RICE UNIVERSITY

The Neural Computations in Spatial Memory From Single Cells to Networks

by

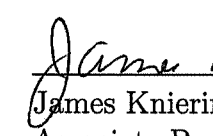
Kathryn Ruth Hedrick

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

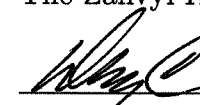
Doctor of Philosophy

Approved, Thesis Committee:

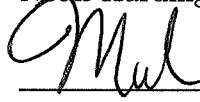
Steven Cox, Chairman
Professor of Computational and Applied Mathematics



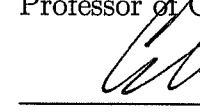
James Knierim
Associate Professor of Neuroscience
The Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University



Danny Sorensen
Noah Harding Professor of Computational and Applied Mathematics



Mark Embree
Professor of Computational and Applied Mathematics



Caleb Kemere
Assistant Professor of Electrical and Computer Engineering

Houston, Texas

April, 2012

Abstract

The Neural Computations in Spatial Memory From Single Cells to Networks

by

Kathryn Ruth Hedrick

Studies of spatial memory provide valuable insight into more general mnemonic functions, for by observing the activity of cells such as place cells, one can follow a subject's dynamic representation of a changing environment. I investigate how place cells resolve conflicting neuronal input signals by developing computational models that integrate synaptic inputs on two scales. First, I construct reduced models of morphologically accurate neurons that preserve neuronal structure and the spatial specificity of inputs. Second, I use a parallel implementation to examine the dynamics among a network of interconnected place cells. Both models elucidate possible roles for the inputs and mechanisms involved in spatial memory.

Place cells combine spatial information carried by afferents from the medial entorhinal cortex (MEC) with contextual information carried by afferents from the lateral entorhinal cortex (LEC). Studies indicate that the structured distribution of

synapses on the dendrites affects how place cells integrate input from MEC, LEC, and neighboring place cells. However, models mapping spatially distributed inputs to the somatic potential are computationally expensive, and existing reduced models sacrifice either the spatial component of the input signal or the electrophysiology of the original model. I use Krylov subspace projection methods to construct reduced models of passive and quasi-active neurons that accurately reproduce the somatic potential, preserve the spatial specificity of inputs and the underlying circuit structure, and are constructed from straightforward transformations dependent on intrinsic dendritic properties. The accuracy, efficiency, and electrophysiological interpretation of the reduced models render them useful in examining dendritic computations.

I further develop a computational platform to study the dynamics among a network of isopotential place cells as a learned environment is modified, placing LEC and MEC inputs in conflict. The experimental paradigm on which the model is based allows me to directly test the respective roles of each input group as well as the mechanisms affecting how place cells resolve the competition in their input signals. Using both the reduced and network models, I demonstrate that accepted mechanisms, such as synaptic plasticity and cell assembly dynamics, can combine in non-intuitive ways to govern the place cell representation of a modified environment.

Acknowledgements

I am very grateful for the guidance and support that has enabled me to complete this thesis. I would first like to thank my adviser, Steve Cox, for his advice, guidance, and encouragement over the past five years. He has stepped beyond the role of adviser to become my mentor, role model, and friend. I would also like to thank my co-adviser, James Knierim, for the tireless hours he has invested to help me design the model and understand its implications. He has greatly contributed to my understanding of the neuroscience underlying the network model, and I look forward to our continued work together in the coming years. I am further grateful to Mark Embree, Danny Sorensen, and Caleb Kemere for their advice and for the expertise in model reduction and spatial memory they bring to the committee.

Many colleagues have contributed directly to this work. I thank Horatiu Voitu and Katherine Johnston for their contributions to the initial stages, and I thank Andrew Wu for working with me to examine the effect of several plasticity rules. I also thank Francesco Savelli and Sachin Deshmuk for their advice and insights in connecting the experimental and simulation results.

The support of my family and friends throughout this process has been invaluable. I would especially like to thank my husband for his support and understanding. He has brought me home-cooked meals on countless long nights in the office, and he contributed directly to this thesis by creating the circuit diagrams. I also give a

special thanks to my parents, who have encouraged me through all my endeavors, and I thank my Father for the talents He has given me and the strength and peace He provides.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Space surrounds us, imposing itself on our behaviors, perceptions, and memories. At times, we directly interact with space, planning routes or gaging the trajectory of a ball we hope to catch. At other times, we subconsciously perceive where we are relative to some fixed framework and where objects are relative to ourselves. The interaction with space extends beyond humans, as a variety of species exhibit highly developed spatial behaviors. For example, many species of birds migrate hundreds to thousands of miles each year to reach their breeding grounds, and the homing pigeon can find its way back to its nest after being displaced by hundreds of miles to an unfamiliar location. Rats have proven themselves to be highly efficient at navigating through mazes, and even fish demonstrate spatial abilities such as learning the distribution of feeding sites (Healy [31]). These are but a few examples of the complex interactions between organisms and space.

The intriguing questions of what space is and how we perceive space have been studied by philosophers, physicists, and epistemologists for centuries, most notably among them Aristotle, Newton, and Kant. In particular, Kant postulated that the concept of space is innately present in the brain *a priori* to experience. This idea was adopted by O'Keefe and Nadel in the 1970's, who argued that many species have the innate ability to access a fixed spatial framework, forming a cognitive map of their external surroundings. This cognitive map can then be used to perform complex spatial tasks, such as self-localization and navigation. Taking advantage of the developments in physiology and neuroscience, O'Keefe and Nadel proceeded to experimentally uncover the neural correlates of this cognitive map, directing us to the brain regions capable of forming a spatial representation and examining the activity of neurons within these regions (O'Keefe and Nadel [65]).

As the understanding of the brain circuitry and functional properties of neurons grows, so does the need for efficient, accurate computational models that can be used to test hypotheses, generate new theories, and design experiments. Indeed, it is largely the combination of experimental and theoretical work that drives discovery forward.

It is from this perspective that I approach spatial memory on two complementary scales. Driven by experimental findings, I model both morphologically accurate neurons and large networks of isopotential neurons with plastic connections. I then use these models to examine the stability of the cognitive map under the graded

environmental perturbations of the experiment. To stay as true to the physiology as possible with minimal computational cost, I use model reduction to investigate the computations performed by individual neurons and a parallel implementation to investigate network dynamics.

1.1 The Hippocampus and Spatial Memory

Numerous experiments and case studies have shown that the hippocampus is an essential region for memory, and the implications of a selection of these studies is presented in §1.1.1. Aware of these studies, O'Keefe and Nadel point to the hippocampus as a region well-equipped to act as a cognitive map, both structurally and functionally. A basic component of this cognitive map is what they term a mapping space consisting of interconnected places receptive to inputs (O'Keefe and Nadel [65]). The functional aspects of this mapping space are provided by pyramidal neurons called place cells, which are described in §1.1.2. Structurally, place cells receive sensory input regarding the environment and are members of a well-studied circuit that processes these inputs, as described in §1.1.3.

1.1.1 Memory Tasks of the Hippocampus

The link between the hippocampus and memory has been firmly established largely due to case studies in which the hippocampus is damaged. The most famous of these studies is a patient known as HM. Having suffered from intractable epilepsy from his

youth, parts of HM's medial temporal lobe, including two-thirds of his hippocampus, were surgically removed. While the surgery stopped the seizures, it also destroyed the functionality of his hippocampus. As a result, HM developed anterograde amnesia. That is, he could remember detailed events from his childhood but was unable to form new long-term memories (Scoville and Milner [80]). The fascinating nature of HM's subsequent impairments were well-studied and documented over the next fifty years, providing a solid foundation for the study of memory tasks performed by the hippocampus.

The anterograde amnesia exemplified by HM is a characteristic impairment caused by bilateral hippocampal damage. As remote memories stored prior to the brain damage are largely unaffected, the hippocampus is not the storage unit for long-term memory. However, the hippocampus is vital for the storage of new declarative memories, or memories of facts (semantic memory) and events (episodic memory) (Squire et al. [86], Rolls and Treves [74]).

Spatial memory is a special subclass of declarative memory, and studies have established the necessity of the hippocampus for spatial memory. For example, rats with hippocampal lesions lose the ability to recognize places using spatial cues (O'Keefe and Nadel [65]), and similar spatial impairments are found in macaques and humans (Rolls and Treves [74]). The anterograde nature of spatial memory loss is retained in humans. In one classic case, a man with damage to his hippocampus and neighboring cortical regions could navigate and construct novel routes within the environment he

grew up in. In his current environment, however, he could not so much as point to the direction of the Pacific Ocean, despite living a mere two miles away (Teng and Squire [87]).

1.1.2 Place Cells and Grid Cells

The hippocampus has been established as a key region for declarative memory tasks, including spatial memory, since the 1950s. However, it was not until 1971 that O'Keefe and Dostrovsky uncovered the functional aspects of the hippocampus giving it the capacity to represent space. They discovered that certain cells in the hippocampus have firing patterns strongly correlated to behavior. Most notably, place cells fire almost exclusively when the subject is in one particular place within an environment (O'Keefe and Dostrovsky [64]). Since their discovery, place cells have provided valuable insight into the mnemonic functions of the hippocampus, for by recording the firing patterns of a population of place cells, one can follow a subject's dynamic spatial representation of its environment (Leutgeb et al., [50]).

Returning to the theory of O'Keefe and Nadel, a mapping space must contain units representing place receptive to inputs. Place cells do indeed receive spatial information, primarily from grid cells in the medial entorhinal cortex. Grid cells were discovered in 2005 and have two-dimensional tesselating firing patterns (Hafting [27]). Figure 1.1 shows the characteristic firing patterns of place cells and grid cells.

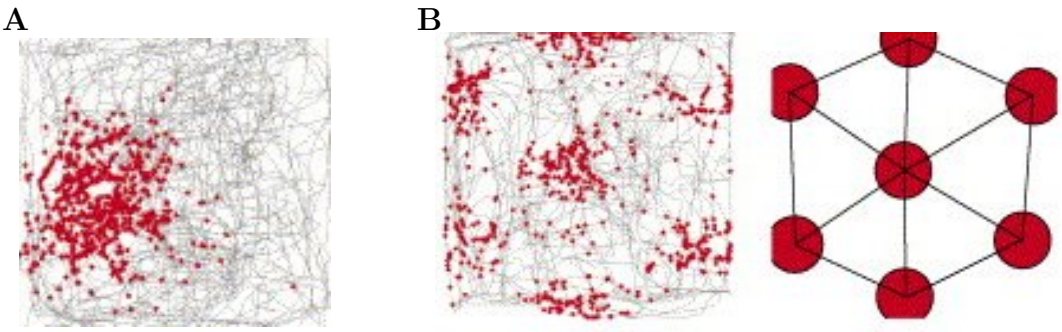


Figure 1.1: Characteristic firing fields of a place cell **(A)** and grid cell **(B)** as a rat explores a rectangular enclosure. **A.** Each red dot represents the rat's location when the place cell fired. Place cells form at most one firing field within a given environment. **B.** Grid cells form firing fields arranged in tesselating patterns. Images from (Witter and Moser [96]).

1.1.3 Basic Circuitry

A network of place cells and grid cells provides a cognitive map of the environment through the cells' spatially tuned firing patterns. These cells are far from isolated, and understanding the dynamics of this map requires an understanding of the underlying circuitry. Grid cells are found in the medial entorhinal cortex (MEC), and place cells are found in the CA1 and CA3 subregions of the hippocampus. These regions are shown in Figure 1.2.

The connections among subregions in the entorhinal cortex and hippocampus form a well-studied loop, as shown in Figure 1.3. I briefly discuss the functional impact of some of these connections below.

Perforant Path: The perforant path refers to the connections from the entorhinal cortex to the hippocampus. The entorhinal cortex is subdivided into the medial entorhinal cortex (MEC) and lateral entorhinal cortex (LEC), and it carries two distinct types of information. As previously described, the MEC provides spatial

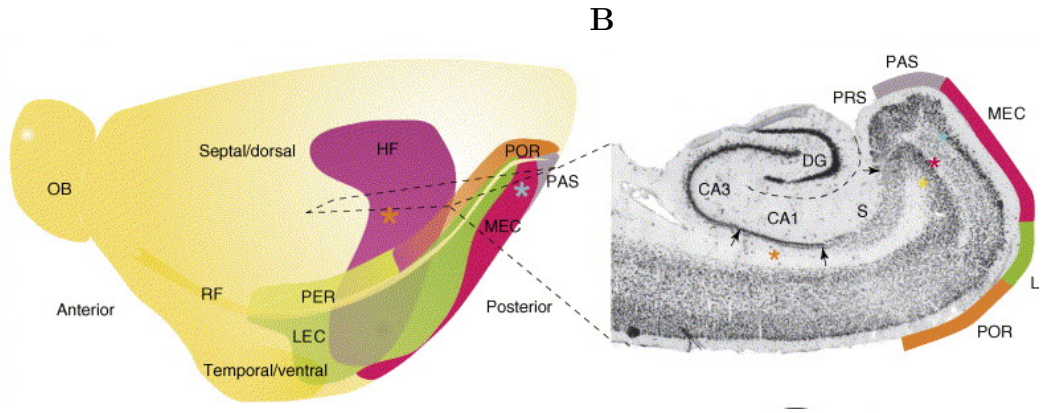


Figure 1.2: Schematic of the hippocampus and entorhinal cortex. **A.** Posterolateral view of the rat brain. The entorhinal cortex drives networks in the hippocampal formation (HF). It is subdivided into the medial entorhinal cortex (MEC), which contains grid cells, and the lateral entorhinal cortex (LEC), whose cells have very weak spatial tuning. **B.** Horizontal slice of **(A)**. Place cells are contained in the CA1 and CA3 subregions of the hippocampus. Image from (Witter and Moser [96]).

information through the activity of cells such as grid cells, and cells in the LEC are tuned to local objects (Hargreaves et al. [28]). Thus, the hippocampus receives information regarding both *where* and *what*.

Mossy Fibers: The CA3 is the only output of the dentate gyrus (DG), and the connections from DG to CA3 are called mossy fibers. While DG cells fire sporadically, the connections from the DG to the CA3 are among the strongest in the brain, giving DG cells the ability to strongly influence CA3 activity (Shepherd [40]).

Associational Connections: One distinct difference between the CA1 and CA3 are the associational connections among CA3 neurons (Shepherd [40]). These connections are strong enough that a burst of spikes from a single CA3 cell can evoke a burst of spikes in another CA3 cell (Traub et al. [88]). This intercommunication enables CA3 cells to form associations between places.

Schaffer Collaterals: The Schaffer collaterals refer to the feedforward connections from the CA3 to the CA1. CA1 cells have a unique perspective as they compare direct input from the entorhinal cortex via the perforant path to input that has been processed by both the DG and CA3 subregions via the Schaffer collaterals. CA1 cells then project directly back to the entorhinal cortex, predominantly to the MEC (Shepherd [40]).

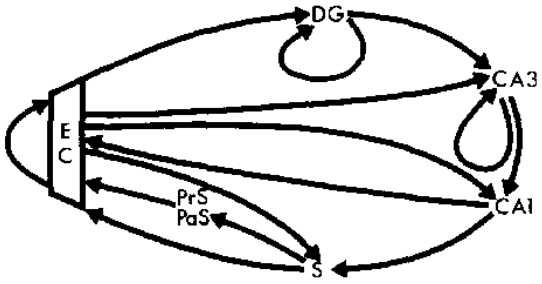


Figure 1.3: Circuitry of entorhinal cortical and hippocampal subregions. Both subregions of the entorhinal cortex, the MEC and LEC, project to the hippocampus, which includes the dentate gyrus (DG), CA1, and CA3 subregions. This input is further processed in the DG and CA3 subregions before converging on the CA1, which projects back to the entorhinal cortex, closing the loop. Other regions shown are the presubiculum (PrS), parasubiculum (PaS), and subiculum (S). These regions contain neurons providing directional information but are not included in this work. Image from (Amaral and Witter [1]).

1.2 Synaptic Plasticity

By studying the firing patterns within a network of grid cells, one can uniquely determine a subject's position within an environment. Similarly, any given place can be uniquely represented by the firing patterns among place cells. The brain thus contains two spatial maps, leading to the question of how these maps relate to one

another. Do changes in the grid cell representation affect changes in the place cell representation, and vice versa? On a more local level, how do the interconnected CA3 place cells relate to one another? If a subject were to repeatedly move from Place A to Place B, how does the place cell network reflect an association between Places A and B?

Theories addressing these questions date back to 1949, when Donald Hebb theorized that assemblies form in the brain through the change in the strength of connections among neurons, a process called synaptic plasticity. Hebb asserted that if two neurons are repeatedly active at the same time, the connection between them strengthens (Hebb [32]). Hebbian plasticity has been coined with the simplistic phrase, *neurons that fire together wire together*.

Experiments such as the classic Morris water maze (Morris [62]) have demonstrated that the ability to learn the spatial position of objects is lost when synaptic plasticity is prevented, and thus synaptic plasticity is a key component for most models of spatial memory. I implemented two types of Hebbian plasticity in the models: rate-based plasticity (§1.2.1) and spike-timing-dependent plasticity (§1.2.2).

1.2.1 Rate-Based Plasticity

Rate-based plasticity models specify that the strength, or weight, of the connection from a presynaptic cell to a postsynaptic cell depends on the respective spike rates. This model is supported by experimental data, such as that shown in Figure 1.4. In

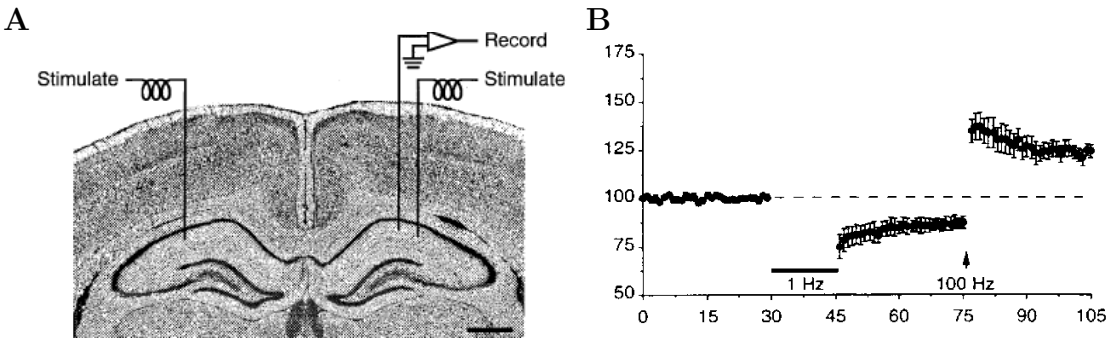


Figure 1.4: Example of experimental results supporting rate-based plasticity. **A.** CA1 cells in the hippocampus are stimulated *in vivo*. **B.** When the cells spike at a rate of 1 Hz, the weights decrease. When the cells spike at a rate of 100 Hz, the weights increase, recovering to a greater value than the baseline weight. The plot shows the percentage of the baseline weight vs. time (min). Images from (Heynen et al. [35]).

this experiment Heynen et al. stimulated cells in the hippocampus to spike at a given rate and measured the average response of postsynaptic cells due to each presynaptic spike. When they stimulated cells to spike at a rate of 1 Hz, the postsynaptic responses decreased, indicating that long-term-depression (LTD), or a weight decrease, had occurred. On the other hand, long-term potentiation (LTP), or a weight increase, occurred when when they stimulated cells to spike at a rate of 100 Hz.

For this thesis postsynaptic-gated rate-based plasticity governs the weights from MEC and LEC cells to CA3 excitatory cells. The plasticity rule has the form

$$\frac{dw}{dt}(t) = k(R_{pre}(t) - \bar{R})R_{post}(t), \tag{1.1}$$

where w is the weight of the given synapse, R_{pre} and R_{post} are the instantaneous spike rates of the presynaptic and postsynaptic cell, respectively, $\bar{R}$ is the threshold rate, and k is a parameter that specifies the learning rate (Gerstner and Kistler [21]). Accordingly, the weight changes only if the postsynaptic cell has a nonzero spike rate,

and the presynaptic spike rate determines the direction of weight change.

1.2.2 Spike-Timing-Dependent Plasticity

Spike-timing-dependent plasticity (STDP) is an asymmetric learning rule in which the weight change depends on the relative timing of individual pre- and postsynaptic spike pairs. The underlying principle of STDP is that LTP (weight increase) occurs when the presynaptic cell spikes before the postsynaptic cell, and LTD (weight decrease) occurs when the postsynaptic cell spikes before the presynaptic cell. The magnitude of the weight change is determined by the exact time difference between the two spikes. STDP is thus a rule for causality, where if a spike from cell A helps to drive cell B to spike, then the weight from A to B increases while the weight from B to A decreases.

Many experiments have shown the existence of STDP (e.g. Markrum et al. [56], Bi and Poo [6]), and theorists soon followed to propose models of the STDP learning rule. The classic STDP model dictates that the weight change due to each spike pair is given by

$$\Delta w(t_{post} - t_{pre}) = \begin{cases} -A_- \exp((t_{post} - t_{pre})/\tau_-), & \text{if } t_{post} \leq t_{pre} \\ A_+ \exp((t_{pre} - t_{post})/\tau_+), & \text{if } t_{post} > t_{pre}. \end{cases} \tag{1.2}$$

Figure 1.5 shows the STDP curve from both experimental data and STDP models.

The primary point of controversy for modeling STDP is in how to combine weight changes from multiple spike pairs. The simplest and most commonly used model

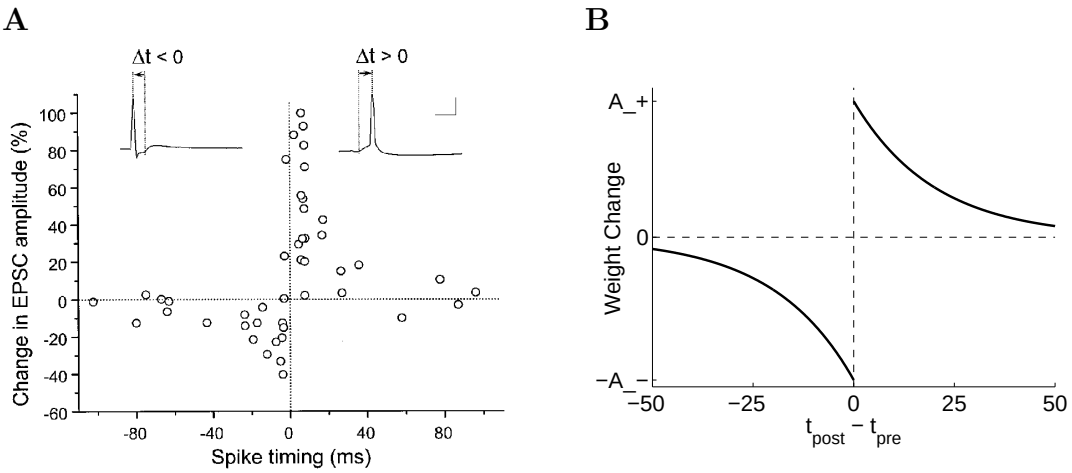


Figure 1.5: Classic STDP curve. **A.** Example of experimental results supporting STDP. Recordings were taken in the hippocampus. Each circle represents the percentage weight change due to an individual spike pair, where the time difference between the post- and presynaptic spikes is plotted on the x-axis. Image from (Bi and Poo [6]). **B.** Mathematical model of the STDP curve. The weight change due to each spike pair obeys an exponential distribution such that the change decays as the two spikes grow further apart.

is that the weight changes sum linearly (Song et al. [84]). Other models include a multiplicative model in which the weight change from Eq. 1.2 is multiplied by $(1-w)$ for LTP and by w for LTD, allowing the weights to smoothly approach their upper and lower bounds (Rubin et al. [76]), and a weighted linear model in which the first few spike pairs within a burst of spikes carries the largest weight in the summation (Froemke and Dan [19]). There is also a biophysical model based on calcium dynamics which exhibits STDP in a certain range of spike rates (Shouval et al. [81]). For the results presented in this thesis, I use the classic linear summation model of Song et al. due to its simplicity and its ability to approximate experimental results.

1.3 The Double Rotation Experiment

My network model simulates an experiment referred to as the *double rotation experiment*, performed by Dr. James Knierim and his lab at Johns Hopkins University. The experiment was originally designed to investigate the coherence among an ensemble of place cells when the external environment is modified by graded amounts. For the past decade, the findings from the experiment have elucidated many properties concerning the dynamics of the cognitive map, including attractor dynamics among place cells (Knierim [44]), functional differences between CA1 and CA3 place cells (Lee et al. [49, 48]), and the role of each subregion in the entorhinal cortical-hippocampal network in pattern separation and pattern completion (J. Neunuebel [manuscript in preparation]).

In the experiment a rat is trained over several days to move clockwise around a track with salient local cues on its surface and salient distal cues hung on curtains surrounding the track. After training the cues are rotated in opposite directions, the local cues counter-clockwise and the distal cues clockwise, where the mismatch angle varies between 45° and 180°. The activity of place cells in both the CA1 and CA3 subregions of the hippocampus is then recorded to determine how the place cell representation changes as a function of the mismatch angle. Figure 1.6 illustrates a 90° mismatch.

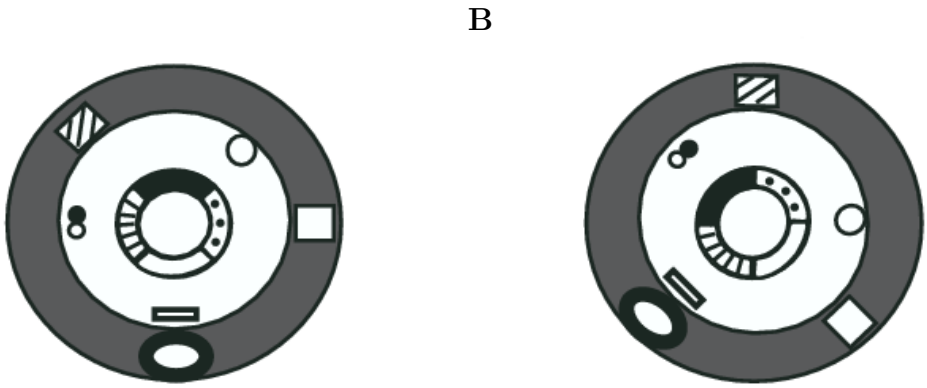


Figure 1.6: Setup for the double rotation experiment. Local cues are placed along the track, and distal cues are hung on dark curtains surrounding the track. **A.** Standard configuration. **B.** 90° mismatch configuration. With respect to **(A)**, local cues are rotated 45° counter-clockwise, and distal cues are rotated 45° clockwise. Images from (Knierim [44]).

1.3.1 Experimental Results

The experiment was among the first to demonstrate a functional difference between the CA1 and CA3, as shown in Figure 1.7. When the cues were rotated, most CA1 cells remapped, meaning that place fields appeared, disappeared, or split their representation. Of the remaining CA1 place cells, roughly 15% rotated with local cues while the other 15% rotated with distal cues. On the other hand, over 50% of place cells in the CA3 rotated with local cues (Lee et al. [49]). This relatively strong coherence in the CA3 may be due to the associational connections among CA3 place cells. As place cells form associations with one another, they are more likely to respond as a coherent unit than are place cells in the CA1, which form very few recurrent connections.

The Knierim lab took the experiment one step further by recording from the two major hippocampal inputs (LEC and MEC) as well as from the hippocampal DG

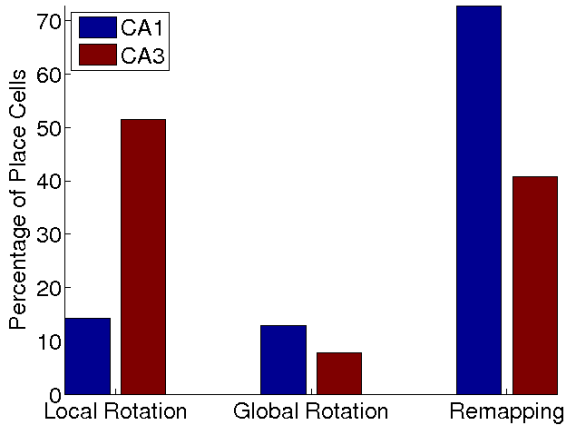


Figure 1.7: Population coherence in CA1 and CA3 subregions. Place cells in the CA3 respond much more coherently when the cues are rotated than do place cells in the CA1. Remapped cells indicate place fields that appear, disappear, or split into two distinct fields. Of place cells in the CA1, 14.3% rotate with local cues, 12.9% rotate with distal cues, and the remaining 62.8% remap. Of CA3 place cells, 51.5% rotate with local cues, 12.9% rotate with distal cues, and the remaining 40.8% remap. Data presented in (Knierim et al. [49]).

and CA3 subregions to tease apart the processes performed by the CA3 and those performed by processing units upstream. They found that the distal cues largely control the MEC population, while the CA3 and LEC populations are more controlled by the local cues (Neunuebel [manuscript in preparation]). These results lead to the search for mechanisms responsible for the coherent response of CA3 cells to follow the weak spatial signal from the LEC in rotating with local cues over the much stronger spatial signal from the MEC.

