour; the number of frames of the video is only relevant to the evaluation of the external velocity, which accounts for a small fraction of the computational cost. In Table 3.1, we demonstrate that increasing video length by a factor of 10 has only a minor impact on processing time. Updating an active contour is generally a local problem — consequently, we observe that algorithm runtime increases linearly with the total number of cells, see Table 3.1. Due to the independence of spatially separate

ROIs in our framework, further performance speed-ups are achievable by parallelising the computation.

Merging and pruning regions

ABLE automatically merges two cells if they are sufficiently close and their interiors sufficiently correlated — a strategy previously employed in the CNMF algorithm of Pnevmatikakis et al. [105]. When two contours are merged, their respective level set functions are replaced with a single level set function, initialized as a signed distance function, with a zero level set that represents the union of the contour interiors. The required proximity for two cells to be merged is one cell radius (the expected cell radius is one of two required user input parameters) and the default correlation threshold is 0.8. A contour is automatically removed (‘pruned’) at the end of the update phase if its area is smaller or greater than adjustable minimum or maximum size thresholds, which, as default, are set at 3 and $3\pi r^2$ pixels, respectively, where r is the expected radius of a cell.

3.5 Experimental methods

In Section 3.6, ABLE’s performance is assessed on a range of real and simulated data. In the following, we describe the experimental methods used to obtain this data and the performance metrics used.

3.5.1 Simulations

Cellular spike trains are generated from mutually independent Poisson processes. A cell’s temporal activity is the sum of a stationary baseline component, the value of which is selected from a uniform distribution, and a spike train convolved with a stereotypical calcium transient pulse shape. Cells are ‘doughnut’ (annulus) shaped to mimic videos generated by genetically encoded calcium indicators, which are excluded from the nucleus. To achieve this, the temporal activity of a pixel in a cell is generated by multiplying the cellular temporal activity vector by a factor in $[0, 1]$ that decreases as pixels are further from the cell boundary. When two cells overlap in one pixel,

we sum the contributions of both cells at that pixel. Spatially and temporally varying background activity, generated independently from the cellular spiking activity, is present in pixels that do not belong to a cell.

3.5.2 Two-photon calcium imaging of hippocampal slices

This dataset was collected by Dr Renaud Schuck, our collaborator in Imperial College London's Department of Bioengineering. Here, we provide an overview of the procedures used. For further details, we refer the reader to [117]. All procedures were performed in accordance with national and institutional guidelines and were approved by the UK Home Office under Project License 70/7355 to Simon R Schultz. Juvenile wild-type mice of either sex (C57Bl6, P13-P21) were anaesthetised using isoflurane prior to decapitation procedure. Brain slices (400 μm thick) were horizontally cut in 1-4°C ventilated (95% O₂, 5% CO₂) slicing Artificial Cerebro-Spinal Fluid (sACSF: 0.5 mM CaCl₂, 3.0 mM KCl, 26 mM NaHCO₃, 1 mM NaH₂PO₄, 3.5 mM MgSO₄, 123 mM Sucrose, 10 mM D-Glucose). Hippocampal slices containing Dentate Gyrus, CA3 and CA1 were taken and resting in ventilated recovery ACSF (rACSF: 2 mM CaCl₂, 123 mM NaCl, 3.0 mM KCl, 26 mM NaHCO₃, 1mM NaH₂PO₄, 2mM MgSO₄, 10mM D-Glucose) for 30min at 37°C. After this the slices were placed in an incubation chamber containing 2.5 mL of ventilated rACSF and 'painted' with 10 μL of the following solution: 50 μg of Cal-520 AM (AAT Bioquest), 2 μL of Pluronic-F127 20% in DMSO (Life Technologies) and 48 μL of DMSO (Sigma Aldrich) where they were left for 30 min at 37°C in the dark. Slices were then washed in rACSF at room temperature for 30 min before imaging. Dentate Gyrus granular cells were identified using oblique illumination prior to being imaged using a standard commercial galvanometric scanner based two-photon microscope (Scientifica Ltd) coupled to a mode-locked Mai Tai HP Ti Sapphire (Spectra-Physics) laser system operating at 810 nm. Functional calcium images of granular cells were acquired with a 40X objective (Olympus) by raster scanning a $180 \times 180 \mu\text{m}^2$ square Field of View at 10 Hz. Electrical stimulation was accomplished with a tungsten bipolar concentric microelectrode (WPI) where the tip of the electrode was placed into the molecular layer of the Dentate Gyrus (20 pulses with a pulse-width of 400 μs and a 60 μA amplitude were delivered into the tissue with a pulse repetition rate of 10 Hz, repeated every 40 sec).

3.5.3 In vivo imaging datasets

We demonstrate algorithm performance on two imaging videos made freely available in online datasets [99, 85]. In the former dataset, the vibrissal cortex was imaged at various depths, from layer 1 to deep layer 3, whilst the mouse performed a pole localization task [100, 49]. Imaging was performed using the genetically-encoded indicator GCaMP6s. Fluorescence was collected through raster scans of a $600 \times 600 \mu\text{m}^2$ region at 7Hz. The latter dataset was recorded from the vibrissal primary somatosensory cortex of an awake head-fixed mouse in a virtual reality setup [123]. Images were obtained using indicator GCaMP6s. Fluorescence was collected through raster scans of a $600 \times 600 \mu\text{m}^2$ region at 8Hz.

3.5.4 Performance metrics

To quantify segmentation accuracy we use two sets of metrics. On simulated data, we compute the success rate of algorithm output compared to the ground truth cell locations. For each cell in the ground truth dataset, we identify the corresponding cell in the algorithm output. We then compute the precision, which is the percentage of pixels in an estimated cell that are also in the corresponding ground truth cell. Similarly, we compute the recall, which is the percentage of pixels in the ground truth cell that are in the estimated cell. The success rate (%) is the harmonic mean of the precision and recall, such that

$$\text{success rate} = 2 \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (3.28)$$

For each simulated imaging video, the output of the metric is a single score, which is the average of the success rates over all cells in the video.

On real data with ground truth stemming from manually labelled cells, we use a different approach. In this case, the ground truth cell locations are not likely to be accurate on the order of single pixels. Therefore, we identify an estimated cell as matching a ground truth cell if their centres are within 5 pixels of one another. We compute the recall as the proportion of ground truth cells that are matched by a detected cell. We also compute the fall-out-rate, which is the complement of the precision. This is the proportion of detected cells that do not match a ground truth cell.

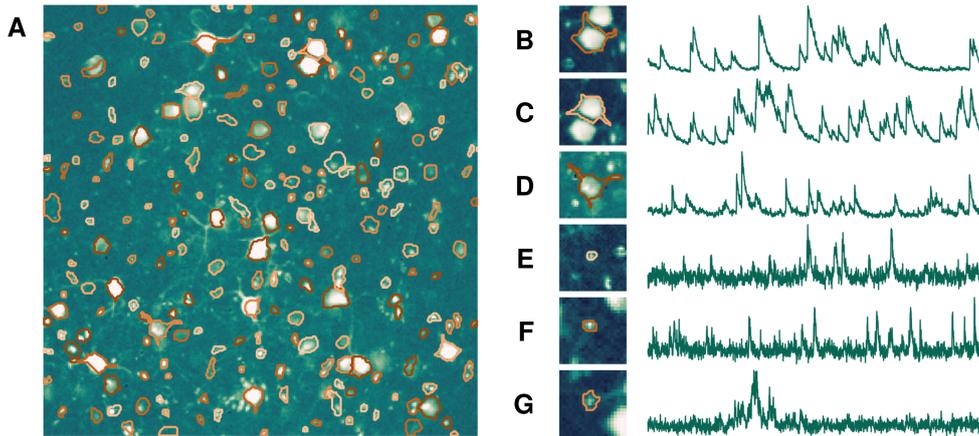


Figure 3.6: Segmentation results on in vivo imaging data. ABLE detects cells with varying size and shape from mouse in vivo imaging data. The detected cell boundaries are superimposed on the correlation image of the imaging video (**A**). Extracted neuropil-corrected time series and corresponding cell boundaries are displayed for a subset of the detected regions: both cell bodies (**B**, **C** and **D**) and cross-sections of dendrites (**E**, **F** and **G**) are detected .

3.6 Results

3.6.1 Robustness to heterogeneity in cell properties

ABLE detected region of interest (ROIs) with diverse properties from the publicly available mouse in vivo imaging dataset of [99], see Fig. 3.6. A video demonstration of segmentation on this data along with the code to generate the results is available at github.com/StephanieRey.

To maintain a versatile framework we included no priors on cellular morphology in the cost function that drives the evolution of an active contour. This allowed ABLE to detect irregularly-shaped ROIs, such as cell bodies with dendrites attached (Fig. 3.6B, C and D). No prior information on stereotypical neuronal temporal activity is included in our framework. As such, ABLE detected cross-sections of dendrites, which do not display stereotypical calcium transient activity (Fig. 3.6E, F and G). Activity from dendritic cross-sections is an essential indicator of neuronal activity in lower cortical layers, at depths where imaging is not possible [98]. We note that, due to ABLE’s treatment of overlapping cells, it was possible to identify the overlapping dendrites belonging to the cells in Fig. 3.6F and G.

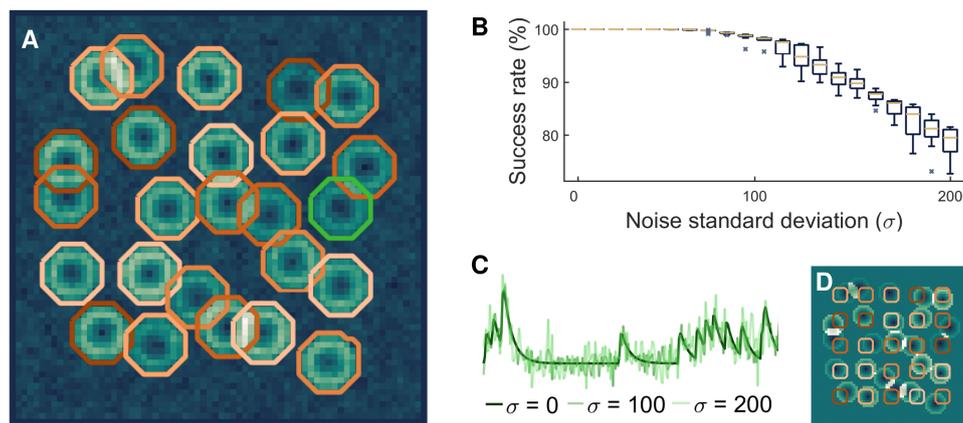


Figure 3.7: Segmentation results on simulated data containing overlapping cells. ABLE detected the true boundaries of overlapping cells from noisy simulated data. The detected contours for one realization of noise with standard deviation (σ) equal to 60 are plotted on the correlation image in **A**. Given an initialisation on a fixed grid, the true cell boundaries are detected with success rate of at least 99% for $\sigma < 90$ (**B**). The central marker and box edges in **B** indicate the median and the 25th and 75th percentiles, respectively. For noise level reference, we plot the average time course from inside the green contour in **A** at various levels (**C**). In **D**, we plot the initialisation, which was on a fixed grid, on the mean image of the video.

3.6.2 Demixing overlapping cells

When imaging through scattering tissue, a two-photon microscope can have relatively low axial resolution (on the order of ten microns) in comparison to its excellent lateral resolution. As a consequence, the photons collected at one pixel can in some cases originate from multiple cells in a range of z-planes. For this reason, cells can appear to overlap in an imaging video. It is crucial that segmentation algorithms can delineate the true boundary of apparently overlapping cells, so that the functional activity of each cell can be correctly extracted and analysed.

On synthetic data containing 25 cells, 17 of which had some overlap with another cell, we measured the success rate of ABLE’s segmentation compared to the ground truth cell locations (Fig. 3.7). For more detail on the performance metric, see Section 3.5.4. Cells were simulated with uneven brightness to mimic the ‘doughnut’ cells generated by some genetically encoded indicators that are excluded from the nucleus. Consequently, pixels within the same cell had the same pattern of temporal activity

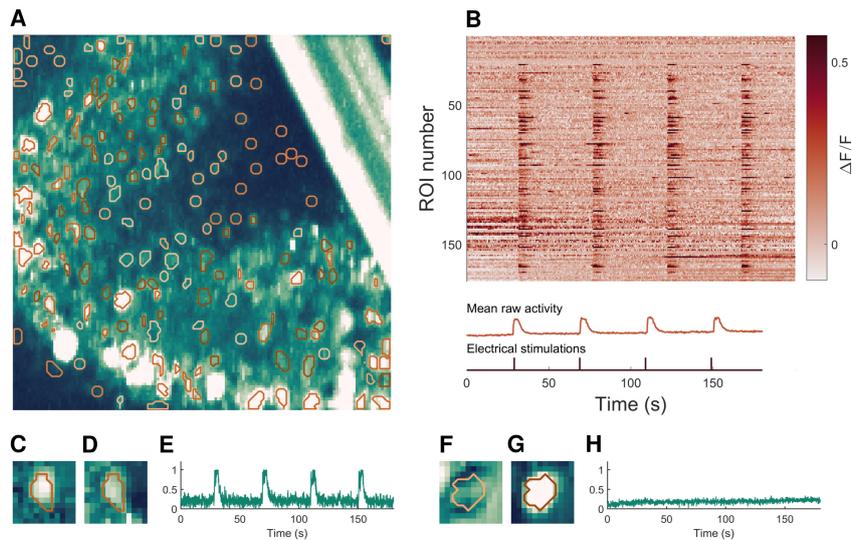


Figure 3.8: Segmentation results on in vitro imaging data containing synchronously spiking cells. The boundaries of the detected ROIs are superimposed on the thresholded maximum intensity image (A). Panel B displays the time courses of the detected ROIs (each plotted as a row of the matrix) along with the video mean raw activity and the electrical stimulations. In B, the time courses are median filtered for visual clarity. ABLE detected both active (C, D and E) and inactive ROIs (F, G and H). We display the contours of the two detected ROIs on the correlation image (C and F), the mean image (D and G) and the cellular time courses (E and H).

but vastly different baseline fluorescence, as is apparent in the mean image in Fig. 3.7D. To segment the data, therefore, we applied ABLE with the correlation-based dissimilarity metric. As can be seen in the example in Fig. 3.7A, despite uneven fluorescence and the prevalence of cellular overlap, pixels were reliably incorporated into the correct cell. We tested segmentation performance as the noise level of the data varied, measuring performance over 10 realisations of noise at each noise level. To eliminate variability due to the quality of the initialisation, we initialised cell estimates on a fixed grid (see Fig. 3.7D). On average, ABLE achieved success rate greater than 99% when the noise standard deviation was less than 90 (Fig. 3.7B). Segmentation performance deteriorated as the noise standard deviation increased; ABLE achieved an average success rate of 79% when the noise standard deviation was equal to 200.

3.6.3 Results on hippocampal brain slices

ABLE detected synchronously spiking, densely-packed cells from mouse in vitro imaging data (Fig. 3.8). As the brain slice was electrically stimulated (at rate 10Hz for 2s every 40s) during imaging, cellular activity in this dataset is highly correlated with that of other cells and the background. To segment the dataset, therefore, we applied ABLE with the Euclidean dissimilarity metric. When the cell interior and narrow-band time courses are highly correlated, the external velocity of an active contour in this version of the algorithm is predominantly driven by the discrepancy in baseline intensities. For the mathematical reasoning behind this statement, see Appendix A.2. As a consequence, the algorithm was able to segment cells whose activity was highly correlated with the background activity (Fig. 3.8B). Furthermore, inactive cells were detected when their baseline fluorescence allowed them to be identified from the background (Fig. 3.8 F-H). In Fig. 3.8C-E, we display an example of an active cell that was detected. We note that each of the four visible activity peaks in this figure corresponds to twenty calcium transients that were induced by the electrical stimulations. The relatively low amplitude of these peaks reflects the low SNR of this data. As a result, only a small amount of spontaneous activity is visible in the activity plot of Fig. 3.8B.

3.6.4 Algorithm comparison on manually labelled dataset

We compared the performance of ABLE with two state of the art calcium imaging segmentation algorithms, CNMF [105] and Suite2p [94], on a manually labelled dataset, see Fig. 3.9. We apply ABLE with the correlation-based dissimilarity metric, Eq. (3.21), which is well suited to neurons with low baseline fluorescence and uneven brightness. As the dataset is large enough (512x512x8000 pixels) to present memory issues on a standard laptop, we ran the patch-based implementation of CNMF, which processes spatially-overlapping patches of the dataset in parallel. We optimised the performance of each algorithm by selecting a range of values for each of a set of tuning parameters and generating segmentation results for all combinations of the parameter set. The results were visualised on the correlation image and the parameter set that presented the best match to the correlation image was selected. This process is representative of what a user may do in practise when applying an algorithm to a new dataset.

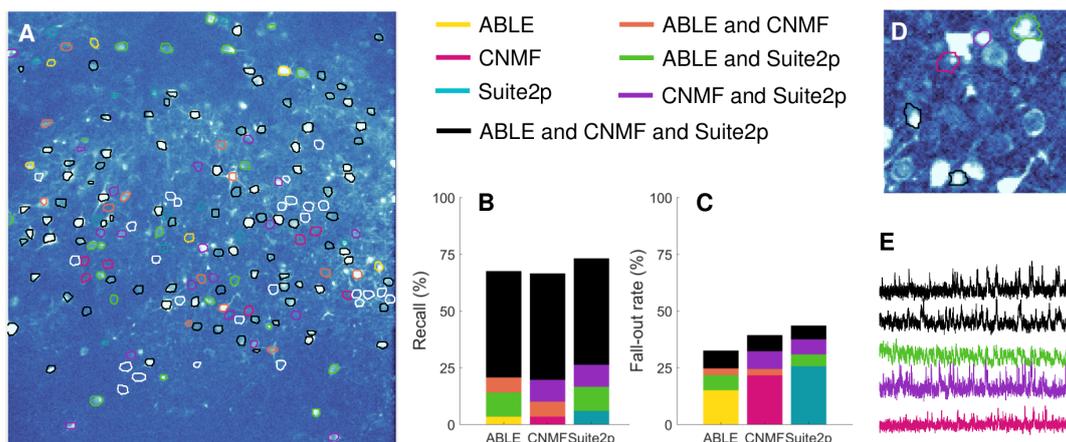


Figure 3.9: Algorithm comparison on a manually labelled dataset. We compare the segmentation results of ABL, CNMF and Suite2p on a manually labelled dataset. On the correlation image we plot the boundaries of the manually labelled cells colour-coded by the combination of algorithms that detected them (**A**), undetected cells are indicated by a white contour. Suite2p detected the highest proportion of manually labelled cells (**B**), whereas ABL had the lowest fall-out rate (**C**), which is the percentage of detected regions not present in the manual labels. Some algorithm-detected ROIs that were not present in the manual labels are detected by multiple algorithms (**D**) and have time courses which exhibit stereotypical calcium transient activity (**E**). The correlation image in **D** is thresholded to enhance visibility of local peaks in correlation. In **E**, we plot the extracted time courses of the ROIs in **D**.

ABLE achieved the highest success rate (67.5%) when compared to the manual labels, see Table 3.2. For a definition of the success rate and other performance metrics used, see Section 3.5.4. ABL achieved a lower fall-out rate than Suite2p and CNMF (Fig. 3.9C) — 67.5% of the ROIs it detected matched with the manually labelled cells. Some of the ‘false detections’ were consistent among algorithms (Fig. 3.9C) and corresponded to local peaks in the correlation image (Fig. 3.9D), whose extracted time courses displayed stereotypical calcium transient activity (Fig. 3.9E). A subset of these ROIs may thus correspond to cells omitted by the manual operator. The highest proportion of the manually labelled cells were detected by Suite2p, which detected the greatest number of cells not detected by any other algorithm (Fig. 3.9B). A small proportion (13.2%) of cells were detected by none of the algorithms. As can be seen from Fig. 3.9A, these do not correspond to peaks in the correlation image, and may reflect inactive cells detected by the manual operator.

	Success rate (%)	Precision (%)	Recall (%)
ABLE	67.5	67.5	67.5
CNMF	63.4	60.7	66.5
Suite2p	63.7	56.5	73.1

Table 3.2: Algorithm success rate on a manually labelled dataset. We compare the performance of three segmentation algorithms: ABLE, CNMF [105] and Suite2p [94], using the manual labels as ground truth.

3.7 Summary

In this chapter, we presented a novel approach to the problem of detecting cells from calcium imaging data. Our approach uses multiple coupled active contours to identify cell boundaries. The core assumption is that the local region around a single cell can be well-approximated by two subregions, the cell interior and exterior. The average time course of the respective subregions is used as a feature with which to classify pixels into either subregion. We assume that pixels in which multiple cells overlap have time courses that are well-approximated by the sum of each cell’s time course. We form a cost function based on these assumptions that is minimised when the active contours are located at the true cell boundaries.

ABLE is a flexible method: we include no priors on a region’s morphology or stereotypical temporal activity. Due to this versatility, ABLE segmented cells with varying size and shape from a mouse in vivo dataset (Fig. 3.6) and both active and inactive cells from brain slices (Fig. 3.8). Furthermore, we demonstrated its ability to detect the true boundaries of overlapping cells on both real and simulated data.

Chapter 4.

CosMIC: A consistent metric for spike inference from calcium imaging

Two-photon calcium imaging does not directly read out a neuron’s spiking activity. Rather, it reads out a secondary variable — the intracellular calcium concentration — from which spiking activity must be inferred. After an experiment, neural coding hypotheses are investigated by comparing the rate and timing of inferred spike trains to properties of stimuli and behavioural variables. To justify such analysis, the ability of spike inference algorithms to generate accurate spike train estimates must first be verified.

Inferred spikes are compared to the electrophysiological ground truth (see Section 2.2) in order to assess their accuracy. Whilst there are a rich collection of ground truth datasets [7], the metrics currently used to compare ground truth and estimated spike trains are not ideal. In this chapter, we highlight the limitations of the two most commonly used spike inference metrics, the spike train correlation and success rate, and propose an alternative, which we refer to as CosMIC. In the following, we introduce CosMIC and demonstrate its application to real and simulated spike trains.

4.1 Introduction

Although the development of spike inference algorithms has received a lot of recent attention, few researchers have examined the metrics used to assess an algorithm’s

The work presented in this chapter led to the following publication [108].

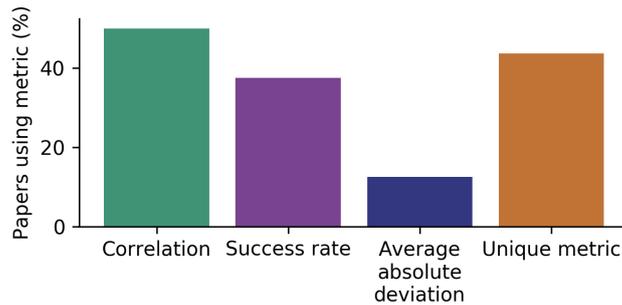


Figure 4.1: Inconsistency in usage of spike inference metrics. This chart illustrates the number of papers, of the 16 presenting a spike inference algorithm, that use each metric to assess their algorithm. Of the surveyed papers, 44% use a metric unique to that paper and 57% use more than one metric. Under the umbrella of the success rate, we include the true and false positive rates, from which the success rate is computed, and the error rate — the complement of the success rate.

performance. At present, there is no consensus regarding the best choice of metric. In fact, 44% of papers presenting a new method assess its performance using a metric unique to that paper (Fig. 4.1). This inconsistency impedes progress in the field — algorithms are not directly comparable and, consequently, data collectors cannot easily select the optimal algorithm for a new dataset. In light of this problem, a recent benchmarking study tested a range of algorithms on a wide array of imaging data [7]. Although informative, the study, which relied heavily on the spike train correlation (STC) to assess algorithm performance, may not provide the full picture.

Despite being the most commonly used metric, the STC is not well-suited to the task. It is invariant under linear transformations of the inputs and is therefore unable to evaluate the similarity of the rates of two spike trains [95]. The second most commonly used metric, the success rate, does not reward increasing temporal precision above a given threshold. Consequently, it is not an appropriate metric for evaluating an algorithm’s performance when the end goal is, for example, to investigate the synchrony of activations within a network.

Due to limiting factors, such as noise and model mismatch, it is improbable that an estimate will match a true spike with infinite temporal precision. In order to incorporate an implicit temporal tolerance, metrics manipulate the sets of spikes before comparing their similarity. One such technique is to discretise the temporal interval into a finite number of ‘bins’ and to compare the corresponding vectors of spike counts. This method, which is employed by the STC, has its limitations. In particular, estimates

the same absolute distance from a true spike can fall into different bins and thus receive different scores. Furthermore, estimates in the same bin are treated equally and, as a consequence, temporal precision above the bin width is not rewarded. An alternative strategy is to convolve spike trains with a smoothing pulse and to assess the similarity of the resulting pulse trains. This type of continuous approach is preferred by metrics assessing the relationship between spike trains from different neurons [144, 116].

In the following, we propose a metric that we refer to as CosMIC (a Consistent Metric for spike Inference from Calcium imaging). CosMIC first convolves the spike trains with a smoothing pulse whose width reflects the temporal precision that an estimate is able to achieve given the limitations of the data. To set this width, we calculate a lower bound on the temporal precision of the estimate of one spike, the Cramér-Rao bound (CRB). In the context of calcium imaging, this bound has been previously used to evaluate detectability of spikes under different imaging modalities [117]. We use the CRB as a performance benchmark; an estimator that achieves the CRB should be awarded a relatively high score.

We formalise the comparison of the resulting pulse trains by leveraging results from fuzzy set theory. In contrast to classical sets, to which an element either belongs or does not belong, fuzzy sets contain elements with a level of certainty represented by a membership function — the higher the value of the membership function, the more certain the membership. We view the pulse trains as the membership functions of the fuzzy sets of true and estimated spikes. These sets represent the original sets of spikes with a level of temporal tolerance set by the pulse width. CosMIC computes the similarity of the original sets of spikes from the size of the intersection of the fuzzy sets as a fraction of their average size. In this context, CosMIC resembles a continuous Sørensen-Dice coefficient — an index commonly used to assess the similarity of discrete, presence/absence data [32, 125].

