Author:
Delaney, Thomas J

Title:
Investigating, implementing, and creating methods for analysing large neuronal ensembles

UNIVERSITY OF BRISTOL

DOCTORAL THESIS

Investigating, implementing, and creating methods for analysing large neuronal ensembles

---

Author:
Thomas J. DELANEY

Supervisors:
Dr. Cian O'DONNELL
Dr. Michael C. ASHBY

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

Biological Intelligence & Machine Learning Unit
Department of Computer Science

August 28, 2020

Word count: 39000

Declaration of Authorship

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

_____

Date:

_____

Acknowledgements

I would like to thank my supervisors, Cian O'Donnell and Mike Ashby, for their help, encouragement, advice, and patience over the last four years. This includes not only helping with research, but also enabling and encouraging me to make the most of my opportunities during that time. Without their help, I would not have grown as much as I have done in those years. I very grateful for their time and effort.

I would also like to thank the members of the Bristol Computational Neuroscience Unit for introducing me to all the various aspects of computer science, neuroscience, and machine learning, of which I otherwise would not have heard. As the first person to introduce me to the concept of mathematical neuroscience during my undergraduate days, and a great source of advice and guidance during my PhD, I would also like to thank Conor Houghton.

Personally, I would to thank my girlfriend Ashley, who has been nothing but helpful since I met her.

Finally, I would like to that my father, mother, brother and sister. I am truly fortunate to have such a good family. I thank them for their love, encouragement, and excellent example.

Abstract

Since the use of multi-electrode recording in neuroscience began, the number neurons being recorded in parallel has been increasing. Recently developed methods using calcium or voltage imaging have also contributed to the growth in neuronal datasets. As datasets grow, the need for new analysis methods also grows. In this research we attempted to address some of the problems associated with reading from large neuronal ensembles using fluorescent calcium indicators, and some of the problems with analysing data read from large neuronal ensembles.

We created a biophysical model for the fluorescence trace produced by a calcium indicator responding to a given spike train. Our model reproduced the characteristics of a real fluorescence trace recognised by spike inference algorithms. This model will be useful for anyone using or considering calcium imaging.

To find order in the correlated behaviour of a large multi-region neuronal ensemble, we applied a novel method from network science to detect structure and communities in correlated behaviour. We investigated the similarities between these communities and their brain anatomy. Our results indicate local correlated networks function at shorter timescales ($<$ 50ms), while multi-region correlated networks function over longer timescales ($>$ 100ms). This result agrees with previous findings from EEG data, but has not been shown before using spiking data.

We developed a statistical model for the number of neurons spiking in a neuronal ensemble based on the Conway-Maxwell-binomial distribution. Our aim was to capture correlated activity in a neuronal population without measuring correlation coefficients directly. The model captured correlated activity at very short timescales better than measuring correlation coefficients. We also replicated one of the findings of Churchland et al. (2010) relating to the quenching of neural variability at stimulus onset. We propose a connection between this result and the changes in association captured by our model.

Contents

List of Figures

List of Tables

List of Abbreviations

COMb	Conway-Maxwell-binomial (distribution)
OASIS	Online active set method to infer spikes
SNR	Signal to noise ratio
NMI	Normalised mutual information
AMI	Adjusted mutual information
VI	Variation of information

List of Symbols

$[Ca^{2+}]$	Free calcium concentration	mol
$[BCa]$	Fluorescent indicator bound calcium	mol
$[ECa]$	Endogenous mobile buffer bound calcium	mol
$[ImCa]$	Immobile mobile buffer bound calcium	mol
$[BCa^*]$	excited fluorescent indicator bound calcium	mol
k_{X_f}	binding (affinity) rate	$\mathrm{mol}^{-1}\,\mathrm{s}^{-1}$
k_{X_b}	unbinding (dissociation) rate	s^{-1}

Chapter 1

Introduction

1.1 Overview

Since Hodgkin and Huxley's squid experiments featuring a single axon (Hodgkin and Huxley, 1939), to more recent research with spike sorted data from $\sim$ 24000 neurons from 34 brain regions from 21 mice (Allen et al., 2019), the number of neurons contributing to electrophysiological datasets has been growing. The number of simultaneously recorded neurons has doubled approximately every seven years since the use of multi-electrode recording in neuroscience began (Stevenson and Kording, 2011). Recording methods using two-photon calcium imaging have also been used to extract data from populations containing over 10000 neurons (Peron et al., 2015). This dramatic growth in the number of neurons available for analysis requires a dramatic change in analysis methods.

There are multiple methods for reading activity from neuronal ensembles: electrophysiology, calcium imaging, and voltage imaging. Electrophysiology involves inserting electrodes into the brain of an animal. The electrodes read extra-cellular membrane potential, and using these readings we observe activity in the ensemble. Calcium imaging and voltage imaging use indicator dyes or fluorescent proteins that emit fluorescence traces that indicate either the concentration of calcium in a neuron's cytoplasm, or the neuron's membrane potential. In this project, we have attempted to address some of the difficulties in collecting data from these large ensembles using fluorescent calcium indicators, and some of the difficulties in analysing the collected data.

The rest of this introductory chapter will give some background about methods of recording from the brain, and some background for the rest of the document. Chapter two describes a biophysical model for the fluorescence trace induced by a given spike train in a cell containing a fluorescent calcium indicator. Our third chapter describes our investigations into

the correlated activity across different regions of a mouse behaving spontaneously. We applied a novel community detection method (Humphries et al., 2019) from network science to correlation based networks of neurons, and observed differences in the structure of these correlations at different timescales. In our fourth chapter, we detail a new statistical model for the number of neurons spiking in a neuronal ensemble at any given moment. With this model, we attempted to capture correlated activity in a new way. The fifth chapter is a brief description of the work that yielded negative results or was abandoned. The final chapter is a discussion of our work and results from the previous chapters and their implications.

1.2 Modelling the fluorescence of calcium indicators

To focus on calcium imaging for a start, a neuron that contains a fluorescent calcium indicator in its cytoplasm will fluoresce when bombarded with photons. The amount that the cell will fluoresce is dependent on the concentration of fluorescent indicator within the cell, and the concentration of calcium within the cell. When a neuron fires an action potential, the influx of free calcium ions causes an increase in fluorescence when those ions bond with the fluorescent indicator and those bounded molecules are bombarded with photons. After the action potential, as calcium is extruded from the cell the fluorescence returns to a baseline level. This is the basic mechanism of fluorescent calcium indicator based imaging.

This method has some advantages over electrophysiology as measure of neuronal ensemble activity. Many of the problems with electrophysiology are within the processes used to isolate spikes in the extracellular voltage readings, and assign these spikes to individual cells. These processes are collectively called ‘spike sorting’. A comparison of many different spike sorting algorithms found that these algorithms only agreed on a fraction of cases (Buccino et al., 2019). Furthermore, because electrodes measure extracellular voltage, neurons that do not spike will not be detected. Isolating individual neurons is easier and more reliable when using calcium imaging data, because cells will emit a baseline level of fluorescence when not firing action potentials. Another advantage is that calcium imaging sites can be re-used for weeks for longitudinal studies (Chen et al., 2013). One of the methods of delivering the fluorescent indicator is by adeno-associated viruses, consequently there can be problems with indicator gradients around the infection site, and expression levels will change in individual cells over weeks (Tian et al., 2009; Chen et al., 2013). This delivery method can also cause cell pathology, and nuclear filling (Zariwala et al., 2012), but these problems can be solved by using lines of transgenic mice (Dana et al., 2014). The fluorescence signal itself can serve a a

good indicator of cell activity, but similarly to electrophysiology, the aim of calcium imaging is often spike detection.

If the imaging data is collected at a high enough frequency, and the signal-to-noise ratio of the fluorescence trace is high enough, it should be possible to infer the spike times to some level of accuracy. For example, the calmodulin based indicator GCaMP6s has a sufficiently high signal-to-noise ratio that isolated action potentials can be detected and inferred (Chen et al., 2013). Many spike inference algorithms exist (Vogelstein et al., 2010; Pnevmatikakis et al., 2016; Friedrich and Paninski, 2016; Pnevmatikakis et al., 2013; Pnevmatikakis et al., 2014; Deneux et al., 2016; Greenberg et al., 2018), and some of these can perform both cell isolation and spike detection simultaneously (Vogelstein et al., 2010; Pnevmatikakis et al., 2016; Pnevmatikakis et al., 2014; Deneux et al., 2016). But the relationship between spiking and fluorescence change is not fully understood. For example, the fluorescent indicator will act like an additional calcium buffer within the cell cytoplasm and will compete with the other endogenous buffers to bind with free calcium ions. Therefore, the concentration of those endogenous buffers, and the binding dynamics of those buffers will have an effect on the change in fluorescence in response to an action potential. Furthermore, the binding dynamics of the fluorescent indicator itself will have an effect on the change in fluorescence. For example, the GCaMP series of fluorescence indicators are based on the calcium buffer protein calmodulin. This protein has four binding sites, whose affinities interact non-linearly. But most of the spike inference algorithms model the fluorescence as a linear function of a calcium trace, and they model this calcium trace as a first or second order autoregression with a pulse input to represent action potentials. Deneux et al. (2016) developed two different calcium fluorescence models behind their spike inference algorithm (MLspike) with a more biological inspiration. For their simpler model, they take a physiological approach and account for baseline calcium indicator dynamics. They end up with a system of first order differential equations defining the dynamics of calcium concentration, baseline fluorescence, and fluorescence. For their more complicated model specifically for genetically encoded calcium indicators, they also took into account indicator binding and unbinding rates, which added another equation to their system of equations. The algorithms that use the autoregression model and the MLspike algorithm are outperformed by the most recently published spike inference algorithm (Greenberg et al., 2018). This algorithm takes into account the binding dynamics of both the endogenous buffers and fluorescent calcium indicator, and the concentrations of free calcium, indicator, and endogenous buffer within the cell cytoplasm. The performance of this algorithm shows that there is value in more biologically inspired

models of fluorescent calcium indicators.

In light of the growing popularity of two-photon calcium imaging, and the lack of biologically inspired spike inference algorithms ((Greenberg et al., 2018) developed their spike inference algorithm in parallel to our work), we decided to develop a biologically inspired model for fluorescent calcium indicator fluorescence. The idea being that our model would take a spike train, or simply spike times, provided by the user, and return the fluorescence trace that would be induced by this spike train or spike times. The model contains parameters for concentrations of indicator and endogenous buffers, as well as affinity and unbinding rates for these buffers. There are also parameters for the baseline concentration of free calcium in the cell cytoplasm, and the cell radius (as a means for calculating the cell volume). With this model, we hoped that experimentalists would be able to test out different calcium indicators on the types of spike trains that they expect to encounter. This way they could decide ahead of time which indicator suited their situation best. Since the output of our model is a fluorescence trace, the spike inference models mentioned above can be applied to the modelled fluorescence. This means that the model could also be used to benchmark the performance of these spike inference algorithms, and to investigate the impact of variations in the model on spike inference accuracy.

1.3 Functional networks

We have outlined some of the advantages that calcium imaging has over electrophysiology. But electrophysiology is more useful in some situations. One particular drawback for two-photon calcium imaging is that usually it can only be used for imaging near to the surface of the brain. This problem can be solved by removing the tissue around the area to be imaged, and custom building a two-photon microscope Dombeck et al., 2010. Imaging with three (or presumably more) photons may solve this problem in the future (Ouzounov et al., 2017). A better option for reading activity from neurons beyond the surface of the brain is to use Neuropixels probes (Jun et al., 2017). These probes can be used to read from thousands of neurons simultaneously in many different areas of the brain (Allen et al., 2019; Stringer et al., 2019; Steinmetz, Carandini, and Harris, 2019; Steinmetz et al., 2019). This brings us to another problem for which we require new innovations in our analysis methods. Specifically, analysing correlated behaviour in neural ensembles consisting of neurons from many different brain regions.

Until the invention of new technologies such as the Neuropixels probes, most electrophysiology datasets read from neurons in only one or two regions. Therefore most of the research on interactions between neurons in different regions is limited to two regions (Wierzynski et al., 2009; Patterson et al., 2014; Girard, Hupé, and Bullier, 2001). In chapters 3 and 4 we used datasets with neurons from 9 and 5 different brain regions respectively. In their review of the interaction between growing the number of neurons in datasets and the analysis methods applied to those dataset, Stevenson and Kording (2011) assert that an important objective of computational neuroscience is to find order in these kinds multi-neuron of datasets. This was our main aim for the research described in chapter 3.

In light of recent findings based on correlated behaviour showing that spontaneous behaviours explain activity in many different parts of the brain that would otherwise be regarded as noise (Stringer et al., 2019), that satiety is represented brain wide (Allen et al., 2019), and that exploratory and non-exploratory states are represented in the amygdala (Gründemann et al., 2019), it was clear that state representation or motor control had an influence on correlated behaviour in areas of the brain not usually associated with these tasks. Also, given differences in timescales of fluctuations in different brain regions (Murray et al., 2014), and different timescales for event representation in different brain regions (Baldassano et al., 2017), we decided to investigate brain wide correlated behaviour at timescales ranging from 5ms up to 3s.

We started off measuring the correlations in spike counts between individual neurons in our ensemble. These measurements induced a weighted undirected graph where each node represented a neuron, and the weight of each edge was the strength of the correlation between the neurons represented by the nodes at either end of that edge. In order to put the neurons into groups with correlated behaviour, we applied a novel community detection algorithm (Humphries et al., 2019) to this graph. We repeated this analysis for timescales from milliseconds to seconds. Bear in mind that our correlation based graph was completely agnostic of the anatomical regions in which our cells resided. We then compared our correlated communities to their anatomy at each timescale. In this way, we used a novel method, never applied neuronal data before, to analyse the makeup of correlated communities across different regions at different timescales.

1.4 A new statistical model for capturing correlated behaviour

Many important findings have been made by measuring the correlations between binned spike counts, but there are some problems with this method of analysis. Firstly, the width of the bins used to bin spike times into spike counts has an effect on the magnitude of the correlations measured. Using a short bin width can cause your measurements to be artificially small (Cohen and Kohn, 2011). This may not be an issue if one is considering relative size of correlations when using the same bin width, but it is still not ideal. Secondly, while pairwise correlations can capture most of the information in a small network (up to 40 cells) of highly correlated cells (Schneidman et al., 2006), a model based on pairwise correlations alone will fail to capture the activity of larger ($\sim$ 100 cells) networks, higher order correlated activity is required (Ganmor, Segev, and Schneidman, 2011). One problem with these higher order correlations is that they may be defined in different ways (Staude, Grün, and Rotter, 2010). Furthermore if we want to include them in a model this usually involves greatly increasing the number of parameters to fit, which increases the dimension of the parameter space leading to the 'curse of dimensionality'. Some models attempt to sidestep these problems while still capturing higher-order correlations. These models attempt to capture the relationship between each individual neuron in the ensemble, and the ensemble as a whole. Okun et al (2015) called the strength of this relationship the 'population coupling', and demonstrated that this quantity can predict an individual neuron's response to optogenetic stimulation of the whole ensemble. They also showed that this quantity was an indicator of the neuron's synaptic connectivity (Okun et al., 2015). With the 'population tracking model', O'Donnell et al. (2016) linked the probability of firing an action potential for each individual neuron with the distribution of the number of active neurons. This allowed model fitting for a large number of neurons, as well as calculation of full pattern probabilities, and population entropy (O'Donnell et al., 2017).

In this work, we also aimed to capture correlated behaviour between the neurons in a neuronal ensemble without measuring correlations directly. Correlation coefficients capture the linear component of the relationship between two random variables, but will not measure any relationship beyond linearity. Also, measuring correlation coefficients using short timebins can be difficult for neuronal data (Cohen and Kohn, 2011). We decided to abandon correlation, and we aimed to quantify a more general concept of association by modelling the number of active neurons in the ensemble using a Conway-Maxwell-binomial (COMb) distribution (Kadane, 2016).

The COMb distribution is a probability distribution over the number of successes in a sequence of Bernoulli trials, where these trials can be associated in some way. The COMb distribution is an extension of the standard binomial distribution, with an additional parameter to model association between the Bernoulli variables. Using this additional parameter the distribution can capture positive association, where the Bernoulli variables tend to take the same value, negative association, where the Bernoulli variables tend to take opposite values, or no association i.e. the standard binomial distribution.

We fit a COMb distribution to spike sorted electrophysiological data taken from five different regions in the brain of an awake mouse exposed to visual stimuli (Steinmetz et al., 2019). We examined whether or not a model based on the COMb distribution was able to capture changes in the number of active neurons in these neuronal ensembles in response to the stimuli. We also investigated the relationship between the changes as captured by the COMb model and the change in neural variability as measured by Churchland et al. in their famous paper (Churchland et al., 2010).

Our overall aim was to investigate some of the challenges in analysing large ensembles of neurons present today. That included collecting the data to analyse (via calcium imaging), and subsequently analysing these data. We felt that this was a worthwhile project because the size of datasets, in terms of numbers of neurons and data collected, is growing rapidly. Consequently these challenges will only become greater unless they are addressed. This is our attempt at addressing them.

Chapter 2

Sensitivity of the spikes-to-fluorescence transform to calcium indicator and neuron properties

Abstract

Fluorescent calcium indicators such as GCaMP are widely used to monitor neuronal activity. However the relationship between the fluorescence signal and the underlying action potential firing is poorly understood. This lack of knowledge makes it difficult for experimenters to decide between different indicator variants for a given application. We addressed this problem by studying a basic biophysical model of calcium dynamics in neuronal soma. We fit the model parameters to publicly available data where GCaMP6s fluorescence and whole-cell electrophysiological recordings were made simultaneously in the same single neurons. We systematically varied the model's parameters to characterise the sensitivity of spike train inference algorithms to the calcium indicator's main biophysical properties: binding rate, dissociation rate, and molecular concentration. This model should have two potential uses: experimental researchers may use it to help them select the optimal indicator for their desired experiment; and computational researchers may use it to generate simulated data to aid design of spike inference algorithms.

2.1 Introduction

Although fluorescent calcium indicators such as GCaMP are widely used to monitor neuronal activity, the relationship between the fluorescence signal and the underlying action potential firing is imperfect. The fluorescence signal has a low signal-to-noise ratio, and most indicators' kinetics are slow relative to the millisecond-timescale dynamics of the membrane voltage (example in figure 2.1A). This makes spike inference difficult. Furthermore, the effects of the indicator and cell properties on the fluorescence signal are unknown. For example, genetically encoded indicators can accumulate within neurons over weeks and months (Chen et al., 2013). Studies using calcium-sensitive fluorescent dyes have shown that indicator concentration has substantial effects on the spike-to-fluorescence relationship (Maravall et al., 2000). Therefore spike rates inferred from GCaMP fluorescence signals may give misleading results if comparing across imaging sessions. More generally, the poor understanding of the spike-to-fluorescence transform means experimenters may not know whether to trust the outputs of spike train inference methods in any given application.

Spike trains are usually inferred from the time series of intensity values of one pixel of the fluorescence image, where the pixel is located at the cell's soma. The problems of identifying these pixels, and inferring spikes from their time series can solved separately or together. When attempting to infer spikes, the fluorescence trace is modelled as a linear combination of calcium concentration dynamics, a baseline calcium concentration, and some Gaussian noise. The calcium concentration dynamics are modelled as an autoregressive process of degree 1 or 2 with a pulse input corresponding to the spike train, or the number of spikes fired in a time step. The model includes no dynamics for the fluorescent indicator itself. Furthermore, in order to make this model into an easily solvable linear programming problem the number of spikes fired in a timestep is not restricted to non-negative integers but to arbitrary non-negative values (Vogelstein et al., 2010; Pnevmatikakis et al., 2016; Friedrich and Paninski, 2016; Pnevmatikakis et al., 2013; Pnevmatikakis et al., 2014). More biologically inspired spike inference models do exist (Deneux et al., 2016), but their fundamentals are very similar. In this work, we investigated the effect of changing dynamics and buffer concentrations on the accuracy of the inference algorithms based on these models.

The aim of this project was to model the fluorescence traces produced by a fluorescent calcium indicator in a neuron soma resulting from a specific spike train, given calcium indicator parameters such as binding rate, dissociation rate, and molecular concentration. Such

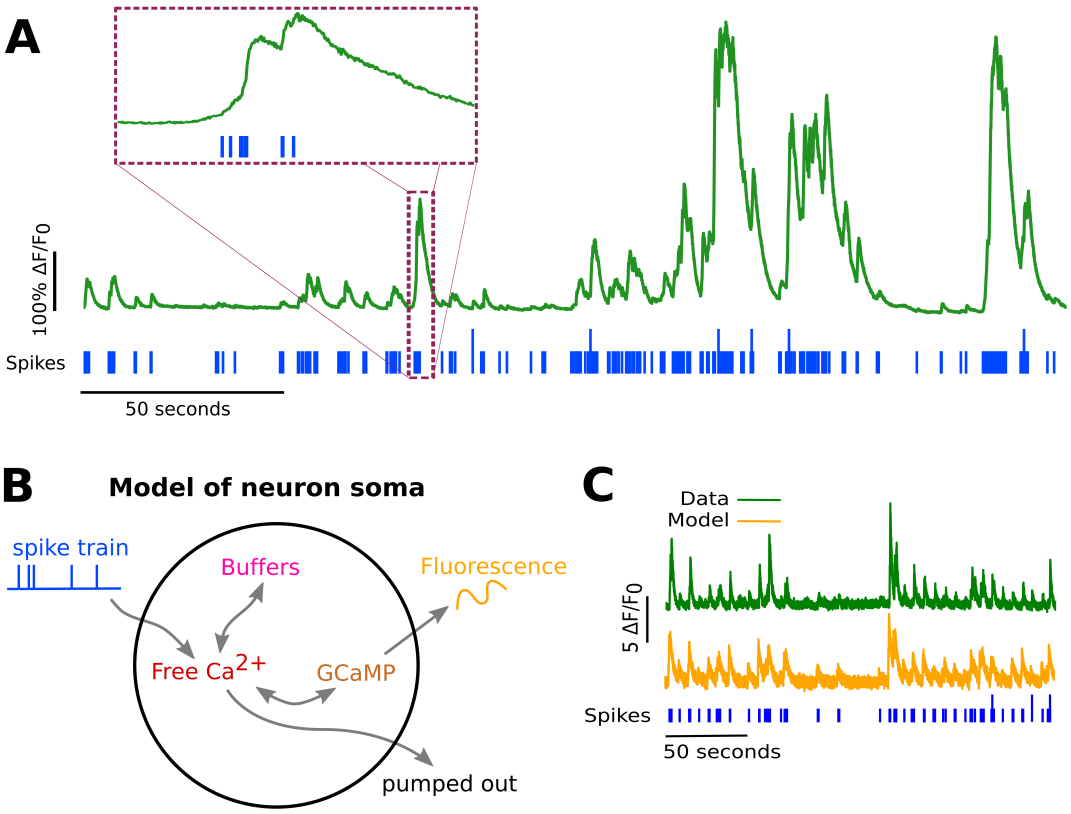


FIGURE 2.1:
A: Example spike train (blue) and the corresponding GCaMP6s fluorescence trace (green), data replotted from (Berens et al., 2018). Inset shows zoomed section of traces to highlight slow decay of GCaMP6s fluorescence relative to spike time intervals.
B: Schematic diagram of the neuron calcium and GCaMP computational model.
C: Good visual match of data fluorescence trace (green) and model simulated fluorescence (orange) in response to an identical spike train (blue).

a model would allow benchmarking of various spike inference algorithms, and enable understanding of how indicator characteristics affect the quality of spike train inference.

The model we developed consisted of free calcium, fluorescent indicator molecules, and mobile and immobile endogenous calcium buffers. The indicator molecules which were bound to a calcium molecule could be either excited, i.e. able to release a photon, or relaxed. In order to reproduce the noise inherent in the data collection, we modelled the release of photons from the excited indicator bound calcium as a stochastic process.

The fluorescence traces produced by the simulation were calibrated to reproduce the signal-to-noise ratio observed in experimental data. Previously published spike inference algorithms were then used to infer spike trains from the experimental fluorescence traces and the modelled fluorescence traces. The parameters of the model were then varied in order to determine the effect on the system dynamics and the effects on spike inference.

2.2 Methods

2.2.1 Calcium dynamics model

We wrote a biophysical model of the calcium dynamics within a cell soma. When a neuron fires an action potential, voltage-dependent calcium ion-channels open up that allow a current of calcium ions (Ca^{2+}) to flow into the neuron (Koch, 1999). The increase in the free calcium ion concentration inside of the cell, along with changes in the concentration of potassium and sodium, causes the change in cell membrane potential, which must be repolarised. The repolarising process consists of free calcium ions leaving the cell through open ion channels, or binding to molecules within the cell called buffers, or calcium storage by organelles such as the endoplasmic reticulum. A diagram illustrating the cell, its channels, and its buffers can be seen in figure 2.1A. There are several different types of calcium buffer, each with different dynamics and different concentrations within different types of excitable cell. The fluorescent calcium indicator is another calcium buffer, with the useful property that when it is bound to a calcium ion, the bound molecule may become excited by a photon and release a photon in return. This is what creates the fluorescence. After the action potential has taken place, the free calcium concentration within the cell will return to a baseline level (Maravall et al., 2000).

We modelled the the dynamics of five molecular concentrations,

- Free calcium ion concentration, $[Ca^{2+}]$

- Fluorescent indicator bound calcium, $[BCa]$
- Endogenous mobile buffer bound calcium, $[ECa]$
- Endogenous immobile buffer bound calcium, $[ImCa]$
- Excited buffered calcium, $[BCa^*]$

The term 'buffering' refers to free calcium ions coming into contact with buffer molecules followed by the binding of those molecules. Diagrammatically:

$$[X][Ca^{2+}] \underset{k_{Xb}}{\overset{k_{Xf}}{\rightleftharpoons}} [XCa]$$

where $[X]$ represents any buffer molecule, and k_{X_f} and k_{X_b} represent the binding (association) and unbinding (dissociation) rates in units of per molar concentration per second (M^{-1} s^{-1}) and per second (s^{-1}) respectively. The speed of this chemical reaction is determined by the binding and unbinding rates.

There are a number different endogenous buffers in any neuron. Which buffers are present, and the buffers' concentrations vary from cell to cell. In order to capture the effects of different kinds of buffers without modelling dozens of different individual buffers, we modelled two different kinds of buffer only. These 'mobile' and 'immobile' buffers were designed to be aggregations of the effects of multiple different buffers into two effective buffers with different concentrations and binding rates. (Bartol et al., 2015) also divide calcium buffers into mobile and immobile varieties. Note that since the model has no spatial component, the mobile and immobile buffers only differ in their binding and unbinding rates.

The fluorescent calcium indicator behaves similarly to the other calcium buffers. The calcium is buffered by the indicator in the same way. But an indicator bound calcium molecule can become excited by absorbing the energy from a photon. An excited indicator bound calcium molecule can then release a photon to go back to its 'relaxed' state.

$$[B][Ca^{2+}] \underset{k_{Bb}}{\overset{k_{Bf}}{\rightleftharpoons}} [BCa] \leftrightsquigarrow [BCa^*]$$

The released photons are captured by a photon collector. This gives us the fluorescence trace.

The system of equations we used to model all of these interactions is as follows:

$$\begin{aligned}\frac{d[Ca^{2+}]}{dt} =& k_{Bb}[BCa] + k_{Eb}[ECa] + k_{Imb}[ImCa] \\ &- k_{Bf}[B][Ca^{2+}] - k_{Ef}[E][Ca^{2+}] - k_{Imf}[Im][Ca^{2+}] \\ &+ \beta([Ca_0^{2+}] - [Ca^{2+}])\end{aligned} \tag{2.1}$$

$$\frac{d[BCa]}{dt} = k_{Bf}[B][Ca^{2+}] - k_{Bb}[BCa] + r[BCa^*] - \eta[BCa] \tag{2.2}$$

$$\frac{d[ECa]}{dt} = k_{Ef}[E][Ca^{2+}] - k_{Eb}[ECa] \tag{2.3}$$

$$\frac{d[ImCa]}{dt} = k_{Imf}[Im][Ca^{2+}] - k_{Imb}[ImCa] \tag{2.4}$$

$$\frac{d[BCa^*]}{dt} = \eta[BCa] - r[BCa^*] \tag{2.5}$$

where $[Ca_0^{2+}]$ is the baseline calcium concentration within the cell soma, β is a rate defining how quickly free calcium enters or leaves the cell in the absence of an action potential, η is the excitation rate for indicator bound calcium, r is the photon release rate for the excited indicator bound calcium, and f and b are used to indicate the forward and backward rates for chemical reactions respectively. The excitation rate defines the proportion of indicator bound calcium that becomes excited at each time step. The photon release rate defines the proportion of excited indicator bound calcium that releases a photon and returns to its relaxed state at each time step. An action potential is modelled as a discontinuous increase in the free calcium concentration to an appropriate value (Maravall et al., 2000).

Note that each of the three pairs of binding and unbinding terms in the first equation has a corresponding pair in one of the subsequent three equations. Binding removes a free calcium molecule and adds a bound calcium molecule, and unbinding does the opposite.

When using this model to simulate a fluorescence trace, the system of equations above are first solved over a period of 25s without action potentials. This lets each of the five tracked chemical concentrations reach their steady state. Then we use the given spike train and the parameters to model the fluorescence trace.

Photon release & capture

We used a simple model for the photon release. The number of photons released at each time step was controlled by the number of excited indicator bound calcium molecules in the cell

and a parameter called the 'release rate'. The release rate is an optimised free parameter of the model.

As for the photon capture, in two-photon excitation microscopy the photons scattered by the fluorescent indicator get scattered in all directions. Therefore the number of photons detected is stochastic. This made the process for capturing photons the natural source of noise in the system. The number of photons captured, and therefore the intensity of the fluorescence, is modelled using a binomial distribution. The number of photons released was used as the number of trials. The probability of success, or 'capture rate' was a free parameter of the model that we optimised.

2.2.2 Parameter optimisation

The free parameters of the model are as follows:

Calcium rate, β Controls how quickly the concentration of free calcium will be driven to the baseline concentration.

Capture rate, p The average proportion of photons captured by the photon detector.

Excitation rate, η The number of indicator bound calcium molecules that become excited by photon bombardment at each time step.

Release rate, r The number of excited indicator bound calcium molecules that release a photon at each time step.

To optimise the free parameters given a fluorescence trace, we applied the following procedure:

1. The frequency power spectrum of the trace was measured.
2. The power spectrum was smoothed using a boxcar smoother (aka. sliding average, box smoother).
3. The log of the smoothed power spectrum was measured.
4. Use the model to create a modelled fluorescence trace.
5. Apply steps 1, 2, and 3 to the modelled fluorescence trace.
6. Calculate the root mean squared difference between the log power of the actual fluorescence trace, and the log power of the modelled fluorescence trace.

7. Calculate the root mean squared difference between the actual fluorescence trace and the modelled fluorescence trace.

8. Use an optimisation algorithm to reapply this process, attempting to minimize the sum of the two root mean squared differences at each iteration.

Using the root mean squared difference of the log power spectra as part of the objective function forces the model to match the noise frequency of the actual fluorescence. Using the root mean squared difference of the traces themselves forces the model to match the amplitude of the fluorescence trace more accurately. Using both of these terms as part of our objective function was designed to make our model match the change in $\Delta F/F_0$ in response to an action potential as well as the signal-to-noise ratio of the the actual fluorescence trace. We weighted both of these terms equally.

In order to minimise the objective function, a suite of meta-heuristic optimisation (aka. black-box optimisation) algorithms were implemented on each of the traces in the dataset. These methods were chosen because they don't require a gradient for the objective function (gradient-free) and they are particularly useful for minimising stochastic objective functions like the one we used here. The free parameters were optimised for each individual fluorescence trace. The most successful method for each trace was recorded. The method that was most often successful was probabilistic descent, and the second most successful method was generating set search. Both of these methods are examples of pattern search. These two methods were the best optimisers on about 75% of the traces in the dataset. The other methods were differential evolution (with and without radius limited sampling, adaptive and not adaptive), natural evolution strategy, and random search for comparison.

Although this optimisation procedure minimises the value of the optimisation function, the value never reaches zero for a number of reasons. Firstly, the fluorescence traces carry low frequency fluctuations that cannot be captured by the model. Secondly, the model assumes that the process of calcium binding to the fluorescent indicator is linear in time (see equation 1), but there are more complicated dynamics involved here. Fluorescent calcium indicators, the GCaMP series for example, are often built upon the calcium binding protein called 'calmodulin'. This protein has four calcium binding sites. These sites are locally split into two pairs. Each pair has a different affinity for calcium, and the affinity of the binding sites is affected by the occupancy of the other binding sites (Kilhoffer et al., 1992). So the calcium to calcium indicator binding process is non-linear, but the model does not take this into account.

Parameter	Description	Value	Source
baseline	The baseline concentration of free calcium within the cell soma	4.5×10^{-8}M	(Maravall et al., 2000)
cell radius	The radius of the soma (assumed to be spherical)	10^{-5}M	(Fiala and Harris, 1999)
endogenous	The concentration of endogenous mobile buffer within the cell soma	10^{-4}M	(Faas et al., 2011)
frequency	The frequency at which the spike trains are sampled.	100Hz	
immobile	The concentration of endogenous immobile buffer within the cell soma	7.87×10^{-5}M	(Bartol et al., 2015)
indicator	The concentration of fluorescent indicator within the cell soma	10^{-4}M	(Maravall et al., 2000)
k_{Bb}	The unbinding rate of the fluorescent calcium indicator	$160s^{-1}$	(Bartol et al., 2015)
k_{Bf}	The binding rate of the fluorescent calcium indicator	$7.77 \times 108\text{s}^{-1}\text{M}^{-1}$	(Bartol et al., 2015)
k_{Eb}	The unbinding rate of the endogenous mobile buffer	10^4s^{-1}	(Bartol et al., 2015)
k_{ef}	The binding rate of the endogenous mobile buffer	$10^8\text{s}^{-1}\text{M}^{-1}$	(Bartol et al., 2015)
k_{Imb}	The unbinding rate of the endogenous immobile buffer	524s^{-1}	(Bartol et al., 2015)
k_{Imf}	The binding rate of the endogenous immobile buffer	$2.47 \times 10^8\text{s}^{-1}\text{M}^{-1}$	(Bartol et al., 2015)
peak	The increase in free calcium concentration within the cell induced by an action potential	2.9×10^{-7}M	(Maravall et al., 2000)

TABLE 2.1: **Fixed parameters** A table of the parameters fixed before optimising the model. The values of these parameters could be changed to model different fluorescent calcium indicators.

Fixed parameters

As well as the optimised parameters mentioned in section 2.2.2, the model also has thirteen fixed parameters. Please see table 2.1 for details of these parameters and their values. In an application of the model, these parameters can be changed in order to model any given fluorescent calcium indicator, or even prospective indicators that only exist in theory.

2.2.3 Julia

The programming language used to write and execute the model was 'Julia'. Julia is a dynamic programming language designed for technical computing. Julia was designed specifically to provide a convenient high-level dynamic language similar to MATLAB, or Python, with improved performance. Julia's type system and Julia's direct interfaces with C and

Fortran allow this aim to be achieved (Bezanson et al., 2012). The Julia version of the 'Sundials' package for ODE solving was used to solve the system of equations above. The `BlackBoxOptim.jl` package for Julia was used to perform the optimisation.

2.2.4 Spike inference

We used spike inference algorithms to compare the quality of spike inference using the modelled traces to the quality of spike inference using the observed traces. We also used the spike inference algorithms to assess the effect of parameter perturbation on the spike inference. Three algorithms were used:

Constrained non-negative matrix deconvolution algorithm (aka CNMD algorithm) The underlying model models the fluorescence as a linear combination of a calcium trace with additional noise. This calcium trace is a first order autoregression with a pulse input to represent action potentials. This algorithm uses a constrained version of non-negative Weiner deconvolution to infer a calcium signal and a 'spiking activity signal' from the fluorescence trace (Vogelstein et al., 2010; Pnevmatikakis et al., 2016). The spiking activity signal is a non-negative vector of real numbers reflecting the cell's activity rather than an actual spike train.

