UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Context-dependent calcium dynamics in the C. elegans nervous system

Permalink

https://escholarship.org/uc/item/4j43w34r

Author

Cecere, Zachary

Publication Date

2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Context-dependent calcium dynamics in the *C. elegans* nervous system

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Neurosciences with a specialization in Computational Neurosciences

by

Zachary Cecere

Committee in charge:

Professor Sreekanth Chalasani, Chair
Professor Nicholas Spitzer, Co-Chair
Professor Richard Daneman
Professor Eran Mukamel
Professor Tatyana Sharpee

2020

Copyright
Zachary Cecere, 2020
All rights reserved.

The dissertation of Zachary Cecere is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California San Diego

2020

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to ackowledge Professors Sreekanth Chalasani and Tatyana Sharpee for their mentorship. Their long-term support made this thesis possible.

My committee has been helpful with their suggestions and generous with their time. I would like to thank Drs. Nicholas Spitzer, Eran Mukamel, Richard Daneman, and Massimo Vergassola.

Chapters 2, 3, and 4 are a manuscript in preparation: Context-dependent calcium dynamics in the *C. elegans* nervous sytem. The full author list is Zachary Cecere, Tatyana Sharpee, and Sreekanth Chalasani. The dissertation author is the primary investigator for this paper.

VITA

2013	B. A. in Computer Science *suma cum laude*, Colby College
2020	Ph. D. in Neurosciences with a specialization in Computational Neurosciences, University of California, San Diego

PUBLICATIONS

Cecere Z, Sharpee S, Chalasani S. Context-dependent calcium dynamics in the *C. elegans* nervous system. In preparation.

ABSTRACT OF THE DISSERTATION

Context-dependent calcium dynamics in the *C. elegans* nervous system

by

Zachary Cecere

Doctor of Philosophy in Neurosciences with a specialization in Computational Neurosciences

University of California San Diego, 2020

Professor Sreekanth Chalasani, Chair
Professor Nicholas Spitzer, Co-Chair

Organisms must adapt their behavior to different environmental contexts. A smart behavior in one environment may be a very dumb behavior in another. Many different parts of the nervous system exhibit context-dependent behavior. Sensory neurons adapt their response properties to efficiently encode the organism's environment. Sensorimotor neurons consider both the stimulus history and the recent history of organismal actions when deciding what motor commands to issue. Both of these concepts are studied using large-scale calcium imaging in the *Caenorhabditis elegans* nervous system. State-space models are used to understand how *C. elegans* sensory neurons change their response properties as a function of recent stimulus history. A variety of models are used to study the combined effects of recent stimulus history and recent motor activity on *C. elegans* motor command neurons.

Chapter 1

Introduction

An organism's response to identical stimuli may vary dramatically depending on the context of its past experiences and actions. For instance, while my dog Sebastian will happily accept a treat in most circumstances, he will be too afraid to do so if the toaster has been recently taken out of its storage location. His recent experience (whether he has seen the toaster being moved) provides a context that frames his responses (whether he will accept a treat). The effect of context on neural processing pervades many different Neuroscience disciplines.

One of the earliest studies of context effect on sensory processing is ED Adrian's work in the rabbit olfactory bulb [2]. When the rabbit is under anaesthesia, its olfactory bulb exhibits consistent odorant responses. As the anaesthesia wears off, the consistency of bulb neuron responses goes away; rhythmic activity dominates at the expense of simple odor encoding. Similarly, in anesthetized mice, many olfactory bulb neurons respond to odorants in a simple and consistent manner. The majority of these neurons cease to exhibit this simple responsiveness when the animal is awake [47]. The fact that wakefulness state has such a profound influence on sensory processing surprised the sensory neuroscience community. In light of these studies, the neuroscience field has become much more cognizant of the environmental context in which it does its experiments.

Adaptation can be thought of as a subset of context-dependent neural behavior. The field of adaptation studies how recent stimuli affect the way a sensory system experiences new stimuli. In this case, the recent stimuli is the context. Responses of sensory neurons to new stimuli depend on this context. Many studies have found that neurons adapt their responses to represent stimuli with optimally efficient coding strategies [55] [49] [17]. For instance, fly H1 neurons modulate their gain in accordance with the standard deviation of Gaussian stimuli [7]. This standard deviation adaptation maximizes information transmission about the Gaussian stimuli. Sensory neurons adapt to more complex stimuli in a similar manner [50].

Attention can also be framed as a context-dependent neural behavior. The attentional state of a nervous system affects the response properties of its sensory neurons [44]. For instance, visual system neurons detect visual features at lower contrasts when an organism is attending to an image [20] [44] [56]. The attentional state is the context that informs neuronal sensitivity. Even more interesting, attentional state is correlated with neural state. For instance, in monkeys, frontal and parietal cortices synchronize at lower firing frequencies when the monkey is attending [9]. This introduces a very important concept: the global activity state of an organism encodes context. Thus, study of global activity state of a nervous system may reveal what kinds of contextual information are important to an organisms. Moreover, the relationship between global brain state and the response properties of individual neurons can elucidate the mechanisms of context-dependent neural behaviors.

Since neural response properties are mediated by global brain-state, it makes sense to view as much of the nervous system as possible. With this in mind, researchers have turned to large-scale imaging of simpler organisms [53]. With their smaller number of neurons, these organisms make for an easier target for state analyses. Large-scale calcium imaging studies of larval zebrafish have identified a number of network states correlating with different behaviors [14] [15]. Interestingly, the neuronal dynamics of one region of the zebrafish brain depend on the state of other brain regions. For instance, the state of the zebrafish's inferrior olive controls the fish's motor command gain in a number of brain regions [4]. Thus, the firing properties in a subset of fish neurons depend on the network state of the larger fish brain.

1.1 Context-dependent behavior in C. elegans

Caenorhabditis elegans is an excellent candidate for large-scale state analysis. *C. elegans* is a translucent, soil-dwelling nematode [8]. With only 302 neurons, a large percentage of the worm's neurons can be analyzed simultaneously. Despite its small number of neurons, *C. elegans* exhibits a number of context-dependent behaviors.

Habituation is perhaps the best studied context-dependent behavior in *C. elegans* [48]. *C. elegans* swim backwards in response to taps on their petri dish. This 'tap withdrawal' response habituates to repeated taps. The rate of habituation and recovery depends on the interstimulus interval. Interestingly, *C. elegans* can learn contextual associations to tapping. If a worm initially habituates in the presence of an odorant, the rate of later habituations is increased by presence of the same odorant.

C. elegans primary sensory neurons also exhibit habituation. The response magnitude of primary sensory cells diminishes with repeated stimulus exposure [28]. This decrease in responsiveness persists across multiple stimulus pulses.

In another example of history-dependent information processing, *C. elegans* search behavior is modified by food experience [10] [22]. When removed from food, *C. elegans* enter their local search behavior. The local search behavior is characterized by an elevated number of reorientation events [22]. After ten to fifteen minutes, the worm transitions from local to global search behavior, reducing its reorientation event probability. Interestingly, the size of the food patch affects worm reorientation probability during local search [10].

Overall, *C. elegans* exhibit a number of history and state dependent behaviors. The ease with which the worm can be imaged makes it a prime target for identifying neural correlates of this history dependence.

1.2 Modeling context-dependence

Although visualizing a greater percentage of the nervous system theoretically helps to elucidate neuronal encoding properties, it adds complexity to encoding analyses. Understanding neuronal encoding is a core goal of computational neuroscience. Any method for understanding encoding must be simultaneously tractable, flexible, and interpretable. Fulfilling all three of these requirements is difficult.

Early methods derived neuronal receptive fields using reverse correlation and spike-triggered covariance methods [46] [13]. These simple methods give 'rigorously interpretable results only for gaussian distributions of inputs' [49]. For more complex stimulus distributions, first and second moments are no longer sufficient to capture the distribution. Information theoretic methods allow for the derivation of receptive fields for stimuli with higher order moments [49]. These information theoretic methods are made tractable by averaging over multiple presentations of the same stimulus [49] [31]. This averaging is problematic if one assumes contexts like network state and recent stimulus history are important.

Another way to handle complex input distributions is to assume a spiking model. This has become common practice with the success of generalized linear models [38] [39] [40]. These methods model neuronal spiking as a combination of linear and nonlinear components. Recent stimuli are passed through an optimizable linear filter. The scalar output of this filter is then passed through a nonlinearity: a logistic function for modeling spike probabilities or a Poisson function for modeling spike rates [40]. Generalized linear models are particularly attractive because of their convex objective functions. Moreover, these models have successfully captured contextual information. In [40], the authors added additional linear filters to capture the spiking activity of nearby neurons. The inclusion of these filters significantly improved the generalized linear model's ability to predict spikes in a target neuron.

Generalized linear models capture state-dependence by linearly combining spike information from adjacent neurons with spatiotemporal stimulus filters. It is unclear whether this assumption of linearity is appropriate. For instance, it is possible that network state fundamentally changes the way in which a given neuron processes information. Input-driven hidden markov models have become increasingly popular for modeling network state dependent receptive fields [16] [32]. These models introduce latent (hidden) states. They learn both the transition probabilities between states and the receptive fields for each hidden state. These input-driven HMMs are flexible enough to model neurons with receptive fields that vary drastically across network states.

Switching linear dynamical systems (a generalization of input-driven HMMs) have successfully modeled *C. elegans* calcium activity [33]. It was found that fourteen latent states best capture *C. elegans* network activity. In other words, there are at least fourteen network states in the worm. In each of these states, the network exhibits different linear dynamics. Thus, *C. elegans* network behavior varies significantly depending on network state context. It is clear that any reasonable analysis of *C. elegans* network dynamics must capture the effect of network state on individual neurons.

Chapter 2

Large-scale imaging in *C. elegans*

2.1 Introduction

Information flows through nervous systems in a brain state dependent manner. For instance, monkey frontal and parietal cortices exhibit different oscillatory behavior when the animal is attending to a stimulus versus when it is not [9]. This change in oscillatory behavior is coupled with an increased sensitivity to low contrast stimuli in sensory neurons [44] [45]. Hence, brain state controls the information processing properties of individual neurons within a neural network. It also affects organismal performance on behavioral tasks [44].

Given the important contextual information of brain state, it makes sense to visualize as much of the brain as possible. While imaging technology has advanced considerably over the past two decades, current imaging technology is still limited to the study of hundreds or thousands of neurons [3]. With this in mind, many scientists have turned to invertebrates (section 2.5.1). With their comparatively small number of neurons, invertebrates like *Drosophila*, larval Zebrafish, and *C. elegans* provide a much easier imaging target. Nearly all of the neurons in the heads of each of these organisms can be imaged with sub-second resolution.

Post-processing techniques have improved alongside advances in calcium imaging technology. Motion correction and image segmentation algorithms have improved to the point that calcium transients from nearby cells can be disentangled. This enables the high fidelity visualization of a huge portion of an invertebrate's total calcium activity (section 2.5.3). It has become possible to resolve how an invertebrate organism's calcium state affects its information processing.

Kato et al [28] is the first *C. elegans* 'full brain' imaging study. The authors image immobilized *C. elegans*. A principal components analysis of the worm's neuronal calcium traces finds the first three components capture 65% of the neuronal variance. This is likely an underestimate as every trace shown in the paper looks like a noisified version of one of three cells: RME, AVA, or SMDV (section 2.5.4). A number of *C. elegans* mutants are created to find the behavioral correlates of RME and AVA rise activity [28]. AVA and its correlated partner neurons rise to their upper baseline value during worm reversals. The RME cell cluster, on the other hand, rises to its upper baseline during forward locomotion. These appear to be the command neuron clusters associated with forward and backward locomotion. Perhaps the most interesting finding of this paper is that *C. elegans* motor commands continue to operate in immobilized worms. Overall, Kato et al identify four neuronal classes: neurons correlated with AVA, neurons correlated with RME, neurons correlated with SMDV, and oxygen-sensing sensory neurons.

The SMDV neurons identified by Kato et al control ventral turning in the worm. Kaplan et al [27] study SMDV and SMDD neurons in more detail. The authors find SMDV and SMDD control ventral and dorsal head casting. Overall, imaging of *C. elegans* in microfluidic chips unveils forward-backward and dorsal-ventral motor commands. Perhaps the most surprising finding is how little other activity there appears to be. The *C. elegans* nervous system is operating in a very low dimensional space in these experiments.

The question becomes: how does the worm's motor command state affect information processing? There are a few questions embedded in this larger goal. What are the different worm command neuron states? What are the possible state transitions? How does command neuron state affect stimulus processing?

2.2 Results

The NeuroPAL (section 2.5.5) (OH15500) strain and another pan-neuronal GCaMP strain (ZIM294) are exposed to a variety of pulse-based stimulus patterns. Food (OP50) is used as the stimulus in all trials (section 2.4.3) [10]. Neuronal calcium traces are acquired using non-negative matrix factorization (section 2.5.3). The same six cell types appear in every worm: 'On' cells, 'Off' cells, anti-correlated moderately fast cells, and anti-correlated bistable cells. The 'On' and 'Off' cells track the stimulus exactly. The 'On' cells are positively correlated with the stimulus, while the 'Off' cells are anti-correlated with the stimulus.

The moderately fast, anticorrelated cell types always appear on opposite sides of the worm. One set of cells can be found on the ventral side of the worm, while the opposing cell type is invariably found on the dorsal side. Cells on the same side of the worm exhibit nearly identical calcium transients. The combination of this spatial and waveform information strongly suggests these cells are SMDV and SMDD [27]. Indeed, where NeuroPAL landmarks are easy to make out, these cells can be positively identified as the SMDV and SMDD pairs [57]. Moreover, the stereotyped waveform and spatial locations enable SMDV and SMDD identification in non-NeuroPAL strains.

The RME cell is invaluable for bistable cell identification. It can be easily identified in NeuroPAL and non-NeuroPAL worms as it appears in relative isolation just in front of the nerve ring [57]. The RME cell cluster is defined as the cells whose calcium transients correlate with the RME calcium transient at 85% or greater. The cell most anti-correlated with RME is used to identify the anti-RME cluster. All cells correlated with this cell at a rate greater than 85% are members of the anti-RME cluster. When AVA identification is possible, it invariably appears in the anti-RME cluster. Hence, the anti-RME cluster is referred to as the AVA cluster throughout the paper. Kato et al [28] identify the same cell types. Moreover, no other cell types could be identified in either the ZIM294 or OH15500 worms. Putative cells not appearing in one of the aforementioned six cell types typically resemble noise or noisier versions of the six primary cell types.

Given the extreme correlation within the SMDV, SMDD, RME, and AVA clusters, averaging within each of the clusters loses very little information. Thus, all subsequent sections use cluster-averaged representations to simplify model structure.

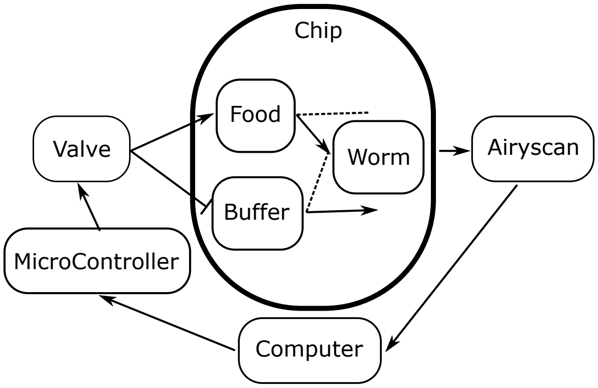

Figure 2.1: Figure 1. Imaging setup schematic. A worm is loaded into a microfluidic chip. Food and buffer are always flowing in two different chip channels. Non-dashed lines indicate food-exposure mode. In this mode, the valve is set so that food flows to the nose of the worm. The buffer flows away from the worm. Triggering the valve causes the buffer to flow to the nose of the worm and food to flow away from the worm. The computer controls the valve and receives images from the Airyscan, allowing for the synchronization of image acquisition and flow control.

2.3 Discussion

In a microfluidic chip, the *C. elegans* neural network operates in low dimensional space. Like in [28], nearly every calcium trace belongs to one of the following cell clusters: forward-backward locomotion controllers, left-right head casting controllers, and sensory neurons. Notably, there do not appear to be any neurons dedicated to encoding statistical features of the stimulus. Thus, it is unlikely that classical receptive field analyses will find anything of interest in the worm.

As in [28], a large number of neurons in each worm correlate with AVA and RME. The

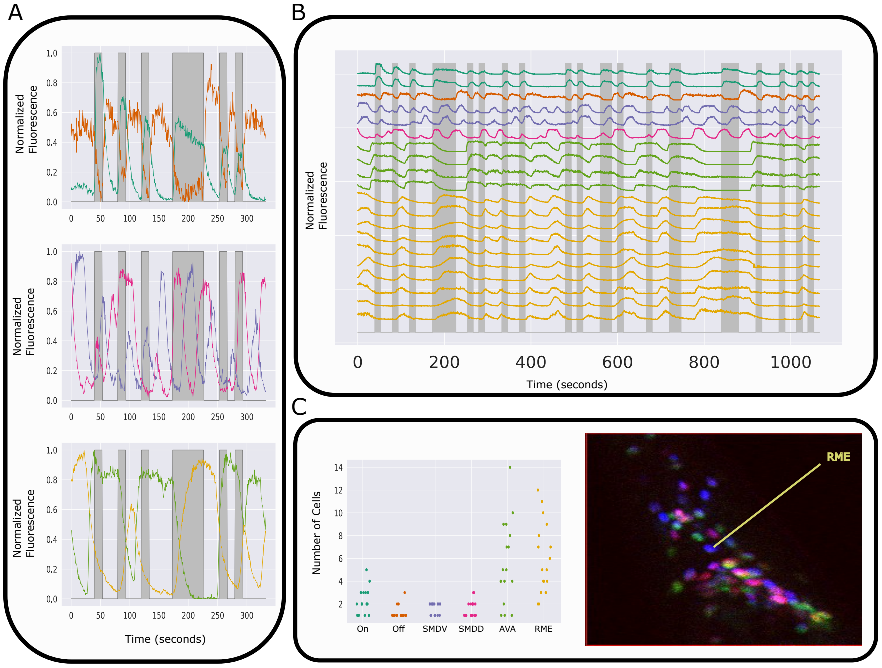

Figure 2.2: Figure 2. Large-scale calcium imaging in C. elegans uncovers six cell types. A: Sample fluorescence traces for the six cell types for a single in a single worm. Gray bars indicate food stimulus is flowing to the nose of the worm. Top: Average fluorescence traces across On and Off sensory cells for single worm. Middle: Average fluorescence traces across SMDV (purple) and SMDD (magenta) for the same worm. Bottom: Average fluorescence traces across AVA (green) and RME (yellow) cell clusters. B: Fluorescence traces for the On, Off, SMDV, SMDD, AVA, and RME cell clusters. Colors match those of A. AVA trace in A is the average of the AVA traces shown in B. C-left: number of cells in each cell cluster identified across worms for OH15500 strain. C-right: sample Neuropal imaging z-slice. A RME cell is annotated.

benefit of this redundancy is unclear. Perhaps this cluster will split into subclusters if multiple independent signals are presented to the worm. Building a setup capable of controlling several stimuli independently is the most obvious chip-based extension to this work. It is also expected that more cell clusters will appear as analysis of freely-moving *C. elegans* matures.