1.3.2 The Backward Shift Phenomenon

In experiments in which a rat repeatedly traverses a given route, place fields shift backward, in the opposite direction of the rat's movement (e.g. Blum and Abbott [7],

Mehta et al. [58]). In the context of the double rotation experiment, place fields shift counter-clockwise during the rat's initial laps around the track in the clockwise direction (Lee et al. [48]). The backward shift may have implications in navigation as it allows one to anticipate the next step along a learned route. This idea agrees with Donald Hebb's theory of a *phase sequence* in which cell assemblies representing one brain state should activate cell assemblies representing the next brain state within a learned sequence of events (Hebb [32]). The ability to neuronally represent the temporal sequence of events rather than simply the events themselves is a characteristic of episodic memory.

Place fields do not shift when synaptic plasticity is blocked (Ekstrom et al. [15]), implying that the shift is a reflection of associations formed among place cells representing locations along the route. Several simulation studies have demonstrated that synaptic plasticity among place cells can account for the backward shift, and two such studies were even done in a simplified setting of the double rotation experiment (Yu et al. [98, 99]). The underlying idea is that if cell A repeatedly fires before cell B, then the weight from A to B strengthens, causing cell A to activate cell B earlier than it would have fired previously. STDP is the most standard plasticity model able to create this feedforward structure in the connections among place cells.

1.4 Dendritic Computations

Neurons receive tens of thousands of inputs throughout their dendrites, or branched projections emanating from the cell body (see §2.1.1). Dendrites vary greatly in size, shape, and electrical properties, and the distribution of synaptic inputs into the dendritic tree is often highly structured. For example, the Lobula Giant Movement Detector (LGMD), a collision-detecting neuron in the locust, tracks an object's movement within the locust's field of vision through the topology of the excitatory inputs onto its dendrites, and the locations of active synapses depend on the object's location in visual space (Krapp and Gabbiani [47], O'Shea and Rowell [66]).

Despite the inherent structure in the spatial component of input signals, the output of a neuron is the generation of action potentials at the cell body, a binary output signal. It should be no surprise, then, that the dendrites often play a crucial role to a neuron's function as they receive inputs and transmit them to the cell body.

Synaptic locations within the dendritic tree of a place cell depend on the distinct brain regions containing the presynaptic neurons, where afferents from the hippocampus project to more proximal regions in the dendritic tree than do afferents from the entorhinal cortex (Shepherd [40]). Both modeling and experimental studies indicate that this spatial distribution endows the dendrites with the ability to perform coincidence detection, as described in §1.4.1. However, retaining the spatial component of input signals is computationally expensive, and network models of such neurons quickly become intractable. §1.4.2 contains an outline of existing reduced models

of the neuron that allow one to incorporate dendritic computations without the additional expense. I provide a more detailed description of these reduced models in §2.2.

1.4.1 Coincidence Detection in Place Cells

The DG and CA3 subregions can be considered as the processing units of the hippocampus. As shown in Figure 1.3, they both process sensory inputs from the entorhinal cortex through their recurrent connections. Furthermore, the projections from both subregions are primarily confined to the hippocampus via the trisynaptic circuit, EC→DG→CA3→CA1. In contrast, the CA1 provides the output of the hippocampus, sending projections both to the subiculum and back to the entorhinal cortex, primarily to the MEC (Shepherd [40]). It is uniquely positioned to compare direct information about the environment (from the EC) to information about the stored representation of this environment (from the DG/CA3). As the environment changes, its spatial representation in CA1 place cells depends on the interaction between inputs from the EC and CA3.

Coincidence Detection in the CA1 Implications of the spatial distribution of inputs onto CA1 neurons have been well-studied, and I present two ways in which CA1 dendrites could act as coincident detectors. First, dendrites act as a filter, and the signal from distal EC synapse activation does not reliably travel to the cell body. These signals travel through dendritic regions innervated by the Schaffer collaterals

and thus have a much greater chance of reaching the cell body if the EC and CA3 inputs are coincident (Jarsky et al. [39]). On the other hand, inhibition at the proximal synapses can effectively shunt the signal from the distal synapses. The proximal synapses therefore gate the distal signal from the EC (Spruston [85]).

Second, Hebbian plasticity rewards the coincident activation of pre- and postsynaptic cells and thus depends on the synapse's awareness of when the postsynaptic cell fires. This awareness is non-trivial as synapses are located in the dendrites, but action potentials are often generated in the soma. The key is in the back-propagating action potential, or a strong signal that travels back through the dendrites when a cell fires. The successful transmission of this signal to the distal dendrites is greatly aided by coincident inputs from the proximal Schaffer collaterals, and thus the degree of coincidence between EC and CA3 inputs affects the strength of the synapses between EC and CA1 cells (Spruston et al. [85]).

On the other hand, the EC can also act as a teaching signal to update the Schaffer collateral synapses. Signals generated through the perforant path travel downstream, depolarizing the synapses from the CA3. This depolarization could induce LTP, even if the postsynaptic CA1 cell has not produced an action potential (Dudman et al. [14]).

Coincidence Detection in the CA3 Similar to pyramidal cells in the CA1, CA3 pyramidal cells receive direct input from the EC at distal synapses and processed inputs from the DG and other CA3 cells more proximally. Although implications of the spatial distribution of CA3 inputs have not been as well-studied, dendrites in

the CA3 should be able to perform the same coincidence detection function as do their neighbors in the CA1 subregion. Accordingly, EC input may drive CA3 activity primarily when it is coincident with input from the DG or CA3. Similarly, synaptic plasticity between EC and CA3 neurons should be greatly aided by the coincidence of EC and hippocampal inputs.

1.4.2 Modeling the Dendritic Tree

While the spatial component of input signals can be crucial, retaining the spatial specificity of inputs necessitates complex computational models. Efforts to reduce this complexity extend back to the 1950s, and model reduction of the neuron remains an active area of research today.

Reduced models of the neuron can be sorted into two broad classes. The first class merges compartments, thereby coarsening the spatial discretization. The simplest method is to model an isopotential cell in which dendrites play no role. Other models divide the neuron into proximal and distal regions and empirically determine the electrical properties of the dendrite within each region. These models are successful in that they can reproduce complex behaviors, such as bursts of activity (Traub et al. [89]) and spike shapes (Bush and Sejnowski [11]). The reduction procedure is intuitive, and the reduced models have a clear biological interpretation that can lead to theories regarding dendritic function (e.g., Pinsky and Rinzel [69]). However, the spatial specificity in the input signal is sacrificed.

The second class employs model reduction techniques, such as Krylov subspace projection methods (Kellems et al. [43], Yan and Li [97]) and SVD-based methods (Kellems et al. [43, 42]). These models retain the spatial component in the input signal and reproduce the potential at a few prescribed locations within the neuron. However, the reduction procedure for many of these models is far from intuitive, and existing reduced models have no apparent interpretation in terms of the electrophysiology underlying the original models.

1.5 Overview

The hippocampus is driven by two distinct input signals from the entorhinal cortex, but the LEC is often overlooked in spatial memory tasks due to the weak spatial tuning of its signal compared to the robust spatial firing patterns of grid cells in the MEC. However, experiments such as the double rotation experiment indicate that the LEC may play a critical role in partial remapping paradigms. I propose that the LEC acts as a teaching signal to train the network upon the rat's initial exposure to a modified environment. As generally accepted, the MEC then provides the primary spatial drive for the place fields. I demonstrate this process and elucidate possible roles for the mechanisms involved using both a reduced CA3 neuron and a network model of the double rotation experiment.

Reduced Model I use Krylov subspace projection methods to construct reduced models of passive and quasi-active neurons that preserve both the spatial component of the input signal and the electrophysiological interpretation of the original models. The reduced models accurately capture the potential at a prescribed spike initiation zone (siz), typically assumed to be the soma, given a reduced system as little as 0.5% of the full system size with a speed-up factor of up to one hundred. Unlike existing reduced models, I adapt established reduction techniques so that the reduced models are represented by RC and RLC circuits similar to those representing the original passive and quasi-active models, respectively. Furthermore, the transformation from the full to the reduced system is straightforward and depends on intrinsic dendritic properties by which input signals are attenuated as they are transmitted to the siz.

I use this technique to reduce morphologically accurate CA3 neurons given currents appropriate for the CA3 and driven by inputs from the LEC, MEC, and neighboring CA3 place cells. The simulations demonstrate both the capability of the reduced model and the processes that could enable the LEC to act as a teaching signal.

Network Model I developed a computational platform to study the network dynamics that occur during the double rotation experiment. The model incorporates the LEC, MEC, and CA3 subregions, and the connections among neurons are governed by Hebbian plasticity. The dynamic connections present a computational challenge in that millions of variables must be stored and updated at each timestep of the

simulation, and I use a parallel implementation to store and update these variables efficiently. I further created a graphical user interface (gui) to accompany the model, allowing one to easily use the model to test a wide variety of hypotheses.

The simulation of the double rotation experiment can be divided into two stages. During training, place cells in the CA3 form single place fields given grid cell input and rate-based plasticity. These place fields then shift backward, as seen experimentally, due to STDP among the associational connections of CA3 cells. In the mismatch stage, the cues are rotated in opposite directions. I model the responses of MEC and LEC cells such that the population responses match those observed experimentally. The place cell responses to the cue rotations then depend on the inputs they receive as well as the plasticity mechanisms involved.

Outline I begin in Chapter 2 by reviewing the electrophysiology of neurons and the common neuronal models. I then provide a condensed history of model reduction techniques applied to the neuron, placing the model reduction portion of this thesis in context. The model reduction technique and results are presented in Chapter 3. The work of Chapter 3 was done in collaboration with Steven Cox and can be found in (Hedrick and Cox [33]). I then shift to the network model, describing both the details of the model and the training dynamics in Chapter 4. I close in Chapter 5 by using both the reduced and network models to demonstrate how certain plasticity mechanisms and behavioral paradigms can combine with the LEC input signal to cause the counter-intuitive response seen in the double rotation experiment.

Chapter 2

Neuronal Structure and Function

Considered by many to be the father of modern neuroscience, Spanish neuroanatomist Santiago Ramon y Cajal pioneered work in identifying the structure of neurons and the synaptic organization of the nervous system. He used (and improved) the Golgi stain, a staining method presented in 1857 by the Italian Camillo Golgi. Prior to the Golgi stain, it was very difficult to infer neuronal structure due to the high density of cell bodies and their branched projections in neural tissue. The genius in the Golgi stain is that it dyes only a few well-separated neurons in their entirety. Cajal systematically sketched many types of neurons, recognizing that while neurons vary in shape and excitability, they share common features and functions (Gabbiani and Cox [20]).

In §2.1, I describe the basic electrophysiology of a neuron and present common neuronal models. Neurons have tens to hundreds of branched projections that ren-

der these models computationally expensive, and many reduced models have been developed to capture certain aspects of a neuron's dynamics using a system of much smaller dimension. I present a condensed history of model reduction techniques applied to the neuron in §2.2 and describe my research goals in finding a minimal model of the neuron in §2.2.4.

2.1 Models of the Neuron

The cell membrane is comprised of a lipid bilayer perforated with pumps, exchangers, and channels regulating the flow of ions in and out of the cell. The pumps and exchangers continuously work to maintain an imbalance in intracellular and extracellular ionic concentrations, where $[K^+]_{in} \gg [K^+]_{out}$, $[Na^+]_{in} \ll [Na^+]_{out}$, $[Cl^-]_{in} \ll [Cl^-]_{out}$, and $[Ca^{2+}]_{in} \ll [Ca^{2+}]_{out}$.

Basic circuit theory dictates that, if allowed, ions will flow across the membrane to eliminate the potential and concentration gradients. The membrane controls this flow through ionic channels, each selective to a particular ion. Numerous ionic channels have been discovered, and a primary distinguishing factor between models of the neuron lies in the choice of ionic channels included. In §2.1.1, I briefly describe the structure of a neuron and the process through which neurons communicate. I then present three well-studied models of the neuron. The passive model of §2.1.2 captures the cell's subthreshold potential, the active model of §2.1.3 captures the full nonlinear dynamics of the cell, and the quasi-active model of §2.1.4 is a linearization

of the active model about rest. For a full treatment of these models, see (Gabbiani and Cox [20]).

2.1.1 Neuronal Structure

Neurons vary in size, shape, and excitability, but all neurons have certain features in common, as illustrated in Figure 2.1. The main components of the neuron are described below.

The *synapse* is the connection formed between two neurons. When an action potential (electrical impulse) from a presynaptic cell arrives at a chemical synapse, neurotransmitter molecules are released and travel across the synaptic cleft, binding to receptors on the postsynaptic cell. This binding opens ionic channels, enabling ions such as Na^+ and K^+ through the cell membrane. Most synapses form between an axon and either the soma or dendrites of the postsynaptic cell, although axo-axonal synapses have been found.

The *dendrites* are branched projections emanating from the cell body. Their primary role is to receive inputs from upstream neurons and transmit the resulting ionic current to the cell body. The distinct shape and excitability of dendrites affect how each type of neuron processes synaptic inputs.

The *soma*, or cell body, is generally considered to be the spike initiation zone (siz) of the neuron. It integrates inputs from all areas of the dendritic tree and, due to the high excitability of the axonal base connected to the soma, will generate an action

potential if sufficiently depolarized.

The *axon* reliably transmits the action potential to downstream neurons, where axons traveling greater distances are surrounded by an insulating myelin sheath. The action potential also back-propagates through the dendrites, providing a feedback signal to the synapses that aided in its creation.

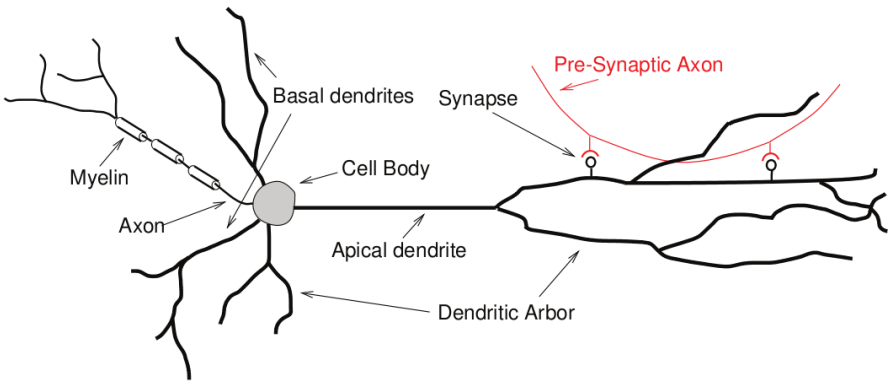


Figure 2.1: Schematic of a pyramidal cell, the neuronal type of a place cell. Pyramidal cells are characterized by bushy basal dendrites and a single apical dendrite that branches to form the apical tuft. A synapse, or connection between neurons, forms between a presynaptic axon and a small protrusion called a spine on the postsynaptic dendrite. The dendrites then transmit the signal to the cell body, where an action potential may be generated. Image from (Gabbiani and Cox [20]).

2.1.2 The Passive Cable

Consider an isolated branch of the dendritic tree modeled as a sealed cable of length ℓ and radius a, and assume the cable has m synapses at positions $\{x_j : j = 1, \ldots, m\}$. The current passing through each synapse j into the cell is given by $g(x_j, t)(\phi(x_j, t) - V(x_j))$, where $\phi(x, t)$ is the potential difference between the interior of the cell and extracellular fluid, and $g(x_j, t)$ and $V(x_j)$ are the conductance and reversal potential

of synapse j, respectively. The synaptic conductance briefly becomes nonzero upon binding of neurotransmitter at the synapse when the presynaptic cell fires an action potential. The time course of the conductance can vary, but it is often modeled as the alpha function,

$$g = \overline{g}((t-t_0)/\tau_\alpha)\exp(1-(t-t_0)/\tau_\alpha)\chi_{[t_0,\infty)}(t), \tag{2.1}$$

where $\overline{g}$ is the peak conductance, t_0 is the time of the presynaptic spike, τ_α is the decay constant, and $\chi_{[t_0,\infty)}(t) = 1$ if $t \geq t_0$ and 0 otherwise.

The resulting electrical signal travels along the cable according to intrinsic properties of the dendrite. The passive properties are described by the dendrite's membrane capacitance C_m $(\mu F/cm^2)$, axial resistivity R_a $(k\Omega\, cm)$, leakage conductance g_L (mS/cm^2), and leakage reversal potential V_L (mV). The nonlinearities in the membrane potential are introduced through the dendrite's active ionic channels characterized by voltage-dependent conductances, creating a nonlinear current denoted $I_{act}(\phi)$. The transmembrane potential is governed by

$$(2\pi a)\{C_m\partial_t\phi + g_L(\phi - V_L)\} + I_{act}(\phi) - \frac{\pi a^2}{R_a}\partial_{xx}\phi = -\sum_{j=1}^{m} g(x_j,t)\delta(x-x_j)(\phi - V(x_j)),$$

$$\phi(x,0) = \overline{\phi}, \quad \partial_x\phi(0,t) = \partial_x\phi(\ell,t) = 0, \tag{2.2}$$

where $\overline{\phi}$ is the resting potential and δ denotes the Dirac-delta function.

The passive model is obtained by setting $I_{act} = 0$, thereby including only passive ionic channels, which are voltage-independent. The resting potential is then given by $\overline{\phi} = V_L$, and the change of variables, $v \equiv \phi - V_L$ and $E(x_j) \equiv V(x_j) - V_L$, leads to

the classic passive cable equation,

$$(2\pi a)\{C_m\partial_t v + g_L v\} - \frac{\pi a^2}{R_a}\partial_{xx} v = -\sum_{j=1}^{m} g(x_j,t)\delta(x-x_j)(v-E(x_j)),$$
$$v(x,0) = 0, \quad \partial_x v(0,t) = \partial_x v(\ell,t) = 0. \tag{2.3}$$

Spatial discretization into compartments of length dx leads to the RC circuit representing the passive cable with $n = \ell/dx$ compartments, shown in Figure 2.2.

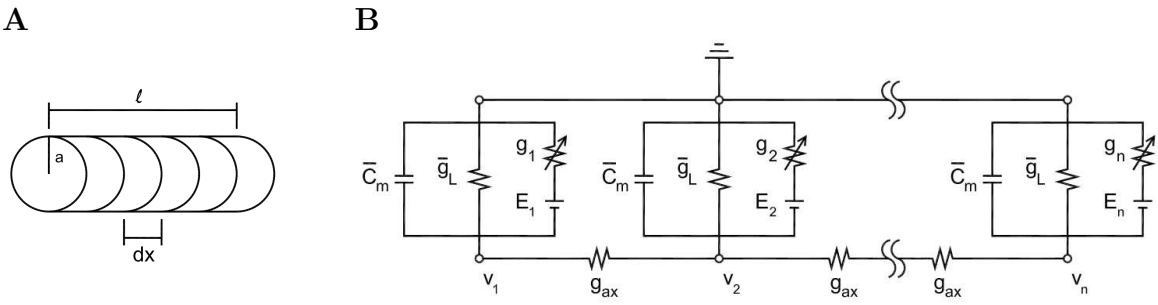


Figure 2.2: RC circuit for the passive cable. **A.** Spatial discretization. The cable has radius a and total length ℓ, and each of the n compartments has compartment length $dx = \ell/n$. The continuous variables of Eq. 2.3 evaluated at compartment j become $v(x_j,t) \approx v_j(t)$, $g(x_j,t) = g_j(t)$, and $E(x_j) = E_j$, where x_j is the center of compartment j. **B.** Schematic of the passive cable with n compartments driven by synaptic input. Each compartment has an identical membrane capacitance and leakage conductance given by $\overline{C}_m = AC_m$ and $\overline{g}_L = Ag_L$, respectively, where A is the compartmental surface area. Axial current flows between neighboring compartments proportional to the axial conductance, $g_{ax} = \pi a^2/(R_a dx)$. Each compartment contains a synapse with synaptic conductance $g_j(t)$ and reversal potential E_j. The transmembrane potential v_j as well as the reversal potential E_j are relative to the resting potential, and ground is considered to be extracellular fluid.

2.1.3 The Active Model

While the passive model describes the fundamental transmission of an electrical signal to the soma, the active model (Eq. 2.2) includes the voltage-dependent currents giving rise to nonlinearities in the transmembrane potential. The exact nature of these active currents were a mystery until the work of Hodgkin and Huxley in the 1950s

(e.g. Hodgkin and Huxley [37]). They individually isolated a potassium and sodium current in the squid giant axon by doctoring the bath to eliminate other ionic currents, applying a space clamp to eliminate the axial current, and applying a voltage clamp to fix the transmembrane potential, allowing them to measure each current over a range of potentials. They then constructed a model for each current and fit the parameters to best match experimental data. Although numerous voltage-dependent ionic currents have since been discovered, they are all well-described by the formalism developed by Hodgkin and Huxley.

Most active currents are Ohmic, given by $I_{act}(\phi) = Ag_{act}(\phi)(\phi - V_{ion})$, where V_{ion} is the reversal potential of the ion specific to the given channel. The conductance of any ionic current can be modeled using N_g gating variables, denoted $\{w_i\}$, via

$$g_{act}(\phi) = g_{ion} \prod_{i=1}^{N_g} w_i^{p_i}, \quad \text{where} \quad \tau_i(\phi)\partial_t w_i = w_{i,\infty}(\phi) - w_i, \quad w_i(x,0) = w_{i,\infty}(\overline{\phi}). \tag{2.4}$$

The term g_{ion} is the peak conductance per unit area of an open channel, and the time course for the activation and inactivation of the conductance is determined by the product of the gating variables, each raised to the power p_i. The general form for the functionals of each gating variable, $\tau_i(\phi)$ and $w_{i,\infty}(\phi)$, follow from the Boltzman equation, where the model parameters are set to match experimental data (Borg-Graham [8]).

For example, Hodgkin and Huxley observed that the potassium current in the squid giant axon monotonically increases to an asymptotic value and inferred the

model

$$I_K = Ag_K n^4(\phi - V_K), \quad \text{where} \quad \tau_n(\phi)\partial_t n = n_\infty(\phi) - n, \quad n(x,0) = n_\infty(\overline{\phi}).$$

Similarly, they observed that the sodium current exhibits a fast rise and slow decay, modeled through an activation gating variable m and inactivation gating variable h, where

$$I_{Na} = Ag_{Na} m^3 h(\phi - V_{Na}),$$

$$\tau_m(\phi)\partial_t m = m_\infty(\phi) - m, \quad m(x,0) = m_\infty(\overline{\phi}),$$

$$\tau_h(\phi)\partial_t h = h_\infty(\phi) - h, \quad h(x,0) = h_\infty(\overline{\phi}).$$

Together, the so-called delayed rectifier potassium current and fast-activating sodium current reproduce the generation of an action potential due to the fast influx of sodium and slow eflux of potassium.

2.1.4 The Quasi-Active Model

The quasi-active model lies between the passive and active models in terms of both its accuracy and its computational cost. While the passive model neglects active channels, the quasi-active model captures the linear perturbations of the active model from rest. It is accurate for small to moderate inputs and, although it cannot generate an action potential, it is more tractable than the active model and retains certain properties of the neuron, such as stability and resonance. I derive the quasi-active model for a particular example, but the linearization procedure is general and can be

applied to any active current.

Consider the active model of Eq. 2.2 for which $I_{act} = I_h$, the so-called h-current transmitting both sodium and potassium. Assume the current is defined by a single gating variable, q, such that $I_h = Ag_h q^2(\phi - V_h)$. The resting potential satisfies $g_L(\overline{\phi} - V_L) + g_h \overline{q}^2(\overline{\phi} - V_h) = 0$, where $\overline{q} = q_\infty(\overline{\phi})$. Assuming the synaptic conductances are small, they can be written as $g = \varepsilon \widetilde{g}$ for some $\varepsilon > 0$. The resulting potential and gating variable can be written as linear perturbations from rest, or

$$\phi = \overline{\phi} + \varepsilon \widetilde{\phi} + O(\varepsilon^2) \quad \text{and} \quad q = \overline{q} + \varepsilon \widetilde{q} + O(\varepsilon^2). \tag{2.5}$$

Substituting Eq. 2.5 into Eq. 2.2 and matching terms of order ε leads to the governing equation for the linear perturbation from rest of the transmembrane potential,

$$\begin{aligned} (2\pi a)\{C_m \partial_t \widetilde{\phi} + g_L \widetilde{\phi} + g_h(\overline{q}^2 \widetilde{\phi} + 2\overline{q}\widetilde{q}(\overline{\phi} - V_h))\} - \frac{\pi a^2}{R_a} \partial_{xx} \widetilde{\phi} \\ = -\sum_{j=1}^{m} \widetilde{g}(x_j, t)\delta(x - x_j)(\overline{\phi} - V(x_j)), \end{aligned} \tag{2.6}$$

$$\widetilde{\phi}(x, 0) = 0, \quad \partial_x \widetilde{\phi}(0, t) = \partial_x \widetilde{\phi}(\ell, t) = 0.$$

By replacing the functionals τ_q and q_∞ with their Taylor expansions about $\overline{\phi}$, the gating equation becomes

$$(\tau_q(\overline{\phi}) + \varepsilon \tau_q'(\overline{\phi})\widetilde{\phi})(\varepsilon \partial_t \widetilde{q}) = (\overline{q} + \varepsilon q_\infty'(\overline{\phi})\widetilde{\phi}) - (\overline{q} + \varepsilon \widetilde{q}) + O(\varepsilon^2).$$

By matching terms of order ε, the linear perturbation from rest of the gating variable is governed by

$$\tau_q(\overline{\phi}) \partial_t \widetilde{q} = q_\infty'(\overline{\phi})\widetilde{\phi} - \widetilde{q}, \quad \widetilde{q}(x, 0) = 0. \tag{2.7}$$

The quasi-active model thus consists of coupled ODEs for the linear perturbations of the potential and gating variables from rest. For small synaptic conductances eliciting a subthreshold response, $\phi \approx \overline{\phi} + \widetilde{\phi}$.

2.1.5 The Dendritic Tree

§2.1.2 through §2.1.4 consider only a single uniform dendritic branch sealed at its endpoints. In reality, dendrites branch profusely, forming dense trees. The models derived above extend naturally to the dendritic tree with the use of additional notation and a few more equations to describe the interaction among branches at junction points. For a thorough description of how to extend Eq. 2.2 to a forked branch, see (Gabbiani and Cox [20]).

Assume a given dendritic tree has N_D dendrites emanating from the soma, where dendrite d has B_d branches. The branches and their respective compartments are ordered according to the Hines ordering (Hines [36]), which specifies that all indeces are ordered from the point furthest from the soma. This ordering has numerical advantages when solving the resulting system as it retains sparsity upon use of an LU factorization. Let ℓ_b denote the length of branch b, and let $a_b(x)$ denote its radius as a function of distance, where $x = 0$ is the branch's endpoint furthest from the soma. Further assume a uniform capacitance per unit area, C_m, and axial resistivity, R_a.

The potential in the interior of each branch b, $\phi_b(x,t)$ for $0 < x < \ell_b$, is governed by Eq. 2.2. The endpoints of each branch are either isolated and therefore sealed or

connect to other branches through a junction point. Assume the dendrite d has J_d junctions, where junction J_d is between the soma and root branch. For $j < J_d$, assume the mother branch, b_j^3, gives rise to two daughter branches, b_j^1 and b_j^2. Continuity at the junction requires

$$\phi_{b_j^1}(\ell_{b_j^1}, t) = \phi_{b_j^2}(\ell_{b_j^2}, t) = \phi_{b_j^3}(0, t),$$

and current balance requires

$$a^2_{b_j^1}(\ell_{b_j^1})\partial_x\phi_{b_j^1}(\ell_{b_j^1}, t) + a^2_{b_j^2}(\ell_{b_j^2})\partial_x\phi_{b_j^2}(\ell_{b_j^2}, t) = a^2_{b_j^3}(0)\partial_x\phi_{b_j^3}(0, t).$$

Continuity at the N_D junctions with the soma requires

$$\phi_s(t) = \phi_{B_d}(\ell_{B_d}, t)$$

for $d = 1, \ldots, D$, where ϕ_s denotes the somatic potential, and branch B_d is the root branch for each dendrite.