The underlying model of the fluoresence trace used by this algorithm requires 5 parameters that are calculated from the data: the standard deviation of the white noise component of the trace, the decay time constant, the order of the autoregressive model, baseline calcium concentration (or equivalently, baseline fluoresence), and the initial calcium concentration (or equivalently, initial fluoresence).

The standard deviation of the noise was estimated by calculating the power spectral density of the fluorescence trace, then taking the exponent of the mean of the log of the density across the frequency range between 0.25 and 0.5 times the Nyquist rate for the fluorescence trace. The time constant of the fluorescence trace was estimated from the autocovariance of the denoised fluorescence trace. If a stable autoregressive model could not be found at the default order for the autoregression ($p = 2$), the order was increased by one, and the decay time constant was re-estimated. If a value was not provided for the baseline concentration and initial concentration they were both assumed to be 0. We did not provide values for these parameters.

We inferred a spike train by choosing an optimised threshold for the spiking activity signal. Whenever the spiking activity signal exceeded that threshold, an action potential was inferred. The threshold was optimised by minimising the difference between the number of spikes in the ground truth spike train and the number of spikes predicted. This kind of calibration gives the algorithm an advantage when predicting the correct number of spikes, but not the timing of those spikes. Because of this, comparisons in performance between different spike inference algorithms are not valid. However, comparisons between the performance of this algorithm applied to different fluoresence traces are still valid.

MLSpike algorithm Deneux et al. (2016) developed two different calcium fluorescence models behind their spike inference algorithm (MLspike) with a more biological inspiration. For their simpler model, they take a physiological approach and account for baseline calcium indicator dynamics. They end up with a system of first order differential equations defining the dynamics of calcium concentration, baseline fluorescence, and fluorescence. For their more complicated model specifically for genetically encoded calcium indicators, they also took into account indicator binding and unbinding rates, which added another equation to their system of equations. This algorithm uses a generalised version of the Viterbi algorithm to return the spike train that maximises the likelihood of producing the given fluorescence trace. The Viterbi algorithm is an algorithm for estimating the most likely sequence of hidden states resulting in a sequence of observed states in a discrete-time finite-state Markov process (Forney, 1973). In this case, each hidden state is defined by the presence or absence of an action potential, and each observed state is the value of the fluorescence trace at each time step. (Deneux et al., 2016).

We used the autocalibration algorithm provided with this spike inference algorithm to estimate values for the standard deviation of the white noise in the fluoresence trace, the increase in fluoresence in response to an action potential, and the decay time constant of the fluoresence trace. The standard deviation of the noise in the trace was estimated as the mean of power spectral density of the trace in the frequency range between 3 and 20Hz. The authors of this algorithm believed that the frequencies lower than this range will contain information about the calcium dynamics, while the frequencies above this range contain correlated noise and are therefore unsuitable for estimating white noise.

The increase in fluoresence in response to an action potential and the decay time constant were estimated in parallel. Firstly, a 'calcium events' trace was extracted from the fluorescence trace using a modified version of the spike inference algorithm with fixed values for the increase in calcium in response to a calcium event and decay time constant. The temporally isolated calcium events with moderate amplitudes ($\Delta F/F \leq .25$) were extracted, and the other events were removed. The idea here was to extract a trace of the calcium concentration produced by isolated groups of action potentials. The event amplitude and the time decay constant were then estimated from this trace of 'good' events. Next, a histogram of the event amplitudes was constructed. An estimate of the number of spikes in each event was estimated from the peaks of this histogram. A number of spikes was then assigned to each calcium event based on this estimate. Finally, the autocalibrated values for the change in fluoresence and the decay time constant were calculated by fitting to the modified calcium event trace.

Many values that may be influential in the autocalibration process were hard-coded by the authors of the algorithm. Particularly, defining 'moderate' calcium events as those that produce $\Delta F/F \leq .25$ may be suitable for indicators with lower signal-to-noise ratios than GCaMP6s. This could be the reason for this algorithm's poor performance in our application (see Results section 2.3.2).

The poor performance in comparison to the other two algorithms could also be down to the other algorithms being calibrated to match the number of spikes in the ground truth spike trains.

Online Active Set method to Infer Spikes (OASIS) This algorithm is once again based on an auto-regressive model of the fluorescence trace, but can be generalised to any order. Both the first and second order versions can be fit to a spike train in a reasonable time. The algorithm itself is a generalisation of the pool adjacent violators algorithm (PAVA) that is used in isotonic regression. The OASIS algorithm works through the fluorescence trace from beginning to end, this combined with the speed of the algorithm means that it could be used for real-time online spike inference (Friedrich and Paninski, 2016).

This algorithm required the standard deviation of the noise component of the fluorescence trace and the fluorescence trace decay time constant to be estimated. The standard deviation of the noise was estimated by calculating the power spectral density of the fluorescence trace, then taking the mean of the density across the frequency

range between 0.25 and 0.5 times the Nyquist rate for the fluorescence trace. These frequencies were judged to be most likely to contain uncorrelated noise, rather than signal or correlated noise, by the developers of the algorithm. The time constant of the fluorescence trace was estimated from the autocovariance of the denoised fluorescence trace. In practice, we used the `estimate_parameters` function provided by the authors of this algorithm to estimate these parameters.

This algorithm also required one more paramter, a constant that indicated the minimal non-zero activity within a bin in order for a spike to be assinged to that bin. When applying the algorithm, we used 10 different values for this parameter, and counted the total number of spikes inferred by the algorithm each time. We then chose the value that gave the smallest difference between the total number of inferred spikes and the total number of actual spikes. Calibrating the algorithm in this way gives the algorithm an advantage for inferring the number spikes, but not for inferring the timing of the spikes. This kind of calibration makes the results of the different algorithms incomparable, but the results from the same algorithm applied to different fluorescence traces are still comparable.

In order to quantify the quality of spike inference for a given algorithm, we ran that algorithm on all of the fluorescence traces in dataset number eight of the spike finder datasets. These datasets contained fluorescence traces from neurons containing a fluorescent calcium indicator (either OGB-1 or GCaMP6s) and spike trains from those cells simultaneously recorded using loose-patch electrophysiological recordings. This provided us with a *ground-truth* for spike inference from the fluorescence traces. Then we measured some binary classification measures on the results. These measures included

- Accuracy
- True positive rate (aka recall, sensitivity, hit rate)
- True negative rate (aka specificity)
- Precision
- Negative predicted value
- False negative rate (aka miss rate)
- False positive rate (aka fall-out)

- False discovery rate
- False omission rate

In making these measurements, we allowed a tolerance of two subsequent time bins for spike prediction. For example, the spike train data is a vector of 0s and 1s, with one element for each time bin, and time bin being 10ms. A '0' denotes inactivity, a '1' denotes the presence of at least one action potential. The inferred spike trains produced by the spike inference algorithms take the same form. In our analysis, if a spike appeared in the inferred spike train up to two time frames after a spike in the observed spike train, that spike was considered correctly inferred i.e. a true positive. However, once a spike in the inferred spike train was matched to a spike from the observed spike train, the inferred spike could not be matched to another observed spike. To illustrate, if two spikes were inferred in the two time bins following an isolated observed spike, the first inferred spike was considered correctly inferred, but the second inferred spike was considered incorrectly inferred, i.e. a false positive.

The most useful measure was the true positive rate. This is because the spiking is sparse and this measurement is sensitive to the number of spikes observed and inferred, but is not affected by the true negative or false negative rates. After optimising the parameters for each fluorescence trace we measured the spike inference quality for the observed fluorescence traces, and compared this to the spike inference quality for the modelled traces.

When measuring the spike inference quality for higher frequency spike train ($1 - 10$Hz), we used the accuracy as our binary classification measure. At these frequencies the variance of the fluorescence trace was much higher than for sparser spiking regimes, therefore we wanted to take into account the number of false negatives inferred by the algorithm.

Comparing spike inference quality

In order to compare spike inference quality we had to use methods for comparing samples. When comparing the true positive rate distributions arising from two different datasets, or two different algorithms on the same dataset, we compared the distributions using a paired t-test.

2.2.5 Perturbation analysis

In order to measure the sensitivity of spike inference to changes in a given model parameter, we perturbed the parameter and compared the quality of spike inference with the perturbed

parameters to the quality of spike inference with the experimental or optimised parameters. In order to maximise the possibility of observing a difference due to the perturbation, we perturbed the chosen parameter by a relatively large amount. For example, the experimental value for the molar concentration of the fluorescent indicator within the cell was 10^{-4}M (Maravall et al., 2000). The perturbed values used for this parameter were 10^{-2}M, 10^{-3}M, 10^{-5}M, and 10^{-6}M. The quality of the inference was compared by measuring the true positive rate for each perturbed value and using a t-test to compare the distributions of the results.

2.2.6 Signal-to-noise ratio

To assess the effect of perturbation on the modelled traces, we measured and compared the signal to noise ratio (SNR) on each of the modelled traces. We calculated the SNR as the peak change in fluorescence divided by the standard deviation of the baseline fluctuation of the fluorescence trace (Tada et al., 2014). We measured these values by running the model on a spike train consisting a long period of inactivity followed by one action potential. We ran the model on this spike train one hundred times. We then measured the mean change in fluorescence and standard deviation of baseline activity across the one hundred modelled fluorescence traces, and calculated the SNR.

2.2.7 Data sources

All of the data used in this project was sourced from the 'Spike Finder' project (spikefinder.codeneuro.org). The data consisted of a collection of datasets with simultaneously measured fluorescence traces and action potentials (Berens et al., 2018).

2.3 Results

2.3.1 A biophysical computational model can generate accurate fluorescence traces from spike trains

To study the relationship between action potential firing and calcium fluorescence, we built a computational model of calcium dynamics in a neuronal soma. The model consisted of four dynamic variables: the concentration of free calcium, two types of endogenous buffer, and the calcium-sensitive fluorescent indicator. Each of the buffers and the indicator could

independently bind and unbind with calcium. These reactions were modelled as

$$[X][Ca^{2+}] \underset{k_{Xb}}{\overset{k_{Xf}}{\rightleftharpoons}} [XCa]$$

where X is the buffer concentration and Ca^{2+} is the calcium concentration. Each species could therefore exist in two states: either bound with calcium or unbound. To model the imaging process, we also added a third, excited state to the indicator. When in the calcium-bound state, the indicator could be converted to an excited state, corresponding to the absorption of a photon. The rate of this excitation process could be interpreted as the intensity of the light illuminating the sample. Once excited, the species decayed back to the unexcited state at a fixed rate, corresponding to the spontaneous emission of photons. The total emitted fluorescence signal was interpreted as proportional to this de-excitation flux. To represent experimental noise in the photon capture process, we drew a random number of captured photons at each time step from a binomial distribution, parameterised by a number p that corresponds to the mean fraction of released photons that are captured.

The model had 17 parameters in total describing the molecules' concentrations and reaction rates (Methods). We set 13 of these parameters to values from the literature. The remaining 4 parameter values we fit to publicly-available data (Berens et al., 2018), briefly explained as follows (see Methods for full details). Single neurons from acute rat cortical slices expressing GCaMP6s were imaged with two-photon microscopy while the membrane potentials of the somata of the same neurons were simultaneously recorded via whole-cell patch clamp electrophysiology. In this dataset, the electrical recordings give unambiguous information about neurons' spike times. To do the parameter fitting, we feed these spike trains as inputs to the computational model. After running, the model returns a simulated fluorescence trace. We aimed to find the model parameter values that give the best match between this simulated fluorescence trace and the real fluorescence time series recorded in the corresponding neuron. To do this we used a suite of optimisation procedures to jointly fit both the real neuron's fluorescence time series and power spectrum, which capture complementary information about the spikes-to-fluorescence mapping (Methods). We performed the fitting procedure independently for each of the 20 neurons in the spikefinder dataset (http://spikefinder.org). After fitting, the model produced realistic-looking fluorescence time series (Figure 2.1).

Given that fluorescence traces are often modelling using a linear combination of a first degree autoregressive process and white noise, it could be argued that a four parameter model

contains some redundancy in the parameter space. But, our aim was to create a biophysical model that could be useful for interpreting the role of cell dynamics behind the production of the fluorescence trace. Because of that we felt that our redundancy was justified.

2.3.2 Spike inference algorithms perform similarly on real data compared with time series simulated from the model

Researchers often pass the fluorescence time series through a spike inference tool before performing further statistical analyses. These spike inference algorithms take the fluorescence trace as input and attempt to estimate the neuronal spike train that triggered them (Vogelstein et al., 2010; Pnevmatikakis et al., 2016; Friedrich and Paninski, 2016; Pnevmatikakis et al., 2013; Pnevmatikakis et al., 2014; Deneux et al., 2016). Part of our motivation for building this model was to allow us to investigate which properties of the cell and the calcium indicator affect the quality of spike inference? In order to trust the conclusions from our model, we should first be confident that spike inference from our simulated fluorescence traces is similar to that from the real data. To test this we passed each of the simulated fluorescence traces through three previously published spike inference algorithms, quantified their performance against the ground-truth electrophysiology data, repeated the procedure for the real calcium fluorescence time series, and compared the accuracy of the inference processes in all cases. The *true positive rate*, also known as the *recall*, the *sensitivity*, or the *probability of detection* of spike inference varied across the three inference algorithms we tried (p value and statistical test here). The constrained non-negative matrix deconvolution algorithm (Pnevmatikakis et al., 2016) (CNMD algorithm) correctly detected approximately 45% of the true spikes, the OASIS algorithm (Friedrich and Paninski, 2016) correctly detected approximately 35% of the true spikes, and the ML spike algorithm (Deneux et al., 2016) correctly detected approximately 15% of the true spikes (see figure 2.2). Notably, for two of the three inference algorithms, the quality of inference was also fairly consistent for individual spike trains, not just the group means ($p > 0.05$, paired t-test). This demonstrates that the models were generating fluorescence time series that were similarly difficult to decode as the real data, in ways that were not specific to any one inference algorithm. This is evidence that the models captured real aspects of the spikes-to-fluorescence transform.

2.3.3 Relative effects of various buffers to the fluorescence signal

One of the benefits of computational models over laboratory experiments is that we can observe all the variables in the simulation to gain insight into the system's dynamics, which

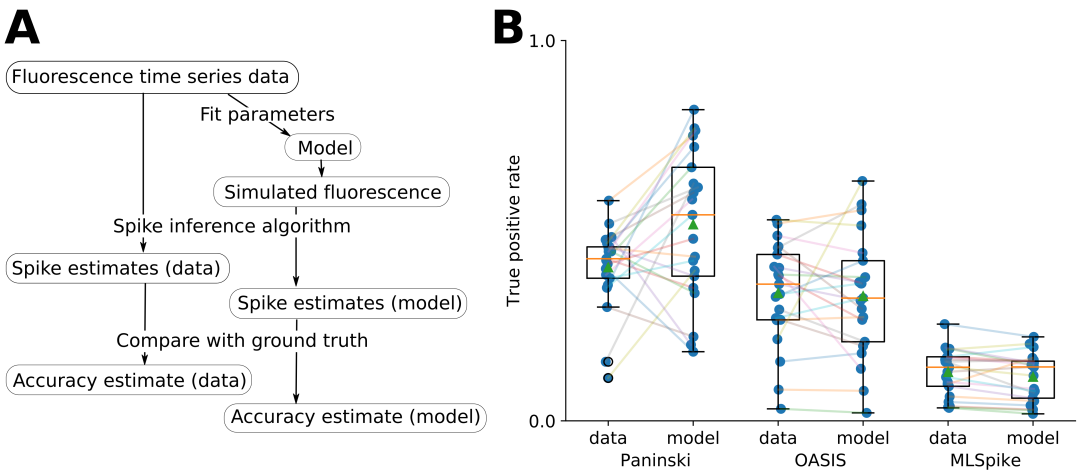


FIGURE 2.2:
A: Workflow to compare spike inference for real versus simulated fluorescence data.
B: True positive rates achieved by three different spike inference algorithms when applied to observed spike trains, and simulated spike trains. Data points overlaid as blue circles. The performance is similar from real and simulated data for each of the algorithms.

quantities

can be difficult to do in the lab. We plotted the concentrations of the various species over time for a version of the model fit to one data set, in response to the same train of spikes used for fitting (figure 2.3). Figure 2.3a shows the absolute values of the species concentrations, summed. Consistent with experimental estimates (Maravall et al., 2000), only a small fraction ($\sim$ 0.1%) of calcium is free and unbound to any buffer. Of the bound calcium, the vast majority, ($\sim$ 96%) is bound to the GCaMP indicator. The two types of endogenous buffer are bound to the remaining calcium ($\sim$ 4%). An influx of calcium from a single spike adds very little to the total calcium, in relative terms (red line in Figure 3a).

When calcium entered the model neuron it was rapidly buffered (Bartol et al., 2015). However the relative fractions of which buffer molecules bound to the influxed calcium was dynamic, and changed over time. Figure 2.3 (b-f) shows the time course of the various species over time in response to a calcium influx event from a single action potential. Crucially, the indicator $[BCa]$ competed with the endogenous buffers $[ImCa]$ and $[ECa]$ – all three bind calcium on similar timescales. This implies that the timecourse and amplitude of the $[BCa]$ variable will also depend on the binding rates and availabilities of the endogenous buffers. For example if we decreased the concentration of an endogenous buffer, we might expect both a faster rise time and greater peak amplitude of the $[BCa]$ signal in response to a calcium influx event. The slowest component of the decay had a similar time constant for $[BCa]$, $[ImCa]$ and $[ECa]$, which in turn matched the $[Ca]$ extrusion time constant in our

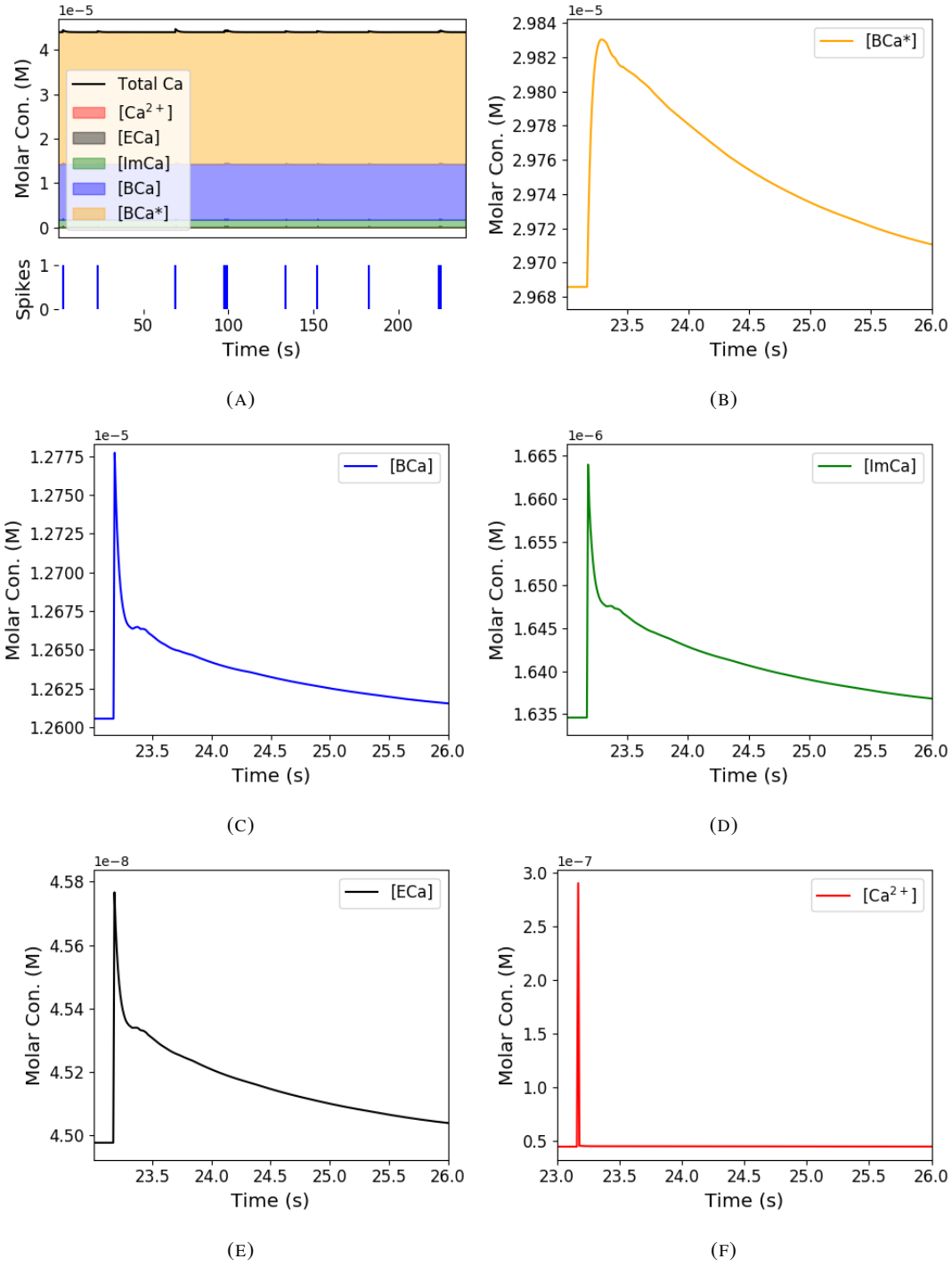


FIGURE 2.3: **Calcium Buffering Dynamics** (A) The proportions of bound and free calcium concentrations within a cell, with the associated spike train. (B)-(F) The dynamics of the concentration of (B) excited indicator bound calcium, (C) indicator bound calcium, (D) immobile endogenous buffer bound calcium, (E) mobile endogenous buffer bound calcium, and (F) free calcium in response to an action potential at ∼23.2s.

model ($\sim 6.29 \times 10^{-22}\text{Ms}^{-1}$). This implies that the buffers and the indicator had reached a dynamic equilibrium and were jointly tracking the free calcium concentration as calcium was slowly extruded from the cell.

Interestingly the excited bound calcium species ($[BCa^*]$) showed a qualitatively different timecourse in response to a calcium influx event. This concentration is subject to the added 'excitation and release' dynamic, where a certain proportion of the concentration absorbs the energy from an incoming photon and goes into an 'excited state' at each time step. A certain proportion of the concentration releases a photon and reverts to a 'relaxed state' at each timestep also. This means that the excited bound calcium lags behind the bound calcium trace. We could think of the excited bound calcium trace as a low pass filtered version of the bound calcium trace.

2.3.4 Spike inference accuracy is sensitive to indicator properties, and likely varies within and between cells

The above results imply that the fluorescence signal depends on the relative properties of both GCaMP and the endogenous buffers. We next used the model to directly ask how sensitive spike inference was to these components. We focused on three key parameters that likely vary from cell to cell and experiment to experiment: GCaMP binding kinetics, GCaMP concentration, and endogenous buffer concentration.

Several variants of GCaMP itself have been made that differ in calcium binding kinetics, baseline fluorescence, fluorescence efficiency, and other factors. For example, GCaMP6f has a decay time constant of $\sim$ 1s, while GCaMP6s has a decay time constant of $\sim$ 2s (Chen et al., 2013). Here we asked how these differences in binding kinetics affect spike inference. We jointly varied the calcium binding and unbinding rates of the indicator by the same factor over a range from 100-fold slower to 100-fold faster from the fitted values, and simulated the fluorescence response for each of the parameter settings in response to the same spike trains as before (figure 2.4). Notably this manipulation does not affect the indicators affinity, and therefore would not affect steady-state responses to prolonged changes in calcium. Instead it is likely to affect its sensitivity to the spike train dynamics. We computed two summary measures from the simulated fluorescence traces: the signal-to-noise ratio for a single spike (Methods, section 2.2.6), and the accuracy of spike inference for each of the spike trains. We observed a reduction in the signal-to-noise ratio and the spike inference quality when we set the binding and unbinding rates were set to one hundredth of their fitted values, and to one tenth of their fitted values. When we increased the value of both binding rates, we observed

no change in these measurements. The reduction in both rates lead to smaller increases in fluorescence in response to an action potential and a longer decay time (figure 2.4a), this caused the reduction in signal-to-noise ratio. As both rates were increased, the change in $\Delta F/F_0$ in response to an action potential increased and the decay time decreased slightly, but the fluorescence trace created by these values was very similar to the trace created by the fitted values.

Second, the overall concentrations of GCaMP often varies from cell to cell. For example different cells, even of the same type in the same tissue, can express different levels of GCaMP, due to proximity to the infection site, or the cell becoming 'nuclear-filled' (Tian et al., 2009; Chen et al., 2013). Also, GCaMP is often used for longitudinal experiments where the same cells are re-imaged across multiple days or weeks. However since GCaMP expression typically ramps up over time (Chen et al., 2013), the accuracy of spike inference may differ across multiple longitudinal recordings in the same cell. We addressed this by varying the concentration of calcium indicator in the model, simulating spike trains and measuring signal-to-noise ratio and spike inference accuracy on the resulting fluorescence traces. Both increasing and decreasing the concentration of the indicator had effects on the fluorescence trace, signal-to-noise ratio, and spike inference. The signal-to-noise ratio and spike inference quality decreased with decreased indicator concentration, and both showed a decrease when the indicator concentration was increased to 100 times its fitted value (figure 2.5). The signal-to-noise ratio showed an increase when the indicator concentration was increased to 10 times its fitted value, but there was no corresponding change in the spike inference quality. The decrease in indicator concentration caused a reduction in the increase in $\Delta F/F_0$ in response to an action potential, and an increase in the decay time of this increase (figure 2.5a). The increase in indicator concentration had the opposite effect, it casued an increase in the change in $\Delta F/F_0$ in response to an action potential, and a decrease in the decay time.

Third, the concentration and types of endogenous calcium buffers also vary from neuron to neuron, both within and between cell types (Bartol et al., 2015; Maravall et al., 2000; Neher and Augustine, 1992). Since the calcium buffer capacity of neurons is high, around 50-70 (Lee et al., 2000) in excitatory hippocampal pyramidal cells, around 100-250 (Lee et al., 2000) in inhibitory hippocampal pyramidal cells, and 900-200 in Purkinje cells (depending on the age of the subject), these endogenous buffers compete with GCaMP for binding to calcium, and variations in endogenous buffer concentration may affect GCaMP signal and therefore spike inference. To address this we varied the concentration of the endogenous buffer in the model neuron over five orders of magnitude from 0.8 to 8000 μM, simulated

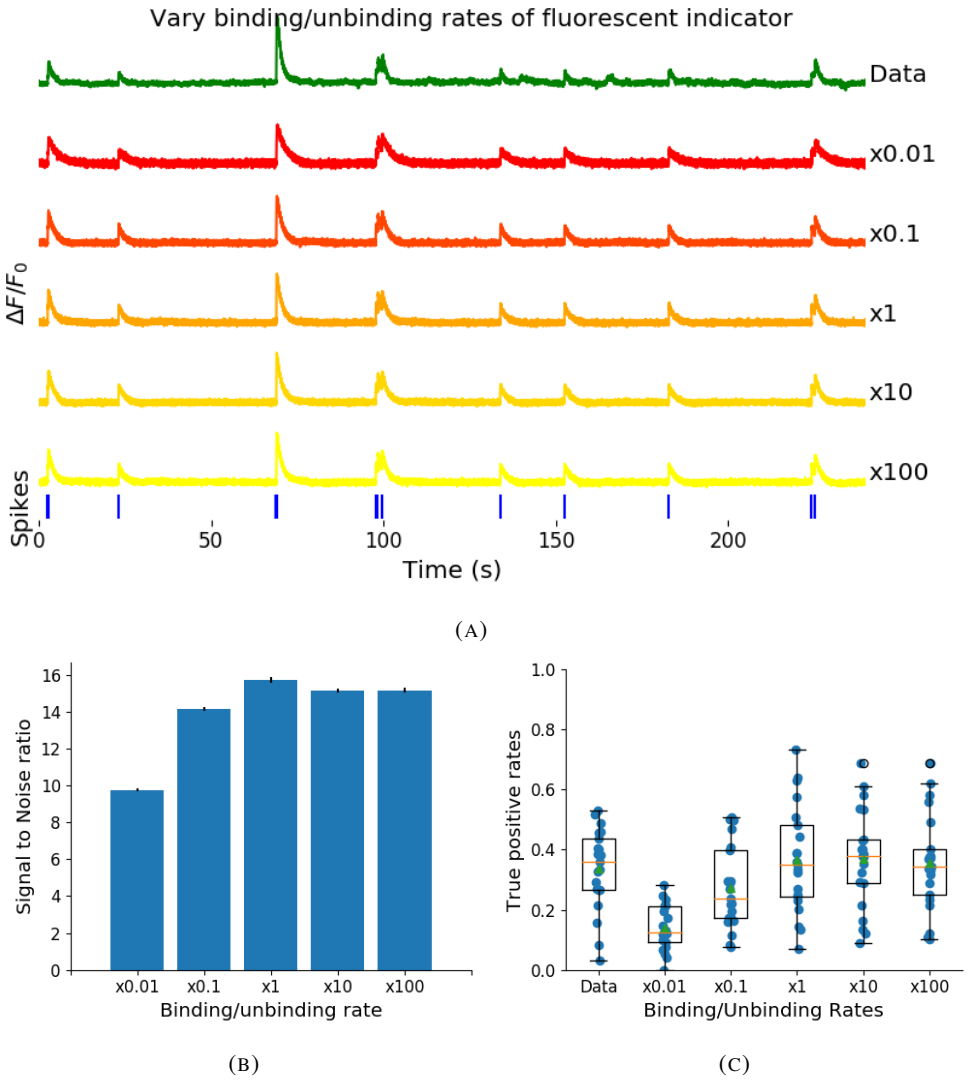


FIGURE 2.4: (A) An example trace for each of the five pairs of values used for the binding and unbinding rates of the fluorescent calcium indicator. (B) The signal-to-noise ratio of the modelled fluorescence traces using each of the four perturbed value pairs, and the experimental value. The SNRs for the value pairs perturbed downward are lower than that for the unperturbed value pair or the higher value pairs. (C) The true-positive rates of the deconvolution algorithm's predictions when inferring from the observed data, and inferring from modelled traces using the perturbed and experimental values. We used the OASIS algorithm for spike inference here. The results from the other spike inference methods were similar, with their true positive rates scales similarly to figure 2.2 B.

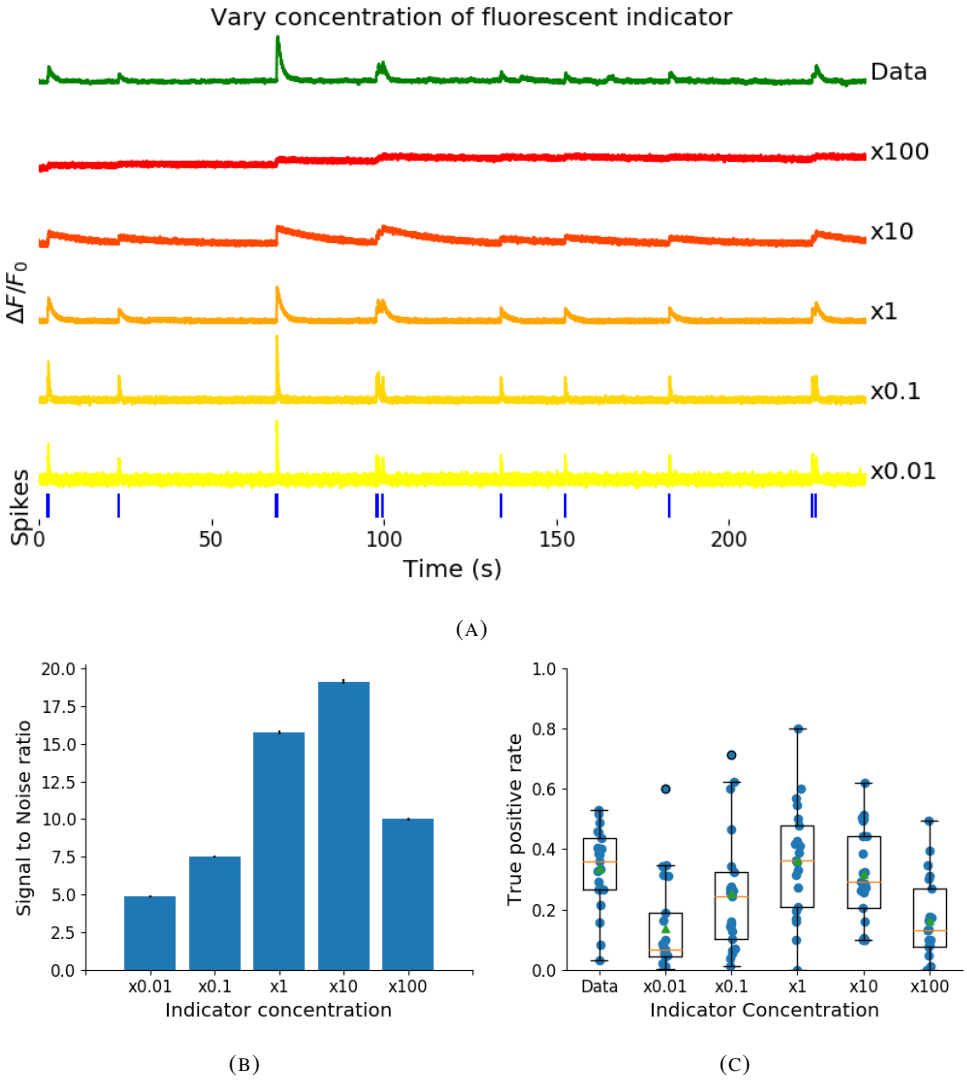


FIGURE 2.5: (A) An example trace for each of the five perturbed values for the concentration of fluorescent calcium indicator. The top two traces are produced by the lower perturbed values, the middle trace is produced by the experimental value, and the lowest two traces are produced when using the higher perturbed values. (B) The signal-to-noise ratio of the modelled fluorescence traces using each of the four perturbed values, and the experimental value. Extreme perturbations of the concentration either above or below the experimental level lowered the SNR. (C) The true-positive rates of the deconvolution algorithm's predictions when inferring from the observed data, and inferring from modelled traces using the perturbed and experimental values. We found that the algorithms performs equally badly on the two most extreme values, and performs equally well on the experimental value, and the next higher perturbed value. We used the OASIS algorithm for spike inference here. The results from the other spike inference methods were similar, with their true positive rates scales similarly to figure 2.2 B.

calcium fluorescence traces in response to the same set of spike trains, and performed spike inference on the resulting fluorescence time series. Increasing the endogenous buffer concentration had a substantial effect on the GCaMP fluorescence signal, both decreasing its amplitude and slowing its kinetics (figure 2.6(a)). This corresponded with a decrease in both single-spike signal-to-noise ratio (figure 2.6(b)) and spike inference accuracy (figure 2.6(c)). In contrast, decreasing endogenous buffer capacity from the fitted value had little effect on either the GCaMP signal or spike inference (figure 2.6).

2.3.5 Single spike inference accuracy drops for high firing rates, but firing rate itself can be estimated from mean fluorescence amplitude

The fluorescence signal recorded from neurons using calcium indicators is typically much slower than changes in membrane potential for two reasons: first, because the calcium and the indicator have slow binding and unbinding kinetics, the signal is a low-pass filtered version of the membrane potential. Second, neuronal two-photon imaging experiments are often performed in scanning mode, which limits their frame rate to $\sim$ 10Hz or slower. This implies that multiple spike events that occur close in time might be difficult to resolve from a calcium indicator time series. Many cells, especially several types of inhibitory interneurons, fire tonically at rates higher than 10Hz. We used the model to test whether spike inference accuracy depended on the neuron's firing frequency by driving the cell with spike trains sampled from a Poisson processes of varying frequency. We simulated a variable firing rate using an Ornstein-Uhlenbeck process, and simulated the spike trains using a Poisson distribution with its rate taken from this process. Because of the high frequency firing rate of these spike trains, we using the accuracy as the measure of spike inference quality. We simulated 30 spike trains at average firing rate of 1, 5, and 10Hz, and measured the spike inference quality of all these traces. Spike inference accuracy decreased with increasing firing rate, for up to 10Hz Poisson spike trains (figure 2.8(left)). Although the accuracy remained above 90% for each of the three frequencies. We also plotted the average $\Delta F/F_0$ as a function of stimulation firing rate. We found that it increased monotonically as a function of firing rate (figure 2.8(right)).