There are several natural questions to ask of this dataset. The sensory cells follow the stimulus nearly perfectly. What are the dynamics of these cells? Do they habituate? The other natural questions focus on the motor command neuron clusters. Kato et al [28] discuss how none of these cell clusters have an obvious relationship to oxygen stimulus. They lack an obvious relationship to food stimulus as well. Part of the complexity is that these motor command clusters activate spontaneously. But, perhaps they are driven by food stimulus in a probabilistic manner. Moreover, it seems likely that they are controlled in a state-dependent manner.

2.4 Methods

2.4.1 Full-head imaging

Day 1 adult *C. elegans* expressing GCaMP5k (Zim294) or GCaMP6s (OH15500) are imaged using a Zeiss Airyscan. Imaging is done in 2 micron z steps. In ‘Fast’ mode, the Airyscan images the entire head of the adult worm at about 1.5 volumes per second.

2.4.2 Microfluidic Chip

Worms are loaded into a microfluidic chip for airyscan imaging [11]. The microfluidic chip accomplishes two goals. First, it controls odor exposure to the worm’s head. Secondly, it constrains worm movement, keeping the worm’s head in a single imaging volume for long trials. Here, worms are typically imaged for about ten minutes. All microfluidic chip channels are loaded with M9 solution. The movement of untreated worms proved too difficult to motion correct. To compensate, 1.5 mMol of tetramisole hydrochloride is added to the loading channel. The combination of the chip and the tetramisole paralytic eliminates most perceivable worm movement.

2.4.3 Food Imaging

C. elegans are raised on OP50 lawns. On day 1 of adulthood, they are washed in M9 and loaded into a microfluidic chip. Each worm is left to sit in the loading chamber for 10 minutes before the beginning of an imaging trial. The microfluidic chip is used to expose the worm to

OP50 in order to study food responses. OP50 is prepared for food stimulus imaging according to [10]. Briefly, two week old OP50 is spun down in a centrifuge. The pellet is then resuspended in M9 solution. All resuspension is done in a 1:1 ratio, meaning that the final volume of resuspended OP50 matched the initial volume of OP50.

2.4.4 Precise Food Exposure Patterns

An arduino is used to control the pattern of food exposure to the worm. The arduino sends ttl pulses to a valve controller. This controller dictates whether the food solution is routed to the nose of the trapped worm or away from the worm. In this way, worms are exposed to binary patterns of food stimuli. Multiple patterns are used in this study. The patterns vary in correlation lengths and transition probabilities.

2.5 Related Work

2.5.1 Imaging Simpler Organisms

Although large-scale calcium imaging has improved significantly in the last two decades, it is still mostly limited to studies of hundreds or thousands of neurons [3]. It also has depth limitations due to light scattering [24]. One approach to ease the multi-neuron calcium imaging task is to use smaller, less-complex organisms. Many simpler organisms have been studied over the years, including leech [30], and crab [23]. However, in recent years, the imaging field has coalesced around three simpler organisms: *Drosophila*, larval Zebrafish, and *Caenorhabditis* elegans. *Drosophila* and *C. elegans* are particularly appealing due to the extensive genetic toolkits developed for these organisms. Although zebrafish genetics are not as well understood, the fish perhaps promises a greater reward due to its complex behavioral repertoire.

The simplicity of these invertebrate species enables current imaging technologies to capture information from many neurons at the same time. The small size and translucent cuticle of *C. elegans* allow off-the-shelf technologies to be used for 'whole-brain' worm imaging. Spinning-disc confocal appears to be sufficients for imaging both trapped [28] and moving [36] *C. elegans*. Also, it is becoming increasingly common for scanning confocal microscopes like the Zeiss Airyscan to scan multiple lines simultaneously. As a result, these off-the-shelf microscopes are becoming fast enough to capture *C. elegans* 'whole-brain' data with sufficient temporal resolution.

Other imaging technologies promise worm imaging at even faster speeds. Prevedel et al successfully use light-field imaging to capture volumetric images from *C. elegans* at extremely fast sampling rates [43]. According to Ahrens and Engert, 'the temporal resolution [of light-field imaging] is limited only by the camera and photon count' [3]. The translucent *C. elegans* cuticle enables the usage of light-sheet microscopy. Light-sheet microscopy uses multiple objectives to quickly obtain imaging slices with excellent spatial resolution.

On top of the increasing availability of microscopes with high spatial and temporal resolution, genetically encoded calcium indicators have improved significantly since the early days of GCaMP [3]. GCaMP is increasingly sensitive at faster temporal resolutions. Moreover, the well-understood genetics of *C. elegans* have allowed further GCaMP improvements. The fusion of nuclear localization tags to GCaMP have significantly aided in *C. elegans* imaging analysis [43]. The nuclear-localized gcamp signal limits the signal overlap between adjacent neurons, easing the task of image segmentation.

2.5.2 Microfluidic Chips

Although simpler organisms provide an easier imaging target, they can be difficult to prepare for imaging. At this point it is difficult to acquire high quality calcium imaging data from moving simple organisms and there is no way to head-fix a worm of a fish. Fortunately, microfluidic chips for both worms and fish have progressed significantly in recent years. According to

[1], ‘microfluidic devices evolved from the microfabrication methods used for electronic chips’. These devices are typically fabricated by shaping a silicone elastomer with a mold. This process creates channels in the silicone through which fluids can flow. The Bargmann lab has developed a number of standard molds for *C. elegans* calcium imaging [11]. The chips created using these molds typically have a number of different flow channels. A loading channel allows worms to be loaded into a chip. Odor and buffer channels allow for the precise exposure of odorants to the tip of the worm’s nose.

Microfluidic chips have a number of benefits on top of the ability to precisely deliver stimulus. The chip channels are shaped to trap the worm. The constraining of worm movement is important for limiting motion artifacts. It also allows for a smaller imaging volume, enabling faster volumetric acquisition. The chip also allows for the continuous flow of salts required for hospitable worm environments. Finally, it allows for the continual flow of paralytic to the worm, further reducing motion artifacts and allowing for long imaging sessions.

2.5.3 Post-Processing

Commensurate with microscopy and calcium indicator improvements, post-processing of neuronal imaging data has seen several recent improvements. Post-processing of calcium imaging data can be separated into two stages: motion-correction (also referred to as registration) and image segmentation. Piece-wise rigid registration has been successfully applied to in vivo mouse imaging data [41]. The use of rigid registration is particularly appealing due to its stability and simplicity. The piece-wide modification to rigid registration allows the algorithm to account for local non-rigid movements. This is particularly important for animals like *C. elegans* that lack both skull and exoskeleton. Even with limited motion, the worm will often squeeze through small spaces, producing thoroughly non-rigid motion artifacts.

Mukamel et al successfully combine independent components analysis (ICA) with image segmentation to obtain separate calcium traces from Purkinje cells [34]. Using both temporal

(ICA) and spatial information is critical. Even with nuclear-localized indicators, it is common for the fluorescence signals of nearby neurons to spatially overlap. Temporal information can be used to separate these cells if they fire independently. The use of spatial information is also important as calcium imaging data (and neural data in general) is often composed of many statistically dependent signals. Nonnegative matrix factorization (NNMF) has been successfully used to segment in vivo imaging data [42]. This approach factorizes imaging slices into two positive matrices: a spatial matrix and a temporal matrix. Pnevmatikakis et al claim their constrained NNMF strategy handles overlapping spatial fields better than the PCA/ICA approach [42].

2.5.4 *C. elegans* Cell Identification

This manuscript, along with most C. elegans papers, repeatedly refers to *C. elegans* cells by three or four lettered identifiers. This is largely due to the Nobel prize winning work of Sulston et al [51]. The authors traced the entire cell lineage of the worm from zygote to newly hatched larva. Sulston et al found cell fates to be highly invariant within the N2 worm strain. This invariance is made possible by the fact that *C. elegans* reproduce via self fertilization in the lab setting. Every studied worm for a particular strain is nearly genetically identical (with some genetic drift through generations). Promoters have been identified for cell-specific genetic expression for most *C. elegans* cells. Thus, when a study refers to the SMDV cell, for instance, they are referring to a specific cell that has been mapped from embryogenesis and is identified by some set of promoters.

2.5.5 NeuroPAL

Large-scale calcium imaging studies rarely expend significant effort generalizing findings across animals. Due to the expense of generating, housing, and analyzing these datasets, large-scale calcium imaging studies rarely have data from more than a few animals. Early *C. elegans* large-scale imaging studies only consider a handful of worms. The Kato et al [28] dataset includes data from 5 worms. However, as technology advances, it is becoming easier to image and analyze large numbers of worms. Thus, it is becoming more important to combine data across worms. Smartly combining data across animals will increase the statistical power of any findings. It will also allow for the discovery of general ‘worm properties’.

The NeuroPAL strain is a very promising tool for combining data across worms [57]. The NeuroPAL strain expresses various fluorophores under a number of different promoters. The fluorophore-promoter combinations are chosen so that each neuron exhibits a fluorescence profile different from its neighbors. This is made possible by the position stereotypy of *C. elegans* neurons. The NeuroPAL strain enables cell identification. If the same cell is assumed to have the properties across worms (within a strain), data can easily be combined across worms.

2.6 Acknowledgements

Chapters 2 is from a manuscript in preparation: Context-dependent calcium dynamics in the *C. elegans* nervous sytem. The full author list is Zachary Cecere, Tatyana Sharpee, and Sreekanth Chalasani. The dissertation author is the primary investigator for this paper.

Chapter 3

Adaptation in sensory neurons

3.1 Introduction

Organisms have a finite capacity to represent information. They cannot efficiently represent the full set of possible stimuli. Instead they only encode stimulus features relevant to their current environment. *C. elegans*, with its limited nervous system, must make especially good choices about how it represents its surroundings. For instance, the worm will sense odorants at drastically different concentrations over its lifetime. Dedicating a set of neural responses to each possible concentration is clearly infeasible. Instead, the worm's sensory neurons adapt to its recent stimulus exposure.

Kato et al [29] fit an adaptation model to *C. elegans* On and Off sensory cells. Their model (the cascade model) assumes the calcium dynamics of *C. elegans* sensory cells can be approximated as the difference between the solution to two ordinary differential equations (ODEs) of the form:

$$\frac{dx}{dt} = \tau * (Inp(t) - x(t)) \tag{3.1}$$

τ is the time constant. It determines the stimulus response waveform. Taking the difference of the solutions to two ODEs of this form produces a cancellation model. That is, in response to a stimulus step change, the calcium level will briefly increase before returning to its baseline value. In this way, the cascade model assumes that *C. elegans* sensory neurons encode changes relative to a locally averaged representation of their environment. The relative timescales in the model dictate the time over which the environment is averaged.

This simple model achieves reasonable performance for sparser stimulus patterns but struggles with more correlated pulse patterns [29]. Kato et al hypothesized that using more ODE equations could remedy these fit issues. It is also possible that some of the fitting difficulties stem from the extreme nonlinearity of the cascade model. It is possible that an autoregressive model might ease some of these difficulties. The autoregressive model linearly relates the current stimulus value and the recent stimulus history to instantaneous changes in calcium level. Within a stimulus pulse, *C. elegans* neurons behave like first order ODEs. Thus, linearly regressing the autoregressive covariates against an approximation of the calcium derivative provides a good initial estimate of parameters. A good initial estimate is crucial to fitting highly nonlinear models.

3.2 Results

3.2.1 Core Model

Both On and Off sensory cells are assumed to follow simple ordinary differential equation kinetics. The change in calcium is linear in terms of the stimulus input and the current calcium level.

$$\frac{dx}{dt} = \vec{\alpha} \cdot \vec{s}(t) - \beta x(t) \tag{3.2}$$

Where $\vec{s}(t)$ refers to the stimulus terms. The stimulus terms are composed of the current stimulus value as well as the several recent windows over which the stimulus is averaged. These window terms are designed to capture the adaptation effect. $\vec{\alpha}$ and β are the parameters to be optimized.

There are two parts to the GCaMP model: the nonlinear transformation and the linear convolution. First, the calcium value is passed through a nonlinearity and taken to the power of p. This transformation is designed to capture both the nonlinear binding kinetics of GCaMP and the limited range of GCaMP. The transformation is very similar to the one used in Kato et al [29]. Finally, the transformed calcium is convolved with the cell-specific GCaMP filter discussed in section 3.4.1.

The model is fit to training data by successive approximation. Initial parameter values are estimated using linear regression (section 3.4.4). The parameter estimates are refined using the combination of an Unscented Kalman Filter (UKF) and a random search algorithm (sections 3.4.2, 3.4.4, 3.4.5).

3.2.2 Model Performance

The model has four hyperparameters: the GCaMP nonlinearity power p, the observation noise variance r, the calcium noise variance q, and the pulse-correlated noise variance eta (section 3.4.3). Values for these hyperparameters are determined using cross-validation. For each hyperparameter combination, the model is fit to the first half of trials for a subset of the cells. The fit models are then used to simulate a number of outcomes for the second half of each trial. Hyperparameter sets are chosen separately for On and Off cells such that the average mean squared error between the simulated and observed fluorescence values is minimized. The GCaMP nonlinearity parameter p has the strongest effect. For any reasonable combination of the noise hyperparameters, changing the value of p from 1 to 3 reduces the cross-validation error by 25% in primary On cells. Using 1 for the value of p overestimates the adaptation effect in On cells. High values for the noise parameters ($> 1e-4$) tend to cause the model to vastly overshoot. Using different values below $1e-4$ for these hyperparameters affected small changes in the cross-validation error.

For the remaining set of cells, the model was fit to the first half of the trial and simulated on the second half of the trial. The hyperparameters gleaned from the cross-validation experiment are used for these fits. Figure 3.1 shows the simulation results for several sample cells. The mean adaptation model produces reasonable results for both On and Off cells. It performs well across a variety of different stimulus patterns. The combination of the low error and low model noise suggests there is relatively little noise in the imaging setup; that worm movement is negligible and the stimulus delivery is consistent.

Figure 3.2 shows the generated calcium and fluorescence values for the same trials as 3.1. Each subplot compares the calcium and fluorescence traces for a single simulation. The most obvious trend is that the observed fluorescence values dramatically overstate the calcium range of the On cells. The time course, on the other hand, is similar between calcium and fluorescence. This is because the volumetric sampling rate is slow relative to the dynamics of GCaMP5k.

Figure 3.3 shows the distribution of window average coefficients across worms. It should be noted that this figure does not include the coefficients for the current input effect, which are obviously strongly positive for On cells and strongly negative for Off cells. On cell window average coefficients exhibit a strong inhibitory bias while Off cell coefficients follow an excitatory trend. This excitatory trend (referred to as the rebound effect) causes both habituation of the inhibitory signal and strong rebound after stimulus removal. These findings match the expectation for a mean adaptation model. That is, the cell's response to a pulse will equilibrate back towards zero over a long enough time period. The On and Off cells differ in their timecourses. The On cells tend to be unaffected by activity within the last ten timepoints (7 seconds), while the Off cells are strongly excited by activity within the last ten timepoints. The inhibition of On cells, overall, stems from activity over the last 7 to 100 seconds.

3.3 Discussion

Overall, the autoregressive model captures much of primary sensory cell behavior. Thus, On and Off cells are well described by mean adaptation. The autoregressive model shows that

Off cell rebound is faster than On cell habituation. These findings roughly match those of Kato et al [29].

The autoregressive model warns about underestimating GCaMP effects. Using a nonlinear filter to describe GCaMP improves cross-validation performance in the On cells. Smaller values for the GCaMP nonlinearity hyperparameter p cause the model to overestimate the strength of habituation. This can be seen most clearly in figure 3.2. For the fastest stimulus pattern (top-left), the On cell never accumulates very much calcium. As a result, the On cell is always operating in the insensitive GCaMP range, and the fluorescence dynamics look very small. This contrasts with the slowest stimulus pattern (one down - left) where calcium is able to accumulate. Cell fluorescence responses to these patterns look very different, while the inferred calcium responses actually operate over a similar range. Overall, it is easy to make mistakes about cell dynamics just by looking at fluorescence traces.

GCaMP inactivation confounds some of the findings for On cells. There is little evidence of permanent photobleaching as the majority of cells return to baseline fluorescence after a period of time. However, it is possible that GCaMP becomes temporarily inactivated as modeled by Kato et al. It is possible the model overstates the level of habituation in the On cells. That being said, GCaMP inactivation does not explain why On cell habituation tends to be slower than Off cell rebound. Similarly, Off cell rebound is incongruous with GCaMP inactivation. In fact, if anything, the model likely underestimates the effect of stimulus history on Off cells.