Assume the soma is isopotential with surface area A_s, total active current $I_{act,s}(\phi_s)$, and a single synapse with conductance g_s and reversal potential V_s. Current balance requires

$$A_s\{C_m\partial_t\phi_s + g_L(\phi_s - V_L)\} + I_{act,s}(\phi_s) - \frac{\pi}{R_a}\sum_{d=1}^{N_D} a^2_{B_d}(\ell_{B_d})\partial_x\phi_{B_d}(\ell_{B_d}, t) = -g_s(\phi_s - V_s),$$

where $\phi_s(0)$ is set to the resting potential at the soma. Finally, let L denote the set of leafs, or branches that have no children. The sealed endpoint condition requires

$$\partial_x\phi_b(0, t) = 0, \quad b \in L.$$

The extension of the passive and quasi-active models to the dendritic tree follows in an identical fashion.

2.2 Reduced Models of the Neuron

As described in §1.4, the morphology and excitability of the dendritic tree as well as the spatial distribution of inputs affect the firing properties of a neuron. However, retaining these properties requires complex computational models with potentially tens to hundreds of thousands of variables. For example, distinguishing between inputs at the level of the spine requires a spatial discretization of 1 μm, giving a single dendritic branch of length 1 mm one thousand compartments. The complexity grows with the number of active ionic channels included, where the transmembrane potential and all gating variables must be tracked within each compartment, and realistically capturing the sodium, potassium, and calcium dynamics requires at least ten gating variables. Models become even more computationally complex if one wishes to include networks, where a small network twenty such neurons requires solving 200,000 ODEs, even in the simplistic case of non-plastic connections.

This high computational complexity has inspired many reduced models of the neuron. Wilfrid Rall was the first to systematically reduce the spatial complexity of the dendritic tree. He showed that, under certain conditions, the passive dendritic tree can be mapped to a single equivalent cable. Consider a single junction, where a mother branch of radius a_3 connects to two daughter branches of radii a_1 and a_2. The junction satisfies Rall's so-called 3/2 law if $a_1^{3/2} + a_2^{3/2} = a_3^{3/2}$. Rall showed that if the junction is symmetric (each daughter has equal electrotonic length), receives symmetric inputs onto each daughter, and satisfies the 3/2 law, then it can be mapped to a single cable.

He extended this idea to include all passive, symmetric trees receiving symmetric inputs and containing junctions which satisfy the 3/2 law (Rall [71]).

The assumptions prescribed by Rall are rarely satisfied in realistic dendrites, and models which followed his work (Schierwagen [79] and Poznanski [70]) were unable to relax the assumptions enough to render the reduction useful to models of the general dendritic tree. The numerous reduced models of the neuron that have since been proposed fall into two classes: those that merge compartments and those that use moment-matching or SVD-based model reduction techniques. In §2.2.1 I highlight a reduced model of the CA3 neuron obtained by merging compartments. I then describe reduced models for the quasi-active and active neurons in §2.2.2 and §2.2.3, respectively, that retain the spatial specificity in the input signal.

2.2.1 Two-Compartment Model of the CA3 Pyramidal Cell

CA3 pyramidal cells exhibit complex bursting patterns dependent on stimulus strength. In the early 1990s, Traub and Miles used voltage-clamp data to isolate the active conductances in different regions of the dendrite. They then constructed a 19-compartment model of the dendritic tree using six active channel types, where they determined the channel densities within each compartment to fit the data. This 19-compartment empirical model accurately captures the variety of bursting patterns seen experimentally, as well as synchronous bursting in networks when neurons mutually excite one another (Traub et al. [88, 89]).

Shortly after Traub and Miles presented their results, Pinsky and Rinzel reduced the model further by observing that most fast-spiking ionic currents were located in the soma and proximal compartments, while most calcium channels were relegated to the distal compartments. This inspired their two-compartment model in which the proximal compartment has fast-spiking currents and the distal compartment has Ca^{2+} dynamics. This model not only reproduced the somatic bursting patterns, but it also led to theories concerning the dependence of bursting on the electrotonic separation between fast, low-threshold currents in proximal regions and slow, high-threshold currents in distal regions (Pinsky and Rinzel [69]).

2.2.2 Reduction of the Quasi-Active Model

While models such as the Pinsky and Rinzel model can capture complex neuronal activity, the spatial specificity of inputs is lost when compartments are merged. I now transition to model reduction which retains the spacial specificity of inputs and faithfully reports the transmembrane potential at a few specified locations on the dendrite or soma. The model reduction techniques applied to this problem are well-studied and fall into two classes: SVD-based methods and moment matching methods. In 2009, Kellems constructed and compared reduced models of the quasi-active neuron using techniques from both classes (Kellems et al. [43]). Building on this work, he then used an SVD-based method to construct a reduced model of the active neuron (Kellems et al. [42]). I briefly present the techniques used below. See (Antoulas et

al. [2]) for a detailed description and comparison of the two classes of methods.

Balanced Truncation Upon spatial discretization of the quasi-active model for the general dendritic tree (§2.1.4 and §2.1.5), the linear system can be written as

$$z'(t) = Az(t) + Bu(t), \quad y(t) = Cz(t), \tag{2.8}$$

where the input vector u contains the synaptic current arriving at each compartment, and the state vector z contains the linear perturbations from rest of the transmembrane potential and gating variables at each compartment. The goal is to produce a reduced model that approximates the output vector, $y(t)$, containing the transmembrane potentials at a few specified locations within the neuron.

Balanced Truncation (BT) requires one to first transform the system into one that is balanced, and then to use the singular value decomposition (SVD) to truncate the system and thus reduce its dimension. A system is *balanced* if its controllability grammian P and observability grammian Q are equal and diagonal. Such a balanced system has the property that states that are difficult to reach are rarely observed. The grammians are computed by solving the Lyapunov equations,

$$AP + PA^* + BB^* = 0 \quad \text{and} \quad A^*Q + QA + C^*C = 0,$$

where A, B, and C are the matrices in the original system of Eq. 2.8.

For any controllable and observable system, there exists a transformation matrix T such that the basis change, $\hat{z} = Tz$, produces a balanced system. T is computed by first factoring the grammians into their Cholesky factors, $P = UU^*$ and $Q = LL^*$,

and then computing the SVD factorization $U^*L = Z\Sigma Y^*$. The transformation matrix is then given by $T = \Sigma^{1/2}Y^*L^*$, which has the inverse $T^{-1} = UZ\Sigma^{-1/2}$.

The system is transformed into its balanced counterpart by substituting the basis change $\hat{z} = Tz$ into Eq. 2.8, where

$$\hat{z}' = T(Az + Bu) = (TAT^{-1})\hat{z} + (TB)u,$$

$$y = Cz = (CT^{-1})\hat{z}.$$

The balanced system is then given by

$$\hat{z}' = \hat{A}\hat{z} + \hat{B}u, \quad y = \hat{C}\hat{z}, \tag{2.9}$$

where $\hat{A} = TAT^{-1}$, $\hat{B} = TB$, and $\hat{C} = CT^{-1}$.

The second stage of BT is to truncate the balanced system of Eq. 2.9. The grammians of this system are given by $TPT^{-1} = TQT^{-1} = \Sigma$, where Σ is a diagonal matrix containing the Hankel singular values of the original system (Eq. 2.8), given by the singular values of U^*QU. These singular values are nonnegative and in descending order, and including only the first k singular values leads to the reduced system

$$\xi' = \hat{A}_{kk}\xi + \hat{B}_k u, \quad \hat{y} = \hat{C}_k\xi,$$

where $\hat{A}_{kk}$ is the initial $k \times k$ submatrix of $\hat{A}$, $\hat{B}_k$ is the initial k rows of $\hat{B}$, and $\hat{C}_k$ is the initial k columns of $\hat{C}$. The reduced state vector ξ has no apparent biological interpretation, but it is transformed through the matrix $\hat{C}_k$ to produce an approximation $\hat{y}$ of the output vector y. The size of the reduced system necessary to guarantee a given error bound depends on the rate of decay of the Hankel singular values.

Moment Matching SVD-based methods produce an accurate reduced system with an attainable error bound, but the reduction process requires solving the Lyapunov equations to obtain the grammians, which becomes intractable for large systems. Furthermore, the system must be transformed before it is truncated, requiring the computation and storage of large matrices. Moment-matching methods do not suffer from such difficulties as they directly project the full system onto a smaller subspace. However, the reduced system is generally not amenable to a well-defined error bound.

Consider again the general quasi-active system of dimension n, given by Eq. 2.8. To construct a reduced system of dimension r, let V_r and $W_r \in \Re^{n\times r}$ such that $\text{Range}(V_r) \cap \text{Null}(W_r^T) = \{0\}$, implying that $W_r^T V_r$ is invertible. The reduced system is constructed such that its output, $\hat{y}$, is the Galerkin approximation of the true output, y. If for some $\tilde{z} \in \text{Range}(V_r)$, $W_r^T(\tilde{z}' - A\tilde{z} - Bu) = 0$, then the Galerkin approximation for y is given by $\hat{y} = C\tilde{z}$ (Gugercin et al. [26]). Let $\xi \in \Re^r$ such that

$$W_r^T(V_r\xi' - AV_r\xi - Bu) = 0.$$

Then,

$$\xi' - (W_r^T V_r)^{-1} W_r^T A V_r \xi - (W_r^T V_r)^{-1} W_r^T B u = 0, \quad \hat{y} = C V_r \xi.$$

Thus, the reduced system is given by

$$\xi' = \hat{A}\xi + \hat{B}u, \quad \hat{y} = \hat{C}\xi$$

for reduced matrices $\hat{A} = (W_r^T V_r)^{-1} W_r^T A V_r$, $\hat{B} = (W_r^T V_r)^{-1} W_r^T B$, and $\hat{C} = CV_r$.

The accuracy of the Galerkin approximation depends on the choice of the matrices

V_r and W_r, which are constructed to minimize the error in the frequency domain. By taking the Laplace transform of both sides of Eq. 2.8, $(sI - A)\mathcal{L}z(s) = B\mathcal{L}u(s)$, and thus the output in the frequency domain is given by

$$\mathcal{L}y(s) = H(s)\mathcal{L}u(s), \quad H(s) = C(sI - A)^{-1}B.$$

As A is negative definite, the region of convergence contains the right half of the complex plane, and y can be found from the inverse Laplace transform, $y(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{i\omega t}\mathcal{L}y(s)ds$, leading to the error bound

$$\begin{aligned}
|y(t) - \hat{y}(t)| &= \frac{1}{2\pi}\left|\int_{-\infty}^{\infty} \mathrm{e}^{i\omega t}(H(i\omega) - \hat{H}(i\omega))\mathcal{L}u(i\omega)\, d\omega\right| \\
&\leq \frac{1}{2\pi}\int_{-\infty}^{\infty} \|H(i\omega) - \hat{H}(i\omega)\|_2 \|\mathcal{L}u(i\omega)\|_2 d\omega \\
&\leq \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} \|H(i\omega) - \hat{H}(i\omega)\|_2^2\, d\omega\right)^{1/2}\left(\int_{-\infty}^{\infty} \|\mathcal{L}u(i\omega)\|_2^2\, d\omega\right)^{1/2}.
\end{aligned}$$

Assuming the input vector u is bounded, the problem reduces to constructing V_r and W_r such that $\hat{H}$ interpolates H and possibly its derivative at specified interpolation points, called shifts. Kellems chose the negative of the poles of $H(s)$ as the interpolation points and, as these poles are not generally known *a priori*, used the iterative rational Krylov algorithm (Gugercin et al. [26]) to iteratively find the poles.

2.2.3 Reduction of the Active Model

Many important dynamics in dendrites, such as the propagation of action potentials, requires the active model incorporating nonlinearities through gating variables (§2.1.3). Kellems also constructed reduced models of the active dendritic tree. Given

n compartments with g gating variables per compartment, let $\phi \in \Re^n$ be the transmembrane potential at each compartment and $w \in \Re^{n \cdot g}$ be the gating variables evaluated at each compartment. The full system is given by the coupled ODEs,

$$\begin{aligned} \phi'(t) &= H\phi(t) - N(\phi(t), w(t)) + G_{syn}(t)(\phi(t) - V_{syn}), \\ \text{diag}(B(\phi))w'(t) &= A(\phi) - w(t), \end{aligned} \tag{2.10}$$

where $H\phi$ encodes the axial and leakage currents, $N : \Re^{n \times (n \cdot g)} \to \Re^n$ is a nonlinear function encoding all active ionic currents, $G_{syn} \in \Re^{n \times n}$ is a diagonal matrix containing the synaptic conductances for each compartment, and $V_{syn} \in \Re^n$ contains the associated synaptic reversal potentials. The second equation governs the dynamics of all gating variables, given by Eq. 2.4, where $B, A : \Re^n \to \Re^{n \cdot g}$.

Kellems reduced the system through a combination of proper orthogonal decomposition (POD) and the discrete empirical interpolation method (DEIM). I briefly describe each method below.

Proper Orthogonal Decomposition POD is the standard SVD-based reduction method for nonlinear systems. Given N_s snapshots of the state variable, $\mathcal{V} = [\phi(t_1) \ \ldots \ \phi(t_{N_s})] \in \Re^{n \times N_s}$, the orthonormal basis $\{\varphi_i\}_{i=1}^r$, where $r \ll n$, is constructed to solve the optimization problem

$$\min_{\{\varphi_i\}_{i=1}^r} \sum_{j=1}^{N_s} \|\phi(t_j) - \sum_{i=1}^{N_s} (\phi(t_j)^T \varphi_i)\varphi_i\|_2^2. \tag{2.11}$$

The basis is computed using the SVD of the snapshot matrix, $\mathcal{V} = U\Sigma X^T$. As the diagonal elements of Σ are nonnegative and in descending order, $\mathcal{V}$ can be truncated

to include only the first r singular values, or $\mathcal{V} \approx U_r \Sigma_r X_r^*$, where U_r and X_r are the initial $n \times r$ submatrices of U and X, respectively, and Σ_r is the initial $r \times r$ submatrix of Σ. The columns of U_r solve the optimization problem of Eq. 2.11 (Liang et al. [53]); therefore, $\varphi = U_r$.

The reduced system is constructed by substituting the projection $\phi = U_r \xi$ into Eq. 2.10 to find

$$\begin{aligned} \xi'(t) &= (U_r^T H U_r)\xi - U_r^T N(U_r\xi, w) + (U_r^T G_{syn} U_r)\xi - (U_r^T G_{syn}) V_{syn} \\ \mathrm{diag}(B(U_r\xi))w'(t) &= A(U_r\xi) - w(t). \end{aligned} \tag{2.12}$$

Although the reduced state vector, ξ, has dimension $r \ll n$, the nonlinear term in Eq. 2.12 depends on $U_r\xi \in \Re^n$ and, even worse, on $w \in \Re^{n \cdot g}$, giving the reduced system the same complexity as the full system. This complexity can be eliminated by combining POD with a relatively new method called the discrete empirical interpolation method (DEIM), described below.

Discrete Empirical Interpolation Method DEIM is a discrete variant of the empirical interpolation method, which was derived for finite element applications, that considers discrete points in space (dendritic compartments for our application) and can be applied to arbitrary ODEs regardless of their origin. I briefly describe ideas underlying the method below. For full details, see (Chaturantabut and Sorensen [12]).

Let f denote the nonlinear term of Eq. 2.12, or $f \equiv N(U_r\xi(t), w(t))$. Given k orthonormal basis vectors $\{W_1, \ldots, W_k\} \subset \Re^n$, the nonlinear term can be approximated

by

$$f(t) \approx Wc(t), \tag{2.13}$$

where $W = [W_1, \ldots, W_k] \in \Re^{n \times k}$ and $c \in \Re^k$ is a coefficient vector. This overdetermined system can be solved uniquely for c if one includes only k rows of W. To be explicit, given k interpolation points $p_1, \ldots, p_k$, let $P = [e_{p_1}, \ldots, e_{p_k}]$, where e_{p_i} is the canonical unit vector corresponding to column p_i of the identity matrix in $\Re^n$. Then,

$$P^T f(t) = (P^T W)c(t)$$

is the analogue of Eq. 2.13, where only the rows corresponding to the interpolation points $\{p_i\}$ have been included. As the columns of W are linearly independent, $P^T W$ is invertible and c is uniquely determined, providing the approximation for the nonlinear term

$$f(t) \approx W(P^T W)^{-1}(P^T f(t)).$$

Note that $P^T f(t) = P^T N(U_r \xi(t), w(t))$ is simply the rows of the nonlinear term corresponding to the chosen interpolation points, or compartments. As N encodes the active ionic currents, row i of N depends only on row i of $U_r \xi$ and on the gating variables for compartment i.

Let $\overline{P} \in \Re^{(n \cdot g) \times (k \cdot g)}$ such that $\overline{P}^T w$ extracts all rows in w corresponding to the gating variables of the interpolation compartments $p_1, \ldots, p_k$. The reduced system of

Eq. 2.12 becomes

$$\xi'(t) = \hat{H}\xi(t) - \hat{N}(P^T U_r \xi, \hat{w}) + (U_r^T G_{syn}(t) U_r)\xi(t) - (U_r^T G_{syn}(t))V_{syn},$$

$$\text{diag}(\hat{B}(P^T U_r \xi))\hat{w}'(t) = \hat{A}(P^T U_r \xi) - \hat{w}(t)$$

where

$$\hat{w} = \overline{P}^T w \in \Re^{k \cdot g}$$

$$\hat{H} = U_r^T H U_r \in \Re^{r \times r}$$

$$\hat{N}(P^T U_r \hat{V}, \hat{w}) = U_r^T W (P^T W)^{-1} P^T N(U_r \hat{V}, w) : \Re^{k \times (k \cdot g)} \rightarrow \Re^r$$

$$\hat{B}(P^T U_r \hat{V}) = \overline{P}^T B(U_r \hat{V}) : \Re^k \rightarrow \Re^{k \cdot g}$$

$$\hat{A}(P^T U_r \hat{V}) = \overline{P}^T A(U_r \hat{V}) : \Re^k \rightarrow \Re^{k \cdot g}.$$

The success of the spatial interpolation depends on the basis W and interpolation points chosen. DEIM specifies that the basis be constructed using POD on snapshots of the nonlinear term as POD provides the optimal basis. DEIM then provides an algorithm for iteratively constructing interpolation points that provide an accurate reduced system and have a global error bound.

2.2.4 Minimal Models of the Neuron

Certain functional aspects of the neuron, such as bursting patterns (Pinsky and Rinzel) and somatic potential (Kellems), have been accurately reproduced using models of much smaller dimension and computational expense. The Pinsky and Rinzel two-compartment model is exemplary in that it has a clear biological interpretation.

Its reduction comes from the distributions of specific ionic channels as reported by Traub and Miles, and the reduced model can be interpreted as the interaction between dynamics in proximal and distal dendritic regions. This link between the reduced and full neuron led to theories concerning the necessary spatial separation between fast and slow dynamics in the dendrite. The drawback in the model is that it loses spatial specificity and would therefore not be able to reproduce all functional aspects of the neuron.

Kellems produced an accurate reduced model of the full active dendritic tree. By distributing the active channels according to the Traub and Miles specifications, this reduced model could reproduce the complex bursting patterns in CA3 neurons while retaining the spatial specificity of its inputs. However, his reduced model sacrifices the underlying electrophysiology of the system, and the reduction procedure is far from intuitive.

I present minimal models of the passive and quasi-active neurons that have an intuitive reduction procedure, reminiscent of the Pinsky and Rinzel model, while retaining the spatial specificity of the Kellems model. Following the lead of Kellems, I begin by constructing a minimal model of the passive and quasi-active dendritic tree. Each reduced model is minimal in the following ways:

- The reduced system has a dimension far smaller than that of the full system, dramatically reducing the computational complexity and simulation time.

- The reduction procedure is minimal in terms of its complexity, computational

expense, and storage requirements.

- The reduced model has an electrophysiological interpretation similar to that of the full model.

There are numerous mathematical techniques proven to be effective for reducing the dimension and computational complexity of the model. To minimize the computational expense and storage requirements for the reduction procedure, I use a moment matching method, as the full system is directly projected onto a reduced subspace without computing grammians or the singular value decomposition. There are still numerous techniques within the class of moment-matching methods, such as the IRKA algorithm used by Kellems, and I choose the technique which, for our application, has the most direct reduction procedure. I find this minimal reduction in the PRIMA algorithm for reducing RLC circuits (Odabasioglu et al. [63]), a Krylov subspace projection method. I then adapt the structure-preserving reduction technique presented in (Li and Bai [52]) to preserve the circuit structure in the original passive and quasi-active models.

Chapter 3

Structure-Preserving Model Reduction

The work presented in this chapter was done in collaboration with Steven Cox and can be found in (Hedrick and Cox [33]). We construct reduced models that capture the subthreshold potential at the prescribed spike initiation zone (siz) while preserving both the spatial specificity of inputs and the electrophysiological interpretation of the original models. We consider three cases: the passive neuron given injected current (§3.1), the quasi-active neuron given injected current (§3.2), and the passive neuron given synaptic conductances (§3.3). For each case we present the full and reduced models, explicitly compare the circuit structure of the reduced model to that of the full model, and determine the accuracy of the reduced system both numerically and theoretically. As the reduced models retain the underlying electrophysiology, they

may be useful not only computationally but also theoretically in helping to understand how dendrites process synaptic inputs.

The passive neuron given injected current is the simplest case as the model neglects all voltage-dependent ionic currents, resulting in a linear, time-invariant system. We use the classic reduction procedure referred to as PRIMA (Odabasioglu et al. [63]). While the reduced model is standard, the link between the electrophysiology of the reduced model to that of the original model has not been established in the literature. We show how the reduction depends on dendritic properties and provide circuit diagrams for the full and reduced models, explicitly defining each circuit element. We also provide a more concise moment-matching proof to analyze the accuracy of the model than those found in the literature (Villemagne and Skelton [92], Grimme [24], Li and Bai [52]). In §3.2 and §3.3, we refer back to the models, circuits, and theory presented in this section as the latter two cases are natural extensions of this simplified setting.

The quasi-active model captures the linear perturbations from rest of the transmembrane potential and can be represented by an RLC circuit. Existing reduced models based on PRIMA (Yan and Li [97]) or SVD-based methods (Kellems et al. [43]) are not amenable to synthesis as electrical circuits, and the electrophysiology of the reduced models is unclear. Techniques have been developed to preserve the block structure of a system's matrices (Freund [18], Li and Bai [52]), simplifying the synthesis as an electrical circuit. However, an RLC circuit representation is not apparent

from the direct application of these methods. We use the structure of the quasi-active neuron to adapt the reduction procedure presented in (Li and Bai [52]). The resulting RLC circuit for the reduced system relates to the original circuit in a similar fashion to the circuits for the passive neuron of §3.1.

The transmembrane potential for the passive neuron given synaptic conductances has a time-varying coefficient, causing the potential to depend nonlinearly on the synaptic conductance. Reduced models using projection methods have been constructed for weakly nonlinear systems by expanding the system as a Taylor series (Roychowdhury [75]), Volterra series (Phillips [67, 68], Bai and Skoogh [4], Lin et al. [54]), or quadratic expression (Gu [25]). The primary issue with these methods is that the reduced system grows in size with the number of synapses, although there are techniques to limit this growth (Li and Pileggi [51]). Yan and Li recently presented a reduced model in which the potential is accurately captured at proximal synapses (Yan and Li [97]). However, distal synapses often carry crucial information. For example, activity of LGMD neurons tracks the trajectory of a looming object, and proximal synapses are not activated until the object is near the locust. Similarly, afferents from the EC carrying information regarding the external environment synapse onto the distal regions of CA1 pyramidal cells. For many practical applications, these distal synapses cannot be neglected.

We present a complementary reduced model by using the same transformation derived for injected current in §3.1. The reduced model is a significant improvement

over the Yan and Li reduced model in that it is independent of the locations and number of synapses. We follow the lead of (Phillips [67]) in determining the accuracy through expanding the system as a Volterra series. We analytically and numerically determine the convergence rate of the Volterra series and the accuracy of the reduced system in capturing the first two Volterra terms.

3.1 Current Injection into the Passive Neuron

We begin by constructing a reduced model for the passive neuron given current injections, a linear, time-invariant system neglecting all voltage-dependent currents. We derive both the full and reduced systems in the context of a cable, or single dendritic branch, for simplicity of notation. The models can be generalized for any neuronal morphology with additional notation and current balance at each junction between branches.

In §3.1.1 we show how injected currents differ from synaptic input currents, deriving the governing equation for the transmembrane potential given current injections. Spatial discretization leads to the full model for the passive cable, presented in §3.1.2, which we reduce using the PRIMA algorithm in §3.1.3 and numerically test in §3.1.4. In §3.1.5 we establish the link between the electrophysiology of the full and reduced systems by explicitly comparing the RC circuits representing both systems. We further interpret the transformation from the full system to the reduced system in terms of dendritic properties, both graphically and analytically. §3.1.6 contains a review of

the theory underlying the reduced models, where we provide a more concise moment-matching proof for our particular application than can be found in the literature. We close the section by extending the reduced model to any general morphology and simulating a reduced CA1 pyramidal cell in §3.1.7.

3.1.1 The Continuous Passive Cable

Consider the passive cable equation given by Eq. 2.3. For small, transient synaptic conductances, the potential is close to rest while the synapse is active, or $gv \approx 0$. We can then approximate Eq. 2.3 by the corresponding system for current injection,

$$\begin{gathered}(2\pi a)\{C_m\partial_t v + g_L v\} - \frac{\pi a^2}{R_a}\partial_{xx} v = \sum_{j=1}^{m} g(x_j,t)\delta(x-x_j)E(x_j), \\ v(x,0) = 0, \quad \partial_x v(0,t) = \partial_x v(\ell,t) = 0,\end{gathered} \tag{3.1}$$

where v is the transmembrane potential relative to rest, a is the radius of the cable, C_m is the membrane capacitance per unit area, g_L is the membrane leakage conductance per unit area, and R_a is the axial resistivity. The cable contains m synapses at locations $\{x_j : j = 1,\ldots,m\}$ with synaptic conductances $g(x_j,t)$ and reversal potentials relative to rest $E(x_j)$.

3.1.2 The Discrete Passive Cable

Upon spatial discretization of Eq. 3.1 into compartments of length dx, we arrive at an RC circuit with $n = \ell/dx$ compartments, shown in Fig. 3.2A. Our interest is in the potential at the spike initiation zone (siz), usually assumed to be the soma, which

is determined through the n-dimensional linear system

$$Cv'(t) + Gv(t) = u(t), \quad y(t) = e_{siz}^T v(t), \tag{3.2}$$

where v_j is the transmembrane potential at the center of compartment j, and y is the potential at the siz. The capacitance and conductance matrices are given by

$$C = AC_m I_n \quad \text{and} \quad G = Ag_L I_n - G_{ax}, \tag{3.3}$$

respectively, where $A \equiv 2\pi adx$ is the compartmental surface area, I_n is the $n \times n$ identity matrix, and G_{ax} encodes the dendritic structure such that the j^{th} element of $G_{ax}v$ is the total axial current entering compartment j. For the uniform cable,

$$G_{ax} = \frac{\pi a^2}{R_a dx} \begin{pmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -1 \end{pmatrix}. \tag{3.4}$$

The input locations x_j now correspond to input compartments $p_j = \text{ceil}(x_j/dx)$, giving rise to the input vector u with nonzero elements

$$u_{p_j}(t) = g_{p_j}(t)E_{p_j}, \quad j = 1, \ldots, m. \tag{3.5}$$

Finally, the output y is obtained through the canonical unit vector e_{siz}, which takes a value of 1 at the element corresponding to the siz.

3.1.3 The Reduced Model

Given a reducer X of size $n \times r$, where $r \ll n$, the reduced system is constructed such that its state vector $\hat{v}$ obeys $X^T(CX\hat{v}' + GX\hat{v} - u) = 0$. That is, the residual of Eq. 3.2 evaluated at $v = X\hat{v}$ is orthogonal to X (Gugercin et al. [26]). This condition is satisfied by the reduced system

$$\hat{C}\hat{v}'(t) + \hat{G}\hat{v}(t) = X^T u(t), \quad \hat{y}(t) = e_{siz}^T X\hat{v}(t), \tag{3.6}$$

where the reduced $r \times r$ capacitance and conductance matrices are given by

$$\hat{C} = X^T CX \quad \text{and} \quad \hat{G} = X^T GX, \tag{3.7}$$

respectively, and the input vector u contains the currents injected into each compartment of the full model, given by Eq. 3.5. The goal of model reduction is to construct a reducer X such that $\hat{y} \approx y$, the siz potential, while the resulting reduced system has a much smaller dimension than does the full system.