We expected lower spike inference quality as the average spiking frequency increased. Since the fluorescence trace, in some sense, is a low pass filtered version of the spike train, a tightly packed groups of spikes will be more difficult to infer than isolated spikes. However, the increasing amplitude of the fluorescence trace with increasing frequency suggests that some spike inference algorithm could be developed based on this amplitude.

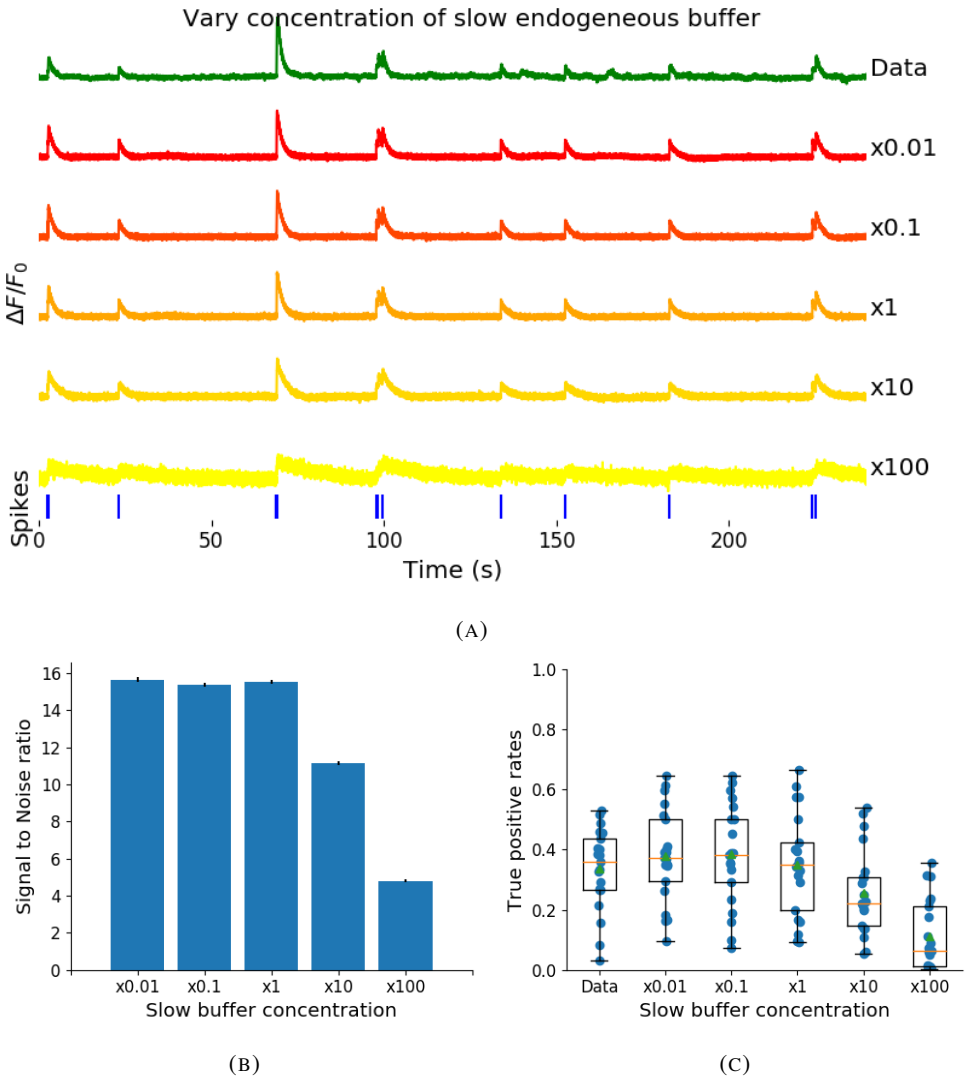


FIGURE 2.6: (A) An example trace for each of the five perturbed values for the concentration of immobile endogenous buffer. (B) The signal-to-noise ratio of the modelled fluorescence traces using each of the four perturbed values, and the experimental value. The lower values for the immobile buffer produce the same SNR as the experimental value. But the higher perturbed values produce fluorescence traces with a lower SNR. (C) The true-positive rates of the deconvolution algorithm's predictions when inferring from the observed data, and inferring from modelled traces using the perturbed and experimental values. We used the OASIS algorithm for spike inference here. The results from the other spike inference methods were similar, with their true positive rates scales similarly to figure 2.2 B.

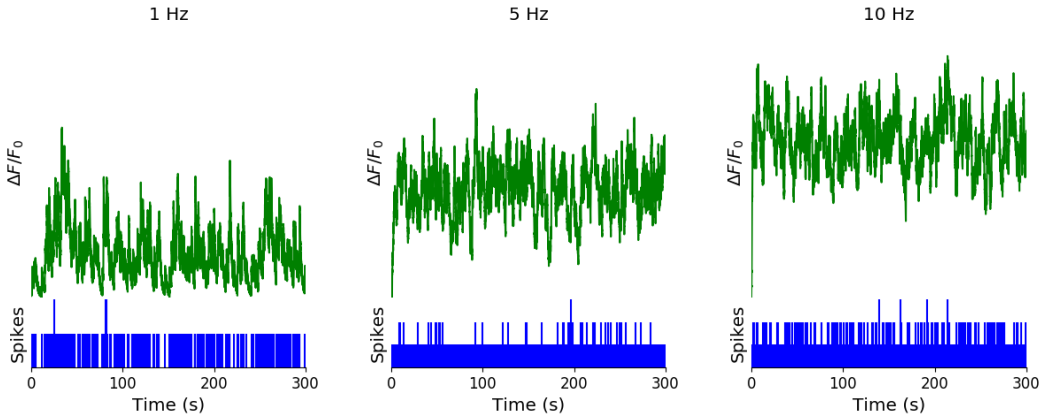


FIGURE 2.7: **Simulating fluorescence traces at different firing rates** Example modelled traces created using simulated spike trains with a mean firing rate of 1Hz (left column), 5Hz (middle column), and 10Hz (right column). Note the difference in amplitude with different mean firing rates.

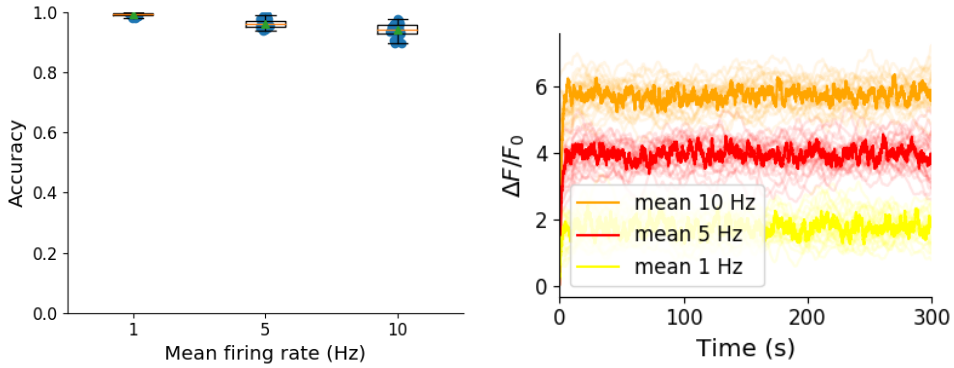


FIGURE 2.8: **Inference quality and $\Delta F/F_0$ vs Firing rate** (left) The spike inference accuracy when applied to 30 traces created using simulated spike trains with mean firing rates of 1, 5, and 10 Hz. (right) The mean $\Delta F/F_0$ across those 30 traces for each frequency.

2.4 Discussion

We designed a biophysical model for the changes in free calcium and bound calcium concentrations within a cell soma with a fluorescent calcium indicator. We used this model to model the fluorescence trace resulting from a spike train in this cell. We fit the free parameters of the model by matching the power spectrum and amplitude of fluorescence traces with simultaneously measured spike trains. We inferred spikes from real fluorescence traces and modelled fluorescence traces, and measured the quality of the spike inference in both cases. We found that the spike inference quality was similar in both cases. We perturbed the concentration of the calcium buffers in the model, and the binding/unbinding rates of those buffers in the model, and measured the effect on the signal-to-noise ratio (SNR) of the modelled fluorescence traces and the spike inference quality.

For the fluorescent calcium indicator, we found that any large perturbation away from the value taken from the literature led to a reduction in SNR, and spike inference quality. For the binding/unbinding rates, we kept the ratio of these rates constant, but altered their values in parallel. The lower values caused a reduction in SNR, and a reduction in spike inference quality. For the endogenous buffer concentration, an increase above the experimental value caused a reduction in SNR and spike inference quality.

We perturbed the concentration of the indicator, the binding/unbinding rates, and the endogenous buffer concentration to values 100 times smaller and 100 times larger than than the value taken from the literature. Given that the indicator concentration can be controlled, at least to some extent, we thought it was worthwhile to simulate these extreme perturbations. For the binding/unbinding rates, such extreme values in these rates are unlikely for fluorescent calcium indicators or endogenous buffers. But these extreme perturbations are still useful for studying what kind of fluorescence trace an indicator with such extreme binding/unbinding rates produce. Similarly, these extreme values in endogenous buffer are also unlikely to occur in neurons. But, using these values in our model allows us to analyse the interplay between the endogeneous buffer concentration and the fluorescence trace produced.

Although the model produced visually similar time series to the real data, there were a few aspects it did not capture. First, the real data featured some low-frequency components that did not appear related to the spike events. These were not captured by the models we used in this study, but could be added in future by adding a suitable low-frequency term to the resulting time series. Second, the real data seemed to have some non-linearities not captured in the model, for example the response to two nearby spikes was greater than expected from the linear sum of two single spikes. This may be due to the co-operative binding of calmodulin to calcium, which gives calmodulin a supra-linear sensitivity to calcium concentration (Faas et al., 2011). The non-linear dynamics of this binding have been included in a recently developed spike inference model (Greenberg et al., 2018). Our model, in contrast, behaved much more linearly but could be extended in future to include such non-linearities. Third, in the real data the fluorescence peak amplitude seemed to vary from spike to spike, even for well-isolated spike events. Recent research has shown that calcium influx due to a single action potential was quite variable in pyramidal cells, and that this variability had a effect on spike inference (Éltes et al., 2019). However in our model we assumed each spike leads to the same fixed-amplitude injection of calcium to the cell, leading to much greater regularity in fluorescence peak amplitudes. This variability could be added in future versions of the model by making the injected calcium peak a random variable. Fourth, we modelled

the soma as a single compartment, but in reality there is likely a non-uniform spatial profile of calcium concentration. This may matter because some endogenous buffers might access calcium right as it influxes from the extracellular space, whereas the majority of the fluorescence signal is more likely coming from the bulk of the cytoplasm. Future models could attempt to model these spatial dependencies to assess whether they affect the overall spike inference procedure.

The concentration of free calcium ions in the neuron cytoplasm enables calcium signalling, which has a vital role in neuronal energy metabolism, and neurotransmission in neurons (Brini et al., 2014). Our model allowed us to examine to which calcium buffer the incoming free calcium ions, due to an action potential, bound. We found that around 95% of the calcium ions bound to the fluorescent calcium indicator. This suggests that the introduction of the fluorescent calcium indicator has a dramatic effect on the free calcium concentration within the cell cytoplasm. This could have a downstream effect on the functionality of the cell.

As well as the optimised parameters, the model has 13 fixed parameters than can be changed to simulate different types of calcium indicators. This model could be used to test the theoretical performance of proposed new types of calcium indicator. The model could also be used by developers of spike inference algorithms to test the effects of changing calcium indicator parameters on spike inference, or to test the affects of changing spiking characteristics on spike inference. For example, high firing rate vs low firing rate, or bursting vs no bursting. Given the increasing amplitude of the fluorescence trace with increasing mean firing rate, it would be possible to build a spike inference algorithm on this principle at least in part.

Our model has already been used as a tool by our colleagues, for simulating fluorescence traces in response to cells that can fire with a continuous rate between 10 and 20Hz, but do not always do so. Our colleagues found that a combination of the amplitude and the variance of the simulated fluorescence trace was the best indicator of firing rate. For example, when a cell was not firing, the amplitude and variance of the fluorescence trace was relatively low. When the cell fired with a low firing rate $\sim$ 1Hz, the mean amplitude was still low but the variance of the fluorescence trace was high, and for high firing rate $10 - 20$Hz, the fluorescence amplitude was high, and the variance was low. In this way, our model may be useful for investigating firing rates underlying real fluorescence traces in response to cells which can fire in these rage ranges.

A recent paper by Greenberg et al (2018) described a biophysical model for spike train

inference called the 'Sequential binding model'. Their model for spike inference was similar to our model for fluorescence traces in that their model included parameters for two types of endogenous buffer. But this model also included dynamics for calcium binding to and unbinding from these endogenous buffers. Furthermore, this model included dynamics for calcium binding to and unbinding from the four binding sites present on a GCaMP6s molecule. In the accuracy measurements specified in that paper, this model performed better than the MLspike algorithm, which is also partially a biophysically model, and it performed better than the constrained non-negative deconvolution algorithm. The sequential binding model also has biophysically interpretable parameters, and its fitted parameters for quantities such as buffering capacity and calcium influx upon action potential firing fall in line with experimental values (Greenberg et al., 2018). Biophysical models like this appear to be the way forward for spike inference algorithms, and would make a good complimentary tool to our fluorescence model.

Chapter 3

Functional networks expand across anatomical boundaries as correlation time-scale increases

Abstract

Decades of research has established that correlated spiking plays a crucial role in representing sensory information. One drawback associated with the recent improvement in recording technology and consequent large datasets is the difficulty in analysing higher order correlations in large neuronal ensembles. One benefit of these datasets that has not yet been explored is the opportunity to compare correlations within anatomical regions to correlations across anatomical regions. In this work, we measured correlations between neurons residing in nine different brains regions in three awake and behaving mice. Using the these correlation measurements, we created weighted undirected graph networks and applied network science methods to detect functional communities in our neural ensembles. We compared these functional communities to their anatomical distribution. We repeated the analysis, using different timescales for our correlation measurements, and found that functional communities were more likely to be dominated by neurons from a single brain region at shorter timescales ($<$ 100ms).

3.1 Introduction

Decades of research has established that correlations play a crucial role in representing sensory information. For example, the onset of visual attention has been shown to have a greater affect on the correlations in the macaque V4 region than on the firing rates in that region (Cohen and Maunsell, 2009). Recent findings show that spontaneous behaviours explain correlations in parts of the brain not associated with motor control (Stringer et al., 2019), that satiety state appears to have a brain wide representation (Allen et al., 2019), and that subject exploratory and non-exploratory states are represented in the amygdala (Gründemann et al., 2019). So, behavioural states are likely represented across many regions of the brain, not just motor related areas. In order to understand the brain, we must understand the interactions between neurons and regions.

Because of limitations in recording technology almost all research has explored correlations between neurons within a given brain region, or within only two regions at most (Wierzynski et al., 2009; Patterson et al., 2014; Girard, Hupé, and Bullier, 2001). Relatively little is known about correlations between neurons in many different brain regions. However, the recent development of 'Neuropixels' probes (Jun et al., 2017) has allowed extracellular voltage measurements to be collected from multiple brain regions simultaneously routinely, and in much larger numbers than traditional methods. In this project we used a publicly-available Neuropixels dataset to analyse correlations between different brain regions (Stringer et al., 2019).

A drawback associated with the improvement in recording technology is an increase in the difficulty in analysing these data. For example, analysing the ith order interactions of N neurons generally requires estimation of N^i parameters. A number that becomes astronomical for large N. New methods are required for analysing these new large datasets. We attempted to address this requirement in this piece of research by applying a cutting-edge network science community detection method to neural data.

Another unexplored area of research is the changes in cell interactions at different timescales. Studies have shown different timescales for fluctuations in spiking activity (Murray et al., 2014), and different time scales for event representation (Baldassano et al., 2017) across different brain regions. Still most studies focus on quantifying interactions at a given timescale. But neurons may interact differently, or may interact with different neurons at different timescales. Here we explore correlated communities of neurons at different timescales.

In this work, we measured correlations between binned spike counts from neurons from

nine different regions of the mouse brain. These measurements induced a weighted undirected graph or network where each neuron is represented by a node, and the strength of the connection between these nodes/neurons is the strength of the correlation between their spike counts. We then applied newly invented network methods (Humphries et al., 2019) to this network to find any community structure, and place the neurons in these correlation based communities. Finally, we compared these functional communities to the anatomical membership of the neurons.

To investigate the functional communities and their relationship with anatomy at different time scales, we repeated these analyses using different length bin widths when binning spike times.

To find and analyse functional networks while controlling for the subject's behaviour, we conditioned the binned spike counts on data from a video of the subject's face, and repeated our analysis for spike count correlations (or noise correlations) and signal correlations.

3.2 Data

The data that we used in this project were collected by Nick Steinmetz and his lab members (Stringer et al., 2019; Steinmetz et al., 2019).

3.2.1 Brain regions

Neuropixels probes were used to collect extracellular recordings (Jun et al., 2017) from three different mice. The mice were awake, headfixed, and engaging in spontaneous behaviour. The mice were of different sexes and different ages. One mouse was 'wild-type', the others were mutants. Details as follows:

1. male, wild type, P73.
2. female, TetO-GCaMP6s, Camk2a-tTa, P113
3. male, Ai32, Pvalb-Cre, P99

Eight probes were used to collect readings from 2296, 2668, and 1462 cells respectively. Data were collected from nine brain regions in each mouse:

- Caudate Putamen (CP)
- Frontal Motor Cortex (Frmoctx)
- Hippocampal formation (Hpf)

- Lateral Septum (Ls)
- Midbrain (Mb)
- Superior Colliculus (Sc)
- Somatomotor cortex (Sommoctx)
- Thalamus (Th)
- Primary visual cortex (V1)

Readings were continuous and lasted for about 1 hour (Stringer et al., 2019; Steinmetz et al., 2019). Locations of each of the probes can be seen in figure 3.1.

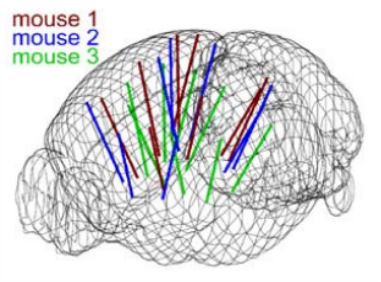


FIGURE 3.1: **Probe Locations:** The locations of the probes in each of the three mouse brains (Stringer et al., 2019).

3.2.2 Video recordings

Video recordings of the mouse's face were taken during the spontaneous behaviour. We had access to the top 500 principal components and top 500 eigenvectors of the processed videos. The frequency of recording was slightly less than 40Hz. Each frame contained 327×561 pixels (Stringer et al., 2019; Steinmetz, Carandini, and Harris, 2019). These principal components were used as behavioural data. We controlled for these components when taking measurements conditioned on behaviour.

3.3 Methods

3.3.1 Binning data

We transformed the spike timing data into binned spike count data by dividing the experimental period into time bins and counting the spikes fired by each cell within the time period covered by each of those bins. The data were divided into time bins of various widths ranging from 0.005s to 4s.

If the total length of the recording period was not an integer multiple of the time bin width, we cut off the remaining time at the end of the recording period. This period was at most 3.99s. This is far less than the total recording time of around 1 hour. So, this detail would not affect our results.

3.3.2 Correlation coefficients

We calculated Pearson's correlation coefficient for pairs of spike counts from pairs of neurons. For jointly distributed random variables X and Y, Pearson's correlation coefficient is defined as:

$$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{3.1}$$

$$= \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{3.2}$$

where E denotes the expected value, μ denotes the mean, and σ denotes the standard deviation. The correlation coefficient is a normalised measure of the covariance. It can take values between 1 (completely correlated) and -1 (completely anti-correlated). Two independent variables will have a correlation coefficient of 0, but having 0 correlation does not imply independence.

If we do not know the means and standard deviations required for equation 3.1, but we have samples from X and Y, Pearson's sample correlation coefficient is defined as:

$$r_{XY} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{3.3}$$

where $\{(x_i, y_i)\}$ for $i \in \{1,\ldots,n\}$ are the paired samples from X and Y, and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ are the sample means.

In practice we used the Python function `scipy.stats.pearsonr` to calculate the correlation coefficients.

Total correlations, r_{SC}

In this context, we defined the total correlation (r_{SC}) of two cells to be the correlation between the spike counts of those cells across the entire period of spontaneous behaviour.

Shuffled total correlations

We measured the shuffled total correlations between two neurons by randomly permuting one of the neuron's spike counts and measuring the total correlations. These shuffled correlations were useful when measuring the effect of time bin width on correlations, and when deciding which correlations should be preserved when creating correlation networks (see section 3.3.5).

Separating Correlations & Anti-correlations

In order to compare the effect of bin width on measures of negative r_{SC} (anti-correlation) and positive r_{SC} separately, we had to separate correlated and anti-correlated pairs. To do this, we simply measured the mean r_{SC}, taking the mean across all the bin widths. If this quantity was positive or zero we regarded the pair as positively correlated. If this quantity was negative we regarded the pair as anti-correlated.

3.3.3 Conditioning on behavioural data

Our behavioural data consisted of the top 500 principal components (PCs) of a processed video recording of the mouse's face (see section 3.2.2). Denoting the spike count of a given cell by X, and the PCs by $Z_1, \ldots, Z_{500}$, we wanted to model X as a function of $Z_1, \ldots, Z_{500}$ in order to estimate

$$E[X|Z_1, \ldots, Z_{500}] = \int_{x \in X} xP(X = x|Z_1, \ldots, Z_{500})dx \tag{3.4}$$

$$= \int_{x \in X} x\frac{P(X = x, Z_1, \ldots, Z_{500})}{P(Z_1, \ldots, Z_{500})}dx \tag{3.5}$$

Given the 500 components, a naïve estimation of $P(Z_1, \ldots, Z_{500})$ or $P(X, Z_1, \ldots, Z_{500})$ by histogramming was impossible. Therefore we modelled X as a linear combination of the PCs.

Linear regression

We modelled the spike count of a given cell, X, as a linear combination of the PCs of the video of the mouse's face, $\mathbf{Z} = Z_1, \ldots, Z_{500}$. We tried three different types of regularization

- $L1$ or 'Lasso'
- $L2$ or 'Ridge regression'

- 'Elastic net' regularisation (a linear combination of both $L1$ and $L2$ regularisation penalties)

The elastic net regularisation performed the best, so we stuck with that.

Elastic net regularisation

Suppose we wish to model n observations of a random variable X, $\mathbf{x} = (x_1, \cdots, x_n)$ using n instances of m predictors $\mathbf{Z} = (Z_1, \cdots, Z_m)$. The naïve elastic net criterion is

$$L(\lambda_1, \lambda_2, \boldsymbol{\beta}) = |\mathbf{x} - \mathbf{Z}\boldsymbol{\beta}|^2 + \lambda_2|\boldsymbol{\beta}|_2 + \lambda_1|\boldsymbol{\beta}|_1 \tag{3.6}$$

where $\boldsymbol{\beta}$ is the vector of linear coefficients and

$$|\boldsymbol{\beta}|_2 = \sum_{j=1}^{m} \beta_j^2 \tag{3.7}$$

$$|\boldsymbol{\beta}|_1 = \sum_{j=1}^{m} |\beta_j| \tag{3.8}$$

The naïve elastic net estimator $\hat{\boldsymbol{\beta}}$ is the minimiser of the system of equations 3.6 (Zou and Hastie, 2005)

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} L(\lambda_1, \lambda_2, \boldsymbol{\beta}) \tag{3.9}$$

We implemented the model using the `ElasticNetCV` method of Python's `sklearn.linear_models` package. We chose to put equal weighting on the $L1$ and $L2$ regression parts of equation 3.6. We used 10-fold cross validation to set an optimised value for $\lambda_1 = \lambda_2$.

As well as using the PCs, we also tried fitting the models using the raw video data reconstructed from the PCs and eigenvectors. These models performed worse than those using the PCs. We expected this because each representation contains the same amount of information, but the raw video representation spreads this information across many more components. This requires more parameter fitting, but given the same information.

Conditional covariance

We calculated the expected value of the conditional covariance using the law of total covariance.

$$\text{cov}(X, Y) = E[\text{cov}(X, Y|Z)] + \text{cov}(E[X|Z], E[Y|Z]) \tag{3.10}$$

where these expected values are calculated with respect to the distribution of Z as a random variable.

The law of total covariance breaks the covariance into two components. The first component $E[\text{cov}(X, Y|Z)]$ is the expected value, under the distribution of Z, of the conditional covariance $\text{cov}(X, Y|Z)$. This covariance could be interpreted as the unnormalised version of what Cohen et al. (2011) call the spike count correlation (Cohen and Kohn, 2011), aka. the noise correlation. In particular, this is the covariance of the spike counts in response to repeated presentation of identical stimuli.

The second component is analogous to what Cohn et al. (2011) call the *signal correlation* (Cohen and Kohn, 2011). In particular, $\text{cov}(E[X|Z], E[Y|Z])$ is the covariance between spike counts in response to different stimuli.

Our linear model gave us 500 coefficients, one for each of $Z_1, \ldots, Z_500$. By summing the linear combination of these coefficients and a set containing one value for each Z_i, we obtained our model's estimate for the spike count of the cell represented by X. We interpreted our model as a function that takes a set of values $\{Z_1 = z_1, \ldots, Z_{500} = z_{500}\}$ as input and returns $E[X|Z_1 = z_1, \ldots, Z_{500} = z_{500}]$.

Using our linear model, we calculated $E[X|Z_1, \ldots, Z_{500}]$ for each cell X and for all sets of values for $\{Z_1, \ldots, Z_{500}\}$ available to us. We used those values to calculate $cov(E[X|Z_1, \ldots, Z_{500}], E[Y|Z_1, \ldots, Z_{500}])$ for each pair of cells (X, Y). Then we proceeded to calculate

$$\begin{aligned} E[\text{cov}(X, Y|Z_1, \ldots, Z_{500})] = &\text{cov}(X, Y) - \\ &\text{cov}(E[X|Z_1, \ldots, Z_{500}], E[Y|Z_1, \ldots, Z_{500}]) \end{aligned} \tag{3.11}$$

Measures of conditional correlation

As a measure of expected correlation, we measured the 'event conditional correlation' (Maugis, 2014)

$$\rho_{XY|Z} = \frac{E[\mathrm{cov}(X,Y|Z)]}{\sqrt{E[\mathrm{var}(X|Z)]E[\mathrm{var}(Y|Z)]}} \tag{3.12}$$

Although this is not an actual correlation, it is an intuitive analogue to the correlation as a normalised version of the covariance.

For comparison, we also measured the 'signal correlation'

$$\rho_{\mathrm{signal}} = \frac{\mathrm{cov}(E[X|Z],E[Y|Z])}{\sqrt{\mathrm{var}(E[X|Z])\mathrm{var}(E[Y|Z])}} \tag{3.13}$$

this is an actual correlation.

3.3.4 Information Theory

We used an information theory based measure to measure the difference between the communities that we detected in the correlation based functional networks that we constructed and the anatomical division of the cells in our data. We treated these as clusterings, and measured the distance between them. We also were planning on using the mutual information between the spike counts of cells as measure upon which to build functional networks. But our measurements turned out to be heavily biased. So we abandoned that approach.

As a result, we have here a lot of background on information theory. This information is still useful for understanding our measure of distance between between clusterings. So I think it is worth keeping.

Entropy $H(X)$

The entropy of a random variable X, with outcomes $x_1,\ldots,x_N$, and corresponding probabilities $p_1,\ldots,p_N$ is defined as

$$H(X) = -\sum_{n=1}^{N} p_n \log_2 p_n \tag{3.14}$$

This quantity is also known as the information entropy or the 'surprise'. It measures the amount of uncertainty in a random variable. For example, a variable with a probability of 1 for one outcome, and 0 for all other outcomes will have 0 bits entropy, because it contains no

uncertainty. But a variable with a uniform distribution will have maximal entropy as it is the least predictable. This quantity is analogous to the entropy of a physical system (Shannon, 1948). Note that any base may be used for the logarithm in equation 3.14, but using base 2 means that the quantity will be measured in 'bits'.

The joint entropy of two jointly distributed random variables X and Y, where Y has outcomes $y_1, \ldots, y_M$, is defined as

$$H(X,Y) = -\sum_{n=1}^{N}\sum_{m=1}^{M} P(X = x_n, Y = y_m) \log_2 P(X = x_n, Y = y_m) \tag{3.15}$$

If X and Y are independent then $H(X,Y) = H(X) + H(Y)$. Otherwise $H(X,Y) < H(X) + H(Y)$. When X and Y are completely dependent and the mapping from X to Y is one-to-one, $H(X,Y) = H(X) = H(Y)$.

The conditional entropy of Y conditioned on X is defined as

$$H(Y|X) = -\sum_{n=1}^{N}\sum_{m=1}^{M} P(X = x_n, Y = y_m) \log_2 \frac{P(X = x_n, Y = y_m)}{P(X = x_n)} \tag{3.16}$$

When X and Y are independent $H(Y|X) = H(Y)$. Intuitively, we learn nothing of Y by knowing X, so Y is equally uncertain whether we know X or not. If Y is totally dependent on X, then the fraction in the logarithm is 1, which gives $H(Y|X) = 0$.

These entropy measures are the basis of the mutual information measure.

Maximum entropy limit

Originally, we intended to measure the mutual information between the spike counts of cells. We included this section to explain why using larger bin widths resulted in potentially larger spike counts, containing potentially more information, corresponding with higher mutual information values. The idea of measuring the mutual information between spike counts was abandoned. But this section was kept as an illustration of how the number of values a random variable can take affects the information conveyed by that variable.

When spiking data is binned into spike counts there is an upper limit on the entropy of these data. The maximum entropy discrete distribution is the discrete uniform distribution. A random variable with this distribution will take values from some finite set with equal probabilities. Binned spike count data will take values between 0 and some maximum observed spike count $n_{\max}$. A neuron with responses that maximises entropy will take these values with equal probability, i.e. if $i \in \{0, \ldots, n_{\max}\}$ then $P(X = i) = \frac{1}{n_{\max}+1}$. The entropy of

this neuron will be

$$\begin{aligned} H(X) &= -\sum_{i=0}^{n_{\max}} P(X=i)\log_2 P(X=i) \\ &= -\sum_{i=0}^{n_{\max}} \frac{1}{n_{\max}+1}\log_2\left(\frac{1}{n_{\max}+1}\right) \\ &= -\log_2\left(\frac{1}{n_{\max}+1}\right) \\ &= \log_2\left(n_{\max}+1\right) \end{aligned}$$

Therefore, the maximum entropy of the binned spike counts of a neuron is $\log_2(n_{\max}+1)$. Of course, it would be very unusual for a neuron to fire in accordance with the discrete uniform distribution. Most measurements of entropy taken on binned spiking data will be much lower than the maximum. See figure 3.2 to see the maximum entropy as a function of the maximum observed spike count.

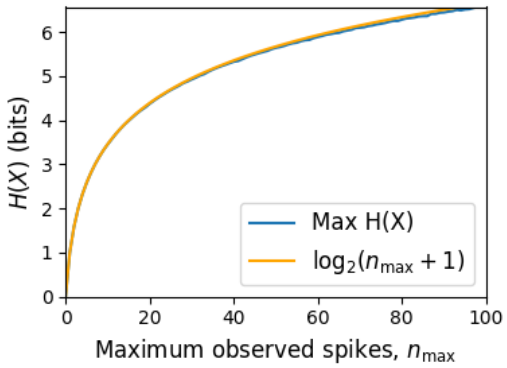


FIGURE 3.2: **Entropy Limit:** The upper limit on entropy of binned spike count data as a function of the maximum observed spike count. The orange line is the analytical maximum. The blue line is the entropy of samples with $N = 1000$ data points taken from the discrete uniform distribution.

Mutual Information $I(X;Y)$

The mutual information can be defined mathematically in a number of ways, all of which are equivalent. These definitions illustrate the different ways of interpreting the mutual information.

For two jointly distributed random variables X and Y, the mutual information $I(X;Y)$ is defined as

$$I(X;Y) = H(Y) - H(Y|X) \tag{3.17}$$

$$= H(X) - H(X|Y) \tag{3.18}$$

Equation 3.17 fits with the following intuition: The mutual information between X and Y is the reduction in uncertainty about X gained by knowing Y, or vice versa. We could also say the mutual information is the amount of information gained about X by knowing Y, or vice versa.

Another useful entropy based definition for the mutual information is

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{3.19}$$

This definition is useful because it does not require the calculation of conditional probabilities.

The mutual information can also be defined in terms of marginal, joint, and conditional distributions. For example,

$$I(X;Y) = -\sum_{n=1}^{N}\sum_{m=1}^{M} P(X = x_n, Y = y_m) \log_2 \frac{P(X = x_n, Y = y_m)}{P(X = x_n)P(Y = y_m)} \tag{3.20}$$

Notice that this can be rewritten as a Kullback–Leibler divergence.

$$I(X;Y) = D_{KL}(P(X,Y)||P(X)P(Y)) \tag{3.21}$$

So, we can also think of the mutual information as a measure of the difference between the joint distribution of X and Y, and the product of their marginal distributions. Since the product of the marginal distributions is the joint distribution for independent variables, we can think of the mutual information as a measure of the variables' dependence on one another.

The minimum value that $I(X;Y)$ can take is 0. This occurs when the random variables X and Y are independent. Then we have $H(X|Y) = H(X)$, and $H(Y|X) = H(Y)$, which according to equation 3.17, gives $I(X;Y) = 0$. We also have that $H(X,Y) = H(X) + H(Y)$ in this case, which according equation 3.19, gives $I(X;Y) = 0$. Finally, we also have $P(X,Y) = P(X)P(Y)$, which leaves us with 1 in the argument for the logarithm in equation 3.20, which again gives $I(X;Y) = 0$.

The mutual information reaches its maximum value when one of the variables X and Y is completely determined by knowing the value of the other. In that case $I(X;Y) = \min\{H(X), H(Y)\}$.

Variation of Information $VI(X,Y)$

The variation of information is another information theoretical quantity based on the mutual information. It is defined as

$$VI(X;Y) = H(X) + H(Y) - 2I(X;Y) \tag{3.22}$$

We can rewrite this as the summation of two positive quantities

$$VI(X;Y) = [H(X) - I(X;Y)] + [H(Y) - I(X;Y)] \tag{3.23}$$

In English, the variation of information is the summation of the uncertainty in the random variables X and Y excluding the uncertainty shared by those variables.

This measure will become more relevant when we go on to talk about clusterings because $VI(X;Y)$ forms a metric on the space of clusterings.

Measuring entropies & mutual information

In practice, we measured the mutual information between spike counts using Python and the python package `pyitlib`. We used the PT-bias correction technique to estimate the bias of our measurements when measuring the mutual information between the spike counts of two cells (Treves and Panzeri, 1995).

When measuring the mutual information between clusterings we used Python, but we used the `mutual_info_score`, `adjusted_mutual_info_score`, and `normalized_mutual_info_score` functions from the `sklearn.metrics` part of the `sklearn` package.

3.3.5 Network analysis

Correlation networks

In order to analyse functional networks created by the neurons in our ensemble, we measured the total correlation between each pair of neurons. These measurements induced an

undirected weighted graph/network between the neurons. The weight of each connection was equal to the total correlation between each pair of neurons.