In Section 4.2, we give a brief introduction to fuzzy set theory. Then, in Section 4.3, we detail the computation of CosMIC. The relationship between the pulse width and the CRB is explained in Section 4.4. In Section 4.6, we demonstrate CosMIC’s ability to discriminate spike train similarity on real and simulated data. CosMIC’s discriminative performance is compared to that of the success rate and the STC, whose definitions we provide for completeness in Section 4.5.

4.2 Fuzzy set theory

Fuzzy set theory was introduced by Zadeh in 1965 to extend classical set-theoretic concepts to areas in which abrupt set boundaries are hard to define [155]. As an illustrative example, Zadeh considered a universe consisting of all men and a set containing all tall men. As with many semantic definitions, it is hard to identify where the set of tall men ends and the set of ‘not tall’ men begins. Instead of forcing elements to either belong or not belong to a set, as in classical set theory, it was proposed that sets should have graded membership. Accordingly, a non-negative value is associated with each element of the universe, representing the degree of certainty to which that element belongs to a set. Here, we provide a brief introduction to fuzzy sets. We refer the interested reader to [39] for an extensive theoretical overview and to [157] for more recent developments. To differentiate between classical sets and fuzzy sets, we use plain font for the former (S) and script-like font for the latter (\mathcal{S}).

In classical set theory, membership of a set, S , is defined for all elements of a universe, X , using a characteristic function:

$$\mathbf{1}_S(x) = \begin{cases} 1 & \text{if and only if } x \in S, \\ 0 & \text{if and only if } x \notin S. \end{cases} \quad (4.1)$$

In fuzzy set theory, the range of the analogous function, $\mu_S(x)$, which we refer to as a membership function, is the non-negative, real numbers. The value of μ_S increases as the certainty of membership does. A fuzzy set is characterised by a set of ordered pairs:

$$\mathcal{S} = \{(x, \mu_S(x)) : x \in X\}, \quad (4.2)$$

each representing an element of the universe and the certainty with which that element belongs to the set. It is not the absolute value of a membership function that is of importance. Rather, it is the implicit ordering induced by a membership function that we are interested in. For example, we may want to compare the values of $\mu_S(x)$ and $\mu_S(y)$ or, alternatively, $\mu_{S_1}(x)$ and $\mu_{S_2}(x)$. We note that the characteristic function of a classical set is a special case of a membership function that only takes the values $\{0, 1\}$.

The operations of intersection and union were introduced to fuzzy sets by Zadeh [155] and, at a later date, justified axiomatically by Bellman and Giertz [6]. The membership

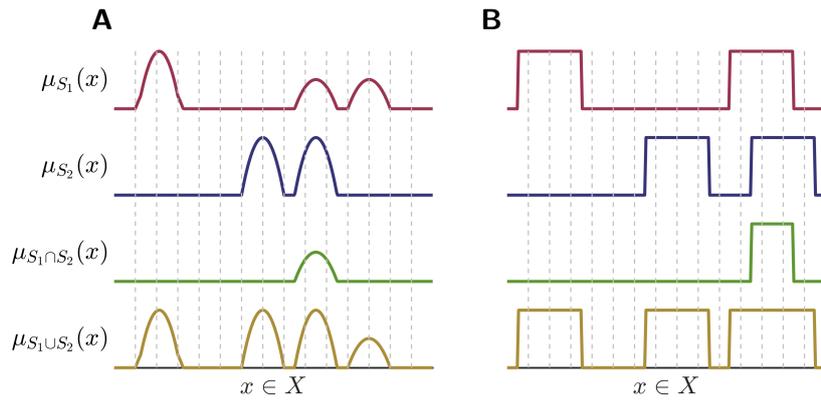


Figure 4.2: Membership functions of the union and intersection of fuzzy sets. In **A**, the membership functions are continuous-valued. In **B**, they are binary and are therefore analogous to the characteristic functions of classical sets.

function of an intersection of fuzzy sets is computed as

$$\mu_{S_1 \cap S_2}(x) = \min(\mu_{S_1}(x), \mu_{S_2}(x)), \quad \forall x \in X. \quad (4.3)$$

Similarly, the membership function of the union is

$$\mu_{S_1 \cup S_2}(x) = \max(\mu_{S_1}(x), \mu_{S_2}(x)), \quad \forall x \in X. \quad (4.4)$$

These definitions coincide with the classical operations of intersection and union in the special case in which the membership functions are characteristic functions. We show an illustrative example of the above operations in Fig. 4.2.

When the universe is a finite set, the cardinality of a fuzzy set, $|\mathcal{S}|_c$, is [39]

$$|\mathcal{S}|_c = \sum_{x \in X} \mu_{\mathcal{S}}(x). \quad (4.5)$$

When the universe is infinite, the cardinality of a set may not exist. In the case that X is a measurable set and P is a measure on X [113], then

$$|\mathcal{S}|_c = \int_X \mu_{\mathcal{S}}(x) \, dP(x). \quad (4.6)$$

In particular, when the universe is the real line, we have

$$|\mathcal{S}|_c = \int_{\mathbb{R}} \mu_{\mathcal{S}}(x) \, dx. \quad (4.7)$$

Finally, the support of \mathcal{S} is: $\text{supp}(\mathcal{S}) = \{x \in X : \mu_{\mathcal{S}}(x) > 0\}$.

4.2.1 Sørensen-Dice coefficient

The Sørensen-Dice coefficient (SDC) was designed to quantify similarity between samples of discrete, presence/absence data [125, 32]. This metric, which is also known as the F1-score, is widely used in many fields, including ecology [102] and image segmentation [159]. Although its initial justification was empirical rather than theoretical, it can be interpreted as a similarity metric for fuzzy sets with binary membership functions belonging to a finite universe [96, 137]. This connection was made after the SDC had become established in numerous fields. For example, an ecologist, Roberts, used fuzzy sets as a theoretical basis for a popular technique for environmental analyses [111]. In the following, we detail the computation of the SDC from the perspective of fuzzy set theory.

Let \mathcal{S}_1 and \mathcal{S}_2 be two fuzzy sets defined on a finite universe, $X = \{x_1, x_2, \dots, x_m\}$, represented by the membership functions $\mu_{\mathcal{S}_1}$ and $\mu_{\mathcal{S}_2}$, respectively. We assume the membership functions take values in the range $\{0, 1\}$. Then, the SDC is

$$\text{SDC}(\mathcal{S}_1, \mathcal{S}_2) = \frac{|\mathcal{S}_1 \cap \mathcal{S}_2|_c}{(|\mathcal{S}_1|_c + |\mathcal{S}_2|_c) / 2} \quad (4.8)$$

$$= \frac{\sum_{i=1}^m \min(\mu_{\mathcal{S}_1}(x_i), \mu_{\mathcal{S}_2}(x_i))}{(\sum_{i=1}^m \mu_{\mathcal{S}_1}(x_i) + \mu_{\mathcal{S}_2}(x_i)) / 2}, \quad (4.9)$$

where Eq. (4.9) follows from Eq. (4.3) and Eq. (4.5).

When \mathcal{S}_1 and \mathcal{S}_2 represent a set of ground truth items and a set of estimates, respectively, the coefficient has two alternative formulations in terms of the types of errors. It is perhaps the intuitive appeal of these expressions that has led to the SDC's popularity. To illustrate these alternative formulations, we consider the example of image segmentation, where we want to quantify the proportion of pixels shared by a ground truth mask and an estimate. The binary formulation of the SDC is well-suited to this situation; if the ground truth mask contains the i^{th} pixel, then $\mu_{\mathcal{S}_1}(x_i) = 1$, otherwise $\mu_{\mathcal{S}_1}(x_i) = 0$. Similarly, if the i^{th} pixel is in the estimated mask then $\mu_{\mathcal{S}_2}(x_i) = 1$, otherwise $\mu_{\mathcal{S}_2}(x_i) = 0$. We write the sets of true positives (TP), false positives (FP)

and false negatives (FN) as:

$$\text{TP} = \{x_i : \mu_{\mathcal{S}_1}(x_i) = 1 \cap \mu_{\mathcal{S}_2}(x_i) = 1\}, \quad (4.10)$$

$$\text{FP} = \{x_i : \mu_{\mathcal{S}_1}(x_i) = 0 \cap \mu_{\mathcal{S}_2}(x_i) = 1\}, \quad (4.11)$$

$$\text{FN} = \{x_i : \mu_{\mathcal{S}_1}(x_i) = 1 \cap \mu_{\mathcal{S}_2}(x_i) = 0\}. \quad (4.12)$$

The number of pixels in the ground truth mask can be calculated from the number of true positives and false negatives, such that

$$\sum_{i=1}^m \mu_{\mathcal{S}_1}(x_i) = |\{x_i : \mu_{\mathcal{S}_1}(x_i) = 1\}|_c \quad (4.13)$$

$$= |\{x_i : \mu_{\mathcal{S}_1}(x_i) = 1 \cap \mu_{\mathcal{S}_2}(x_i) = 1\} \cup \{x_i : \mu_{\mathcal{S}_1}(x_i) = 1 \cap \mu_{\mathcal{S}_2}(x_i) = 0\}|_c \quad (4.14)$$

$$= |\text{TP}|_c + |\text{FN}|_c. \quad (4.15)$$

Eq. (4.15) follows from the fact that the cardinality of the union of disjoint sets is the sum of the cardinalities. By analogous logic, it follows that

$$\sum_{i=1}^m \mu_{\mathcal{S}_2}(x_i) = |\text{TP}|_c + |\text{FP}|_c, \quad (4.16)$$

$$\text{and } \sum_{i=1}^m \min(\mu_{\mathcal{S}_1}(x_i), \mu_{\mathcal{S}_2}(x_i)) = |\text{TP}|_c. \quad (4.17)$$

Then, substituting these identities into Eq. (4.9), the SDC can be expressed in equivalent form:

$$\text{SDC}(\mathcal{S}_1, \mathcal{S}_2) = \frac{2 |\text{TP}|_c}{2 |\text{TP}|_c + |\text{FP}|_c + |\text{FN}|_c}. \quad (4.18)$$

Alternatively, defining the precision as the proportion of pixels (or, more generally, items) in the estimated mask that are true positives and the recall as the proportion of ground truth pixels that are true positives, the SDC is, equivalently, the harmonic mean of the precision and recall:

$$\frac{\text{precision} * \text{recall}}{(\text{precision} + \text{recall}) / 2} = \frac{\frac{|\text{TP}|_c}{|\text{TP}|_c + |\text{FP}|_c} * \frac{|\text{TP}|_c}{|\text{TP}|_c + |\text{FN}|_c}}{\left(\frac{|\text{TP}|_c}{|\text{TP}|_c + |\text{FP}|_c} + \frac{|\text{TP}|_c}{|\text{TP}|_c + |\text{FN}|_c} \right) / 2} \quad (4.19)$$

$$= \frac{2|\text{TP}|_c}{2|\text{TP}|_c + |\text{FP}|_c + |\text{FN}|_c} \quad (4.20)$$

$$= \text{SDC}(\mathcal{S}_1, \mathcal{S}_2). \quad (4.21)$$

A generalised SDC that quantifies common abundance is used to assess similarity of non-binary data [13]. The only difference in the calculation of this coefficient from Eq. (4.8) is that the entries of $\mu_{\mathcal{S}_1}$ and $\mu_{\mathcal{S}_2}$ can assume any non-negative value. It is common that the i^{th} entries of $\mu_{\mathcal{S}_1}$ and $\mu_{\mathcal{S}_2}$ represent counts of the corresponding objects at the i^{th} location. Interpreting these frequency counts as probabilities of object occurrence is consistent with an empirical approach to probability. In the fuzzy set literature, this coefficient has been further generalised to compare non-quantitative data [137].

4.3 Constructing the metric

A spike inference metric compares the similarity between two sets of spikes: a ground truth set, $\mathcal{S} = \{t_k\}_{k=1}^K$, and a set of estimates, $\hat{\mathcal{S}} = \{\hat{t}_k\}_{k=1}^{\hat{K}}$. In the following, we define two fuzzy sets, \mathcal{S}_ϵ and $\hat{\mathcal{S}}_\epsilon$, which represent the original sets of spikes, \mathcal{S} and $\hat{\mathcal{S}}$, with a level of temporal tolerance defined by a parameter ϵ . We set ϵ to reflect the temporal precision that an estimate is able to achieve given the statistics of a dataset (see Section 4.4.4). The corresponding membership functions are defined on the real line and are denoted $\mu(t)$ and $\hat{\mu}(t)$. The membership functions are obtained through convolution of the spike trains,

$$x(t) = \sum_{k=1}^K \delta(t - t_k) \quad \text{and} \quad \hat{x}(t) = \sum_{k=1}^{\hat{K}} \delta(t - \hat{t}_k), \quad (4.22)$$

with a triangular pulse. The resulting functions have local maxima at the locations of the respective sets of spikes (Fig. 4.3A). As $x(t)$ and $\hat{x}(t)$ are analogous to the membership functions of the classical sets of spikes, we can think of the convolution as a temporal smoothing of the membership. The pulse that we employ is a triangular B-spline (Fig. 4.3B),

$$q_\epsilon(t) = \begin{cases} \frac{\epsilon - |t|}{\epsilon} & |t| \leq \epsilon, \\ 0 & \text{otherwise.} \end{cases} \quad (4.23)$$

Using this triangular pulse means that, the further a time point is from a spike, the less weight the membership function receives at that point. Past a certain distance, ϵ , the membership function receives no weight. Many pulse shapes could be chosen to introduce this grading of temporal precision, we select a triangular pulse as it is

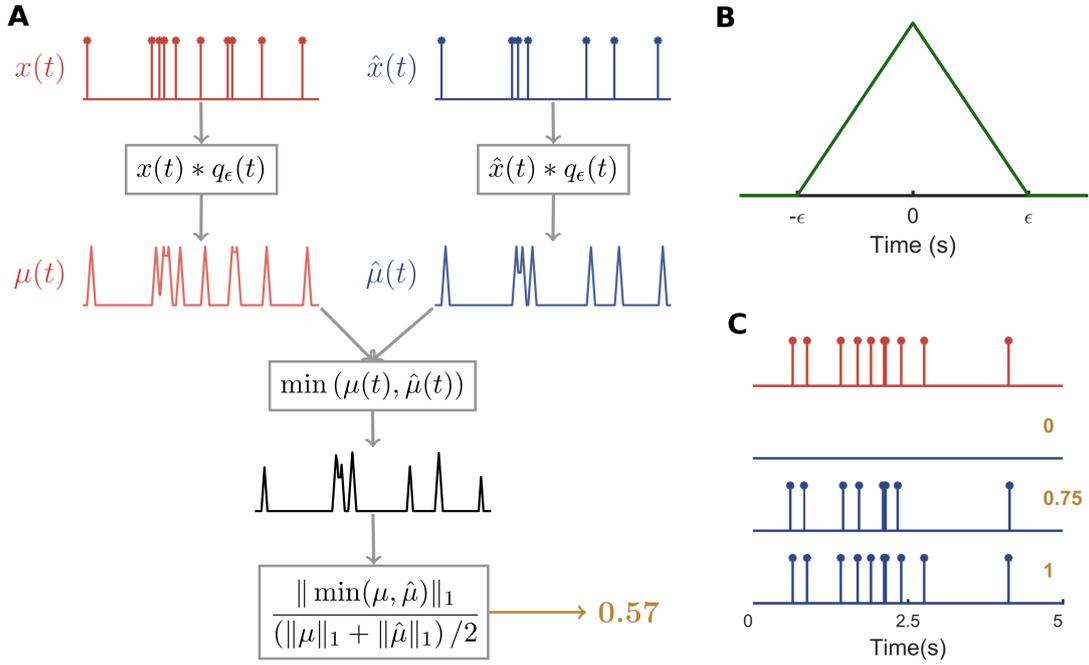


Figure 4.3: Flow diagram of the proposed metric. The ground truth spike train and estimated spike train are convolved with a triangular pulse (**B**), whose width is determined by the statistics of the data. The metric compares the difference between the resulting pulse trains (**A**). Metric scores are in the range $[0,1]$ — a perfect estimate achieves score 1 and an empty spike train is scored 0 (**C**).

straightforward to examine analytically and implement computationally.

We design the proposed metric to quantify the size of the intersection of the fuzzy sets of true and estimated spikes with respect to the average size of the sets, such that

$$M(S, \hat{S}) = \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{(|\mathcal{S}_\epsilon|_c + |\hat{\mathcal{S}}_\epsilon|_c) / 2}. \quad (4.24)$$

From Eq. (4.8), the analogy between this metric and the SDC is clear. Whereas the SDC operates on binary membership function defined on finite universes, CosMIC operates on smooth membership functions defined on the real line. From the definitions of intersection, Eq. (4.3), and cardinality, Eq. (4.7), we have

$$M(S, \hat{S}) = 2 \frac{\int_{\mathbb{R}} \min(\mu(t), \hat{\mu}(t)) dt}{\int_{\mathbb{R}} \mu(t) dt + \int_{\mathbb{R}} \hat{\mu}(t) dt} = 2 \frac{\|\min(\mu, \hat{\mu})\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1}, \quad (4.25)$$

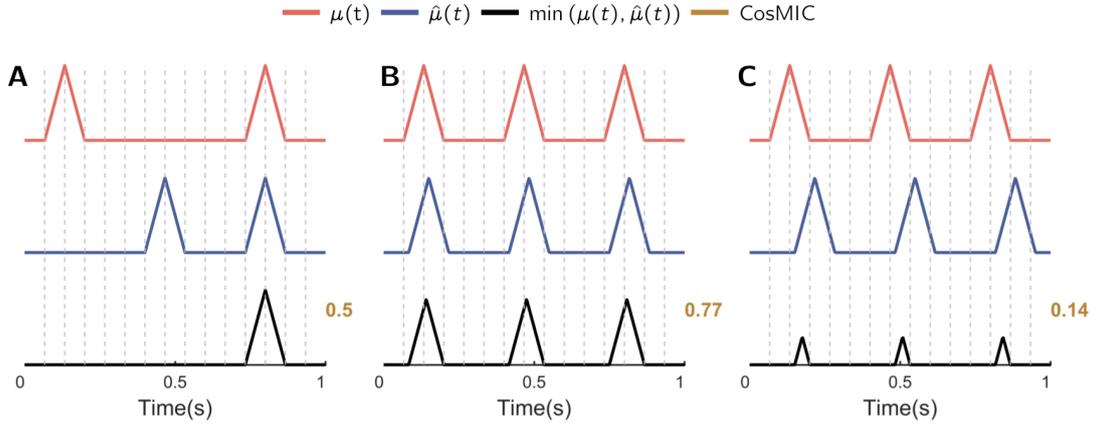


Figure 4.4: Proposed metric example. The proposed metric quantifies the commonalities of the sets of true and estimated spikes as a proportion of the average size of those sets. Commonalities are found by taking the minimum of the pulse trains — as such, spikes that appear in only one pulse train are excluded (**A**) and estimates with lower temporal precision receive a lower score (**B** and **C**).

where we denote the continuous L1-norm with $\|f\|_1 := \int_{\mathbb{R}} |f(t)| dt$. We note that the absolute value is not necessary in the integrand of Eq. (4.25) as the membership functions are non-negative. Taking the minimum of the membership functions produces a conservative representation of the intersection of two sets; in our context, spikes that appear in one spike train and not in the other are removed (Fig. 4.4A) and spikes that are detected with poor temporal precision are assigned less weight (Fig. 4.4B and 4.4C).

Defining $D^+ = \{t \in \mathbb{R} : \mu(t) > \hat{\mu}(t)\}$ and $D^- = \{t \in \mathbb{R} : \mu(t) \leq \hat{\mu}(t)\}$, we have

$$\|\mu - \hat{\mu}\|_1 = \int_{\mathbb{R}} |\mu(t) - \hat{\mu}(t)| dt = \int_{D^+} \mu(t) - \hat{\mu}(t) dt + \int_{D^-} \hat{\mu}(t) - \mu(t) dt, \quad (4.26)$$

and

$$\|\mu\|_1 + \|\hat{\mu}\|_1 - \|\mu - \hat{\mu}\|_1 = 2 \left(\int_{D^+} \hat{\mu}(t) dt + \int_{D^-} \mu(t) dt \right) \quad (4.27)$$

$$= 2 \int_{\mathbb{R}} \min(\mu(t), \hat{\mu}(t)) dt \quad (4.28)$$

$$= 2\|\min(\mu, \hat{\mu})\|_1. \quad (4.29)$$

It follows that an alternative expression for the metric is

$$M(S, \hat{S}) = \frac{2\|\min(\mu, \hat{\mu})\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1} \quad (4.30)$$

$$= \frac{\|\mu\|_1 + \|\hat{\mu}\|_1 - \|\mu - \hat{\mu}\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1} \quad (4.31)$$

$$= 1 - \frac{\|\mu - \hat{\mu}\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1}. \quad (4.32)$$

From Eq. (4.32), it is clear that the maximal score of 1 is achieved when the membership functions, and therefore the sets of true and estimated spikes, are equivalent. The minimal score of 0 is achieved when the support of the membership functions do not overlap, i.e. no estimates are within the tolerance of the metric.

4.3.1 Ancestor metrics

Like the SDC, CosMIC can alternatively be derived from a pair of metrics, which we refer to as ancestor metrics. The first of these metrics measures the proportion of ground truth spikes that were detected within the precision of the pulse width, such that

$$R_{\text{CosMIC}} = \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\mathcal{S}_\epsilon|_c} = \frac{\|\min(\mu, \hat{\mu})\|_1}{\|\mu\|_1}. \quad (4.33)$$

This score is analogous to the recall of a spike train estimate, one of the ancestor metrics from which the SDC is formed. The second of CosMIC's ancestor metrics measures the proportion of estimated spikes that detect a ground truth spike within the precision of the pulse width, such that

$$P_{\text{CosMIC}} = \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\hat{\mathcal{S}}_\epsilon|_c} = \frac{\|\min(\mu, \hat{\mu})\|_1}{\|\hat{\mu}\|_1}. \quad (4.34)$$

This is analogous to the precision, the second metric used to compute the SDC. Finally, computing the harmonic mean of the two ancestor metrics and rearranging, we obtain CosMIC:

$$2 \frac{R_{\text{CosMIC}} * P_{\text{CosMIC}}}{R_{\text{CosMIC}} + P_{\text{CosMIC}}} = 2 \frac{\frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\mathcal{S}_\epsilon|_c} \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\hat{\mathcal{S}}_\epsilon|_c}}{\frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\mathcal{S}_\epsilon|_c} + \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\hat{\mathcal{S}}_\epsilon|_c}} = 2 \frac{|\mathcal{S}_\epsilon \cap \hat{\mathcal{S}}_\epsilon|_c}{|\mathcal{S}_\epsilon|_c + |\hat{\mathcal{S}}_\epsilon|_c} = M(S, \hat{S}). \quad (4.35)$$

4.4 Temporal error tolerance

The width of the triangular pulse with which the spike trains are convolved reflects the accepted tolerance of an estimated spike’s position with respect to the ground truth. To set this width, we calculate a lower bound on the temporal precision of the estimate of one spike — the Cramér-Rao bound (CRB) — from the statistics of the data. The CRB is a lower bound on the mean square error of any unbiased estimator. It is therefore useful as a benchmark; an estimator that achieves the CRB should be awarded a relatively high metric score. In the following, we discuss the mean square error of a parameter estimator (Section 4.4.1), introduce the CRB (Section 4.4.2) and calculate this bound in the context of calcium imaging (Section 4.4.3). We relate this bound to CosMIC’s pulse width in Section 4.4.4.

4.4.1 Parameter estimation

Let $g(t, \theta)$ be a signal evolving in time, t , which is dependent on a univariate, deterministic parameter, $\theta \in \mathbb{R}$. We consider the situation where the parameter θ is unknown and the goal is to infer its value from noisy samples of $g(t, \theta)$. We have

$$\tilde{g}[n] = g[n, \theta] + \xi[n] \quad \text{for } n \in \{0, \dots, N - 1\}, \quad (4.36)$$

where $g[n, \theta] = g(nT, \theta)$ are deterministic samples of $g(\cdot, \theta)$ at time resolution T and $\xi[n]$ are samples of the noise.

Denoting the vector of data samples with $\tilde{\mathbf{g}} = (\tilde{g}[0], \tilde{g}[1], \dots, \tilde{g}[N - 1])$, we now consider the error of the estimator, $\hat{\theta}(\tilde{\mathbf{g}})$, constructed to infer the value of θ . The average accuracy of the estimator can be represented by the mean square error (MSE):

$$\text{MSE}(\hat{\theta}) = \mathbb{E} \left[(\hat{\theta} - \theta)^2 \right]. \quad (4.37)$$

The MSE can be decomposed into two distinct types of error, in what is known as the bias-variance tradeoff [62], such that

$$\text{MSE}(\hat{\theta}) = \left(\mathbb{E}[\hat{\theta}] - \theta \right)^2 + \text{var}(\hat{\theta}). \quad (4.38)$$

The bias, $\mathbb{E}[\hat{\theta}] - \theta$, is a systematic error in the estimator’s output compared to the true value. The variance, $\text{var}(\hat{\theta})$, represents the degree of variability in the estimator’s

output. We restrict our attention to unbiased estimators, which are those that, on average, attain the correct value for all values of θ :

$$\mathbb{E} \left[\hat{\theta} \right] = \theta. \quad (4.39)$$

Within this class of estimators, the one that achieves the minimum MSE is the one that has the smallest variance. One approach to identifying an optimal estimator is, therefore, to identify within the class of unbiased estimators the one with the smallest variance. To do this, one can compute a lower bound on the variance of an unbiased estimator. This bound can subsequently be used as a benchmark; the quality of an estimator is judged by how close it is to attaining the bound.

4.4.2 Cramér-Rao bound

Each sample, $\tilde{g}[n]$, is the sum of a deterministic signal component whose value is dependent on the parameter θ and a stochastic noise component, see Eq. (4.36). The vector of samples, $\tilde{\mathbf{g}}$, can thus be modelled by a PDF parametrised by θ : $p(\tilde{\mathbf{g}}; \theta)$. From these components, a lower bound on the variability of an unbiased estimator, $\hat{\theta}$, can be calculated. There are many lower bounds on the variance of unbiased estimators, for example [158, 77]. Of all the bounds, the CRB is the most straightforward to compute and is accompanied by a set of conditions under which it is satisfied [62]. In Theorem 4.1, we present, without proof, the CRB for a scalar, deterministic parameter, $\theta \in \mathbb{R}$. For a more detailed exposition of the theory, including the result for vector-valued parameters and a proof, we refer the reader to [62].