Although one could try to explicitly model GCaMP inactivation, it is unlikely to make any difference. The only way to conclusively account for GCaMP inactivation is experimentally. The best way to do this would be to image and electrophysiologically record from neurons at the same time. Unfortunately, this is very difficult in *C. elegans*. Another way one could get at this question would be to test whether primary sensory cell habituation affects the way primary sensory cells drive downstream neurons.

The usage of the Unscented Kalman Filter enables analysis of the noise in *C. elegans* sensory cells. The observation and process noise appear to be very small. However, the fluorescence value reached in response to a particular pulse is difficult to predict. It might be the result of some deterministic process (better modeled by autoregressive terms) or it might be simply unpredictable (better modeled by eta noise process). This study addresses this tradeoff. However, this tradeoff can be analyzed further by regularizing the autoregressive terms and assessing the effect of regularization through cross-validation. This experiment is an important follow up.

One of the benefits of the autoregressive model is that it can easily incorporate signals from downstream neurons. This enables the asking of questions like: are On cells more sensitive when RME is in its high state? This is an important long term goal.

3.4 Methods

3.4.1 Estimating GCaMP Parameters

The GCaMP filter is modeled as a difference of exponentials with parameters matching Chen et al [12]. This procedure is complicated by the volumetric nature of the imaging data. For example, imagine a cell in slice 0 and a cell in slice 7 have identical response kinetics. The cell in slice 7 will appear to have much faster response kinetics due it being acquired over half a second later than the cell in slice 0. Thus, the slice in which the cell appears needs to be considered in the creation of its GCaMP filter:

$$\begin{aligned} f(t) &= \int_{t1}^{t2} g(x)dx \\ t1 &= (t + \frac{z}{n_z}) \\ t2 &= (t + \frac{z+1}{n_z}) \end{aligned}$$

$g(x)$ refers to the difference of exponentials and n_z refers to the number of z-slices. The normalized $f(t)$ filter is used in a convolution to map from calcium to fluorescence.

3.4.2 Incorporating GCaMP

Calcium dynamics are assumed to follow a linear, differential equation where the inputs are the various stimulus window terms. At each timepoint, 3.2 is solved using Euler's method. Gaussian noise is added to the ODE solution. The varianceof this Gaussian noise is the q hyperparameter. The noisified calcium is run through the relu transform and taken to the p power. Finally, the transformed state along with the transformation of the recent states are convolved with the GCaMP filter. Gaussian additive noise is again added to the fluorescence estimate. The variance of this distribution is the r hyperparameter.

$$\begin{aligned}
x_0(t) &= soln(\frac{dx}{dt}) + n_q \\
x_i(t) &= x_{i-1}(t-1) \\
y_i(t) &= relu(x_i(t))^p + n_r \\
f(t) &= \sum f_i y_i(t)
\end{aligned}$$

3.4.3 Eta Noise

The core model has two sources of noise: calcium noise controlled by the q hyperparameter and observation noise controlled by the r hyperparameter. Neither of these noise sources are temporally correlated. This is problematic given the appearance of the primary sensory neuron traces. These neurons exhibit very little deviance when at baseline. They also appear fairly stable at their peaks. There is, however, some variability in the saturation point across stimulus pulses. That is, a given primary sensory neuron saturates at different fluorescence values within a trial. Some of this can likely be explained by the adaptation terms but some of the variability may simply be noise.

The pulse-correlated noise system requires the introduction of the input delta covariate. The input delta term is 1 when the stimulus turns on, -1 when it turns off, and 0 everywhere else. At pulse onset, noise is injected into the pulse-noise state. The variance of this noise is equal to the hyperparameter eta. The pulse-noise state is updated according to $x_n(t) = x_n(t-1)$ until pulse offset is detected. At pulse offset, the pulse-noise state returns to 0. The pulse-noise state is combined with the stimulus terms to drive the calcium state $x(t)$. Thus, if the noise term is positive, the calcium response will saturate at a higher value.

3.4.4 Initial Fit

The combination of the linearity of the differential equation and the linearity of the GCaMP filter enables a simple initial estimate. This is important because even linear, Gaussian state space models have many local minima. Thus, having a good initial parameter estimate limits the size of the search space for the final nonlinear optimization. The Richardson-Lucy algorithm is used to deconvolve the normalized fluorescence signal [18]. The deconvolved fluorescence is taken to the (1/p) power to get an estimate for the underlying calcium value.

The calcium derivative is approximated by finite differences. A linear regression is performed between this approximate derivative, the associated calcium value (giving the initial beta estimate), and the various stimulus terms. This regression is done across all worms using a mixed effects formulation. There is assumed to be a mean worm (the fixed effects). The coefficients for each worm are assumed to be the sum of the mean worm's coefficients and its own regularized coefficients. The regularization strength is chosen to be the maximum strength where the model performance is within 95% of that of the unregularized model performance.

3.4.5 UKF and Random Search Fitting

The initial fit produces mediocre results when simulated. This is because the linear model does not account for error accumulation. The simulated model is fit to the observed fluorescence traces using a random search procedure. First, the procedure searches the parameter space near the initial estimate for the N parameter sets that produce the smallest mean error between a noiseless version of the model and the observed fluorescence traces. Then, the Unscented Kalman Filter (UKF) is used to assess the log-likelihood of each of these parameter sets. The set that produces the maximum log-likelihood is chosen. The fitting of the noiseless version of the model is done because simulating the noiseless model is faster than running the UKF. Thus, a much larger region of parameter space can be explored. It was found that the mean error of the noiseless fit correlates with the log-likelihood calculated by the UKF.

3.5 Related Work

3.5.1 C. elegans Sensory Neurons

C. elegans primary sensory cells are the best studied *C. elegans* neurons. Their ability to track odorant stimuli with high fidelity makes them good candidates for study. It is common practice to characterize a worm's response to an odorant by computing the change in fluorescence of its primary sensory cells to odor onset [10] [31]. Interestingly, the majority of these studies only consider the fluorescence change for a single pulse of odorant. This is done because *C. elegans* primary sensory cells habituate to odorants. Their responsiveness decreases with successive odorant exposure. The dynamics of this habituation are still only partially understood.

C. elegans encode stimuli using both 'On' and 'Off' primary sensory cells. The 'On' cells are excited by stimuli. The 'Off' cells are inhibited by stimuli. The 'Off' cells also exhibit a strong rebound effect. Calcium levels in these cells often rise well above baseline when a stimulus is removed [31]. Historically, it was thought that each primary sensory cell encodes a single stimulus (labeled-line coding). Leinwand et al found that several odors are encoded combinatorially [31]. Leinwand et al also showed primary sensory cells cannot be identified solely through pulse-based calcium imaging. The AWB neuron appears to be a primary sensory neuron for benzaldehyde. It exhibits a robust and transient response to benzaldehyde removal. However, this response requires signalling from ASE. Hence, this study focuses on 'sensory' cells: cells that strongly correlated with stimulus or with removal of stimulus. Some of these cells may be secondary encoders.

3.5.2 Modeling GCaMP

The cascade model highlights one of the biggest difficulties in modeling calcium activity: the nonlinear binding properties of GCaMP. GCaMP fluoresces when it is hit with light while binding four calcium ions. The nonlinear binding kinetics of GCaMP make events more obvious while complicating their analysis. For instance, Chen et al, finds that a nonlinear model is necessary to map from spikes in mouse neurons to calcium transients [12]. Unfortunately, it becomes difficult to say whether this effect is due to nonlinearities in the calcium transduction pathway or due to the nonlinear binding kinetics of GCaMP.

GCaMP is also problematic for analyzing sub-baseline activity. It cannot track calcium activity at low calcium concentrations with much fidelity. Kato et al handled the nonlinearity and range limitations of GCaMP by pushing the output of the cascade model through an exponential nonlinearity. Thus, calcium in their model can go well below zero while the model's fluorescence value is strictly positive. In fact, this nonlinearity is crucial to the functionality of their model.

3.5.3 Noise and the Unscented Kalman Filter

Any model of *C. elegans* fluorescence activity must battle multiple sources of noise. These sources include observation noise inherent to the microscope. They also include more complex noise like worm movement and imperfect odorant flow in the microfluidic chip. These latter sources of noise are more complex as they are likely to create noise structures that are correlated in time. For instance, if the chip becomes partially blocked, odorant exposure to the nose of the worm may be decreased for a period of time. This, in turn, will decrease primary sensory cell responsiveness for a period of time.

State-space models explicitly model noise. State-space models probabilistically describe the relationship between latent states and an observed state. Basic state-space models are built on the following two equations [54]:

$$\begin{aligned} x_{k+1} &= F(x_k, v_k) \\ y_k &= H(x_k, n_k) \end{aligned}$$

F is the transition equation and H is the observation equation. v_k is the process noise and n_k is the observation noise.

The process of estimating the current latent state from the history of observations is referred to as filtering. More specifically, filtering estimates $p(x_t|y_0,...y_t)$. Assuming $\hat{x}_{k-1}$ and y_k are Gaussian Random Variables, the following recursion provides the optimal estimate [54].

$$\hat{x}_k = x_k^- + K_k[y_k - y_k^-] \tag{3.3}$$

where

$$
\begin{aligned}
x_k^- &= E[F(\hat{x}_{k-1}, v_{k-1})] \\
K_k &= P_{x_k, y_k} P_{\tilde{y}_k, \tilde{y}_k}^{-1} \\
y_k^- &= E[H(x_k^-, n_k)]
\end{aligned}
$$

Where v and n are the process and observation noise terms respectively and $\tilde{y}_k$ is the difference between the optimal observation prediction and the true observation. When F and H are linear, the Kalman Filter optimally infers the latent state. Unfortunately, there is no reasonable linear model of GCaMP dynamics.

The Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF) generalize the Kalman Filter for nonlinear models. The EKF linearizes the dynamics equations in order to estimate the posterior covariance matrices. It is the de facto system for a number of well understood nonlinear state space models. The UKF represents the state distribution as a Gaussian Random Variable. It draws a set of sigma points from this distribution and propagates these points through the nonlinear transition function. The mean and variance are then calculated for the transformed points, producing the updated Gaussian distribution. A similar procedure is used for the observation transformation. The approximate mean and covariance matrices are used to find the optimal state using the above Kalman gain equations. The UKF often gives more robust estimates than the EKF [54] [26].

The combined usage of the autoregressive model and the Unscented Kalman Filter allow for the simultaneous analysis of dynamics and noise of the C. elegans sensory cells.

3.6 Acknowledgements

Chapters 3 is from a manuscript in preparation: Context-dependent calcium dynamics in the *C. elegans* nervous sytem. The full author list is Zachary Cecere, Tatyana Sharpee, and Sreekanth Chalasani. The dissertation author is the primary investigator for this paper.

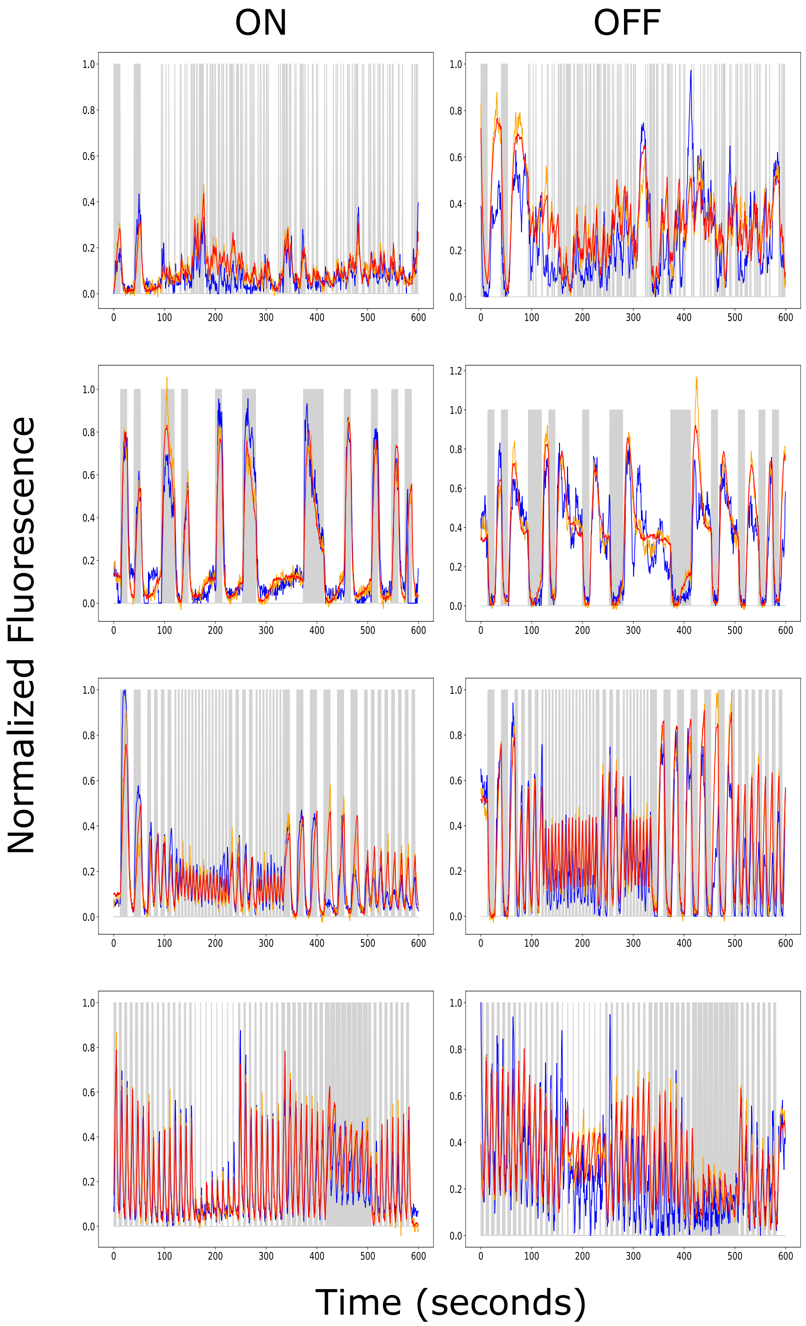

Figure 3.1: Test set performance of autoregressive sensory neuron model. In all subplots, blue is the observed fluorescence. Yellow is a single model simulation on the test set. Red is the average across a number of model simulations. All cells in the left column are On sensory cells. All cells in the right column are Off sensory cells. Each row corresponds to a different stimulus paradigm. The top two rows correspond to random pulse protocols with different block sizes. The bottom two rows correspond to correlated pulse protocols.

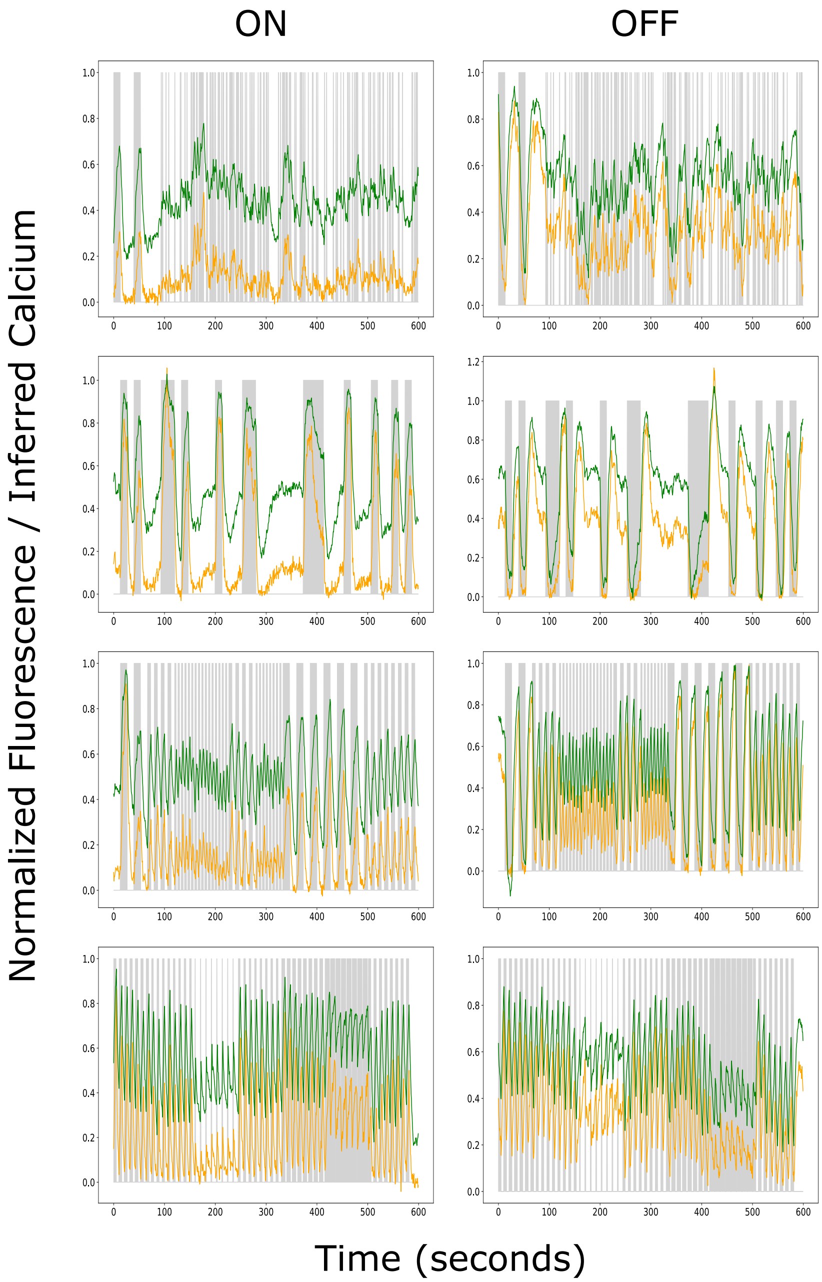

Figure 3.2: Comparison of model-inferred calcium to model-inferred fluorescence. There is a 1:1 correspondence between the traces in figure 1 and the traces in figure 2. Yellow traces are inferred fluorescence. Green traces are inferred calcium.