We followed the same procedure for total correlations 3.3.2, spike count correlations, and signal correlations 3.3.3.

Rectified correlations

At the time of writing, the community detection method outlined in (Humphries et al., 2019) could only be applied to networks with positively weighted connections. But many neuron pairs were negatively correlated. To apply the community detection method, we *rectified* the network, by setting all the negative weights to zero.

We also looked for structure in the network created by negative correlations by reversing the signs of the correlations, and rectifying these correlations before applying our network analysis.

Finally, we used the absolute value of the correlations as the weights for the graph/network. By doing this, we hoped to identify both correlated and anti-correlated functional communities of neurons.

Sparsifying data networks

When creating our correlation networks, we wanted to exclude any correlations that could be judged to exist 'by chance'. To do this, we measured the 5th and 95th percentile of the shuffled correlations (see section 3.3.2) for the given mouse and time bin width. We then set all the data correlations between these two values to 0. This excluded any 'chance' correlations from our network, and created a sparser network. This allowed us to make use of the 'sparse weighted configuration model' as described in section 3.3.5.

Communities

Given some network represented by an adjacency matrix $\mathbf{A}$, a community within that network is defined as a collection of nodes where the number of connections within these nodes is higher than the expected number of connections between these nodes. In order to quantify the 'expected' number of connections, we need a model of expected networks. This is analogous to a 'null model' in traditional hypothesis testing. We test the hypothesis that our data network departs from the null network model to a statistically significant degree. For undirected unweighted networks, the canonical model of a null network is the configuration

model (Fosdick et al., 2016). Since we are working with weighted sparse networks, we used more suitable null models, described below.

Weighted configuration model

The *weighted configuration model* is a canonical null network model for weighted networks. Given some data network, the weighted configuration model null network will preserve the degree sequence and weight sequence of each node in the data network. But the edges will be distributed randomly (Fosdick et al., 2016). Any structure in the data network beyond its degree sequence and weight sequence will not be captured in the weighted configuration model. So, this model can be used in testing the hypothesis that this extra structure exists.

Sparse weighted configuration model

The *sparse weighted configuration model* is another null network model. Similar in nature to the weighted configuration model (see section 3.3.5), but the sparsity of the data network is preserved in the null network. This is achieved by sampling from a probability distribution for the creation or non-creation of each possible connection, then distributing the weight of the data network randomly in this sparse network (Humphries et al., 2019). This is the null network that we used when searching for additional structure in our data networks.

Spectral rejection

We made use of the spectral rejection algorithm as outlined in (Humphries et al., 2019). The spectral rejection algorithm is a method for finding structure in a network not captured by a supposed null model, if such structure exists.

To describe the method, we denote our data network matrix $\mathbf{W}$, we denote the expected network of our null network model as $\langle \mathbf{P} \rangle$. Then the departure of our data network from the null network can be described by the matrix

$$\mathbf{B} = \mathbf{W} - \langle \mathbf{P} \rangle \tag{3.24}$$

a common choice for $\langle \mathbf{P} \rangle$ in community detection is the 'configuration model' (Fosdick et al., 2016; Humphries, 2011). The matrix $\mathbf{B}$ is often called the configuration matrix, in this context we will use the term 'deviation matrix' as it captures the deviation of $\mathbf{W}$ from the null model.

To test for structure in the network represented by **W**, we examine the eigenspectrum of **B** and compare it to the eigenspectrum of our null model. Firstly, note that since our data model doesn't allow self loops, and is not directed, the matrix representing the network will be symmetric and positive semi-definite, and will therefore be invertible with real eigenvalues. We selected a null model with the same characteristics.

To find the eigenspectrum of the null model, we generated N samples from our null model $P_1, \ldots, P_N$, and we measured their deviation matrices $B_1, \ldots, B_N$. We then calculated the eigenspectrum of each of those samples. We calculated the upper bound of the null model eigenspectrum by taking the mean of the largest eigenvalues of $B_1, \ldots, B_N$. We calculated a lower bound on the null model eigenspectrum by taking the mean of the smallest eigenvalues of $B_1, \ldots, B_N$.

We then calculated the eigenspectrum of **B**, our data network deviation matrix. If any of those eigenvalues lay outside of the upper or lower bounds of the null model eigenspectrum, this is evidence of additional structure not captured by the null model. If we chose the sparse weighted configuration model (see section 3.3.5) as our null network model, then eigenvalues lying below the lower bound indicate k-partite structure in the network. For example, if one eigenvalue lay below the lower bound, this would indicate some bipartite structure in the data network. If any eigenvalues lay above the upper bound of the null model eigenspectrum, this is evidence of community structure in the data network. For example, one eigenvalue of **B** lying above the upper bound of the null model eigenspectrum indicates the presence of two communities in the network (Humphries, 2011).

Node rejection

If there are d data eigenvalues lying outside of the null network eigenspectrum, the d eigenvectors corresponding to these eigenvalues will form a vector space. If we project the nodes of our network into this vector space, by projecting either rows or columns of the data matrix, we can see how strongly each node contributes to the vector space. Nodes that contribute strongly to the additional structure will project far away from the origin, nodes that do not contribute to the additional structure will project close to the origin. We want to use this information to discard those nodes that do not contribute.

We can test whether a node projects *far* away from the origin or *close* to the origin using the eigenvalues and eigenvectors of $B_1, \ldots, B_N$. The jth eigenvector and eigenvalue of B_i gives a value for a null network's projection into the jth dimension of the additional structure vector space. The matrices $B_1, \ldots, B_N$ give N projections into that dimension.

These projections are a distribution of the null networks' projections. If the data node's projection exceeds that of the null network projections this node is judged to project *far* from the origin, and therefore contribute to the additional structure. Otherwise, the node is judged to project *close* to the origin, and is therefore rejected (Humphries et al., 2019).

Community detection

Another application for this d dimensional space is community detection. We first project all of the nodes into this d-dimensional space, then perform the clustering in this space. The clustering and community detection procedure is described in (Humphries, 2011).

In practice, the procedure is carried out n times (we chose $n = 100$ times), this returns n clusterings. We resolve these n clusterings to one final clustering using *consensus clustering*. We used the consensus clustering method that uses an explicit null model for the consensus matrix, as outlined in (Humphries et al., 2019).

3.3.6 Clustering Comparison

A clustering $\mathcal{C}$ is a partition of a set D into sets $C_1, C_2, \ldots, C_K$, called clusters, that satisfy the following for all $k, l \in \{1, \ldots, K\}$:

$$C_k \cap C_l = \emptyset \tag{3.25}$$

$$\bigcup_{k=1}^{K} C_k = D \tag{3.26}$$

If we consider two clusterings, $\mathcal{C}$ with clusters $C_1, C_2, \ldots, C_K$ and $\mathcal{C}'$ with clusters $C'_1, C'_2, \ldots, C'_K$. There are a number of measurements we can use to compare $\mathcal{C}$ and $\mathcal{C}'$. In the following, the number of elements in D is denoted by n, and the number of elements in cluster C_k is n_k.

Adjusted Rand Index

The *adjusted Rand Index* is a normalised similarity measure for clusterings based on pair counting.

If we consider the clusterings $\mathcal{C}$ and $\mathcal{C}'$, and denote

- the number of pairs in the same cluster in $\mathcal{C}$ and $\mathcal{C}'$ by N_{11}
- the number of pairs in different clusters in $\mathcal{C}$ and $\mathcal{C}'$ by N_{00}

- the number of pairs in the same cluster in $\mathcal{C}$ and different clusters in $\mathcal{C}'$ by N_{10}
- the number of pairs in different clusters in $\mathcal{C}$ and the same cluster in $\mathcal{C}'$ by N_{01}

then the *Rand Index* is defined as

$$RI = \frac{N_{11} + N_{00}}{N_{11} + N_{00} + N_{10} + N_{01}} = \frac{N_{11} + N_{00}}{\binom{n}{2}} \tag{3.27}$$

The Rand Index is 1 when the clusterings are identical, and 0 when the clusterings are completely different.

The *adjusted Rand Index* intends on correcting the Rand Index for chance matching pairs. This is defined as

$$ARI = \frac{2\left(N_{00}N_{11} - N_{01}N_{10}\right)}{\left(N_{00} + N_{01}\right)\left(N_{01} + N_{11}\right) + \left(N_{00} + N_{10}\right)\left(N_{10} + N_{11}\right)} \tag{3.28}$$

The adjusted Rand Index is 1 when the clusterings are identical, and 0 when the Rand Index is equal to its expected value.

When the number of clusters in each clustering is different, the concept of a pair of being in the same cluster in both clusterings becomes difficult to define. To address this, two clusterings are drawn randomly with $N_{\mathcal{C}}$ and $N_{\mathcal{C}'}$ number of clusters respectively, and a fixed number of elements in each cluster corresponding to the number of elements in each cluster in $\mathcal{C}$ and $\mathcal{C}'$. Then the adjusted Rand Index is the normalised difference between the Rand Index of $\mathcal{C}$ and $\mathcal{C}'$ and the mean value of the Rand Index measured using many pairs of these 'random' clusterings.

Clusterings as random variables

If we take any random element of D, the probability that this element is in cluster C_k of clustering $\mathcal{C}$ is

$$P(K = k) = \frac{n_k}{n} \tag{3.29}$$

this defines a probability distribution, which makes the clustering a random variable. Any clustering can be considered as a random variable this way.

This means that we can measure any of the information theoretic quantities defined in section 3.3.4 with respect to clusterings. For example, the entropy of a clustering is

$$H(\mathcal{C}) = -\sum_{k=1}^{K} \frac{n_k}{n} \log \frac{n_k}{n} \tag{3.30}$$

If we have two clusterings, the joint probability distribution of these clusterings is defined as

$$P(K = k, K' = k') = \frac{|C_k \cap C'_{k'}|}{n} \tag{3.31}$$

The joint distribution allows us to define the mutual information between two clusterings, $I(\mathcal{C};\mathcal{C}')$ (Meilă, 2007).

Information based similarity measures

The mutual information between two clusterings is a similarity measure, with $I(\mathcal{C};\mathcal{C}') = 0$ if $\mathcal{C}$ and $\mathcal{C}'$ are completely different, and $I(\mathcal{C};\mathcal{C}') = H(\mathcal{C}) = H(\mathcal{C}')$ if $\mathcal{C}$ and $\mathcal{C}'$ are identical. This can be normalised in a number of different ways to make more similarity measures (Vinh, Epps, and Bailey, 2010)

$$NMI_{joint} = \frac{I(\mathcal{C};\mathcal{C}')}{H(\mathcal{C},\mathcal{C}')} \tag{3.32}$$

$$NMI_{max} = \frac{I(\mathcal{C};\mathcal{C}')}{\max\{H(\mathcal{C}), H(\mathcal{C}')\}} \tag{3.33}$$

$$NMI_{sum} = \frac{2I(\mathcal{C};\mathcal{C}')}{H(\mathcal{C}) + H(\mathcal{C}')} \tag{3.34}$$

$$NMI_{sqrt} = \frac{I(\mathcal{C};\mathcal{C}')}{\sqrt{H(\mathcal{C})H(\mathcal{C}')}} \tag{3.35}$$

$$NMI_{min} = \frac{I(\mathcal{C};\mathcal{C}')}{\min\{H(\mathcal{C}), H(\mathcal{C}')\}} \tag{3.36}$$

We can control for chance similarities between the two clusterings by measuring the *adjusted mutual information* between the clusterings. This is defined as

$$AMI_{sum} = \frac{I(\mathcal{C};\mathcal{C}') - E\{I(\mathcal{C};\mathcal{C}')\}}{\frac{1}{2}\left[H(\mathcal{C}) + H(\mathcal{C}')\right] - E\{I(\mathcal{C};\mathcal{C}')\}} \tag{3.37}$$

The first term in the denominator, taking the average of the marginal entropies, can be replaced by taking the maximum, minimum, or the geometric mean (Vinh, Epps, and Bailey, 2010).

Information based metrics

The variation of information between two clusterings $VI(\mathcal{C};\mathcal{C}')$ (see section 3.3.4) is a metric on the space of clusterings (Meilă, 2007). That is,

$$VI(\mathcal{C};\mathcal{C}') \geq 0 \tag{3.38}$$

$$VI(\mathcal{C};\mathcal{C}') = 0 \iff \mathcal{C} = \mathcal{C}' \tag{3.39}$$

$$VI(\mathcal{C};\mathcal{C}') = VI(\mathcal{C}';\mathcal{C}) \tag{3.40}$$

$$VI(\mathcal{C};\mathcal{C}'') \leq VI(\mathcal{C};\mathcal{C}') + VI(\mathcal{C}';\mathcal{C}'') \tag{3.41}$$

Another metric is the *information distance* (Vinh, Epps, and Bailey, 2010)

$$D_{max} = \max\{H(\mathcal{C}), H(\mathcal{C}')\} - I(\mathcal{C};\mathcal{C}') \tag{3.42}$$

Both of these can be normalised

$$NVI(\mathcal{C};\mathcal{C}') = 1 - \frac{I(\mathcal{C};\mathcal{C}')}{H(\mathcal{C},\mathcal{C}')} \tag{3.43}$$

$$d_{max} = 1 - \frac{I(\mathcal{C};\mathcal{C}')}{\max\{H(\mathcal{C}), H(\mathcal{C}')\}} \tag{3.44}$$

Comparing detected communities and anatomical divisions

In order to quantify the difference or similarity between the communities detected in our correlation network and the anatomical classification of the cells in that network, we considered the communities and the anatomical regions as clusters in two different clusterings, $\mathcal{C}_{comm}$ and $\mathcal{C}_{anat}$, respectively. We then measured the similarity between the clusterings using the mutual information, the adjusted mutual information, and the normalised mutual information. We measured the difference between, or the distance between, the clusterings using the variation of information, the normalised variation of information, and the normalised information distance. We also measured the difference between the clusterings using the adjusted Rand Index, just to use a non-information based measure.

We took all of these measures for communities detected using different time bin widths. This gave us an idea of the effect of time bin width on correlation networks in neural ensembles relative to anatomical regions within those ensembles.

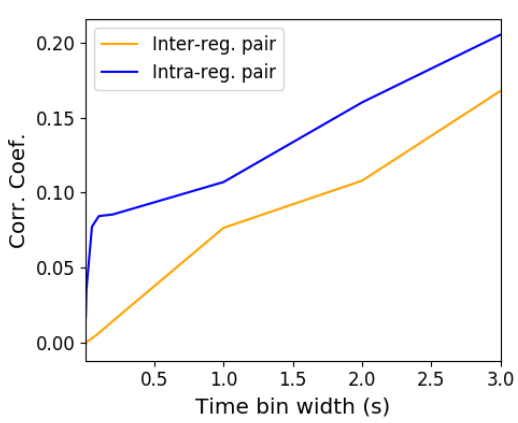
3.4 Results

Note that in the following text, we refer to the correlation coefficient between two sequences of spike counts from two different cells as the *total correlation*. We refer to the correlation between spike counts in response to a certain stimulus as the *spike count correlation* aka *noise correlation*, and we refer to the correlation between mean or expected responses to different stimuli as the *signal correlation* (Cohen and Kohn, 2011).

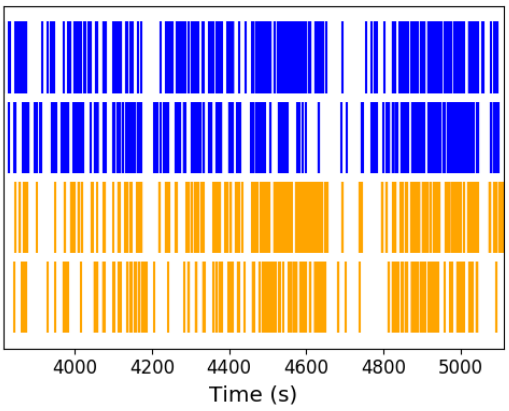
The nine different brain regions from which we had data were the caudate putamen (CP), frontal motor cortex (FrMoCtx), hippocampus (HPF), lateral septum (LS), midbrain (MB), primary visual cortex (V1), superior colliculus (SC), somatomotor cortex (SomMoCtx), and thalamus (TH).

3.4.1 Average correlation size increases with increasing time bin width

First we inspected the affect of time bin width on total correlations. We know that using short time bins results in artificially small correlation measurements (Cohen and Kohn, 2011), so we expected to see an increase in correlation amplitude with increasing time bin width. That is exactly what we observed. Taking 50 cells at random, we calculated the total correlation between every possible pair of these cells, using different time bin widths ranging from 0.005s to 3s. We found that the longer the time bin width, the greater the correlations (see figure 3.4a).



(A) Correlation coefficient as a function of bin width.



(B) Raster plots for the four cells making up our example pairs.

FIGURE 3.3: (A) An example of the correlation coefficients between two different pairs of cells, one where both cells are in the same brain region (intra-regional pair), and one where both cells are in different brain regions (inter-regional pair). The correlation coefficients have been measured using different time bin widths, ranging from 5ms to 3s. Note the increasing amplitude of the correlations with increasing bin width. (B) A raster plot showing the spike times of each pair of cells.

We also separated the positively correlated pairs from the negatively correlated pairs using the mean correlation of each pair across all bin widths (see section 3.3.2). We found that the positively correlated pairs become more positively correlated with increasing time bin width, and the negatively correlated pairs become more negatively correlated with increasing time bin width (see figures 3.4b and 3.4c).

In figure 3.3a we plot correlations from two example pairs, one pair from within a region, and one pair between regions. It can be seen that the correlation coefficient increases with bin width. The correlations can be observed by eye in the raster plot for these cells in figure 3.3b.

When taking the mean across all pairs, the positively correlated pairs dominate in terms of both number of pairs, and amplitude of correlations. Therefore the mean across all pairs is positive.

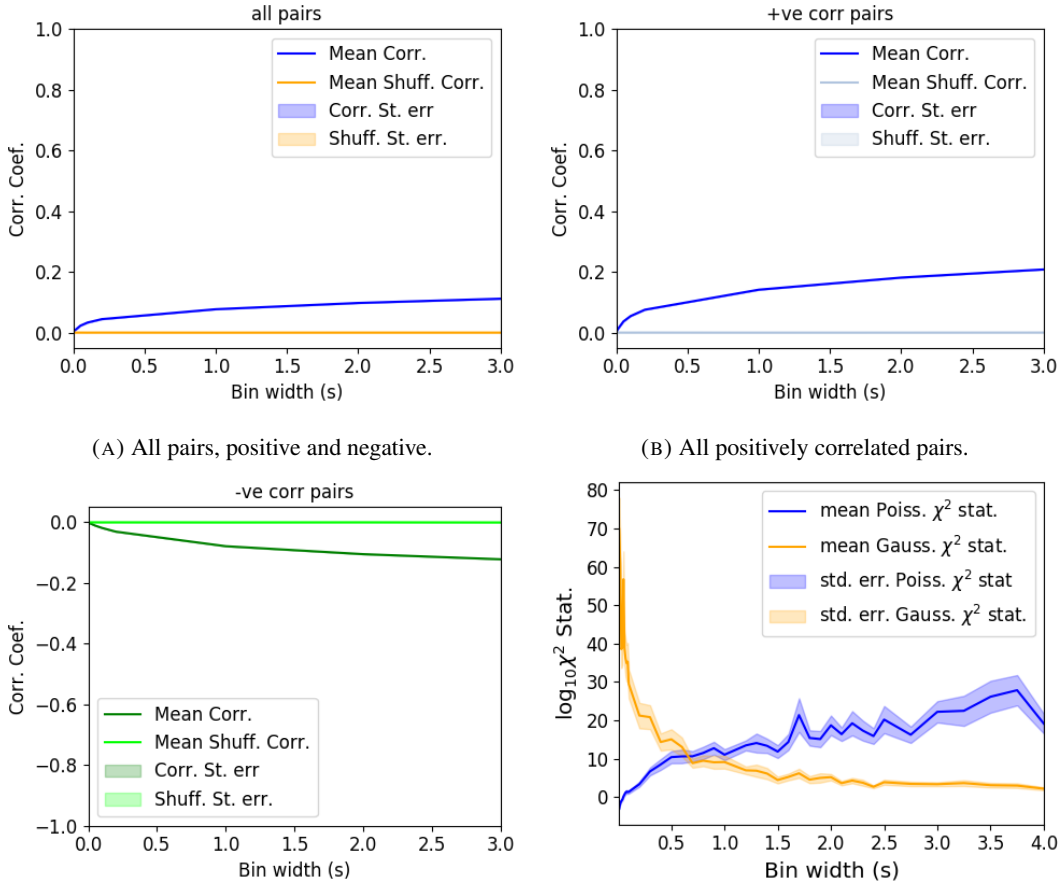
These results were observed in each of the three mouse subjects from which we had data.

3.4.2 Goodness-of-fit for Poisson and Gaussian distributions across increasing time bin widths

We wanted to investigate if the width of the time bin used to bin spike times into spike counts had an effect on the distribution of spike counts. We used the χ^2 statistic as a goodness-of-fit measure for Poisson and Gaussian (normal) distributions to the spike count of 100 randomly chosen neurons for a number of bin widths ranging from 0.01s to 4s. For the χ^2 statistic, the higher the value, the worse the fit.

We expected a Poisson distribution to be a better fit for shorter time bin widths because spike counts must be non-negative, therefore any distribution of spike counts with mass distributed at or close to 0 will be skewed. The distribution of spike counts is more likely to be distributed close to 0 when the time bin widths used to bin spike times into spike counts are small relative to the amount of time it takes for a neuron to fire an action potential ($\sim$ 1ms in the case of non-burst firing neurons).

We expected a Gaussian distribution to be a better fit for longer time bin widths, because a Poisson distribution with a large rate is well approximated by a Gaussian distribution with mean and variance equal to the Poisson rate. Therefore, a Gaussian distribution would approximate the mean of a collection of large spike counts, and have more flexibility than a Poisson distribution to fit the variance.



(A) All pairs, positive and negative.

(B) All positively correlated pairs.

(C) All negatively correlated pairs.

(D) χ^2 test statistics as goodness-of-fit.

FIGURE 3.4: Mean correlation coefficients measured from pairs of 50 randomly chosen neurons. (A) All possible pairs, (B) positively correlated pairs, and (C) negatively correlated pairs. (D) Mean and standard error of χ^2 test statistics for Poisson and Gaussian distributions fitted to neuron spike counts.

We found that that a Poisson distribution is the best fit for shorter time bins less than 0.7s in length. Then a Gaussian distribution is a better fit for time bins greater than 0.7s in length (see figure 3.4d).

3.4.3 Differences between and inter- and intra- regional correlations decrease with increasing bin width

We investigated the differences in distribution between inter-regional correlations, i.e. correlations between neurons in different brain regions, and intra-regional correlations, i.e. correlations between neurons in the same brain region.

Firstly, we investigated these quantities for all possible pairs of $\sim$ 500 neurons taken from across all the 9 brain regions from which we had data. We distributed these neurons as evenly as possible across all of the regions, so that cells from one region would not dominate our data. We observed that the mean intra-regional correlations were always higher than the mean inter-regional correlations for every value of time bin width used. We also observed that as the time bin width increased these mean correlations increased and the difference between the mean inter-regional and intra-regional correlations grew (see figure 3.5 (Left)).

Stringer et al. (2019) had a similar finding using the same data. They used only one value for the time bin width, 1.2s. Using this time bin width to bin spike times and measure total correlations, they found that the mean 'within-region' correlations were always greater than the 'out-of-region' correlations (Stringer et al., 2019). The figure from their paper showing this result can be seen in figure 3.5 (Right).

Examples of the correlations of one intra-regional pair and one inter-regional pair can be seen in figure 3.3.

Secondly, we separated those pairs into intra-regional and inter-regional groups. We noted that the mean intra-regional correlations (coloured dots in figures 3.6a and 3.6b) for a given region tended to be higher than the mean inter-regional correlations (black dots in figures 3.6a and 3.6b) involving cells from that region. However, in contrast with our previous result, we noted that the difference between the mean intra-regional correlations and most highly correlated inter-regional correlations reduced as we increased the time bin width (see figures 3.6a and 3.6b). This shows that the mean correlations showin in figure 3.5 are not distributed evenly across all region pair combinations.

Finally, to see these regional mean correlations in a bit more detail, to examine the individual pair combinations in particular, we displayed these data in a matrix of mean correlations (see figure 3.7), showing the mean intra-regional correlations on the main diagonal, and

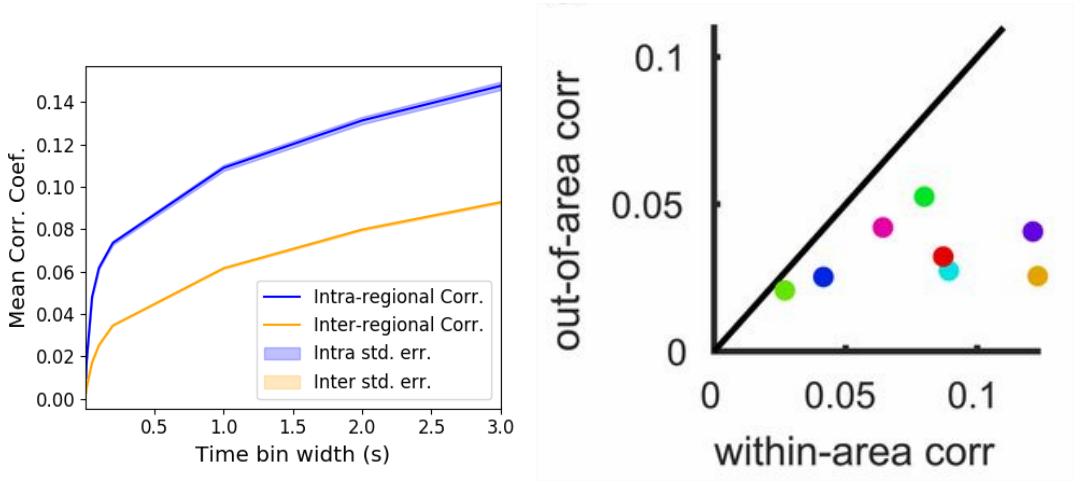


FIGURE 3.5: (Left)The mean intra-region and inter-region correlations using all possible pairs of $\sim$ 500 neurons, spread across 9 different brain regions. (Right) Courtesy of Stringer et al. (2019), mean inter-regional (out-of-area) correlation coefficients vs mean intra-regional (within-area) correlation coefficients for a bin width of 1.2s. Note that the intra-regional coefficients are higher in each case.

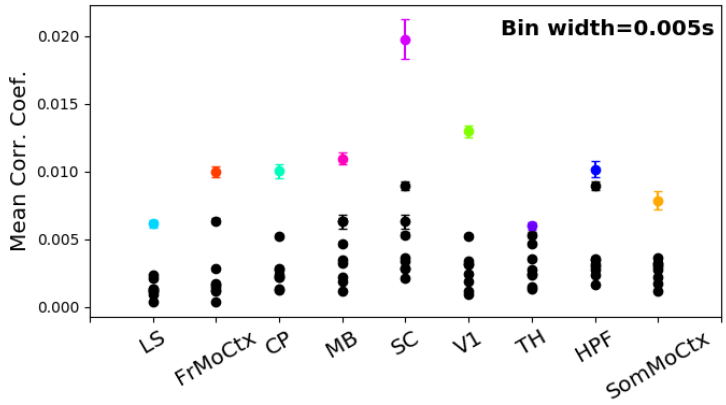
the mean inter-regional correlations off diagonal. Comparing a version of this figure created using a short time bin width of 5ms (figure 3.7a) and a version using a longer time bin width of 1s (figure 3.7b) we observed that the mean intra-regional correlations are always relatively high in comparison to the mean inter-regional correlations, but the mean correlations in some inter-regional pairs are relatively much higher when using the longer time bin width.

This could indicate information being processed quickly at a local or within-region level, and the local representations of this information spreading between regions at longer timescales.

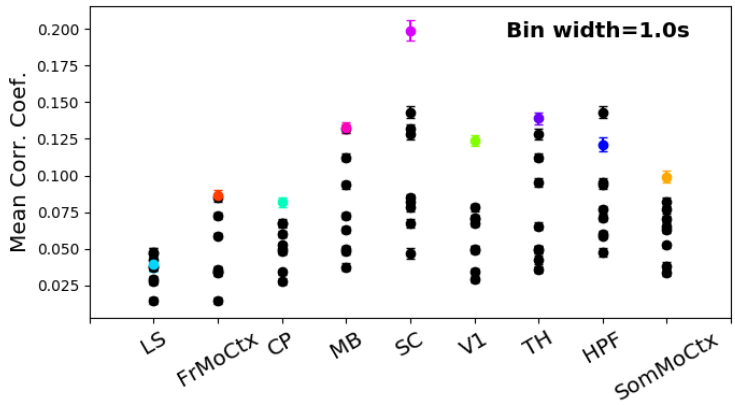
These results were consistent across the three mouse subjects. But, the relative magnitudes of the mean intra-regional and inter-regional correlations were not consistent. For example, the region with the highest mean intra-regional correlations when using 1s bin widths for subject one is the superior colliculus (SC), but for subject two it is the midbrain (MB).

3.4.4 Connected and divided structure in correlation based networks reduces in dimension with increasing bin width

We used the correlation measurements to create weighted undirected graphs/networks where each node represents a neuron, and the weight of each edge is the pairwise correlation between those neurons represented by the nodes at either end of that edge. We aimed to find communities of neurons within these networks, and compare the structure of these communities to the anatomical division of those neurons. The first step of this process involved applying the 'spectral rejection' technique developed by Humphries et al (2019) (Humphries

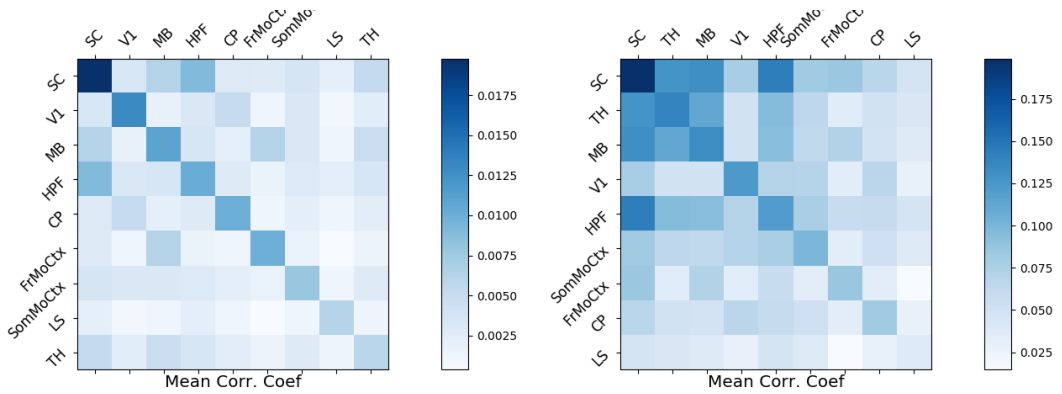


(A) Mean inter-regional and intra-regional correlations using a time bin width of 5ms.



(B) Mean inter-regional and intra-regional correlations using a time bin width of 1s.

FIGURE 3.6: The mean intra-regional correlations (coloured dots) and mean inter-regional correlations (black dots) for a given region, indicated on the x-axis, for different time bin widths. Each black dot represents the mean inter-regional correlations between the region indicated on the x-axis and one other region. (A) shows these measurements when we used a time bin width of 5ms. (B) shows these measurements when we used a time bin width of 1s. Note that the difference between the mean inter-regional correlations and mean intra-regional correlations is smaller for 1s bins.



(A) Time bin width 0.005s.

(B) Time bin width 1s.

FIGURE 3.7: Mean inter-regional (main diagonal) and intra-regional (off diagonal) correlation coefficients. (A) Shows these measurements when spike times were binned using 5ms time bins. (B) Shows the same, using 1s time bins. Note that the matrices are ordered according to the main diagonal values, therefore the ordering is different in each subfigure.

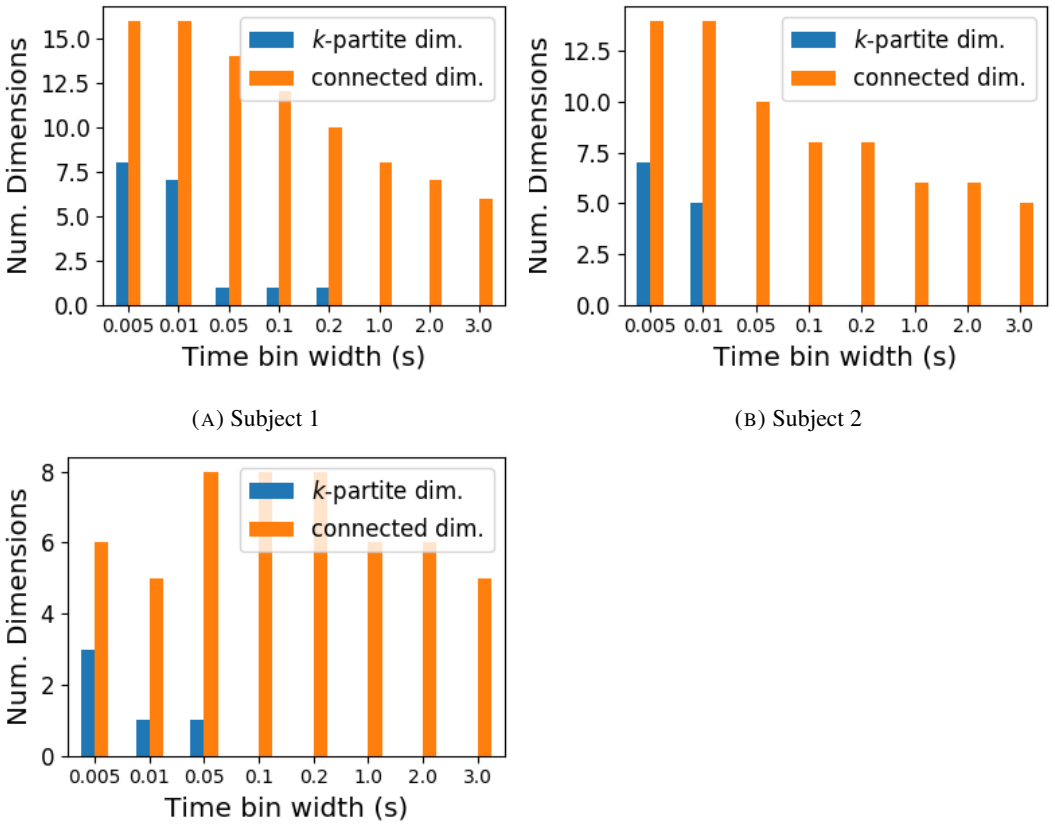
et al., 2019). This technique compares our data network to a chosen null network model, and finds any additional structure in the data network beyond that which is captured in the null network model (if there is any such structure).

By comparing the eigenspectrum of the data network to the eigenspectrum of many samples from the null network model, this technique allows us to estimate the dimensionality of the additional structure in the data network, and gives us a basis for that vector space. It also divides the additional structure into connected structure, and k-partite (or divided) structure. For example, if our algorithm found two dimensions of additional connected structure, and one dimension of additional divided structure. We might expect to find three communities, that is groups more strongly connected within group than without, and we might expect to find bi-partite structure, that is two sets that are more strongly connected between groups than within groups.

The technique also finds which nodes contribute to this additional structure, and divides our data network into signal and noise networks. The details of spectral rejection and node rejection can be found in sections 3.3.5 and 3.3.5 respectively, and a full overview can be found in (Humphries et al., 2019).

We chose the sparse weighted configuration model (see section 3.3.5) as our null network model. This model matches the sparsity and the total weight of the original network but distributes the weight at random across the sparse network.