As the passive neuron is well represented by an RC circuit, we use a reduction procedure shown to be effective for reducing RC circuits (Odabasioglu et al. [63]). The idea is to project the full system onto a particular Krylov subspace of dimension r. For any given matrix A and vector b, let $\mathrm{sp}(A)$ denote the range of A, or the span of its columns, and define the Krylov subspace $K_r(A, b) \equiv \mathrm{sp}(b, Ab, A^2b, \ldots, A^{r-1}b)$. If $r \geq 3$ and $\mathrm{sp}(X) = K_r(G^{-1}, G^{-1}e_{siz})$, then the reduced system accurately captures the siz potential both in theory and in practice. Such a reducer can be generated

from the basis vectors

$$X_{basis} = \begin{pmatrix} G^{-1}e_{siz} & G^{-2}e_{siz} & \dots & G^{-r}e_{siz} \end{pmatrix}. \tag{3.8}$$

For numerical stability one may use the Arnoldi procedure to orthogonalize the columns of X_{basis}, thereby obtaining a well-conditioned system in which $X^T X = I_r$. Algorithm 1 outlines this procedure. The algorithm is classic and can be found in many texts, including (Trefethen and Bau [90]).

Algorithm 1: Arnoldi procedure for the uniform cable
Input: $G,\ e_{siz},\ r$ **Output:** X
$X_1 = G^{-1}e_{siz}/\|G^{-1}e_{siz}\|$ **for** $j = 1 : r-1$ $\quad x = G^{-1}X_j$ $\quad$ **for** $k = 1 : j$ $\quad\quad x = x - (X_k^T x)X_k$ $\quad X_{j+1} = x/\|x\|$ $X = (X_1 \ \ X_2 \ \ \dots \ \ X_r)$

Algorithm 2: Arnoldi procedure for the cable with a variable radius
Input: $G,\ D,\ e_{siz},\ r$ **Output:** X
$Y_1 = D^{1/2}G^{-1}e_{siz}/\|D^{1/2}G^{-1}e_{siz}\|$ $X_1 = D^{-1/2}Y_1$ **for** $j = 1 : r-1$ $\quad y = D^{1/2}G^{-1}DX_j$ $\quad$ **for** $k = 1 : j$ $\quad\quad y = y - (Y_k^T y)Y_k$ $\quad Y_{j+1} = y/\|y\|$ $\quad X_{j+1} = D^{-1/2}Y_{j+1}$ $X = (X_1 \ \ X_2 \ \ \dots \ \ X_r)$

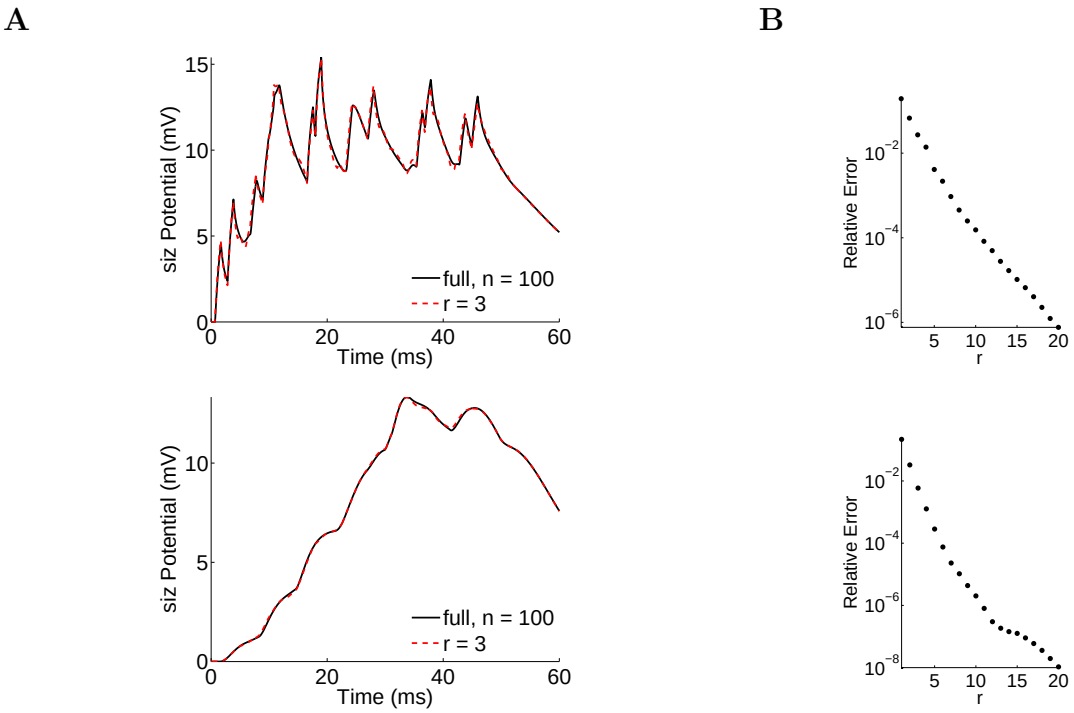


Figure 3.1: Testing of the reduced passive cable driven by current injections. Fifty synapses are randomly distributed throughout the cable with square pulse conductances (top) given by Eq. 3.10 or with alpha conductances (bottom) given by Eq. 3.11. **A.** Response at the siz in a 100-compartment cell and the corresponding 3-compartment reduced cell to identical current injections. **B.** Relative 2-norm error in the siz potential given the reduced system with r compartments. The reduced systems require 40% to 60% of the full simulation time.

3.1.4 Numerical Examples

Consider the passive uniform cable for which

$$\ell = 1\,mm,\ dx = 0.01\ mm,\ a = 1\,\mu m,\ C_m = 1\,\mu F/cm^2,$$

$$R_a = 0.3\,k\Omega\,cm,\ \text{and}\ g_L = 1/15\,mS/cm^2. \tag{3.9}$$

Unless otherwise specified, these parameters are used throughout the chapter, although the results would be similar given any realistic parameter set. Assume each synapse j is excitatory, where $E_{p_j} = 50\ mV$. For example 1 we drive the cable with

square pulse conductances, each of duration 1 ms with onset times t_{p_j} and the peak conductance $\overline{g}$. The conductance at synapse j is given by

$$g_{p_j}(t) = \bar{g}\chi_{[t_{p_j}, t_{p_j}+1]}(t), \quad \bar{g} = 1\, nS, \tag{3.10}$$

where $\chi_{[t_{p_j}, t_{p_j}+1]}(t) = 1$ if $t \in [t_{p_j}, t_{p_j} + 1]$ and 0 otherwise. For example 2 we apply the alpha conductances

$$g_{p_j}(t) = \overline{g}((t - t_{p_j})/\tau_\alpha) \exp(1 - (t - t_{p_j})/\tau_\alpha)\chi_{[t_{p_j}, \infty)}(t), \tag{3.11}$$

where $\bar{g} = 0.2\, nS$ and $\tau_\alpha = 2\, ms$. For both examples the synaptic compartments $\{p_j\}$ and onset times $\{t_{p_j}\}$ are selected randomly. The numerics indicate that the siz potential is well-approximated using a reduced system no greater than 5% of the full system size, as shown in Fig. 3.1. Example 1 can be considered as a worst case scenario as sharp input pulses elicit jagged responses in the potential. The results of these two examples are representative, and the reduction is similarly effective for any general injected currents.

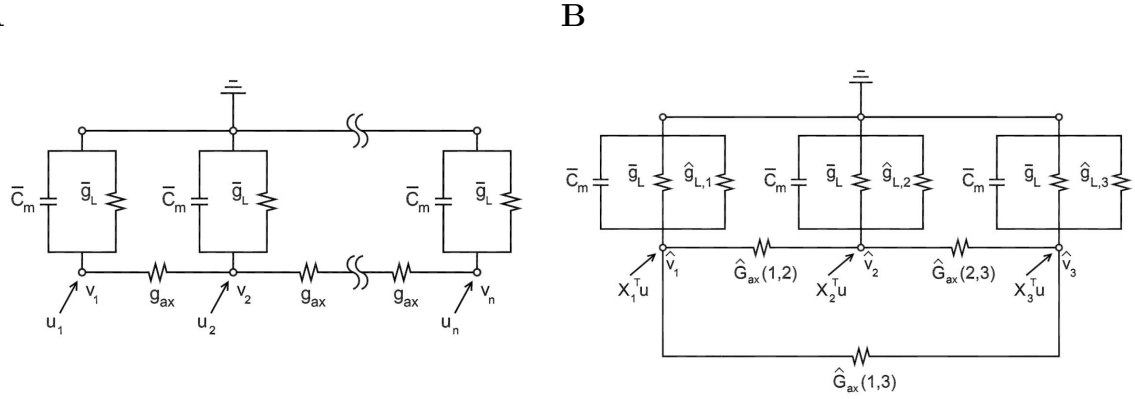
3.1.5 RC Circuit Structure

Since X is orthonormal, by Eqs. 3.3 and 3.7,

$$C = AC_mI_n \;\Rightarrow\; \hat{C} = AC_mI_r, \quad \text{and}$$

$$G = Ag_LI_n - G_{ax} \;\Rightarrow\; \hat{G} = Ag_LI_r - X^TG_{ax}X,$$

where I_n and I_r denote the $n \times n$ and $r \times r$ identity matrices, respectively. As C and G are symmetric positive definite, so too are $\hat{C}$ and $\hat{G}$, the capacitance and



$$y = v_1$$
$$\overline{C}_m = 6.28 \times 10^{-7}\ \mu F,\ \overline{g}_L = 0.042\ nS$$
$$g_{ax} = 104.7\ nS$$

$$\hat{y} = X_1(1)\hat{v}_1 + X_2(1)\hat{v}_2 + X_3(1)\hat{v}_3$$
$$\hat{g}_{L,1} = 0.057\ nS,\ \hat{g}_{L,2} = -0.209\ nS,\ \hat{g}_{L,3} = 0.785\ nS$$
$$\hat{G}_{ax}(1,2) = -0.081\ nS,\ \hat{G}_{ax}(1,3) = 0.108\ nS,$$
$$\hat{G}_{ax}(2,3) = -0.35\ nS$$

Figure 3.2: RC circuit structure for the passive cable given current injections.
A. Schematic of the full cable with n compartments, each containing the membrane capacitance $\overline{C}_m = AC_m$ and membrane leakage conductance $\overline{g}_L = Ag_L$, where A is the compartmental surface area. Ground is considered to be extracellular fluid. Axial current flows between neighboring compartments proportional to $g_{ax} = \pi a^2/(R_a dx)$. Each compartment j is driven by the injected current u_j given by Eq. 3.5, and the first compartment contains the siz. **B.** Schematic of the reduced cable with three compartments, each containing a membrane capacitance and leakage conductance identical to those in **(A)**. Each compartment j also has an additional membrane leakage conductance $\hat{g}_{L,j}$ given by Eq. 3.14. Axial current flows between any two compartments j and k proportional to $\hat{G}_{ax}(j,k)$, given by Eq. 3.13. Each reduced compartment j is driven by the injected current $X_j^T u$, where X_j denotes the j^{th} column of X, and the siz potential is approximated by the summation of the compartmental potentials weighted by the first row of X.

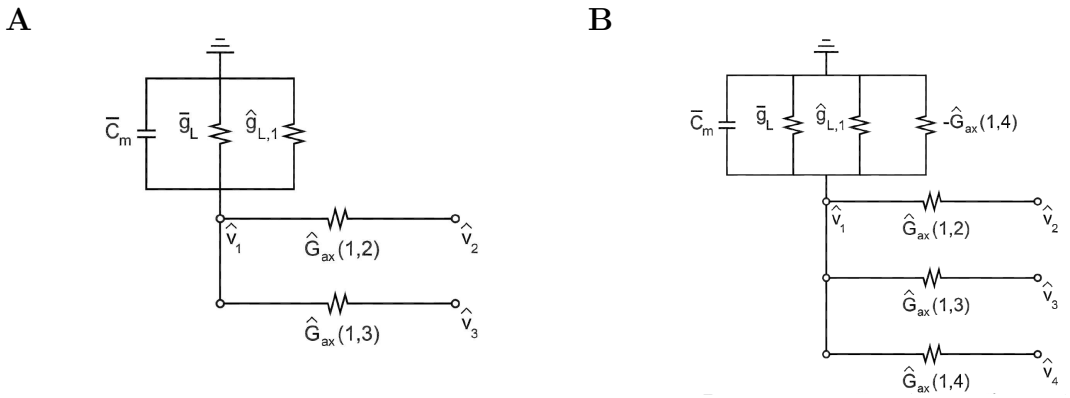


Figure 3.3: Augmentation of the reduced circuit. Given a reduced RC circuit of size r, the circuit of size $r+1$ can be constructed with minimal changes to the existing circuit. **A.** Subcircuit of Fig. 3.2B showing all components connected to compartment 1 in a reduced circuit with three compartments. **B.** All components connected to compartment 1 when a fourth compartment is added to the existing circuit. A resistor is added to ground with conductance $-\hat{G}_{ax}(1,4)$ and between compartments 1 and 4 with conductance $\hat{G}_{ax}(1,4)$. All components present in **(A)** are unchanged.

conductance matrices for the reduced system. The full and reduced systems can thus be interpreted as RC circuits in which each compartment has a membrane capacitance, AC_m, in parallel with a membrane leakage conductance, Ag_L. To resolve the second term of $\hat{G}$, we decompose it into an axial term and an additional leakage term,

$$-X^T G_{ax} X = -\hat{G}_{ax} + \mathrm{diag}(\hat{g}_L), \tag{3.12}$$

where the j^{th} element of $\hat{G}_{ax}\hat{v}$ is the total axial current entering reduced compartment j, and $A\hat{g}_{L,j}$ is an additional leakage conductance for reduced compartment j. Since the sum of each row in $\hat{G}_{ax}$ must be zero, Eq. 3.12 implies that for any j, k,

$$\hat{G}_{ax}(j,k) = X_j^T G_{ax} \begin{cases} X_k & \text{if } j \neq k, \\ -\sum_{i \neq j} X_i & \text{if } j = k; \end{cases} \tag{3.13}$$

where X_j denotes column j of X. The diagonal of Eq. 3.12 reveals that for any j,

$$\hat{g}_{L,j} = -X_j^T G_{ax} \sum_{k=1}^{r} X_k. \tag{3.14}$$

The circuits for the full and reduced systems are compared in Fig. 3.2.

Given a reduced system of size r, the reduced system of size $r+1$ can be easily constructed in the sense of both the underlying circuit and the reduced matrices. When adding compartment $r+1$ to the reduced system, two resistors are added for each existing compartment $j \leq r$. The first resistor is placed between compartment j and ground and has conductance $-\hat{G}_{ax}(j, r+1)$, and the second resistor is placed between compartments j and $r+1$ and has conductance $\hat{G}_{ax}(j, r+1)$. All existing components connected to compartment j remain unchanged, and the initial $r \times r$ submatrices of $\hat{C}_{r+1}$ and $\hat{G}_{r+1}$ are simply $\hat{C}_r$ and $\hat{G}_r$, respectively, where the subscript specifies the reduced system size. This can be seen mathematically, where if the n x r reducer X is augmented by X_{r+1} via one more step of the Arnoldi procedure, then the reduced capacitance and conductance matrices become

$$\hat{C}_{r+1} = \begin{pmatrix} X^T \\ X_{r+1}^T \end{pmatrix} C(X \quad X_{r+1}) = \begin{pmatrix} \hat{C}_r & 0 \\ 0 & AC_m \end{pmatrix}$$

and

$$\hat{G}_{r+1} = \begin{pmatrix} \hat{G}_r & X^T G X_{r+1} \\ X_{r+1}^T G X & X_{r+1}^T G X_{r+1} \end{pmatrix},$$

where $X^T G X_{r+1} = -\hat{G}_{ax}(1:r, r+1)$. Fig. 3.3 illustrates the augmentation of the reduced circuit.

One unconventional nuance of the reduced circuit is that roughly half the conductances are negative. From a circuit standpoint, negative conductances can be interpreted as a reversal in polarity, which can be implemented through op-amps. From a neuronal standpoint, positive axial conductances connect two compartments that excite one another, and negative axial conductances connect two compartments that inhibit one another. Similarly, positive leakage conductances imply that when the compartment is depolarized, positive current leaks out of the cell, as is the case for the full neuron. On the other hand, negative leakage conductances imply that positive current flows into depolarized compartments, causing a small depolarization to lead to large potential changes. However, the passivity of the overall circuit is maintained, a well-established result of the model reduction procedure (Freund [17]).

Input Mapping Given the input vector u containing the currents injected into each compartment of the full neuron, the input current $X_j^T u$ is injected into each compartment j of the reduced neuron. The columns of X thus map the true inputs to their reduced representation. The element in each column corresponding to the siz compartment then weights each potential to produce the siz approximation. Fig. 3.4 shows the columns and basis vectors of X.

As the only degree of freedom in constructing the reduced model is the choice of X_{basis}, we pause to derive its analytical expression. Assuming compartment 1 contains

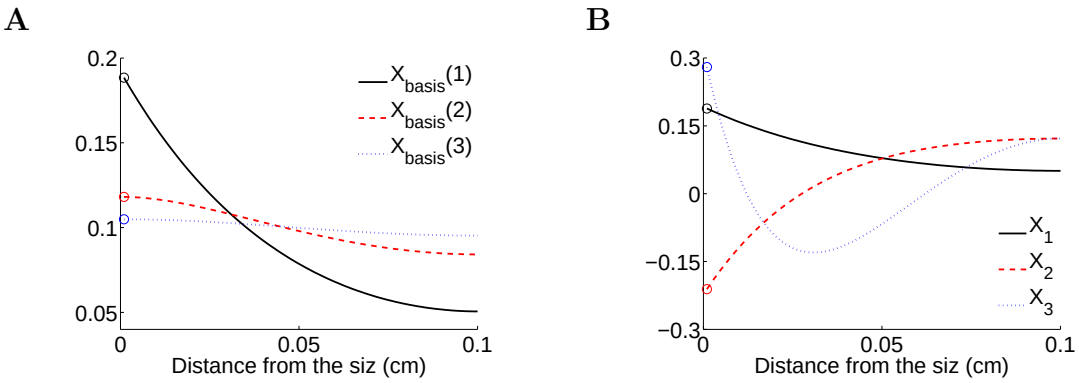


Figure 3.4: Input mapping. **A.** Columns of X_{basis} (Eq. 3.8) presented as a function of distance from the siz. The columns are orthogonalized to construct the columns of X. **B.** Orthonormal columns of X. Column 1 (solid black) specifies the relative impact that current injected at each location has on reduced compartment 1. The remaining two curves specify the impact that injected current has on reduced compartments 2 and 3. The circled elements weight the compartmental potentials as they are summed to approximate the siz potential.

the siz, the first column of X is the normalized solution of $GX_1 = e_{siz}$, i.e.,

$$(Ag_L I_n - g_{ax} S)X_1 = (1\ 0\ \cdots 0)^T,$$

where S is the second difference matrix in Eq. 3.4 and $g_{ax} = \pi a^2/(R_a dx)$. We divide by dx and let $dx \to 0$ to arrive at the boundary value problem for the continuous analogue of $X_{basis}(1)$,

$$-\lambda^2 f_1''(x) + f_1(x) = 0, \quad f_1'(0) = -R_a/(\pi a^2), \quad f_1'(\ell) = 0,$$

where $\lambda^2 \equiv a/(2R_a g_L)$ is the commonly used space constant. The solution is proportional to $\cosh((\ell - x)/\lambda)$. Hence X_1, which maps the true inputs into the input for reduced compartment 1, acts as the dendritic filter to attenuate each input according to the electrotonic distance between the synapse and siz. The subsequent basis

vectors are given by the inverse iteration on G, where for any j,

$$f_{j+1}(x) = \cosh((\ell - x)/\lambda) \int_0^x \cosh(y/\lambda) f_j(y)\, dy + \cosh(x/\lambda) \int_x^\ell \cosh((\ell - y)/\lambda) f_j(y)\, dy.$$

The sequence converges to the constant eigenvector associated with the smallest eigenvalue of G.

3.1.6 Theory

In Fig. 3.1 we illustrate that if X is generated from the basis vectors given by Eq. 3.8, then the siz potential is well-approximated by the reduced system. We now provide the theory that justifies our choice for X_{basis}. According to the classic theory for moment-matching methods, the accuracy of the reduced system can be analyzed by examining the transfer functions, which provide the input-output maps in the frequency domain,

$$\mathcal{L}y(s) = H(s)\mathcal{L}u(s) \quad \text{and} \quad \mathcal{L}\hat{y}(s) = \hat{H}(s)\mathcal{L}u(s),$$

where $\mathcal{L}$ denotes the Laplace transform. The transfer functions are given by

$$H(s) = e_{siz}^T (G + sC)^{-1} \quad \text{and} \quad \hat{H}(s) = e_{siz}^T X (\hat{G} + s\hat{C})^{-1} X^T.$$

The inverse Laplace transform leads to the error bound

$$\begin{aligned} |y(t) - \hat{y}(t)| &= \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} \mathrm{e}^{i\omega t} (H(i\omega) - \hat{H}(i\omega)) \mathcal{L}u(i\omega)\, d\omega \right| \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|H(i\omega) - \hat{H}(i\omega)\| \|\mathcal{L}u(i\omega)\| d\omega \\ &\leq \frac{1}{2\pi} \left(\int_{-\infty}^{\infty} \|\mathcal{L}u(i\omega)\|^2\, d\omega \right)^{1/2} \left(\int_{-\infty}^{\infty} \|H(i\omega) - \hat{H}(i\omega)\|^2\, d\omega \right)^{1/2}. \quad (3.15) \end{aligned}$$

Hence the goal is to minimize the error between the transfer functions along the imaginary axis. For the passive system, H has only real, negative poles and is a low-pass filter, as illustrated in Fig. 3.5A. We thus seek a reduced system for which $\hat{H} \approx H$ at low frequencies and compute the Taylor expansions about the origin,

$$H(s) = \sum_{j=0}^{\infty} (-s)^j M_j, \quad M_j = e_{siz}^T (G^{-1}C)^j G^{-1};$$
$$\hat{H}(s) = \sum_{j=0}^{\infty} (-s)^j \hat{M}_j, \quad \hat{M}_j = e_{siz}^T X (\hat{G}^{-1}\hat{C})^j \hat{G}^{-1} X^T. \tag{3.16}$$

The reducer X is constructed to match the leading r moments, for if $\hat{M}_j = M_j$ for $0 \leq j < r$, then $\hat{H}(s) = H(s) + O(s^r)$. We modify Theorem 3.3 of (Li and Bai [52]) to obtain Prop. 1.

Prop. 1 If $K_r(G^{-1}C, G^{-1}e_{siz}) \subseteq \text{sp}(X)$, then $\hat{M}_j = M_j$ for $0 \leq j < r$.

Proof: Define $P \equiv X\hat{G}^{-1}X^T G$. Since $P^2 = P$ and $PX = X$, P is an oblique projector onto $\text{sp}(X)$. Thus,

$$\hat{M}_0^T = X\hat{G}^{-1}X^T e_{siz} = PG^{-1}e_{siz} = G^{-1}e_{siz} = M_0^T$$

by Eq. 3.16, where the penultimate equality follows from the assumption that $G^{-1}e_{siz} \in \text{sp}(X)$ and the fact that P acts like the identity on $\text{sp}(X)$. By substituting $\hat{C} = X^T C X$ into Eq. 3.16,

$$\hat{M}_1^T = X\hat{G}^{-1}X^T C X \hat{G}^{-1}X^T e_{siz} = PG^{-1}CPG^{-1}e_{siz} = G^{-1}CG^{-1}e_{siz} = M_1^T,$$

where the penultimate equality follows from the assumption that $(G^{-1}C)G^{-1}e_{siz} \in$

sp(X). The remaining moment equalities follow in an identical fashion. ■

Note that for the uniform cable, $C = AC_mI_n$, implying that $K_r(G^{-1}C, G^{-1}e_{siz}) = K_r(G^{-1}, G^{-1}e_{siz})$. In this case a reducer generated from Eq. 3.8 satisfies the assumptions of Prop. 1. In Fig. 3.5 we illustrate that the transfer function H is a low-pass filter well-approximated by $\hat{H}$ given small frequencies.

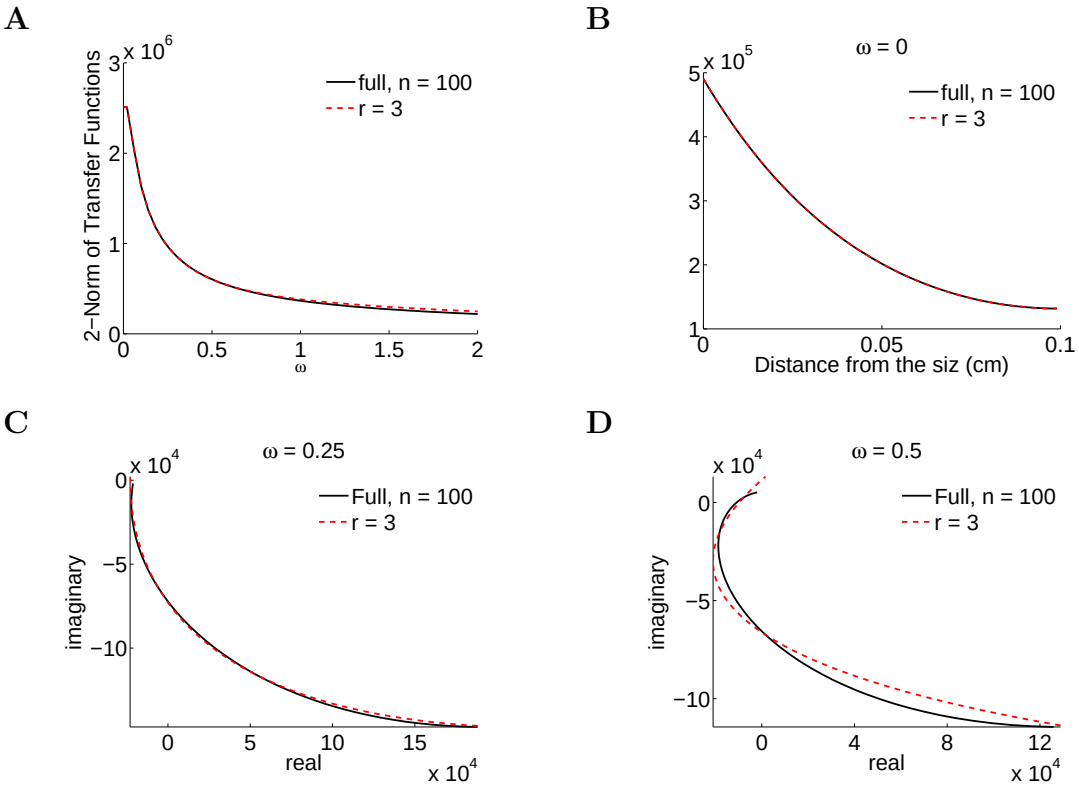


Figure 3.5: Agreement between the transfer functions for the full and reduced systems. **A.** 2-norm of the transfer functions evaluated at $i\omega$. Both transfer functions are low-pass filters symmetric about the real axis. **B-D.** Transfer functions evaluated at 0, $0.25i$, and $0.5i$, respectively. For $\omega = 0$, the transfer functions are real, and the distance from the siz is plotted on the x-axis. For $\omega = 0.25$ and 0.5, the imaginary vs. real components of the transfer functions are plotted.