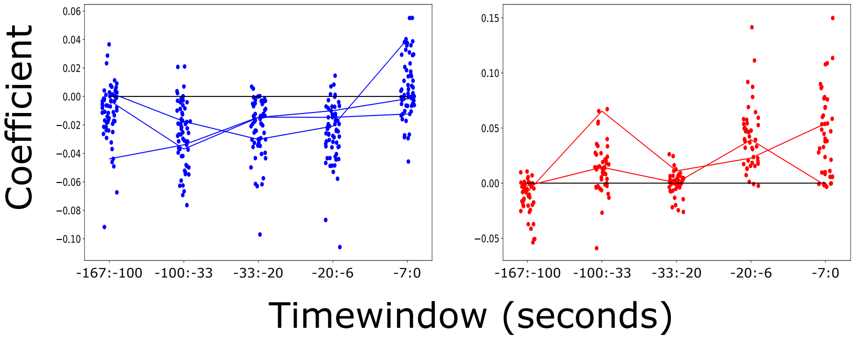

Figure 3.3: Distribution of autoregressive coefficients across sensory cells. Timewindow refers to the time period over which the stimulus is averaged. For instance, the -20 to -6.6 timewindow refers to the average stimulus from 20 seconds before a target timepoint to 6.6 seconds before. The lines connect coefficients derived for the same cell. Only a subset of the lines are shown for clarity. Timewindows are truncated to integers for clarity. The actual values for the timewindow endpoints are -166.67, -100, -33.33, -20, -6.67, and 0.

Chapter 4

State dependence in *C. elegans* command neurons

4.1 Introduction

Changes in *C. elegans* calcium levels are context dependent. *C. elegans* On sensory cells habituate to stimulus as a function of recent stimulus presentations. Similarly, the worm's Off sensory cells rebound more strongly with more stimulus exposure (section 3.2.2). *C. elegans* neuronal calcium fluctuations also depend on the network state. Some trivial examples follow from the boundedness of certain cell types. For instance, the RME cell cannot decrease in its low stable state and it cannot increase in its high stable state. Less trivially, a switching linear dynamical systems analysis found the *C. elegans* neuronal network occupies fourteen distinct states. In each of these states, the network obeys different dynamical rules [33]. Knowing both the state of the *C. elegans* nervous system and its stimulus exposure history is crucial to predicting *C. elegans* circuit behavior.

The goal of this section is to understand how a specific neuron's dynamics depend on the calcium activity of the whole *C. elegans* network; how changes in some sets of neurons cause calcium changes in a target neuron. Calcium imaging of a restrained worm is limited in its ability to resolve this communication. The true dimensionality of the data is small compared to the number of observed cells [28] [33]. The RME and AVA cells have many, strongly correlated partners (seciton 2.2). RME and AVA are strongly anticorrelated. Do they act on each other directly? Do they communicate through one of their correlated partner neurons? The calcium data in the restrained worm is ill posed to answer these questions. It is much better equipped to answer how the RME cell cluster is related to the AVA cell cluster.

There are six cell clusters in restrained C. elegans imaging data: On sensory cells, Off sensory cells, SMDV, SMDD, the RME cluster, and the AVA cluster. The cell clusters that do not contain sensory cells, correlate with worm movements (section 2.2). While the RME and AVA cell clusters correlate with forward and backward locomotion respectively [28], the SMDV and SMDD cells correlate with ventral and dorsal head movements respectively [27]. Notably, each of these cell clusters exhibits stereotyped waveforms. The AVA and RME cell clusters are both bistable and nearly perfectly anticorrelated. The SMDV and SMDD cells are also heavily anticorrelated. Though, they each have one, low stable state.

The specific relationship between the SMDV/D cell cluster and the AVA/RME cell cluster is unclear. For instance, It is unclear whether changes in one cell cluster cause direct changes in another or whether they alter the probability of changes. This work seeks to predict event probabilities due to it being more general than assuming local linear relationships.

In this chapter, a number of modeling techniques are used to assess the functional relationship between cell clusters. All techniques model cell cluster fluctuation probability as a function of network state. The methods (generalized linear models, neural networks, and soft decision trees), vary in complexity. Some methods model network state as being linearly related to cell cluster fluctuation probability. Some allow the relationship between two cell clusters' calcium fluctuations to vary dramatically as a function of network state. All of the models allow for heterogeneity across worms via fixed or random effects (section 4.4.3).

4.2 Results

The primary goal is to understand how changes in the RME, AVA, and SMDV/D cell clusters depend on network state. The initial approach is to predict whether these cell clusters increase or decrease as a function of recent network history. To this end, the data is binned in units of 5 imaging volumes (˜3 seconds). A timebin length of 5 is reasonable because of the profile of stimulus pulse durations. Most worms are only exposed to stimulus pulses with a multiple of 5 imaging volumes for a duration. Therefore, if one assumes cell cluster event timing is in some way tied to stimulus timing, relatively little information will be lost. Moreover, estimating the change in fluorescence over 5 volumes is less noisy than calculating the change in fluorescence over a single volume.

The various cell clusters have, at best, a weak correlation with the input stimulus. Since some stimulus onsets produce no obvious network change, it is more likely that stimulus affects the probability of a network change. With this in mind, a logistic regression model is developed to predict network change. Logistic regression requires discrete outcomes. A blind thresholding system is used to convert changes in cell cluster fluorescence over a timebin to discrete data (section 4.4.1). The goal of the model becomes predicting the probability of a given cell cluster increasing above a threshold (or decreasing below a threshold) as a function of recent network history.

Two different strains are imaged. The NeuroPAL (OH15500) strain is designed to allow for cell identification and pan-neuronal imaging simultaneously [57]. A different pan-neuronal GCaMP strain is also imaged (ZIM294) [43]. The OH15500 animals have lower brood sizes. They also move much less in the microfluidic chip. Given these differences, the two strains should be analyzed separately. It is possible that they obey different rules.

4.2.1 Making generalized linear models context-dependent

C. elegans calcium imaging data poses two difficult problems: it is inherently nonlinear and it combines data from possibly dissimilar worms. The nonlinearity stems from the fact that worm neurons behave differently in different network contexts. In order to capture this nonlinearity, a clustering approach is used (section 4.4.2). A Gaussian Mixture Model is used to divide network states into different clusters. Based on the preceding 16 seconds of network activity, each timebin is assigned to a single cluster. The one-hot encoding of cluster identity enables the linear logistic regression module to capture nonlinear features.

Depending on the number of clusters, there may be substantial network variation within a cluster. Consider a cluster in which RME is decreasing. The probability of whether RME will continue to decrease depends heavily on the fluorescence value of RME just before the target timebin. If RME is closer to its high stable state, the probability of it decreasing is high. If it is close to its low stable state, the probability will be low. One could divide this cluster into two more clusters, but it is also possible that there is a linear relationship between network state and RME decrease probability within the cluster. Within each network state cluster, the network state is projected onto the top N eigenvectors of the cluster's covariance matrix. In this way, the network history space is effectively divided into different subspaces on which Principal Components Analysis is performed. Each timebin only has access to the eigenvector projections corresponding to a single cluster (the cluster to which it is assigned). The balance between the number of clusters and the number of eigenvectors controls the degree of nonlinearity in the model. The linear combination of the eigenvectors scaled by the coefficients learned by the model is referred to as the receptive field. There is a receptive field associated with each cluster.

Worm to worm variability is likely small compared to more complex organisms. In the lab setting, *C. elegans* reproduce through self-fertilization [51]. All members of the same strain are nearly genetically identical. With that being said, different worms may still have different sensitivities and different neuronal connectivities. Random effects are added to the logistic regression module through a regularization system (section 4.4.3). Each model coefficient is able to vary across worms. The degree of variation is controlled by L2 regularization.

Overall, the model is governed by four hyperparameters: the number of clusters, the number of eigenvectors per cluster, the inter-worm variation regularization strength, and an additional regularization parameter for all coefficients. In general, the model will be fit and cross-validated using only network history. Then, the stimulus will be added to assess model performance improvement.

4.2.2 The RME cell cluster fluctuates in a state dependent manner

A cross-validation approach is used to study the tradeoff between the number of clusters and the number of eigenvectors in predicting RME cluster decrease probability. This is initially done while ignoring the stimulus and with only OH15500 animals. For a number of cluster - eigenvector combinations, the out-of-sample bootstrap is used to assess model generalization [52]. This approach specifically measures the out-of-sample log-likelihood. It should also be mentioned that different regularization strengths are tested for each cluster - eigenvector combination. For each cluster - eigenvector combination, the regularizers that produce the highest average out-of-sample log likelihood are chosen. Figure 4.1 shows the log-likelihoods for these chosen regularizers. The model performance is fairly similar as long as it uses more than one cluster and at least twenty-four eigenvectors.

It is difficult to understand how well the model is performing from log-likelihood alone. The logistic regression scores are assessed for each of these hyperparameter combinations. The scores range from 80% success (with 1 cluster) to low 90%s for the better hyperparameter combinations. Thus, RME cell cluster decrease can be predicted fairly well without knowledge of the stimulus. This finding is likely the result of the stereotyped waveform of RME. RME is restricted to certain, predictable calcium trajectories. The stimulus may be able to influence how RME moves through its trajectories but it cannot produce significant changes in RME's waveform.

RME cell cluster fall probability is modulated by recent RME and AVA activity. The red cluster is the only cluster that increases RME fall probability (figure 4.1). It is also the only cluster in which the RME cell cluster is in its high state and is decreasing. Interestingly, this cluster is also the only cluster where neither SMDV nor SMDD are increasing. The blue cluster has the strongest inhibitory effect. This is unsurprising as RME is in its low state in this cluster. There is an interesting comparison between the green and purple clusters. In both clusters, the RME cell cluster is increasing towards its high stable state. One might expect the green cell cluster to be associated with a high probability of RME decrease because RME is higher in the green cluster than in the purple cluster. The fact that this is not the case suggests that either RME transitions can be interrupted or that SMDV and SMDD might have some effect.

The latter hypothesis is confirmed by the receptive fields (figure 4.1). The receptive fields are nearly identical across network history clusters. For every cluster, AVA increase, RME decrease, and SMDV/D decrease raise the RME fall probability. Thus, SMDV and SMDD gate RME high to low transitions. They must return to baseline for RME to return to its lower baseline.

The receptive fields shown in figure 4.1 are averages across all of the OH15500 animals. Plotting the receptive fields for each worm on top of these averages barely changes the plot. This occurs because the cross-worm regularization strength is very high in the best performing models. The coefficients for the network history clusters also vary little across worms. The clusters are distributed fairly equally across all worms. Each cluster has a number of assigned datapoints from each worm. Most of the cross-worm variability is captured in the intercept term. Altogether, most of the worms are well described by the clusters and receptive fields in figure 4.1.

The cross-validation approach is repeated for RME fall prediction in ZIM294 animals. The results are similar to the OH15500 results. Once again, the best performing model uses four network history clusters. The clusters tell a fairly similar story. When the RME cell cluster is decreasing from its high state, it is more likely that RME will continue to decrease (blue and purple clusters). The receptive fields look nearly identical to those of the OH15500 animals. Also like the OH15500 animals, there is very little cross-worm variation in model coefficients.

The stimulus is considered for the best performing RME fall prediction model. Two stimulus windows are added to the model. The latest window averages the mean-centered stimulus from ˜3 seconds before timebin onset until the end of the timebin. Similar stimulus windows are introduced in ˜6 second increments before the latest window. Out-of-sample bootstrapping is performed to assess the change in log-likelihood due to stimulus inclusion. Inclusion of the stimulus improves model performance in every sample for both the OH15500 and ZIM294 worms (figures 4.2 and 4.4).

For both OH15500 and ZIM294 animals the stimulus negatively correlates with RME decrease probability. The last stimulus timewindow has a strongly negative coefficients for both strains. The earlier stimulus window seems to provide little benefit as it is effectively zero. Altogether, stimulus removal increases RME fall probability. This effect seems to be mostly immediate. In summary, RME cell cluster decrease is a function of SMDV, SMDD, and stimulus history. RME falls to its lower baseline value in concert with SMDV and SMDD quiescence and stimulus removal.

4.2.3 RME rise dynamics are symmetric to RME fall dynamics

RME rise probability is modeled in a similar way to RME fall probability. The trials are broken up into timebins of 5 volume durations. The blind thresholding system finds .01 to be a reasonable threshold choice for RME increase. That is, if the RME cell cluster increases by at least .01 (units of normalized fluorescence) over a timebin, the timebin is scored as positive.

The out-of-sample bootstrap finds fewer clusters are needed for RME increase prediction. Two clusters performed best for the OH15500 animals, while only one cluster is needed for ZIM294 animals. The two OH15500 clusters show RME rise dynamics are governed by the exact opposite forces as RME fall dynamics (figure 4.5). Unsurprisingly, increases in RME predict more increases in RME. Increases in SMDV and SMDD also increase the probability of a RME cell cluster increase. The receptive fields are nearly identical across all clusters and across both worm strains. They support the same finding: RME increases in concert with SMDV and SMDD.

Two stimulus windows are added to the best performing models for the OH15500 and ZIM294 animals. These are the same stimulus windows introduced in the RME fall prediction models. The inclusion of these stimulus windows improves performance in every out-of-sample bootstrap for the OH15500 animals. The inclusion of stimulus windows improves performance in all but one of the bootstraps for ZIM294 animals. For both strains, the coefficient associated with the latest stimulus window is strongly positive. The earlier stimulus windows are weakly negative, suggesting the stimulus effect on RME cell cluster increase is strongest at pulse onset. Altogether, RME fall and RME rise seem to be governed by symmetric dynamics. If the stimulus is off when SMDV and SMDD fall back to their baselines, RME will tend to fall back to its lower stable state. If the stimulus turns on while SMDV and SMDD are rising, RME will tend to rise to its high stable state.

4.2.4 SMDV and SMDD exhibit anti-correlated dynamics

The rise prediction procedure is used for the SMDV cell cluster. The SMDV cell cluster is composed of the worm's two SMDV cells. A logistic regression model is built to predict whether the SMDV cell cluster fluorescence will increase by at least .01 over a ˜3 second duration. Initially, only the recent network history is available to the model. As in the RME prediction models, the model uses network history clusters to capture nonlinear features and worm-specific random effect terms to capture cross-worm variability.

Out-of-sample bootstrap cross-validation produces very differing results between the two worm strains. The ZIM294 animals are best described by 6 network history clusters, while the OH15500 animals are best described by 2. The OH15500 network history clusters bear similarities to the ZIM294 network history cluster. The blue OH15500 cluster resembles the blue, purple, and yellow clusters of the ZIM294 animals. In all of these clusters, SMDV is increasing and the RME cell cluster is increasing (or has recently increased). These network history clusters are all associated with increased probability of SMDV rise. The red OH15500 cluster resembles the orange ZIM294 network history cluster. In these clusters, RME and SMDV fall predicts reduced probability of SMDV rise.

The receptive fields are similar across all clusters and across all worm strains. RME rise, SMDV rise, and SMDD fall increase the probability of SMDV increasing over the target timebin. The inhibitory relationship between SMDV and SMDD is unsurprising as these two cells almost never increase at the same time. The effect of RME cell cluster history on SMDV is more surprising. The receptive fields suggest SMDV events will tend to be triggered when RME transitions from its low to high state.

Although all of the network history receptive fields are similar, there are some important differences as well. In the ZIM294 red and blue clusters, sudden increases in SMDV will increase the probability of further SMDV rise. However, it is important to note that the SMDV receptive fields for these clusters are mostly negative. Thus, if SMDV begins to saturate in these clusters, the probability of SMDV rise will decrease. The clustering system is designed to capture this kind of nonlinear behavior. When SMDV begins its rising phase, it will tend to continue rising because SMDV activity is highly autocorrelated. However, as SMDV increases, the probability of it continuing to rise decreases.

The rise prediction procedure is applied to the SMDD cell cluster. In both strains, the network history cell clusters with the strongest positive effect on SMDD rise probability are those in which SMDD is already rising. Similarly, the clusters with the strongest negative effect on SMDD rise probability are those in which SMDD has recently peaked. The similarity of effect between the green and purple clusters in the ZIM294 worms suggests RME/AVA state plays a negligible role in SMDD rise probability.

The receptive fields tell a similar story to that of the SMDV neurons. In most clusters, the strongest receptive field components are the SMDV and SMDD components. Typically, sudden increase in SMDD and/or sudden decrease in SMDV portends further increase in SMDD. SMDV and SMDD are mutually inhibitory. Once again, the SMDD receptive field components are largely negative. If SMDD rise begins to slow, it is unlikely to continue increasing.

4.2.5 A neural network confirms stimulus importance

The generalized linear model suggests OP50 exposure increases RME rise probability and stimulus removal increases RME fall probability. However, it is possible that this effect is a function of artificial modeling conditions or poor null model performance. A temporal convolutional neural network is used to address this possibility (section 4.4.5). With its state-of-the-art performance on multi-class classification problems, it is much better equipped to capture *C. elegans* network dynamics than the modified, generalized linear model.