We applied the spectral rejection method to our networks based on total correlations using different values for the time bin width. We observed that for smaller time bin widths, our data



(A) Subject 1

(B) Subject 2

(C) Subject 3

FIGURE 3.8: The number of dimensions in the k-partite and connected structure in the correlation based networks beyond the structure captured by a sparse weighted configuration null network model (see section 3.3.5), shown for different time bin widths. Note that the k-partite structure disappears for time bin width greater than 200ms for all three subjects. The dimension of the connected structure reduces with increasing bin width for 2 of the 3 subjects (top row).

networks had both k-partite structure, and community structure. As the width of the time bin increased, we found that the k-partite structure disappeared from our data networks, and the dimension of the community structure reduced in two of the three mice from which we had data (see figure 3.8).

The reduction in dimensionality of the connected structure with larger bin widths could indicate information or activity being integrated through the merging of smaller functionaly communities over longer timescales.

The k-partite structure that we found when using small bin widths could be an indication of physical connections between neurons. This is supported by the fact that these k-partite communities are not found over longer timescales. The effect of physical connections through axodendritic or dendrodendritic synapses would only be noticeable at short timescales.

3.4.5 Detecting communities in correlation based networks

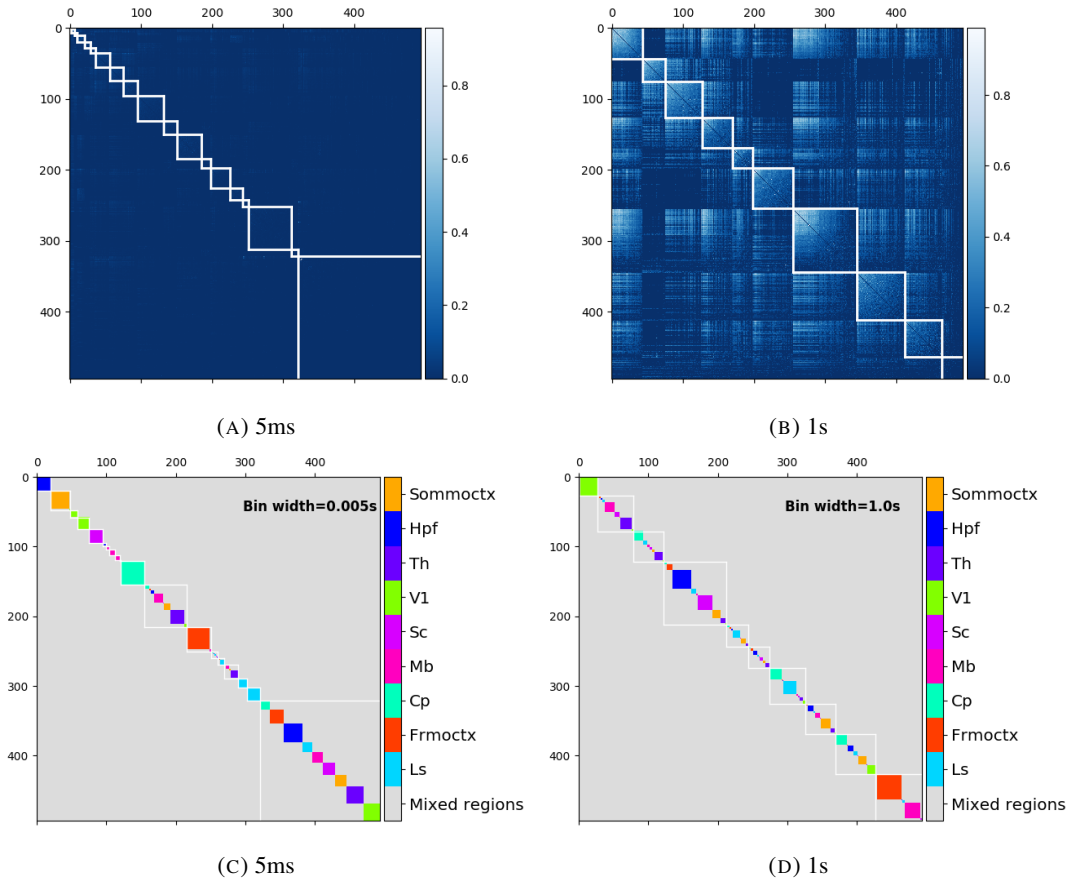
We applied the community detection procedure described in section 3.3.5 to our signal networks for our various time bin widths. We detected a greater number of smaller communities at shorter time bin widths, and a smaller number of larger communities for longer time bin widths (see figure 3.9). This was expected after the results found in section 3.4.4. We found more dimensions of additional structure at shorter time bin widths, therefore we found more communities at shorter time bin widths.

The number of communities that we detected was always greater than the dimensionality of the additional structure that we found by applying spectral rejection.

We also noticed that at short time bin widths the communities detected tended to be dominated by cells from one region. Whereas communities existing in networks created using wider time bin widths tended to contain cells from many different brain regions. More on this in the next section.

3.4.6 Functional communities resemble anatomical division at short timescales

In order to quantify the similarity of the communities detected to the anatomical division of the cells. We treated both the anatomical division and the communities as clusterings of these cells. We then used measures for quantifying the difference or similarity between clusterings to quantify the difference or similarity between the detected communities and the anatomical division. Details of these measures can be found in section 3.3.6 or in (Vinh, Epps, and Bailey, 2010).



(A) 5ms (B) 1s

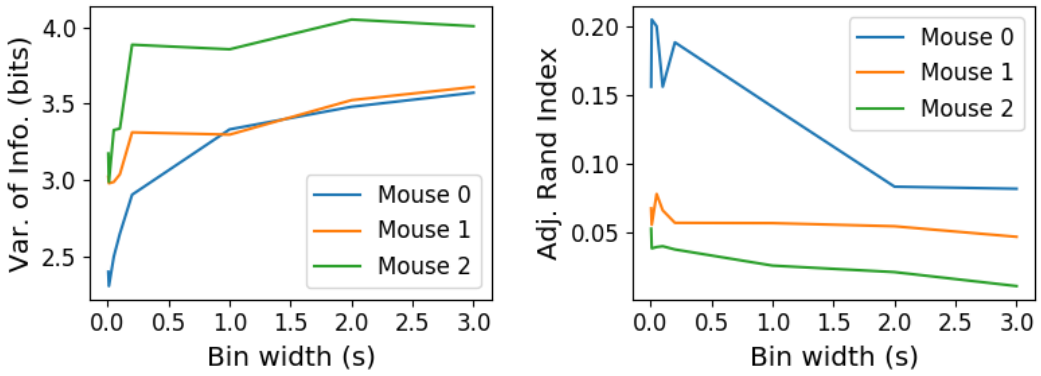
(C) 5ms (D) 1s

FIGURE 3.9: (A-B) Correlation matrices with detected communities indicated by white lines. Each off main diagonal entry in the matrix represents a pair of neurons. Those entries within a white square indicate that both of those neurons are in the same community as detected by our community detection procedure. Matrices shown are for 5ms and 1s time bin widths respectively. Main diagonal entries were set to 0. (C-D) Matrices showing the anatomical distribution of pairs along with their community membership. Entries where both cells are in the same region are given a colour indicated by the colour bar. Entries where cells are in different regions are given the grey colour also indicated by the colour bar.

We used two different types of measures for clustering comparison; information based measures (see section 3.3.6) and pair counting based measures (see section 3.3.6). We include one example of each in figure 3.10.

The variation of information is the information based measure included in figure 3.10a. This measure forms a metric on the space of clusterings. The larger the value for the variation of information, the more different the clusterings.

The adjusted Rand index is the pair counting based measure included in figure 3.10b. In contrast with the variation of information, the adjusted Rand index is a normalised similarity measure. The adjusted Rand index takes value 1 when the clusterings are identical, and takes value 0 when the clusterings are no more similar than chance.



(A) Variation of information

(B) Adjusted Rand index

FIGURE 3.10: (A) The variation of information is a measure of distance between clusterings. The distance between the anatomical 'clustering' and community detection 'clustering' increases with increasing time bin width. (B) The adjusted Rand index is a normalised similarity measure between clusterings. The anatomical and community detection clusterings become less similar as the time bin width increases.

Both measures indicated that the detected communities and the anatomical division of the cells were more similar when we used shorter time bins widths (see figure 3.10). This indicates that correlated behaviour in neuronal ensembles is more restricted to individual brain regions at short timescales ($<$ 250ms), and the correlated activity spreads out across brain regions over longer time scales.

3.4.7 Conditional correlations & signal correlations

In light of the excellent research of Stringer et al (2019) showing that spontaneous behaviours can drive activity in neuronal ensembles across the visual cortex and midbrain (Stringer et al., 2019), we decided to control for the mouse's behaviour when performing our analyses. It is possible that our community detection process may be detecting communities across

multiple brain regions at longer time scales due to aggregating neuronal activity driven by several spontaneous behaviours occurring during the time interval covered by a given time bin. A time bin of 1s, for example, could contain a spike count where those spikes were driven by different spontaneous behaviours. We aimed to investigate this possibility by applying our community detection analysis to conditional correlation measures.

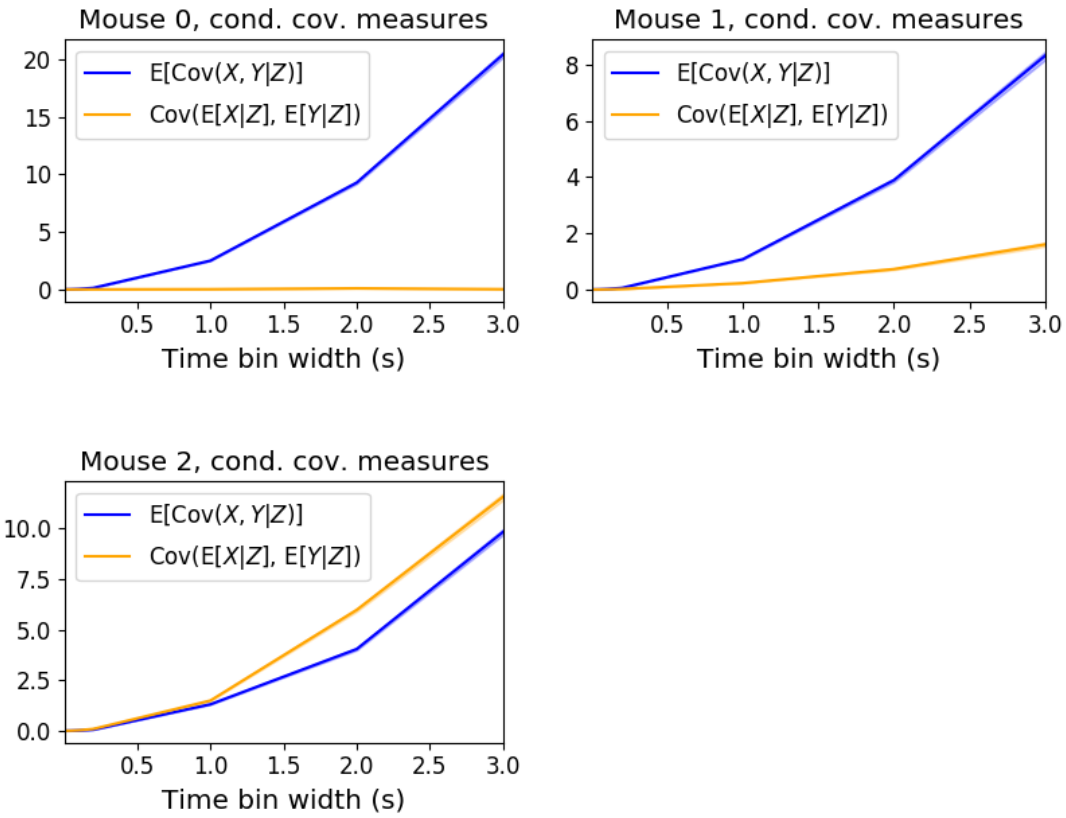


FIGURE 3.11: Comparing the components of the total covariance across different values for the time bin width. We observed a consistent increase in $E[\text{cov}(X,Y|Z)]$ as the time bin width increased. But we saw different trends for $\text{cov}(E[X|Z], E[Y|Z])$ for each mouse.

We used the top 500 principal components of a video of the mouse's face as a measure of the mouse's behaviour (see section 3.2.2). We modelled the spike counts as a linear combination of the principal components using linear regression with ElasticNet regularisation (see section 3.3.3). Using this model, we quantified the expected spike count given the mouse's behaviour $E[X|Z_1, \ldots, Z_{500}]$.

We used these expected values to measure $\text{cov}(E[X|Z], E[Y|Z])$, and we used that value, the covariance $\text{cov}(X,Y)$, and the *law of total covariance* (see section 3.3.3) to measure

$E[\text{cov}(X, Y|Z)]$. Here X and Y represent spike counts from individual cells, and Z is shorthand for the 500 principal components mentioned above. The two components of the covariance, $\text{cov}(E[X|Z], E[Y|Z])$ and $E[\text{cov}(X, Y|Z)]$, represent a 'signal covariance' and expected value of a 'spike count covariance' respectively, analagous to the signal correlation and spike count correlation (Cohen and Kohn, 2011).

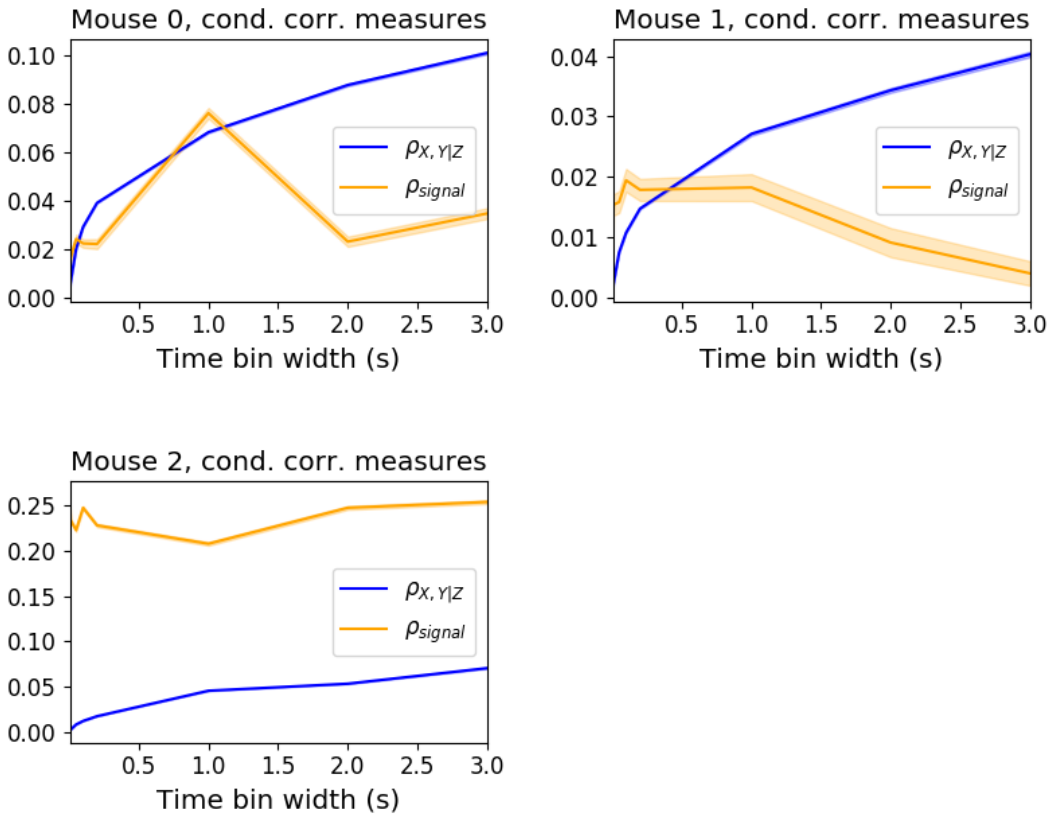
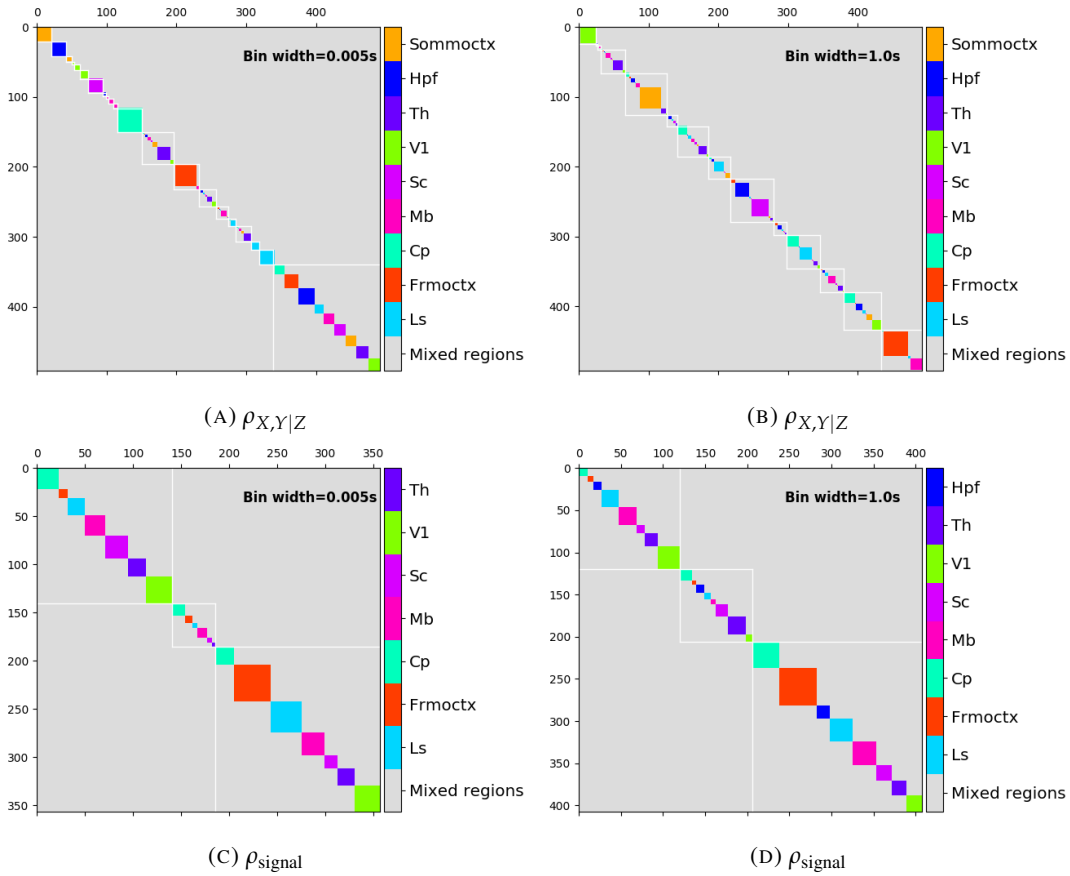


FIGURE 3.12: Comparing the components of the total covariance across different values for the time bin width. We saw a consistent increase in $\rho_{X,Y|Z}$ as the time bin width increased in all three subjects. But we saw different trends in ρ_{signal} for each of the subjects.

We examined the means of these components for different values of the time bin width (see figure 3.11). We observed a consistent increase in $E[\text{cov}(X, Y|Z)]$ as the time bin width increased. But we saw different trends for $\text{cov}(E[X|Z], E[Y|Z])$ for each mouse.

Using $\text{cov}(E[X|Z], E[Y|Z])$ we measured the signal correlation, ρ_{signal}, and using $E[\text{cov}(X, Y|Z)]$ we measured the event conditional correlation, $\rho_{X,Y|Z}$ (see section 3.3.3 for more details). We saw a consistent increase in $\rho_{X,Y|Z}$ as the time bin width increased, this corresponds to the result for $E[\text{cov}(X, Y|Z)]$. We observed different trends for ρ_{signal} for each mouse, this corresponds to the result for $\text{cov}(E[X|Z], E[Y|Z])$.

We applied our network noise rejection and community detection process to networks

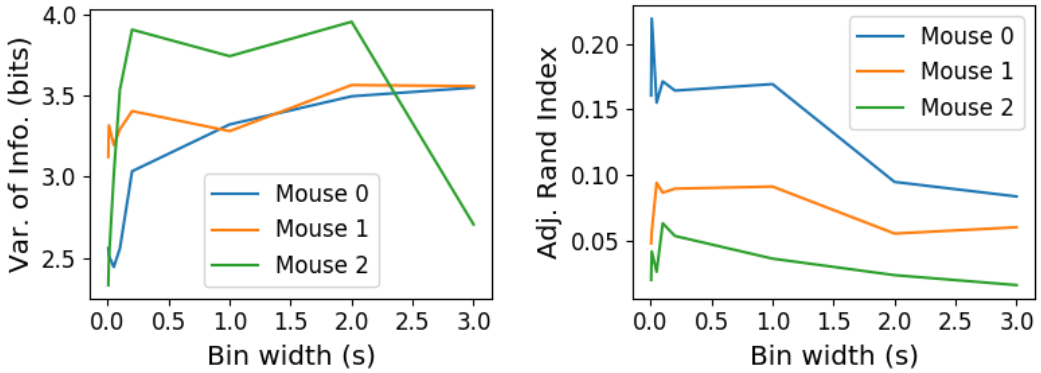


(A) $\rho_{X,Y|Z}$ (B) $\rho_{X,Y|Z}$

(C) ρ_{signal} (D) ρ_{signal}

FIGURE 3.13: Matrices showing the regional membership of pairs by colour, and the communities in which those pairs lie. (A-B) Detected communities and regional membership matrix for network based on rectified spike count correlation $\rho_{X,Y|Z}$, using time bin widths of 0.005s and 1s respectively. (C-D) Detected communities and regional membership matrix for network based on rectified signal correlation ρ_{signal}, using time bin widths of 0.005s and 1s respectively.

based on the spike count correlations $\rho_{X,Y|Z}$ and the signal correlations ρ_{signal}. We noted that the community detection on $\rho_{X,Y|Z}$ behaved similarly to the community detection on the total correlation. We can see this in figures 3.13a and 3.13b. At very short time bin widths, we detect more communities, and those communities often contain cells from one brain region only. At longer time bin widths, we detect fewer communities, and those communities tend to contain cells from multiple brain regions. When we examine the distance between (or similarity between) the anatomical division of the cells, and the detected communities we notice that the two clusterings are more similar at shorter time bin widths (see figure 3.14).

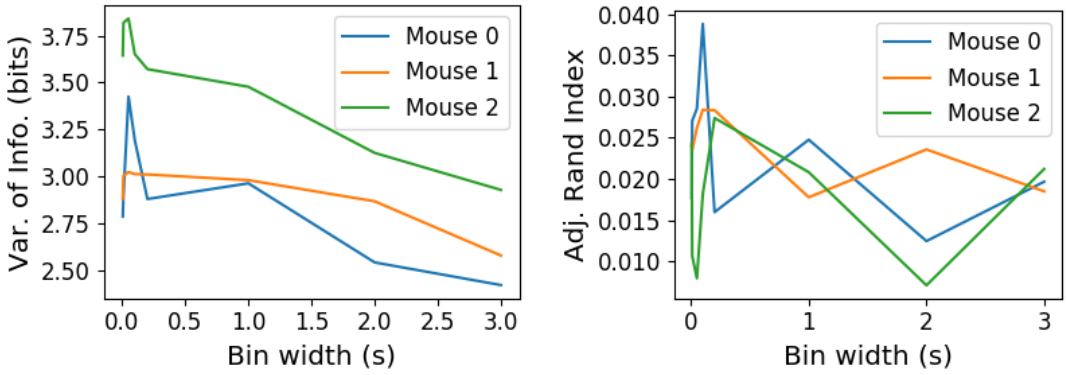


(A) $\rho_{X,Y|Z}$ Variation of information.

(B) $\rho_{X,Y|Z}$ Adjusted Rand Index.

FIGURE 3.14: Distance and similarity measures between the anatomical division of the neurons, and the communities detected in the network based on the spike count correlations $\rho_{X,Y|Z}$. (A) The variation of information is a 'distance' measure between clusterings. The distance between the anatomical 'clustering' and the community clustering increases as the time bin width increases. (B) The adjusted Rand index is a similarity measure between clusterings. The detected communities become less similar to the anatomical division of the cells as the time bin width increases.

When we applied the network noise rejection and community detection process to the networks based on the signal correlations ρ_{signal} we found the number of communities we detected reduced with increasing time bin width. But the number of communities detected was less than that for the total correlations or the spike count correlations. The communities detected always tended to contain cells from multiple regions at both short and long timescales (see figures 3.13c and 3.13d). The communities detected bore very little relation to the anatomical division of the cells. The adjusted Rand index between the community clustering and the anatomical 'clustering' is close to zero for every time bin width (see figure 3.15b). This indicates that the similarity between the clusterings is close to chance. We did observe a slight downward trend in the variation of information with increasing bin width (see figure 3.15a), but this is more likely due to a decrease in the number of communities



(A) ρ_{signal} Variation of information.

(B) ρ_{signal} Adjusted Rand Index.

FIGURE 3.15: Distance and similarity measures between the anatomical division of the neurons, and the communities detected in the network based on the signal correlations ρ_{signal}. (A) The variation of information is a 'distance' measure between clusterings. The distance between the anatomical 'clustering' and the community clustering increases as the time bin width increases. (B) The adjusted Rand index is a similarity measure between clusterings. The detected communities become less similar to the anatomical division of the cells as the time bin width increases.

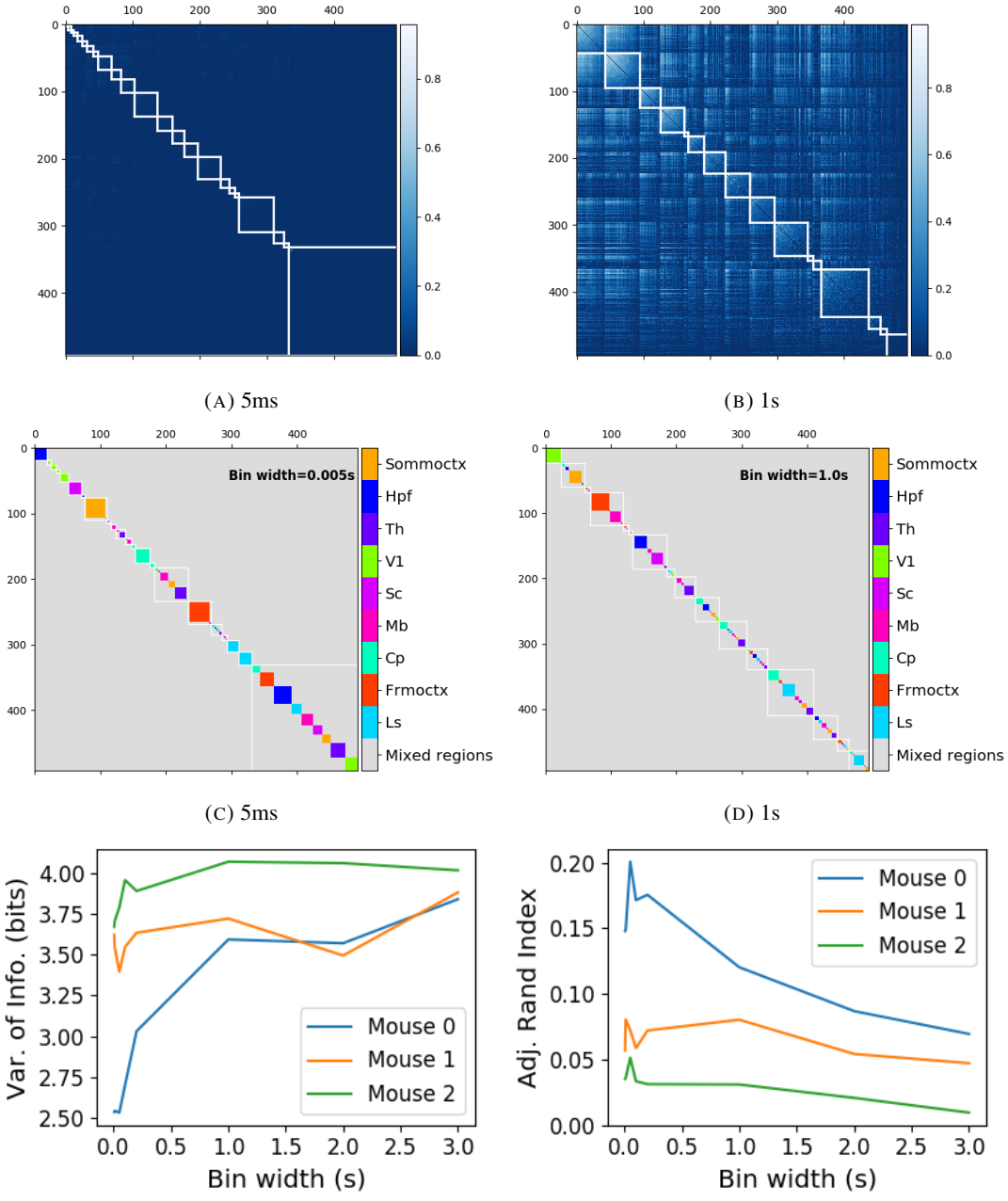
detected rather than any relationship with anatomy.

We also observed that the network noise rejection process rejected some of the cells when applied to the network based on the signal correlations. This means that those cells did not contribute to the additional structure of the network beyond that captured by the sparse weighted configuration model. This is why the matrices in figures 3.13c and 3.13d are smaller than their analogues in figures 3.13a and 3.13b.

The communities detected in the signal correlation based networks indicate that there are groups of cells from different brain regions that react similarly to different activity states. These groups also exist at all timescales from milliseconds to seconds. This indicates that there are subsets of cells in each region that respond to activty states regardless of the timescale of these activties.

3.4.8 Absolute correlations and negative rectified correlations

At the moment, the network noise rejection protocol can only be applied to weighted undirected graphs with non-negative weights. This meant that we had to rectify our correlated networks before applying the network noise rejection and community detection process. We wanted to investigate what would happen if instead of rectifying the correlations, we used the absolute value, or reversed the signs of the correlations and then rectified.



(A) 5ms

(B) 1s

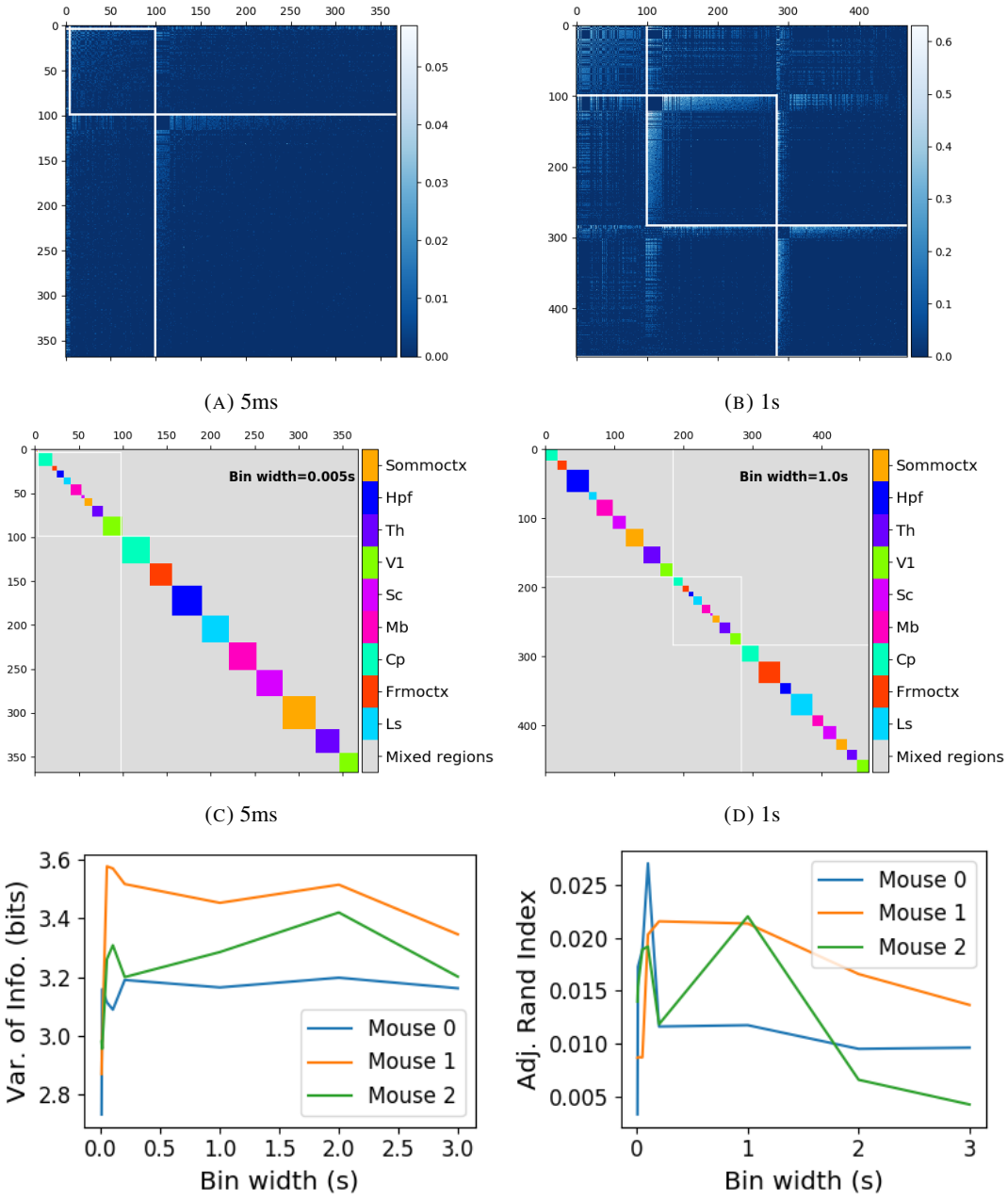
(C) 5ms

(D) 1s

(E) Variation of information

(F) Adjusted Rand index

FIGURE 3.16: (A-B) Absolute correlation matrices with detected communities indicated by white lines. These communities are based on the absolute value of the total correlation between each pair of cells. Those entries within a white square indicate that both of those neurons are in the same community. Matrices shown are for 5ms and 1s time bin widths respectively. Main diagonal entries were set to 0. (C-D) Matrices showing the anatomical distribution of pairs along with their community membership. Regional membership is indicated by the colour bar. (E) Variation of information between the anatomical division of the cells, and the detected communities. (F) Adjusted Rand index between the anatomical division, and the detected communities.



(A) 5ms

(B) 1s

(C) 5ms

(D) 1s

(E) Variation of information

(F) Adjusted Rand index

FIGURE 3.17: (A-B) Sign reversed rectified correlation matrices with detected communities indicated by white lines. Those entries within a white square indicate that both of those neurons are in the same community. Matrices shown are for 5ms and 1s time bin widths respectively. Main diagonal entries were set to 0. (C-D) Matrices showing the anatomical distribution of pairs along with their community membership. Regional membership is indicated by the colour bar. (E) Variation of information between the anatomical division of the cells, and the detected communities. (F) Adjusted Rand index between the anatomical division, and the detected communities.

When we used the absolute value of the correlations, we found very similar results to those shown above for the rectified total correlations and the rectified spike count correlations. We detected more communities using shorter bin widths, and these communities were more similar to the brain's anatomy than those communities detected using a longer bin width (see figure 3.16). The only exception being that we detected more communities. This could indicate that we detected both positively and negatively correlated communities, but we haven't done any further investigation so we cannot say for sure.

When we used the sign reversed rectified correlated networks, we tended to find fewer communities. Each community contained cells from many different anatomical regions, at both long and short bin widths (see figures 3.17a, 3.17b, 3.17c, 3.17d). The communities bore little relation to the anatomical distribution of the cells, this can be seen in figure 3.17f, the values close to zero indicate that the similarity between the two clusterings are around chance level. This indicates that there was not much structure in the negatively correlated networks beyond that captured by the sparse weighted configuration model.