3.1.7 Extension to the Dendritic Tree

Consider a dendritic tree with a general morphology in which the radius varies in space but C_m, g_L, and V_L are uniform. The formulation and structure of the reduced model extends easily to this more general case. Let $A_j = 2\pi a_j dx$ be the surface area of compartment j, and define the diagonal matrix D such that for each j,

$$D_{jj} = A_j/\overline{A}, \quad \text{where} \quad \overline{A} = (1/n)\sum_{j=1}^{n} A_j. \tag{3.17}$$

The capacitance and conductance matrices are slightly modified, given by

$$C = \overline{A}C_m D \quad \text{and} \quad G = \overline{A}g_L D - G_{ax}, \tag{3.18}$$

where G_{ax} encodes the dendritic morphology such that the j^{th} element of $G_{ax}v$ is again the axial current entering compartment j. Prop. 1 holds for any given G and C such that G is invertible, and for any $k \leq r$, the k^{th} basis vector for X becomes

$$X_{basis}(k) = (G^{-1}D)^{k-1}G^{-1}e_{siz}.$$

We construct X via the modified Arnoldi procedure detailed in Algorithm 2 such that $X^T DX = I_r$, leading to the reduced matrices

$$\hat{C} = \overline{A}C_m I_r \quad \text{and} \quad \hat{G} = \overline{A}g_L I_r - X^T G_{ax} X.$$

Using this procedure, each compartment j of the full neuron has a membrane capacitance $A_j C_m$ and membrane leakage conductance $A_j g_L$, while each compartment j of the reduced neuron has a membrane capacitance $\overline{A}C_m$ and membrane leakage conductance $\overline{A}g_L$. All other circuit elements are identical to those in Fig. 3.2. We provide the numerical example of reducing a CA1 pyramidal cell in Fig. 3.6.

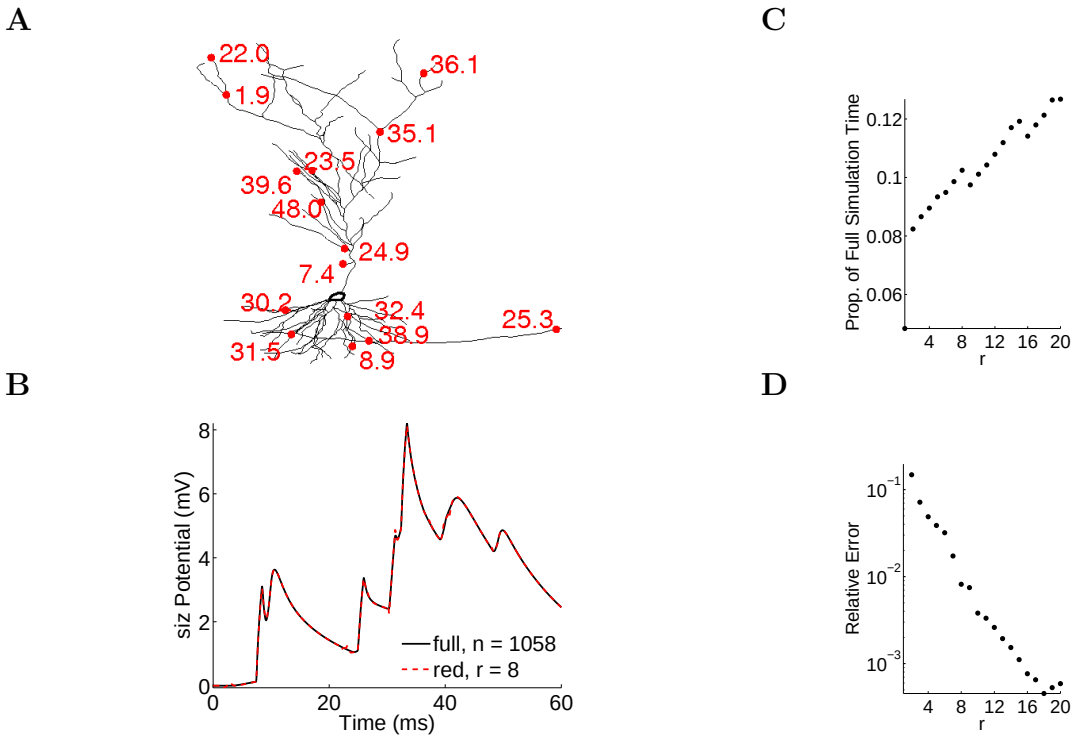


Figure 3.6: Reduction of a morphologically accurate pyramidal cell in the CA1 driven by fifteen square pulse current injections with conductances given by Eq. 3.10. **A.** Times and locations for synaptic inputs shown on a 2D projection of the dendritic tree. **B.** Potential at the siz (soma) for the full system of size $n = 1058$ (solid black) and reduced system of size $r = 8$ (dashed red). **C.** Proportion of the simulation time for the reduced system with r compartments to that of the full system. **D.** Relative 2-norm error in the siz potential.

3.2 Reduction of the Quasi-Active Neuron

The quasi-active model linearizes the active, or voltage-dependent, currents about rest, capturing the linear perturbation of the transmembrane potential from its resting state. Hodgkin and Huxley first introduced the quasi-active model in their seminal work (Hodgkin and Huxley [37]), and it has been widely used as it retains properties of the neuron such as resonance and stability but is more tractable than the active model.

We derive the quasi-active model in §2.1.4 using the instructive example of $I_{act} =$

I_h, the so-called h-current that allows both sodium and potassium ions to pass and activates upon hyperpolarization. In §3.2.1, we use a change of variables to present the model in terms of the underlying RLC circuit, as was done in Chapter 10 of (Koch [46]). Existing reduced models are not amenable to a circuit representation, but we adapt the reduction technique presented in (Li and Bai [52]) to synthesize the reduced circuit, presenting the reduced model in §3.2.2 and comparing the full and reduced circuits in §3.2.3.

The theory concerning matching moments of the transfer functions is divided into three propositions, presented in §3.2.4. Prop. 2 is the standard moment-matching proposition for the reduced quasi-active system derived from PRIMA. Prop. 3 establishes the accuracy of the reducer presented in (Li and Bai [52]), which preserves the block structure in the system's matrices. In Prop. 4 we simplify the Li and Bai reducer using the structure of the quasi-active system so that we can synthesize the reduced system as an RLC circuit.

We then extend the results by modeling multiple ionic currents within a neuron with a general dendritic morphology in §3.2.5. We go beyond what is typically presented in the literature to determine the accuracy of the reduced system as a function of the neuron's resonance and stability as well as the point about which the Taylor series of each transfer function is expanded.

3.2.1 The Discrete Quasi-Active Cable

Consider the h-current given by

$$I_h = 2\pi a g_h q^2(\phi - V_h), \quad \tau_q(\phi)\partial_t q = q_\infty(\phi) - q, \tag{3.19}$$

where g_h is the peak conductance (mS/cm^2) and $V_h = -40\ mV$ is the reversal potential. If the transmembrane potential ϕ were constant, the gating variable q would approach $q_\infty(\phi)$ with a rate determined by $\tau_q(\phi)$. The formulas and plots of both functionals can be found in Chapter 5 of (Gabbiani and Cox [20]). The potential is governed by Eq. 2.2 with $I_{act} = I_h$, and the resting potential $\overline{\phi}$ satisfies

$$g_L(\overline{\phi} - V_L) + g_h\overline{q}^2(\overline{\phi} - V_h) = 0, \quad \text{where} \quad \overline{q} = q_\infty(\overline{\phi}).$$

Given the conductances $g(x_j, t) = \varepsilon\widetilde{g}(x_j, t)$ for some small $\varepsilon > 0$, the potential and gating variable can be written as $\phi(t) = \overline{\phi} + \varepsilon\widetilde{\phi} + O(\varepsilon^2)$ and $q(t) = \overline{q} + \varepsilon\widetilde{q}(t) + O(\varepsilon^2)$. We write τ_q and q_∞ as Taylor expansions about $\overline{\phi}$ and expand Eqs. 2.2 and 3.19, matching terms of order ε to obtain the linear perturbations from rest,

$$(2\pi a)\{C_m\partial_t\widetilde{\phi} + (g_L + g_{q_1})\widetilde{\phi} + i\} - \frac{\pi a^2}{R_a}\partial_{xx}\widetilde{\phi} = -\sum_{j=1}^{m}\widetilde{g}(x_j, t)\delta(x - x_j)(\overline{\phi} - V(x_j)),$$

$$L_q\partial_t i - \widetilde{\phi} + (1/g_{q_2})i = 0, \widetilde{\phi}(x, 0) = i(x, 0) = 0, \quad \partial_x\widetilde{\phi}(0, t) = \partial_x\widetilde{\phi}(\ell, t) = 0. \tag{3.20}$$

The cable has two additional resistors and one additional inductor, given by

$$g_{q_1} \equiv g_h\overline{q}^2, \quad g_{q_2} \equiv 2g_h\overline{q}q'_\infty(\overline{\phi})(\overline{\phi} - V_h), \quad \text{and} \quad L_q \equiv \tau_q(\overline{\phi})/g_{q_2}, \tag{3.21}$$

where g_{q_1} and g_{q_2} have the units mS/cm^2, and L_q has the units $H \cdot cm$. The n-dimensional current i is defined

$$i \equiv (g_{q_2}/q'_\infty(\overline{\phi}))\widetilde{q} \quad (\mu A/cm^2). \tag{3.22}$$

We discretize in space with compartment length dx to model the cable using $n = \ell/dx$ compartments. Given the state vector of size $2n$, $z \equiv (i_1 \ \ldots \ i_n \ \widetilde{\phi}_1 \ \ldots \ \widetilde{\phi}_n)^T$, the siz potential is determined by

$$\mathcal{C}z' + \mathcal{G}z = \mathcal{B}u, \quad y = L^T z, \tag{3.23}$$

where

$$\mathcal{C} = \begin{pmatrix} AL_qI_n & 0 \\ 0 & C \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} (A/g_{q_2})I_n & -AI_n \\ AI_n & G + Ag_{q_1}I_n \end{pmatrix},$$
$$\mathcal{B} = \begin{pmatrix} 0 \\ I_n \end{pmatrix}, \quad \text{and} \quad L = \begin{pmatrix} 0 \\ e_{siz} \end{pmatrix}. \tag{3.24}$$

C and G are given by Eq. 3.3, and the compartments $\{p_j\}$ containing synapses give rise to the nonzero elements of the n-dimensional input vector u, given by

$$u_{p_j}(t) = \widetilde{g}_{p_j}(t)V_{p_j}, \quad j = 1, \ldots, m.$$

3.2.2 The Reduced Model

Analogous to the construction of the reduced model for the passive system, the reduced quasi-active model is constructed from a reducer $\mathcal{X}$ of size $2n \times 2r$, where the

siz potential approximation is determined by

$$\begin{gathered}\hat{\mathcal{C}}\hat{z}'(t) + \hat{\mathcal{G}}\hat{z}(t) = \hat{\mathcal{B}}u(t), \quad \hat{y}(t) = \hat{L}^T\hat{z}(t), \\ \hat{\mathcal{C}} = \mathcal{X}^T\mathcal{C}\mathcal{X},\ \hat{\mathcal{G}} = \mathcal{X}^T\mathcal{G}\mathcal{X},\ \hat{\mathcal{B}} = \mathcal{X}^T\mathcal{B}, \quad \text{and} \quad \hat{L} = \mathcal{X}^T L.\end{gathered} \tag{3.25}$$

For the general quasi-active system with g gating variables and n compartments, PRIMA specifies that the reducer be derived from $X_{basis} = \{(\mathcal{G}^{-1}\mathcal{C})^k\mathcal{G}^{-1}e_{siz}\}_{k=0}^{r-1}$ for some reduced system size r. While the resulting reduced system is accurate for small r (Yan and Li [97]), the reduction requires inverting an ill-conditioned matrix of size $(g+1)n \times (g+1)n$. We present a reduced system that not only retains the circuit structure but also requires inverting a matrix of size $n \times n$ similar to the conductance matrix given in Eq. 3.3.

Returning to the example of $I_{act} = I_h$, we first generate the $n \times r$ matrix X from the basis vectors

$$X_{basis} = \begin{pmatrix} G_{eff}^{-1}e_{siz} & G_{eff}^{-2}e_{siz} & \dots & G_{eff}^{-r}e_{siz} \end{pmatrix} \tag{3.26}$$

such that $X = \text{orth}(X_{basis})$, where

$$G_{eff} = A(g_L + g_{q_1} + g_{q_2})I_n - G_{ax}. \tag{3.27}$$

The $2n \times 2r$ reducer $\mathcal{X}$ is then given by

$$\mathcal{X} = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}. \tag{3.28}$$

X can be generated using Algorithm 1 with the input matrix G_{eff}. Since the h-current is activated upon hyperpolarization, we test the reduced model using inhibitory synap-

tic inputs and plot the results in Fig. 3.7. See §3.2.4 for the theory underlying Eqs. 3.26-3.28.

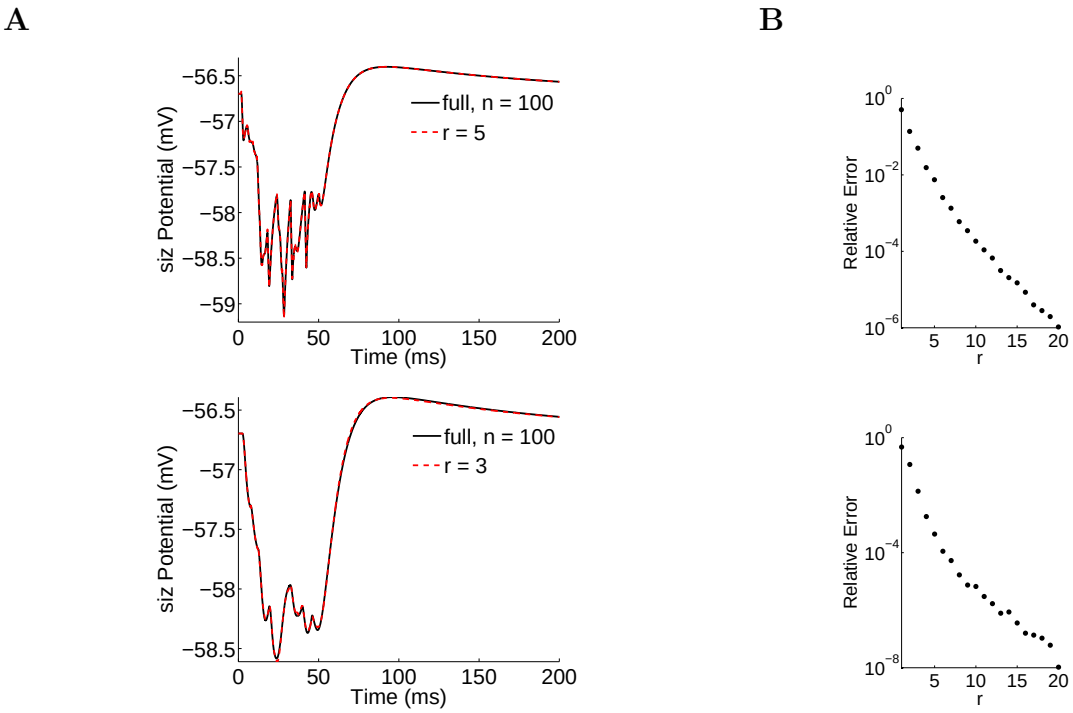


Figure 3.7: Testing of the reduced quasi-active cable for which $I_{act} = I_h$. The cable is driven by fifty inhibitory inputs with square pulse conductances (top) given by Eq. 3.10 or with alpha conductances (bottom) given by Eq. 3.11. The siz is located in the first compartment, $g_h = 2\ mS/cm^2$, and $V_{p_j} = -70\ mV$ for each synapse j. **A.** Response at the siz in a 100-compartment cell and the corresponding r-compartment reduced cell to identical synaptic conductances distributed randomly in space and time. Upon removal of the stimulus at 50 ms, the potential rises above rest before decaying to the resting potential, a characteristic of the quasi-active model not seen in the passive case. **B.** Relative 2-norm error in the siz potential given the reduced system with r compartments. The full system has dimension 200, and each reduced system has dimension $2r$. The reduced systems require 25% to 40% of the full simulation time.

3.2.3 RLC Circuit Structure

By the last n equations of Eq. 3.23, $AC_m\widetilde{\phi}' + A(g_L + g_{q_1})\widetilde{\phi} + Ai - G_{ax}\widetilde{\phi} = u$, implying that the quasi-active model is represented by the passive circuit for injected input

augmented by two currents to ground placed in parallel. The first current passes through a resistor with conductance Ag_{q_1}, and the second current is given by Ai. By the first n equations of Eq. 3.23, $L_q i' + i/g_{q_2} = \widetilde{\phi}$, implying that Ai describes the current flowing through an inductor and resistor in series with inductance L_q/A and conductance Ag_{q_2}.

Since $X^T X = I_r$, the reduced matrices are given by

$$\hat{\mathcal{C}} = \begin{pmatrix} AL_q I_r & 0 \\ 0 & AC_m I_r \end{pmatrix}, \ \hat{\mathcal{B}} = \begin{pmatrix} 0 \\ X^T \end{pmatrix}, \ \hat{L} = \begin{pmatrix} 0 \\ X^T e_{siz} \end{pmatrix},$$

$$\text{and } \hat{\mathcal{G}} = \begin{pmatrix} (A/g_{q_2}) I_r & -AI_r \\ AI_r & A(g_L + g_{q_1}) I_r + \text{diag}(\hat{g}_L) - \hat{G}_{ax} \end{pmatrix};$$

where $\hat{G}_{ax}$ and $\hat{g}_L$ are given by Eqs. 3.13 and 3.14, respectively. The reduced state vector of size $2r$ is given by $\hat{z} = (\hat{i}_1, \ \ldots, \ \hat{i}_r, \ \hat{\phi}_1, \ \ldots, \ \hat{\phi}_r)^T$, where for each reduced compartment k, $A\hat{i}_k$ is the current to ground passing through an inductor and resistor in series with inductance L_q/A and conductance Ag_{q_2}, and $\hat{\phi}_k$ is the linear perturbation from rest of the potential. Therefore, all currents flowing to ground in the full neuron are preserved in the reduced neuron, and both the axial conductances and the additional leakage conductances in the reduced neuron are identical to those found in the reduced passive neuron. The full and reduced circuits for the quasi-active model are compared in Fig. 3.8.

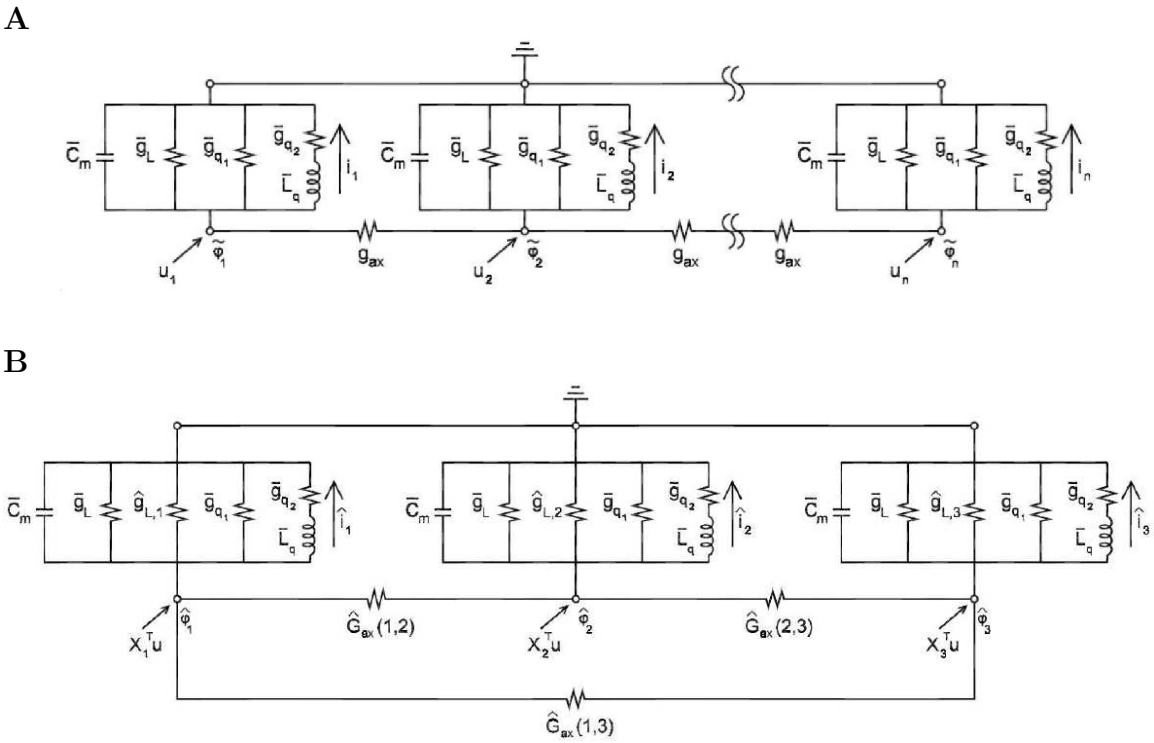


Figure 3.8: RLC circuit structure of the quasi-active cable. **A.** Full model with n compartments. The passive RC circuit in Fig. 3.2A is augmented by two parallel currents to ground for each compartment. The first flows through a resistor with conductance $\overline{g}_{q_1} = Ag_{q_1}$, and the second flows through an inductor with inductance $\overline{L}_q = L_q/A$ and a resistor with conductance $\overline{g}_{q_2} = Ag_{q_2}$, where g_{q_1}, g_{q_2}, and L_q are given by Eq. 3.21. **B.** Reduced model with three compartments. The reduced circuit preserves all currents to ground in **(A)**. The additional membrane leakage and axial conductances are identical to those found in the reduced passive neuron in Fig. 3.2B.

3.2.4 Theory

The theory underlying the reduced quasi-active model uses the block structure in the system's matrices to extend the theory presented in §3.1.6. The Laplace transforms of Eqs. 3.23 and 3.25 provide the linear mappings

$$\mathcal{L}y(s) = H(s)\mathcal{L}u(s) \quad \text{and} \quad \mathcal{L}\hat{y}(s) = \hat{H}(s)\mathcal{L}u(s)$$

in the frequency domain, where the transfer functions are given by

$$H(s) = L^T(\mathcal{G} + s\mathcal{C})^{-1}\mathcal{B} \quad \text{and} \quad \hat{H}(s) = \hat{L}^T(\hat{\mathcal{G}} + s\hat{\mathcal{C}})^{-1}\hat{\mathcal{B}}. \tag{3.29}$$

As in Eq. 3.16, each transfer function can be expanded as a Taylor series about $s = 0$ with the moments

$$M_j = L^T(\mathcal{G}^{-1}\mathcal{C})^j\mathcal{G}^{-1} \quad \text{and} \quad \hat{M}_j = \hat{L}^T\mathcal{X}(\hat{\mathcal{G}}^{-1}\hat{\mathcal{C}})^j\hat{\mathcal{G}}^{-1}\mathcal{X}^T.$$

The reducer is constructed to match as many of the leading moments as possible given the reduced system size. Prop. 2 is the quasi-active analogue of Prop. 1.

Prop. 2 If $K_r(\mathcal{G}^{-T}\mathcal{C}, \mathcal{G}^{-T}L) \subseteq \text{sp}(\mathcal{X})$, then $\hat{M}_j = M_j$ for $0 \leq j < r$.

Proof: The proof follows in an identical fashion to that of Prop. 1.

To establish the accuracy of the reduced system given by Eqs. 3.25-3.28, we first consider a reducer in Prop. 3 which is based on the structure-preserving reduction

technique presented in §4 of (Li and Bai [52]). We use this result in Prop. 4 to show the effectiveness of the reducer given by Eqs. 3.26-3.28.

Prop. 3 Let the $2n \times r$ matrix $\widetilde{W}$ be such that $K_r(\mathcal{G}^{-T}\mathcal{C}, \mathcal{G}^{-T}L) \subseteq \mathrm{sp}(\widetilde{W})$, and assume the $n \times r$ matrices X and Y satisfy

$$\widetilde{W} = \begin{pmatrix} \widetilde{X} \\ \\ \widetilde{Y} \end{pmatrix}, \ \mathrm{sp}(\widetilde{X}) \subseteq \mathrm{sp}(X), \quad \text{and} \quad \mathrm{sp}(\widetilde{Y}) \subseteq \mathrm{sp}(Y).$$

If the $2n \times 2r$ reducer $\mathcal{X}$ is given by

$$\mathcal{X} = \begin{pmatrix} X & 0 \\ \\ 0 & Y \end{pmatrix},$$

then $\hat{M}_j = M_j$ for $0 \leq j < r$. Furthermore, the matrix structure is retained in the sense that if

$$\mathcal{C} = \begin{pmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} \\ \\ \mathcal{C}_{21} & \mathcal{C}_{22} \end{pmatrix}, \quad \text{then} \quad \hat{\mathcal{C}} = \begin{pmatrix} \hat{\mathcal{C}}_{11} & \hat{\mathcal{C}}_{12} \\ \\ \hat{\mathcal{C}}_{21} & \hat{\mathcal{C}}_{22} \end{pmatrix},$$

where $\hat{C}_{jk} = X_j^T C_{jk} X_k$. The block structures of $\mathcal{G}$, $\mathcal{B}$, and L are similarly retained.

Proof: The proof follows from Prop. 2 and the fact that $\mathrm{sp}(\widetilde{W}) \subseteq \mathrm{sp}(\mathcal{X})$.

Prop. 4 Let G_{eff} be given by Eq. 3.27, and assume that $K_r(G_{eff}^{-1}, G_{eff}^{-1} e_{siz}) \subseteq \mathrm{sp}(X)$. If $\mathcal{X}$ is a $2n$ x $2r$ block diagonal matrix with X in each $n \times r$ diagonal sub-block, then $\hat{M}_j = M_j$ for $0 \leq j < r$.

Proof: Define the $2n$ x r matrix $\widetilde{W}$ such that for any $k \leq r$, both the first n elements and the last n elements in column k of $\widetilde{W}$ are given by $G_{eff}^{-k} e_{siz}$. By Eq. 3.24,

$$\mathcal{G}^{-1} = \begin{pmatrix} -g_{q_2}^2 G_{eff}^{-1} + (g_{q_2}/A) I_n & g_{q_2} G_{eff}^{-1} \\ \\ -g_{q_2} G_{eff}^{-1} & G_{eff}^{-1} \end{pmatrix}.$$

Thus,

$$\mathcal{G}^{-T} L = \begin{pmatrix} g_{q_2} G_{eff}^{-1} e_{siz} \\ \\ G_{eff}^{-1} e_{siz} \end{pmatrix},$$

and

$$\mathcal{G}^{-T} \mathcal{C} = \begin{pmatrix} g_{q_2} L_q (-A g_{q_2} G_{eff}^{-1} + I_n) & A g_{q_2} C_m G_{eff}^{-1} \\ \\ -A g_{q_2} L_q G_{eff}^{-1} & A C_m G_{eff}^{-1} \end{pmatrix}.$$

Therefore, $K_r(\mathcal{G}^{-T}\mathcal{C}, \mathcal{G}^{-T} L) \subseteq \text{sp}(\widetilde{W})$, and Prop. 3 holds if $K_r(G_{eff}^{-1}, G_{eff}^{-1} e_{siz}) \subseteq \text{sp}(X)$ and $Y = X$. ■

3.2.5 Extension to the General Quasi-Active System

Dendritic Tree Consider the dendritic tree for which $I_{act} = I_h$, the radius varies in space, and C_m, g_L, and g_h are uniform. Given D and $\overline{A}$ of Eq. 3.17, the capacitance and conductance matrices of the full system become

$$\mathcal{C} = \begin{pmatrix} \overline{A} L_q D & 0 \\ \\ 0 & C \end{pmatrix} \quad \text{and} \quad \mathcal{G} = \begin{pmatrix} (\overline{A}/g_{q_2}) D & -\overline{A} D \\ \\ \overline{A} D & G + \overline{A} g_{q_1} D \end{pmatrix},$$

respectively, where C and G are given by Eq. 3.18. $\mathcal{B}$ and L are given by Eq. 3.24. One can show that the assumptions of Prop. 3 are satisfied if $Y = X$ and $\text{sp}(X) =$

$\mathrm{sp}(X_{basis})$, where for any $k \leq r$,

$$X_{basis}(k) = (G_{eff}^{-1}D)^{k-1}G_{eff}^{-1}e_{siz} \quad \text{and} \quad G_{eff} = \overline{A}(g_L + g_{q_1} + g_{q_2})D - G_{ax}.$$

Furthermore, if X is orthogonal in the weighted sense that $X^T DX = I_r$, then the reduced system preserves the RLC circuit structure. The $n \times r$ matrix X can be generated from Algorithm 2 using the input matrices G_{eff} and D.