Changes in RME fluorescence over 5 volume timebins are discretized into five classes. The class boundaries are -.03, -.01, .01, .03. For instance, a given timebin falls into class two if the RME cell cluster average fluorescence changes by between -.03 and -.01 over the timebin. Moving from a single-class to a multi-class system allows for a greater amount of RME fluctuation variance to be captured. Neural network performance is compared with and without consideration of sensory cells. Sensory cell inclusion leads to significantly improved model performance for both OH15500 and ZIM294 strains (figure 4.13). The full model neural networks are pretrained without sensory cell information. This was found to improve model performance, suggesting that the neural network can easily overfit to stimulus. Overall, it appears that stimulus does have a significant effect on RME (and, by transitive property, AVA).

4.2.6 Soft decision trees find useful network states

In an earlier section, generalized linear modeling was found to be improved by the introduction of network state variables. One might expect the effect of stimulus on RME cell cluster fluctuations to be state dependent. To this end, the RME datasets are split into different datasets according to network state. Surprisingly, consideration of the stimulus had a similar effect on network performance in each of the filtered datasets. Moreover, the stimulus coefficients are found to be relatively similar across the different filtered datasets. This finding makes little sense. Stimulus removal should not have a large effect on RME fall probability when RME is in its low state.

The lack of a differential filter effect is likely the result of the core issue with the generalized linear model framework: the network history clusters are derived in a pre-processing step. A superior model would find network states that directly maximize RME prediction performance. Neural networks do this implicitly, but they are uninterpretable. Decision trees, on the other hand, can learn nonlinear features in an interpretable manner. They make predictions by segmenting data into subsets. For instance, a two-level decision tree divides a dataset into two subsets. It then makes class predictions for each of the subsets. In this way, the upper levels of the decision tree can be thought of as network state representations.

A soft decision tree (section 4.4.6) [21] is fit to the multi-class RME OH15500 dataset described in section 4.2.5. Soft tree cross-validation performance is compared to generalized linear multinomial performance for a number of tree depths. For all models, worm identity is fed into the model via fixed effects. Shallow trees (tree depth of 2) are nearly always worse than the linear model (figure 4.14). Trees of depth 3 are nearly always better than the linear model and trees of depth 4 are always better than the multinomial linear model. Thus, the hierarchical tree structure is beneficial for RME prediction.

Figure 4.15 visualizes a depth four decision tree fit to the whole dataset. Two interesting leaves are annotated as A and B. Leaf A follows directly from the branch with the largest magnitude Off cell increase. When this branch is exposed to a large magnitude OFF cell increase and low SMDV/D activity, leaf A becomes activated (it receives a higher probability value from its parent). This leaf predicts the largest magnitude decrease in RME. Thus, leaf A agrees with earlier analyses that found RME falls in accordance with low SMDV/D and high Off cell activity. Going up the tree, it can be seen that the branch leading to leaf A will be activated when RME is falling (top level). However, if RME is too low, this branch will not be evaluated (second level). This also agrees with earlier analysis. If RME is falling, it will continue to fall. If it is nearing its low stable state, it will tend to stop decreasing.

Leaf B tells a story opposite to that of leaf A. The branch leading to leaf B will be evaluated if RME, SMDV, and/or SMDD are increasing (second level). This branch will also tend to be evaluated if the On cell has been high for the last 5-10 imaging volumes (second level). This branch, in turn, will cause leaf B to be evaluated if RME, SMDV/D, and the Off cells are low. Since leaf B predicts the largest magnitude increase in RME, it is once again found that large increases in SMDV/D are coupled with large increases in RME. The one difference from earlier findings is that this tree predicts the On cells have a delayed excitatory effect on RME increase.

4.2.7 Macro View: SMDV and SMDD gate RME fluctuations

RME rise correlates with SMDV and SMDD rise. How many RME rise events co-occur with SMDV and SMDD events? To answer this question, SMDV/D and RME cluster events are annotated. A cluster event initiates when its average fluorescence value eclipses 0.2. The cluster's event endures until its fluorescence falls below 0.2 or for 7 seconds (whichever is longer). SMDV and SMDD 'cover' a timepoint if they are undergoing an event during that timepoint. A trial's 'coverage' is the percent of the trial's timepoints that are covered by SMDV and SMDD. Similarly, an RME event is covered if the RME cell cluster eclipses its 0.2 threshold in a covered timepoint.

Nearly every RME cell cluster event is covered (figure 4.16). That is, nearly every RME cell cluster event co-occurs with either an SMDD or SMDV event. In ZIM294 animals, SMDV and SMDD typically only cover about 40 - 80% of trials. If RME cell cluster events are assumed to be randomly distributed, the probability of obtaining the number of observed covered RME events from the observed percent coverages is astronomically small. The same is true for OH15500 animals even though SMDV and SMDD cover a greater percentage of the trials in these worms. This greater coverage matches the observation that SMDV and SMDD events tend to be slower in OH15500 animals. A hierarchical bootstrap is used to estimate the exact probabilities, but no random sample ever covers the observed number of RME events in tens of thousands of samples.

4.2.8 Stimulus effect on RME is gated by SMDV and SMDD

Since, RME cell cluster transitions are heavily dependent on the state of SMDV and SMDD, it is possible to use SMDV and SMDD events to divide imaging trials into discrete windows. For instance, since RME typically enters its falling phase when SMDV/D are falling or quiescent, one can ask whether a given SMDV/D event will be the last one before RME falls to its lower baseline.

In order to test this, SMDV/D and RME cluster events are annotated. A cluster event initiates when that cluster's fluorescence eclipses 0.2. The cluster's event endures until the cluster's fluorescence falls back below 0.2 or for 7 seconds (whichever is longer). The RME fall dataset consists of every SMDV/D event that initiates within an RME event. A given SMDV/D event is scored as a 1 if it is the last event before RME transitions back to its low state.

The question becomes: does the stimulus activity during a given SMDV/D event predict whether it will be the last SMDV/D event before RME falls to its lower baseline. Since, SMDV and SMDD events typically last about 20-25 seconds, the stimulus average from SDMV event onset to ˜13 seconds after onset and from ˜13 seconds after onset to ˜26 seconds after onset are used as covariates. The simplicity of this model enables the usage of traditional mixed effects models. The lme4 package is used to fit random-intercept, binomial mixed effects models [5]. The likelihood-ratio test is used to assess whether the different stimulus average covariates are significant as fixed effects (section: 4.4.4). For the OH15500 animals, the 0-13 stimulus average window is not significant (1 degree of freedom, chi-squared value of 0.4085, p-value of 0.5227) and the 13-26 window is negative and significant at the .05 level (1 degree of freedom, chi-squared value of 7.1068, p-value of 0.007679). Similar results are obtained for the ZIM294 animals. The earlier stimulus bin is insignificant (1 degree of freedom, chi-squared value of 0.0128, p-value of 0.9099) and the later stimulus bin is negative and significant (1 degree of freedom, chi-squared value of 22.245, p-value of 2.4e-06).

Thus, the activity of the stimulus during SMDV and SMDD rising phases does not seem to matter. However, if the stimulus is off during SMDV and SMDD falling phases, it increases the probability that RME will transition to its lower baseline. In this way, SMDV and SMDD gate the effect of the stimulus on RME.

To test whether the effect of the stimulus varies significantly across worms, the later stimulus window is introduced as a random effect. These more complex models do not significantly improve model performance, producing likelihood ratio p-values of 0.7371 for the ZIM294 animals and 0.9169 for the OH15500 animals. This analysis agrees with the earlier finding that worms are homogenous in terms of their RME fall dynamics.

4.3 Discussion

Fluctuations in *C. elegans* motor command cells depend on network state context. The RME cell cluster changes in accordance with the SMDV and SMDD cells. RME only enters its rising phase during SMDV or SMDD events. It only enters its fall phase when SMDV and SMDD are either returning to their baseline or quiescent. Similarly, the effect of food stimulus on RME cell cluster fluctuations depends on network state. If RME is near its high state, stimulus removal increases RME fall probability. Symmetrically, if RME is near its low state, addition of food stimulus increases RME rise probability. The stimulus effect is further gated by SMDV and SMDD activity. Stimulus has little effect on RME fall probability when SMDV or SMDD is rising. Removal of stimulus only increases RME fall probability when SMDV or SMDD are in their falling phase.

SMDV and SMDD can be used to section *C. elegans* neural network activity into discrete events. RME/AVA fluctuate in accordance with SMDV/D phase. Similarly, stimulus effect on RME/AVA is gated by SMDV/D phase. Using SMDV and SMDD to break up worm traces into discrete units could be very useful for future coding studies.

The sensitivity of *C. elegans* cells to external stimuli depends on network state. Thus, many of the food pulses in this study have little effect on the network. By experimentally tying stimulus exposure to network state, one could increase the event rate of the *C. elegans* network. This, in turn, would increase the amount of information gleaned from worm recordings. Such an effort would require significant improvements to calcium data processing as the current motion correction and segmentation methods are too slow for live fluorescence trace acquisition.

4.4 Methods

4.4.1 Blind scoring

In order to perform logistic regression on the binned data, each bin must be assigned a score. For each cell cluster, each bin is assigned a 1 or a 0 depending on whether that cell cluster's fluorescence change across the bin exceeds a certain threshold. A proper choice for the threshold is critical. Consider the scoring of RME cluster increases. When the RME cell cluster transitions from its low state to its high state, it increases slowly before rising rapidly. If the chosen threshold is too high, the scoring system might totally miss the initial, slow rise phase of RME low to high transitions. Such a failure arguably misses the most important part of the RME trace. In order to handle this issue, a blind scoring system is used. When scoring the RME cluster, for instance, only the RME cluster is considered. The initiation point of each RME low to high transition is found manually. The associated time bin is considered for each of these initiation points. The RME cell cluster increases by at least .01 for the vast majority of these initiation points. This threshold captures SMDV and SMDD initiation points as well. A similar procedure is used for the RME fall phase (high to low transition). A threshold of -.01 is reasonable for capturing where RME falls over a given timebin.

4.4.2 Pre-processing: Gaussian Mixture Modeling and eigenvector projection

Clustering is used to capture nonlinear features. First a network history dataset is assembled. For each timebin beginning at time t, the fluorescence traces of RME, AVA, SMDV, and SMDD from time t - 16 seconds to time t are combined into a single vector. The set of all of these vectors comprises the network history dataset. Gaussian Mixture Modeling is used on this dataset, producing cluster centers that represent network trajectories. The mixture model is used to assign each timebin to a single cluster center. The one-hot encoding of these cluster identities are used as covariates in downstream modeling. This one-hot encoding enables the downstream generalized linear model to approximate nonlinear features. For instance, it is possible that similar network trajectories have drastically different effects on whether RME is about to initiate an event. The model can capture this if these similar trajectories fall into different clusters.

Gaussian Mixture Modeling generates a covariance matrix for each cluster. These covariance matrices are ready-made for capturing the variance within each network history cell cluster. Eigenvectors and their associated eigenvalues are derived for each cluster center. The network history vectors are then projected onto the eigenvectors with the largest magnitude eigenvalues. A time bin only has access to the eigenvectors of cluster center i if it is assigned to cluster center i. In this way, the one-hot encoding is extended to the eigenvector projection covariates. Essentially, this clustering system divides the space of network trajectories into different subspaces and then performs principal components analysis (PCA) on each subspace. The eigenvector projection system reduces the dimensionality of the network history representation in a principled manner.

4.4.3 Pseudo mixed effects model

A simplified mixed effects model is used to predict cell cluster changes. If worm identity (which worm a particular timebin comes from) is a grouping factor, each worm will get its own copy of each random effect. Each random effect is assumed to be sampled from a normal distribution with mean zero. Typical mixed effects models [5] find the value of the random effect for each worm and the variance of the distribution from which the random effects are sampled (section 4.5.1). This is done via a nonlinear optimization that struggles for complex models with lots of random effects. Here, a simpler approach is used. Every covariate is treated as both a random and fixed effect. For covariate i and worm j, the model multiplies covariate i by a coefficient that applies across all worms and by a coefficient specific to worm j. All coefficients are constrained by an L2 penalty. The magnitude of this penalty is the regularization strength. The worm-specific coefficients are controlled by an additional L2 penalty. The random effects regularization strength refers to the factor by which the regularization strength is increased for worm-specific coefficients. For instance, a regularization strength of 2 means that the L2-norm of the cross-worm coefficients is multiplied by 2 to form an L2 penalty. A random effect regularization strength of 3 means the L2-norm of the worm-specific coefficients are multiplied by 2*3 to form the random effects L2 penalty. A cross-validation approach is used to determine good choices for these regularization strengths.

4.4.4 Likelihood Ratio Test

The likelihood ratio test is a generally accepted test for assessing significance in mixed effects models [5]. The model is fit with and without a target covariate. Assuming the null hypothesis is true, the likelihood ratio between the fit without the covariate and the fit with the covariate asymptotically converges to the chi-squared distribution.

4.4.5 Temporal Convolutional Neural Network

Temporal convolutional neural networks have achieved state-of-the-art or near state-of-the-art performance on a number of time series prediction tasks [37]. They can capture arbitrarily complex features in a dataset. A temporal convolutional neural network is built by connecting a number of 1D convolutional layers and a number of dense neural network layers. A network with three convolutional layers followed by two dense layers is sufficiently complex to predict RME fluctuation probabilities from RME history, SMDV/D history, and present stimulus levels. A network more complex than this tends to overfit the dataset. Worm identity is submitted to the network via a one-hot vector.

4.4.6 Soft Decision Trees

A soft-decision tree is used to predict RME fluctuations from RME, SMDV/D, and stimulus history as well as from a one-hot encoding of worm identity. The soft-decision tree is a type of oblique tree [35] (section 4.5.2). The soft-decision tree architecture developed in [21] is used. Each branch of the soft-decision tree learns a hyperplane that separates the target classes (in this case, whether RME will increase by at least .01, by at least .03, etc.). Each branch outputs a probability of the left-side of the sub-tree being true for a particular datapoint. These probabilities are multiplied recursively down the tree structure. In this way, the outputs of the leaves (the endpoints of the tree that explicitly predict class probabilities) are scaled by the probabilities output by the above branches. With this construction, the soft tree is end-to-end differentiable. It is trained using stochastic gradient descent.

4.5 Related Work

4.5.1 Mixed Effects Models

The problem of combining large-scale calcium imaging data across animals remains unsolved. This is not surprising because most large-scale calcium imaging datasets consist of only a few animals. Solving this problem is crucially important for two reasons. First, an overall goal of systems neuroscience is to find information processing properties that generalize across animals and species. The second reason is more practical and especially relevant to the worm. Events may be sparse within a worm. For instance, a 10-20 minute imaging session may only have 10 or so RME cell cluster events. Thus, it makes sense to combine information across worms to increase statistical power in finding predictors of these events.

Mixed effects models are designed to combine information across subjects. Since their introduction in the early 20th century, they have been used in many different disciplines [19]. Mixed effects models use both fixed and random effects. Fixed effects are constant across subjects. Random effects vary across subjects according to a normal distribution with mean 0 and a parameter variance.

These models are very powerful for testing specific statistical hypotheses. However, they tend to have convergence problems for complex models with lots of covariates [5]. Thus, they are not typically used for model specification. This is problematic for predictions using worm network history as the important network history features are not known a priori. As an initial step, using a regularization strategy to both select important features and to constrain variation across animals is preferable. In this way, the regularized model serves as a pseudo mixed effects model.

4.5.2 Oblique Decision Trees

Decision trees are a powerful and interpretable tool for solving multi-class classification problems. Much of their power stems from the fact that they can represent nonlinear features with an interpretable, hierarchical tree structure. Decision tree variants have achieved state-of-the-art performance on a variety of classification takes [25] [6]. Decision trees typically select a single covariate from the dataset. They find a threshold for this covariate that maximally separates the target classes. They apply this procedure recursively until the leaves of the tree are mostly pure. Leaf purity means that applying the branch thresholds to the dataset separates one class from the rest.

Typical decision trees are not good candidates for predicting RME fluctuations from network history. This is because the typical decision tree does not handle the dimensionality of this task appropriately. Predicting RME fluctuation from the last 24 timepoints involves (24 timepoints) x (3 cells + 2 input dimensions) covariates (120 input dimensions). Due to the smoothness of the calcium traces, many of these covariates are highly correlated. Thus, a model that uses these covariates well should reduce the dimensionality of the input in some smart way. Randomly picking one of these covariates at each level of the tree is not a smart way. Oblique decision trees were developed for this case [35]. At each level of the tree, an oblique decision tree selects a hyperplane in the full input space that separates the different target classes.

4.6 Acknowledgements

Chapters 4 is from a manuscript in preparation: Context-dependent calcium dynamics in the *C. elegans* nervous sytem. The full author list is Zachary Cecere, Tatyana Sharpee, and Sreekanth Chalasani. The dissertation author is the primary investigator for this paper.