3.5 Discussion

It is well established that the brain uses correlated behaviour in neuronal ensembles to represent the information taken in through sensation (Cohen and Maunsell, 2009; Litwin-Kumar, Chacron, and Doiron, 2012; deCharms and Merzenich, 1996). However, most studies that examine the nature of these correlations in-vivo, study an ensemble of cells from only one ot two brain regions (Cohen and Kohn, 2011; Wierzynski et al., 2009; Patterson et al., 2014; Girard, Hupé, and Bullier, 2001). Furthermore, recent results have shown that behaviour can drive correlated activity in multiple brain regions, including those not normally associated with motor control (Stringer et al., 2019; Gründemann et al., 2019; Allen et al., 2019). In this study, we utilised one of the newly recorded large datasets containing electrophysiological recordings from multiple brain regions simultaneously. We investigated correlated behaviour in these different brain regions and we investigated correlated behaviour between neurons in different regions, during spontaneous behaviour.

A number of studies have found that the timescale of correlated behaviour induced by a stimulus can be modulated by the stimulus structure and behavioural context. For example, the spike train correlations between cells in weakly electric fish are modulated by the spatial extent of the stimulus (Litwin-Kumar, Chacron, and Doiron, 2012), and neurons in the marmoset primary auditory cortex modulate their spike timing (and therefore correlation) in

response to stimulus features without modulating their firing rate (deCharms and Merzenich, 1996). Furthermore, the width of the time bins over which spike counts are measured has been shown to have an effect on the magnitude of those correlations (Cohen and Kohn, 2011). Despite this, very little research has been done comparing correlation measures from the same dataset at different timescales. We investigated this by varying the time bin width used to bin spike times into spike counts from as short as 5ms up to 3s.

In order to further investigate the effect of these correlations at different timescales, we regarded our neuronal ensemble as a weighted undirected graph, where each neuron is represented by a node, and the weight on each edge is the correlation between the neurons connected by that edge. We then applied a novel clustering method from network science (Humphries et al., 2019) to identify communities in these networks. Communities in a network graph refer to sets of nodes that are more strongly connected to each other than the nodes outside of their set. Another way to put this is to say that the nodes in a community are more strongly connected than *expected*. What connection strength might be expected is defined by a null network model. We chose a null network model that matched the sparsity and total strength of our correlation based data networks. So, if two cells were in the same community, those cells were more correlated than would be expected given the correlation strength of their ensemble.

These networks, and the community detection process, were completely agnostic of the anatomical division of the cells in our ensemble. When we compared the detected communities with the anatomical division of the cells using distance and similarity measures for clusterings, we found that the detected communities were more similar to the anatomical division at shorter timescales. That is, when we used a wider time bin to count spikes, and computed pairwise correlations with these spike counts, the correlated communities tended to exist within anatomical regions at shorter timescales, and tended to span anatomical regions at longer timescales. This could reflect localised functional correlations at short time scales rippling outwards across brain regions at longer timescales. The brain may be processing some information quickly locally, and carrying out further, perhaps more detailed, representation over a longer timescale across many regions using the representations that were just built locally.

These changes in communities across timescales could also be driven by the anatomy of the individual cells. For example, it may simply take longer to transmit action potentials over longer distances, hence correlated activity over longer timescales will exist between

anatomical regions, rather than within. However, the switch to almost exclusively multi-regional functional networks at 1s bin widths, rather than a mixture of multi-region, and single-region suggests that the inter-regional correlations either overpower, or inhibit the local correlations. So there may be more at play than just timescales.

We acknowledged that the region spanning correlated communities that we detected at longer time scales could exist due to collating activity driven by distinct spontaneous activities. In order to account for this, we modelled the spike counts as a linear function of the top 500 principal components of a video of the mouse's face filmed simultaneously with the electrophysiological readings. We applied our network noise rejection and community detection process to the weighted undirected networks formed by the spike count correlations (or noise correlations) and the signal correlations that we calculated using our model. For the spike count correlation networks, we found much the same results as for the total correlations as described above. For the signal correlations, the communities detected in these networks bore little relation to the anatomical division of the cells. Recent findings have shown that behavioural data accounts for correlations in many brain regions that would otherwise be dismissed as noise (Stringer et al., 2019), our finding to shows that these correlations are still governed by the timescale division between local communication and across-region communication.

There is a lot of room for further investigation based on this research. For a start, the data that we used here were collected from nine different regions in the mouse brain, but none of these regions were part of the somatosensory cortex. Given that a mouse experiences so much of its environment through its sense of smell, some data from this region would be interesting to investigate. On the same theme, the mice in the experiment from which the data were collected were headfixed and placed on a rotating ball, but were otherwise behaving spontaneously. Had these mice been exposed to a visual, aural, or olfactory stimulus, we could have examined the responses of the cells in the brain regions corresponding to vision, hearing, and olfaction, and compared these responses to the responses from the other brain regions. Furthermore, we could have investigated the interaction between the sets of responses.

Another space for further investigation is the community detection. The algorithm that we used here never detects overlapping communities. But functional communities could indeed have overlaps. Clustering methods that detect overlapping clusters do exist (Baadel, Thabtah, and Lu, 2016). Applying one of those algorithms could yield some interesting results. Also, the community detection algorithm that we used here cannot process graphs with negative

weights, this forced us to separate positive and negative correlations before applying our network noise rejection and community detections process, or use the absolute value of our correlations. A community detection algorithm that can work on weighted undirected graphs with negative weights could yield some interesting results here.

Chapter 4

A simple two parameter distribution for modelling neuronal activity and capturing neuronal association

Abstract

Recent developments in electrophysiological technology have lead to an increase in the size of electrophysiology datasets. Consequently, there is a requirement for new analysis techniques that can make use of these new datasets, while remaining easy to use in practice. In this work, we fit some one or two parameter probability distributions to spiking data collected from a mouse exposed to visual stimuli. We show that the Conway-Maxwell-binomial distribution is a suitable model for the number of active neurons in a neuronal ensemble at any given moment. This distribution fits these data better than binomial or beta-binomial distributions. It also captures the correlated activity in the primary visual cortex induced by stimulus onset more effectively than simply measuring the correlations, at short timescales ($<$ 10ms). We also replicate the finding of Churchland et al (2010) relating to stimulus onset quenching neural variability in cortical areas, and we show a correspondence between this quenching and changes in one of the parameters of the fitted Conway-Maxwell-binomial distributions.

4.1 Introduction

Recent advances in electrophysiological technology, such as 'Neuropixels' probes (Jun et al., 2017) have allowed extracellular voltage measurements to be collected from larger numbers of cells than traditional methods, in multiple brain regions simultaneously, and routinely. These larger datasets require innovative methods to extract information from the data in a reasonable amount of time, 'reasonable' being subjective in this case.

Theoretically, all the information at any given moment in an electrophysiological dataset with n neurons could be captured by calculating the probability distribution for every possible spiking pattern. This would require defining a random variable with 2^n possible values, a task that quickly becomes impossible as n increases. Attempts at approximating this random variable often involve measuring pairwise or higher order correlations (Schneidman et al., 2006; Flach, 2013; Ganmor, Segev, and Schneidman, 2011). But pairwise correlations may not be enough to characterise instantaneous neural activity (Tkačik et al., 2014). Furthermore, these kinds of models tend to ignore the temporal structure of neuronal data, in favour of smaller model size, and scalability.

Higher order correlations would be helpful here, but defining and quantifying these correlations can be tricky (Staude, Grün, and Rotter, 2010). If we use the interaction parameters arising from the exponential family model as measures of higher order correlations, measuring these correlations becomes computationally impractical quite quickly (the number of 'three neuron correlations' to measure scales with $\binom{n}{3}$). In this work, we dispense with measuring correlations directly, and we attempt to characterise correlated behaviour using a parameter in statistical model.

In this work, we examined the ability of simple distributions to model the number of active (spiking) neurons in a neuronal ensemble at any given timepoint. We compared a little-known distribution named the Conway-Maxwell-binomial distribution to the binomial distribution and the beta-binomial distribution. The binomial distribution is a probability distribution over the number of successes is a sequence of independent and identical Bernoulli trials. The beta-binomial distribution is similar, but allows for a bit more flexibility while still being a model for heterogeneity. Similar to the binomial and beta-binomial, the Conway-Maxwell-binomial distribution is a probability distribution over the number of successes in a series of Bernoulli trials, but allows over- and under-dispersion relative to the binomial distribution. This distribution should therefore be a good candidate for our purposes. We found that Conway-Maxwell-binomial distribution was usually the best candidate of the three that

we examined.

We also observed some interesting changes in the number of active neurons in the primary visual cortex and hippocampus at stimulus onset and some changes in this activity in the thalamus which were sustained for the full duration of the stimulus presentation. This let us know that there were some responses to model.

We found that fitting a Conway-Maxwell-binomial distribution was a better method of capturing association between neurons than measuring the spike count correlation for the short time bins that we used ($<$ 10ms).

Finally, we also wanted to investigate parallels between the parameters of the Conway-Maxwell-binomial distribution and quantities that have been established as relevant to sensory processing. So, we replicated the findings made by Churchland et al. (2010) relating to a reduction in neural variability at stimulus onset in the macaque cortical regions, but for data taken from the mouse primary visual cortex. We compared these findings to the values of the fitted Conway-Maxwell-binomial distribution parameters.

4.2 Data

We used data collected by Nick Steinmetz and his lab 'CortexLab at UCL' (Steinmetz, Carandini, and Harris, 2019). The data can be found online [1] and are free to use for research purposes.

Two 'Phase3' Neuropixels (Jun et al., 2017) electrode arrays were inserted into the brain of an awake, head-fixed mouse for about an hour and a half. These electrode arrays recorded 384 channels of neural data each at 30kHz and less than 7μV RMS noise levels. The sites are densely spaced in a 'continuous tetrode'-like arrangement, and a whole array records from a 3.8mm span of the brain. One array recorded from visual cortex, hippocampus, and thalamus, the other array recorded from motor cortex and striatum. The data were spike-sorted automatically by Kilosort and manually by Nick Steinmetz using Phy. In total 831 well-isolated individual neurons were identified.

4.2.1 Experimental protocol

The mouse was shown a visual stimulus on three monitors placed around the mouse at right angles to each other, covering about $\pm$135 degrees azimuth and $\pm$35 degrees elevation.

[1] http://data.cortexlab.net/dualPhase3/

The stimulus consisted of sine-wave modulated full-field drifting gratings of 16 drift directions ($0°, 22.5°, \ldots, 337.5°$) with 2Hz temporal frequency and 0.08 cycles/degree spatial frequency displayed for 2 seconds plus a blank condition. Each of these 17 conditions were presented 10 times in a random order across 170 different trials. There were therefore 160 trials with a drifting-grating visual stimulus present, and 10 trials with a blank stimulus.

4.3 Methods

4.3.1 Binning data

We converted the spike times for each cell into spike counts by putting the spike times into time bins of a given 'width' (in milliseconds). We used time bins of 1ms, 5ms, and 10ms. We used different time bin widths to assess the impact of choosing a bin width.

4.3.2 Number of *active* neurons

To count the number of active neurons in each neuronal ensemble, we split the time interval for each trial into bins of a given width. We counted the number of spikes fired by each cell in each bin. If a cell fired *at least* one spike in a given bin, we regarded that cell as active in that bin. We recorded the number of active cells in every bin, and for the purposes of further analysis, we recorded each cell's individual spike counts.

It should be noted that when we used a bin width of 1ms, the maximum number of spikes in any bin was 1. For the wider time bins, some bins had spike counts greater than 1. Consequently when using a bin width of 1ms, the number of active neurons and the total spike count of a given bin were identical. But for wider bin widths, the total spike count was greater than the number of active neurons.

So for the 1ms bin width, the activity of a neuron and the number of spikes fired by that neuron in any bin can be modelled as a Bernoulli variable. But for wider time bins, only the activity can be modelled in this way.

4.3.3 Moving windows for measurements

When taking measurements (e.g. moving average over the number of active neurons) or fitting distributions (eg. the beta binomial distribution) we slid a window containing a certain number of bins across the data, and made our measurements at each window position. For example, when analysing 1ms bin data, we used a window containing 100 bins, and we slid

Bin width (ms)	Window size (bins)	Window size (ms)	Windows per trial
1ms	100	100ms	296
5ms	40	200ms	286
10ms	40	400ms	266

TABLE 4.1: Details of the different bin width and analysis window sizes used when binning spike times, and analysing those data.

the window across the time interval for each trial moving 10 bins at a time. So that for 3060ms of data, we made 296 measurements.

For the 5ms bin width data, we used windows containing 40 bins, and slid the window 2 bins at a time when taking measurements.

For the 10ms bin width data, we used windows containing 40 bins, and slid the window 1 bin at a time when taking measurements (see table 4.1 for concise details).

By continuing to use windows containing 40 bins, we retained statistical power but sacrificed the number of measurements taken.

There was an interval between each trial with a grey image in place of the moving bar stimulus. This interval varied in time. But we included some of this interval when recording the data for each trial. We started recording the number of active neurons, and the number of spikes from each neuron from 530ms before each trial until 1030ms after each trial. This way, we could see the change in our measurements at the onset of a stimulus and the end of stimulus presentation.

As mentioned in section 4.3.2, we recorded the number of active neurons in each bin, and the spike count for each neuron in each bin. The measurements we took using these data in each window were as follows:

Moving average The average number of active cells in each window.

Moving variance The variance of the number of active cells in each window.

Average correlation We measured the correlation between the spike counts of each pair of cells in the ensemble, and took the average of these measurements.

Binomial p We fitted a binomial distribution to the data in each window and recorded the fitted probability of success, p in each case.

Beta-binomial α, β We fitted a beta-binomial distribution to the data in each window, and recorded the values of the fitted shape parameters, α and β, of each distribution.

Conway-Maxwell-binomial distribution p, ν We fitted a Conway-Maxwell-binomial distribution to the data in each window, and recorded the fitted values of p and ν for each distribution.

Log-likelihoods We also recorded the log-likelihood of each of the fitted distributions for each window.

4.3.4 Fano factor

The *Fano factor* of a random variable is defined as the ratio of the variable's variance to its mean.

$$F = \frac{\sigma^2}{\mu} \tag{4.1}$$

We measured the Fano factor of the spike count of a given cell by measuring the mean and variance of the spike count across trials, and taking the ratio of those two quantities. When calculated in this way the Fano factor can be used as a measure of neural variability that controls for changes in the firing rate. This is similar to the calculation used in (Churchland et al., 2010).

4.3.5 Probability Distributions suitable for modelling ensemble activity

We present here three different probability distributions that could be suitable to model the number of active neurons in an ensemble. Each distribution has the set $\{0, \ldots, n\}$ as its support, where n is the number of neurons in the ensemble. These are simple distributions with either two or three parameters each. However, we regard n as known when using these distributions for modelling, so in effect each distribution has either one or two free parameters.

Association

Association between random variables is similar to the correlation between random variables but is more general in concept. The correlation coefficient is a measure of association; and association doesn't necessarily have a mathematical definition like correlation does. Essentially, an association between two random variables is a dependency of any kind. Positively associated variables tend to take the same value, and negatively associated variables tend to take different values. In this research, we work with probability distributions of the number of successes in a set of Bernoulli trials. These Bernoulli variables may or may not be associated.

A probability distribution over the number of successes in n Bernoulli trials, where the Bernoulli variables may be associated, could constitute a good model for the number of active neurons in an ensemble of n neurons. As long as the observation period is divided into time bins short enough so that any neuron is unlikely to fire more than spike in any time bin.

Binomial distribution

The binomial distribution is a two parameter discrete probability distribution that can be thought of as a probability distribution the number of successes from n independent Bernoulli trials, each with the same probability of success. The parameters of the binomial distribution are n the number of trials, and $0 \leq p \leq 1$, the probability of success for each of these trials. A random variable with the binomial distribution can take values from $\{0, \ldots, n\}$. The probability mass function of the distribution is

$$P(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{4.2}$$

As a model for the activity of a neuronal ensemble, the main problem with the binomial distribution is that it treats each neuron, represented as a Bernoulli trial, as independent. It is well know that neurons are not independent, and that correlated behaviour between neurons is vital for representing sensory information (Cohen and Maunsell, 2009). The binomial distribution falls short in this regard, but it is useful as performance benchmark when assessing the performance of other models.

Beta-binomial distribution

The beta distribution is the conjugate distribution of the binomial distribution. The beta-binomial distribution is the combination of the beta distribution and the binomial distribution, in that the probability of success for the binomial distribution is sampled from the beta distribution. This allows the beta-binomial distribution to capture some over dispersion relative to the binomial distribution.

The beta-binomial distribution is a three parameter distribution, n the number of Bernoulli trials, and $\alpha \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}_{>0}$ the shape parameters of the beta distribution. The probability mass function for the beta-binomial distribution is

$$P(k; n, \alpha, \beta) = \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)} \tag{4.3}$$

where $B(\alpha, \beta)$ is the beta function.

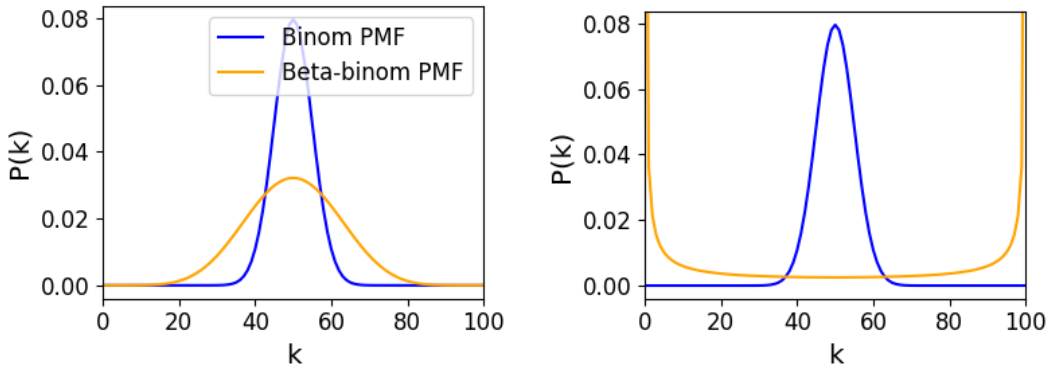
This probability distribution can be reparametrised in a number of ways. One of which defines new parameters π and ρ by

$$\pi = \frac{\alpha}{\alpha + \beta} \tag{4.4}$$

$$\rho = \frac{1}{\alpha + \beta + 1} \tag{4.5}$$

This reparametrisation is useful because π acts as a location parameter analogous to the p parameter of a binomial distribution. A value of $\rho > 0$ indicates over-dispersion relative to a binomial distribution.

As a model for the activity of a neuronal ensemble, the beta-binomial distribution is more suitable than a binomial distribution because the over-dispersion of the beta-binomial distribution can be used to model positive association between the neurons. An extreme example of this over-dispersion/positive association can be seen in figure 4.1b. In this figure, the neurons are positively associated and so tend to take the same value, consequently the probability mass of the beta-binomial distribution builds up close to $k = 0$ and $k = n$. It is worth noting that the location parameter for each distribution has the same value, $p = \pi = 0.5$.



(A) $n = 100, p = 0.5, \alpha = \beta = 10$ (B) $n = 100, p = 0.5, \alpha = \beta = 0.3$

FIGURE 4.1: Figures showing the over-dispersion possible for a beta-binomial distribution relative to a binomial distribution. Parameters are shown in the captions.

Conway-Maxwell-binomial distribution

The Conway-Maxwell-binomial distribution (COMb distribution) is a three parameter generalisation of the binomial distribution that allows for over dispersion and under dispersion

relative to the binomial distribution. The parameters of the distribution are n the number of Bernoulli trials, and two shape parameters $0 \leq p \leq 1$, and $\nu \in \mathbb{R}$.

The probability mass function of the COMb distribution is

$$P(k; n, p, \nu) = \frac{1}{S(n, p, \nu)} \binom{n}{k}^{\nu} p^k (1-p)^{n-k} \tag{4.6}$$

where

$$S(n, p, \nu) = \sum_{j=0}^{n} \binom{n}{k}^{\nu} p^j (1-p)^{n-j} \tag{4.7}$$

The only difference between this PMF and the PMF for the standard binomial is the introduction of ν and the consequent introduction of the normalising function $S(n, p, \nu)$.

Indeed, if $\nu = 1$ the COMb distribution is identical to the binomial distribution with the same values for n and p. We can see in figure 4.2d that the KL-divergence $D_{KL}(P_{COMb}(n, p, \nu)||P_{Bin}(n, p)) = 0$ along the line where $\nu = 1$. The analytical expression for the divergence is

$$D_{KL}(P_{COMb}(k; n, p, \nu)||P_{Bin}(k; n, p)) = (\nu - 1) E_{P_{COMb}(k;n,p,\nu)} \left[\log \binom{n}{k} \right] \tag{4.8}$$

$$- \log S(n, p, \nu) \tag{4.9}$$

At $\nu = 1$, we have $S(n, p, 1)$ which is just the sum over the binomial PMF, so $S(n, p, 1) = 1$ and therefore $D_{KL}(P_{COMb}(n, p, \nu)||P_{Bin}(n, p)) = 0$.

If $\nu < 1$ the COMb distribution will exhibit over-dispersion relative to the binomial distribution. If $p = 0.5$ and $\nu = 0$ the COMb distribution is the discrete uniform distribution, and if $\nu < 0$ the mass of the COMb distribution will tend to build up near $k = 0$ and $k = n$. This over-dispersion represents positive association in the Bernoulli variables. An example of this over-dispersion can be seen in figure 4.2b.

If $\nu > 1$ the COMb distribution will exhibit under-dispersion relative to the binomial distribution. The larger the value of ν the more probability mass will build up at $n/2$ for even n, or at $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ for odd n. This under-dispersion represents negative association in the Bernoulli variables. An example of this under-dispersion can be seen in figure 4.2a.

It should be noted that the p parameter of the COMb distribution does not correspond to the mean of the distribution, as is the case for the binomial p parameter, and beta-binomial π parameter. That is, the COMb p parameter is not a location parameter. An illustration of this can be seen in figure 4.2c. This is because an interaction between the p and ν parameters skews the mean. There is no analytical expression for the mean of the COMb distribution.

ν	Relative dispersion	Association between neurons/variables
< 1	over	positive
1	none	none
> 1	under	negative

TABLE 4.2: Relative dispersion of the COMb distribution, and association between Bernoulli variables as represented by the value of the ν parameter.

Since the COMb distribution has the potential to capture positive and negative associations between the neurons/Bernoulli variables, it should be an excellent candidate for modelling the number of active neurons in a neuronal ensemble.

We wrote a dedicated Python package to enable easy creation and fitting of COMb distribution objects. The format of the package imitates the format of other distribution objects from the `scipy.stats` Python package. The COMb package can be found here:
https://github.com/thomasjdelaney/Conway_Maxwell_Binomial_Distribution

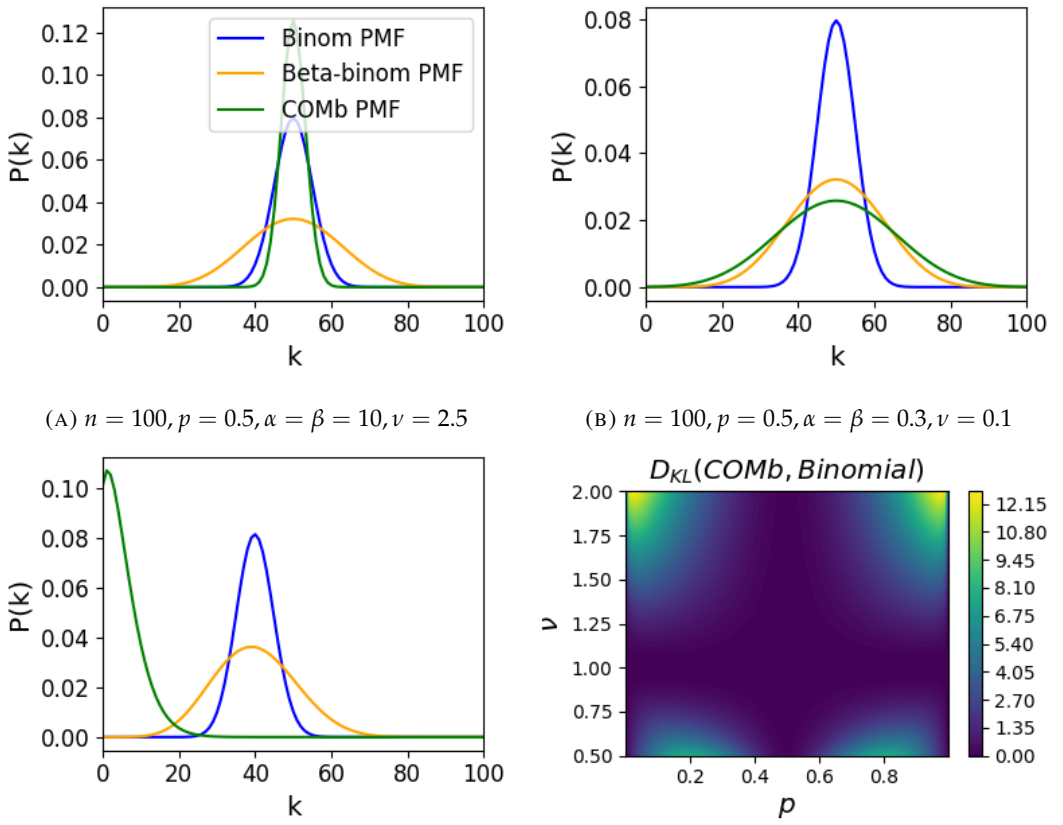
4.3.6 Fitting

We fitted binomial, beta-binomial, and Conway-Maxwell-binomial (COMb) distributions to the neural activity in each of the overlapping windows covering each trial. To fit the distributions we minimised the appropriate negative log likelihood function using the data from the window.

There is an analytical solution for maximum likelihood estimate of the binomial distribution's p parameter.

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{N} k_i \tag{4.10}$$

We minimised the negative log likelihood function of the beta-binomial distribution numerically. We calculated the negative log likelihood for a sample directly, by taking the sum of the log of the probability mass function for each value in the sample. We minimised the negation of that function using the `minimise` function of the `scipy.optimize` Python package.



(A) $n = 100, p = 0.5, \alpha = \beta = 10, \nu = 2.5$

(B) $n = 100, p = 0.5, \alpha = \beta = 0.3, \nu = 0.1$

(C) $n = 100, p = 0.4, \alpha = 10, \beta = 15, \nu = 0.1$

(D) KL-Divergence as a function of p and ν. $n = 100$.

FIGURE 4.2: Figures showing (A) the under-dispersion and (B) over-dispersion permitted by the COMb distribution relative to a binomial distribution. (C) illustrates that the p parameter of the COMb distribution does not correspond to the mean of the distribution, as it does for the binomial and beta-binomial distributions. (D) shows a heatmap for the value of the Kullback-Liebler divergence between the COMb distribution and the standard binomial distribution with same value for n, as a function of p and ν. The point of this figure is to give the reader a sense of how the values of p and ν influence the difference between the COMb distribution and the binomial distribution. The divergence is smallest when $\nu \approx 1$, or when $p \approx 0.5$. When $\nu = 1$, the PMF for the COMb distribution is the same as the PMF for the binomial distribution. When $p = 0.5$ the mass of the distribution is centred around $n/2$ for both the COMb and the binomial distribution. The difference between the two distributions is controlled by the ν parameter. The further the p and ν parameters are from 0.5 and 1 respectively, the greater the difference between the COMb distribution and the binomial distribution. Parameters for all figures are shown in the captions.

The log likelihood function of the COMb distribution given some sample $\{k_1, \ldots, k_N\}$ is

$$\ell(p, \nu | k_1, \ldots, k_N) = N\left[n \log(1-p) - \log S(n, p, \nu)\right] \tag{4.11}$$

$$+ \log \frac{p}{1-p} \sum_{i=1}^{N} k_i \tag{4.12}$$

$$+ \nu \sum_{i=1}^{N} \log \binom{n}{k_i} \tag{4.13}$$

We minimised the negation of this function using numerical methods. More specifically, we used the `minimise` function of the `scipy.optimize` Python package.

4.3.7 Goodness-of-fit

After fitting, we measured the goodness-of-fit of each model/distribution with their log likelihood. We calculated this directly using the `logpmf` functions of the distribution objects in Python.

4.4 Results

We defined a neuron as *active* in a time bin if it fires at least one spike during the time interval covered by that bin. We measured the number of active neurons in the primary visual cortex of a mouse in 1ms bins across 160 trials of a moving bar visual stimulus. We then slid a 100ms window across these 1ms bins taking measurements, and fitting distributions along the way. We did the same for neurons in the thalamus, hippocampus, striatum, and motor cortex. We repeated the analysis for 5ms time bins with 40 bin windows, and 10ms time bins with 40 bin windows.

4.4.1 Increases in mean number of active neurons and variance in number of active neurons at stimulus onset in some regions

We measured the average number of active neurons, and the variance of the number of active neurons in a 100ms sliding window starting 500ms before stimulus onset until 1000ms after stimulus onset. We found differences in the response across regions. There were no observed changes in response to the stimulus in the motor cortex or the striatum. The changes in the other regions are detailed below.

Primary visual cortex

We found a transient increase in both the average and variance of the number of active neurons at stimulus onset, followed by a fall to pre-stimulus levels, followed by another transient increase (see figure 4.3). The oscillation in both of these measurements appear to reflect the frequency of the stimulus (see Data section 4.2.1), and it is known that stimulus structure can influence response structure(Litwin-Kumar, Chacron, and Doiron, 2012). We see a similar but lower amplitude oscillation at the end of the stimulus presentation.

Hippocampus

In the hippocampus we observed a transient increase in the average number of active neurons and in the variance of the number of active neurons at stimulus onset (see figure 4.4). The increase lasted about 125ms, and the subsequent fall to baseline took the a similar amount of time.

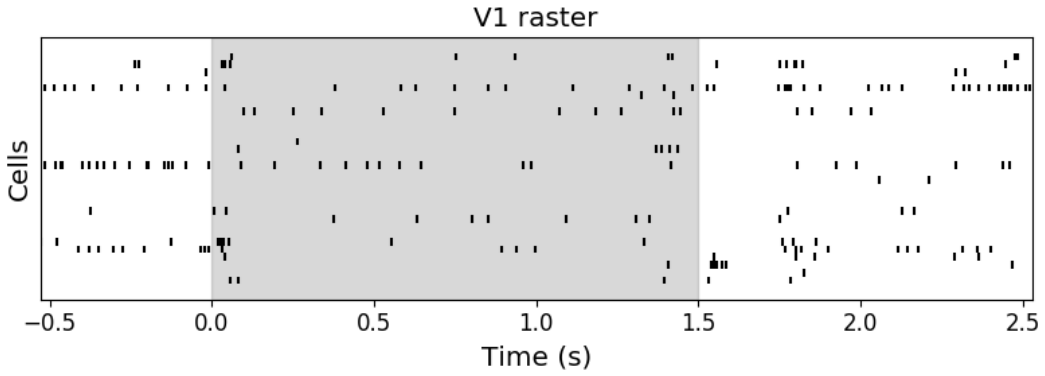
Thalamus

In the thalamus we observed a transient increase in the both the average and variance of the number of active neurons on stimulus onset, followed by a fall to pre-stimulus levels, followed by a sustained increase until the stimulus presentation ends.

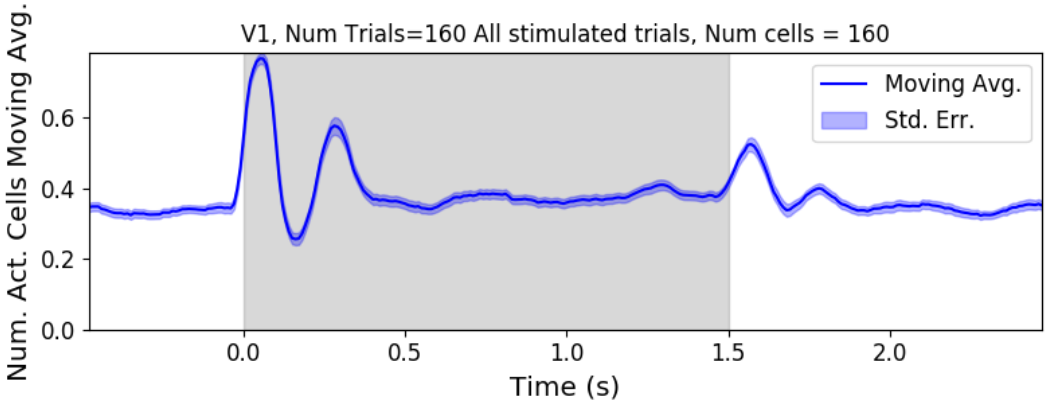
As one you might expect for a visual stimulus, the change in the average number of active neurons was greatest in the primary visual cortex. In this region, this quantity doubled on stimulus onset. In contrast, in the hippocampus and the thalamus, the average number of active neurons only increased by a fraction of the unstimulated baseline value. The duration of the response in V1 and the hippocampus at stimulus onset was $300 - 400$ms, but the response in the thalamus appeared to last for the duration of stimulus presentation. The V1 also showed a change in the average number of active neurons at stimulus end. The change was similar to that observed at stimulus onset, but smaller in magnitude (see figures 4.3, 4.4, and 4.5)

4.4.2 Conway-Maxwell-binomial distribution is usually a better fit than binomial or beta-binomial

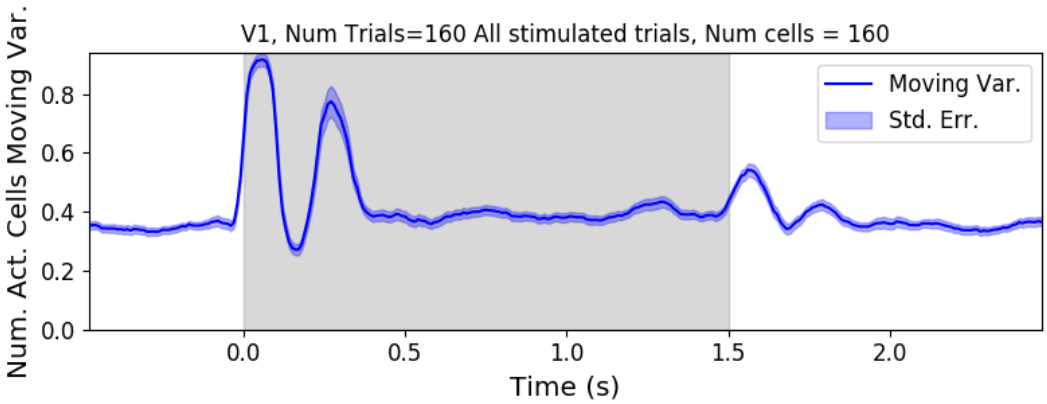
Since the Conway-Maxwell-binomial distribution has not been fitted to neuronal data before, it is not clear that it would be a better fit than the binomial or beta-binomial distributions. In order to find out which parametric distribution was the best fit for the largest proportion



(A) Raster.

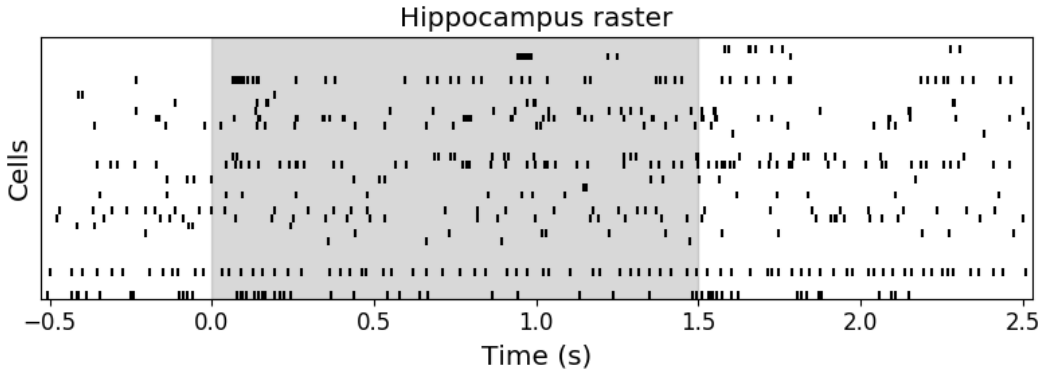


(B) Moving average.