Multiple Channel Types The reduction procedure generalizes naturally for any active channel types. To demonstrate the generality of the method as well as its robustness to resonance and instability, we test the reduced quasi-active neuron for the dendritic tree with the classic Hodgkin and Huxley sodium and potassium currents as well as the h-current. For this system, $I_{act} = I_K + I_{Na} + I_h$, where I_h is given by Eq. 3.19;

$$I_K = g_K b^4(\phi - V_K),\ g_K = 36, \quad \text{and} \quad V_K = -77; \quad \text{and}$$

$$I_{Na} = g_{Na} m^3 h(\phi - V_{Na}),\ g_{Na} = 120, \quad \text{and} \quad V_{Na} = 56;$$

with the units mS/cm^2 and mV. Let $w \in \{b, m, h, q\}$ denote any gating variable, governed by $\tau_w(\phi)\partial_t w = w_\infty(\phi) - w$. The formulas and plots for all functionals can be found in Ch. 4 of (Gabbiani and Cox [20]). As was done in §3.2.1, the potential and each gating variable can respectively be written as $\phi = \overline{\phi} + \varepsilon\widetilde{\phi} + O(\varepsilon^2)$ and $w = \overline{w} + \varepsilon\widetilde{w} + O(\varepsilon^2)$, and the linear perturbations from rest are determined by matching terms of order ε. With respect to the passive system, each gating variable adds two resistors and one inductor, where g_{q_1}, g_{q_2}, and L_q are given by Eq. 3.21,

and the expressions for g_{w_1}, g_{w_2}, and L_w are derived analogously for any w. As in Eq. 3.22, each gating variable gives rise to the transmembrane current given by $i_w = (g_{w_2}/w'_\infty(\overline{\phi}))\widetilde{w}$. The state vector has size $5n$ and is governed by Eq. 3.23, where

$$z = \begin{pmatrix} i_b \\ i_m \\ i_h \\ i_q \\ \widetilde{\phi} \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} \frac{\overline{A}}{g_{b_2}} D & & & & -\overline{A}D \\ & \frac{\overline{A}}{g_{m_2}} D & & & -\overline{A}D \\ & & \frac{\overline{A}}{g_{h_2}} D & & -\overline{A}D \\ & & & \frac{\overline{A}}{g_{q_2}} D & -\overline{A}D \\ \overline{A}D & \overline{A}D & \overline{A}D & \overline{A}D & \mathcal{G}_{55} \end{pmatrix},$$

$\mathcal{G}_{55} = \overline{A}(g_L + \sum_w g_{w_1})D - G_{ax}$, and

$$\mathcal{C} = \begin{pmatrix} \overline{A}L_b D & & & & \\ & \overline{A}L_m D & & & \\ & & \overline{A}L_h D & & \\ & & & \overline{A}L_q D & \\ & & & & \overline{A}C_m D \end{pmatrix}.$$

As in Eq. 3.24, only the last n rows of $\mathcal{B}$ and L are nonzero, given by I_n and e_{siz}, respectively.

The reduced matrices are given by Eq. 3.25, where $\mathcal{X}$ is a $5n \times 5r$ block diagonal matrix with X on each $n \times r$ diagonal sub-block, and X satisfies $\mathrm{sp}(X) = \mathrm{sp}(X_{basis})$ and $X^T D X = I_r$. For any $k \leq r$,

$$X_{basis}(k) = (G_{eff}^{-1} D)^{k-1} G_{eff}^{-1} e_{siz} \quad \text{and} \quad G_{eff} = \overline{A}(g_L + \sum_w (g_{w_1} + g_{w_2}))D - G_{ax}.$$

A numerical example is provided in Fig. 3.9.

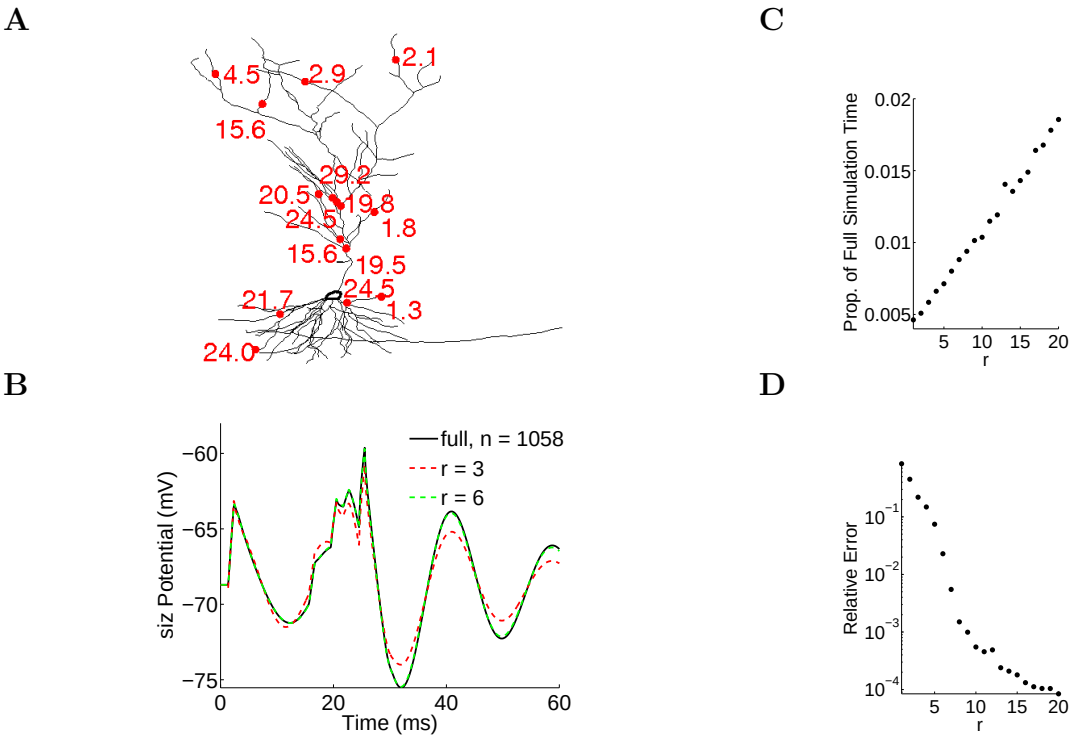


Figure 3.9: Quasi-active Hodgkin and Huxley tree augmented by the h-current, where $g_h = 0.3\ mS/cm^2$. **A.** Onset times and locations for synaptic inputs shown on a 2D projection of a morphologically accurate pyramidal cell in the CA1. **B.** Response at the siz to square pulse conductances given by Eq. 3.10 presented between 0 and 30 ms according to **(A)**. Since there are five variables for each compartment, the full system (solid black) has dimension 5290, while the reduced systems have dimensions 15 (dashed red) and 30 (dashed green). **C.** Proportion of the simulation time for the reduced system with r compartments to that of the full system. **D.** Relative 2-norm error in the siz potential.

Theory: Resonance and Instability Unlike the passive system, the quasi-active system can become unstable under certain conditions, requiring the modification in the error bound (Eq. 3.15)

$$|y(t) - \hat{y}(t)|^2 \leq c(u) \int_{-\infty}^{\infty} \|H(\omega_0 + i\omega) - \hat{H}(\omega_0 + i\omega)\|^2 \, d\omega, \tag{3.30}$$

where c depends only on the input u, and ω_0 is greater than the real component of all poles. Regardless of stability, the transfer function is a band-pass filter that peaks along the imaginary axis near $\pm i\omega_{res}$, where ω_{res} is the cell's resonant frequency

(rad/ms) and is generally nonzero. Since the reduced model is constructed such that the transfer functions agree near s_0, the center of their Taylor expansions, both resonance and instability cast doubt on the effectiveness of setting $s_0 = 0$.

Consider the quasi-active cable with a uniform radius for which $I_{act} = I_K + I_{Na} + I_h$. In Fig. 3.10 we show the agreement between the transfer functions along the line of integration in Eq. 3.30 when $s_0 = 0$. As g_h increases, the resonant frequency increases, causing the transfer function's peak to shift away from the origin and become sharper. For each value of g_h tested, however, ω_{res} is sufficiently small for the reduced transfer function to capture this peak given a relatively small reduced dimension. In general, if the temporal resonant frequency f_{res} is less than 150 Hz, then $\omega_{res} = 2\pi f_{res}/1000$ is less than one, implying that $\hat{H}(i\omega_{res}) \approx H(i\omega_{res})$ for sufficiently large r since $\hat{H}(i\omega) = H(i\omega) + O(\omega^r)$. The convergence slows as ω_{res} increases, as illustrated in Fig. 3.10D.

We next examine the reduced output itself for a range of s_0 values. As shown in (Li and Bai [52]), $H(s) = L^T(\mathcal{G} + s\mathcal{C})^{-1}\mathcal{B} = L^T(\mathcal{G}(s_0) + \widetilde{s}\mathcal{C})^{-1}\mathcal{B}$, where $\mathcal{G}(s_0) \equiv \mathcal{G} + s_0\mathcal{C}$ and $\widetilde{s} \equiv s - s_0$. Therefore, the method used to generate the reducer $\mathcal{X}$ for $s_0 = 0$ can still be applied for a nonzero s_0 if $\mathcal{G}$ is replaced by $\mathcal{G}(s_0)$, s is replaced by $\widetilde{s}$, and $H(\widetilde{s})$ is expanded about $\widetilde{s} = 0$. The reduced system is then equivalent to that derived for $s_0 = 0$ with G_{eff} replaced by

$$G_{eff}(s_0) \equiv Ag_L + s_0AC_m + A\sum_w g_{w_1}I_n + A\sum_w g_{w_2}/(1 + s_0\tau_w(\overline{\phi}))I_n - G_{ax},$$

the summations being taken over all gating variables. If s_0 is nonreal, then the

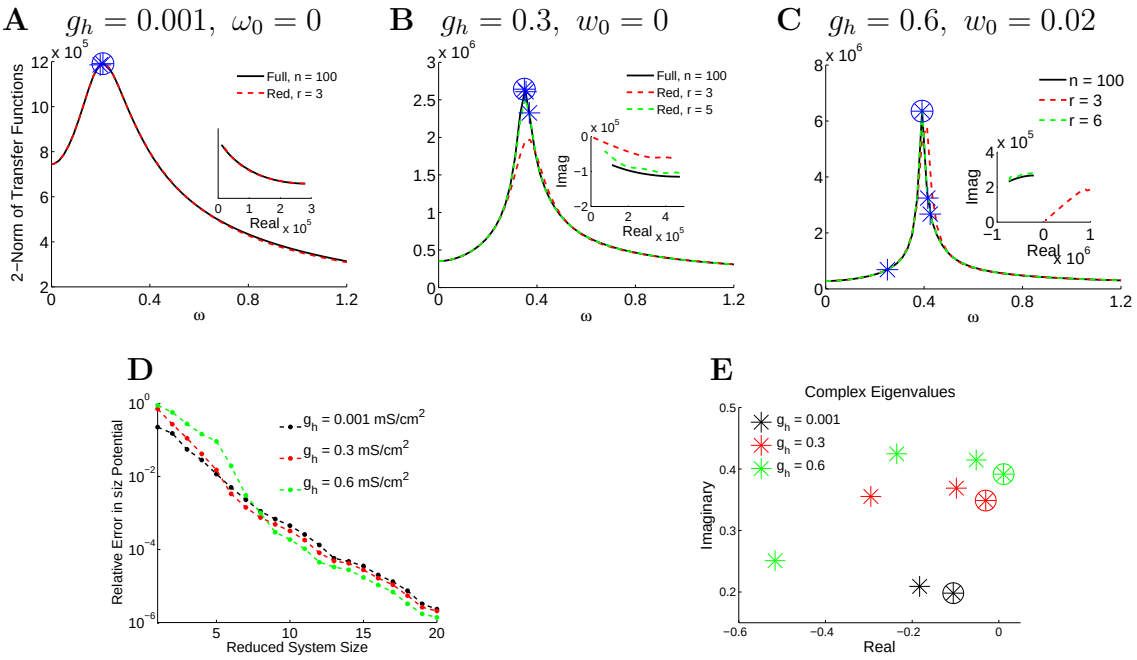


Figure 3.10: Effectiveness of setting $s_0 = 0$ in reducing the quasi-active cable for which $I_{act} = I_{Na} + I_K + I_h$. The potential oscillations grow as g_h increases, as shown in Fig. 3.11, and the system is unstable when $g_h = 0.6\ mS/cm^2$. **A-C.** 2-norm of the transfer functions evaluated at $\omega_0 + i\omega$ for the full (solid black) and reduced (dashed) systems, where w_0 is greater than the real component of all poles. All plots are symmetric about the real axis. Stars indicate the imaginary component of each pole, and the inset shows the transfer functions evaluated at $\omega_0 + i\omega_{res}$ (circled). **D.** Relative 2-norm error in the siz potential given ten inputs with square-pulse conductances given by Eq. 3.10. **E.** Complex eigenvalues of $-\mathcal{C}^{-1}\mathcal{G}$, which are poles of the transfer functions. The spectrum contains the complex conjugate of each eigenvalue. The resonant frequency is given by the imaginary component of the complex eigenvalue with largest real part, which is circled.

reduced system is nonreal, creating nonreal potentials. We thus confine s_0 to the real axis. Fig. 3.11 illustrates that although $s_0 = 0$ is not always the optimal choice, it is effective, even if the original system is unstable.

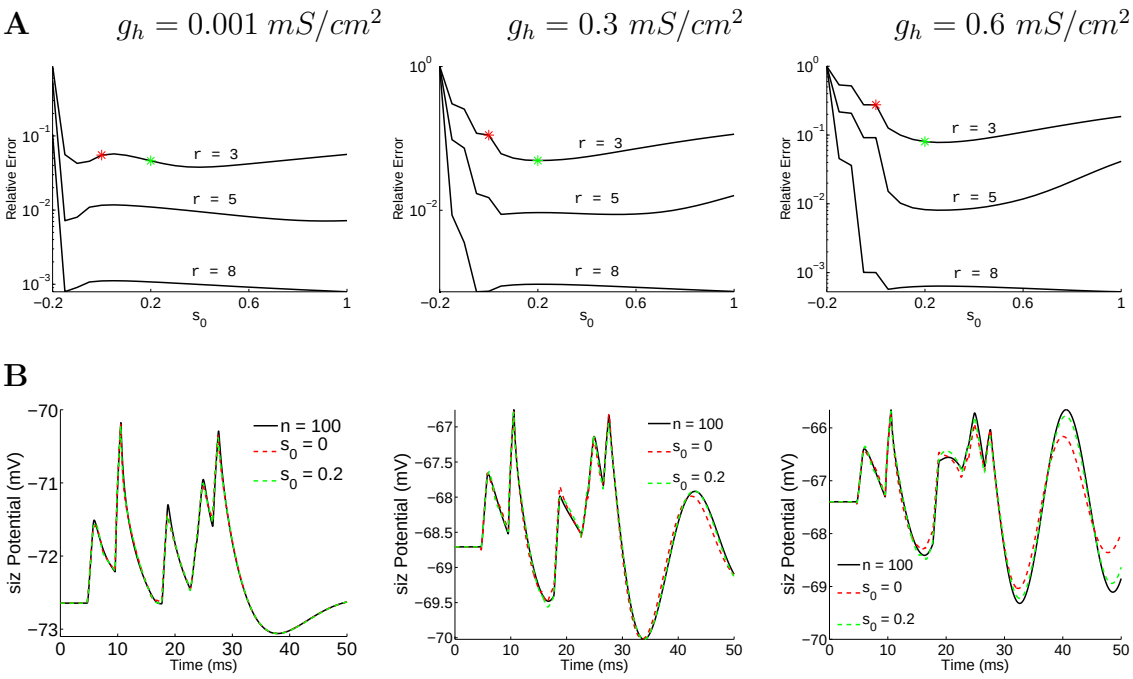


Figure 3.11: Accuracy of the reduced quasi-active model as s_0 varies along the real axis. The cable and its synaptic inputs are identical to those used in Fig. 3.10. **A.** Relative 2-norm error in the siz potential as a function of s_0 for three reduced system sizes. **B.** Potential at the siz for the full system (solid black) and reduced systems (dashed) for $r = 3$, where $s_0 = 0$ (red) and $s_0 = 0.2$ (green), as indicated by the stars in **(A)**. Each value of s_0 included in **(A)** gives rise to an accurate reduced system, but the optimal region seems to be between 0 and 0.3.

3.3 Synaptic Drive of the Passive Neuron

The siz potential for the passive cable given synaptic input, described by Eq. 2.3, has a nonlinear dependence on the synaptic conductances as the input current depends on the products $g(x_j, t)v(x_j, t)$. This nonlinearity causes an effect known as sublinear

summation in which the potential given two inputs is less than the summation of the potentials given each input individually.

In §3.3.1 we present a reduced model that retains the RC circuit structure, is independent of the number and locations of synaptic inputs, and approximates the siz potential, capturing much of the sublinear summation effect. In §3.3.2 we expand Eq. 2.3 as a series of linear ODEs known as the Volterra series. We evaluate the convergence rate of the series for our application, which has not been done in the literature, and we derive the transfer functions and corresponding moments for each term. We close in §3.3.3 by numerically and analytically comparing our reduced system to that derived from PRIMA for which the potential is captured at the synapses, presented in (Yan and Li [97]). Much of the theory and simulations presented in this section are in the context of monosynaptic input; however, the results can be generalized for polysynaptic input with additional notation.

3.3.1 Methods and Numerical Results

Upon spatial discretization of Eq. 2.3 with compartment length dx, we arrive at the n-dimensional model determining the siz potential driven by m synaptic conductances,

$$Cv'(t) + Gv(t) = u(t) - \sum_{j=1}^{m} g_{p_j}(t) N^{(p_j)} v(t), \quad y(t) = e_{siz}^T v(t), \tag{3.31}$$

where for each synapse j at compartment p_j, the $n \times n$ matrix $N^{(p_j)}$ has the single nonzero element $N^{(p_j)}(p_j, p_j) = 1$. All other terms are defined in §3.1.2. The model is identical to that derived for injected input (Eq. 3.2) with one additional term

containing a time-varying coefficient of the membrane potential.

The reduced model of dimension r is again constructed from a reducer X of size $n \times r$, where the siz potential approximation is determined by

$$\hat{C}\hat{v}'(t) + \hat{G}\hat{v}(t) = X^T u(t) - \sum_{j=1}^{m} g_{p_j}(t)\hat{N}^{(p_j)}\hat{v}(t), \quad \hat{y}(t) = e_{siz}^T X\hat{v}(t). \tag{3.32}$$

$\hat{N}^{(p_j)} = X^T N^{(p_j)} X$, and $\hat{C}$ and $\hat{G}$ are given by Eq. 3.7. A reduced model of Eq. 3.32 was recently presented in which the reducer X is constructed from the basis vectors, $X_{basis} = \{G^{-k}B\}_{k=0}^{r-1}$ for a given reduced system size r, where $B = [e_{p_1}, \ \ldots, \ e_{p_m}]$ (Yan and Li [97]). This reduced system captures the potential at all synapses, but the reducer and thus the reduced system depend on synaptic locations and grow with the number of synapses. We take a complementary approach by constructing the reducer from the basis vectors derived for injected input, given by Eq. 3.8. The resulting reduced system is independent of synaptic inputs and captures much of the sublinear summation, despite being derived from a reducer incorporating only the linear components of the system.

In Fig. 3.12 we provide two numerical examples to illustrate the effectiveness of the reduced model. For these examples each compartment contains a synapse, and thus $p_j = j$ and $\sum g_j \hat{N}^{(j)} = X^T \text{diag}(g) X$, the computation of which can annihilate the gain in efficiency of the reduced system. For certain conductances, this product can be updated at each timestep to avoid the computational expense. For example, given the square pulse conductances of Eq. 3.10, $\overline{g} X(j,:)^T X(j,:)$ is added to the product if t_j is in the current timestep and subtracted from the product if $t_j + 1$ is in the current

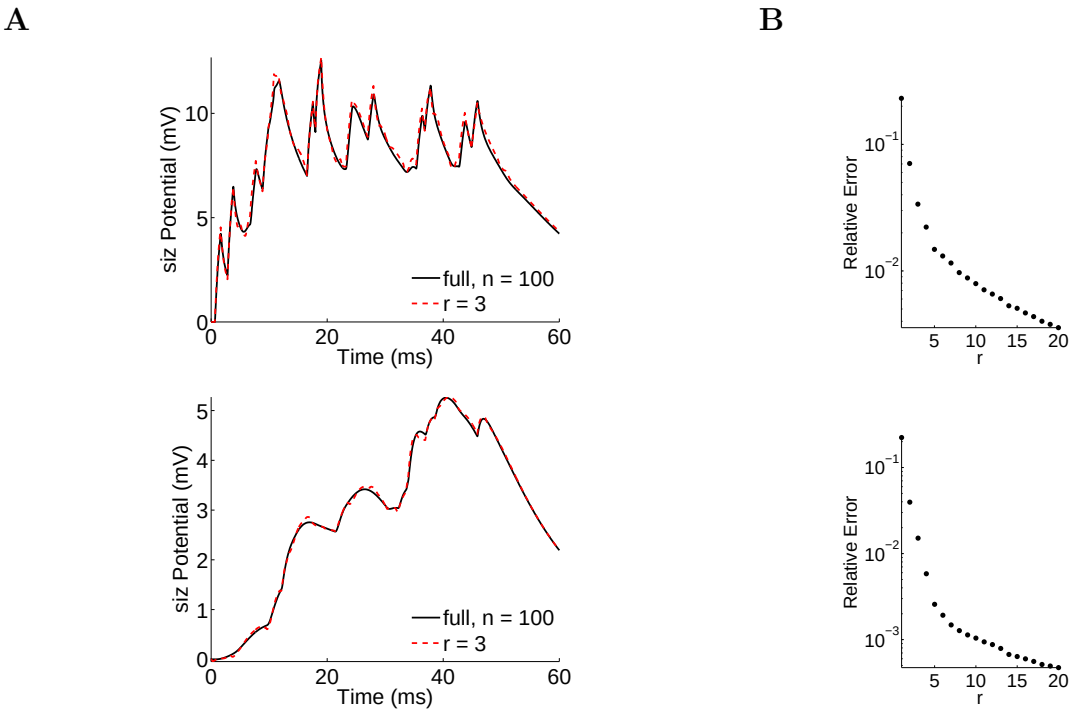


Figure 3.12: Testing of the reduced passive cable driven by synaptic inputs with square pulse conductances (top) given by Eq. 3.10 or with transient conductances (bottom) given by Eq. 3.33. **A.** Response at the siz of a 100-compartment cell (solid black) and the corresponding 3-compartment reduced cell (dashed red) to fifty identical synaptic currents distributed randomly in space and time. For the top plot, the synaptic conductances are identical to those used to generate Fig. 3.1A (top), and the sublinear summation is seen in the reduction of the siz potential. **B.** Relative 2-norm error in the siz potential given the reduced system with r compartments. The reduced systems require 60% to 75% of the full simulation time.

timestep, where $X(j,:)$ denotes row j of X. A similar update rule can be applied for transient conductances with an instantaneous rise time, given by

$$g_j(t) = \overline{g}\exp((t-t_j)/\tau)\chi_{[t_j,\infty)}. \tag{3.33}$$

The circuit representation of the full system given by Eq. 3.31 is that of Fig. 3.2A augmented by a time-varying membrane conductance, g_{p_j}, at each compartment containing a synapse. Analogous to Eq. 3.12, we uncover the circuit structure of the

reduced system by forming the decomposition

$$\sum_{j=1}^{m} g_{p_j}(t)\hat{N}^{(p_j)} = -\hat{W}_{ax} + \mathrm{diag}(\hat{g}), \tag{3.34}$$

implying that for any $i, k \leq r$,

$$\hat{W}_{ax}(i,k) = X_i^T \sum_{j=1}^{m} g_{p_j}(t) N^{(p_j)} \begin{cases} -X_k. & \text{if } i \neq k, \\ \sum_{q \neq i} X_q & \text{if } i = k. \end{cases} \tag{3.35}$$

By the diagonal of Eq. 3.34, for each $i \leq r$,

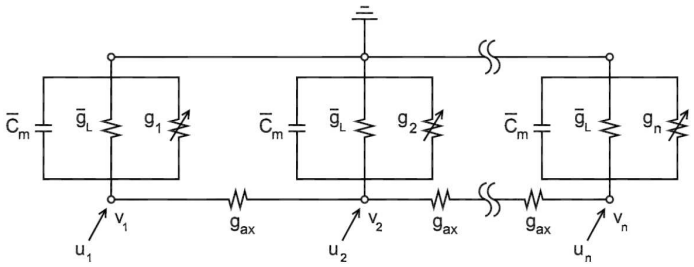
$$\hat{g}_i(t) = X_i^T \sum_{j=1}^{m} g_{p_j}(t) N^{(p_j)} \sum_{k=1}^{r} X_k. \tag{3.36}$$

Therefore, a synapse is located at each compartment i of the reduced neuron with a synaptic conductance $\hat{g}_i$ dependent on the original synaptic conductances $\{g_{p_j}\}$. An additional time-varying resistor also forms between any two compartments i and k of the reduced neuron with a conductance given by $\hat{W}_{ax}(i,k)$. The full and reduced RC circuits are compared in Fig. 3.13.

3.3.2 Volterra Series Expansion for Monosynaptic Input

The theory of §3.1.6 no longer applies as the time-varying coefficients render the transfer functions more difficult to compute. Following the lead of (Phillips [67]), we expand Eq. 2.3 as a series of linear ODEs known as the Volterra series. The theory is presented in the context of monosynaptic input for simplicity of notation, but it can be generalized for polysynaptic input.

A



B

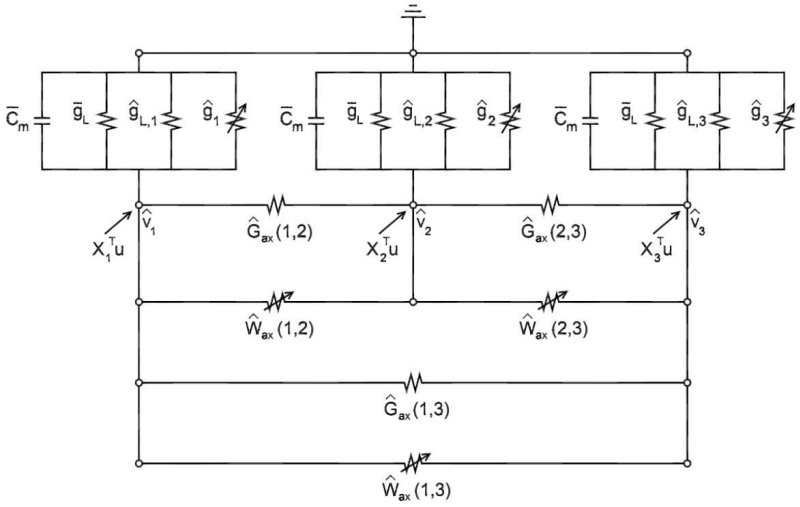


Figure 3.13: RC circuit structure of the passive cable for synaptic input. **A.** Schematic of the full system with n compartments. Given the circuit for injected input of Fig. 3.2A, a resistor is added across the membrane of each compartment p_j containing a synapse with the synaptic conductance $g_{p_j}(t)$. **B.** Schematic of the reduced system with three compartments. Each reduced compartment i has the synaptic conductance $\hat{g}_i(t)$, given by Eq. 3.36, and an additional leakage conductance $\hat{g}_{L,i}$, given by Eq. 3.14. Axial resistors connect any two compartments according to $\hat{G}_{ax}$ and $\hat{W}_{ax}(t)$, given by Eqs. 3.13 and 3.35, respectively.

Assume the cable is driven by a single synaptic input at location x_{syn} with conductance $g(t)$ and reversal potential E. Let $v^1(x,t)$ be the solution of Eq. 3.1, where the synaptic input of Eq. 2.3 has been approximated by the injected input, $g(t)\delta(x-x_{syn})E$. The nonlinearities are incorporated through the iterative sequence $\{v^k\}$, where for any $k \geq 2$,

$$(2\pi a C_m)\partial_t v^k(x,t) + (2\pi a g_L)v^k(x,t) - \frac{\pi a^2}{R_a}\partial_{xx}v^k(x,t) = -g(t)\delta(x-x_{syn})v^{k-1}(x,t)$$
$$v^k(x,0) = 0, \quad \partial_x v^k(0,t) = \partial_x v^k(\ell,t) = 0. \tag{3.37}$$

As shown in Appendix A, Eq. 2.3 is solved by the infinite series, $v = \sum v^k$. Upon spatial discretization,

$$C(v^1)'(t) + Gv^1(t) = g(t)Ee_p \quad \text{and} \quad C(v^k)'(t) + Gv^k(t) = -g(t)Nv^{k-1}(t)$$

for $k \geq 2$, where compartment p contains the synapse, e_p is a unit vector such that $e_p(p) = 1$, and C, G, and N are the matrices in Eq. 3.31. The reduced system can be similarly expanded by the infinite series $\hat{v} = \sum \hat{v}^k$, governed by

$$\hat{C}(\hat{v}^1)'(t) + \hat{G}\hat{v}^1(t) = g(t)X^T E e_p \quad \text{and} \quad \hat{C}(\hat{v}^k)'(t) + \hat{G}\hat{v}^k(t) = -g(t)\hat{N}\hat{v}^{k-1}(t),$$

for $k \geq 2$, where $\hat{C}$, $\hat{G}$, and $\hat{N}$ are the matrices in Eq. 3.32. Define $y^k \equiv e_{siz}^T v^k$ and $\hat{y}^k \equiv e_{siz}^T X \hat{v}^k$. Then, the siz potentials for the full and reduced systems are respectively given by $y = \sum y^k$ and $\hat{y} = \sum \hat{y}^k$.