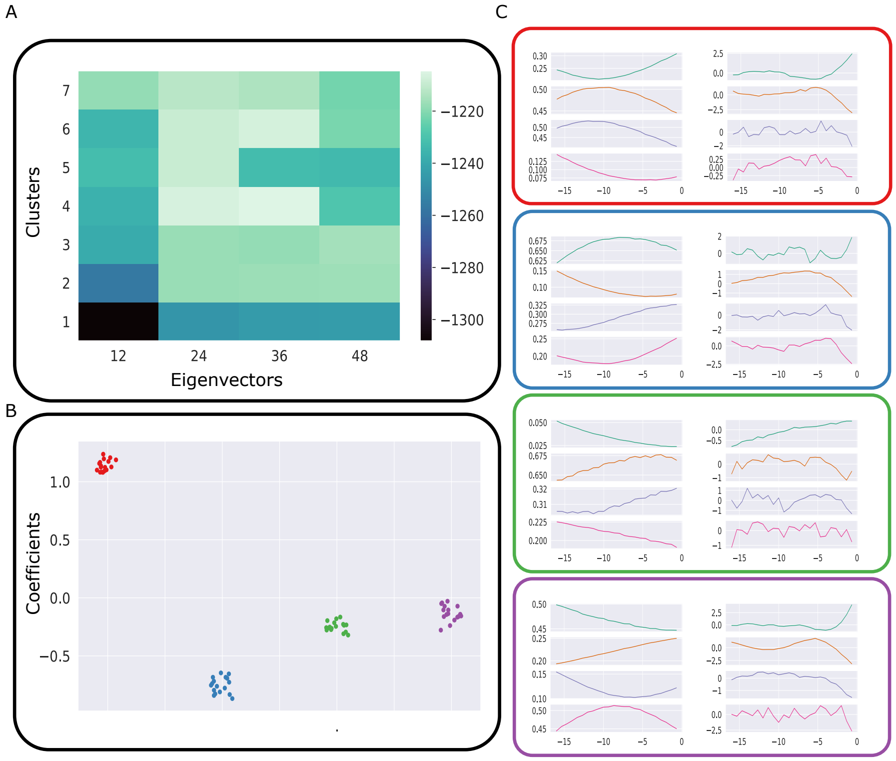

Figure 4.1: Logistic regression analysis on RME fall probability for OH15500 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. Performance is measured using log-likelihood on the cross-validation set. B. Regression coefficients for the four chosen clusters. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. Receptive field is defined as the linear combination of the eigenvectors and their regression coefficients. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

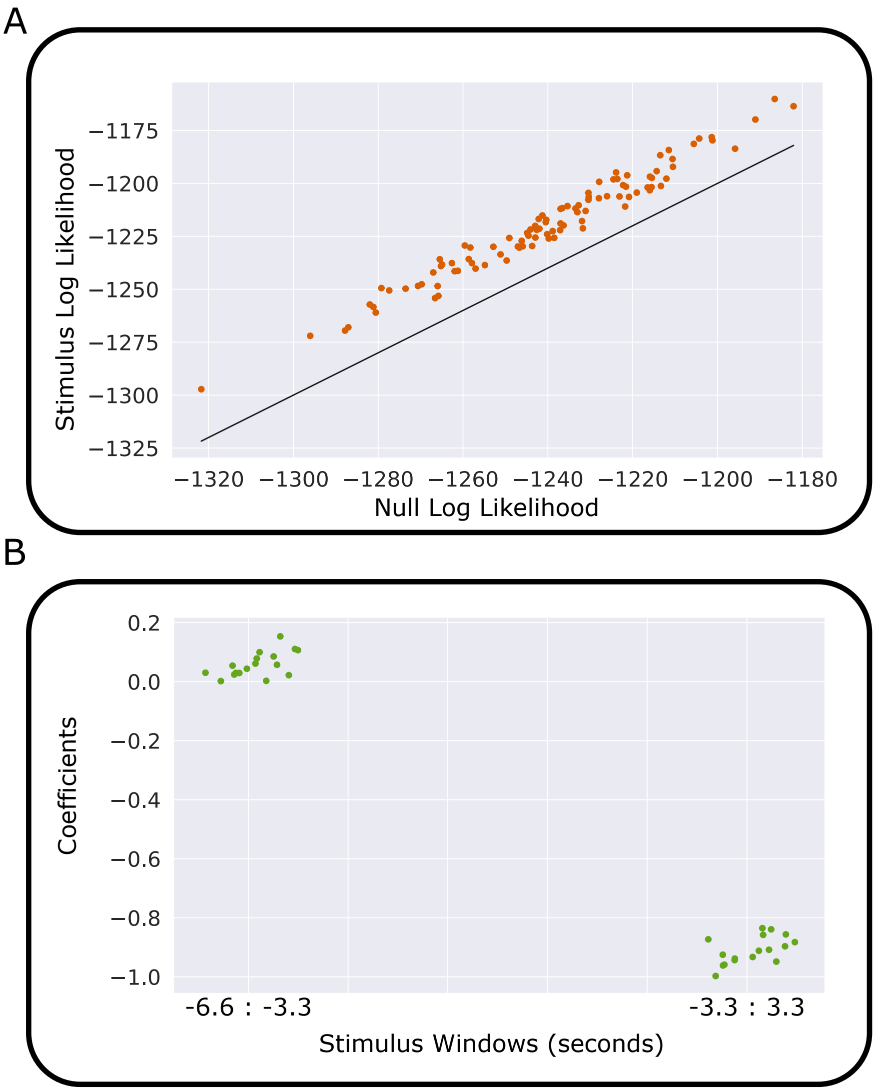

Figure 4.2: Stimulus effect on RME cell cluster fall probability for OH15500 animals. A. Model performance improvement due to stimulus feature inclusion. The null model is the best performing network history model found in the cross-validation analysis. Stimulus windows are added to this model to obtain the stimulus model. The stimulus model outperforms the null model on every cross-validation sample. B. The learned coefficients for the two stimulus windows. The -3.3 to 3.3 stimulus window refers to the stimulus average from the 3.3 seconds before a target timepoint to 3.3 seconds after a target timepoint. The distribution of coefficients across worms is shown.

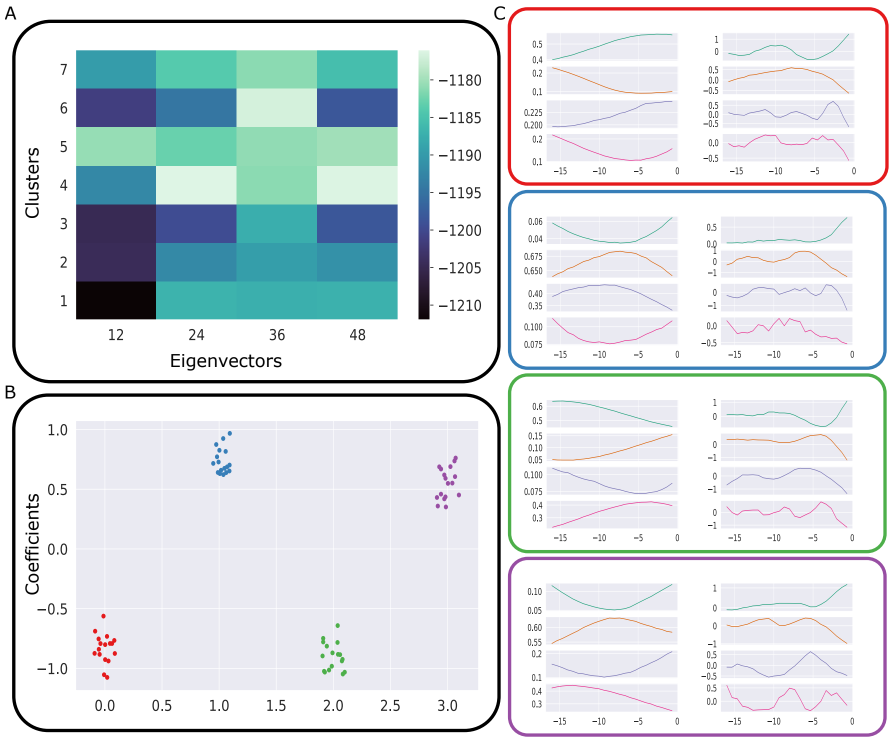

Figure 4.3: Logistic regression analysis on RME fall probability for ZIM294 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the four chosen clusters. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

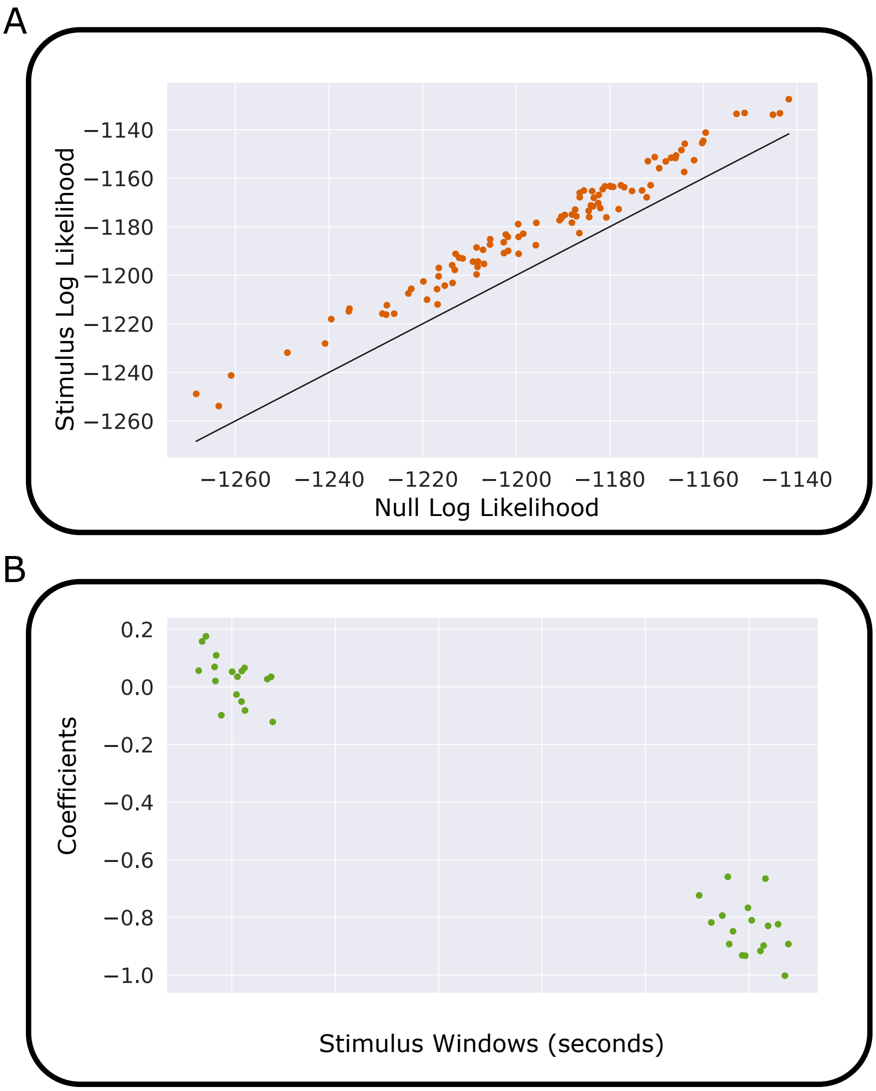

Figure 4.4: Stimulus effect on RME cell cluster fall probability for ZIM294 animals. A. Model performance improvement due to stimulus feature inclusion. The null model is the best performing network history model found in the cross-validation analysis. Stimulus windows are added to this model to obtain the stimulus model. The stimulus model outperforms the null model on every cross-validation sample. B. The learned coefficients for the two stimulus windows. The distribution of coefficients across worms is shown.

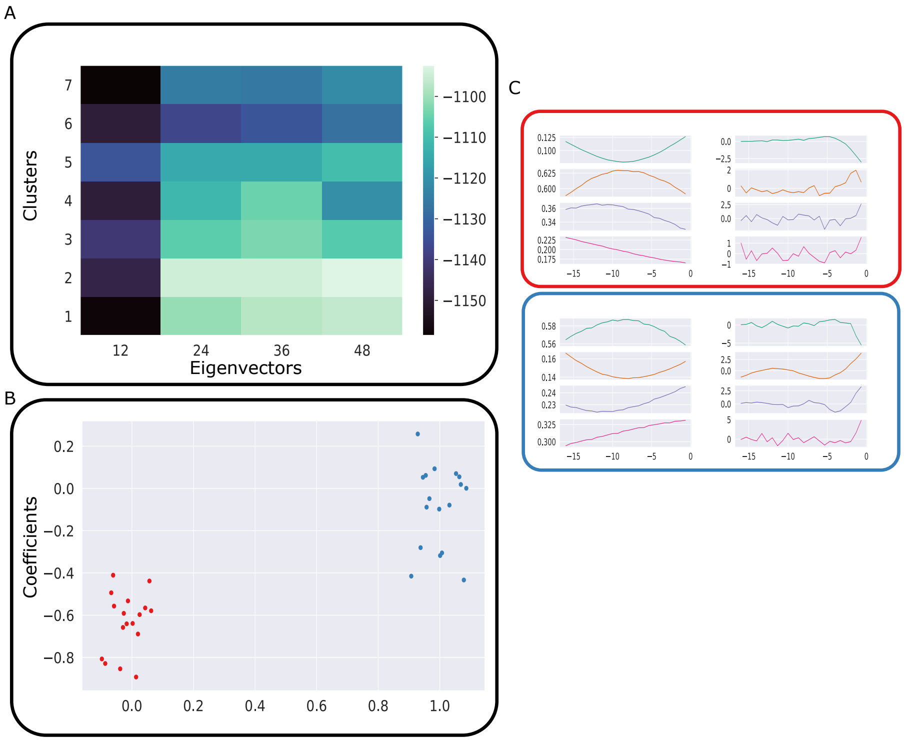

Figure 4.5: Logistic regression analysis on RME rise probability for OH15500 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the two chosen clusters. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

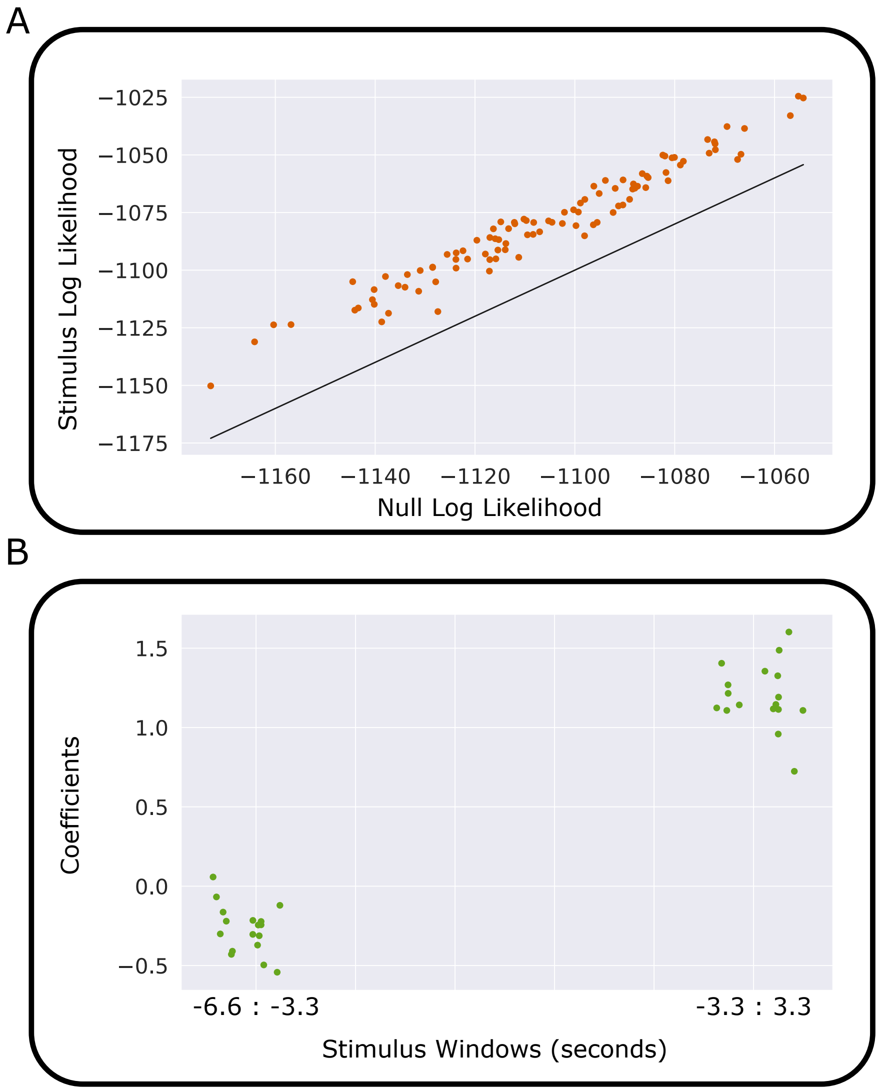

Figure 4.6: Stimulus effect on RME cell cluster rise probability for OH15500 animals. A. Model performance improvement due to stimulus feature inclusion. The null model is the best performing network history model found in the cross-validation analysis. Stimulus windows are added to this model to obtain the stimulus model. The stimulus model outperforms the null model on every cross-validation sample. B. The learned coefficients for the two stimulus windows. The distribution of coefficients across worms is shown.