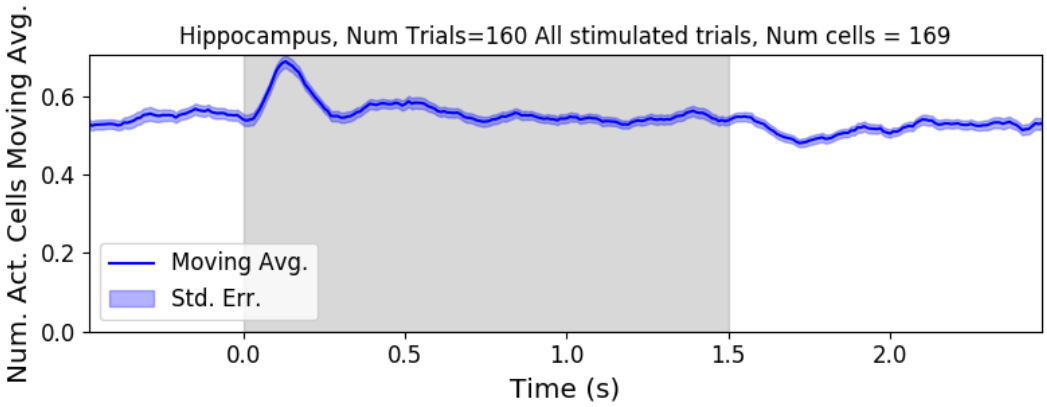


(C) Moving variance.

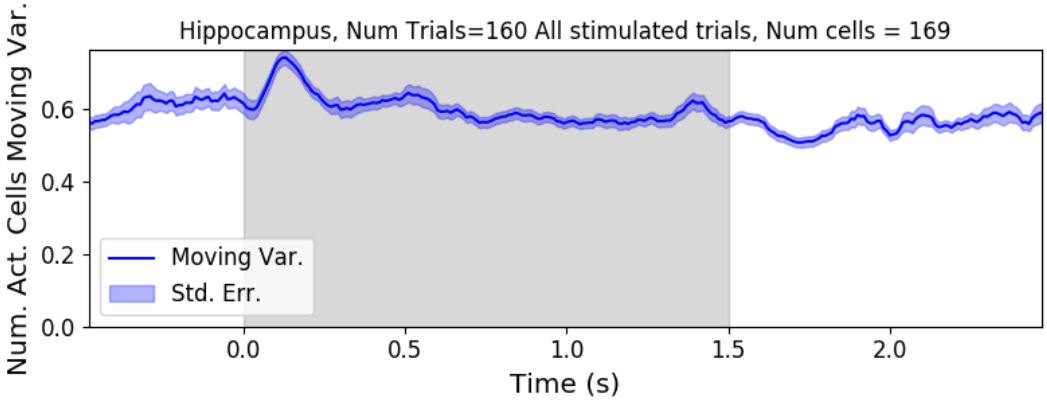
FIGURE 4.3: (A) Raster plot showing the spikes fired by 33 randomly chosen neurons in the primary visual cortex. (B-C) (B) average and (C) variance of the number of active neurons, measured using a sliding window 100ms wide, split into 100 bins. The midpoint of the time interval for each window is used as the timepoint (x-axis point) for the measurements using that window. The grey shaded area indicates the presence of a visual stimulus. The opaque line is an average across the 160 trials that included a visual stimulus of any kind. We can see a transient increase in the average number of active neurons and the variance of this number, followed by a fluctuation and another increase.



(A) Raster.

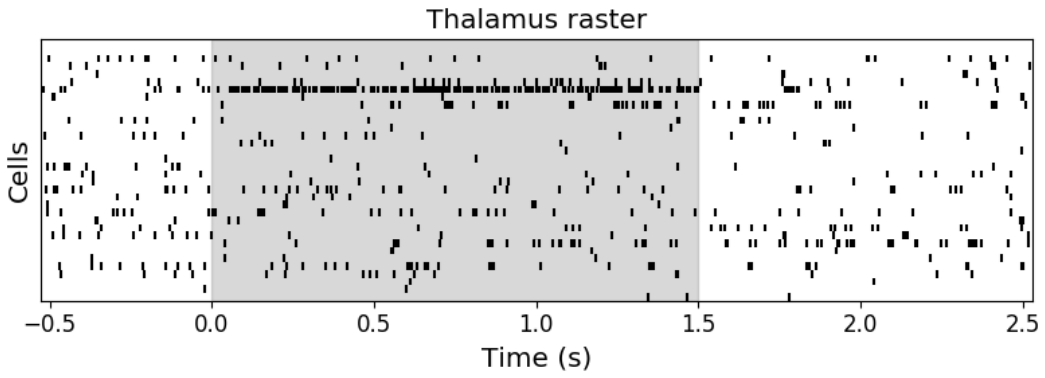


(B) Moving average.

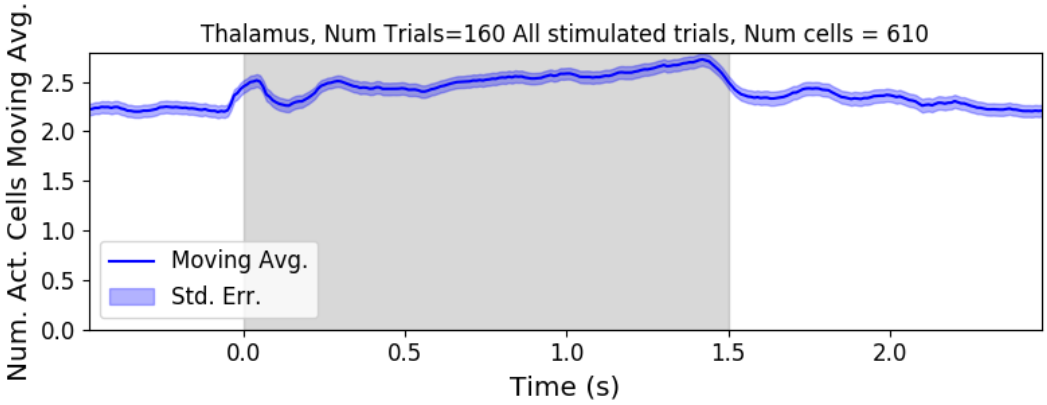


(C) Moving variance.

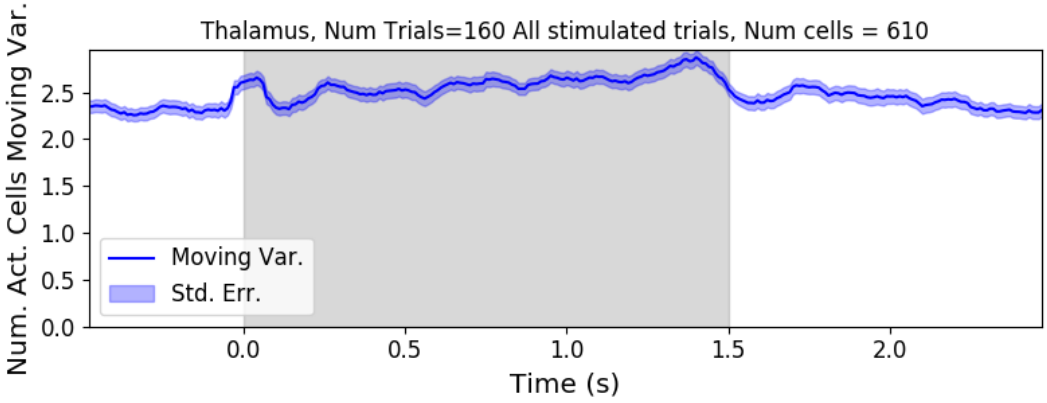
FIGURE 4.4: (A) Raster plot showing the spikes fired by 33 randomly chosen neurons in the hippocampus. (B-C) (B) average and (C) variance of the number of active neurons, measured using a sliding window 100ms wide, split into 100 bins. The midpoint of the time interval for each window is used as the timepoint (x-axis point) for the measurements using that window. The grey shaded area indicates the presence of a visual stimulus. The opaque line is an average across the 160 trials that included a visual stimulus of any kind. We can see a transient increase in the average number of active neurons and the variance of this number at stimulus onset.



(A) Raster.



(B) Moving average.

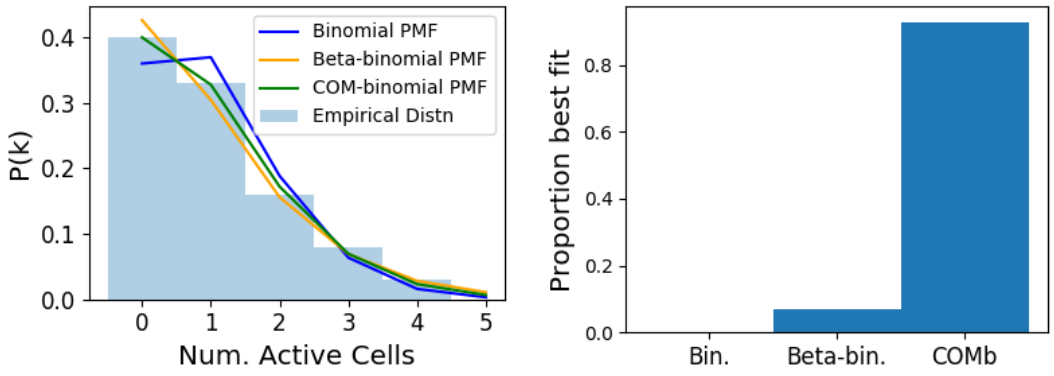


(C) Moving variance.

FIGURE 4.5: (A) Raster plot showing the spikes fired by 33 randomly chosen neurons in the thalamus. (B-C) (B) average and (C) variance of the number of active neurons, measured using a sliding window 100ms wide, split into 100 bins. The midpoint of the time interval for each window is used as the timepoint (x-axis point) for the measurements using that window. The grey shaded area indicates the presence of a visual stimulus. The opaque line is an average across the 160 trials that included a visual stimulus of any kind. We can see in immediate increase at stimulus onset, a subsequent fall, and another sustained increased until the stimulus presentation ends.

of our data, we fit a binomial, a beta-binomial, and a Conway-Maxwell-binomial (COMb) distribution to each window for each bin width, and each region. Then we assessed the goodness-of-fit of each distribution by calculating the log-likelihood of each fitted distribution using the associated sample. We measured the proportion of samples for which each distribution was the best fit, for each bin width value and each region.

We found that the COMb distribution was the best fit for most of the samples regardless of bin width or region. The bin width had an effect on the number of samples for which the COMb distribution was the best fit. The results are summarised in table 4.3. For a bin width of 1ms, the COMb distribution was the best fit for over 90% of samples, the beta-binomial distribution was the best fit for less than 10% of samples, and the binomial distribution was the best fit for less that 1% of samples, across regions. For 5ms bins, the COMb distribution was the best fit for $70-80\%$ of samples, the beta-binomial distribution was the best fit for $20-30\%$ of the samples, and again the binomial distribution was the best fit for less that 1% of samples, across regions. Finally, for 10ms bins, the COMb distribution was the best fit for $53-80\%$ of samples, the beta-binomial distribution was the best fit for $20-47\%$ of the samples, and the binomial distribution was the best fit for less that 0.1% of samples, across regions.



(A) Example of fitted distributions.

(B) Proportion of best fit

FIGURE 4.6: (A) An example of the binomial, beta-binomial, and Conway-Maxwell-binomial distributions fitted to a sample of neural activity. The Conway-Maxwell-binomial distribution is the best fit in this case. The histogram shows the empirical distribution of the sample. The probability mass function of each distribution is indicated by a different coloured line. (B) Across all samples in all trials, the proportion of samples for which each fitted distribution was the the best fit. The Conway-Maxwell-binomial distribution was the best fit for 93% of the samples taken from V1 using a bin width of 1ms.

Bin Width (ms)	Binomial	Beta-binomial	COMb
1ms	$< 1\%$	$< 10\%$	$> 90\%$
5ms	$< 0.1\%$	$20 - 30\%$	$70 - 80\%$
10ms	$< 0.1\%$	$20 - 47\%$	$53 - 80\%$

TABLE 4.3: Proportion of samples for which each distribution was the best fit, grouped by bin width. The COMb distribution is the best fit most of the time.

4.4.3 Relative goodness-of-fit for binomial, beta-binomial, and COMb distributions

In the previous section we showed that the COMb distribution was usually a better option than the binomial or beta-binomial distributions when attempting to fit a distribution to a sample of the number of a active neurons. In this section, we aim to illustrate typically how much better the COMb distribution is.

Log likelihoods of distributions fitted to stimulated and unstimulated responses

We fitted binomial, beta-binomial, and COMb distributions to two windows in each of the 160 trials with a visual stimulus. One of the windows was the last full window before stimulus onset. The other window was the first full window after stimulus onset. We measured the log likelihood for each fitted distribution. For the histograms of these log likelihood values for data from the primary visual cortex using 1ms bin width, see figure 4.7.

Comparing unstimulated to stimulated windows, we observed that the log likelihood values were greater for the unstimulated windows for all three distributions. This shows that the distributions were fitting better to the number of active neurons in the unstimulated windows. This might be due to a greater diversity in the distributions of number of active neurons in the stimulated windows (see figure 4.3). We saw similar results for the primary visual cortex when using 5ms and 10ms bin widths. For other regions, the histograms for unstimulated and stimulated windows were more similar, covering similar ranges of log likelihoods.

We observed a marginal increase in the log likelihoods from binomial to beta-binomial to COMb distributions for both unstimulated and stimulated windows. But, the distribution of the log likelihoods looked quite similar for all three distributions. So, the COMb distribution only fits a little bit better than the other two distributions. We observed similar results when using different time bin widths, and data from different brain regions.

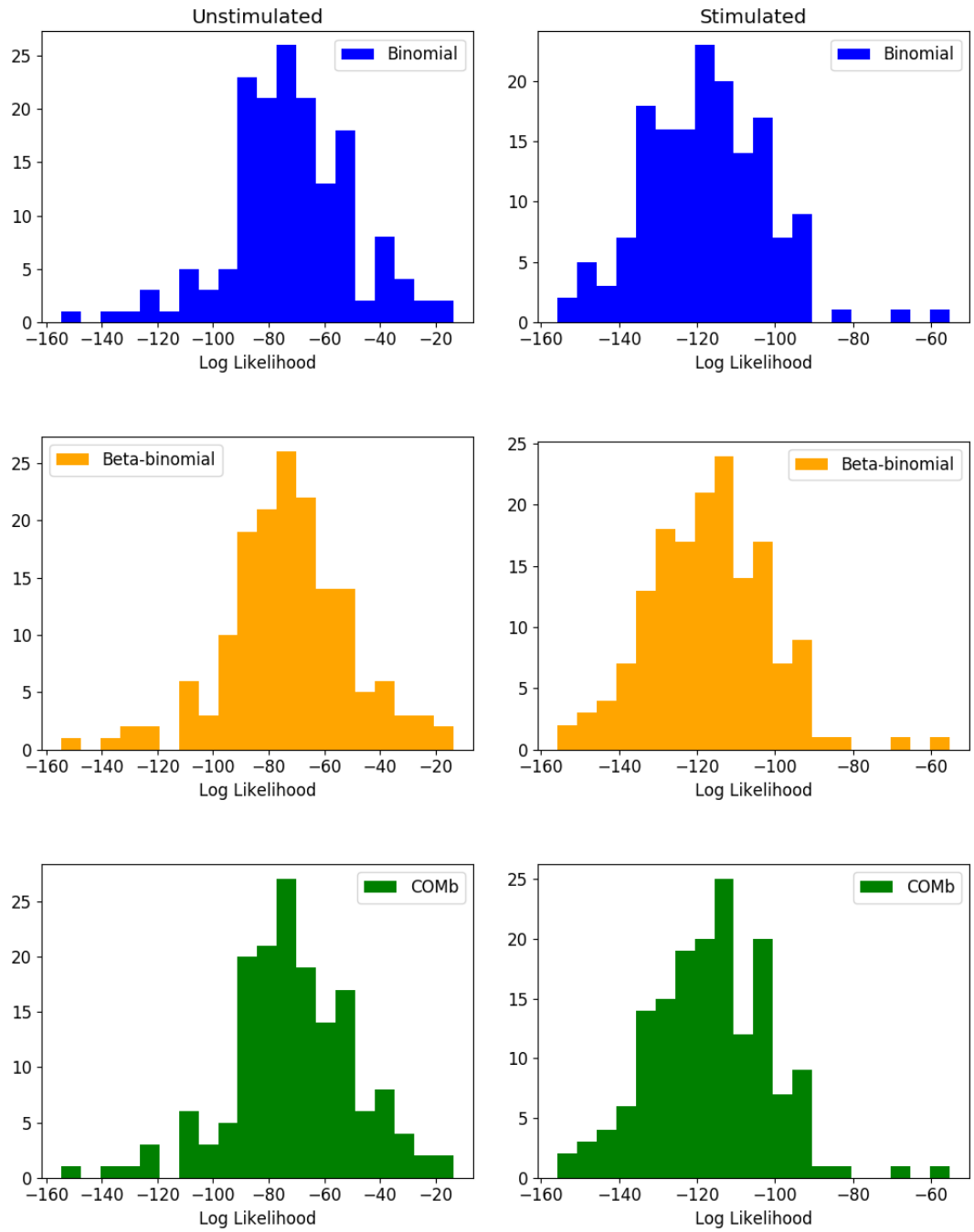


FIGURE 4.7: Number of active neurons data from the primary visual cortex, 1ms bin widths. (Left column) Histograms of log likelihoods for the binomial, beta-binomial, and COMb distributions fitted to windows where no visual stimulus was present. 160 trials. (Right column) Similar histograms for windows where a visual stimulus was present. 160 trials. In both cases, there are marginal increases in the log likelihoods from binomial to beta-binomial to COMb. The log likelihoods are larger for the unstimulated windows.

Distribution of fitted parameters for stimulated and unstimulated responses

We recorded the fitted parameters of each of the three distributions fitted to both the unstimulated and stimulated windows mentioned in section 4.4.3. We noticed an increase in the binomial distribution's parameter for the stimulated windows (figure 4.8 top row). When we used wider bin widths, we noticed the same relative increase from unstimulated to stimulated windows, and the values over which the parameter was distributed increased. The relative increase from unstimulated to stimulated windows was also visible in data from the hippocampus, but not in other regions.

For the beta-binomial distribution, we converted the fitted α and β parameters to π and ρ parameters (see Methods section 4.3.5) before examining their histograms. In this form, the $0 \leq \pi \leq 1$ parameter is a location parameter, and the $\rho > 0$ parameter is a shape parameter than encodes over-dispersion in the distribution relative to a binomial distribution. Comparing distribution of parameters fitted to the unstimulated windows to the distribution of those fitted to the stimulated windows, we observed slight increases in both the π and ρ paramters (figure 4.8 middle row). When we used a wider bin width, the paramter values were distributed across larger absolute values. The results comparing unstimulated to stimulated distributions were similar to the 1ms case. For the other regions, we observed similar results in the hippocampus. All the other regions show little difference between unstimulated and stimulated histograms in a given bin width. As we increased the size of the bin width used the absolute values across which the parameters were distributed increased.

For the COMb distribution, the ν paramter was distributed around 1 for the unstimulated windows. For the stimulated windows ν was more tightly distributed and centred around 0.75. This reflects a positive association between the neurons present immediately after stimulus onset. We also observed that the p parameter was distributed over slightly greater values for the stimulated windows as compared to the unstimulated (figure 4.8 bottom row). For increased bin widths, we observed that the ν parameter was distributed lower than for the 1ms bin width case. The reduction in ν at for the stimulated windows was still present. We observed this drop in the ν parameter in the hippocampal data also, albeit to a lesser extent than in the primary visual cortex. We didn't observe differences in the the data from the other regions.

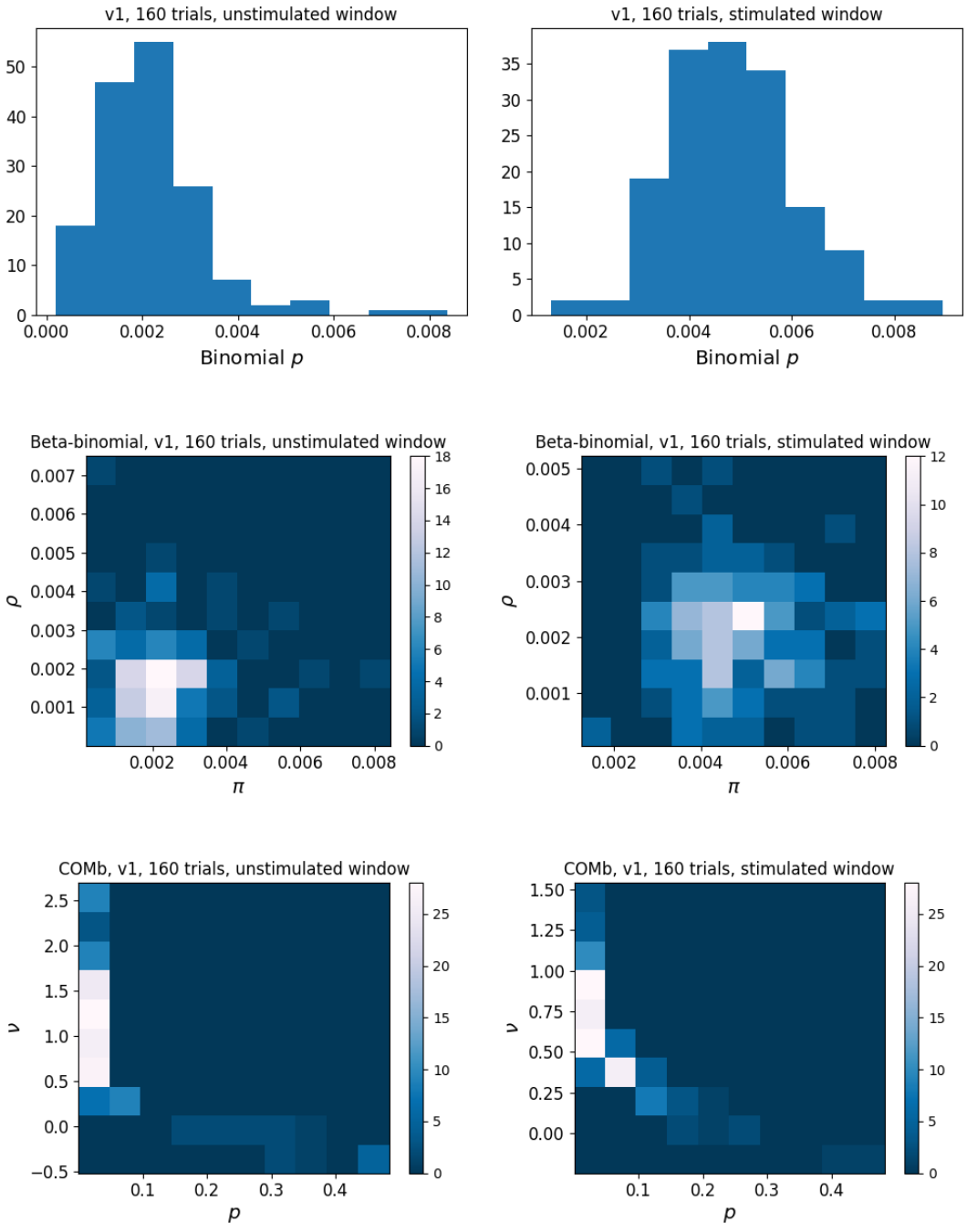


FIGURE 4.8: Histograms of fitted parameters for binomial, beta-binomial, and COMb distributions. Distributions fitted to data from the primary visual cortex, using 1ms bin widths. (Left column) The distributions were fitted a window before the onset of the visual stimulus. (Right column) The distributions were fitted to a window immediately after the onset of the visual stimulus.

Examples of empirical distributions and fitted distributions

In figure 4.9 there are some examples of fitted binomial, beta-binomial, and COMb distributions alongside the empirical distributions to which they are fitted. We can see that these fitted distributions over distributed their probability mass to $P(0)$ (or $P(0)$ and $P(1)$ for the thalamus) and underdistributed their probability mass elsewhere. Also, each of the three fitted distributions look similar in each example. This is in agreement with our observations in section 4.4.3.

4.4.4 Conway-Maxwell-binomial distribution captures changes in association at stimulus onset

We fit a Conway-Maxwell-binomial (COMb) distribution to the number of active neurons in the 1ms time bins in a 100ms sliding window. We also measured the correlation coefficient between the spike counts of all possible pairs of neurons, and took the average of these coefficients. We did this for all the trials with a visual stimulus. We observed a reduction in the COMb distribution's ν parameter at stimulus onset from around 1 to between 0 and 1 (see figure 4.10a). A value of ν less than 1 indicates positive association between the neurons (see section 4.3.5). We might expect to see this positive association reflected in the correlation coefficients, but this is not the case. We see no change in the time series of average correlation measures at stimulus onset.

This may be due to the very short time bin we used in this case. We know that using small time bins can artificially reduce correlation measurements (Cohen and Kohn, 2011). In this case, fitting the COMb distribution may be a useful way to measure association in a neuronal ensemble over very short timescales ($<$ 10ms).

4.4.5 Replicating stimulus related quenching of neural variability

Churchland et al. (2010) inspected the effect of a stimulus on neural variability. One of the measures of neural variability that they employed was the Fano factor of the spike counts of individual cells (see section 4.3.4). They found a reduction in neural variability as measured by the Fano factor in various cortical areas in a macaque at the onset of various visual stimuli, or a juice reward (Churchland et al., 2010).

We measured the Fano factor of the spike count of each cell in each brain region, during each trial. We measured the mean and standard error of these Fano factors from 500ms before stimulus onset until 1000ms after stimulus end. For the primary visual cortex, we found a

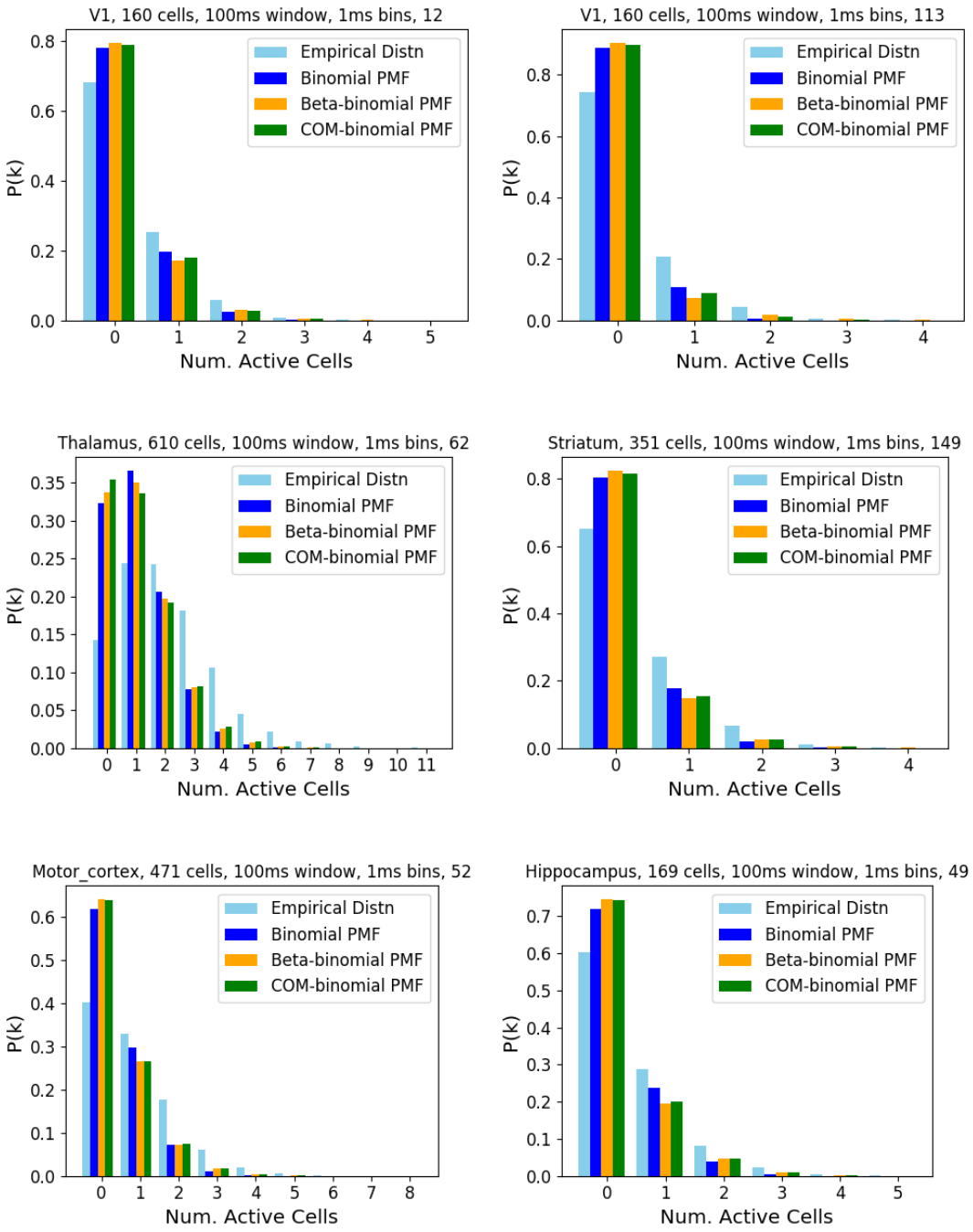
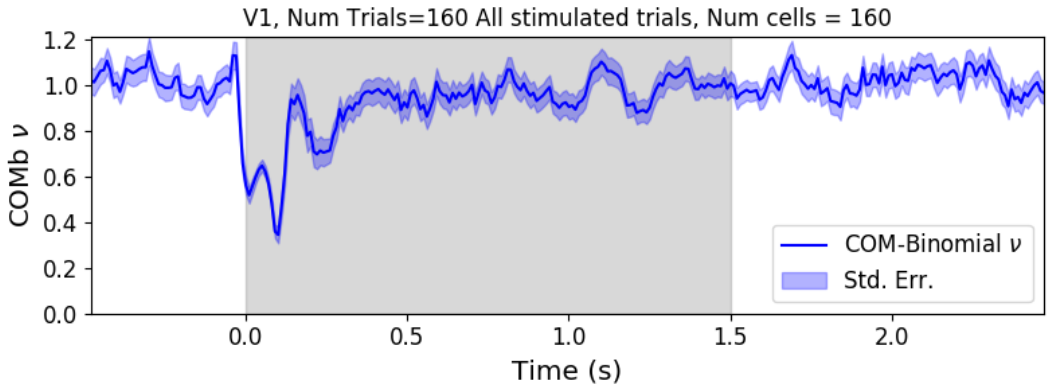
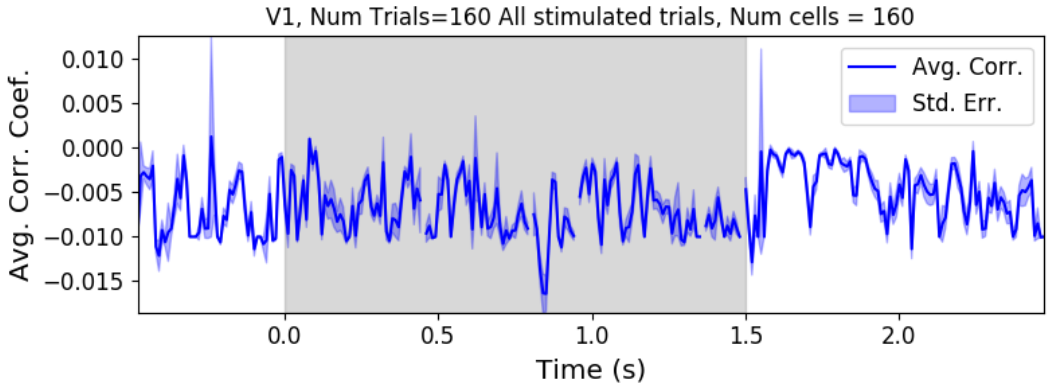


FIGURE 4.9: Examples of empirical and fitted distributions. At least one example from each of the 5 brain regions from which we had data.



(A) COMb ν parameter.

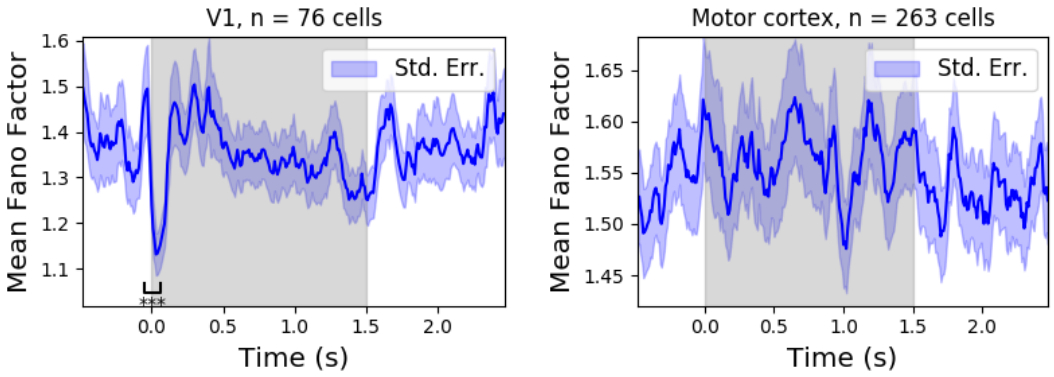


(B) Average correlation coefficient.

FIGURE 4.10: (A) We fit a Conway-Maxwell-binomial distribution to the number of active neurons in 1ms time bins of a 100ms sliding window. We did this for all trials with a visual stimulus and took the average across those trials. We see a transient drop in value for the distribution's ν parameter at stimulus onset. This shows an increase in positive association between the neurons. (B) We measured the correlation coefficient between the spike counts of all possible pairs of neurons in the same sliding window. The took the average of those coefficients. We also did this for every visually stimulated trial, and took the average across trials. The increase in positive association is not reflected with an increase in average correlation.

transient reduction in the Fano factor immediately after stimulus onset. We used a Mann-Whitney U test to check that the Fano factors measured in a window starting at stimulus onset and ending 100ms later were significantly lower than the factors measured in a window ending at stimulus onset ($p < 0.001$, see figure 4.11a). We did not get this statistically significant result in any other region.

Our findings agree with those of Churchland et al. for the primary visual cortex. However Churchland also found a reduction in the Fano factor in the dorsal premotor cortex (PMd) at stimulus onset. Our measurements from the mouse motor cortex show no change at stimulus onset (see figure 4.11b). This could indicate some difference in the functionality of the motor cortex in a macaque and the motor cortex of a mouse.



(A) Primary visual cortex.

(B) Motor cortex.

FIGURE 4.11: (A) The mean Fano factor of the spike counts of the cells in the primary visual cortex. Means were taken across cells first, then across trials. There was a significant decrease in the Fano factors immediately after stimulus onset. (B) The mean Fano factor of the spike counts of the cells in the motor cortex. No significant change in measurements at any point.

Similar to these findings in the Fano factor, we found a reduction in the ν parameter of the COMB distribution on stimulus onset in V1 (figure 4.10a) and in no other region from which we had data. Specifically, the ν parameter reduced from around 1, to between 1 and 0. This represents a change from no association between the neurons, to a positive association. It is possible that this positive association may be responsible for the reduction in the Fano factor.