Convergence Rate We next argue that $v \approx v^1 + v^2$ given fast, small synaptic conductances. The theory is developed for the infinite-length cable for which the

boundary conditions of Eqs. 3.1 and 3.37 are replaced by $v^k(-\infty, t) = v^k(\infty, t) = 0$. The results obtained under this simplification can be applied to the finite cable, as we show numerically. Without loss of generality, assume $x_{syn} = 0$.

Upon dividing Eqs. 3.1 and 3.37 by $2\pi a g_L$, the governing equation for each Volterra term becomes

$$\tau \partial_t v^1 + v^1 - \lambda^2 \partial_{xx} v^1 = c(t)\delta(x)E \quad \text{and} \quad \tau \partial_t v^k + v^k - \lambda^2 \partial_{xx} v^k = -c(t)\delta(x)v^{k-1}$$

for $k \geq 2$. The familiar space and time constants are given by

$$\lambda^2 = \frac{a}{2R_a g_L} \ \ (cm) \quad \text{and} \quad \tau = \frac{C_m}{g_L} \ \ (ms), \tag{3.38}$$

respectively, and the normalized synaptic conductance is given by

$$c(t) = \frac{g(t)}{2\pi a g_L} \ \ (cm).$$

The heat equation governs $z^1 \equiv v^1 \exp(t/\tau)$ via

$$(\lambda^2/\tau)\partial_{xx} z^1 = \partial_t z^1 - c(t)\exp(t/\tau)\delta(x)E/\tau.$$

Thus,

$$z^1(x,t) = \frac{E}{\tau}\int_0^t \int_{-\infty}^{\infty} K(x-y, t-s)c(s)\mathrm{e}^{s/\tau}\delta(y)\, dy\, ds,$$

given the Green's function

$$K(x,t) = \frac{\mathrm{e}^{-(x/\lambda)^2/(4t/\tau)}}{\sqrt{4\pi\lambda^2 t/\tau}}.$$

At the synapse $x = 0$, and

$$z^1(0,t) = (E/\tau)\int_0^t \frac{c(s)\mathrm{e}^{s/\tau}}{\sqrt{4\pi\lambda^2(t-s)/\tau}}\, ds.$$

By the change of variables $s = t - u^2$,

$$z^1(0,t) = \frac{E}{\lambda\sqrt{\tau\pi}} \int_0^{\sqrt{t}} c(t-u^2)\mathrm{e}^{(t-u^2)/\tau}\, du.$$

Therefore, the synaptic potential of the first Volterra term is given by

$$v^1(0,t) = \mathrm{e}^{-t/\tau} z^1(0,t) = \frac{E}{\lambda\sqrt{\tau\pi}} \int_0^{\sqrt{t}} c(t-u^2)\mathrm{e}^{-u^2/\tau}\, du. \tag{3.39}$$

One can similarly show that for any $k \geq 2$,

$$v^k(0,t) = -\frac{1}{\lambda\sqrt{\tau\pi}} \int_0^{\sqrt{t}} c(t-u^2)\mathrm{e}^{-u^2/\tau} v^{k-1}(0,t-u^2)\, du. \tag{3.40}$$

Prop. 5 establishes the contribution of each Volterra term to the potential at the synapse. Since the cable is passive and has a single source at the synapse, the synaptic potential is the upper bound for the potential throughout the cable, or $|v^k(x,t)| \leq |v^k(0,t)|$ for any x, k, and t.

Prop. 5 Let $c(t) \equiv \bar{c}\,\chi_{[0,w]}(t)$ be a normalized square-pulse conductance of amplitude $\bar{c}$ and width w. Each Volterra term obeys

$$\max_t |v^k(0,t)| \leq \gamma^{k-1} \max_t |v^1(0,t)|, \quad \text{where} \quad \gamma = \frac{\bar{c}\sqrt{w/\tau}}{\lambda\sqrt{\pi}}.$$

Proof: By Eq. 3.39, $v^1(0,t)$ is proportional to

$$\int_0^{\sqrt{t}} c(t-u^2)\mathrm{e}^{-u^2/\tau}\, du = \begin{cases} \bar{c}\displaystyle\int_0^{\sqrt{t}} \mathrm{e}^{-u^2/\tau}\, du & \text{if } t \leq w, \\ \bar{c}\displaystyle\int_{\sqrt{t-w}}^{\sqrt{t}} \mathrm{e}^{-u^2/\tau}\, du & \text{if } t > w. \end{cases}$$

Thus, $v^1(0,t)$ attains its unique maximum at $t = w$. As all potentials decay exponentially to zero upon removal of the stimulus, each $|v^k(0,t)|$ attains its maximum when $t \leq w$. By Eq. 3.40, if $k \geq 2$ and $t \leq w$,

$$v^k(0,t) = \frac{-\bar{c}}{\lambda\sqrt{\tau\pi}} \int_0^{\sqrt{t}} \mathrm{e}^{-u^2/\tau} v^{k-1}(0, t-u^2)\, du,$$

which has the time derivative

$$\partial_t v^k(0,t) = \frac{-\bar{c}}{\lambda\sqrt{\tau\pi}} \int_0^{\sqrt{t}} \mathrm{e}^{-u^2/\tau} \partial_t v^{k-1}(0, t-u^2)\, du.$$

As $v^1(0,t) > 0$ and $\partial_t v^1(0,t) > 0$ for $t < w$, $v^2(0,t) < 0$ and $\partial_t v^2(0,t) < 0$ for $t < w$. This pattern continues for all k, and thus $|v^k(0,t)|$ attains its unique maximum when $t = w$ for any k, leading to the error bound

$$\begin{aligned}
\max_t |v^k(0,t)| &= \frac{\bar{c}}{\lambda\sqrt{\pi\tau}} \int_0^{\sqrt{w}} \mathrm{e}^{-u^2/\tau} |v^{k-1}(0, w-u^2)|\, du \\
&\leq \frac{\bar{c}}{\lambda\sqrt{\pi}} |v^{k-1}(0,w)| \int_0^{\sqrt{w/\tau}} \mathrm{e}^{-s^2}\, ds \\
&\leq \frac{\bar{c}\sqrt{w/\tau}}{\lambda\sqrt{\pi}} |v^{k-1}(0,w)| \\
&= \gamma \max_t |v^{k-1}(0,t)|,
\end{aligned}$$

given the change of variables $s = u/\sqrt{\tau}$. Therefore, $\max_t |v^k(0,t)| \leq \gamma^{k-1} \max_t |v^1(0,t)|$ for any $k \geq 1$. ■

Within the physiological range, $\gamma \ll 1$. For example, if $\overline{g} = 1\ nS$, $w = 2\ ms$, and all other parameters are given by Eq. 3.9, then $\gamma \approx 0.098$. Therefore, only the first few terms in the series make noticeable contributions to the synaptic potential and

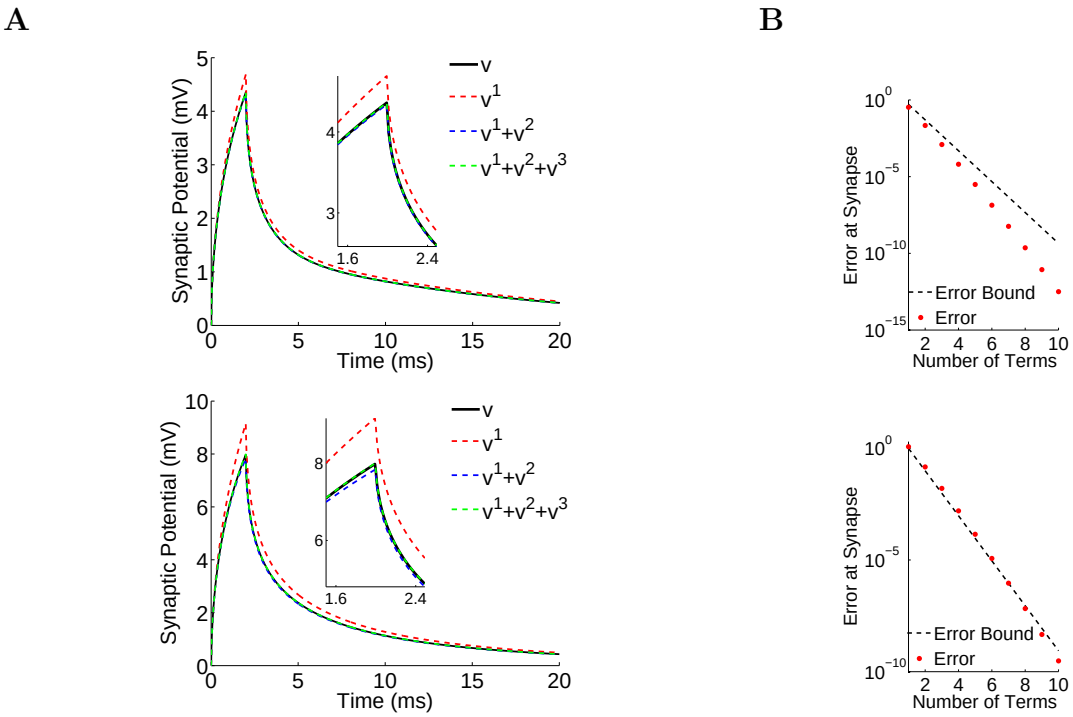


Figure 3.14: Convergence of the Volterra series. **A.** Synaptic potential for a 100-compartment cell driven by the single synaptic conductance, $g(t) = \overline{g}\,\chi_{[0,w]}(t)$, at the cable's midpoint (top) or endpoint (bottom). The potential at the synapse, $v(0,t)$ (solid black), is compared to the partial summation, $\sum_{k=1}^{\overline{k}} v^k(0,t)$ (dashed). All parameters are given by Eq. 3.9, and $w = 2$ ms. The potentials plotted between 1.6 and 2.4 ms are magnified in the inset. **B.** Error in the synaptic potential, given by $\max_t |v(0,t) - \sum_{k=1}^{\overline{k}} v^k(0,t)|$, where $\overline{k}$ is the number of terms included. The dashed line shows the error bound, given by Eq. 3.41 and derived for the infinite-length cable.

thus to the potential throughout the cable. Explicitly,

$$\max_t \left| v(0,t) - \sum_{k=1}^{\overline{k}} v^k(0,t) \right| \leq \frac{\gamma^{\overline{k}}}{1-\gamma} \max_t |v^1(0,t)|. \tag{3.41}$$

In Fig. 3.14 we illustrate the convergence of the series evaluated at the synapse when the synapse is at the cable's midpoint, in which case Eq. 3.41 applies, and at an endpoint, in which case the boundary conditions can no longer be ignored.

Moment-Matching Since each Volterra term is governed by a linear system, its transfer function provides a linear mapping from the input, g or gv^{k-1}, to the output, y^k, in the frequency domain. As was done in §3.1.6, one can use the inverse Laplace transform to relate the error, $\max_t |y^k(t) - \hat{y}^k(t)|$, to the error in the transfer functions evaluated along the imaginary axis. As for the injected input case, each transfer function has only real, negative poles and is a low-pass filter. We thus write each transfer function as a Taylor expansion about the origin and analyze the accuracy of the reduced system by its ability to match the leading moments.

We derive the transfer functions in Appendix B. The moments for the first Volterra term of the full and reduced systems are respectively given by

$$M_j = e_{siz}^T (G^{-1}C)^j G^{-1} E e_p \quad \text{and} \quad \hat{M}_j = e_{siz}^T X (\hat{G}^{-1}\hat{C})^j \hat{G}^{-1} X^T E e_p. \tag{3.42}$$

The transfer functions for the second Volterra terms are two-dimensional and have the moments

$$\begin{aligned} M_{ij} &= e_{siz}^T (G^{-1}C)^i G^{-1} N (G^{-1}C)^j G^{-1} E e_p \quad \text{and} \\ \hat{M}_{ij} &= e_{siz}^T X (\hat{G}^{-1}\hat{C})^i \hat{G}^{-1} \hat{N} (\hat{G}^{-1}\hat{C})^j \hat{G}^{-1} X^T E e_p. \end{aligned} \tag{3.43}$$

As the accuracy of the reduced system largely depends on the ability of $\hat{y}^1$ and $\hat{y}^2$ to approximate y^1 and y^2, we focus in §**??** on the ability of the reduced system to match the leading two moments, M_j and M_{ij}.

3.3.3 Theory

We compare the ability of three reduced models to approximate the siz potential given synaptic input. The first reduced model was recently presented in (Yan and Li [97]). Given a small number of synaptic inputs at compartments $\{p_j\}$, let the reducer Z and corresponding reduced matrices be given by

$$\begin{gathered} Z = \text{orth}\begin{pmatrix} G^{-1}B & G^{-2}B & \dots & G^{-r}B \end{pmatrix}, \quad B = \begin{pmatrix} e_{p_1} & \dots & e_{p_m} \end{pmatrix}, \\ \hat{C} = Z^T C Z,\ \hat{G} = Z^T G Z, \quad \text{and} \quad \hat{N} = Z^T N Z. \end{gathered} \tag{3.44}$$

Such a reduced model is accurate but is limited by the number of synaptic inputs. The second reduced model is constructed from the reducer derived for injected input, presented in §3.3.1 and given by

$$\begin{gathered} X = \text{orth}\begin{pmatrix} G^{-1}e_{siz} & G^{-2}e_{siz} & \dots & G^{-r}e_{siz} \end{pmatrix}, \\ \hat{C} = X^T C X,\ \hat{G} = X^T G X, \quad \text{and} \quad \hat{N} = X^T N X. \end{gathered} \tag{3.45}$$

The final reduced model we consider approximates y by y^1, where the synaptic input is replaced by the corresponding injected input approximation. The resulting reduced model is given by Eq. 3.45 with $N = 0$. We compare the accuracy of the three models theoretically in Props. 6 and 7 and numerically in Figs. 3.16 and 3.17. The theory is again presented in the context of monosynaptic input but can be generalized for polysynaptic input.

Prop. 6 If the cable is driven by a single synaptic input at compartment p and the reduced model is given by Eq. 3.44, then $\hat{H}_k^{reg}(s_1, \dots, s_k) = H_k^{reg}(s_1, \dots, s_k) +$

$O(s_1^r \cdots s_k^r)$ for any $k \geq 1$.

Proof: For the uniform cable, $C = AC_m I_n$, and thus $K_r(G^{-1}C, G^{-1}e_p) \subseteq \text{sp}(Z)$. Assume $0 \leq i < r$. By Eq. 3.42, one can show that $M_i = \hat{M}_i$ similarly to the proof of Prop. 1. By substituting $N = e_p e_p^T$ into Eq. 3.43,

$$\begin{aligned} M_{ij} &= [e_{siz}^T (G^{-1}C)^i G^{-1} e_p][e_p^T (G^{-1}C)^j G^{-1} E e_p] \\ &= M_i e_p^T (G^{-1}C)^j G^{-1} e_p. \end{aligned} \tag{3.46}$$

Similarly, since $M_i = \hat{M}_i$,

$$\hat{M}_{ij} = M_i e_p^T Z (\hat{G}^{-1}\hat{C})^j \hat{G}^{-1} Z^T e_p.$$

Using the projector $Q \equiv Z\hat{G}^{-1}Z^T G$ onto $\text{sp}(Z)$,

$$\begin{aligned} \hat{M}_{ij} &= M_i e_p^T (Z\hat{G}^{-1}Z^T C)^j Z\hat{G}^{-1}Z^T e_p \\ &= M_i e_p^T (QG^{-1}C)^j QG^{-1} e_p \\ &= M_i e_p^T (G^{-1}C)^j G^{-1} e_p = M_{ij} \end{aligned}$$

for $0 \leq j < r$. One can continue the proof iteratively for any $k > 2$. ■

To theoretically examine the accuracy of the reduced model given by Eq. 3.45, we focus on the ability of the reduced model to match the two leading moments, M_i and M_{i0}, for $0 \leq i < r$, which provides a good indication of the accuracy in the siz approximation for fast, small inputs.

Prop. 7 Assume $0 \leq i < r$. If the cable is driven by a single synaptic input into compartment p and the reduced model is given by Eq. 3.45, then $\hat{M}_i = M_i$, and

$$\frac{|M_{i0} - \hat{M}_{i0}|}{|M_{i0}|} = 1 - \sum_{k=1}^{r} (e_p^T q_k)^2 / (e_p^T G^{-1} e_p) < 1, \tag{3.47}$$

where $\mathrm{sp}(q_1, \ldots, q_r) = \mathrm{sp}(X)$ and $q_j^T G q_k = \delta_{jk}$, the Kronecker delta function, for $1 \leq j, k \leq r$. The relative error given by Eq. 3.47 converges to zero monotonically as $r \to n$, and $M_{i0} = \hat{M}_{i0}$ for any r if the synapse and siz are located in the same compartment.

Proof: Assume $0 \leq i < r$. One can show that $\hat{M}_i = M_i$ similarly to the proof of Prop. 1. Since $N = e_p e_p^T$ and $C = AC_m I_n$, Eqs. 3.42 and 3.43 imply that

$$\begin{aligned}
\hat{M}_{ij} &= M_i e_p^T X (\hat{G}^{-1}\hat{C})^j \hat{G}^{-1} X^T e_p \\
&= (AC_m)^j M_i e_p^T (X \hat{G}^{-1} X^T)^{j+1} e_p \\
&= (AC_m)^j M_i e_p^T (P G^{-1})^{j+1} e_p,
\end{aligned}$$

where $P \equiv X \hat{G}^{-1} X^T G$ is a projector onto $\mathrm{sp}(X)$. By combining this result with Eq. 3.46,

$$\frac{\left|M_{ij} - \hat{M}_{ij}\right|}{\left|M_{ij}\right|} = \frac{\left|e_p^T \left[(G^{-1})^{j+1} - (PG^{-1})^{j+1}\right] e_p\right|}{\left|e_p^T G^{-(j+1)} e_p\right|}. \tag{3.48}$$

Thus, the ability of $\hat{y}^2$ to approximate y^2 depends on the action of the projector P on G^{-1}. We focus on the error in the leading moments for which $j = 0$. As described in §3.1.6, for dx sufficiently small,

$$G^{-1}(c_1, c_2) \approx \frac{\lambda R_a}{\pi a^2 \sinh(\ell/\lambda)} \cosh\left(\frac{\ell - x_{c_1}}{\lambda}\right) \cosh\left(\frac{x_{c_2}}{\lambda}\right), \tag{3.49}$$

for any two compartments c_1 and c_2 at locations $x_{c_1} \geq x_{c_2}$, where λ is given by Eq. 3.38. Thus, $e_p^T G^{-1} e_p > 0$, and Eq. 3.48 can be written in terms of the Rayleigh quotient R by

$$\frac{|M_{i0} - \hat{M}_{i0}|}{|M_{i0}|} = |1 - R|,$$

where

$$R = \frac{(G^{-1/2}e_p)^T (G^{1/2} P G^{-1/2})(G^{-1/2}e_p)}{(G^{-1/2}e_p)^T (G^{-1/2}e_p)}.$$

Let $\{\theta_k\}$ and $\{\widetilde{q}_k\}$ be the n eigenvalues and eigenvectors of $G^{1/2} P G^{-1/2}$. Since P is a projector of rank r, each of its eigenvalues, θ_k, and eigenvectors, $q_k = G^{-1/2}\widetilde{q}_k$, obey

$$\theta_k = 1 \quad \text{and} \quad q_k \in \text{sp}(X) \quad \text{for} \quad k = 1, \ldots, r;$$

$$\theta_k = 0 \quad \text{and} \quad G q_k \in \text{null}(X^T) \quad \text{for} \quad k = r+1, \ldots, n;$$

where we used the fact that if $P q_k = 0$, then $X^T G q_k = 0$ since both X and $\hat{G}^{-1}$ have full column rank. Therefore, $0 \leq R \leq 1$.

By the symmetry of $G^{1/2} P G^{-1/2}$, assume its eigenvectors $\{\widetilde{q}_k\}$ form an orthonormal basis for $\mathbb{R}^n$, which permits the decomposition

$$G^{-1/2} e_p = \sum_{k=1}^{n} (e_p^T G^{-1/2} \widetilde{q}_k) \widetilde{q}_k = \sum_{k=1}^{n} (e_p^T q_k) \widetilde{q}_k,$$

where $q_j^T G q_k = \widetilde{q}_j^T \widetilde{q}_k = \delta_{jk}$ for $1 \leq j, k \leq n$. Hence,

$$(G^{-1/2}e_p)^T (G^{1/2} P G^{-1/2})(G^{-1/2}e_p) = \sum_{k=1}^{n} (e_p^T q_k)^2 \theta_k = \sum_{k=1}^{r} (e_p^T q_k)^2,$$

and

$$(G^{-1/2}e_p)^T (G^{-1/2}e_p) = e_p^T G^{-1} e_p = \sum_{k=1}^{n} (e_p^T q_k)^2.$$

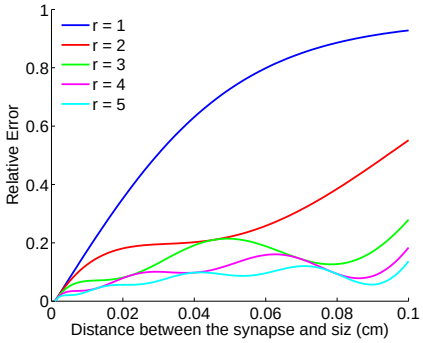


Figure 3.15: Relative error between M_{i0} and $\hat{M}_{i0}$, given by Eq. 3.47 and plotted as a function of the distance between the synapse and siz. The error is independent of i if $i < r$, and we numerically find that the error decreases for $j > 0$.

Therefore, $R \to 1$ monotonically as $r \to n$, and

$$R = \sum_{k=1}^{r} (e_p^T q_k)^2 / (e_p^T G^{-1} e_p).$$

Since q_1 is proportional to $G^{-1} e_{siz}$, $e_p^T q_1 > 0$ by Eq. 3.49. Hence $R > 0$, and Eq. 3.47 holds. Finally, $G^{-1/2} e_{siz} = G^{1/2} G^{-1} e_{siz} \propto \widetilde{q}_1$, implying that $R = 1$ and $M_{i0} = \hat{M}_{i0}$ if $x_p = x_{siz}$. ■

Fig. 3.15 shows the relative error given by Eq. 3.47 for a uniform cable driven by a single synaptic input. Although the reduced model given by Eq. 3.45 does not capture the higher dimensional moments as well as does the reduced model given by Eq. 3.44, it is independent of the synaptic input and does not grow with the number of synapses. Furthermore, since the relative error between M_{i0} and $\hat{M}_{i0}$ is less than one, it captures some of the nonlinearity incorporated by y^2, providing a better approximate for the siz potential than does y^1, the approximation given only injected input for which $\hat{M}_{ij} = 0$. This is illustrated in Fig. 3.16.

We close by stimulating a CA1 pyramidal cell with fifty transient synaptic inputs, 70% of which arrive at distal excitatory synapses and 30% of which arrive at proximal

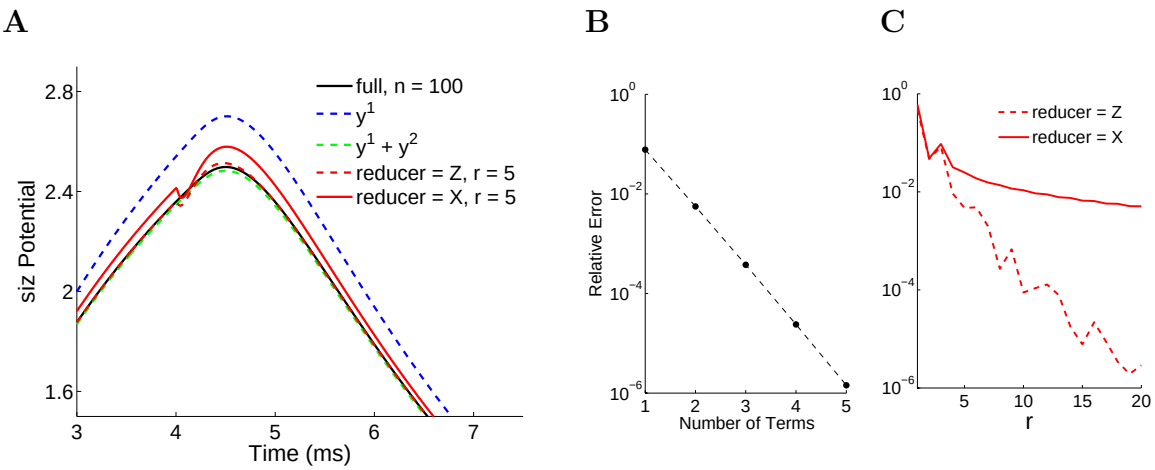


Figure 3.16: Comparison of the reduced passive systems and the partial summations in the Volterra series given a single synaptic input at the midpoint of the cable. **A.** Potential at the siz for the full system containing 100 compartments (solid black), the sum of the leading Volterra terms (dashed blue and green), and the reduced systems generated from Eq. 3.44 (dashed red) and Eq. 3.45 (solid red). **B.** Relative 2-norm error in the siz potential for the sum of the leading Volterra terms (left) and for the two reduced systems (right). The output of the reduced system constructed from X converges to the sum of the first two Volterra terms.

inhibitory synapses. The reducer Z can no longer be used, and the results again demonstrate that the reduced system generated from X incorporates much of the sublinear summation. Results are shown in Fig. 3.17.

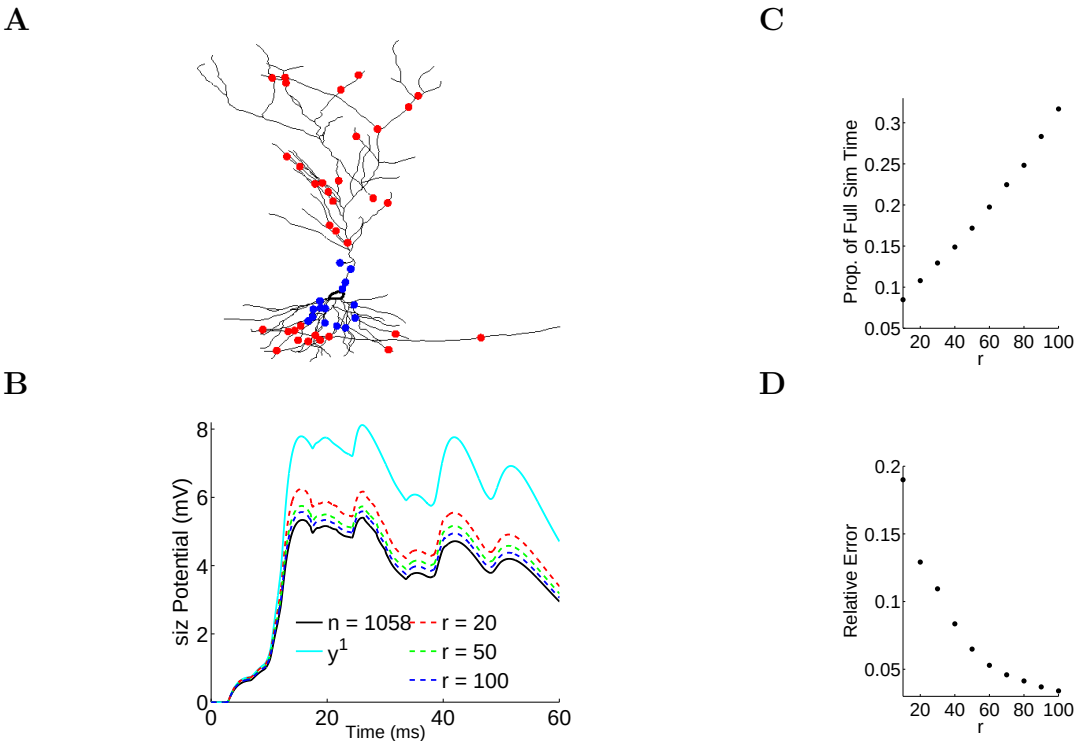


Figure 3.17: Integration of synaptic input within the full dendritic tree of a CA1 pyramidal cell. **A.** Spatial distribution of fifty transient conductances given by Eq. 3.33 within a 2D-projection of the dendritic tree, where $\overline{g} = 3\ nS$ and $\tau = 3\ ms$. Red dots indicate excitatory synapses for which $E_{syn} = 50\ mV$, and blue dots indicate inhibitory synapses for which $E_{syn} = 0$. **B.** Potential at the siz for the full system of size 1058 (solid black), the first Volterra term (solid cyan), and the reduced system (dashed) given by Eq. 3.45. The potential y^1 incorporates only the response to injected input, while the reduced system demonstrates sublinear summation. **C.** Proportion of the simulation time for the reduced system with r compartments to that of the full system. **D.** Relative 2-norm error in the siz potential.