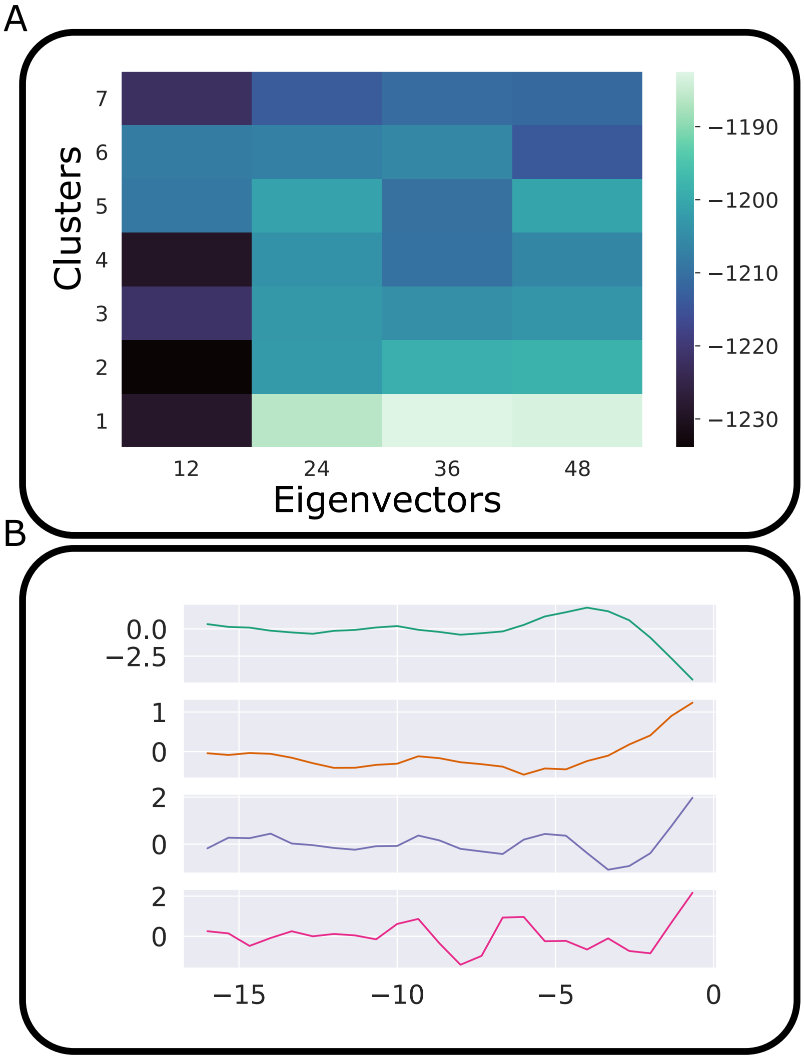

Figure 4.7: Logistic regression analysis on RME rise probability for ZIM294 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. Since the best performing model only used one network history cluster, it is well described by a single receptive field (B). The receptive field subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

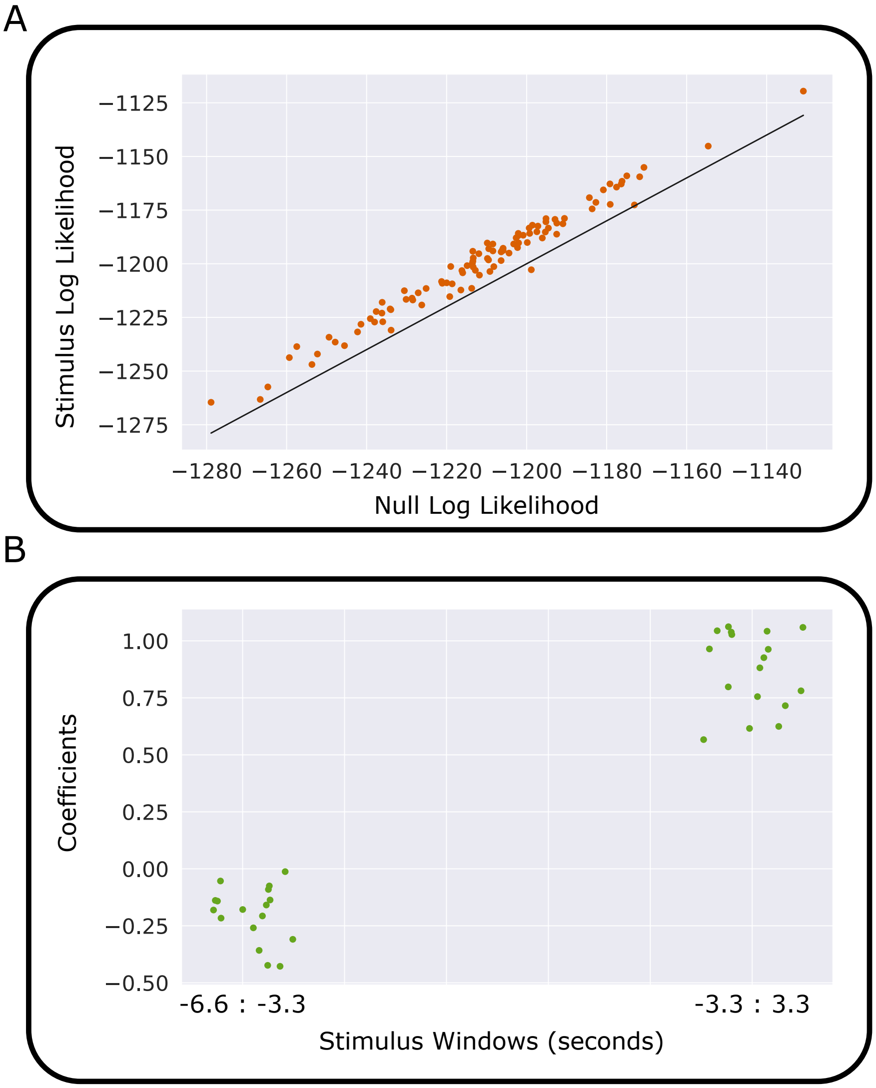

Figure 4.8: Stimulus effect on RME cell cluster rise probability for ZIM294 animals. A. Model performance improvement due to stimulus feature inclusion. The null model is the best performing network history model found in the cross-validation analysis. Stimulus windows are added to this model to obtain the stimulus model. The stimulus model outperforms the null model on every cross-validation sample except for one. B. The learned coefficients for the two stimulus windows. The distribution of coefficients across worms is shown.

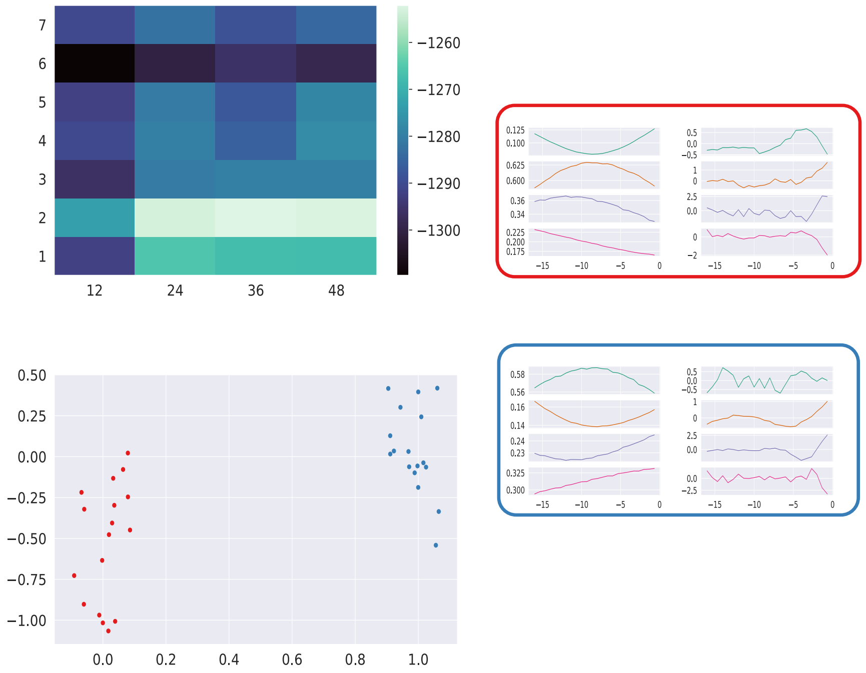

Figure 4.9: Logistic regression analysis on SMDV rise probability for OH15500 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the clusters of the best performing model. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

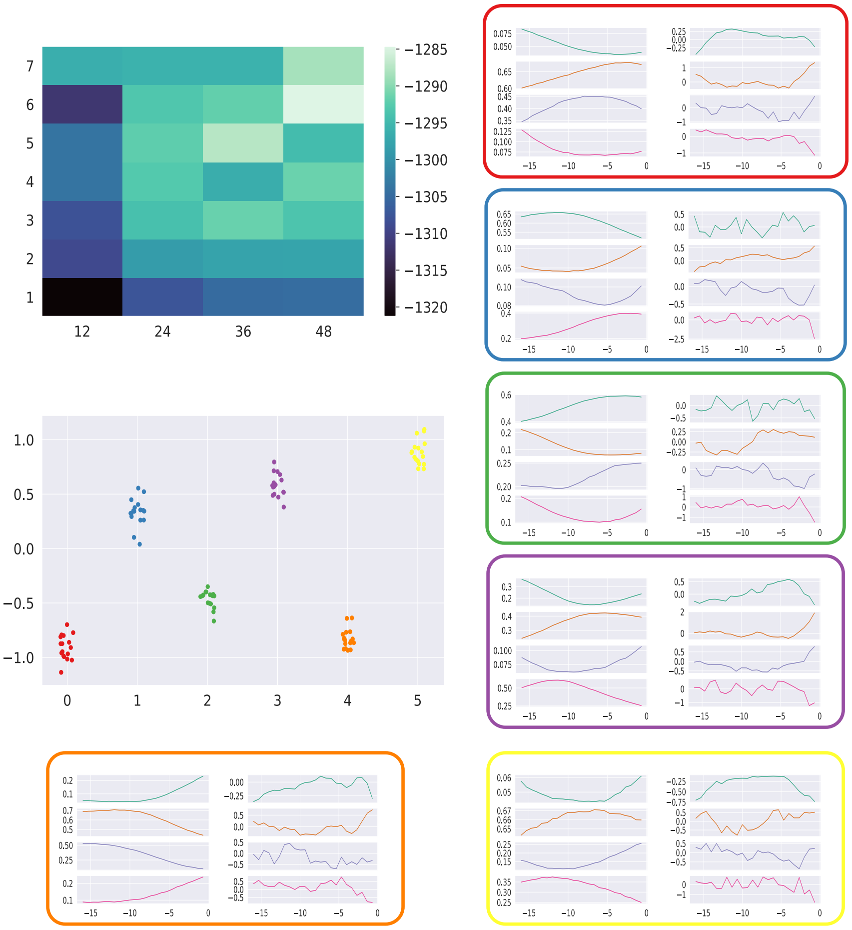

Figure 4.10: Logistic regression analysis on SMDV rise probability for ZIM294 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the clusters of the best performing model. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

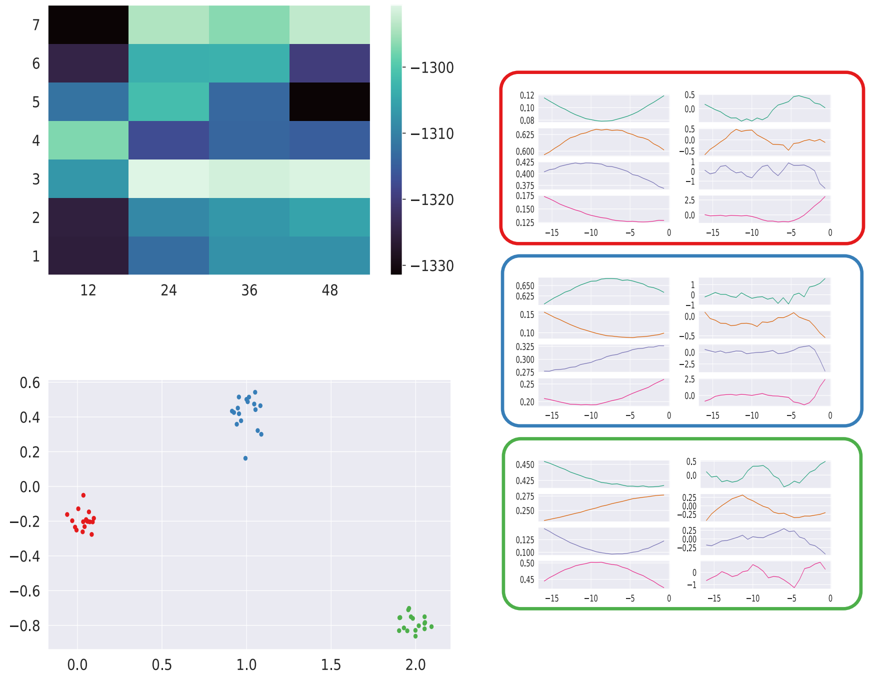

Figure 4.11: Logistic regression analysis on SMDD rise probability for OH15500 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the clusters of the best performing model. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

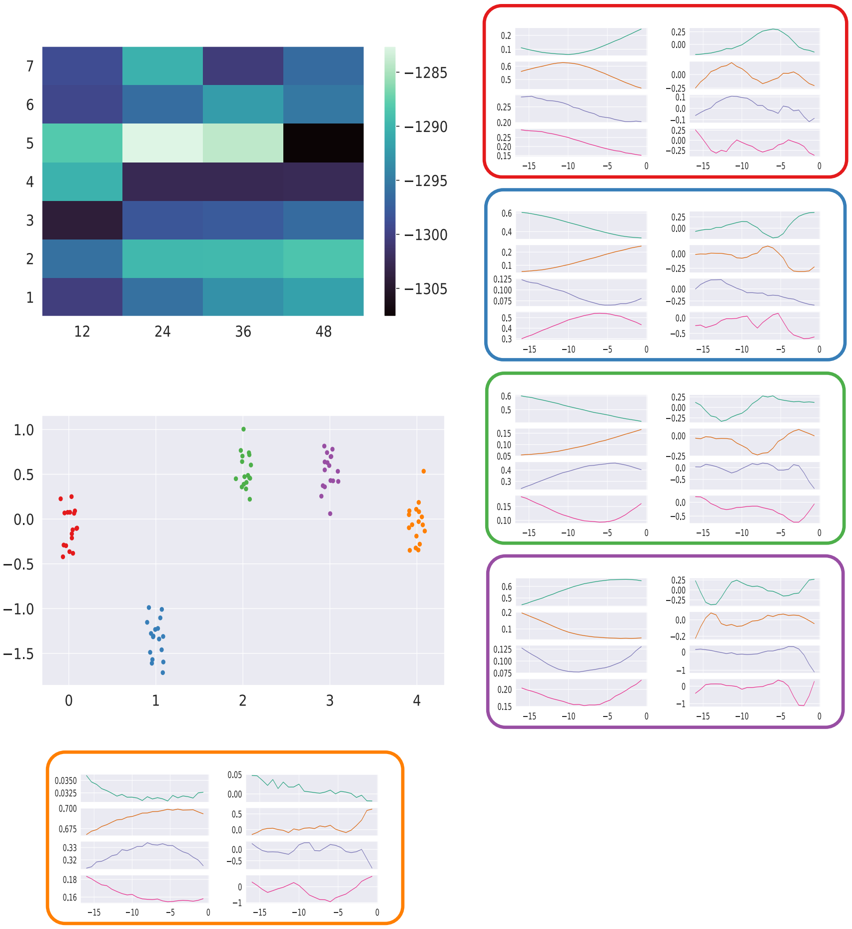

Figure 4.12: Logistic regression analysis on SMDD rise probability for ZIM294 animals. A. Cross-validation model performance as a function of number of clusters and eigenvector projection dimensionality. B. Regression coefficients for the clusters of the best performing model. The distribution of the coefficients across worms is shown. C. Visualization of the clusters and receptive fields. The box colors in C match the cluster colors in B. The left subplot in each box is the network history cluster. The right subplot in each box is the receptive field for that cluster. Each subplot consists of four more subplots. These correspond to the four cell clusters. In order: AVA, RME, SMDV, SMDD.

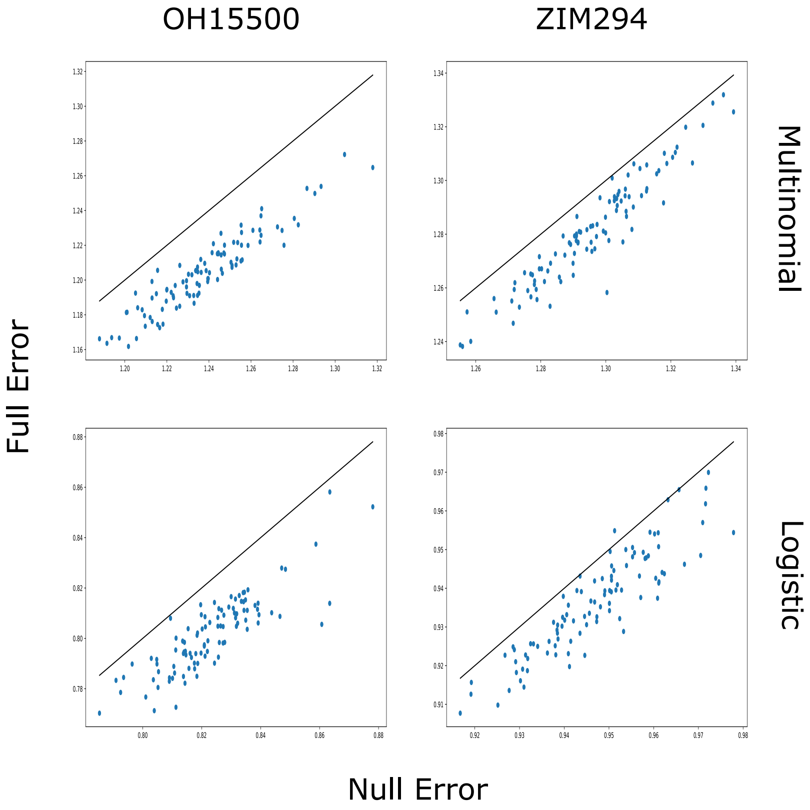

Figure 4.13: Inclusion of stimulus improves neural network performance on RME fluctuation prediction. A temporal convolutional neural network is used to predict RME fluctuations with and without knowledge of the stimulus. Full error is the negative log-likelihood cross-validation performance of neural networks with access to stimulus information. The null error is model performance without access to stimulus information. In complex prediction, RME fluctuations are binned into five categories. In simple prediction, they are binned into three categories.