Theorem 4.1. *The variance of an unbiased estimator satisfies*

$$\text{var} \left(\hat{\theta} \right) \geq \frac{1}{-\mathbb{E} \left[\frac{\partial^2 \ln p(\tilde{\mathbf{g}}; \theta)}{\partial \theta^2} \right]}, \quad (4.40)$$

if the following regularity condition holds:

$$\mathbb{E} \left[\frac{\partial \ln p(\tilde{\mathbf{g}}; \theta)}{\partial \theta} \right] = 0, \quad \forall \theta. \quad (4.41)$$

Here, the expectation is taken with respect to $\tilde{\mathbf{g}}$ and the derivatives are evaluated at the true value of θ .

We now derive the specific form of the CRB when the samples are independent and

identically distributed (i.i.d.) and are corrupted by white Gaussian noise. In the case of independence, we have

$$p(\tilde{\mathbf{g}}; \theta) = \prod_{i=0}^{N-1} p_i(\tilde{g}[i]; \theta), \quad (4.42)$$

where p_i is the PDF of $\tilde{g}[i]$. When the samples are identically distributed, we have $p_i = q$ for all $i \in \{0, 1, \dots, N-1\}$ and some PDF q . In particular, when $\xi[n]$ are i.i.d. samples of white Gaussian noise with variance σ^2 , we have

$$\ln p(\tilde{\mathbf{g}}; \theta) = \ln \left(\prod_{i=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(\tilde{g}[i] - g[i; \theta])^2}{2\sigma^2} \right) \right) \quad (4.43)$$

$$= -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=0}^{N-1} (\tilde{g}[i] - g[i; \theta])^2. \quad (4.44)$$

It follows that

$$\frac{\partial \ln p(\tilde{\mathbf{g}}; \theta)}{\partial \theta} = \frac{1}{\sigma^2} \sum_{i=0}^{N-1} (\tilde{g}[i] - g[i; \theta]) \frac{\partial g[i; \theta]}{\partial \theta}. \quad (4.45)$$

As $\mathbb{E}[\tilde{g}[i] - g[i; \theta]] = 0$, it is clear from Eq. (4.41) and the linearity of the expectation that the regularity condition is satisfied when the derivative, $\frac{\partial g[i; \theta]}{\partial \theta}$, exists. Differentiating again, we have

$$\frac{\partial^2 \ln p(\tilde{\mathbf{g}}; \theta)}{\partial \theta^2} = \frac{1}{\sigma^2} \sum_{i=0}^{N-1} (\tilde{g}[i] - g[i; \theta]) \frac{\partial^2 g[i; \theta]}{\partial \theta^2} - \left(\frac{\partial g[i; \theta]}{\partial \theta} \right)^2. \quad (4.46)$$

Taking the expectation removes the first term of the above formula, which has expectation equal to 0. As a result, the variance of an unbiased estimator in this setting satisfies:

$$\text{var}(\hat{\theta}) \geq \frac{\sigma^2}{\sum_{i=0}^{N-1} \left(\frac{\partial g[i; \theta]}{\partial \theta} \right)^2}, \quad (4.47)$$

as long as the regularity condition in Eq. (4.41) is satisfied.

4.4.3 Bound for spike inference

We consider the problem of estimating the location of one spike, t_1 , from noisy calcium imaging data. We assume that we have access to noisy samples, $\tilde{f}[n] = f[n] + \xi[n]$, for $n \in \{0, 1, \dots, N-1\}$. Here, $f[n] = f(nT)$ are samples of the fluorescence signal with time resolution T and $\xi[n]$ are i.i.d samples of white Gaussian noise with variance σ^2 .

The fluorescence signal is modelled as

$$f(t) = \begin{cases} Ae^{-(t-t_1)/\tau_{\text{off}}} (1 - e^{-(t-t_1)/\tau_{\text{on}}}) & \text{for } t > t_1 \\ 0 & \text{otherwise,} \end{cases} \quad (4.48)$$

where τ_{on} and τ_{off} define the speed of the rise and decay of the fluorescence pulse, respectively, and A is the amplitude.

From Eq. (4.47), the CRB of the parameter estimator, \hat{t}_1 , is

$$\text{var}(\hat{t}_1) \geq \frac{\sigma^2}{\sum_{i=0}^{N-1} \left(\frac{\partial f[n]}{\partial t_1} \right)^2}. \quad (4.49)$$

To compute the CRB, we must therefore evaluate the partial derivative of $f(t)$ with respect to t_1 at $t \in \{0, T, 2T, \dots, (N-1)T\}$. For $t > t_1$, we have

$$\frac{\partial f}{\partial t_1} = Ae^{-(t-t_1)/\tau_{\text{off}}} \left(\frac{1}{\tau_{\text{off}}} - \left(\frac{1}{\tau_{\text{off}}} + \frac{1}{\tau_{\text{on}}} \right) e^{-(t-t_1)/\tau_{\text{on}}} \right). \quad (4.50)$$

For $t < 0$, the derivative is equal to 0. To evaluate the derivative at $t = t_1$, we consider the left and right derivatives [134], starting with the former:

$$\lim_{t \rightarrow t_1^-} \frac{f(t_1) - f(t)}{t_1 - t} = \lim_{t \rightarrow t_1^-} \frac{0 - 0}{t_1 - t} = 0. \quad (4.51)$$

The right derivative is

$$\lim_{t \rightarrow t_1^+} \frac{f(t) - f(t_1)}{t - t_1} = \lim_{h \rightarrow 0^+} \frac{f(t_1 + h) - f(t_1)}{h} \quad (4.52)$$

$$= \lim_{h \rightarrow 0^+} \frac{Ae^{-h/\tau_{\text{off}}} (1 - e^{-h/\tau_{\text{on}}})}{h} \quad (4.53)$$

$$= A \lim_{h \rightarrow 0^+} \frac{\left(1 - \frac{h}{\tau_{\text{off}}} + O(h^2)\right) \left(\frac{h}{\tau_{\text{on}}} + O(h^2)\right)}{h} \quad (4.54)$$

$$= A \lim_{h \rightarrow 0^+} \frac{\frac{h}{\tau_{\text{on}}} + O(h^2)}{h} \quad (4.55)$$

$$= \frac{A}{\tau_{\text{on}}}. \quad (4.56)$$

As the left and right derivatives of $f(t)$ are not equal at $t = t_1$, the derivative does not exist at that point. When we compute the CRB, therefore, we ensure that no samples are taken at $nT = t_1$.

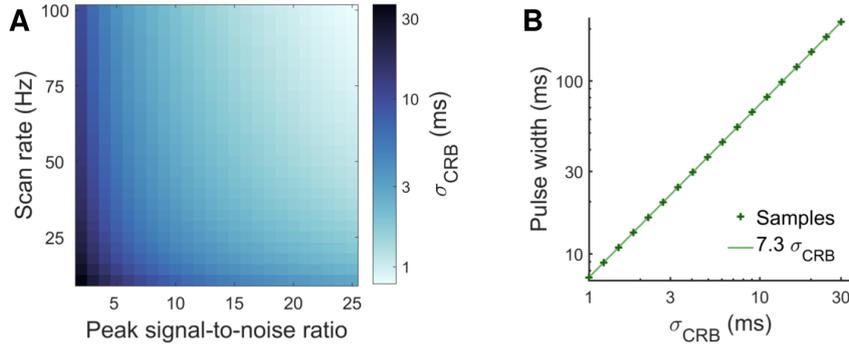


Figure 4.5: Metric pulse width. The pulse width is set to reflect the temporal precision achievable given the statistics of the dataset. We calculate the Cramér-Rao bound (CRB), σ_{CRB}^2 , a lower bound on the mean square error of the estimated location of one spike from calcium imaging data (**A**). This bound decreases as the scan rate (Hz) and peak signal-to-noise ratio (squared calcium transient peak amplitude/noise variance) increase. We set the pulse width to ensure that an estimate of one spike at the temporal precision of the CRB achieves, on average, a score of 0.8. This results in a pulse width of approximately $7.3 \sigma_{\text{CRB}}$ (**B**).

Substituting the value of the derivative into Eq. (4.49), we have

$$\text{var}(\hat{t}_1) \geq \left[\frac{A^2}{\sigma^2} \sum_{n=\lceil t_1/T \rceil}^{N-1} e^{-2(t-t_1)/\tau_{\text{off}}} \left(\frac{1}{\tau_{\text{off}}} - \left(\frac{1}{\tau_{\text{off}}} + \frac{1}{\tau_{\text{on}}} \right) e^{-(t-t_1)/\tau_{\text{on}}} \right)^2 \right]^{-1}. \quad (4.57)$$

In order that the CRB holds for an arbitrarily placed spike, we remove the dependency on the true spike time by averaging the result over several values of t_1 . Denoting $\text{CRB}(t_1)$ as the right hand side of Eq. (4.57), we compute $\sigma_{\text{CRB}}^2 = \frac{1}{M} \sum_{m=1}^M \text{CRB}(t_1^m)$, where t_1^m are evenly placed in the interval $(nT, (n+1)T)$ for a fixed n .

In Fig. 4.5, we plot σ_{CRB} as the sampling rate and peak signal-to-noise ratio (PSNR) of the data vary. The PSNR is computed as $A_{\text{peak}}^2/\sigma^2$, where A_{peak} is the peak amplitude of the fluorescence signal (the maximum) and σ is the standard deviation of the noise. For this example, we use $\tau_{\text{on}} = 32\text{ms}$ and $\tau_{\text{off}} = 314\text{ms}$; the parameters for a Cal-520 pulse [133]. We see that the CRB decreases as either the scan rate or the PSNR of the data increases.

4.4.4 Pulse width

The CRB can be used as a benchmark for temporal precision of any unbiased estimator. As such, we set the pulse width to ensure that, on average, an estimate at the precision of the CRB achieves a relatively high score. We set the benchmark metric score at 0.8, as this represents a relatively high value in the range of the metric, which is between 0 and 1. The importance of this score is not the particular benchmark value — there are a range of values that give similar performance — but rather that it is a reproducible number with a clear interpretation. In this chapter, we characterise the discrimination performance of CosMIC with a benchmark value of 0.8, so that its scores can be interpreted when applied to spike inference algorithms on real data. The benchmark value was set lower than the metric’s maximum value, 1, so that the score does not saturate when the model assumptions are not ideally satisfied. On real data, the noise may not be stationary (σ may vary), and so it may be possible for algorithms to exceed the theoretical CRB. A benchmark score of 0.8 allows leeway to be exceeded in this scenario.

We consider a true spike at t_1 and an estimate, U , normally distributed around it at the precision of the CRB, such that $U \sim \mathcal{N}(t_1, \sigma_{\text{CRB}}^2)$. Then, we fix the pulse width so that, on average, $\mathbb{E}[M(t_1, U)] = 0.8$. In Appendix B, we show that this condition is satisfied when

$$0.4 = (\Phi(1/\beta) - 0.5) (\beta^2 + 1) + \frac{\beta}{\sqrt{2\pi}} (\exp(-1/2\beta^2) - 2), \quad (4.58)$$

where $\beta = \sigma_{\text{CRB}}/w$, w is the pulse width and Φ denotes the cumulative distribution function of the standard normal distribution. We observe empirically that the pulse width that solves this equation is approximately equal to $7.3\sigma_{\text{CRB}}$ (Fig. 4.5B).

4.5 Experimental methods

In Section 4.6, we compare CosMIC with the two most commonly used metrics in the spike inference literature, the success rate and STC, which we define in Sections 4.5.1 and 4.5.2, respectively. We fix the parameters of both metrics with respect to CosMIC’s pulse width, to ensure that the scores are comparable (Fig. 4.6). The metrics are assessed on both simulated and real data. The latter dataset, which was

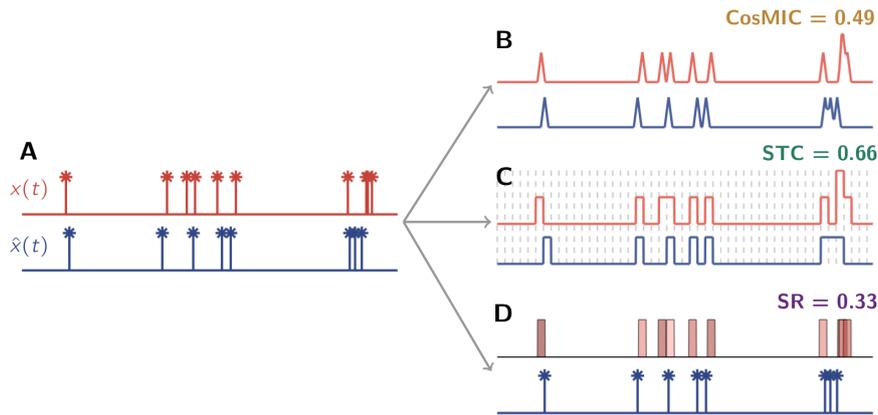


Figure 4.6: Illustration of common metrics. In this chapter, we compare the scores of three metrics: CosMIC, the spike train correlation (STC) and the success rate (SR). None of the metrics compute scores directly from the true and estimated spike trains, $x(t)$ and $\hat{x}(t)$, shown in **A**. Rather, CosMIC initially convolves the spike trains with a triangular pulse (**B**). The STC first discretises the temporal interval and utilises the counts of spikes in each time bin, the bin edges and counts are plotted in **C**. The SR uses a bin centered around each true spike — an estimate in that bin is deemed a true detection (**D**). In order that the metric scores are comparable, we fix the STC and SR bin widths to be equal to CosMIC’s pulse width.

collected by Dr T. Abrahamsson and Dr P. J. Sjöström, our collaborators at McGill University, is described in Section 4.5.3.

4.5.1 Success rate

The success rate, which is defined as a function of the true and false positive rates or, alternatively, as a function of precision and recall, appears in various forms in the literature. Spike inference performance has been assessed using true and false positive rates [106], precision and recall analysis [107] and using the complement of the success rate, the error rate [30]. We study this class of metrics under the umbrella of the success rate, which we define here.

A ground truth spike is deemed to have been ‘detected’ if there is an estimate within $\delta_1/2$ (s) of that spike, where δ_1 is a free parameter. Only one estimate can be deemed to detect one ground truth spike. The recall is the percentage of ground truth spikes that were detected. The precision is the percentage of estimates that detect a ground

truth spike. Then, the success rate is the harmonic mean of the precision and recall, such that

$$\text{success rate} = 2 \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (4.59)$$

From this formula and Eq. (4.19), the relationship between the success rate and the SDC is apparent. A binary true detection region centred around each ground truth spike is analogous to an implementation of CosMIC with a box function pulse. To ensure that the success rate ‘pulse’ has the same width as CosMIC’s pulse, we set $\delta_1 = w$, where w is CosMIC’s pulse width.

4.5.2 Spike train correlation

The first step in the calculation of the STC is the discretisation of the temporal interval into bins of width δ_2 . Two vectors of spike counts, \mathbf{c} and $\hat{\mathbf{c}}$, are subsequently produced, whose i^{th} elements equal the number of spikes in the i^{th} time bin for the true and estimated spike trains, respectively. The STC is the Pearson product-moment correlation coefficient of the resulting vectors, i.e.

$$\text{STC} = \frac{\langle \mathbf{c} - m(\mathbf{c}), \hat{\mathbf{c}} - m(\hat{\mathbf{c}}) \rangle}{\sqrt{v(\mathbf{c})} \sqrt{v(\hat{\mathbf{c}})}}, \quad (4.60)$$

where $\langle \cdot, \cdot \rangle$, $m(\cdot)$ and $v(\cdot)$, represent the inner product, sample mean and sample variance, respectively. To remain consistent with the success rate, in all numerical experiments, we take $\delta_2 = \delta_1 = w$.

The STC takes values in the range $[-1, 1]$. In practice, however, it is rare for a spike inference algorithm to produce an estimate that is negatively correlated with the ground truth [7]. Moreover, an estimate with maximal negative correlation is equally as informative as one with maximal positive correlation. We utilise the normalised STC, the absolute value of the STC. This ensures that the range of each metric that we analyse is equivalent (and equal to $[0,1]$) and that, as a consequence, the distribution of metric values are comparable.

4.5.3 Two-photon calcium imaging of cortical slices

This data was collected by Dr T. Abrahamsson and Dr P. J. Sjöström, our collaborators at McGill University. P11-P15 mice were anesthetized with isoflurane, decapitated, and

the brain was rapidly dissected in 4°C external solution consisting of 125 mM NaCl, 2.5 mM KCl, 1 mM MgCl₂, 1.25 mM NaH₂PO₄, 2 mM CaCl₂, 26 mM NaHCO₃, and 25 mM dextrose, bubbled with 95% O₂/5% CO₂ for oxygenation and pH. Quadruple whole-cell recordings in acute visual cortex slices were carried out at 32°C-34°C with internal solution consisting of 5 mM KCl, 115 mM K-gluconate, 10 mM K-HEPES, 4 mM MgATP, 0.3 mM NaGTP, 10 mM Na-phosphocreatine, and 0.1% w/v biocytin, adjusted with KOH to pH 7.2-7.4. On the day of the experiment, 20 μM Alexa Fluor 594 and 180 μM Fluo-5F pentapotassium salt (Invitrogen) were added to the internal solution. Electrophysiology amplifier (Dagan Corporation BVC-700A) signals were recorded with a National Instruments PCI-6229 board, using in-house software running in Igor Pro 6 (WaveMetrics). The two-photon imaging workstation was custom built as previously described [16]. Briefly, two-photon excitation was achieved by raster-scanning a Spectraphysics MaiTai BB Ti:Sa laser tuned to 820 nm across the sample using an Olympus 40x objective and galvanometric mirrors (Cambridge Technologies 6215H, 3 mm, 1 ms/line, 256 lines). Substage photomultiplier tube signals (R3896, Hamatsu) were acquired with a National Instruments PCI-6110 board using ScanImage 3.7 running in MATLAB (MathWorks). Layer-5 pyramidal cells were identified by their prominent apical dendrites using infrared video Dodt contrast.

4.6 Results

To investigate metric properties, we simulated estimated and ground truth spike trains and analysed the metric scores. To mimic the temporal error in spike time estimation, unless otherwise stated, estimates were normally distributed about the true spike times. In the following, we refer to the standard deviation of the normal distribution as the jitter of the estimates.

4.6.1 CosMIC penalises overestimation

As opposed to the STC, CosMIC and the success rate penalised overestimation of spikes (Fig. 4.7). We simulated spike train estimates that were normally distributed about the true spike times. When the number of detected spikes (\hat{K}) was less than the number of true spikes (K), the locations about which the estimates were distributed were chosen without replacement. When $\hat{K} > K$, the set of locations included all the true spikes plus a subset of extras chosen with replacement. The overestimation

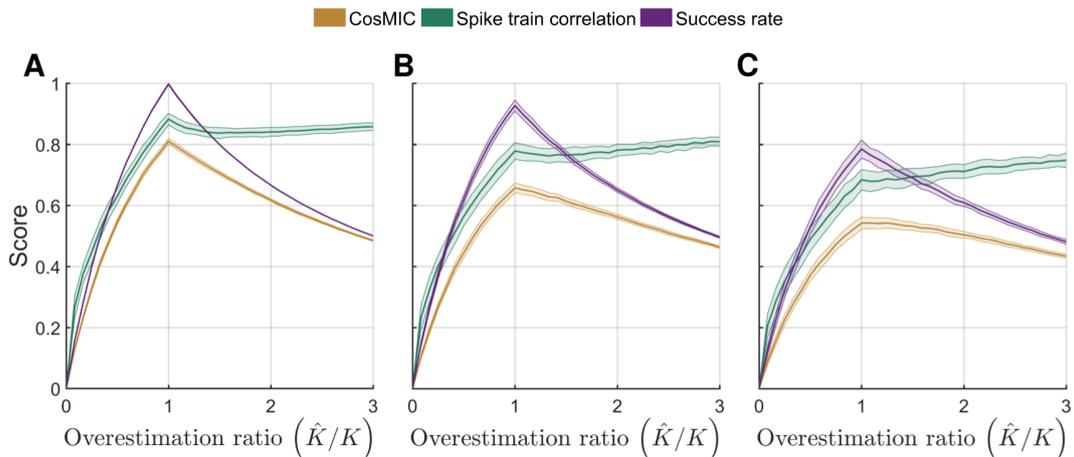


Figure 4.7: Results on spike trains with varying overestimation ratio. In contrast to the spike train correlation, CosMIC and the success rate were maximised when the correct number of spikes were detected. We display the distribution of metric scores as the number of estimated spikes (\hat{K}) varies with respect to the number of true spikes (K). The true spike train, which was identical throughout, consisted of 200 spikes simulated from a Poisson process with spike rate 1Hz. Estimated spikes were normally distributed about the true spikes, with jitter σ_{CRB} (**A**), $2 \sigma_{\text{CRB}}$ (**B**) and $3 \sigma_{\text{CRB}}$ (**C**), respectively, where $\sigma_{\text{CRB}} = 20\text{ms}$. When the number of estimated spikes was greater than the number of true spikes, estimates were distributed around a set of locations including all true spikes plus an extra subset chosen with replacement. For each metric we plot the mean (darker central line) and standard deviation (edges of shaded region) of metric scores on a set of 100 spike train estimates generated at each overestimation and jitter combination.

ratio (\hat{K}/K) reflects the degree of accuracy to which an estimate matches the rate of a ground truth spike train. We observed that, rather than penalising overestimation, the STC increased with the overestimation ratio. In contrast, CosMIC and the success rate were maximised when the correct number of spikes were detected. This behaviour was consistent as the jitter of the estimated spikes varied; in this example, the jitter was σ_{CRB} (Fig. 4.7A), $2 \sigma_{\text{CRB}}$ (Fig. 4.7B) and $3 \sigma_{\text{CRB}}$ (Fig. 4.7C), respectively.

It is the type of normalisation used by the STC that causes it to be insensitive to overestimation. Scaling factors present in the spike count vectors cancel out in the numerator and denominator, see Eq. (4.60), rendering the STC invariant under scalar transformations of the inputs. Even when the STC was adapted to the continuous-time

assessment of spike train similarity, by first convolving spike trains with a smoothing pulse, this flaw persisted [95].

4.6.2 CosMIC rewards high temporal precision

CosMIC is more sensitive to temporal precision than the STC or success rate (Fig. 4.8). First, we investigated this characteristic at the level of estimates of a single spike, t_1 . CosMIC depends only on the absolute difference between the estimate, \hat{t}_1 , and the true spike — the further the distance, the smaller the score. The relationship between CosMIC and the temporal error, $\delta = t_1 - \hat{t}_1$, is

$$M(\{t_1\}, \{\hat{t}_1\}) = \begin{cases} \left(\frac{|\delta|}{w} - 1\right)^2 & \text{if } |\delta| < w \\ 0 & \text{otherwise,} \end{cases} \quad (4.61)$$

where w is the width of the pulse. The derivation of this result is given in Appendix B. The success rate, on the other hand, does not reward increasing temporal precision above the bin width; an estimate is assigned a score of 1 or 0, when its precision is above or below the bin width, respectively. Moreover, the STC is asymmetric in the temporal error; estimates the same distance from the true spike are not guaranteed to be awarded the same score, see Fig. 4.8A. This asymmetry stems from the correlation's temporal discretisation. The temporal interval is first discretised into time bins and the number of spikes in each bin are counted (Fig. 4.6). It follows that estimated spikes that are the same absolute distance from a true spike can fall into different time bins, thus achieving a different score.

On simulated data, we investigated the effect of these properties when spike train estimates, rather than single spikes, were evaluated. In particular, we analysed the metric scores when spike train estimates contained the correct number of spikes but their temporal precision varied. We simulated the ground truth spike train as a Poisson process with rate 1Hz over 200s. The corresponding calcium transient signal was generated assuming a Cal-520 pulse shape (see Table 2.1) and a sampling rate of 30 Hz. White Gaussian noise was added to the calcium transient signal to generate two fluorescence signals, one with low and the other with relatively high noise (Fig. 4.8B). The corresponding metric pulse widths, as calculated from the CRB, were 33ms and 78ms, or 1 and 2.3 sample widths, respectively. Spike train estimates were normally distributed about the true spikes with varying jitter. The metric scores were then

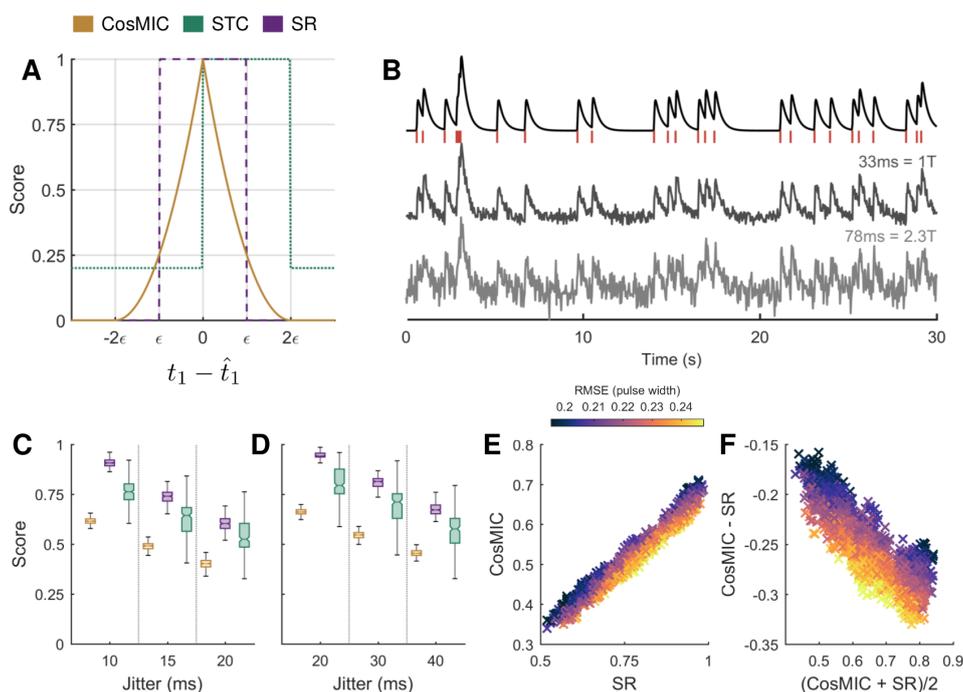


Figure 4.8: Results on spike train estimates with varying temporal precision. CosMIC was more sensitive to the temporal precision of estimates than the spike train correlation (STC) or success rate (SR). Unlike the STC, CosMIC awards estimated spikes (\hat{t}_1) with the same proximity to the true spike (t_1) the same score (A). In contrast to both the STC and SR, CosMIC rewards increasing precision above the pulse width (2ϵ) with strictly increasing scores. In (C) and (D), we plot the distribution of scores awarded to estimates that detect the correct number of spikes at varying temporal precision, in a low and high noise setting, respectively. In (B), a sample of each of the following signals are plotted: the ground truth spike train, simulated as a Poisson process at rate 1Hz over 200s; the corresponding calcium transient signal, sampled with interval $T = 1/30$ s; the low and high noise fluorescence signal and the corresponding pulse widths. At each noise and jitter level, 100 realisations of spike train estimates normally distributed about the true spike times were generated. In both the low (C) and high noise (D) settings, the STC exhibited a relatively large variation in the scores awarded to estimates of the same jitter. CosMIC and the SR were roughly linearly related (E). CosMIC was boosted with respect to the success rate when temporal error, represented by the root mean square error (RMSE) of estimates as a fraction of the pulse width, was low (F). Conversely, CosMIC was relatively low with respect to the SR when temporal error was relatively high. The colour map in (E) and (F) is thresholded at the 1st and 99th percentiles of the RMSE for visual clarity.

calculated for 100 realisations of spike train estimates at each jitter level in both the low and high noise settings (Fig. 4.8C and D, respectively).