4.4.6 Effects of greater bin widths

Using a greater bin width (10ms) affected the scale and shape of the most of the measurements taken (described in section 4.3.3). The average correlation coefficient is the exception to this. The 10ms bin width is still so small that the average correlation coefficient was also small (similar scale as figure 4.10b).

Using the greater bin width acted as a low pass filter on the other measurements taken and the parameters of the fitted distributions. For example compare figure 4.10a to figure 4.12a. Note also that the ν parameter varies between 0.6 and 0.2 when using the 10ms bin width. This indicates some positive association between the neurons at all times, even when not stimulated or when adapted to the stimulus. This may be caused by the wider bin width resulting in more neurons classified as active. The change is association at stimulus onset is still captured by the change in the ν parameter.

The mean number of active neurons was about 10 times greater for a 10ms bin width compared to a 1ms bin width. The variance in the number of active neurons was also greater

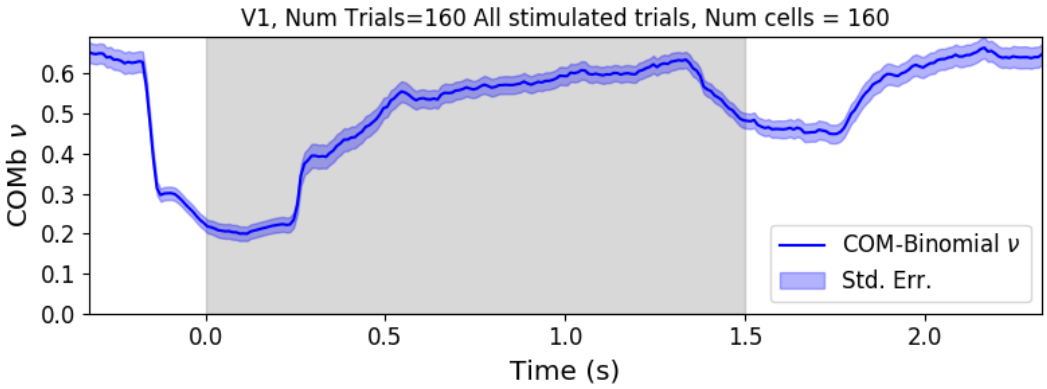
for the wider bin width. This caused a change in the scale of the Fano factor (see figure 4.12b compared to figure 4.11a). Also, the reduction in the mean Fano factor at stimulus onset is not significant when using the 10ms bin width (Mann-Whitney U test, $p = 0.07$). This is likely due to greater variance in the Fano factors of the individual cells. Bear in mind that we are using the activity in each bin (either 0 or 1) rather than the actual spike counts in each bin. When using a 1ms bin width, there is no difference between the spike count and the activity but when using a wider bin width, there may be more than one spike per bin.

4.5 Discussion

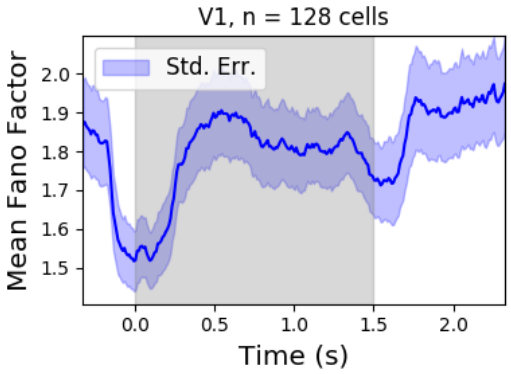
Our aim in this research was to develop a new statistical method for analysing the activity of a neuronal ensemble at very short timescales. We wanted our method to use information taken from the whole ensemble, but we also wanted the method to be quick and easy to implement. It is likely that analysis methods with these characteristics will become valuable as electrophysiological datasets include readings from more cells over longer time periods. In this case, we used the number of active, or spiking, neurons in a very short time bin (< 10ms) as a measure of ensemble activity.

First of all, we showed that there were changes in response that we could model at these very short time scales in some of the brain regions from which we had recordings. We observed changes in the average number of active neurons, and the variance of the number of active neurons in three different brain regions in response to visual stimuli. Since we know that correlated behaviour is associated with sensory perception (deCharms and Merzenich, 1996), we might hope to measure the pairwise correlations within the neuronal population in order to further investigate these responses. But, using such short time bins can produce artificially small spike count correlation measurements (Cohen and Maunsell, 2009). Overcoming this limitation was one of our objectives for our new method. In order to do this, we abandoned the idea of measuring the correlations directly and embraced the concept of *association*. In order to quantify the association between neurons, we used the Conway-Maxwell-binomial distribution to model the number of active (spiking) neurons in an ensemble as a sum of possibly associated Bernoulli random variables.

We showed that the Conway-Maxwell-binomial distribution performed better than the more common options of the binomial and beta-binomial distributions. Furthermore, we showed that the positively associated behaviour between neurons in the primary visual cortex could be captured by fitting a Conway-Maxwell-binomial distribution, but was not captured



(A) Primary visual cortex ν parameter, bin width is 10ms.



(B) Primary visual cortex Fano factor, bin width is 10ms.

FIGURE 4.12: (A) The mean ν parameter of the COMb distribution fitted to activity in the primary visual cortex. Mean taken across all stimulated trials. A bin width of 10ms was used to classify cells as active or inactive. The change in association at stimulus onset is still captured. Some high frequency fluctuations are filtered out by using the wider bins (compare to figure 4.10a) (B) The mean Fano factor of the cells in the primary visual cortex. The change in the mean Fano factor at stimulus onset is not significant when using a bin width of 10ms.

by the more standard approach of measuring the spike count correlation. The associated behaviour could not be measured using spike count correlations, because of the very short bins required to capture short timescale behaviour.

We replicated a famous result from Churchland et al (2010) relating to the quenching of neural variability in cortical areas at stimulus onset, and in doing so, we established a correspondence between the association quantifying parameter of the Conway-Maxwell-binomial (COMb) distribution and the neural variability as measured by the Fano factor. We found a reduction in the ν parameter of the COMB distribution at stimulus onset, indicating a change from no association to positive association between neurons in V1. We found a corresponding reduction in the Fano factor of the individual cells in V1. The positive association between neurons induced by the stimulus would constrain the neurons to fire at the same time. The stimulus also induced a larger number of neurons to spike. These two actions combined could cause an increase in the firing rate of individual cells that is greater in magnitude than the increase in firing rate variability. If this is indeed the case, then the association as captured by the COMB distribution could be regarded as one of the 'natural parameters' of the ensemble response for short timescales. That is, a quantity that directly measures some aspect of the behaviour of the ensemble. In this case, it the correlated behaviour of the individual neurons is captured.

This work could be just a first step in creating analysis methods based on the Conway-Maxwell-binomial distribution, or similar statistical models. One way to extend the method would be to pair it up with the 'Population Tracking model' (O'Donnell et al., 2017). This model attempts to characterise the interaction between an ensemble and each member of the ensemble by quantifying the probability of spiking for a given a cell, given the number of active cells in the whole population. Combining this model with the COMB distribution would give us a model that could accurately fit the number of active neurons at any moment, and that gives a probability of firing for each cell, and therefore probabilities for full spiking patterns, without adding a huge number of parameters to fit.

A more complex way to extend the model would be to fit a Conway-Maxwell-binomial distribution to data recorded from multiple brain regions simultaneously, with a different fit for each region, then to analyse the temporal relationship between the fitted parameters of each region. If we analysed the time series of the COMB distribution parameters from the different regions, looking at cross-correlations between regions, this may give some results relating to the timescales in which information is processed in different brain regions.

Chapter 5

Studies with practical limitations & negative results

Abstract

Here I will present some details on research topics that I started, but that unfortunately did not lead anywhere useful. There are two pieces of research, based on two papers. Each paper is related to the overall theme of my PhD of analysing and modelling behaviours of populations of neurons. The first part is based on a model of parallel spike trains including higher order interactions by Shimazaki et al (2012). The second part is based on a multiscale model for making inferences on hierarchical data.

5.1 Dynamic state space model of pairwise and higher order neuronal correlations

In their paper Shimazaki et al (2012) aimed to model spike trains from populations of neurons in parallel, with pairwise and higher order dynamic interactions between the trains. They modelled the spike trains as multi-variate binary processes using a log-linear model, and they extracted spike interaction parameters using a Bayesian filter/EM-algorithm. They developed a goodness-of-fit measure for the model to test if including these higher order correlations is necessary for an accurate model. Their measure was based on the Bayes factor but they also assessed the suitability of higher order models using the AIC and BIC. So the increase in the number of parameters associated with fitting higher order interactions was taken into account. They tested the performance of the model on synthetic data with known higher order correlations. They used the model to look for higher order correlations in data from awake behaving animals. They use the model to demonstrate dynamic appearance of higher order correlations in the monkey motor cortex (Shimazaki et al., 2012).

We used the available Python repository to implement the model, and we successfully worked through the tutorial provided. But we found that the model did not scale well to larger populations. We attempted to fit the model to a population of 10 neurons and found we didn't manage to finish the process. Since, our goal was to find a model to scale to hundreds or thousands of neurons, we decided that this model was no longer worth pursuing.

5.2 A multiscale model for hierarchical data applied to neuronal data

In their paper Kolaczyk et al (2001) developed a framework for a modelling hierarchically aggregated data, and making inferences based on a model arising from this framework. They assumed that a hierarchical aggregation existed on the data in question, where each element at each level of the hierarchy had some associated measurements, an associated mean process, which was the expected value of these measurements. They also assumed that the measurements of each parent were equal to the sum of the measurements from all of its children. They showed that these assumptions gave rise to a relationship between parent and child measurements across all levels of the hierarchy, where the product of the likelihood of the parameters of the lowest level of the hierarchy can be expressed as products of conditional likelihoods of the elements of higher levels of the hierarchy (Kolaczyk and Huang, 2010).

We hoped that the hierarchical structure of the brain (regions to subregions to cells) and a high level activity measure (fMRI or EEG) could be combined with this model to infer lower level activity from a high level measure.

They gave examples of these expressions for measurements sampled from Gaussian distributions, and Poisson distributions, and showed the definitions of the hierarchical parameters which reparametrise the distribution of these data taking the hierarchy into account. They go on to suggest prior distributions for this multiscale model, and integrate these priors to give posterior distributions for the measurements from each element at each level in the hierarchy, and expressions for the MAP estimated parameters of each the associated processes (Kolaczyk and Huang, 2010).

We implemented their model in Python by creating some synthetic data from Poisson distributions, and defining a hierarchy by agglomerating these data. We calculated the MAP estimates using our knowledge of the hierarchy, and using the expressions given in the paper. We found that the MAP estimates were far less accurate than would be achieved by simply ignoring the hierarchy during estimation, and using a maximum likelihood approach. After that, we decided to move on.

Chapter 6

Discussion

In this project, we attempted to address some of the challenges in data collection from large neuronal ensembles (specifically with calcium imaging) and some of the problems in analysing the data collected from large neuronal ensembles.

Firstly, we investigated the relationship between cell biochemistry, action potentials and the fluorescence traces produced by fluorescent calcium indicators. We did this by building a biophysical model that takes in a spike train and returns the fluorescence trace that that spike would induce. The model included mechanics for the binding of calcium to fluorescent and endogenous mobile and immobile buffers, and the consequent changes in concentration of free and bounded calcium. The model consisted of 17 parameters, 13 of which were set according to data from the literature, and 4 of which were free parameters. We trained the model using simultaneously collected spiking and calcium imaging data (Berens et al., 2018). We fitted the model by matching the $\Delta F/F_0$ in response to an action potential, and by matching the power spectrum of the actual fluorescence trace. This meant that our model would include the correct amount of noise as well as return the correct change in amplitude in response to an action potential.

Since our model produced fluorescence traces, we could apply spike inference algorithms to the modelled fluorescence traces that our model produced after training, and compare the performance of the algorithms on the modelled traces to their performance on the real traces. We used three spike inference algorithms, two of which were based on modelling the calcium trace as an autoregression (Friedrich and Paninski, 2016; Pnevmatikakis et al., 2016), and another inference algorithm that was a little more biologically inspired, but amounted to a very similar algorithm (Deneux et al., 2016). We compared the performance of the algorithms by using them to infer spikes from 20 real and modelled fluorescence traces induced by 20 corresponding real spike trains. We then used several binary classification measures (true positive rate, accuracy etc.) to asses the quality of the spike inference for the real and

modelled fluorescence traces. We found that the spike inference algorithms performed similarly on real and modelled traces, showing that our model is capturing at least some of the characteristics of the real fluorescence traces.

In order to investigate the effect of indicator characteristics on the modelled fluorescence trace and spike inference quality, we perturbed the indicator's affinity and dissociation rate in parallel, keeping the ratio of the two the same for all perturbations. We measured the SNR of the trace, and the true positive rate of the spike inference algorithms at each perturbed value pair. We found that perturbing the values lower caused in decrease in SNR and spike inference quality. This shows that our model could be used to test theoretical fluorescent calcium indicators without having to actually manufacture them. Experimental neuroscientists could also use our model to judge which indicator characteristics are most influential in their experimental context.

We then investigated the effect of perturbing buffer concentration, and indicator concentration, on the signal-to-noise ratio of the modelled fluorescence trace and spike inference quality. This was a worthwhile experiment because endogenous buffer concentrations vary from cell to cell (Bartol et al., 2015; Maravall et al., 2000; Neher and Augustine, 1992), as does indicator expression (Chen et al., 2013). We found that extreme perturbations away from the indicator concentration taken from the literature lowered the SNR of the trace, and the spike inference quality. We also found that increases in the concentration of endogenous buffer above the value taken from the literature caused a decrease in the SNR and spike inference quality. This reiterates that the indicator and endogenous buffers compete to bind with free calcium molecules, and this has an effect on fluorescence and consequently on spike inference.

We then created some synthetic spike trains with controlled mean firing rates sampled the rates from an Ornstein-Uhlehnbeck process. We found that the higher the firing rate, the lower the accuracy of the spike inference algorithms. But the mean firing could perhaps be inferred from the amplitude of the fluorescence traces. The higher firing rate, the higher the amplitude. Calibrating the model to facilitate and accurate measurement would require some kind of ground truth, but relative comparisons could be made without any other knowledge of the underlying spiking process.

One obvious limitation to our model is the lack of binding mechanics for both the indicator and endogenous buffers. Greenberg et al included these mechanics in their successful spike inference model. We felt that the timescale of these binding mechanics was so small in comparison to the fluorescence dynamics that omitting them would make no difference. But

it is possible that their inclusion would improve our model.

After investigating the difficulties with inferring spiking data from calcium imaging data, we moved from data collection to analysis and we decided to implement a new network analysis method on data from a neuronal ensemble. Using an electrophysiological dataset with spike sorted data from 9 different brain regions in 3 mice (Steinmetz, Carandini, and Harris, 2019), we binned the spike times for each cell into spike counts for each cell and measured the correlation coefficients between these spike counts for a selection of cells evenly distributed across the 9 regions. We repeated these measurements for time bin widths ranging from 5ms to 3s. We rectified these measurements and, for a given time bin width, used them as weights for a weighted undirected graph where each node represents a neuron, and the weight of each edge is the correlation between the neurons represented by the nodes on that edge. We applied a novel spectral analysis and community detection method (Humphries et al., 2019) to this network. This clustered the nodes in our ensemble into communities whose behaviour was more correlated than expected. Our measure of 'expected correlation strength' were based on a random network that matched our data network's sparsity and total weight. We compared the detected communities to the anatomical division of our cells using clustering comparison measures. We then conditioned the binned spike counts on the behaviour of the mouse using the principal components of a video of the mouses face recorded simultaneously with the electrophysiology. We broke the total covariance down into 'spike count covariance' and 'signal covariance' components conditioning on the behavioural data and using the law of total covariance. We then repeated our analysis for spike count correlations, and signal correlation. Finally, since our community detection method was only valid on graphs with non-negative weights, we used different methods for creating a non-negative graph from our total correlations, and we repeated our analysis on those graphs.

Our first finding was that the time bin width used to bin spike times into spike counts had a effect on the mean magnitude of the correlations measured. The wider the bin, the higher the correlations. Not only that, we separated the pairs into positively and negative correlated pairs, and we found that positively correlated pairs have greater correlation coefficients when using a wider bin, and negatively correlated pairs have more negative correlation coefficients when using a wider bin. We also found that the width of the bin used had an effect on the distribution of the spike counts. For smaller bin widths, the distribution of spike counts was better represented by a skewed distribution like the Poisson distribution. For wider bins, the spike counts were better represented by a Gaussian distribution.

Next we investigated the differences between correlations within regions and between

regions. When we divided the pairs according to those two groups, we found that the mean within-region correlations were higher at every bin width, and the difference between the two means grew with increasing bin width. When we split the pairs of cells according to their regions, we found that the mean within-region correlations in any given region were usually greater than the mean between-region correlations for any region pair involving that region. The difference between the mean within-region correlation and the highest between-region correlations involving that region grew smaller with increasing bin width. To investigate this further, we plotted these mean correlations in matrices. Although the mean within-region correlations were usually the highest value in their row or column, as the bin width increased, the mean between-region correlations grew in magnitude relative to the within-region figure.

Next we chose a null network model, and we used the 'Network Noise Rejection' process (Humphries et al., 2019) to check for additional structure in our correlation based data network that was not captured by the null model. We found additional structure for any bin width that we used. We also found that the dimensionality of the additional structure reduced as we increased the bin width. This could mean that the processes or representations that take place over longer timescales within the brain also take place in a lower dimensional space.

We applied a community detection method (Humphries, 2011) to the signal correlation networks arising from the network noise rejection. We found that the number of communities detected decreased with increasing bin width. We also noticed that at shorter bin widths, the detected communities were more likely to consist of cells from one brain region only. We investigated this further by using clustering comparison methods to compare the detected communities with the anatomical division of the cells. We found that for short timescales < 50ms correlated communities tended to exist within anatomical regions, and for longer timescales > 100ms, the correlated communities tended to exist across anatomical regions. This is broadly in agreement with a similar finding for EEG data from humans performing semantic or memory tasks (Stein and Sarnthein, 2000). Von Stein et al. (2000) found that visual processing taking place locally in the visual system was captured in the gamma frequency range ($25 - 70$Hz), while semantic processing was captured in the beta range ($12 - 18$Hz), and tasks involving mental imagery and working memory retention were captured in the theta and alpha ranges ($4 - 8$Hz, and $8 - 12$Hz respectively).

We then conditioned our correlation measures on the the mouse's behaviour. This allowed us to create spike count correlation (or noise correlation) networks, and signal correlation networks (Cohen and Kohn, 2011). We applied our analysis to these networks. For the

spike count correlation networks we found very similar results to the total correlation networks. More communities at smaller bin widths, and communities resembled the anatomical division at smaller bin widths. Given that recent findings show that behaviour can account for correlated spiking in many areas of the brain (Stringer et al., 2019), these results for the spike count correlation show that this correlated behaviour is still processed locally at short timescales, while processes and representations that take more time make use of correlated activity across multiple regions.

For the signal correlations, we still found additional structure in these networks. But we always detected a smaller number of communities. These communities also had no relation to the anatomical division of the cells. This result shows that there are groups of cells across multiple brain regions that are activated similarly by certain behaviours.

All of the networks so far were based on rectified correlation measures, because the network noise rejection and community detection processing is (currently) only valid for networks with non-negative weights. For the final part of our analysis, we tried different ways of transforming our total correlations into non-negative quantities before applying our analysis. First of all we took the absolute value of our correlation measures. Our results were very similar to those for the rectified correlations with the exception that we detected more communities consistently. It is possible that using this method detects both positively and negatively correlated communities.

We also tried reversing the sign of all the correlations, then rectifying the network. We hope that this would allow us to detect the negatively correlated communities. We did detect communities in these networks, but never more than three, and these communities bore no relationship with the anatomical distribution of the cells.

There is a lot of potential for network science applications in computational neuroscience. For example, some pairwise measure other than correlation coefficients could be used as the weights of the graph. The synaptic connections between cells can be isolated in-vitro (Okun et al., 2015). A map of these synaptic connections could be used as the basis for directed graphs. The analysis methods applicable to directed graphs could give insights about the formation of synaptic connections, or the dynamic changes in these connections over time. Other methods of community detection could be used on directed or undirected graphs. We used a 'hard' clustering method in our research, that is, each neuron could be a member of one cluster/community only. 'Fuzzy-clustering' methods do exist, where each element of the set to be clustered could be a member of more than one cluster (Baadel, Thabtah, and Lu, 2016).

Having spent much time investigating correlated behaviour using coefficients of spike counts, we decided to try another method for capturing correlated behaviour in neuronal ensembles. We used electrophysiological data taken from 5 brain regions of an awake mouse exposed to visual stimuli (Steinmetz et al., 2019). We modelled the number of active neurons in a given brain region as the number of successes in a collection of dependent Bernoulli random variables using the Conway-Maxwell-binomial distribution. To avoid violating the Bernoulli assumption, we binned the spike times using 1ms bins. The Conway-Maxwell-binomial distribution is a two parameter extension of the standard binomial distribution. The extra parameter allows the distribution to capture possible positive or negative association between the Bernoulli trials (Kadane, 2016). This means that we are assuming that all the neurons are dependent in the same way. This is not an accurate assumption, but it allows us model the data in a simple way.

First of all we established that there were changes in the number of active neurons in response to the visual stimuli. This was the case in the hippocampus, thalamus, and primary visual cortex. Each region had its own signature response. We measured the mean and variance of the number of active neurons in a sliding window starting before stimulus onset, and finishing after the end of stimulus presentation.

As well as the Conway-Maxwell-binomial distribution, we also fitted binomial, and beta-binomial distributions to the number of active neurons in a sliding window. We found that the Conway-Maxwell-binomial distribution was the best fit for over 90% of the samples. This means that the COMb distribution is capturing some dependency between the neurons, because the binomial distribution assumes independence. Also the COMb distribution captures this dependence more accurately than the beta-binomial distribution, which does have some capacity for over dispersion.

Next we showed that the Conway-Maxwell-binomial distribution captured the change in association at stimulus onset better than the correlation coefficient. The extremely small bin width artificially shrunk the correlation coefficient to the point where this measurement didn't detect any correlated activity. But the association parameter of the COMb distribution detected some positive association between the neurons at stimulus onset. So, for particularly short time bins, where neurons can be treated as Bernoulli random variables, the Conway-Maxwell-binomial distribution is a good way to capture correlated behaviour. There are other measurements for capturing association to which this distribution should be compared. Cross-correlograms could be used for some measure of synchrony, for example.

Finally, we replicated a famous finding of Churchland et al. (2010) relating to the quenching of neural variability at stimulus onset, thereby finding a parallel between this reduction in the Fano factor and a reduction in the association parameter of the COMb distribution.

We showed that computational neuroscientists can make progress by being inventive with their statistical models. A similar distribution to investigate would be the Conway-Maxwell-Poisson distribution. This is similar to the standard Poisson distribution, but with an additional parameter that allows for over- or under- dispersion relative to a Poisson distribution. This might be ideal for modelling firing rates of individual neurons. Some interaction between the fitted parameters could capture the association between neurons.

There is one technology that has the potential to take over from both electrophysiology and calcium imaging. The technique of voltage imaging has become more useful in recent years. The aim for neuroscience would be to develop a voltage imaging dye or protein that images the membrane potential of a neuron with enough spatial and temporal resolution to detect action potentials. The voltage imaging dyes that have been developed so far do not have high enough spatial resolution to single out individual cells in-vivo using staining (Bando et al., 2019). But, genetically encoded voltage indicators have been developed that have high enough resolution to indicate individual spikes and subthreshold activity from small numbers of cells in the striatum, hippocampus, and cortex of awake behaving mice (Piatkevich et al., 2019). These indicators have the potential to take over from calcium imaging, and if imaging deep within the brain becomes possible, electrophysiology could also be replaced. This is speculation, but the potential is there.

Bibliography

Allen, William E. et al. (2019). “Thirst regulates motivated behavior through modulation of brainwide neural population dynamics”. In: *Science* 364.6437.

Baadel, S., F. Thabtah, and J. Lu (2016). “Overlapping clustering: A review”. In: *2016 SAI Computing Conference (SAI)*, pp. 233–237.

Baldassano, Christopher et al. (2017). “Discovering Event Structure in Continuous Narrative Perception and Memory”. In: *Neuron* 95.3, 709 –721.e5.

Bando, Yuki et al. (2019). “Genetic voltage indicators”. In: *BMC Biology* 17.1, p. 71.

Bartol, Thomas M. et al. (2015). “Computational reconstitution of spine calcium transients from individual proteins”. In: *Frontiers in Synaptic Neuroscience* 7, p. 17.

Berens, Philipp et al. (May 2018). “Community-based benchmarking improves spike rate inference from two-photon calcium imaging data”. In: *PLOS Computational Biology* 14.5, pp. 1–13.

Bezanson, Jeff et al. (Sept. 2012). “Julia: A Fast Dynamic Language for Technical Computing”. In: *MIT*.

Brini, Marisa et al. (2014). “Neuronal calcium signaling: function and dysfunction”. In: *Cellular and Molecular Life Sciences* 71.15, pp. 2787–2814.

Buccino, Alessio P. et al. (2019). “SpikeInterface, a unified framework for spike sorting”. In: *bioRxiv*.

Chen, Tsai-Wen et al. (July 2013). “Ultrasensitive fluorescent proteins for imaging neuronal activity”. In: *Nature* 499, 295–300.

Churchland, Mark M. et al. (2010). “Stimulus onset quenches neural variability: a widespread cortical phenomenon”. eng. In: *Nature neuroscience* 13.3. 20173745[pmid], pp. 369–378.

Cohen, Marlene R. and Adam Kohn (2011). “Measuring and interpreting neuronal correlations”. In: *Nature Neuroscience* 14.7, pp. 811–819.

Cohen, Marlene R. and John H. R. Maunsell (2009). “Attention improves performance primarily by reducing interneuronal correlations”. eng. In: *Nature neuroscience* 12.12. 19915566[pmid], pp. 1594–1600.

Dana, Hod et al. (Sept. 2014). "Thy1-GCaMP6 Transgenic Mice for Neuronal Population Imaging In Vivo". In: *PloS one* 9, e108697.

deCharms, R. Christopher and Michael M. Merzenich (1996). "Primary cortical representation of sounds by the coordination of action-potential timing". In: *Nature* 381.6583, pp. 610–613.

Deneux, Thomas et al. (July 2016). "Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo". In: *Nature Communications* 7.1.

Dombeck, Daniel A. et al. (2010). "Functional imaging of hippocampal place cells at cellular resolution during virtual navigation". In: *Nature Neuroscience* 13.11, pp. 1433–1440.

Éltes, Tímea et al. (2019). "Improved spike inference accuracy by estimating the peak amplitude of unitary [Ca2+] transients in weakly GCaMP6f-expressing hippocampal pyramidal cells". In: *The Journal of Physiology* 597.11, pp. 2925–2947.

Faas, Guido C. et al. (2011). "Calmodulin as a direct detector of Ca2+ signals". In: *Nature Neuroscience* 14.3, pp. 301–304.

Fiala, John C. and Kristen M. Harris (1999). *Dendrite Structure*.

Flach, Boris (Sept. 2013). "A Class of Random Fields on Complete Graphs with Tractable Partition Function". In: *IEEE transactions on pattern analysis and machine intelligence* 35, pp. 2304–6.

Forney, G. D. (1973). "The viterbi algorithm". In: *Proceedings of the IEEE* 61.3, pp. 268–278.

Fosdick, Bailey et al. (Aug. 2016). "Configuring Random Graph Models with Fixed Degree Sequences". In: *SIAM Review* 60.

Friedrich, Johannes and Liam Paninski (2016). "Fast Active Set Methods for Online Spike Inference from Calcium Imaging". In: *Advances in Neural Information Processing Systems 29*. Ed. by D. D. Lee et al. Curran Associates, Inc., pp. 1984–1992.

Ganmor, Elad, Ronen Segev, and Elad Schneidman (2011). "Sparse low-order interaction network underlies a highly correlated and learnable neural population code". In: *Proceedings of the National Academy of Sciences* 108.23, pp. 9679–9684.

Girard, P., J. M. Hupé, and J. Bullier (2001). "Feedforward and Feedback Connections Between Areas V1 and V2 of the Monkey Have Similar Rapid Conduction Velocities". In: *Journal of Neurophysiology* 85.3. PMID: 11248002, pp. 1328–1331.

Greenberg, David et al. (Nov. 2018). "Accurate action potential inference from a calcium sensor protein through biophysical modeling:" in:

Gründemann, Jan et al. (2019). "Amygdala ensembles encode behavioral states". In: *Science* 364.6437.

Hodgkin, A. L. and A. F. Huxley (1939). "Action Potentials Recorded from Inside a Nerve Fibre". In: *Nature* 144.3651, pp. 710–711.

Humphries, Mark et al. (Jan. 2019). "Spectral rejection for testing hypotheses of structure in networks". In:

Humphries, Mark D. (2011). "Spike-Train Communities: Finding Groups of Similar Spike Trains". In: *Journal of Neuroscience* 31.6, pp. 2321–2336.

Jun, James J. et al. (2017). "Fully integrated silicon probes for high-density recording of neural activity". In: *Nature* 551.7679, pp. 232–236.

Kadane, Joseph B. (June 2016). "Sums of Possibly Associated Bernoulli Variables: The Conway–Maxwell-Binomial Distribution". In: *Bayesian Anal.* 11.2, pp. 403–420.

Kilhoffer, Marie-Claude et al. (Sept. 1992). "Use of Engineered Proteins With Internal Tryptophan Reporter Groups and Pertubation Techniques to Probe the Mechanism of Ligand-Protein Interactions: Investigation of the Mechanism of Calcium Binding to Calmodulin". In: *Biochemistry* 31.34, pp. 8098–8106.

Koch, Christoff (1999). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press.

Kolaczyk, Eric and Haiying Huang (Sept. 2010). "Multiscale Statistical Models for Hierarchical Spatial Aggregation". In: *Geographical Analysis* 33, pp. 95 –118.

Lee, Suk-HO et al. (July 2000). "Differences in Ca2+ buffering properties between excitatory and inhibitory hippocampal neurons from the rat". In: *The Journal of Physiology* 525.

Litwin-Kumar, Ashok, Maurice Chacron, and Brent Doiron (Sept. 2012). "The Spatial Structure of Stimuli Shapes the Timescale of Correlations in Population Spiking Activity". In: *PLoS computational biology* 8, e1002667.

Maravall, M et al. (May 2000). "Estimating intracellular calcium concentrations and buffering without wavelength ratioing". In: *Biophysical Journal* 78.5, pp. 2655–2667.

Maugis, Pa (Jan. 2014). "Event Conditional Correlation: Or How Non-Linear Linear Dependence Can Be". In:

Meilă, Marina (2007). "Comparing clusterings—an information based distance". In: *Journal of Multivariate Analysis* 98.5, pp. 873 –895.

Murray, John D. et al. (2014). "A hierarchy of intrinsic timescales across primate cortex". In: *Nature Neuroscience* 17.12, pp. 1661–1663.

Neher, E. and G. J. Augustine (1992). "Calcium gradients and buffers in bovine chromaffin cells". eng. In: *The Journal of physiology* 450. 1331424[pmid], pp. 273–301.

O'Donnell, Cian et al. (Jan. 2017). "The population tracking model: a simple, scalable statistical model for neural population data". English. In: *Neural Computation* 29.1, pp. 50–93.

Okun, Michael et al. (2015). "Diverse coupling of neurons to populations in sensory cortex". In: *Nature* 521.7553, pp. 511–515.

Ouzounov, Dimitre G. et al. (2017). "In vivo three-photon imaging of activity of GCaMP6-labeled neurons deep in intact mouse brain". eng. In: *Nature methods* 14.4. 28218900[pmid], pp. 388–390.

Patterson, Carlyn A. et al. (2014). "Similar adaptation effects in primary visual cortex and area MT of the macaque monkey under matched stimulus conditions". In: *Journal of Neurophysiology* 111.6. PMID: 24371295, pp. 1203–1213.

Peron, Simon P. et al. (2015). "A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior". In: *Neuron* 86.3, pp. 783–799.

Piatkevich, Kiryl D. et al. (2019). "Population imaging of neural activity in awake behaving mice". In: *Nature* 574.7778, pp. 413–417.

Pnevmatikakis, E.A. et al. (Jan. 2016). "Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data". In: *Neuron* 89.2, pp. 285–299.

Pnevmatikakis, Eftychios et al. (Nov. 2013). "Bayesian spike inference from calcium imaging data". In: *Conference Record - Asilomar Conference on Signals, Systems and Computers*.

Pnevmatikakis, Eftychios et al. (Sept. 2014). "A structured matrix factorization framework for large scale calcium imaging data analysis". In:

Schneidman, Elad et al. (2006). "Weak pairwise correlations imply strongly correlated network states in a neural population". eng. In: *Nature* 440.7087. 16625187[pmid], pp. 1007–1012.

Shannon, C. E. (1948). "A Mathematical Theory of Communication". In: *Bell System Technical Journal* 27.3, pp. 379–423.

Shimazaki, Hideaki et al. (2012). "State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data". eng. In: *PLoS computational biology* 8.3. 22412358[pmid], e1002385–e1002385.

Staude, Benjamin, Sonja Grün, and Stefan Rotter (2010). "Higher-Order Correlations and Cumulants". In: *Analysis of Parallel Spike Trains*. Ed. by Sonja Grün and Stefan Rotter. Boston, MA: Springer US, pp. 253–280.

Stein, Astrid von and Johannes Sarnthein (2000). “Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization”. In: *International Journal of Psychophysiology* 38.3, pp. 301 –313.

Steinmetz, Nick, Matteo Carandini, and Kenneth D. Harris (2019). *”Single Phase3” and ”Dual Phase3” Neuropixels Datasets*.

Steinmetz, Nick et al. (Mar. 2019). “Eight-probe Neuropixels recordings during spontaneous behaviors”. In:

Stevenson, Ian H. and Konrad P. Kording (2011). “How advances in neural recording affect data analysis”. In: *Nature Neuroscience* 14.2, pp. 139–142.

Stringer, Carsen et al. (2019). “Spontaneous behaviors drive multidimensional, brainwide activity”. In: *Science* 364.6437.

Tada, Mayumi et al. (2014). “A highly sensitive fluorescent indicator dye for calcium imaging of neural activity in vitro and in vivo”. In: *European Journal of Neuroscience* 39.11, pp. 1720–1728.

Tian, Lin et al. (2009). “Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators”. eng. In: *Nature methods* 6.12. 19898485[pmid], pp. 875–881.

Tkačik, Gašper et al. (2014). “Searching for collective behavior in a large network of sensory neurons”. eng. In: *PLoS computational biology* 10.1. 24391485[pmid], e1003408–e1003408.

Treves, Alessandro and Stefano Panzeri (1995). “The Upward Bias in Measures of Information Derived from Limited Data Samples”. In: *Neural Computation* 7.2, pp. 399–407.

Vinh, Nguyen Xuan, Julien Epps, and James Bailey (Dec. 2010). “Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance”. In: *J. Mach. Learn. Res.* 11, 2837–2854.

Vogelstein, Joshua T. et al. (Oct. 2010). “Fast nonnegative deconvolution for spike train inference from population calcium imaging”. In: *Journal of neurophysiology* 104.6, 295–300.

Wierzynski, ”Casimir M. et al. (2009). “”State-Dependent Spike-Timing Relationships between Hippocampal and Prefrontal Circuits during Sleep””. In: *”Neuron”* ”61”.”4”, ”587 –596”.

Zariwala, Hatim A. et al. (2012). “A Cre-dependent GCaMP3 reporter mouse for neuronal imaging in vivo”. eng. In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 32.9. 22378886[pmid], pp. 3131–3141.

Zou, Hui and Trevor Hastie (2005). "Regularization and variable selection via the Elastic Net". In: *Journal of the Royal Statistical Society, Series B* 67, pp. 301–320.