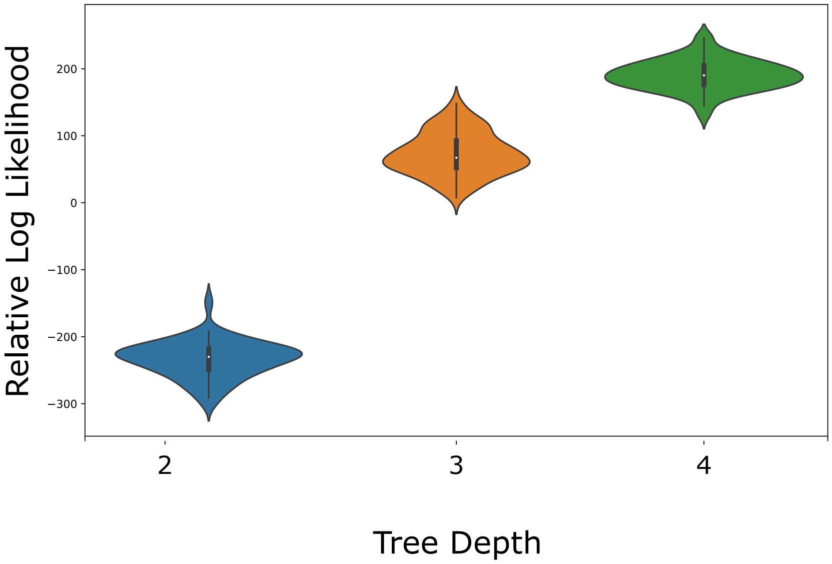

Figure 4.14: Deeper trees outperform a generalized linear model in RME fluctuation prediction. Trees of varying depths and a multinomial linear model are fit to various bootstraps of the 5-state RME fluctuation dataset. Cross-validation log-likelihood relative to the multinomial linear model is reported.

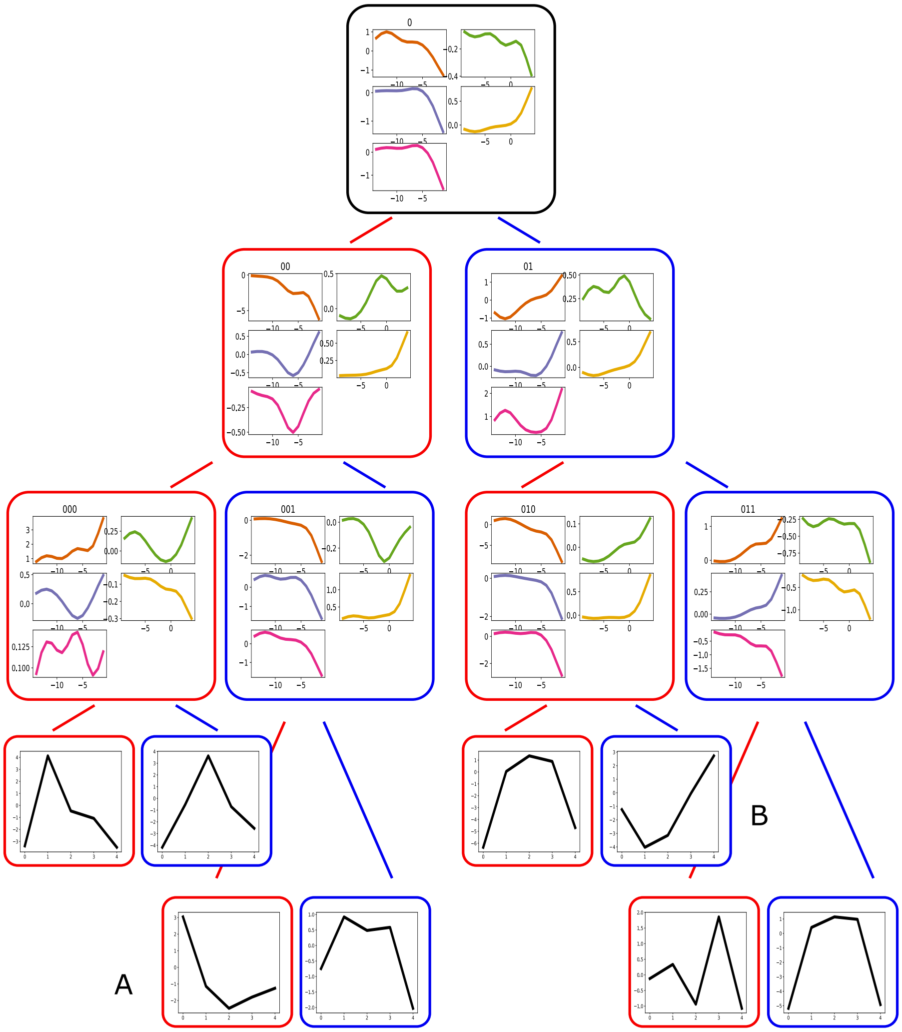

Figure 4.15: A tree of depth four is visualized. The tree has three levels of branches and a final level of leaves. Each of the branches has a receptive field in terms of RME, SMDV, and SMDD history (left three traces from top to bottom) and a receptive field in terms of On and Off sensory neurons (right two traces from top to bottom). The probability output by a branch is passed down to its left sub-tree (red). (1 - probability) output by a branch is passed down to its right sub-tree (blue). Each leaf has a distinct set of log-odds for RME fluctuation prediction. The first category (labeled 0) predicts the probability of strong RME decrease. The last category (labeled 4) predicts the probability of strong RME increase.

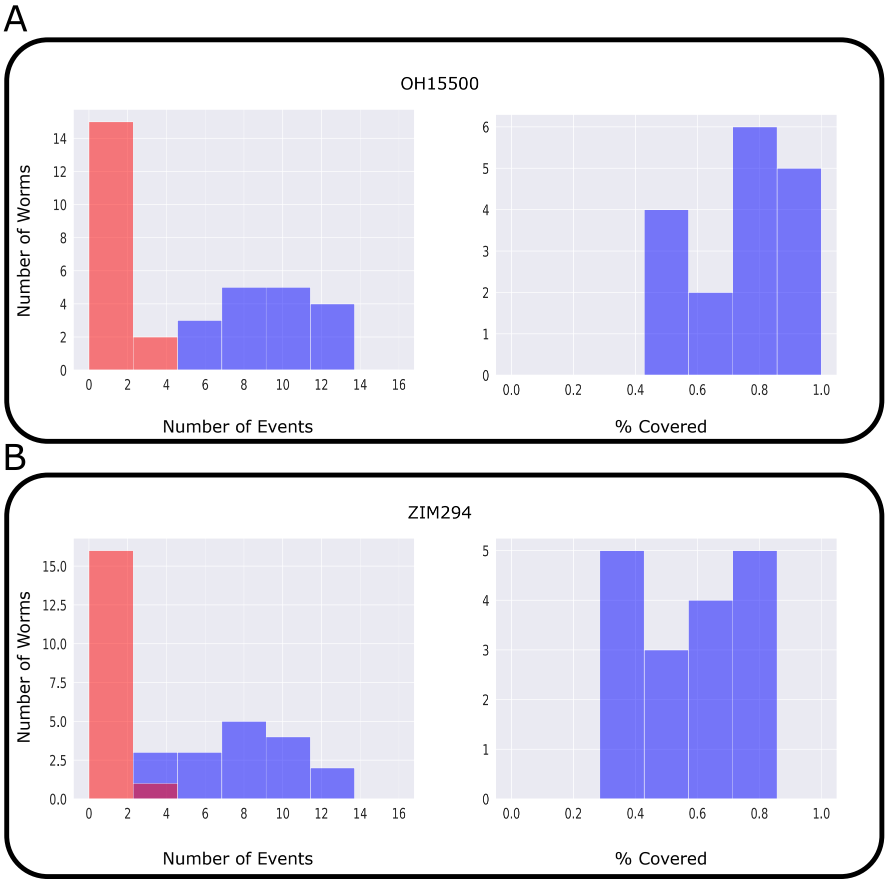

Figure 4.16: Relationship between RME and SMDV/D cell cluster events. The left subplots compare the number of RME cluster events that co-occur with SMDV/D events (blue) to the number of RME events that do not. The distributions of these two quantities across animals is shown for OH15500 (A) and ZIM294 (B). % coverage refers to the percentage of trial timepoints over which SMDV or SMDD is undergoing an event. The distribution of this quantity across animals is shown in the right subplots for (A) OH15500 animals and (B) ZIM294 animals.

Bibliography

[1] Microfluidics as a tool for c. elegans research. WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.162.1, http://www.wormbook.org.

[2] ADRIAN, E. The electrical activity of the mammalian olfactory bulb. *Electroencephalography and clinical neurophysiology* (1950), 377–88.

[3] AHRENS, M., AND ENGERT, F. Large-scale imaging in small brains. *Curr Opin Neurobiol* (2015), 78–86.

[4] AHRENS, M., LI, J., ORGER, M., ROBSON, D., SCHIER, A., ENGERT, F., AND PORTUGUES, R. Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* (2012), 471–7.

[5] BATES, D., MÄCHLER, M., BOLKER, B., AND WALKER, S. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* (2014).

[6] BREIMAN, L. Random forests. *Machine Learning* (2001), 5–32.

[7] BRENNER, N., BIALEK, W., AND VAN STEVENINCK, R. Adaptive rescaling maximizes information transmission. *Neuron* (2000), 695–702.

[8] BRENNER, S. The genetics of caenorhabditis elegans. *Genetics* (1974), 71–94.

[9] BUSCHMAN, T., AND MILLER, E. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* (2007), 1860–2.

[10] CALHOUN, A., TONG, A., POKALA, N., FITZPATRICK, J., SHARPEE, T., AND CHALASANI, S. Neural mechanisms for evaluating environmental variability in caenorhabditis elegans. *Neuron* (2015), 428–41.

[11] CHALASANI, S., CHRONIS, N., TSUNOZAKI, M., GRAY, J., RAMOT, D., GOODMAN, M., AND BARGMANN, C. Dissecting a circuit for olfactory behaviour in caenorhabditis elegans. *Nature* (2007), 63–70.

[12] CHEN, T., WARDILL, T., SUN, Y., PULVER, S., RENNINGER, S., BAOHAN, A., SCHREITER, E., KERR, R., ORGER, M., JAYARAMAN, V., AND LOOGER, L. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* (2013), 295–300.

[13] Chichilnisky, E. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems* (2001), 199–213.

[14] Dunn, T., Gebhardt, C., Naumann, E., Riegler, C., Ahrens, M., Engert, F., and Del Bene, F. Neural circuits underlying visually evoked escapes in larval zebrafish. *Neuron* (2016), 613–28.

[15] Dunn, T., Mu, Y., Narayan, S., Randlett, O., Naumann, E., Yang, C., Schier, A., Freeman, J., Engert, F., and Ahrens, M. Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion. *Elife* (2016), e12741.

[16] Escola, S., Fontanini, A., Katz, D., and Paninski, L. Hidden markov models for the stimulus-response relationships of multistate neural system. *Neural computation* (2011), 1071–132.

[17] Fairhall, A., Lewen, G., Bialek, W., and van Steveninck, R. Efficiency and ambiguity in an adaptive neural code. *Nature* (2001), 787–92.

[18] Fish, D., Brinicombe, A., Pike, E., and Walker, J. Blind deconvolution by means of the richardson–lucy algorithm. *JOSA A.* (1995), 58–65.

[19] Fisher, R. The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh* (1918), 399–433.

[20] Fontanini, A., and Katz, D. 7 to 12 hz activity in rat gustatory cortex reflects disengagement from a fluid self-administration task. *Journal of neurophysiology* (2005), 2832–40.

[21] Frosst, N., and Hinton, G. Distilling a neural network into a soft decision tree. *arXiv preprint arXiv:1711.09784* (2017).

[22] Gray, J., Hill, J., and CI, B. A circuit for navigation in caenorhabditis elegans. *Proceedings of the National Academy of Sciences* (2005), 3184–91.

[23] Gutierrez, G., O'Leary, T., and Marder, E. Multiple mechanisms switch an electrically coupled, synaptically inhibited neuron between competing rhythmic oscillators. *Neuron* (2013), 845–858.

[24] Helmchen, F., and Denk, W. Deep tissue two-photon microscopy. *Nature Methods* (2005), 932–940.

[25] Ho, T. K. Random decision forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition* (1995), 278–282.

[26] Kandepu, R., Foss, B., and Imsland, L. Applying the unscented kalman filter for nonlinear state estimation. *Journal of process control* (2008), 753–68.

[27] KAPLAN, H., THULA, O., KHOSS, N., AND ZIMMER, M. Nested neuronal dynamics orchestrate a behavioral hierarchy across timescales. *Neuron* (2020), 562–76.

[28] KATO, S., KAPLAN, H., SCHRÖDEL, T., SKORA, S., LINDSAY, T., YEMINI, E., LOCKERY, S., AND ZIMMER, M. Global brain dynamics embed the motor command sequence of caenorhabditis elegans. *Cell* (2015), 656–69.

[29] KATO, S., XU, Y., CHO, C., ABBOTT, L., AND BARGMANN, C. Temporal responses of c. elegans chemosensory neurons are preserved in behavioral dynamics. *Neuron* (2014), 616–28.

[30] KL, B. Optical imaging of neuronal populations during decision-making. *Science* (2005), 896–901.

[31] LEINWAND, S., AND CHALASANI, S. Neuropeptide signaling remodels chemosensory circuit composition in caenorhabditis elegans. *Nature neuroscience* (2013), 1461–7.

[32] LINDERMAN, S., JOHNSON, M., MILLER, A., ADAMS, R., BLEI, D., AND PANINSKI, L. Bayesian learning and inference in recurrent switching linear dynamical systems. *Artificial Intelligence and Statistics* (2017), 914–922.

[33] LINDERMAN, S., NICHOLS, A., BLEI, D., ZIMMER, M., AND PANINSKI, L. Hierarchical recurrent state space models reveal discrete and continuous dynamics of neural activity in c. elegans. *bioRxiv* (2019), 621540.

[34] MUKAMEL, E., NIMMERJAHN, A., AND SCHNITZER, M. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* (2012), 747–760.

[35] MURTHY, S., KASIF, S., AND SALZBERG, S. A system for induction of oblique decision trees. *Journal of artificial intelligence research* (1994), 1–32.

[36] NGUYEN, J., SHIPLEY, F., LINDER, A., PLUMMER, G., LIU, M., SETRU, S., SHAEVITZ, J., AND LEIFER, A. Whole-brain calcium imaging with cellular resolution in freely behaving caenorhabditis elegans. *Proceedings of the National Academy of Sciences* (2016), E1074–81.

[37] OORD, A., DIELEMAN, S., ZEN, H., SIMONYAN, K., VINYALS, O., GRAVES, A., KALCHBRENNER, N., SENIOR, A., AND KAVUKCUOGLU, K. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499* (2016).

[38] PANINSKI, L., PILLOW, J., AND SIMONCELLI, E. Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural computation* (2004), 2533–61.

[39] PILLOW, J., PANINSKI, L., UZZELL, V., SIMONCELLI, E., AND CHICHILNISKY, E. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience* (2005), 11003–13.

[40] PILLOW, J., SHLENS, J., PANINSKI, L., SHER, A., LITKE, A., CHICHILNISKY, E., AND SIMONCELLI, E. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* (2008), 995–9.

[41] PNEVMATIKAKIS, E., AND GIOVANNUCCI, A. Normcorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods* (2017), 83–94.

[42] PNEVMATIKAKIS, E., SOUDRY, D., GAO, Y., MACHADO, T., MEREL, J., PFAU, D., REARDON, T., MU, Y., LACEFIELD, C., YANG, W., AND AHRENS, M. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* (2016), 285–99.

[43] PREVEDEL, R., YOON, Y., HOFFMANN, M., PAK, N., WETZSTEIN, G., KATO, S., SCHRÖDEL, T., RASKAR, R., ZIMMER, M., AND BOYDEN, E. E. A. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature Methods* (2014).

[44] REYNOLDS, J., AND CHELAZZI, L. Attentional modulation of visual processing. *Annu. Rev. Neuroscic* (2004), 611–47.

[45] REYNOLDS, J., PASTERNAK, T., AND DESIMONE, R. Attention increases sensitivity of v4 neurons. *Neuron* (2000), 703–14.

[46] RIEKE, F., WARLAND, D., VAN STEVENINCK, R., AND BIALEK, W. *Spikes: exploring the neural code*. Cambride: MIT press, 1999.

[47] RINBERG, D., KOULAKOV, A., AND GELPERIN, A. Sparse odor coding in awake behaving mice. *Journal of Neuroscience* (2006), 8857–65.

[48] ROSE, J., AND RANKIN, C. Analyses of habituation in caenorhabditis elegans. *Learning and Memory* (2001), 63–9.

[49] SHARPEE, T., RUST, N., AND BIALEK, W. Analyzing neural responses to natural signals: maximally informative dimensions. *Neural computation* (2004), 223–50.

[50] SHARPEE, T., SUGIHARA, H., KURGANSKY, A., REBRIK, S., STRYKER, M., AND KD, M. Adaptive filtering enhances information transmission in visual cortex. *Nature* (2006), 936–42.

[51] SULSTON, J., SCHIERENBERG, E., WHITE, J., AND THOMSON, J. The embryonic cell lineage of the nematode caenorhabditis elegans. *Developmental biology* (1983), 64–119.

[52] TSAMARDINOS, I., GREASIDOU, E., AND BORBOUDAKIS, G. Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. *Machine Learning* (2018), 1895–922.

[53] VANWALLEGHEM, G., AHRENS, M., AND SCOTT, E. Integrative whole-brain neuroscience in larval zebrafish. *Current opinion in neurobiology* (2018), 134–45.

[54] WAN, E., AND VAN DER MERWE, R. The unscented kalman filter for nonlinear estimation. *InProceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium* (2000), 153–158.

[55] WARK, B., LUNDSTROM, B., AND FAIRHALL, A. Sensory adaptation. *Current opinion in neurobiology* (2007), 423–9.

[56] WOMELSDORF, T., ANTON-ERXLEBEN, K., PIEPER, F., AND TREUE, S. Dynamic shifts of visual receptive fields in cortical area mt by spatial attention. *Nature neuroscience* (2006), 1156–60.

[57] YEMINI, E., LIN, A., NEJATBAKHSH, A., VAROL, E., SUN, R., MENA, G., SAMUEL, A., PANINSKI, L., VENKATACHALAM, V., AND HOBERT, O. *biorxiv* (2020).