As the correct number of spikes were always estimated, the level of jitter represented the quality of a spike train estimate in this setting. Ideally, a metric would reliably reward spike train estimates of the same quality with the same score. The STC, however, took a relatively large range of values for estimates of the same jitter (Fig. 4.8C and D), despite having the same range as CosMIC and the success rate. This inconsistency is a consequence of the edge effects introduced by binning. Here, we use the term consistency in line with its semantic rather than mathematical definition.

We observed a roughly linear trend in the scores of CosMIC and the success rate (Fig. 4.8E). As expected, CosMIC was boosted with respect to the success rate when the root mean square error (RMSE) of detected spikes was relatively low when measured as a fraction of the pulse width. In each case, the RMSE was computed empirically from the estimated spikes within the precision of the pulse width. Conversely, CosMIC was relatively low with respect to the success rate when the RMSE was relatively high. This trend is visible in the Bland-Altman plot [3], in which the mean of the two methods is plotted against the difference (Fig. 4.8F). We conclude that CosMIC is more sensitive to the temporal precision of detected spikes, as, unlike the success rate, it discriminates precision above the bin width.

4.6.3 Application to real imaging data

On imaging data of the mouse visual cortex at a frame rate of 13 Hz, CosMIC was more sensitive than the success rate to the temporal precision of detected spikes. For a detailed description of the imaging data, see Section 4.5.3. As detailed in Section 4.4, the metric's pulse width was set with respect to the CRB. On this dataset, the pulse widths were concentrated between 1 and 3 sample widths — this range encompassed 92% of the data, see (Fig. 4.9F). As the noise level of the data increases, so does the pulse width. Consequently, the tolerance of the metric with regards to the temporal precision of estimates also increases. As a result, estimates on noisier data (Fig. 4.9B) were scored with more lenience than those on less noisy data (Fig. 4.9A). Spikes were detected from this data using the version of the FRI spike inference algorithm presented in Chapter 5 of this thesis.

As was found on simulated data in Section 4.6.2, there was a linear trend between

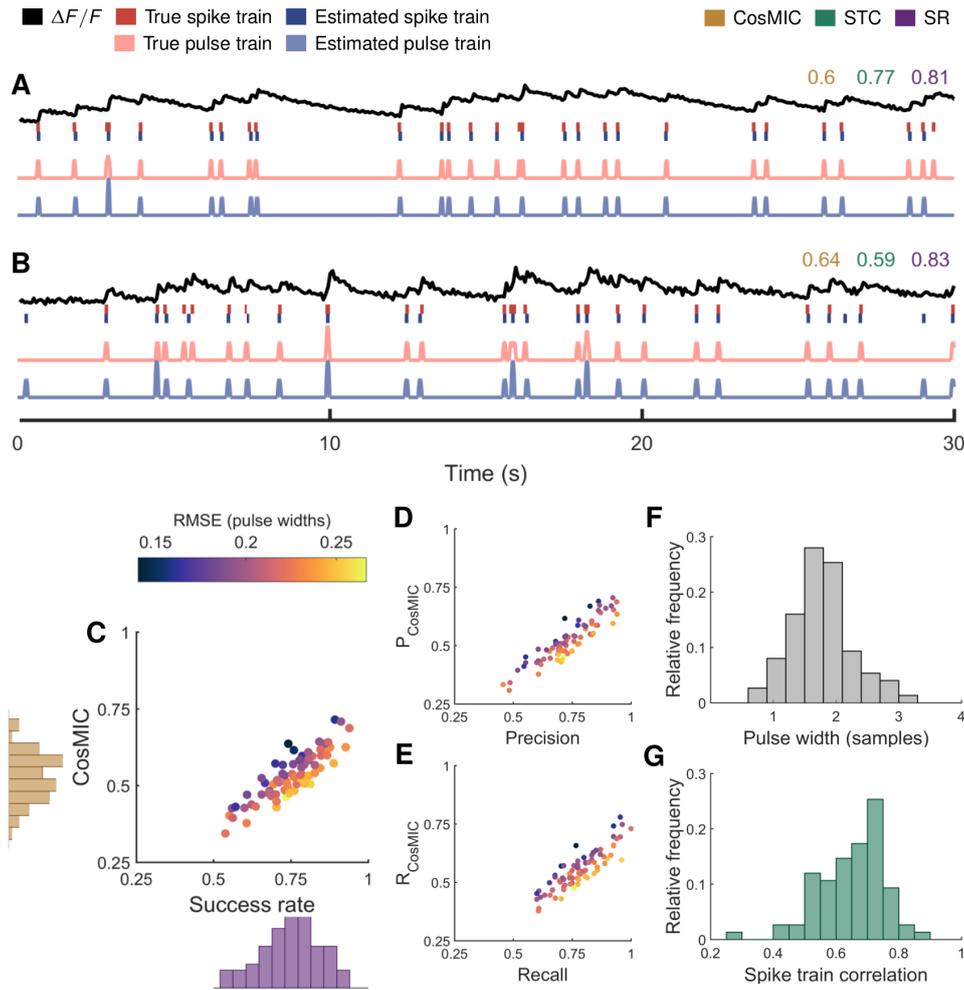


Figure 4.9: Results on real imaging data. On mouse in vitro imaging data, CosMIC was more sensitive than the success rate (SR) to the temporal precision of detected spikes. Spikes were detected using the FRI algorithm from 83 traces sampled from visual cortex slices at 13Hz. In **A** and **B**, we display from top to bottom: an example fluorescence trace ($\Delta F/F$), ground truth and detected spike trains, and the corresponding pulse trains. There was an approximately linear relationship between CosMIC and the SR (**C**). CosMIC was relatively high with respect to the SR when temporal error, represented by root mean square error (RMSE) as a fraction of the pulse width, was relatively low. Conversely, CosMIC was low with respect to the SR when temporal error was relatively high. This pattern was conserved in the relationship between the precision and CosMIC’s analogous ancestor metric, P_{CosMIC} , (**D**) and between the recall and R_{CosMIC} (**E**). The range of pulse widths as computed from the Cramér-Rao bound (**F**) and the range of spike train correlation (STC) scores (**G**) are also shown.

the scores of CosMIC and the success rate (Fig. 4.9C). CosMIC was relatively high with respect to the success rate when the temporal error represented by RMSE as a fraction of the pulse width, was relatively low. Conversely, CosMIC was low with respect to the success rate when the temporal error was relatively high. The RMSE was computed empirically from the estimated spikes within the precision of the pulse width. This pattern was conserved when CosMIC’s ancestor metrics, P_{CosMIC} and R_{CosMIC} (Section 4.3.1), were compared to the precision and recall (Fig. 4.9D and E). The average RMSE over all traces was 27ms, or 0.37 sample widths. As CosMIC is able to discriminate precision above the pulse width, it is more able to reward this super-resolution performance than the success rate or STC.

4.6.4 CosMIC discriminates precision and fall-out rate of spike trains

By construction, CosMIC bears a strong resemblance to the Sørensen-Dice coefficient and, consequently, the success rate. Both metrics can be written as the harmonic mean of the precision and recall, two intuitive metrics which represent the proportion of estimates that detect a ground truth spike and the proportion of true spikes detected, respectively. In this section, we demonstrate that CosMIC can accurately discriminate both the precision and recall of spike train estimates.

When a spike train estimate detects exactly a subset of the true spikes, plus no remainders, CosMIC and the success rate depend only on the percentage of true spikes detected (the recall) and not the location of that subset, see Fig. 4.10A and 4.10D. Denoting the size of the subset of true detections as $K - R$, with K the number of true spikes and $0 < R \leq K$, we have

$$M(\mathbb{S}, \hat{\mathbb{S}}) = 1 - \frac{1}{2K/R - 1}, \quad (4.62)$$

see Appendix B for a proof. Thus, CosMIC depends only on the proportion of ‘missing’ spikes, R/K , not their location. In contrast, the STC exhibits significant variation at each level of recall. This is illustrated in Fig. 4.10A, in which we plot the distribution of CosMIC, success rate and correlation scores over 100 realizations of spike train estimates at each level of recall. It can be seen that, in this setting, CosMIC and the success rate are fixed with the recall of the spike train estimates.

When all the true spikes were exactly detected together with $R > 0$ surplus spikes, CosMIC and the success rate depend only on the level of precision not the location of

the surplus spikes, see Fig. 4.10B and 4.10E. We have

$$M(S, \hat{S}) = \frac{1}{1 + R/2K}, \quad (4.63)$$

where K is the number of true spikes, see Appendix B for a proof. The fall-out rate, which is the complement of the precision, is the proportion of estimates that were not deemed to have detected a ground truth spike. It is apparent from Eq. (4.63) that, in this setting, CosMIC depends only on the fall-out rate, R/K . The correlation, on the other hand, varied with the location of the surplus spikes. In Fig. 4.10E, we plot the distribution of the correlation scores for 100 realizations of spike train estimates at each level of precision. CosMIC and the success rate, which were constant (and identical) at a given precision, in this scenario, are also shown.

4.7 Summary

In this chapter, we presented a new metric, CosMIC, with which to assess the similarity of spikes inferred from calcium imaging data to the ground truth spiking activity. Rather than operating on the true and estimated spike trains directly, the proposed metric assesses the similarity of the pulse trains obtained from convolution of the spike trains with a smoothing pulse. The pulse width is derived from the Cramér-Rao bound, which, in turn, is computed from the statistics of the imaging data. As a result, the metric quantifies the accuracy of the inferred spikes with respect to the limitations of the data. The final metric score is the size of the commonalities of the pulse trains as a fraction of their average size.

We demonstrated the advantage of CosMIC over the two most commonly used metrics, the STC and the success rate. Unlike the STC, which appears to reward overestimation, CosMIC is maximised when the correct number of spikes have been detected. Moreover, CosMIC is more sensitive to the temporal precision of estimated spikes than the success rate. We consider a single summary score to be practical for users who do not have the time or desire to analyse multi-dimensional trade-offs. Alternatively, CosMIC's ancestor metrics, R_{CosMIC} and P_{CosMIC} , can be used to determine the extent to which errors stem from undetected or falsely-detected spikes.

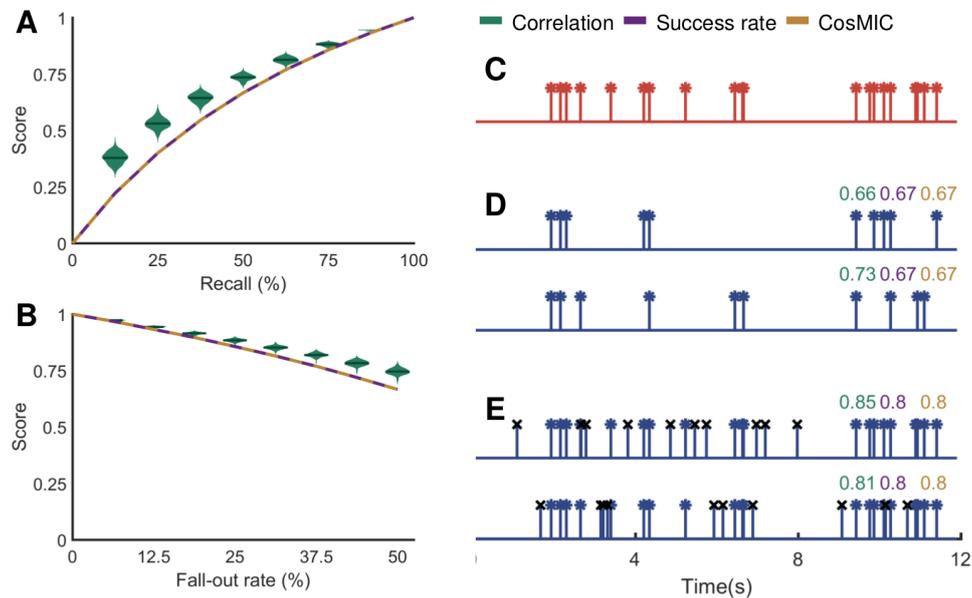


Figure 4.10: Results on spike train estimates with varying precision and fall-out rate. CosMIC scored estimated spike trains of the same recall and fall-out rate consistently, unlike the spike train correlation (STC). When a spike train estimate detected precisely the location of a subset of spikes from a true spike train, the scores of CosMIC and the success rate depended only on the percentage of spikes detected (the recall), not the location of the detected spikes (**A**, **D**). In contrast, the STC varied with the subset of spikes that were detected. When a spike train estimate detected all the true spikes precisely plus a number of surplus spikes, the STC varied with the placement of the surplus spikes (**B**, **E**). In contrast, the success rate and CosMIC depended only on the percentage of estimated spikes that did not correspond to ground truth spikes (the fall-out rate, also known as the false positive rate). The distribution of correlation scores plotted in **D** and **E** stem from 100 realizations of estimated spike trains at each recall and fall-out rate. In **C**, we plot an example of a true spike train. In **D** and **E**, we plot estimated spike trains, with a recall and fall-out rate of 50% and 33%, respectively, along with the corresponding metric scores. The spikes with a black ‘x’ marker in **E** indicate the surplus spikes.

Chapter 5.

Spike inference using FRI theory

In two-photon calcium imaging, when a neuron fires a spike, it also emits a characteristic pulse of fluorescence. The fluorescence signal from one neuron over time is thus modelled as a stream of pulses, each initiated by a spike. Neuronal network activity is analysed using spike times inferred from the imaging data. Due to the challenges inherent in imaging through living tissue, this data is typically noisy. Furthermore, sampling rates are relatively low (on the order of 10Hz) with respect to the temporal precision required for network connectivity analysis. Robust signal processing methods that are able to infer spike times from noisy data with precision above the sample width are thus required.

In previous work, it was identified that a neuron's fluorescence signal belongs to the class of signals with finite rate of innovation (FRI). Using FRI theory and a simplified model of the pulse shape, spike times were inferred with high accuracy and temporal precision. The simplified pulse model is, however, not ideal for a subset of fluorescent indicators, including some of the most promising, genetically-encoded indicators. In this chapter, we extend the FRI framework to encompass a wider class of pulse shapes. We also introduce strategies to increase the robustness of the algorithm, including a procedure to whiten the statistics of the noise and an updated approach to the estimation of the number of spikes in a window. We demonstrate on simulated data that the updated signal model eliminates estimation bias due to model mismatch. Additionally, we compare the performance of the updated FRI algorithm and a state of the art deconvolution algorithm on mouse in vitro imaging data.

The work presented in this chapter led to the following publication [109].

5.1 Introduction

The goal of spike inference is to recover a neuron’s spiking activity from its fluorescence signal (Fig. 5.1). To this end, several authors proposed supervised learning algorithms that were trained on datasets with ground truth spiking activity [114, 135, 127]. Despite promising results on familiar data, it has been shown that state of the art supervised learning approaches impose spurious auto-correlation patterns onto test data [93]. This type of error, which occurs when patterns of spiking activity in the training data are implicitly learned by an algorithm, could precipitate misleading conclusions about neuronal activity. Until the nature of errors of supervised learning algorithms are fully characterised, it is likely that model-based algorithms will be predominately used.

In contrast to supervised learning approaches, model-based algorithms take advantage of the known structure of the fluorescence signal to perform spike inference [153, 148, 47, 87, 103, 30, 106, 93]. These algorithms can be divided into two groups, according to whether their model is in discrete or continuous time. The former group model the signal as an auto-regressive process, whereby fluorescence intensity at a given sample depends on past fluorescence values and the (unknown) number of spikes that occurred in that sample [148, 147, 103, 105, 44]. With the exception of a Bayesian algorithm [103], these approaches do not infer the locations of spikes in continuous time. Rather, the output is a vector of equal length to the fluorescence signal, where each entry reflects the number of spikes fired in that sample. Vogelstein et al. argued that identification of the most likely vector of spiking activity is intractable when its entries are integer-valued [147]. Instead, they relaxed their problem to infer a vector with non-negative, real entries, whereby each entry loosely represents the probability that a spike is fired in that sample. The work of Vogelstein et al. inspired a family of non-negative deconvolution methods that infer vectors of spiking probabilities under certain sparsity constraints [105, 44]. When the SNR of traces is low and, consequently, temporal precision is not likely to surpass the sample width, non-negative deconvolution approaches can perform well. They are computationally relatively efficient and convey uncertainty in the spike time estimates. However, as temporal resolution is implicitly limited in this framework, they are not ideal when the end goal of an experiment is, for example, to analyse temporal coding hypotheses and causal network activity.

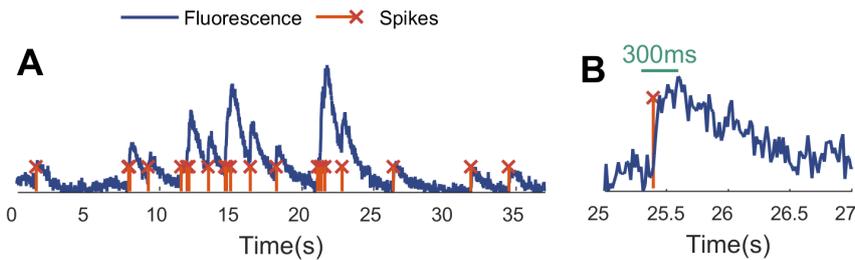


Figure 5.1: Example of a slow-rise fluorescence signal. The signal from one neuron is modelled as a stream of pulses, each initiated by a spike. This figure shows data sampled at 60Hz generated using the genetically encoded indicator GCaMP6s [132, 26]. In **B**, we plot a single calcium transient from the signal in **A**. The rise of this indicator’s pulse is relatively slow — the time from baseline to peak is approximately 300ms.

Continuous-time models used for spike inference vary in the degree of detail with which they model calcium dynamics. Rahmati et al. formulated a detailed biophysical model relating the fluorescence signal to neuronal membrane potential, considering properties of ion channels and sub-threshold voltage fluctuations [106]. This approach, while potentially cumbersome for bulk image analysis, allows inference of a cell’s biophysical properties as well as spiking activity. As such, it could be fruitful when used in combination with pharmacological interventions, in which the aim is to alter those properties [114]. Typically, continuous-time models use a simpler framework, in which the fluorescence signal is a convolution of a spike train with a characteristic pulse [47, 87, 30]. A variety of methods have been used to infer spike times in this setting, including template matching [47] and maximum likelihood estimation [30]. Of these two, the latter was found to have the highest accuracy on a wide array of real data with electrophysiological ground truth [30].

Oñativia et al. inferred the timing of spikes by exploiting the fact that the fluorescence signal belongs to the class of signals with finite rate of innovation [87]. A signal belongs to this class if it admits a parametric representation with a finite number of free parameters per unit of time [145]. In the case of neuronal fluorescence signals, the free parameters are the unknown spike times and amplitudes. By mapping the problem of spike inference to the FRI problem of reconstructing a stream of Diracs, Oñativia et al. were able to utilise well-established methods to detect spike locations. The mapping incorporated assumptions about the fluorescence pulse, which Oñativia et al. modelled with an instantaneous rise and exponential decay. This simplification

is common [147, 30] and well-suited to situations in which the pulse rise time is not significantly longer than the sampling period. However, scan rates are consistently increasing [117] and some of the most promising fluorescent indicators have relatively slow rise times [26]. For example, the genetically-encoded indicator GCaMP6s, which has a high dynamic range and brightness, rises over a period of 200-300ms (Fig. 5.1). In these cases, there is a risk of estimation bias due to model mismatch.

In this chapter, we extend the FRI framework for spike inference from calcium imaging data to encompass a slow-rise pulse shape. In the FRI framework, the sampling process of the continuous-time spike train is modelled explicitly, so we start by introducing relevant aspects of sampling theory in Section 5.2. In order to increase the algorithm's robustness to noise, we introduce a pre-whitening step [141], so that the statistics of the noise are suited to the parameter estimation techniques used. Furthermore, we include a model-based procedure to estimate the number of spikes in a window. We present these contributions in Section 5.3. Finally, we assess the performance of the updated algorithm on real and simulated data.

5.2 Sampling theory

Let $x(t)$ be a continuous-time signal, which is sampled by a generic acquisition device. The modifications that occur to the signal during the acquisition process are modelled by convolution of $x(t)$ with a filter, $h(t)$. In digital cameras, for example, this represents the blurring of the original image by a low-quality lens. The actual signal from which samples are obtained is thus

$$y(t) = x(t) * h(t) = \int_{\mathbb{R}} x(\tau)h(t - \tau) d\tau. \quad (5.1)$$

Discrete samples, which are obtained with uniform time resolution, T , are denoted

$$y[n] = y(t)|_{t=nT} = \int_{\mathbb{R}} x(\tau)h(nT - \tau) d\tau, \quad (5.2)$$

for $n \in \{0, 1, \dots, N - 1\}$. Defining $h(t) = \varphi(-\frac{t}{T})$ and using the inner product notation, the samples are expressed as

$$y[n] = \int_{\mathbb{R}} x(\tau)\varphi(\frac{\tau}{T} - n) d\tau = \langle x(t), \varphi(\frac{t}{T} - n) \rangle, \quad (5.3)$$

where we refer to φ as the sampling kernel.

Whereas in some applications, such as standard digital photography, it can be sufficient to retain the digital samples, it is typically desirable to use the samples to recover aspects of the original, continuous-time signal, $x(t)$. Indeed, under certain conditions on the signal and filter, it is possible to perfectly reconstruct the original signal from digital samples. In the mid 20th century, it was established that a bandlimited signal can be completely determined by sufficiently finely-spaced samples [119, 152, 67]. Whilst an important development, the corresponding reconstruction formula is only approximately applicable to real-world signals, which are limited in time and thus not truly bandlimited [121]. In 2002, Vetterli et al. established conditions for perfect reconstruction of a wider class of signals [145]. In the following, we introduce this framework, which will later be used to infer the locations of spikes from calcium imaging data.

5.2.1 Signals with finite rate of innovation

We consider a class of signals that can be represented as a linear combination of shifts of known functions, $\{g_r(t)\}_{r=0}^R$, such that

$$x(t) = \sum_{k \in \mathbb{Z}} \sum_{r=1}^R a_{r,k} g_r(t - t_k), \quad (5.4)$$

where t_k and $a_{r,k}$ are free parameters for $r \in \{1, \dots, R\}$ and $k \in \mathbb{Z}$. In this case, the problem of reconstructing $x(t)$ is equivalent to that of retrieving the unknown parameters. When the parameters are known, $x(t)$ is defined everywhere. The complexity of the reconstruction problem is thus reflected by the number of free parameters per unit of time. To formalise this concept, a counting function $C_x(t_a, t_b)$ is used that counts the number of free parameters in $x(t)$ over the temporal interval $[t_a, t_b]$. Then, the rate of innovation of $x(t)$ is defined as

$$\rho = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} C_x(-\frac{\tau}{2}, \frac{\tau}{2}). \quad (5.5)$$

A signal has finite rate of innovation if it admits a parametric representation as in Eq. (5.4) with a finite ρ . This class of signals, which includes bandlimited signals, encompasses a wide range of physical processes [139, 35, 84]. FRI methods exploit the knowledge of these signals' parametric structure in order to retrieve the unknown

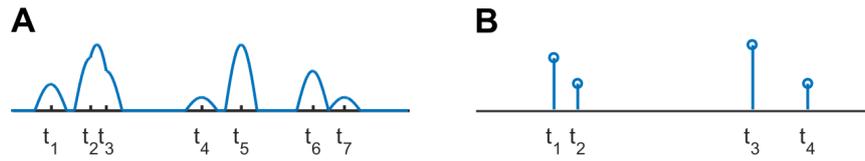


Figure 5.2: Example of signals with finite rate of innovation. Both signals are linear combinations of a function with varying shifts and amplitudes. In **A**, the constructive function is smoothly varying whereas in **B** it is a point impulse.

parameters and, hence, reconstruct the signal.

An example of an FRI signal, which is a linear combination of shifts of a single function, $g(t)$, is illustrated in Figure 5.2A. In this case, $x(t)$ consists of a finite number, K , of shifts of $g(t)$ and can be written

$$x(t) = \sum_{k=1}^K a_k g(t - t_k). \quad (5.6)$$

A further example, which is used to model neuronal spike trains, is a stream of point impulses (Fig. 5.2B). A point impulse is expressed mathematically using the Dirac Delta function, $\delta(t)$, which is zero at all locations on the real line except at $t = 0$. The Delta function is defined with respect to its effect on the integral of any continuous function, $s(t)$, such that [75]

$$\int_{\mathbb{R}} s(t) \delta(t) dt = s(0). \quad (5.7)$$

A stream of Diracs at locations $\{t_k\}_{k=1}^K$ with amplitudes $\{a_k\}_{k=1}^K$ is written

$$x(t) = \sum_{k=1}^K a_k \delta(t - t_k). \quad (5.8)$$

Due to its connection with our neuroscience problem, in the following, we focus on how FRI theory can be used to sample and reconstruct streams of Diracs.

5.2.2 Sampling kernels

The reconstruction scheme is tailored to the sampling kernel used in the acquisition process. Among others, reconstruction schemes have been proposed for sinc and Gaussian kernels [145], kernels with rational Fourier transform [38] and polynomial and

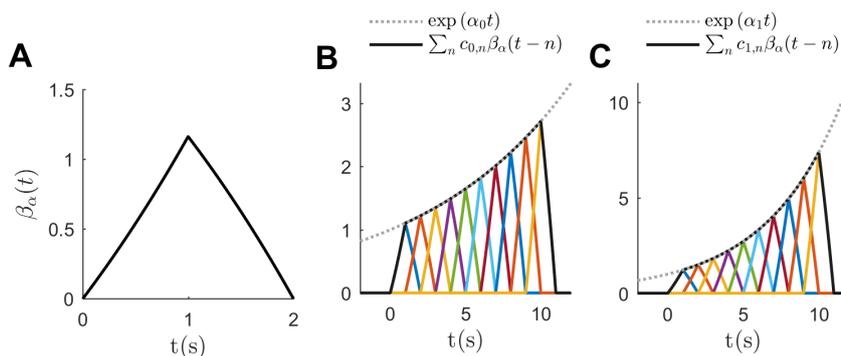


Figure 5.3: Exponential reproduction with an E-spline of order 2. In **A**, we plot the E-spline in the time domain. In **B** and **C**, the exponentials $e^{\alpha_0 t}$ and $e^{\alpha_1 t}$ are reproduced, respectively, over the interval $[1, 10]$.

exponential reproducing kernels [38]. Here, we focus on the latter class of kernels.

An exponential reproducing kernel is a function that, when linearly combined with its uniformly shifted versions, can reproduce exponentials. More formally, it is a function, $\varphi(t)$, for which there exists a set of coefficients, $c_{m,n} \in \mathbb{C}$, and exponents, $\alpha_m \in \mathbb{C}$, such that

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t} \quad \text{for } m \in \{0, \dots, P\}, \quad (5.9)$$

almost everywhere. A function reproduces exponentials if and only if it satisfies the generalised Strang-Fix conditions [149]:

$$\mathcal{L}\{\varphi\}(\alpha_m) \neq 0 \quad \text{and} \quad \mathcal{L}\{\varphi\}(\alpha_m + 2\pi il) = 0, \quad (5.10)$$

for $m \in \{0, \dots, P\}$ and $l \in \mathbb{Z} \setminus 0$, where $\mathcal{L}\{\varphi\}$ represents the double-sided Laplace transform of $\varphi(t)$.

A family of functions that satisfy the exponential reproduction property are the exponential B-splines [140], otherwise known as E-splines, whose name refers to their relation to polynomial-reproducing B-splines. A zero order E-spline is defined as follows:

$$\beta_\alpha(t) = \begin{cases} e^{\alpha t} & 0 \leq t \leq 1, \\ 0 & \text{otherwise,} \end{cases} \quad \xrightarrow{\mathcal{F}} \quad \hat{\beta}_\alpha(\omega) = \frac{1 - e^{\alpha - i\omega}}{i\omega - \alpha}. \quad (5.11)$$

In the case of the zero-order E-spline, the parameter α indicates the exponent of the

exponential that can be reproduced. More generally, an E-spline of order $P \geq 0$ is associated with a vector of exponents $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \dots, \alpha_P\}$ and is obtained through convolution of zero order E-splines, such that

$$\beta_{\boldsymbol{\alpha}}(t) = \beta_{\alpha_0}(t) * \beta_{\alpha_1}(t) * \dots * \beta_{\alpha_P}(t) \xrightarrow{\mathcal{F}} \hat{\beta}_{\boldsymbol{\alpha}}(\omega) = \prod_{m=0}^P \frac{1 - e^{\alpha_m - i\omega}}{i\omega - \alpha_m}. \quad (5.12)$$

This E-spline can reproduce functions in the subspace spanned by $\{e^{\alpha_0 t}, e^{\alpha_1 t}, \dots, e^{\alpha_P t}\}$, as long as no two exponents are separated by a multiple of $2\pi i$, which would violate the Strang-Fix conditions. In Fig. 5.3, we illustrate the exponential reproduction property for an E-spline of order two. To reproduce the exponentials for all $t \in \mathbb{R}$ would require the sum in Eq. (5.9) to be evaluated over all $n \in \mathbb{Z}$. In this example, we truncate the sum and reproduce the exponentials over the interval $[1, 10]$.

5.2.3 Reconstruction of a stream of Diracs

We now consider the situation in which a stream of Diracs, $x(t)$, has been sampled with an exponential reproducing kernel, $h(t) = \varphi(-\frac{t}{T})$, at time resolution T to produce samples $\{y[n]\}_{n=0}^{N-1}$. To retrieve the unknown spike times and amplitudes, the samples are linearly combined using the exponential reproduction coefficients of the sampling kernel, $c_{m,n}$, such that

$$s[m] = \sum_{n \in \mathbb{Z}} c_{m,n} y[n] = \sum_{n \in \mathbb{Z}} c_{m,n} \langle x(t), \varphi(\frac{t}{T} - n) \rangle = \langle x(t), \sum_{n \in \mathbb{Z}} c_{m,n} \varphi(\frac{t}{T} - n) \rangle, \quad (5.13)$$

for $m \in \{0, 1, \dots, P\}$, where the last step follows from the linearity of the inner product. We refer to $s[m]$ as the sample moments. Due to the exponential reproduction property, see Eq. (5.9), it follows that

$$s[m] = \langle x(t), e^{\alpha_m t/T} \rangle = \int_{\mathbb{R}} x(t) e^{\alpha_m t/T} dt = \sum_{k=1}^K a_k e^{\alpha_m t_k/T}, \quad (5.14)$$

where the last step follows from Eq. (5.7). When the exponents are evenly spaced, that is $\alpha_m = \alpha_0 + \lambda m$, the sample moments can be written in power-sum form, such that

$$s[m] = \sum_{k=1}^K b_k u_k^m, \quad (5.15)$$

with $b_k = a_k e^{\alpha_0 t_k/T}$ and $u_k = e^{\lambda t_k/T}$.

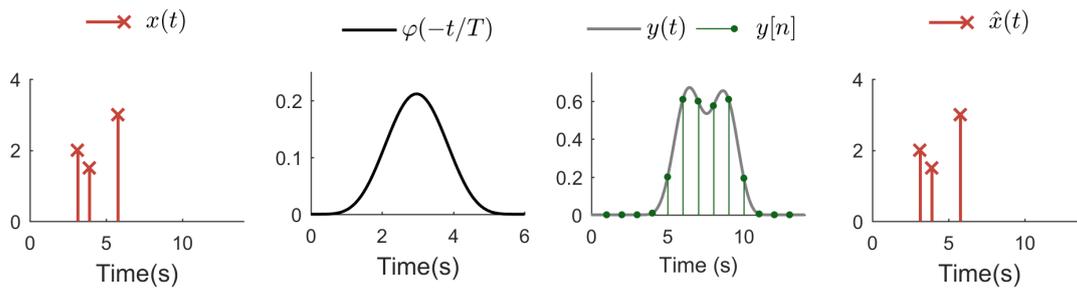


Figure 5.4: Example of sampling and reconstruction of a stream of Diracs.

A stream of 3 Diracs, $x(t)$, is filtered with an exponential reproducing kernel, $h(t) = \varphi(-t/T)$, to produce $y(t) = x(t) * h(t)$. Samples $y[n]$ are obtained at time resolution T . The spike times and amplitudes are recovered from the samples $y[n]$.

The formation of the sample moments can also be written in matrix vector form, such that $\mathbf{s} = \mathbf{C}\mathbf{y}$, where \mathbf{y} and \mathbf{s} represent the vectors of samples and sample moments, respectively, and the $(i, j)^{\text{th}}$ entry of \mathbf{C} is $c_{i,j}$. We select α_m to ensure that the coefficient matrix \mathbf{C} is well-conditioned. Urigüen et al. demonstrated that E-splines with

$$\alpha_m = \frac{\pi}{P+1} (2m - P) i \quad \text{and} \quad c_{m,n} = e^{\alpha_m n} (\hat{\varphi}(\alpha_m))^{-1}, \quad (5.16)$$

enabled relatively stable reconstruction of streams of Diracs [141]. This is the formulation we use in the rest of this chapter. In Fig. 5.4, we display an example of such an E-spline. We note that the kernel itself is real, as the exponents occur in complex conjugate pairs.

In order to retrieve the spike times and amplitudes, one needs to recover $\{b_k\}_{k=1}^K$ and $\{u_k\}_{k=1}^K$ from the sample moments. This is a common problem in spectral estimation theory, for which there are many methods [129]. For example, Prony's method can be used to identify the filter $h[m]$ that annihilates $s[m]$, such that

$$h[m] * s[m] = \sum_{l=0}^K h[l] s[m-l] = 0. \quad (5.17)$$

One can show that the filter $h[m]$ that satisfies Eq. (5.17) has roots that correspond to the parameters u_k [38]. Thus, the parameters u_k are obtained as the roots of the annihilating filter. Using Prony's method it is possible to perfectly retrieve u_k , and thus t_k , using the minimum number of measurements, $P + 1 = 2K$, see Fig. 5.4.

In practice, the samples and, consequently, the sample moments are corrupted by

noise. In this case, it is preferable to use $P \geq 2K$ moments in order to increase the robustness of the parameter estimation procedure. One can also use Prony's method with denoising strategies, such as Cadzow denoising [19]. In this thesis, we use the robustified matrix pencil method [53, 54], which is as effective as Cadzow denoising with Prony's method but has the advantage that it is not iterative.

Matrix pencil method

The matrix pencil method identifies u_k using the Toeplitz matrix of sample moments:

$$\mathbf{S} = \begin{pmatrix} s[M] & s[M-1] & \dots & s[0] \\ s[M+1] & s[M] & \dots & s[1] \\ \vdots & \vdots & \ddots & \vdots \\ s[P] & s[P-1] & \dots & s[P-M] \end{pmatrix}, \quad (5.18)$$

where $M = \lfloor P/2 \rfloor$. First, the singular value decomposition of this matrix is computed, such that $\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Here, $\mathbf{\Sigma}$ is diagonal and \mathbf{U} and \mathbf{V} contain the left and right singular vectors of \mathbf{S} , respectively. A truncated matrix \mathbf{U}_K is subsequently formed from the left singular vectors corresponding to the K largest singular values. When the noise level is sufficiently low, these singular vectors form a basis of the signal subspace. Defining \mathbf{S}_0 and \mathbf{S}_1 as the matrix \mathbf{U}_K with the first and last row removed, respectively, the parameters $\{u_k\}_{k=1}^K$ are identified as the eigenvalues that solve the following generalised eigenvalue problem

$$\left(\mathbf{S}_1^\dagger \mathbf{S}_0 - \mu \mathbf{I} \right) v = 0, \quad (5.19)$$

where \mathbf{S}_1^\dagger denotes the Moore-Penrose psuedoinverse of \mathbf{S}_1 .

5.3 Spike inference from calcium imaging data

In the following, we outline how the problem of spike inference from calcium imaging data is mapped to the FRI problem of reconstructing a stream of Diracs. We start by re-introducing the mathematical model of a neuronal fluorescence signal. For a more detailed discussion on the origins of the signal model, we refer the reader to Section 2.3.

When a spike is fired from a neuron, a characteristic pulse of fluorescence is emitted, such that

$$p(t) = c \left(1 - e^{-t/\tau_{\text{on}}} \right) e^{-t/\tau_{\text{off}}} \mathbb{1}_{t>0}, \quad (5.20)$$

where c is a normalisation constant that ensures that $\max\{p(t)\} = 1$ and $\tau_{\text{on}}, \tau_{\text{off}}$ are parameters defining the speed of the rise and decay, respectively. The parameters τ_{on} and τ_{off} are largely determined by the calcium indicator that was used to generate the data (see Section 2.3) and are straightforward to estimate [148, 103, 105, 30]. In the following, therefore, we consider these to be known parameters. A neuron's fluorescence signal is modelled as a stream of pulses at spike times $\{t_k\}_{k=1}^K$ with amplitudes $\{a_k\}_{k=1}^K$, such that

$$f(t) = \sum_{k=1}^K a_k p(t - t_k), \quad (5.21)$$

where we assume the baseline fluorescence has been subtracted. There are many procedures for removing the baseline component, typically a parametric model is fit to the signal and subsequently subtracted [30, 106, 44, 104].

For the moment, we consider the number of spikes (the model order) to be known. We discuss the estimation of this value in Section 5.4.2. The fluorescence signal, therefore, has $2K$ unknown parameters — the spike times and amplitudes — and also belongs to the class of FRI signals. Denoting $x(t) = \sum_{k=1}^K a_k \delta(t - t_k)$, we have

$$f(t) = \sum_{k=1}^K a_k p(t - t_k) = p(t) * \sum_{k=1}^K a_k \delta(t - t_k) = p(t) * x(t). \quad (5.22)$$

We assume that the fluorescence signal has been filtered with an exponential reproducing kernel, $h(t) = \varphi(-\frac{t}{T})$ and sampled with time resolution T , such that we have samples

$$y[n] = \langle f(t), \varphi(\frac{t}{T} - n) \rangle = \langle x(t) * p(t), \varphi(\frac{t}{T} - n) \rangle, \quad (5.23)$$

for $n \in \{0, \dots, N\}$. Then, expanding the integrals,

$$y[n] = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} x(\tau) p(t - \tau) d\tau \right] \varphi\left(\frac{t}{T} - n\right) dt \quad (5.24)$$

$$= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \varphi\left(\frac{t}{T} - n\right) p(t - \tau) dt \right] x(\tau) d\tau \quad (5.25)$$

$$= \int_{\mathbb{R}} x(\tau) (p(-\tau) * \varphi\left(\frac{\tau}{T} - n\right)) d\tau \quad (5.26)$$

$$= \langle x(t), p(-t) * \varphi\left(\frac{t}{T} - n\right) \rangle, \quad (5.27)$$

where, in Eq. (5.25), we are able to exchange the order of integration as $y[n]$ is finite (the integrand is bounded and has compact support).

Practical implementation

So far, we have assumed that the fluorescence signal was sampled with an exponential reproducing kernel. While not impossible, two-photon microscopes do not currently implement such a sampling kernel. In practice, we have access to $f[n] = f(nT)|_{t=nT}$. We compute the samples $y[n]$ through discrete convolution of $f[n]$ and discretised sampling kernel, such that

$$y[n] = \sum_{m \in \mathbb{Z}} f[m] h[n - m]. \quad (5.28)$$

To enable the theoretical analysis, in the following we use the continuous-time formulation. The discrete implementation of the sampling kernel has been shown to be effective in a range of setups [88], and has been tested on real data from a range of indicators with sampling rate ranging from 8Hz [87] to 60Hz [109].

5.3.1 Mapping samples to the FRI setting

In the following proposition, we establish the filtering operations that transform samples of slow-rise calcium transients into samples of the underlying spike train filtered with an exponential reproducing kernel. We define a discrete filter, $g[n]$, such that, when convolved with $y[n]$, we obtain

$$w[n] = y[n] * g[n] = \langle x(t), \psi\left(\frac{t}{T} - n\right) \rangle. \quad (5.29)$$

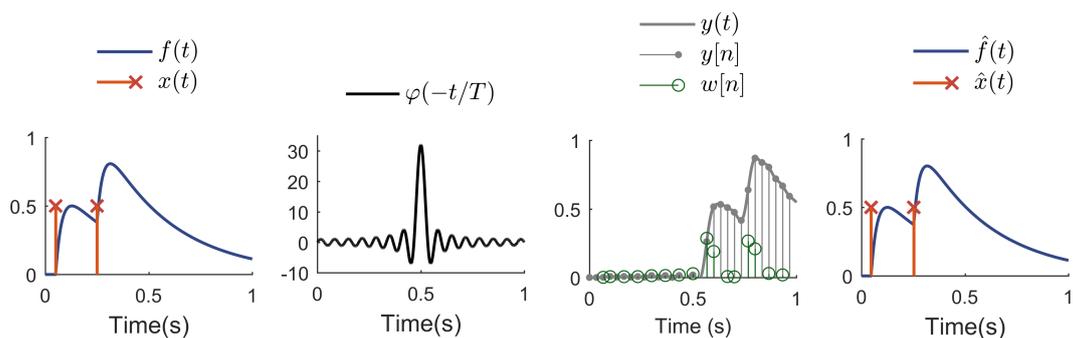


Figure 5.5: Example of spike inference with the FRI algorithm. A fluorescence signal, $f(t)$, is sampled with kernel $\varphi(-t/T)$ to produce samples $y[n]$. The samples are subsequently convolved with a discrete-time filter to produce $w[n]$, from which the unknown spike times can be retrieved.

The analogous filter was identified under the assumption of pulses with an instantaneous rise and exponential decay in [87]. In that case, $g[n]$ had two non-zero coefficients whose value depended on the decay parameter of the pulse. The resulting kernel, $\psi(t)$, was a convolution of the original sampling kernel, $\varphi(t)$, and a zero order E-spline. In our setting, assuming the samples are as in Eq. (5.27), $g[n]$ has three non-zero coefficients and

$$\psi(t) = \beta \alpha_{\tau} \left(-\frac{t}{T}\right) * \varphi(t), \quad (5.30)$$

where $\alpha_{\tau} = \{-T/\tau_{\text{off}}, -T/\tau_{\text{off}} - T/\tau_{\text{on}}\}$. The coefficients of both the filter and the E-spline depend on the pulse parameters, τ_{on} and τ_{off} . As the exponential reproduction property is conserved through convolution [140], $\psi(t)$ is also an exponential reproducing kernel and, in particular, reproduces the exponentials that are reproduced by $\varphi(t)$. Although the exponents are conserved, the reproduction coefficients are, in general, different. We compute the sample moments using the reproduction coefficients of $\psi(t)$, such that

$$s[m] = \sum_{n \in \mathbb{Z}} d_{m,n} w[n]. \quad (5.31)$$

Then, the methods of Section 5.2.3 can be used to reconstruct the unknown spike times and amplitudes (see Fig. 5.5).

Proposition 5.1. *Samples $y[n]$ are convolved with filter $g[n] = \frac{\tau_{\text{on}}}{T} g_1[n] * g_2[n]$, where*

$$g_1[n] = \delta[n] - \exp\left(-\frac{T}{\tau_{\text{off}}}\right) \delta[n-1], \quad (5.32)$$

$$g_2[n] = \delta[n] - \exp\left(-\frac{T}{\tau_{\text{off}}} - \frac{T}{\tau_{\text{on}}}\right) \delta[n-1], \quad (5.33)$$

and $\delta[n]$ indicates the discrete delta function. This produces

$$w[n] = y[n] * g[n] = \langle x(t), \beta_{\alpha_\tau}(-t/T) * \varphi(\frac{t}{T} - n) \rangle, \quad (5.34)$$

with $\alpha_\tau = \{-T/\tau_{\text{off}}, -T/\tau_{\text{off}} - T/\tau_{\text{on}}\}$.

Proof. From the definition of $w[n]$, we have

$$w[n] = \sum_{m \in \mathbb{Z}} g[m]y[n - m]. \quad (5.35)$$

Substituting the value of $y[n]$ from Eq. (5.27) and using the linearity of the inner product and convolution, $w[n]$ becomes

$$w[n] = \left\langle x(t), p(-t) * \sum_{m \in \mathbb{Z}} g[m] \varphi\left(\frac{t}{T} - (n - m)\right) \right\rangle. \quad (5.36)$$

Representing the filter coefficients in continuous time with $g(t) = \sum_{m \in \mathbb{Z}} g[m] \delta(t - mT)$, we have

$$g(-t) * \varphi\left(\frac{t}{T} - n\right) = \int_{\mathbb{R}} g(-\tau) \varphi\left(\frac{t-\tau}{T} - n\right) dt \quad (5.37)$$

$$= \int_{\mathbb{R}} \sum_{m \in \mathbb{Z}} g[m] \delta(-\tau - mT) \varphi\left(\frac{t-\tau}{T} - n\right) dt \quad (5.38)$$

$$= \sum_{m \in \mathbb{Z}} g[m] \varphi\left(\frac{t}{T} - (n - m)\right). \quad (5.39)$$

It follows that

$$w[n] = \langle x(t), p(-t) * g(-t) * \varphi\left(\frac{t}{T} - n\right) \rangle. \quad (5.40)$$

Decomposing the continuous-time Fourier transform of $g(-t)$ into the discrete-time Fourier transforms of filters $g_1[n]$ and $g_2[n]$, we have

$$\mathcal{F}\{g(-t)\}(\omega) = \sum_{m \in \mathbb{Z}} g[m] e^{i\omega m T} = \frac{\tau_{\text{on}}}{T} G_1(e^{-i\omega T}) G_2(e^{-i\omega T}). \quad (5.41)$$

Writing $c_1 = -\frac{1}{\tau_{\text{off}}}$ and $c_2 = -\frac{1}{\tau_{\text{off}}} - \frac{1}{\tau_{\text{on}}}$, this becomes

$$\mathcal{F}\{g(-t)\} = \frac{1}{T(c_1 - c_2)} (1 - e^{c_1 T} e^{i\omega T}) (1 - e^{c_2 T} e^{i\omega T}) \quad (5.42)$$

$$= \frac{1}{T(c_1 - c_2)} (1 - e^{c_1 T + i\omega T}) (1 - e^{c_2 T + i\omega T}). \quad (5.43)$$

Noting that

$$\mathcal{F}\{p(-t)\}(\omega) = \frac{c_1 - c_2}{(c_1 + i\omega)(c_2 + i\omega)}, \quad (5.44)$$

and denoting $\alpha_\tau = \{\alpha_1, \alpha_2\} = \{c_1 T, c_2 T\}$, we have

$$\mathcal{F}\{p(-t) * g(t)\} = \mathcal{F}\{p(-t)\} \mathcal{F}\{g(t)\} \quad (5.45)$$

$$= \frac{1}{T} \prod_{m=1}^2 \frac{1 - e^{c_m T + i\omega T}}{c_m + i\omega} \quad (5.46)$$

$$= T \prod_{m=1}^2 \frac{1 - e^{\alpha_m + i\omega T}}{-\alpha_m - i\omega T} \quad (5.47)$$

$$= \mathcal{F}\{\beta_{\alpha_\tau}(-t/T)\}. \quad (5.48)$$

Finally, it follows that

$$w[n] = \langle x(t), \beta_{\alpha_\tau}(-\frac{t}{T}) * \varphi(\frac{t}{T} - n) \rangle. \quad (5.49)$$

□

5.3.2 Amplitude estimation

The fluorescence signal has $2K$ unknown parameters: K spikes times and K amplitudes. Accurate computation of the amplitudes, which are of limited practical interest, is essential for the procedure that we use to estimate the model order (Section 5.4.2). To estimate the amplitudes, we use the following matrix-vector formulation of the fluorescence signal model:

$$\begin{pmatrix} f[0] \\ f[1] \\ \vdots \\ f[N-1] \end{pmatrix} = \begin{pmatrix} p_1[0] & p_2[0] & \cdots & p_K[0] \\ p_1[1] & p_2[1] & \cdots & p_K[1] \\ \vdots & \vdots & \ddots & \vdots \\ p_1[N-1] & p_2[N-1] & \cdots & p_K[N-1] \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{pmatrix}, \quad (5.50)$$

where $p_j[n] = p(t - t_j)|_{t=nT}$. Writing Eq. (5.50) more compactly as $\mathbf{f} = \mathbf{P}\mathbf{a}$, we estimate the amplitude vector as

$$\hat{\mathbf{a}} = \min_{\mathbf{a}} \|\mathbf{P}\mathbf{a} - \mathbf{f}\|^2 \quad \text{where} \quad \frac{1}{2}A \leq \hat{\mathbf{a}} \leq \frac{3}{2}A. \quad (5.51)$$

Here, we constrain the entries of $\hat{\mathbf{a}}$ to be within a certain range of physiologically plausible amplitudes, where A is the expected amplitude of a calcium transient.

5.4 Spike inference from noisy data

In practice, sampling a signal with any acquisition device generates measurement noise. We model the samples as corrupted by additive noise, such that

$$\tilde{y}[n] = y[n] + \epsilon_y[n], \quad (5.52)$$

where $\epsilon_y[n]$ are i.i.d. noise samples with distribution $\mathcal{N}(0, \sigma^2)$. As discussed in Section 2.4.4, noise generated in fluorescence microscopy is predominantly photon shot noise, which is typically modelled in this way [147, 105, 106, 44, 127]. As noted earlier, in practice, we obtain $\tilde{y}[n]$ from the raw signal samples by applying the convolution with an exponential reproducing kernel in post-processing. The model of the noise in Eq. (5.52), therefore, only approximately holds.

The filtering operations that are used to map the 2PCI spike inference problem to that of reconstructing a stream of Diracs cause the noise to become correlated. We have

$$\tilde{w}[n] = \sum_{m \in \mathbb{Z}} g[m] \tilde{y}[n - m] \quad (5.53)$$

$$= \sum_{m \in \mathbb{Z}} g[m] y[n - m] + \sum_{m \in \mathbb{Z}} g[m] \epsilon_y[n - m] \quad (5.54)$$

$$= w[n] + \epsilon_w[n], \quad (5.55)$$

where the noise samples, $\epsilon_w[n]$, are still Gaussian but are no longer independent. When we compute the sample moments, the noise distribution is further altered, such that

$$\tilde{s}[m] = \sum_{n \in \mathbb{Z}} d_{m,n} \tilde{w}[n] = \sum_{n \in \mathbb{Z}} d_{m,n} w[n] + \sum_{n \in \mathbb{Z}} d_{m,n} \epsilon_w[n] = s_m + \epsilon_s[m], \quad (5.56)$$

where

$$\epsilon_s[m] = \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_{m,n} g[k] \epsilon_y[n - k]. \quad (5.57)$$

Finally, the Toeplitz matrix of noisy sample moments, which is analogous to Eq. (5.18), is denoted $\tilde{\mathbf{S}} = \mathbf{S} + \mathbf{B}$, where \mathbf{B} contains the contributions from the noise.

When the data is not corrupted by noise, the sample moment matrix has rank K . Determining the number of innovations is therefore straightforward. In the noisy setting, $\tilde{\mathbf{S}}$ is no longer rank-deficient and the estimation of the number of innovations is more challenging. In Section 5.4.2, we outline the method we use to estimate this parameter. Additionally, the performance of the methods that infer spike times from the noisy sample moment matrix are negatively affected by correlated noise. In Section 5.4.1, we discuss a method to whiten the statistics of the noise.

5.4.1 Pre-whitening

The matrix pencil method, which we use to estimate the spike times from the sample moment matrix, aims to separate the signal subspace from the noise subspace. To do so, it assumes that the singular vectors corresponding to the largest K singular values characterise the signal subspace. When the noise is correlated, however, it is likely that some of the dominant singular vectors characterise trends in the noise rather than the signal. In this case, some singular vectors corresponding to the signal subspace will be discarded and the unknown parameters will not all be retrieved correctly.

To mitigate this problem, a pre-whitening transform can be applied to $\tilde{\mathbf{S}}$ [141], such that

$$\tilde{\mathbf{S}}' = \tilde{\mathbf{S}} \mathbf{W} = (\mathbf{S} + \mathbf{B}) \mathbf{W} = \mathbf{S} \mathbf{W} + \mathbf{B} \mathbf{W}, \quad (5.58)$$

where \mathbf{W} is designed to ensure that $\mathbf{B} \mathbf{W}$ contains white noise. Specifically, we require that

$$\mathbb{E} \left[(\mathbf{B} \mathbf{W})^H (\mathbf{B} \mathbf{W}) \right] = \mathbf{I}, \quad (5.59)$$

where \mathbf{I} is the identity matrix. This whitening condition is satisfied by $\mathbf{W} = \mathbf{R}_{\mathbf{B}}^{\dagger/2}$, where $\mathbf{R}_{\mathbf{B}} = \mathbb{E} [\mathbf{B}^H \mathbf{B}]$ [41, 142] and the superscript $\dagger/2$ denotes the square root of the pseudo inverse. When \mathbf{W} is full-rank, the rank of $\mathbf{S} \mathbf{W}$ is equal to K . It can be shown that the same parameters that satisfy the generalised eigenvalue problem in Eq. (5.19)

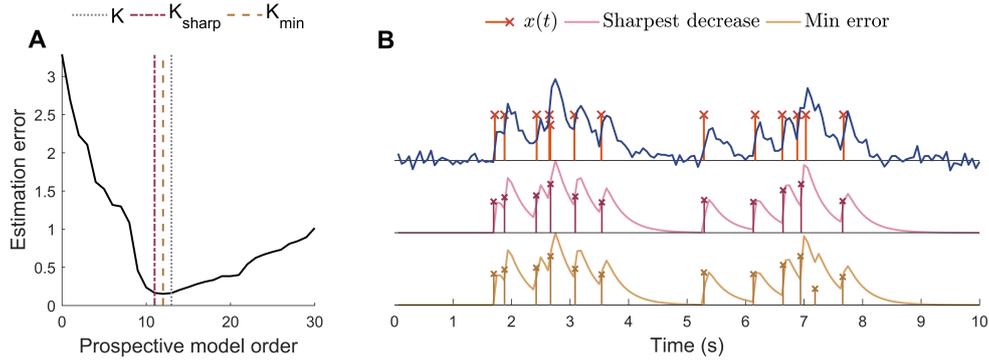


Figure 5.6: Model order estimation procedure. For a range of prospective model orders, the unknown spike times and amplitudes are estimated and the estimation error is computed between the data samples and the resynthesised fluorescence signal. The selected model order, K_{sharp} , is the one that causes the sharpest decrease in estimation error and is in the vicinity of K_{min} , the order that achieves minimum error. In **A**, these values are indicated on the estimation error curve along with the true model order, K . In **B**, we plot the raw data and the resynthesised fluorescence signals for the estimates with K_{min} and K_{sharp} spikes.

for **S** also do so for **SW** [142]. The pre-whitened matrix, $\tilde{\mathbf{S}}'$, can therefore be used to retrieve the unknown parameters.

To calculate **W**, we note that $\boldsymbol{\epsilon}_s = \mathbf{D}\mathbf{G}\boldsymbol{\epsilon}_y$, with

$$\boldsymbol{\epsilon}_s = \begin{pmatrix} d_{0,0} & d_{0,1} & \dots & d_{0,N-3} \\ d_{1,0} & d_{1,1} & \dots & d_{1,N-3} \\ \vdots & \vdots & \ddots & \vdots \\ d_{P,0} & d_{P,1} & \dots & d_{P,N-3} \end{pmatrix} \begin{pmatrix} g[0] & g[1] & g[2] & \dots & 0 \\ 0 & g[0] & g[1] & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & g[2] \end{pmatrix} \begin{pmatrix} \epsilon_y[0] \\ \epsilon_y[1] \\ \dots \\ \epsilon_y[N-1] \end{pmatrix}. \quad (5.60)$$

Writing $\mathbf{A} = \mathbf{D}\mathbf{G}$, the $(k, l)^{\text{th}}$ entry of \mathbf{R}_B is [142]

$$[\mathbf{R}_B]_{k,l} = \sigma^2 \langle \mathbf{A}_{[M-k:P-k,:]}, \mathbf{A}_{[M-l:P-l,:]} \rangle_F, \quad (5.61)$$

where $\mathbf{A}_{[a:b,:]}$ denotes the submatrix of \mathbf{A} consisting of rows a to b and $\langle \cdot, \cdot \rangle_F$ denotes the Frobenius inner product.

5.4.2 Model order estimation

To estimate the model order, K , we adapt the method of Doğan et al. [36] to our setting. For a range of prospective model orders, $J \in \{1, \dots, J_{\max}\}$, we estimate the unknown parameters of the spike train, $\{\hat{t}_k, \hat{a}_k\}_{k=1}^J$. We set $J_{\max} = \min(\frac{3}{2} \lambda NT, \frac{P}{2})$, where $\lambda(\text{Hz})$ is the expected neuronal firing rate. We then compute the estimation error between the fluorescence signal samples, $f[n]$, and the samples resynthesised from the parameter estimates, $\hat{f}_J[n] = \hat{f}_J(t)|_{t=nT}$ where

$$f_J(t) = \sum_{k=1}^J \hat{a}_k p(t - \hat{t}_k). \quad (5.62)$$

The estimation error is computed as

$$E_T(J) = \|\hat{f}_J[n] - f[n]\|^2. \quad (5.63)$$

The estimation error exhibits a U-curve shape, which indicates the transition from underfitting (too few spikes) to overfitting (too many), see Fig. 5.6A. The optimal fitting region lies between these two regions. The fitting level is subsequently defined as the minimum estimation error,

$$\hat{\sigma}^2 = \min_J E_T(J). \quad (5.64)$$

Prospective estimates of the model order are narrowed down to those in the vicinity of the model order that minimises the estimation error, such that

$$N(\hat{\sigma}^2) = \{J : \frac{1}{2}\hat{\sigma}^2 \leq E_T(J) \leq \frac{3}{2}\hat{\sigma}^2\}. \quad (5.65)$$

Ideally, this region encompasses those in between the underfitting and overfitting regions. Finally, the estimate of the model order, \hat{K} , is the order in $N(\hat{\sigma}^2)$ that causes the largest drop in the estimation error. This is a more conservative estimate since the model order that yields the minimum estimation error tends to overfit the data, adding spikes to compensate for noise (Fig. 5.6B).

Whereas Doğan et al. compute the estimation error from the sample moments, we use the raw signal samples. This is because we prefer to perform the model fitting in the native domain of the spike inference problem, rather than in the FRI domain. We note that the amplitude fitting procedure (Section 5.3.2) includes an upper and lower

Algorithm 1: Procedure by which spike times and amplitudes are estimated from a noisy fluorescence signal.

Input: $\tilde{y}[n]$ for $n \in \{0, 1, \dots, N - 1\}$

Output: \hat{t}_k and \hat{a}_k for $k \in \{1, 2, \dots, \hat{K}\}$

- 1 Apply discrete filter to samples: $\tilde{w}[n] = \tilde{y}[n] * g[n]$.
 - 2 Compute sample moments: $\tilde{s}[m] = \sum_n d_{m,n} \tilde{w}[n]$.
 - 3 Compute matrix $\tilde{\mathbf{S}}$ from sample moments.
 - 4 Perform pre-whitening: $\tilde{\mathbf{S}}' = \tilde{\mathbf{S}}\mathbf{W}$.
 - 5 **for** $J \in \{1, \dots, J_{max}\}$ **do**
 - 6 Use matrix pencil method to estimate $\{\hat{t}_j\}_{j=1}^J$ from $\tilde{\mathbf{S}}'$.
 - 7 Estimate $\{\hat{a}_j\}_{j=1}^J$ by least squares from samples and signal model.
 - 8 Compute estimation error.
 - 9 **end**
 - 10 Select model order, \hat{K} from $\{1, \dots, J_{max}\}$, by assessing estimation error.
 - 11 Retrieve corresponding estimates, $\{\hat{t}_k, \hat{a}_k\}_{k=1}^{\hat{K}}$, from Steps 6 and 7 with $J = \hat{K}$.
-

bound on the amplitudes that are obtained from the calcium transient model. Without a lower bound, the estimation error would be likely to be minimised by overfitting with a large number of spikes with relatively small amplitudes. The amplitude lower bound implicitly acts as a penalty on the number of spikes, meaning that spikes are not typically inserted where there is a high degree of noise or fluctuations in the baseline fluorescence.

5.4.3 Spike inference algorithm

In Algorithm 1, we summarise the procedure with which spikes are detected from calcium imaging data. When we detect spikes from long streams of data, we use the sliding window approach of Oñativia et al. [87], which we describe briefly here. In this approach, the fluorescence signal has length N samples but only $N_{win} < N$ consecutive samples are analysed at one time. Firstly, the unknown parameters are estimated in the stretch of fluorescence signal defined by samples $\{y[0], y[1], \dots, y[N_{win}]\}$. Then, the sliding window is advanced one sample, so that the parameters are estimated for the stretch corresponding to $\{y[1], y[N], \dots, y[N_{win} + 1]\}$. This process is repeated until the sliding window reaches the end of the data. A histogram is then computed of the estimates from all the sliding windows. Typically, estimates that correspond to true spikes are detected consistently in numerous sliding windows. The final spike locations are then estimated from the peaks of the histogram.

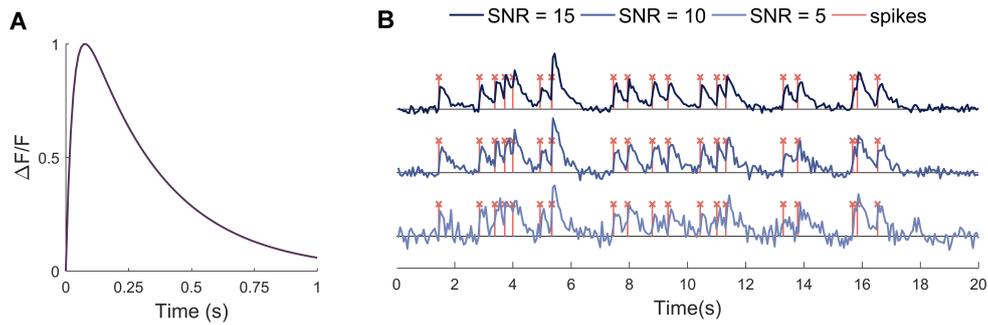


Figure 5.7: Example of simulated data. We simulate calcium transients with pulse shape corresponding to the calcium indicator Cal-520 (**A**). In **B**, we plot examples of fluorescence signals produced from the same underlying spike train at SNRs equal to 5 (dB), 10 (dB) and 15 (dB).

5.5 Experimental methods

In Section 5.6, we demonstrate the performance of the spike inference algorithm on real and simulated data. The real data, which was described in detail in Section 4.5.3, was collected by Dr T. Abrahamsson and Dr P. J. Sjöström. In the following, we present the procedures used to generate the simulated data.

5.5.1 Simulated data

The spike trains are generated from a Poisson distribution with varying rate. We generate the fluorescence pulse with parameters corresponding to the commonly-used synthetic indicator Cal-520 [133], see Fig. 5.7. As is typical for synthetic indicators, the amplitude of Cal-520 pulses decrease as the local spike rate increases. In the simulated data, we compute the amplitude of a spike based on the number of spikes in the previous 250ms time bin, see Table 5.1. This procedure was derived from trends observed in real data by Tada et al. [133]. After the spike train has been convolved with the pulse to generate the fluorescence signal, white Gaussian noise is added to the discrete samples, such that $\tilde{f}[n] = f[n] + \epsilon[n]$, where $f[n]$ is the fluorescence signal sampled at rate 16Hz and $\epsilon[n]$ are samples of white Gaussian noise.

We vary the SNR of the fluorescence signals in order to fully characterise algorithm performance. We use the decibel format of SNR, where $\text{SNR} = 10 \log_{10} P_{\text{sig}}/P_{\text{noise}}$ and P_{sig} and P_{noise} are the powers of the noiseless fluorescence signal samples, $f[n]$,

Number of spikes	0	1	2	3	4
Amplitude	0.27	0.18	0.18	0.14	0.1

Table 5.1: Amplitude of simulated calcium transients. The amplitude of simulated spikes varies with the local spike rate. On real data, it has been observed that amplitudes of calcium transients generated by synthetic indicators decrease as the local spike rate increases [133]. In our simulations, we compute the amplitude of a spike based on the number of spikes in the previous 250ms time bin.

and the noise samples, $\epsilon[n]$, respectively. We calibrate the standard deviation of the noise that attains a given SNR on a fluorescence signal containing one spike at $t_1 = 0$ observed over the range $[0, 1]$ s (as in Fig. 5.7A). This standard deviation is then used to attain that SNR in all simulations. This prevents noise power at a given SNR from varying with the spike rate. In Fig. 5.7B, we show an example of fluorescence signals at different noise levels.

5.6 Results

5.6.1 Comparison with instantaneous-rise algorithm

On simulated data, we examined the temporal error between the true spike times and spike times inferred by the FRI algorithm. The fluorescence signals, which lasted 10s and contained 7 spikes, were generated at three noise levels: 5 dB, 10 dB and 15 dB. Fluorescence signals were generated over 1000 realisations of noise at each SNR. Spikes were detected using the FRI algorithm with both the instantaneous-rise and slow-rise models. The only difference between our implementation of these versions of the algorithm is that the coefficients of the filter, $g[n]$, and the entries of the pre-whitening matrix, \mathbf{W} , differ in the slow-rise and instantaneous-rise case. As the simulated fluorescence signals stem from a slow-rise pulse model, the instantaneous-rise version of the algorithm suffers from model mismatch in this setting. In this example, as the model order estimation procedure of both implementations are identical, we assume that K is known. This allows us to restrict the comparison to the temporal error of each implementation of the algorithm.

In Fig. 5.8, we plot the distribution of the temporal error between estimates and true spikes as a fraction of the sampling period, which is $T = 1/16$. The instantaneous-

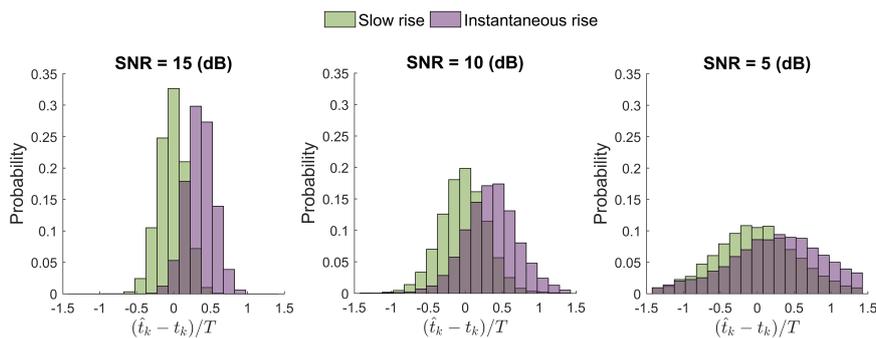


Figure 5.8: Comparison of the temporal error of the FRI algorithm with instantaneous-rise and slow-rise models. The fluorescence signals, which are simulated over 10s and contain 7 spikes, were generated at three noise levels: 5 dB, 10 dB and 15 dB. We plot the temporal error between estimates and true spikes (each estimate is paired with only one true spike) as a fraction of the sample width, T . The distribution of temporal error is plotted over 1000 realisations of noise at each SNR. The instantaneous-rise model exhibits an average bias of 0.33 samples, due to the mismatch between the algorithm model and the data.

rise algorithm exhibits estimation bias due to mismatch between the algorithm model and the data. For data with SNR equal to 15 dB, 10 dB and 5 dB, the average estimation bias is 14s, 21ms and 22ms, respectively. In contrast, using the correct pulse model enables average estimation bias of 1ms in all cases. Using the slow-rise model, 99.3%, 89.2% and 62.7% of estimates, respectively, have super-resolution precision, i.e. precision above the sample width. While using the correct pulse model seems to have rectified the estimation bias, it does not seem to have affected the variability in the estimates. The standard deviations of the temporal error are 10ms, 20ms, and 40ms, respectively, for both the slow and instantaneous-rise algorithm. We conclude that, using the slow-rise pulse model, which is a more accurate representation of a calcium transient, enables detection of spikes with higher temporal precision.

5.6.2 Model order estimation

We tested the accuracy of the model order estimation procedure on simulated data. Specifically, we simulated fluorescence signals over 8s with varying numbers of spikes and SNRs. We used the procedure presented in Section 5.4.2 to estimate the number of spikes in the trace. In Fig. 5.9, we display the distribution of the discrepancy between the true number of spikes, K , and the estimated number, \hat{K} . At each noise level and

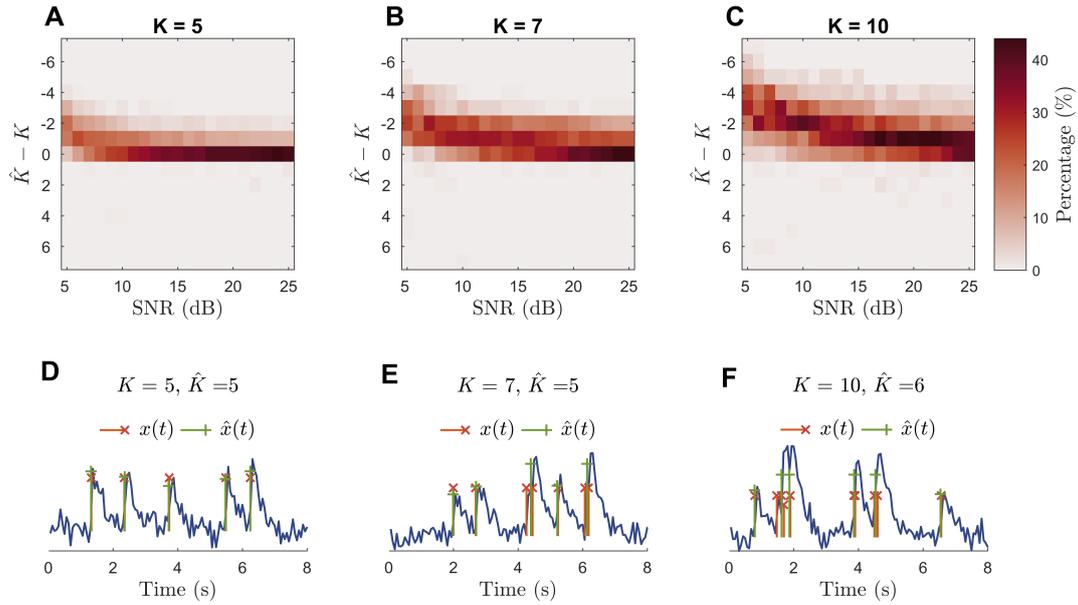


Figure 5.9: Accuracy of the model order estimation procedure on simulated data. A fluorescence signal was simulated over 8s with varying numbers of spikes, K , and SNRs. At each combination of K and SNR, 100 realisations of fluorescence signals were generated. In **A**, **B** and **C**, we display the distribution of the model order estimation error, $\hat{K} - K$, under each set of conditions. In **D**, **E** and **F**, we plot an example true spike train, $x(t)$, and estimated spike train, $\hat{x}(t)$, at SNR equal to 10 (dB) for each value of K .

number of true spikes, the procedure is repeated over 100 realisations of noise and spike trains.

The estimated model order is selected as the one that causes the sharpest decrease in the estimation error, whilst being in the vicinity of the model order that minimises the estimation error. The estimated value is largely accurate; when SNR is sufficiently high, the estimated model order is predominantly equal to or one less than the true value. However, the estimate is generally overly conservative — the number of spikes is more commonly underestimated than overestimated. Furthermore, as the number of spikes in the time period increases, the accuracy decreases. This can be attributed to two factors. Firstly, as spike rates increase, the amplitudes of spikes decrease. These lower amplitude spikes are harder to detect above the noise. Secondly, as the spike rate increases, it is more likely to have closely neighbouring spikes. As is illustrated in Fig. 5.9, these are sometimes identified by the algorithm as a single, higher-amplitude spike.

5.6.3 Comparison with state of the art

On imaging data of the mouse visual cortex acquired at a frame rate of 13 Hz, we compare the performance of the FRI algorithm and a state of the art, non-negative deconvolution algorithm [44]. Whereas the FRI algorithm uses a continuous time formulation of the fluorescence signal, the deconvolution algorithm models the signal as an auto-regressive process. At each sample, the fluorescence intensity is assumed to depend upon previous intensity values and the (unknown) number of spikes fired in that sample. From the data and signal model, a spiking activity vector is inferred, in which each entry represents the probability that a spike was fired in a given sample. Spike inference is performed by solving an isotonic regression problem under sparsity constraints on the spiking activity vector. By subsequently applying a threshold, the spiking probability is converted into the number of spikes per sample.

The dataset contains fluorescence signals from 8 cells with ground truth spiking activity, some of which is plotted in Fig. 4.9. We used one training example per cell with which to fit the parameters for each algorithm. In the case of the FRI algorithm, we fitted the expected amplitude, A , and rise and decay parameters, τ_{on} and τ_{off} . For the deconvolution algorithm, we identified the most suitable threshold, which is related to the calcium transient amplitude, and the optimal lag of the estimates, which is related to the rise-time. Then, we applied the algorithms to the remaining 75 traces in the dataset, see Fig. 5.10. We observed that, on average, the FRI algorithm outperformed the deconvolution algorithm; the algorithms achieved mean CosMIC scores of 0.53 and 0.47, respectively. We also investigated the source of errors by computing CosMIC’s ancestor metrics, P_{CosMIC} and R_{CosMIC} , which are analogous to the precision and recall. On average, the FRI algorithm was awarded higher scores by both of these metrics (Fig. 5.10B). The lower scores of the deconvolution algorithm are likely due to its quantisation of the temporal interval. As spikes are estimated per sample, the temporal precision of the algorithm is limited by the bin width. Deconvolution algorithms may perform relatively better on higher sampling rate or noisier data, on which precision above the bin width is unlikely.

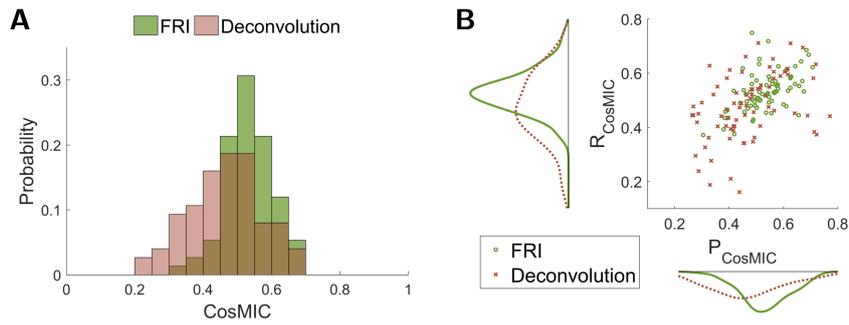


Figure 5.10: Algorithm comparison on real data. On a mouse in vitro imaging dataset with ground truth spiking activity, we compare the performance of the FRI algorithm and a state of the art non-negative deconvolution algorithm. In **A**, we display the distribution of each algorithm’s respective CosMIC scores over all 75 traces. In **B**, we display the distribution of CosMIC’s ancestor metric scores (P_{CosMIC} and R_{CosMIC}) over all 75 traces.

5.7 Summary

We have presented an extension to the FRI framework for spike inference from calcium imaging data. In the original framework, calcium transients were modelled with an instantaneous-rise and exponential decay. While this is suitable for calcium indicators with fast rises and low sampling rate data, in other cases it can lead to estimation bias due to model mismatch. We have extended the framework to encompass calcium indicators that produce slow-rise pulses of fluorescence. On simulated data, we demonstrated that this increases the temporal precision with which spikes are detected. To map the fluorescence signal to the FRI setting, we apply a discrete filter to the samples. This, in turn, alters the statistics of the noise. As the parameter estimation methods we use are better suited to white noise, in the modified algorithm, we implement a procedure to whiten the statistics of the noise. On real data from slices of the mouse visual cortex, we demonstrate that the updated algorithm outperforms a state of the art deconvolution algorithm.

Chapter 6.

Conclusion

6.1 Summary

In this thesis, we considered the problem of detecting cells and cellular activity from two-photon calcium imaging data. Firstly, we presented an algorithm for detecting the locations of cells within a video. Next, we considered the problem of spike inference, in which the timing of spikes are inferred from a cell’s fluorescence signal. Our first contribution in this area was to develop a spike inference metric, which compares the similarity between inferred and ground truth spike trains. Finally, we presented a spike inference algorithm, whose accuracy we assessed with the aforementioned metric.

In Chapter 3, we presented an algorithm, which we refer to as ABLE, for detecting the location of cells within an imaging video. In our framework, multiple coupled active contours evolve, guided by a model-based cost function, to identify cell boundaries. An active contour seeks to partition a local region into two subregions, a cell interior and exterior, in which all pixels have maximally ‘similar’ time courses. This simple, local model allows contours to be evolved predominantly independently. ABLE is a flexible method: we include no priors on a region’s morphology or stereotypical temporal activity. Due to this versatility, ABLE segmented cells with varying size and shape from a mouse in vivo dataset and both active and inactive cells from brain slices. Furthermore, we demonstrated its ability to detect the true boundaries of overlapping cells on both real and simulated data.

In Chapter 4, we presented a metric, referred to as CosMIC, for comparing the similarity between inferred spikes and ground truth spiking activity. Rather than operating

on the true and estimated spike trains directly, the proposed metric assesses the similarity of the pulse trains obtained from convolution of the spike trains with a smoothing pulse. The pulse width is derived from the Cramér-Rao bound (CRB), a lower bound on the variance of any unbiased estimator, which, in turn, is computed from the statistics of the imaging data. As such, spike trains receive a score that is implicitly adjusted for the difficulty of the spike inference problem. The final metric score is the size of the commonalities of the pulse trains as a fraction of their average size. Viewed through the lens of fuzzy set theory, CosMIC resembles a continuous Sørensen-Dice coefficient – an index commonly used to assess the similarity of discrete, presence/absence data. Unlike the spike train correlation, which appears to reward overfitting, the proposed metric score is maximised when the correct number of spikes have been detected. Furthermore, we show that CosMIC is more sensitive to the temporal precision of estimates than either the success rate or the spike train correlation.

The problem of spike inference can be considered to be one of recovering a finite number of unknown parameters from a noisy signal with known parametric structure. This is because neuronal fluorescence signals belong to the class of signals with finite rate of innovation (FRI). In Chapter 5, we extended the FRI framework for spike inference from calcium imaging data. In the original framework, calcium transients were modelled with an instantaneous-rise and exponential decay. While this is suitable for calcium indicators with fast rises and low sampling rate data, in other cases it can lead to estimation bias due to model mismatch. We extended the framework to encompass calcium indicators that produce slow-rise pulses of fluorescence. On simulated data, we demonstrated that this increases the temporal precision with which spikes are detected. We also introduced procedures to improve the robustness of the algorithm to noise. On real data from slices of the mouse visual cortex, we demonstrated that the updated algorithm outperforms a state of the art deconvolution algorithm.

6.2 Future research

We conclude this chapter with a discussion of potential directions for future research.

- In our proposed segmentation algorithm, an active contour moves toward a pixel if that pixel’s time course is more similar to the average time course of the cell interior than the narrowband. As techniques for in vivo imaging advance

(for a review, see [98]), the length of imaging videos, and thus the length of the corresponding activity vectors, are consistently increasing. While we have demonstrated that ABLE’s runtime is virtually unaffected by video length, this does not hold when a video is too large to load into the workstation. It may therefore become necessary to consider techniques by which the temporal dimension of the video can be reduced prior to segmentation. For example, Pachitariu et al. perform a singular value decomposition (SVD) [94] and Spaen et al. [126] represent each pixel in ‘correlation space’ by a vector of correlation coefficients. Rather than using a generic dimensionality reduction technique such as the SVD, it may be beneficial to develop a bespoke technique, which considers both the known parametric structure of the data and the manner in which temporal activity vectors are compared in our segmentation algorithm. This could both extend the range of videos to which ABLE can be applied and, potentially, increase the segmentation accuracy.

- An active contour in our proposed segmentation algorithm is advanced using the level set method. One of the virtues of the level set method is that it is relatively straightforward to extend to higher spatial dimensions. For example, Dufour et al. developed a segmentation algorithm for biological imaging data with three spatial dimensions and one temporal dimension [40], which is closely related to Chan and Vese’s active contours without edges [24]. With the advent of widefield methods such as lightsheet microscopy [1], it is now possible to collect (3+1)-D calcium imaging data. Accordingly, future work may focus on the extension of ABLE to higher spatial dimensions. When extending the framework, it may be useful to consider the impact of photon scattering, which affects widefield microscopy to a greater extent than two-photon microscopy, due to the less-localised excitation of fluorescence.
- Due to the manner in which calcium ions interact with calcium indicators, the amplitude of calcium transients from one neuron are not constant. Rather, they vary with the local spike rate of the neuron. At high spike rates, calcium transient amplitudes can change substantially (see, for example [26]), which can negatively affect spike detection performance. In future, it could thus be beneficial to incorporate a model of the variation in spike amplitudes into the FRI spike detection algorithm. In the FRI algorithm, locations of spikes are inferred sequentially using a sliding window that advances through the data. It may therefore be beneficial to use the knowledge of the number of spikes estimated in previous

windows, to obtain a better estimate of calcium transient amplitudes and, consequently, the number and location of spikes.

Appendix A.

ABLE: supplemental analytical results

A.1 External velocity

In Section 3.4.2, we presented the external energy for $M \geq 1$ cells. In the following, we compute the external velocity of the m^{th} cell. We recall that $\bar{\Omega}$ represents the local spatial domain, which is equal to the union of all contour interiors and their narrowbands. Given two sets $A \subset \bar{\Omega}$ and $B \subset \bar{\Omega}$, we can decompose this as

$$\bar{\Omega} = (A \cap B) \cup (A^c \cap B) \cup (A \cap B^c) \cup (A^c \cap B^c), \quad (\text{A.1})$$

where A^c denotes the complement of A with respect to $\bar{\Omega}$. In particular, when A represents the interior of the m^{th} cell, $A = \Omega^{\text{in},m}$, and B represents the union of the interiors of all other cells, $B = \bigcup_{j \neq m} \Omega^{\text{in},j}$, we have

$$\bar{\Omega} = \left(\Omega^{\text{in},m} \cap \bigcup_{j \neq m} \Omega^{\text{in},j} \right) \cup \left((\Omega^{\text{in},m})^c \cap \bigcup_{j \neq m} \Omega^{\text{in},j} \right) \quad (\text{A.2})$$

$$\cup \left(\Omega^{\text{in},m} \cap \bigcap_{j \neq m} (\Omega^{\text{in},j})^c \right) \cup \left((\Omega^{\text{in},m})^c \cap \bigcap_{j \neq m} (\Omega^{\text{in},j})^c \right) \quad (\text{A.3})$$

$$= S_1 \cup S_2 \cup S_3 \cup S_4, \quad (\text{A.4})$$

where we use De Morgan's laws to convert the complement of the union of sets into the intersection of complements and, in the final line, we introduce a compact notation for the sets. Writing $C_m(\mathbf{x})$ as the function which reports the cell interiors that are present at $\mathbf{x} \in \Omega$ excluding the m^{th} cell, we can decompose the integral in the external

energy over the four disjoint sets, such that

$$\mathcal{E}_{\text{ext}}(\phi_1, \phi_2, \dots, \phi_M) = \int_{\text{outside}} D(I(\mathbf{x}), \mathbf{f}^{\text{out}}) d\mathbf{x} + \int_{\text{inside}} D(I(\mathbf{x}), \sum_{m \in C(\mathbf{x})} \mathbf{f}^{\text{in},m}) d\mathbf{x} \quad (\text{A.5})$$

$$= \int_{S_1} D(I(\mathbf{x}), \mathbf{f}^{\text{in},m} + \sum_{j \in C_m(\mathbf{x})} \mathbf{f}^{\text{in},j}) d\mathbf{x} \quad (\text{A.6})$$

$$+ \int_{S_2} D(I(\mathbf{x}), \sum_{j \in C_m(\mathbf{x})} \mathbf{f}^{\text{in},j}) d\mathbf{x} \quad (\text{A.7})$$

$$+ \int_{S_3} D(I(\mathbf{x}), \mathbf{f}^{\text{in},m}) d\mathbf{x} + \int_{S_4} D(I(\mathbf{x}), \mathbf{f}^{\text{out}}) d\mathbf{x} \quad (\text{A.8})$$

$$= \int_{S_1} c_1(\mathbf{x}) d\mathbf{x} + \int_{S_2} c_2(\mathbf{x}) d\mathbf{x} + \int_{S_3} c_3(\mathbf{x}) d\mathbf{x} + \int_{S_4} c_4(\mathbf{x}) d\mathbf{x}, \quad (\text{A.9})$$

where, in the final step, we introduce a compact notation for the integrands. Using the Heaviside function of the LSFs as an implicit indicator function, we have

$$\mathcal{E}_{\text{ext}}(\phi_1, \phi_2, \dots, \phi_M) = \int_{\bar{\Omega}} c_1(\mathbf{x}) H(\phi_m) \prod_{j \in C_m(\mathbf{x})} H(\phi_j) \prod_{k \notin C_m(\mathbf{x})} (1 - H(\phi_k)) \quad (\text{A.10})$$

$$+ c_2(\mathbf{x}) (1 - H(\phi_m)) \prod_{j \in C_m(\mathbf{x})} H(\phi_j) \prod_{k \notin C_m(\mathbf{x})} (1 - H(\phi_k)) \quad (\text{A.11})$$

$$+ c_3(\mathbf{x}) H(\phi_m) \prod_{j \neq m} (1 - H(\phi_j)) \quad (\text{A.12})$$

$$+ c_4(\mathbf{x}) (1 - H(\phi_m)) \prod_{j \neq m} (1 - H(\phi_j)) d\mathbf{x}. \quad (\text{A.13})$$

As in Section 3.3.1, to obtain the derivative we substitute the regularised version of the Heaviside function and differentiate with respect to ϕ_m . Then, the external velocity is

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m}(\mathbf{x}) = \delta_\epsilon(\phi_m(\mathbf{x})) (c_1(\mathbf{x}) - c_2(\mathbf{x})) \prod_{j \in C_m(\mathbf{x})} H(\phi_j) \prod_{k \notin C_m(\mathbf{x})} (1 - H(\phi_k)) \quad (\text{A.14})$$

$$+ \delta_\epsilon(\phi_m(\mathbf{x})) (c_3(\mathbf{x}) - c_4(\mathbf{x})) \prod_{j \neq m} (1 - H(\phi_j)). \quad (\text{A.15})$$

Substituting back in the values of $c_i(\mathbf{x})$, the external velocity is equal to

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m}(\mathbf{x}) = \delta_\epsilon(\phi_m(\mathbf{x})) (D(I(\mathbf{x}), \mathbf{f}^{\text{in},m}) - D(I(\mathbf{x}), \mathbf{f}^{\text{out}})), \quad (\text{A.16})$$

when \mathbf{x} is not in another cell and

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi_m}(\mathbf{x}) = \delta_\epsilon(\phi_m(\mathbf{x})) \left(D(I(\mathbf{x}), \mathbf{f}^{\text{in},m} + \sum_{j \in C_m(\mathbf{x})} \mathbf{f}^{\text{in},j}) - D(I(\mathbf{x}), \sum_{j \in C_m(\mathbf{x})} \mathbf{f}^{\text{in},j}) \right) \quad (\text{A.17})$$

otherwise.

A.2 Euclidean dissimilarity metric

In this section, we consider the external velocity of an active contour driven by the Euclidean dissimilarity metric. From Eq. (3.12), we have

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi}(\mathbf{x}) = \frac{\delta_\epsilon(\phi(\mathbf{x}))}{m} (\|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{in}}\|^2 - \|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{out}}\|^2), \quad (\text{A.18})$$

where $I(\mathbf{x})$ is the time course of pixel \mathbf{x} , and \mathbf{f}^{in} and \mathbf{f}^{out} are the average time courses from the contour interior and exterior, respectively. We model the time courses as the sum of a stationary baseline component, the resting fluorescence, and a time-varying activity component, which is zero in the absence of activity, such that

$$\mathbf{f}^{\text{out}} = \mathbf{b}^{\text{out}} + \mathbf{a}^{\text{out}}, \quad \mathbf{f}^{\text{in}} = \mathbf{b}^{\text{in}} + \mathbf{a}^{\text{in}} \quad \text{and} \quad \mathbf{I}(\mathbf{x}) = \mathbf{b} + \mathbf{a}. \quad (\text{A.19})$$

We have

$$\|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{in}}\|^2 = \|\mathbf{b} + \mathbf{a} - (\mathbf{b}^{\text{in}} + \mathbf{a}^{\text{in}})\|^2 \quad (\text{A.20})$$

$$= \|(\mathbf{b} - \mathbf{b}^{\text{in}}) + (\mathbf{a} - \mathbf{a}^{\text{in}})\|^2 \quad (\text{A.21})$$

$$= \|\mathbf{b} - \mathbf{b}^{\text{in}}\|^2 - 2\langle \mathbf{b} - \mathbf{b}^{\text{in}}, \mathbf{a} - \mathbf{a}^{\text{in}} \rangle + \|\mathbf{a} - \mathbf{a}^{\text{in}}\|^2. \quad (\text{A.22})$$

Similarly, we have

$$\|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{out}}\|^2 = \|\mathbf{b} - \mathbf{b}^{\text{out}}\|^2 - 2\langle \mathbf{b} - \mathbf{b}^{\text{out}}, \mathbf{a} - \mathbf{a}^{\text{out}} \rangle + \|\mathbf{a} - \mathbf{a}^{\text{out}}\|^2. \quad (\text{A.23})$$

When cellular activity is correlated with the background and the cell's baseline fluorescence is distinguishable above the background, the elements of $\mathbf{a} - \mathbf{a}^{\text{in}}$ and $\mathbf{a} - \mathbf{a}^{\text{out}}$ are small compared to $\mathbf{b} - \mathbf{b}^{\text{in}}$ and $\mathbf{b} - \mathbf{b}^{\text{out}}$. Then, we can approximate the Euclidean

distance as

$$\|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{in}}\|^2 \approx \|\mathbf{b} - \mathbf{b}^{\text{in}}\|^2 \quad \text{and} \quad \|\mathbf{I}(\mathbf{x}) - \mathbf{f}^{\text{out}}\|^2 \approx \|\mathbf{b} - \mathbf{b}^{\text{out}}\|^2. \quad (\text{A.24})$$

In this case, the Euclidean distance in Eq. (A.18) is driven by the discrepancies in baseline intensity rather than the time-varying activity, such that

$$\frac{\partial \mathcal{E}_{\text{ext}}}{\partial \phi}(\mathbf{x}) \approx \frac{\delta_{\epsilon}(\phi(\mathbf{x}))}{m} (\|\mathbf{b} - \mathbf{b}^{\text{in}}\|^2 - \|\mathbf{b} - \mathbf{b}^{\text{out}}\|^2). \quad (\text{A.25})$$

Appendix B.

CosMIC: proof of analytical results

In this section we provide proofs of some analytical results presented in Section 4. Proposition B.1 presents a closed-form expression for the score of an estimate of a single spike. We see that, for estimates within the precision of the pulse width, CosMIC monotonically decreases as estimates are further from the true spike. The result of B.1 is then used to establish the pulse width that, on average, awards a score of 0.8 to an estimate at the precision of the CRB. This result and the corresponding derivation are presented in Proposition B.2. To improve the readability of the formulae, we denote the metric score of sets of single spikes $S = \{t_1\}$ and $\hat{S} = \{\hat{t}_1\}$ as $m(t_1, \hat{t}_1) := M(S, \hat{S})$. Finally, in Propositions B.3 and B.4, we derive an alternative expression for the metric score when estimates exactly detect the location of true spikes, but the correct number of spikes is not detected.

Proposition B.1. *Let there be a single ground truth spike at t_1 and an estimate, $\hat{t}_1 = t_1 + u$. The metric score is*

$$m(t_1, t_1 + u) = \begin{cases} \left(\frac{|u|}{2\epsilon} - 1\right)^2 & \text{if } |u| < 2\epsilon \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.1})$$

where ϵ is half the width of the pulse.

Proof. Without loss of generality, we let the true spike location be at $t_1 = 0$, as the metric score depends on the relative rather than absolute locations of the estimated

and ground truth spikes. From Eq. (4.25), we have

$$m(0, u) = 2 \frac{\|\min(q_\epsilon(t), q_\epsilon(t-u))\|_1}{\|q_\epsilon(t)\|_1 + \|q_\epsilon(t-u)\|_1}. \quad (\text{B.2})$$

When $|u| > 2\epsilon$, the pulses do not overlap and, consequently, the numerator is equal to 0. Therefore, the metric score is zero for all $|u| > 2\epsilon$. For $|u| \leq 2\epsilon$, we write

$$m(0, u) = \frac{1}{\epsilon} \left(\int_A q_\epsilon(t) dt + \int_B q_\epsilon(t-u) dt \right), \quad (\text{B.3})$$

which follows from $\|q_\epsilon\| = \epsilon$, $A = \{t \in \mathbb{R} : q_\epsilon(t) < q_\epsilon(t-u)\}$ and $B = \{t \in \mathbb{R} : q_\epsilon(t) \geq q_\epsilon(t-u)\}$. From the change of variables $v = t + u$, we see that $m(0, u) = m(0, -u)$. As M is even in the second argument, we must only evaluate $m(0, u)$ for $0 < u < 2\epsilon$. To identify the support of A and B , we must identify the value of t at which $q_\epsilon(t) = q_\epsilon(t-u)$. We have

$$q_\epsilon(t) = q_\epsilon(t-u) \Leftrightarrow 1 - \frac{|t|}{\epsilon} = 1 - \frac{|t-u|}{\epsilon} \Leftrightarrow |t| = |t-u|. \quad (\text{B.4})$$

For $0 < u < 2\epsilon$, the relevant intersection point is at $t = u/2$. Omitting the range of integration at which the integrand is zero, Eq. (B.3) becomes

$$m(0, u) = \frac{1}{\epsilon} \left(\int_{u/2}^\epsilon q_\epsilon(t) dt + \int_{u-\epsilon}^{u/2} q_\epsilon(t-u) dt \right) \quad (\text{B.5})$$

$$= \frac{1}{\epsilon} \left(\int_{u/2}^\epsilon q_\epsilon(t) dt + \int_{-\epsilon}^{-u/2} q_\epsilon(v) dv \right) \quad (\text{B.6})$$

$$= \frac{2}{\epsilon} \int_{u/2}^\epsilon q_\epsilon(t) dt, \quad (\text{B.7})$$

which follows from the change of variables $v = t + u$ and the symmetry of $q_\epsilon(t)$ about 0. Evaluating the integral, we obtain $m(0, u) = (|u|/2\epsilon - 1)^2$ for $|u| < 2\epsilon$.

□

Proposition B.2. *Let t_1 denote the location of the true spike. The spike estimate, modelled with the random variable $U \sim \mathcal{N}(t_1, \sigma_{CRB}^2)$, is normally distributed about the true spike at the precision of the CRB. We denote $\beta = \sigma_{CRB}/w$, where $w = 2\epsilon$ is the pulse width. Then, the expectation of the metric score is 0.8 if β satisfies*

$$0.4 = (\Phi(1/\beta) - 0.5) (\beta^2 + 1) + \frac{\beta}{\sqrt{2\pi}} (\exp(-1/2\beta^2) - 2), \quad (\text{B.8})$$

where Φ denotes the cumulative distribution function of the standard normal distribution.

Proof. Without loss of generality, we consider the case where $t_1 = 0$. Denoting $f_U(u)$ as the PDF of U , we have

$$\mathbb{E}_U [m(0, U)] = \int_{\mathbb{R}} m(0, u) f_U(u) du \quad (\text{B.9})$$

$$= 2 \int_0^w \left(\frac{u}{w} - 1 \right)^2 f_U(u) du \quad (\text{B.10})$$

$$= \frac{2}{w^2} \int_0^w u^2 f_U(u) du - \frac{4}{w} \int_0^w u f_U(u) du + 2 \int_0^w f_U(u) du \quad (\text{B.11})$$

$$= \frac{2}{w^2} I_1 - \frac{4}{w} I_2 + 2I_3, \quad (\text{B.12})$$

where Eq. (B.10) follows from Proposition B.1. Applying integration by parts to I_1 , we obtain

$$I_1 = \sigma_{\text{CRB}}^2 \mathbb{P}(U \in [0, w]) - \sigma_{\text{CRB}}^2 f_U(w) w. \quad (\text{B.13})$$

The remaining integrals are $I_2 = -\sigma_{\text{CRB}}^2 (f_U(w) - f_U(0))$ and $I_3 = \mathbb{P}(U \in [0, w])$, respectively. Putting the integrals together yields

$$\mathbb{E} [m(0, U)] = 2 \left[\mathbb{P}(U \in [0, w]) \left(\frac{\sigma_{\text{CRB}}^2}{w^2} + 1 \right) + \frac{\sigma_{\text{CRB}}^2}{w} (f_U(w) - 2f_U(0)) \right]. \quad (\text{B.14})$$

Writing $\beta = \sigma_{\text{CRB}}/w$ and expanding the distribution function of U , we have

$$\mathbb{E}_U [m(0, U)] = 2 \left((\Phi(1/\beta) - 0.5) (\beta^2 + 1) + \frac{\beta}{\sqrt{2\pi}} (\exp(-1/2\beta^2) - 2) \right). \quad (\text{B.15})$$

□

Proposition B.3. *Let S be a set of K true spikes and \hat{S} a set of \hat{K} estimates. If \hat{S} contains a subset of the ground truth spike times with the exception of R missing spikes and no extras, such that $\hat{K} = K - R$ with $0 < R \leq K$, then*

$$M(S, \hat{S}) = 1 - \frac{1}{2K/R - 1}. \quad (\text{B.16})$$

Proof. We denote with $x_r(t)$ and $r(t)$ the spike train and pulse train, respectively, of the spikes present in S but missing from \hat{S} . Due to the distributivity of the convolution

operation, we have

$$\hat{\mu}(t) = \hat{x}(t) * q_\epsilon(t) = (x(t) - x_r(t)) * q_\epsilon(t) = \mu(t) - r(t).$$

From the metric expression in Eq. (4.32), the metric score becomes

$$M(S, \hat{S}) = 1 - \frac{\|\mu - \hat{\mu}\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1} \quad (\text{B.17})$$

$$= 1 - \frac{\|\mu - (\mu - r)\|_1}{\|\mu\|_1 + \|\mu - r\|_1} \quad (\text{B.18})$$

$$= 1 - \frac{R\|q_\epsilon\|_1}{K\|q_\epsilon\|_1 + (K - R)\|q_\epsilon\|_1} \quad (\text{B.19})$$

$$= 1 - \frac{1}{2K/R - 1}. \quad (\text{B.20})$$

□

Proposition B.4. *Let S be set of K spikes and \hat{S} a set of \hat{K} estimates containing all the ground truth spikes together with $R > 0$ extras, such that $\hat{K} = K + R$. Then*

$$M(S, \hat{S}) = \frac{1}{1 + R/2K}. \quad (\text{B.21})$$

Proof. We denote with $x_r(t)$ and $r(t)$ the spike train and pulse train, respectively, of the spikes present in \hat{S} but missing from S . Due to the distributivity of the convolution operator, the estimated pulse train can be written

$$\hat{\mu}(t) = \hat{x}(t) * q_\epsilon(t) = (x(t) + x_r(t)) * q_\epsilon(t) = \mu(t) + r(t).$$

The metric score is

$$M(S, \hat{S}) = \frac{2\|\min(\mu, \hat{\mu})\|_1}{\|\mu\|_1 + \|\hat{\mu}\|_1} \quad (\text{B.22})$$

$$= \frac{2\|\min(\mu, \mu + r)\|_1}{\|\mu\|_1 + \|\mu + r\|_1} \quad (\text{B.23})$$

$$= \frac{2\|\mu\|_1}{2\|\mu\|_1 + \|r\|_1} \quad (\text{B.24})$$

$$= \frac{1}{1 + \|r\|_1/2\|\mu\|_1}, \quad (\text{B.25})$$

where the penultimate line follows from the non-negativity of μ and r . As $\|\mu\|_1 =$

$K\|q_\epsilon\|_1$ and $\|r\|_1 = R\|q_\epsilon\|_1$, it follows that

$$M(S, \hat{S}) = \frac{1}{1 + R/2K}. \quad (\text{B.26})$$

□

Bibliography

- [1] M. B. Ahrens, M. B. Orger, D. N. Robson, J. M. Li, and P. J. Keller. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nature methods*, 10:413–420, 2013.
- [2] J. Akerboom, T.-W. Chen, T. J. Wardill, L. Tian, J. S. Marvin, S. Mutlu, N. C. Calderón, F. Esposti, B. G. Borghuis, X. R. Sun, A. Gordus, M. B. Orger, R. Portugues, F. Engert, J. J. Macklin, A. Filosa, A. Aggarwal, R. A. Kerr, R. Takagi, S. Kracun, E. Shigetomi, B. S. Khakh, H. Baier, L. Lagnado, S. S.-H. Wang, C. I. Bargmann, B. E. Kimmel, V. Jayaraman, K. Svoboda, D. S. Kim, E. R. Schreiter, and L. L. Looger. Optimization of a GCaMP calcium indicator for neural activity imaging. *Journal of Neuroscience*, 32(40):13819–13840, 2012.
- [3] D. G. Altman and J. M. Bland. Measurement in medicine: the analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 32(3):307–317, 1983.
- [4] N. J. Apthorpe, A. J. Riordan, R. E. Aguilar, J. Homann, Y. Gu, D. W. Tank, and H. S. Seung. Automatic neuron detection in calcium imaging data using convolutional networks. In *Proceedings of the 29th International Conference on Neural Information Processing Systems, NIPS’16*, pages 3278–3286, USA, 2016. Curran Associates Inc.
- [5] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer-Verlag New York, second edition, 2006.
- [6] R. Bellman and M. Giertz. On the analytic formalism of the theory of fuzzy sets. *Information Sciences*, 5:149–156, 1973.
- [7] P. Berens, J. Freeman, T. Deneux, N. Chenkov, T. McColgan, A. Speiser, J. H. Macke, S. C. Turaga, P. Mineault, P. Rupprecht, S. Gerhard, R. W. Friedrich,

- J. Friedrich, L. Paninski, M. Pachitariu, K. D. Harris, B. Bolte, T. A. Machado, D. Ringach, J. Stone, L. E. Rogerson, N. J. Sofroniew, J. Reimer, E. Froudarakis, T. Euler, M. Romn Rosn, L. Theis, A. S. Tolias, and M. Bethge. Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *PLOS Computational Biology*, 14(5):1–13, 2018.
- [8] J. Berent and P. L. Dragotti. Plenoptic manifolds. *IEEE Signal Processing Magazine*, 24(6):34–44, 2007.
- [9] M. J. Berridge, P. Lipp, and M. D. Bootman. The versatility and universality of calcium signalling. *Nature Reviews Molecular Cell Biology*, 1:11 – 21, 2000.
- [10] V. Bonin, M. H. Histed, S. Yurgenson, and R. C. Reid. Local diversity and fine-scale organization of receptive fields in mouse visual cortex. *Journal of Neuroscience*, 31(50):18506–18521, 2011.
- [11] J. G. G. Borst and B. Sakmann. Calcium current during a single action potential in a large presynaptic terminal of the rat brainstem. *The Journal of Physiology*, 506(Pt 1):143–157, 1998.
- [12] M. B. Bouchard, V. Voleti, C. S. Mendes, C. Lacefield, W. B. Grueber, R. S. Mann, R. M. Bruno, and M. C. Hillman. Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms. *Nature Photonics*, 9:113–119, 2015.
- [13] J. R. Bray and J. T. Curtis. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, 27(4):325–349, 1957.
- [14] R. Brette and A. Destexhe. *Handbook of Neural Activity Measurement*. Cambridge University Press, 2012.
- [15] P. Brigger, J. Hoeg, and M. Unser. B-spline snakes: a flexible tool for parametric contour detection. *IEEE Transactions on Image Processing*, 9(9):1484–1496, 2000.
- [16] K. A. Buchanan, A. V. Blackman, A. W. Moreau, D. Elgar, R. P. Costa, T. Lalanne, A. A. Tudor Jones, J. Oyrer, and P. J. Sjöström. Target-specific expression of presynaptic NMDA receptors in neocortical microcircuits. *Neuron*, 75(3):451–466, 2012.

-
- [17] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10:186–198, 2009.
- [18] G. Buzsáki. Large-scale recording of neuronal ensembles. *Nature Neuroscience*, 7:446–451, 2004.
- [19] J. A. Cadzow. Signal enhancement—a composite property mapping algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(1):49–62, 1988.
- [20] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.
- [21] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
- [22] T. Chan and L. Vese. An active contour model without edges. In M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision*, pages 141–151, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.
- [23] T. F. Chan, B. Sandberg, and L. A. Vese. Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11(2):130–141, 2000.
- [24] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
- [25] T. F. Chan and L. A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *Proceedings IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 161–168, Vancouver, BC, 2001. IEEE.
- [26] T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, and D. S. Kim. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, 499:295–300, 2013.
- [27] A. Danielyan, Y.-W. Wu, P.-Y. Shih, Y. Dembitskaya, and A. Semyanov. Denoising of two-photon fluorescence images with block-matching 3D filtering. *Methods*, 68(2):308–316, 2014.

- [28] R. Delgado-Gonzalo, V. Uhlmann, D. Schmitter, and M. Unser. Snakes on a plane: A perfect snap for bioimage analysis. *IEEE Signal Processing Magazine*, 32(1):41–48, 2015.
- [29] R. Delgado-Gonzalo and M. Unser. Spline-based framework for interactive segmentation in biomedical imaging. *IRBM*, 34(3):235–243, 2013.
- [30] T. Deneux, A. Kaszas, G. Szalay, K. Katona, T. Lakner, A. Grinvald, B. Rózsa, and I. Vanzetta. Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. *Nature Communications*, 7:12190, 2016.
- [31] W. Denk, J. H. Strickler, and W. W. Webb. Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76, 1990.
- [32] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [33] A. S. Dickey, A. Suminski, Y. Amit, and N. G. Hatsopoulos. Single-unit stability using chronically implanted multielectrode arrays. *Journal of Neurophysiology*, 102(2):1331–1339, 2009.
- [34] F. Diego Andilla and F. A. Hamprecht. Sparse space-time deconvolution for calcium image analysis. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, NIPS’14, pages 64–72, USA, 2014. Curran Associates Inc.
- [35] Z. Doğan, T. Blu, and D. Van De Ville. Detecting spontaneous brain activity in functional magnetic resonance imaging using finite rate of innovation. In *11th International Symposium on Biomedical Imaging*, ISBI, pages 1047–1050. IEEE, 2014.
- [36] Z. Doğan, C. Gilliam, T. Blu, and D. Van De Ville. Reconstruction of finite rate of innovation signals with model-fitting approach. *IEEE Transactions on Signal Processing*, 63(22):6024–6036, 2015.
- [37] R. J. Douglas and K. A. C. Martin. Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, 27(1):419–451, 2004.
- [38] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Transactions on Signal Processing*, 55(5):1741–1757, 2007.

-
- [39] D. Dubois and H. Prade. *Fuzzy sets and systems: theory and applications*. Academic Press, New York, 1980.
- [40] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J. C. Olivo-Marin, and C. Zimmer. Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Transactions on Image Processing*, 14(9):1396–1410, 2005.
- [41] Y. C. Eldar and A. V. Oppenheim. MMSE whitening and subspace whitening. *IEEE Transactions on Information Theory*, 49(7):1846–1851, 2003.
- [42] M. A. T. Figueiredo, J. M. N. Leitao, and A. K. Jain. Unsupervised contour representation and estimation using B-splines and a minimum description length criterion. *IEEE Transactions on Image Processing*, 9(6):1075–1087, 2000.
- [43] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, 2008.
- [44] J. Friedrich, P. Zhou, and L. Paninski. Fast online deconvolution of calcium imaging data. *PLOS Computational Biology*, 13(3):1–26, 2017.
- [45] A. Giovannucci, J. Friedrich, M. Kaufman, A. Churchland, D. Chklovskii, L. Paninski, and E. A. Pnevmatikakis. OnACID: Online analysis of calcium imaging data in real time. In *Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS’17*, pages 2378–2388, USA, 2017. Curran Associates Inc.
- [46] W. Göbel and F. Helmchen. In vivo calcium imaging of neural network function. *Physiology*, 22(6):358–365, 2007.
- [47] B. F. Grewe, D. Langer, H. Kasper, B. M. Kampa, and F. Helmchen. High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision. *Nature Methods*, 7:399–405, 2010.
- [48] C. Grienberger and A. Konnerth. Imaging calcium in neurons. *Neuron*, 73(5):862–885, 2012.
- [49] Z. V. Guo, N. Li, D. Huber, E. Ophir, D. Gutnisky, J. T. Ting, G. Feng, and K. Svoboda. Flow of cortical activity underlying a tactile decision in mice. *Neuron*, 81(1):179–194, 2014.

- [50] O. P. Hamill, A. Marty, E. Neher, B. Sakmann, and F. J. Sigworth. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv*, 391(2):85–100, 1981.
- [51] K. D. Harris, R. Q. Quiroga, J. Freeman, and S. Smith. Improving data quality in neuronal population recordings. *Nature Neuroscience*, 19:1165–1174, 2016.
- [52] F. Helmchen and W. Denk. Deep tissue two-photon microscopy. *Nature Methods*, 2:932–940, 2005.
- [53] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, 1990.
- [54] Y. Hua and T. K. Sarkar. On SVD for estimating generalized eigenvalues of singular matrix pencil in noise. In *IEEE International Symposium on Circuits and Systems*, volume 5, pages 2780–2783, 1991.
- [55] D. Huber, D. A. Gutnisky, S. Peron, D. H. O’Connor, J. S. Wiegert, L. Tian, T. G. Oertner, L. L. Looger, and K. Svoboda. Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*, 484:473–478, 2012.
- [56] H. Inan, M. A. Erdogdu, and M. Schnitzer. Robust estimation of neural signals in calcium imaging. In *Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS’17*, pages 2901–2910, USA, 2017. Curran Associates Inc.
- [57] M. Jacob, T. Blu, and M. Unser. Efficient energies and algorithms for parametric snakes. *IEEE Transactions on Image Processing*, 13(9):1231–1244, 2004.
- [58] S. Jehan-Besson, M. Barlaud, and G. Aubert. DREAM2S: Deformable regions driven by an Eulerian accurate minimization method for image and video segmentation. *International Journal of Computer Vision*, 53(1):45–70, 2003.
- [59] N. Ji, J. Freeman, and S. L. Smith. Technologies for imaging neural activity in large volumes. *Nature Neuroscience*, 19:1154–1164, 2016.
- [60] P. Kaifosh, J. D. Zaremba, N. B. Danielson, and A. Losonczy. SIMA: Python software for analysis of dynamic fluorescence imaging data. *Frontiers in Neuroinformatics*, 8:80, 2014.

-
- [61] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- [62] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [63] A. M. Kerlin, M. L. Andermann, V. K. Berezovskii, and R. C. Reid. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron*, 67(5):858–871, 2010.
- [64] J. N. D. Kerr, D. Greenberg, and F. Helmchen. Imaging input and output of neocortical networks in vivo. *Proceedings of the National Academy of Sciences*, 102(39):14063–14068, 2005.
- [65] A. Klibisz, D. Rose, M. Eicholtz, J. Blundon, and S. Zakharenko. Fast, simple calcium imaging segmentation with fully convolutional networks. In M. J. Cardoso et al., editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, DLMIA 2017, ML-CDS 2017. Lecture Notes in Computer Science, vol 10553. Springer, Cham, 2017.
- [66] J. J. Knierim. Chapter 19 - Information processing in neural networks. In J. H. Byrne, R. Heidelberger, and M. N. Waxham, editors, *From Molecules to Networks*, pages 563 – 589. Academic Press, Boston, 2014.
- [67] V. A. Kotelnikov. On the transmission capacity of the ether and wire in electro-communications. *Izd. Red. Upr. Svyazzi RSKA (Moscow)*, 1933.
- [68] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore. A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Transactions on Image Processing*, 20(7):2007–2016, 2011.
- [69] C. Li, C. Xu, C. Gui, and M. D. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19(12):3243–3254, 2010.
- [70] A. Longstaff. *Neuroscience*. BIOS instant notes series. Taylor and Francis, London, third edition, 2011.
- [71] L. L. Looger and O. Griesbeck. Genetically encoded neural activity indicators. *Current Opinion in Neurobiology*, 22(1):18–23, 2012.

- [72] Z. Lu, G. Carneiro, and A. P. Bradley. Automated nucleus and cytoplasm segmentation of overlapping cervical cells. In *Medical Image Computing and Computer-Assisted Intervention*, MICCAI 2013, pages 452–460, Berlin, 2013. Springer.
- [73] H. Lütcke, F. Gerhard, F. Zenke, W. Gerstner, and F. Helmchen. Inference of neuronal network spike dynamics and topology from calcium imaging data. *Frontiers in Neural Circuits*, 7:201, 2013.
- [74] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.
- [75] S. Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, third edition, 2009.
- [76] R. Maruyama, K. Maeda, H. Moroda, I. Kato, M. Inoue, H. Miyakawa, and T. Aonishi. Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Networks*, 55:11–19, 2014.
- [77] R. McAulay and E. Hofstetter. Barankin bounds on parameter estimation. *IEEE Transactions on Information Theory*, 17(6):669–676, 1971.
- [78] A. Miri, K. Daie, R. D. Burdine, E. Aksay, and D. W. Tank. Regression-based identification of behavior-encoding neurons during large-scale optical imaging of neural activity at cellular resolution. *Journal of Neurophysiology*, 105(2):964–980, 2011.
- [79] A. Mitchie, R. Feghali, and A. Mansouri. Motion tracking as a spatio-temporal motion boundary detection. *Robotics and autonomous systems*, 43:39–50, 2003.
- [80] A. Miyawaki, J. Llopis, R. Heim, J. M. McCaffery, J. A. Adams, M. Ikura, and R. Y. Tsien. Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature*, 388:882–887, 1997.
- [81] V. B. Mountcastle. An organizing principle for cerebral function: The unit model and the distributed system. In G. M. Edelman and V. V. Mountcastle, editors, *The Mindful Brain*, pages 7–50. MIT Press, Cambridge, MA, 1978.
- [82] E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, 63(6):747–760, 2009.

-
- [83] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [84] J. Murray-Bruce and P. L. Dragotti. Estimating localized sources of diffusion fields using spatiotemporal sensor measurements. *IEEE Transactions on Signal Processing*, 63(12):3018–3031, 2015.
- [85] <http://neurofinder.codeneuro.org/>. Accessed: 2018-01-04.
- [86] K. Ohki, S. Chung, P. Kara, M. Hübener, T. Bonhoeffer, and R. C. Reid. Highly ordered arrangement of single neurons in orientation pinwheels. *Nature*, 442:925–928, 2006.
- [87] J. Oñativia, S. R. Schultz, and P. L. Dragotti. A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *Journal of Neural Engineering*, 10(4):046017, 2013.
- [88] J. Oñativia Bravo. *Sampling and reconstruction of finite rate of innovations signals with applications in neuroscience and sparse representation*. PhD thesis, Imperial College London, 2015.
- [89] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153 of *Applied Mathematical Sciences*. Springer-Verlag New York, 2003.
- [90] S. Osher and N. Paragios. *Geometric level set methods in imaging, vision, and graphics*. Springer-Verlag New York, 2003.
- [91] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988.
- [92] M. Pachitariu, A. M. Packer, N. Pettit, H. Dalgleish, M. Häusser, and M. Sahani. Extracting regions of interest from biological images with convolutional sparse block coding. In *Proceedings of the 26th International Conference on Neural Information Processing Systems, NIPS’13*, pages 1745–1753, USA, 2013. Curran Associates Inc.
- [93] M. Pachitariu, C. Stringer, and K. D. Harris. Robustness of spike deconvolution for calcium imaging of neural spiking. *bioRxiv*, 2017. doi:10.1101/156786.

- [94] M. Pachitariu, C. Stringer, S. Schröder, M. Dipoppa, L. F. Rossi, M. Carandini, and K. D. Harris. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*, 2016. doi:10.1101/061507.
- [95] A. R. C. Paiva, I. Park, and J. C. Príncipe. A comparison of binless spike train measures. *Neural Computing and Applications*, 19(3):405–419, 2010.
- [96] C. P. Pappis and N. I. Karacapilidis. A comparative assessment of measures of similarity of fuzzy values. *Fuzzy Sets and Systems*, 56(2):171–174, 1993.
- [97] J. B. Pawley. *Handbook of Biological Confocal Microscopy*. Springer US, third edition, 2006.
- [98] S. Peron, T.-W. Chen, and K. Svoboda. Comprehensive imaging of cortical networks. *Current Opinion in Neurobiology*, 32:115–123, 2015.
- [99] S. Peron, J. Freeman, V. Iyer, C. Guo, and K. Svoboda. Volumetric calcium imaging data recorded during performance of a single whisker object localization task, sampling activity in the majority of the relevant superficial barrel cortex neurons (75%, 12,000 neurons per mouse). CRCNS.org. doi:10.6080/K0TB14TN, 2015.
- [100] S. P. Peron, J. Freeman, V. Iyer, C. Guo, and K. Svoboda. A cellular resolution map of barrel cortex activity during tactile behavior. *Neuron*, 86(3):783–799, 2015.
- [101] D. S. Peterka, H. Takahashi, and R. Yuste. Imaging voltage in neurons. *Neuron*, 69(1):9–21, 2011.
- [102] E. C. Pielou. *The interpretation of ecological data: A primer on classification and ordination*. John Wiley and Sons, Inc., 1984.
- [103] E. A. Pnevmatikakis, J. Merel, A. Pakman, and L. Paninski. Bayesian spike inference from calcium imaging data. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 349–353. IEEE, 2013.
- [104] E. A. Pnevmatikakis and L. Paninski. Sparse nonnegative deconvolution for compressive calcium imaging: algorithms and phase transitions. In *Proceedings of the 26th International Conference on Neural Information Processing Systems, NIPS'13*, pages 1250–1258, USA, 2013. Curran Associates Inc.

-
- [105] E. A. Pnevmatikakis, D. Soudry, Y. Gao, T. A. Machado, J. Merel, D. Pfau, T. Reardon, Y. Mu, C. Lacefield, W. Yang, et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron*, 89(2):285–299, 2016.
- [106] V. Rahmati, K. Kirmse, D. Markovi, K. Holthoff, and S. J. Kiebel. Inferring neuronal dynamics from calcium imaging data using biophysical models and bayesian inference. *PLOS Computational Biology*, 12(2):1–42, 2016.
- [107] S. Reynolds, T. Abrahamsson, R. Schuck, P. J. Sjöström, S. R. Schultz, and P. L. Dragotti. ABLE: An activity-based level set segmentation algorithm for two-photon calcium imaging data. *eNeuro*, 4(5), 2017.
- [108] S. Reynolds, T. Abrahamsson, P. J. Sjöström, S. R. Schultz, and P. L. Dragotti. CosMIC: A consistent metric for spike inference from calcium imaging. *Neural Computation*, In press.
- [109] S. Reynolds, C. S. Copeland, S. R. Schultz, and P. L. Dragotti. An extension of the FRI framework for calcium transient detection. In *13th International Symposium on Biomedical Imaging, ISBI*, pages 676–679. IEEE, 2016.
- [110] M. Ristivojevic and J. Konrad. Space-time image sequence analysis: object tunnels and occlusion volumes. *IEEE Transactions on Image Processing*, 15(2): 364–376, 2006.
- [111] D. W. Roberts. Ordination on the basis of fuzzy set theory. *Vegetatio*, 66(3): 123–131, 1986.
- [112] N. L. Rochefort, H. Jia, and A. Konnerth. Calcium imaging in the living brain: prospects for molecular medicine. *Trends in Molecular Medicine*, 14(9):389–399, 2008.
- [113] H. L. Royden. *Real analysis*. Prentice Hall, Englewood Cliffs, third edition, 1988.
- [114] T. Sasaki, N. Takahashi, N. Matsuki, and Y. Ikegaya. Fast and accurate detection of action potentials from somatic calcium fluctuations. *Journal of Neurophysiology*, 100(3):1668–1676, 2008.
- [115] M. Scanziani and M. Häusser. Electrophysiology in the age of light. *Nature*, 461: 930–939, 2009.

- [116] S. Schreiber, J. Fellous, D. Whitmer, P. Tiesinga, and T. Sejnowski. A new correlation-based measure of spike timing reliability. *Neurocomputing*, 52-54 (Supplement C):925–931, 2003.
- [117] R. Schuck, M. A. Go, S. Garasto, S. Reynolds, P. L. Dragotti, and S. R. Schultz. Multiphoton minimal inertia scanning for fast acquisition of neural activity signals. *Journal of Neural Engineering*, 15(2):025003, 2018.
- [118] S. R. Schultz, K. Kitamura, A. Post-Uiterweer, J. Krupic, and M. Häusser. Spatial pattern coding of sensory information by climbing fiber-evoked calcium signals in networks of neighboring cerebellar purkinje cells. *Journal of Neuroscience*, 29(25):8005–8015, 2009.
- [119] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IEEE*, 37(1):10–21, 1949.
- [120] M. E. J. Sheffield and D. A. Dombeck. Calcium transient prevalence across the dendritic arbour predicts place field properties. *Nature*, 517:200–204, 2015.
- [121] D. Slepian. On bandwidth. *Proceedings of the IEEE*, 64(3):292–300, 1976.
- [122] S. L. Smith and M. Häusser. Parallel processing of visual space by neighboring neurons in mouse visual cortex. *Nature Neuroscience*, 13:1144–1149, 2010.
- [123] N. J. Sofroniew, J. D. Cohen, A. K. Lee, and K. Svoboda. Natural whisker-guided behavior by head-fixed mice in tactile virtual reality. *Journal of Neuroscience*, 34(29):9537–9550, 2014.
- [124] N. J. Sofroniew, D. Flickinger, J. King, and K. Svoboda. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife*, 5:e14472, 2016.
- [125] T. Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Biologiske Skrifter*, 5:1–34, 1948.
- [126] Q. Spaen, D. S. Hochbaum, and R. Asín-Achá. HNCcorr: A Novel Combinatorial Approach for Cell Identification in Calcium-Imaging Movies. *ArXiv e-prints*, 2017. arXiv:1703.01999.
- [127] A. Speiser, J. Yan, E. W. Archer, L. Buesing, S. C. Turaga, and J. H. Macke. Fast amortized inference of neural activity from calcium imaging data with variational

- autoencoders. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'17, pages 4024–4034, USA, 2017. Curran Associates Inc.
- [128] J. M. Squirrell, D. L. Wokosin, J. G. White, and B. D. Bavister. Long-term two-photon fluorescence imaging of mammalian embryos without compromising viability. *Nature Biotechnology*, 17:763–767, 1999.
- [129] P. Stoica and R. Moses. *Spectral Analysis of Signals*. Prentice Hall, 2005.
- [130] C. Stosiek, O. Garaschuk, K. Holthoff, and A. Konnerth. In vivo two-photon calcium imaging of neuronal networks. *Proceedings of the National Academy of Sciences*, 100(12):7319–7324, 2003.
- [131] K. Svoboda and S. M. Block. Biological applications of optical forces. *Annual Review of Biophysics and Biomolecular Structure*, 23:247–285, 1994.
- [132] K. Svoboda (contact). GENIE project, Janelia Farm Campus, HHMI. (2015). Simultaneous imaging and loose-seal cell-attached electrical recordings from neurons expressing a variety of genetically encoded calcium indicators. CRCNS.org. doi:10.6080/k02r3pmm.
- [133] M. Tada, A. Takeuchi, M. Hashizume, K. Kitamura, and M. Kano. A highly sensitive fluorescent indicator dye for calcium imaging of neural activity in vitro and in vivo. *European Journal of Neuroscience*, 39(11):1720–1728, 2014.
- [134] A. E. Taylor and W. R. Mann. *Advanced calculus*. John Wiley and Sons, Inc., USA, third edition, 1983.
- [135] L. Theis, P. Berens, E. Froudarakis, J. Reimer, M. R. Rosón, T. Baden, T. Euler, A. S. Tolias, and M. Bethge. Benchmarking spike rate inference in population calcium imaging. *Neuron*, 90(3):471–482, 2016.
- [136] L. Tian, S. A. Hires, T. Mao, D. Huber, M. E. Chiappe, S. H. Chalasani, L. Petreanu, J. Akerboom, S. A. McKinney, E. R. Schreiter, C. I. Bargmann, V. Jayaraman, K. Svoboda, and L. L. Looger. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods*, 6:875–881, 2009.
- [137] Y. A. Tolias, S. M. Panas, and L. H. Tsoukalas. Generalized fuzzy indices for similarity matching. *Fuzzy Sets and Systems*, 120(2):255–270, 2001.

- [138] R. Y. Tsien. New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures. *Biochemistry*, 19(11):2396–2404, 1980.
- [139] R. Tur, Y. C. Eldar, and Z. Friedman. Innovation rate sampling of pulse streams with application to ultrasound imaging. *IEEE Transactions on Signal Processing*, 59(4):1827–1842, 2011.
- [140] M. Unser and T. Blu. Cardinal exponential splines: part I - theory and filtering algorithms. *IEEE Transactions on Signal Processing*, 53(4):1425–1438, 2005.
- [141] J. A. Urigüen, T. Blu, and P. L. Dragotti. FRI sampling with arbitrary kernels. *IEEE Transactions on Signal Processing*, 61(21):5310–5323, 2013.
- [142] J. A. Urigüen Garaizabal. *Exact and approximate Strang-Fix conditions to reconstruct signals with finite rate of innovation from samples taken with arbitrary kernels*. PhD thesis, Imperial College London, 2013.
- [143] I. Valmianski, A. Y. Shih, J. D. Driscoll, D. W. Matthews, Y. Freund, and D. Kleinfeld. Automatic identification of fluorescently labeled brain cells for rapid functional imaging. *Journal of Neurophysiology*, 104(3):1803–1811, 2010.
- [144] M. Van Rossum. A novel spike distance. *Neural Computation*, 13(4):751–763, 2001.
- [145] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.
- [146] I. Vlahinić, E. Andò, G. Viggiani, and J. E. Andrade. Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images. *Granular Matter*, 16(1):9–21, 2014.
- [147] J. T. Vogelstein, A. M. Packer, T. A. Machado, T. Sippy, B. Babadi, R. Yuste, and L. Paninski. Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of Neurophysiology*, 104(6):3691–3704, 2010.
- [148] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynek, and L. Paninski. Spike inference from calcium imaging using sequential monte carlo methods. *Biophysical Journal*, 97(2):636–655, 2009.

-
- [149] C. Vonesch, T. Blu, and M. Unser. Generalized Daubechies wavelet families. *IEEE Transactions on Signal Processing*, 55(9):4415–4429, 2007.
- [150] G. Wang, D. R. Wyskiel, W. Yang, Y. Wang, L. C. Milbern, T. Lalanne, X. Jiang, Y. Shen, Q.-Q. Sun, and J. J. Zhu. An optogenetics- and imaging-assisted simultaneous multiple patch-clamp recording system for decoding complex neural circuits. *Nature Protocols*, 10(3):397–412, 2015.
- [151] B. Weber and F. Helmchen. *Optical Imaging of Neocortical Dynamics*. Neuromethods. Humana Press, 2014.
- [152] J. M. Whittaker. The Fourier theory of the cardinal functions. *Proceedings of the Edinburgh Mathematical Society*, 1:169–176, 1928.
- [153] E. Yaksi and R. W. Friedrich. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved ca2+ imaging. *Nature Methods*, 3:377–383, 2006.
- [154] W. Yang, J. K. Miller, L. Carrillo-Reid, E. Pnevmatikakis, L. Paninski, R. Yuste, and D. S. Peterka. Simultaneous multi-plane imaging of neural circuits. *Neuron*, 89(2):269–284, 2016.
- [155] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.
- [156] C. Zimmer and J. C. Olivo-Marin. Coupled parametric active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1838–1842, 2005.
- [157] H.-J. Zimmermann. Fuzzy set theory. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3):317–332, 2010.
- [158] J. Ziv and M. Zakai. Some lower bounds on signal parameter estimation. *IEEE Transactions on Information Theory*, 15(3):386–391, 1969.
- [159] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells, F. A. Jolesz, and R. Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index: Scientific Reports. *Academic Radiology*, 11(2):178–189, 2004.