3.4 Conclusion

We construct morphologically accurate reduced models of the passive neuron given current injection, the quasi-active neuron given current injection, and the passive neuron given synaptic conductances. In each case the reduced model accurately reproduces the transmembrane potential at the siz, increases the efficiency of numerical simulations, can be represented by a circuit similar to that of the full model, and is constructed from a reducer that is input-independent and formed through a simple application of the Arnoldi procedure. Our numerical simulations illustrate that the reduced models are accurate given any general set of inputs, including the worst case scenario of square pulses with random start times. We analyze their accuracy by examining the transfer functions evaluated along the imaginary axis to show that the reduced models perform well in theory as well as in practice.

Transient synaptic input currents can be approximated by the corresponding injected input currents for which the sublinear summation effect is ignored. The passive model then becomes a linear mapping from the injected input currents to the siz potential, and a classic Krylov subspace projection method gives rise to a reduced model represented by an RC circuit similar to the circuit representation of the full model. For both the full and reduced systems, the state variables represent the transmembrane potential within each compartment, and currents cross the membrane of each compartment according to the membrane capacitance and leakage conductance. The circuits are not identical, however, as the reduced circuit has an additional trans-

membrane current, and the reduced compartments are fully connected. Despite this increased density of connections, the reduced model increases the efficiency of simulations since accuracy is obtained using far fewer compartments than are needed in the full model. For example, in our simulations of a CA1 pyramidal cell, the reduced model accurately reproduces the siz potential using less than 1% of the number of compartments in the full model and has a speed-up factor of over ten.

The reducer is a matrix used to transform the full system into the reduced system and is chosen such that the transfer functions for the full and reduced systems agree at small frequencies, a desirable property since the transfer function for the passive neuron is a low-pass filter. Each column k of the reducer maps the inputs for the full system into the input for compartment k of the reduced system. The first column acts as the dendritic filter to attenuate each input as a function of the electrotonic distance between the corresponding synapse and siz. The subsequent columns are generated such that the reducer is orthonormal and spans the space determined by a particular set of basis vectors for which we derive analytical expressions. As the basis vectors depend only on the location of the siz and the intrinsic properties of the neuron (the membrane capacitance and conductance, the axial conductance, and the compartmental connectivity), the reducer and thus the reduced model are input-independent.

The quasi-active system is represented by an RLC circuit. If the neuron has n compartments, then the last n state variables represent the transmembrane potential

within each compartment, and each prior subgroup of n state variables represents the transmembrane current within each compartment dependent on the corresponding gating variable. We use a model reduction technique designed to preserve the block structure in the system's matrices (Li and Bai [52]) to preserve this RLC circuit structure. The state variables of the reduced system are analogous to those of the full system, and the reduced compartments again become fully connected. As was found in the passive case, however, accuracy in the siz potential requires relatively few compartments, and the reduced system is highly efficient despite the increase in density. For our simulations of the CA1 pyramidal cell consisting of sodium and potassium channels as well as the h-current, the reduced model accurately reproduces the siz potential using only 0.5% of the number of compartments in the full model and has a speed-up factor of almost one hundred.

The theory underlying the reduction of the quasi-active system is similar to the theory for the reduced passive system, but the transfer function is no longer a low-pass filter. Quasi-active neurons generally have a nonzero resonant frequency, which shifts the peak of the transfer function evaluated along the imaginary axis away from zero to create a band-pass filter. Furthermore, the poles of the transfer function shift closer to the real axis as the potential oscillations grow, and they cross into the right half of the complex plane for unstable systems. Despite these nuances, approximating the transfer function at small frequencies still produces an accurate reduced system given the physiological range of resonant frequencies.

The reduced model easily generalizes to the dendritic tree for which the radius varies in space. When the peak conductance of an active channel varies, however, the reduced model generated from Eqs. 3.25-3.28 no longer applies as Prop. 4 no longer holds. A reduced model of size r can be constructed in this general case using the basis vectors $X_{basis} = \{(\mathcal{G}^{-1}\mathcal{C})^k\mathcal{G}^{-1}e_{siz}\}_{k=0}^{r-1}$, as was done in (Yan and Li [97]). Although accurate, the resulting reduced system does not lend itself to a circuit representation. Structure-preserving model reduction of the general quasi-active system is an important future direction of this work.

The passive system given synaptic input is the most computationally complex system we consider as the siz potential has a nonlinear dependence on the synaptic input conductances. Despite this complexity we find that the siz potential is well-approximated by applying the reducer derived for the passive system given injected input. To analyze the accuracy of the resulting reduced system, we expand the siz potential for each system as a Volterra series in which each term is governed by a system of linear ODEs. The leading terms dominate the series, and for fast, small input conductances, only the first few terms make a noticeable contribution to the summation. We thus focus on the ability of the reduced model to capture the first two Volterra terms.

The reduced model quickly captures the first Volterra term as this term is nothing more than the response to injected inputs. Although the convergence is slower for subsequent Volterra terms, we show through both theory and numerical simulations

that the reduced model captures much of the sublinear summation incorporated by the second Volterra term, always providing a better approximation of the siz potential than does the full model for which the synaptic input is approximated by injected input. Given a small number of fixed synapses, the reduced model presented in (Yan and Li [97]) is preferable as it requires the potential to be accurate at each synapse. This reduced model, however, depends on the synaptic locations and grows with the number of synapses, and thus our reduced model should be used for the general case in which synaptic input may arrive at any compartment, a necessary assumption for retaining the spatial specificity in the input signal. The next step in this work is to construct a reduced model which captures all Volterra terms but is input-independent and preserves the circuit structure. Such a reduced model would be an important step toward structure-preserving model reduction of the active neuron.

The reduced models we present are unique as they not only accurately reproduce the siz potential, but they also preserve the underlying electrophysiology, have simple construction procedures, can be applied to any neuron, and are independent of the locations and number of inputs. A reduced model with these properties is a powerful tool for modeling neuronal activity as the input-output map of the reduced model can be analyzed and interpreted in terms of neuronal function. When modeling neurons for which the inputs are well understood, such as the LGMD or CA1 pyramidal cell, the reduced model can be used not only for efficient numerical simulations but also to help elucidate functions performed in dendrites.

Chapter 4

Network Model of the Double Rotation Experiment

Place cells encode a spatial map sensitive to context. That is, place fields responsive to a specific location within a learned environment are modulated by stimuli which occurred at that location, such as the interaction with external landmarks. Given that the entorhinal cortex provides the primary input signal to the hippocampus through its two subregions, theories abound that the spatial representation of place cells originates in the absolute, context-independent reference frame of the MEC, while the context-dependence of place fields is due to sensory feedback regarding external landmarks from the LEC (Hargreaves et al. [28], Knierim et al. [45], Lisman [55], Hayashi and Nonaka [30], Silkis [82]).

This theory has been applied in models of path integration with error correction.

Given a starting position, the spatial metric of the MEC coupled with input from cells whose activity is modulated by velocity enable place cells to continuously update a subject's position along a trajectory. The error which unavoidably accumulates during path integration necessitates a sensory feedback signal to perform error correction, allowing a subject to navigate a path through internally tracking its position with occasional glances to external landmarks to localize itself (Hasselmo [29], Monaco et al. [61]). The LEC may provide this error correction signal.

Similarly, the LEC may contribute to a subject's recognition of a learned environment upon re-entry, helping to reset the starting position. While the relative roles of the MEC and LEC in self-localization are usually speculative, the double rotation experiment provides a paradigm in which these roles can be tested. According to the experimental results, the fixed spatial metric of the MEC is tuned to distal cues, or the global framework, while the LEC signal is weakly tuned to local cues, which can be interpreted as objects or landmarks (Lee et al. [49], J. Neunuebel [manuscript in preparation]).

I propose that upon the rat's initial exposure to the modified track, the LEC provides a teaching signal strong enough to initially cause place fields to fire with local cues. The MEC then provides the primary spatial drive as the rat circles the track. I use computational models to demonstrate that mechanisms and theories generally accepted in the literature, such as synaptic plasticity and cell assembly dynamics, can combine in non-intuitive ways to cause place fields to rotate with the

weak LEC signal for local cues over the relatively strong MEC signal for distal cues.

In this chapter I describe the setup for the network model of the double rotation experiment, providing an overview of the model in §4.1, describing the method for spike generation in §4.2, and illustrating the place field dynamics apparent in the model before the cues are rotated in §4.3. The insights gained by modeling the cue mismatch phase are presented in §5.2.

4.1 Network Model Overview

The network model illustrated in Figure 4.1 is a simplification of the entorhinal cortical and hippocampal circuitry shown in Figure 1.3, where only the LEC, MEC, and CA3 subregions are included. The dentate gyrus is excluded due to its sporadic activity and to experimental results indicating that DG cells show no preference for the local or distal cue set (J. Neunuebel [manuscript in preparation]). As this study focuses on the activity of CA3 cells, the CA1 is also excluded, although the model could easily be extended to investigate the response in the CA1. The activity of cells in the entorhinal cortex mimics the activity seen experimentally and depends solely on the rat's position; therefore, the hippocampus provides no feedback to the entorhinal cortex in the model.

Between 20% and 40% of place cells are active within any given region, and it is necessary to simulate large networks of cells to investigate the effect of recurrent connections among place cells while maintaining this sparsity. Such a model comes at a

high computational cost as three coupled ODEs must be solved for each hippocampal cell and all plastic weights must be updated. To isolate the network dynamics without incurring further expense due to a spatial discretization, hippocampal cells in the model are isopotential. Even so, the results presented in this thesis required solving a system with 6.3 million variables to be updated at each timestep of the backward Euler scheme. I handle this computational expense through a parallel implementation, running the simulations on a high-performance cluster. The computational capability of the model is illustrated in Figure 4.2.

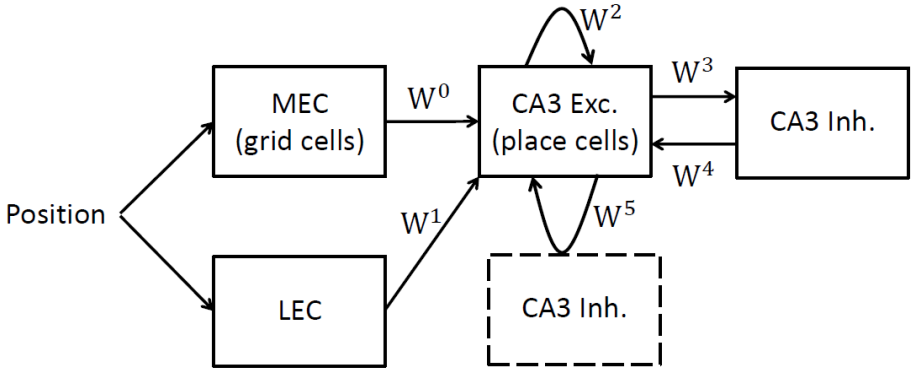


Figure 4.1: Schematic of the simplified network modeled in the double rotation simulations. MEC and LEC cells are given nonhomogeneous Poisson spike trains with a rate dependent on the position of the rat. CA3 excitatory and inhibitory cells integrate synaptic input according to a conductance-based integrate and fire model. The strength of each synapse is stored in a weight matrix, where W^0, W^1, and W^5 are governed by rate-based plasticity; W^2 is governed by STDP; and W^3 and W^4 are constant. CA3 inhibitory cells marked with a solid box provide non-specific feedback inhibition to control the overall activity level. CA3 inhibitory cells marked with a dashed box provide targeted feedback inhibition and are modeled only implicitly. When a given place cell is active, it excites place cells with place fields overlapping its own according to W^2 while effectively inhibiting all other place cells according to W^5.

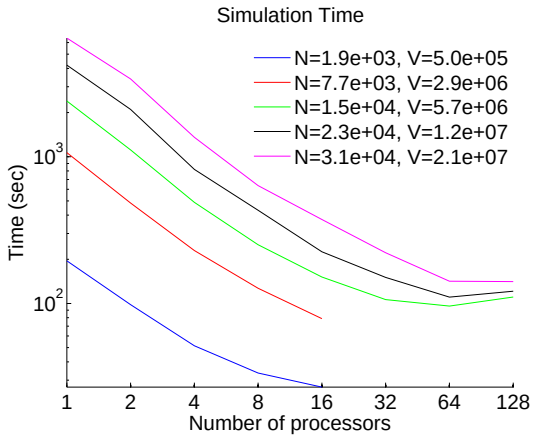


Figure 4.2: Computational capability of the model. Due to its parallel implementation, the model can efficiently simulate large networks of cells. N denotes the total number of cells, and V denotes the corresponding number of dynamic variables. The time taken to run 15 laps is plotted against the number of processors used.

4.1.1 Single-Cell Models

Cells in the MEC are modeled as grid cells with hexagonal firing patterns (see Figure 1.1). This is again a simplification as the MEC contains a variety of cells, grid cells comprising only a minority. However, grids are undersampled as the rat is confined to the track, and thus the results would be unchanged were the model to include any spatially tuned cell in the MEC. LEC cells are modeled to have noisy activity weakly tuned to local cues, which correspond to the rat's position in the standard session (before cue rotations).

In both the model and in reality, hippocampal cells receive input regarding the rat's position after it has been processed by the entorhinal cortex. In the model place cells initially have no preference for any given location on the track. Rather, they integrate synaptic inputs from MEC, LEC, and other CA3 cells via the conductance-

based integrate and fire model.

The number of cells within each subregion is set to maintain accepted ratios among subregions while providing enough active place fields to examine the effect of recurrent connections. For all simulations presented in this thesis, the network contains 540 LEC cells, 1080 MEC cells, 2400 CA3 excitatory cells, and 192 CA3 inhibitory cells.

4.1.2 Synaptic Plasticity Models

The strength of the connections among cells in two connecting subregions is stored in a weight matrix in which the synapses, or elements allowed to be nonzero, are set randomly according to a prescribed density. Since the CA3 contains over 200,000 cells (Ascoli et al. [3]), the density among groups of relatively few cells is set to a much higher value than anatomical data would suggest. One should keep in mind, however, that the majority of weights are driven to zero by the synaptic plasticity mechanism applied. All weights are bounded between zero and a set maximum weight, constant weights are set randomly within the prescribed range, and weights within each plastic matrix are initialized to the same small value such that LEC and CA3 input cannot cannot drive place cells to spike, while MEC input drives weak place cell activity. The parameters used for the simulations presented in this thesis are provided in Appendix B.

There is no initial structure in any weight matrix since synapses are set randomly, as are the MEC and LEC parameters determining their firing fields. Rather, structure

emerges as synaptic plasticity is applied. The plasticity models are phenomenological, chosen for their ability to reproduce place field dynamics observed in experiments, although these dynamics may in fact be due to other mechanisms in the brain.

Postsynaptically gated rate-based plasticity (see §1.2.1) governs the weights from MEC and LEC cells to CA3 excitatory cells, stored in W^0 and W^1, respectively. This plasticity mechanism enables grid cells to beget place cell activity (Savelli and Knierim [78]). According to the data from the double rotation experiment, LEC cells are weakly tuned to the local cues. To investigate the influence of this weak LEC signal on place cell activity, a given LEC cell must be coupled, at least weakly, to place cells responsive to a location on the track at which the LEC cell has increased activity. This is achieved in the model through the same rate-based plasticity rule applied to connections from the MEC, although the learning and threshold rates are smaller. While plasticity is known to differ between afferents from the MEC and LEC into the CA3 (Martinez et al. [57]), the results of the model would be similar given any plasticity mechanism that enables LEC cells to provide a noisy input signal to place cells weakly correlated with their place fields.

The model includes non-specific feedback inhibition through which place cells excite random inhibitory cells, which in turn inhibit random place cells. This controls the level of overall activity and causes place fields to spread evenly throughout the track. All weights within W^3 and W^4 are constant, and the maximum weights are set to obtain an appropriate level of place cell activity.

As described in §4.4, the model incorporates the theory of a cell assembly in which a place cell excites place cells with correlated activity while inhibiting place cells with uncorrelated activity. The cell assembly must form dynamically as place fields form in the training session, and it must adapt to the partial remapping seen in the mismatch session. This is accomplished in the model through synaptic plasticity applied to W^2 and W^5. The asymmetric STDP rule governs the excitatory weights of W^2, encoding the repetitive sequence of activated place fields among the cell assembly. Rate-based plasticity governs the inhibitory weights of W^5, creating a targeted inhibition that enforces stability on the cell assembly.

Targeted inhibition is feedback in nature, the idea being that place cells excite CA3 inhibitory cells, which in turn inhibit specific place cells. The implementation, however, simplifies this process by using a single weight matrix, W^5, to effectively bypass the inhibitory CA3 cells. Given any pair of place cells, W^5_{jk} is the peak synaptic conductance at an inhibitory synapse of cell k given activity in cell j (see §4.2.3). This simplification could affect the results since inhibitory cells typically have a higher firing rate than do excitatory cells. However, the effect in the model is that co-active place cells excite one another while inhibiting place cells that are either responsive to other locations on the track or silent.

The update rule is modified from Eq. 1.1, where the weight from cell j to cell k is governed by

$$\frac{dW^5_{jk}}{dt} = -k(R_k - \overline{R})R_j. \tag{4.1}$$

This modified rule drives the weight between cells with correlated activity to zero and the weight between cells with uncorrelated activity to the upper bound. In reality most place fields are apparent upon the rat's initial exploration of an environment and are not created over time. Since the weight update is gated by the presynaptic firing rate, targeted inhibition ensures that place cells silent after the first few laps remain silent, creating a sparse representation of the environment. When place fields disappear in the mismatch session, new place fields form due to the decrease in targeted inhibition, helping to maintain a balance in the population activity level.

4.2 Spike Generation

Two methods are used to generate spikes. The firing patterns for MEC and LEC cells are determined empirically and directly depend on the rat's position. Hippocampal cells integrate synaptic inputs from the MEC, LEC, and other CA3 cells.

4.2.1 Grid Cells in the MEC

Grid cells in the model fire in hexagonal patterns according to three parameters: the tilt θ, base b, and offset $\delta = (r, \phi)$. As demonstrated in Figure 4.3, the tilt specifies the angle from the Cartesian x-axis to the grid axis, the base specifies the distance from one grid point to another along a grid lateral, and the offset specifies both the magnitude r and direction ϕ of the difference between the grid center and the origin, where the grid center is given by $\mathbf{c} \equiv (r\cos(\phi), r\sin(\phi))$. The base, or scale, for each

grid cell is set randomly within 25 to 60 *cm* in agreement with experimental data from dorsal MEC (Hafting et al. [27], Brun et al. [10]), the region containing afferents to place cells recorded from in the double rotation experiment. The tilt and offset of each grid are set randomly.

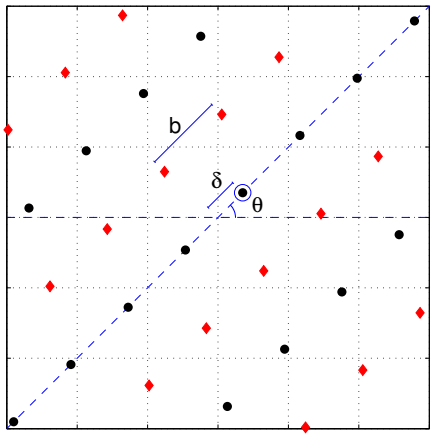


Figure 4.3: Graphical demonstration of grid cell parameters. This plot shows an example of a grid, $G(\theta, b, \delta)$, where $\theta = \pi/4$ rad, $b = 34.4$ cm, and $\delta = (14.9 \text{ cm}, \theta)$. The black circles and red diamonds represent the elements of G_1 and G_2, respectively. The grid axis is shown by a dotted blue line, and a circle is drawn around the grid center.

Let the set of grid points be given by

$$G(\theta, b, \delta) = G_1(\theta, b, \delta) \cup G_2(\theta, b, \delta),$$

where

$$G_1(\theta, b, \delta) = \{\mathbf{c} + kb(\cos\theta, \sin\theta) + 2jh(-\sin\theta, \cos\theta) : j, k \in \mathbb{Z}\},$$

$$G_2(\theta, b, \delta) = \{\mathbf{c} + (k + (1/2))b(\cos\theta, \sin\theta) + (2j - 1)h(-\sin\theta, \cos\theta) : j,\ k \in \mathbb{Z}\},$$

and h is the grid height, given by $h = (b/2)\tan(\pi/3)$. Figure 4.3 shows an example of a grid, illustrating the grid's tilt, base, offset, and center.

Grid cells are topologically organized in the brain, where neighboring cells have similar tilts but random offset values, and the base of the grids and size of the grid fields increases with distance from the dorsal border (Brun et al. [10]). The model does not retain this topological organization as the weights from grid cells to place cells are randomized, providing no sense of which grid cells are located where along the axis. Rather, all parameters are chosen randomly within the given ranges.

Each grid cell generates a spike train represented by a vector containing the cell's spike times, which follow a Poisson distribution (Rodieck [73]). As done in (Savelli and Knierim [78]), I generate a thinned, nonhomogeneous Poisson spike train for each grid cell according to the algorithm below. As the interval between Poisson events follows an exponential distribution, an interspike interval (ISI) is first selected from an exponential distribution with mean $1/f_{max}$, where f_{max} is the expected firing rate were the rat to remain fixed at one of its grid points. The cell's refractory period, t_{ref}, provides a lower bound for the ISIs, modeling the reduction in excitability seen in a neuron immediately following a spike.

After the ISI expires, the probability of accepting a spike follows a Gaussian distribution dependent on the rat's distance to the nearest grid point, given by the metric

$$\mathrm{d}(\mathbf{x},\ G) = \min\{|\mathbf{x} - \mathbf{y}| : \mathbf{y} \in G\},$$

where $\mathbf{x}$ denotes the position of the rat.

Algorithm 3: Generation of a thinned, nonhomogeneous Poisson spike train

1. Initialize $t_0 = 0$.

2. Propose an ISI from an exponential distribution with mean $1/f_{max}$ and lower bound t_{ref}, given in MATLAB by

$$\Delta t = \max(\text{exprnd}(1/f_{max}), t_{ref}). \tag{4.2}$$

3. At $t = t_0 + \Delta t$, accept a spike according to the probability

$$P = \exp\left(\frac{-\text{d}^2(\mathbf{x}, G)}{\gamma b^2}\right), \tag{4.3}$$

   where $\mathbf{x}$ denotes the position of the rat, and γ determines the spread of the firing fields.

4. Set $t_0 \rightarrow t_0 + \Delta t$; return to step 2.

The size of each grid field increases with b, the distance between fields, a characteristic captured in the model by the dependence of the Gaussian spiking probability P on the base b in Eq. 4.3. Figure 4.4 demonstrates the effect of γ on the spread of the firing fields as a rat explores a rectangular enclosure. I set γ to be 0.018 in the model to best match experimental data (Hafting et al. [27], Savelli and Knierim [78]). All other grid cell parameters are provided in Appendix B.

The uniform grid is undersampled in the double rotation experiment as the rat is constrained to a circular track. As a result, grid fields appear less periodic, and

the number, size, and peak rate of firing fields intersecting the track vary across grid cells. Figure 4.6 provides three examples of grid cell activity on the track.

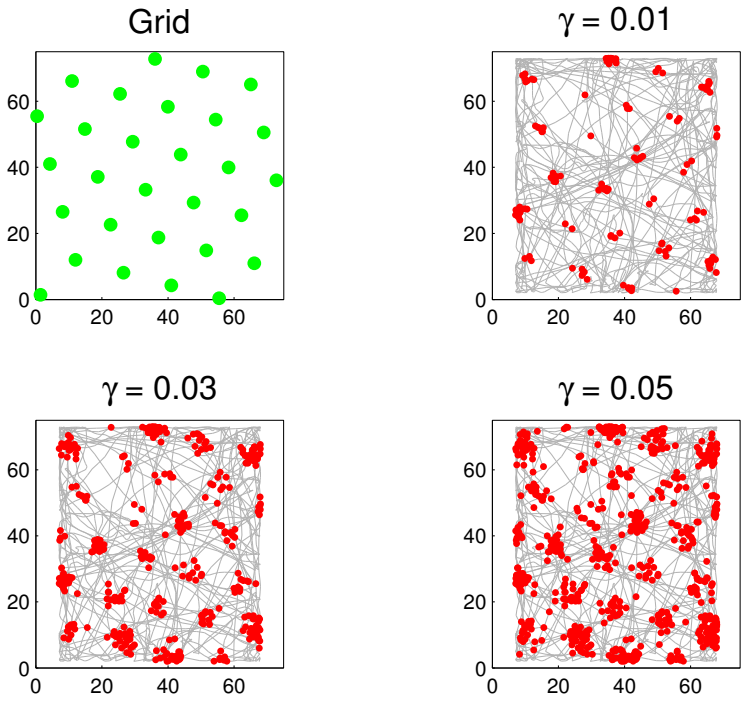


Figure 4.4: Effect of γ on the size of grid cell firing fields. For each value of γ, the firing fields form a hexagonal pattern corresponding to the underlying grid (top left). As γ increases, the width of the Gaussian probability given by Eq. 4.3 increases, causing the spread of the firing fields to increase. For all plots, $\theta = \frac{\pi}{4}$ radians, $b = 15$ cm, and $\delta = (2$ cm, θ rad). The rat runs the same trajectory for all values of γ.

4.2.2 LEC Cells

Similar to MEC cells, LEC cells generate Poisson spike trains with a rate dependent on the rat's position. However, their fields are much more noisy and lack the spatial organization of a grid. Algorithm 3 is also used to generate LEC spike trains, but

Eq. 4.3 is replaced with

$$P = P_0 + \sum_{i=1}^{N_{hot}} P_{D,i} \exp(-d_i^2/\gamma^2), \tag{4.4}$$

where P_0 is the base probability, N_{hot} is the number of hot spots for the particular cell, $P_{D,i}$ is the probability difference for hot spot i, and d_i is the distance between the rat's position and the center of hot spot i. As for the MEC model, γ determines the spread of the firing fields. The spiking probability curve is demonstrated in Figure 4.5.

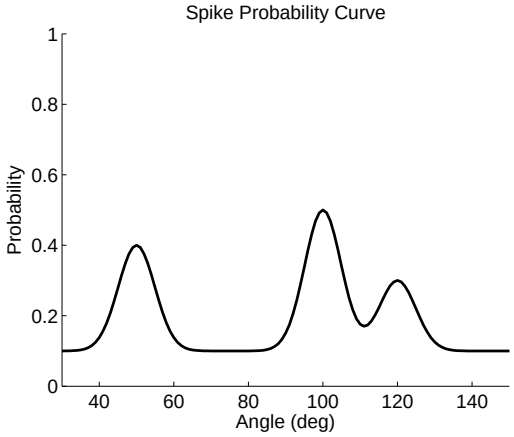


Figure 4.5: Demonstration of the spiking probability curve for LEC cells. For this example, the cell has three hot spots centered at 50°, 100°, and 120°. The curve has the base probability $P_0 = 0.1$ and probability differences $P_D = [0.3,\ 0.4,\ 0.2]$.

For all simulations presented in this thesis, $N_{hot} = 3$, $P_0 \in [0.05,\ 0.15]$, and $P_{D,i} \in [0,\ 0.3]$. Since $P_0 > 0$ and $P_{D,i} < 1$, LEC firing patterns have more noise and less spatial selectivity than do grid fields. Figure 4.6 compares MEC and LEC activity on the track.

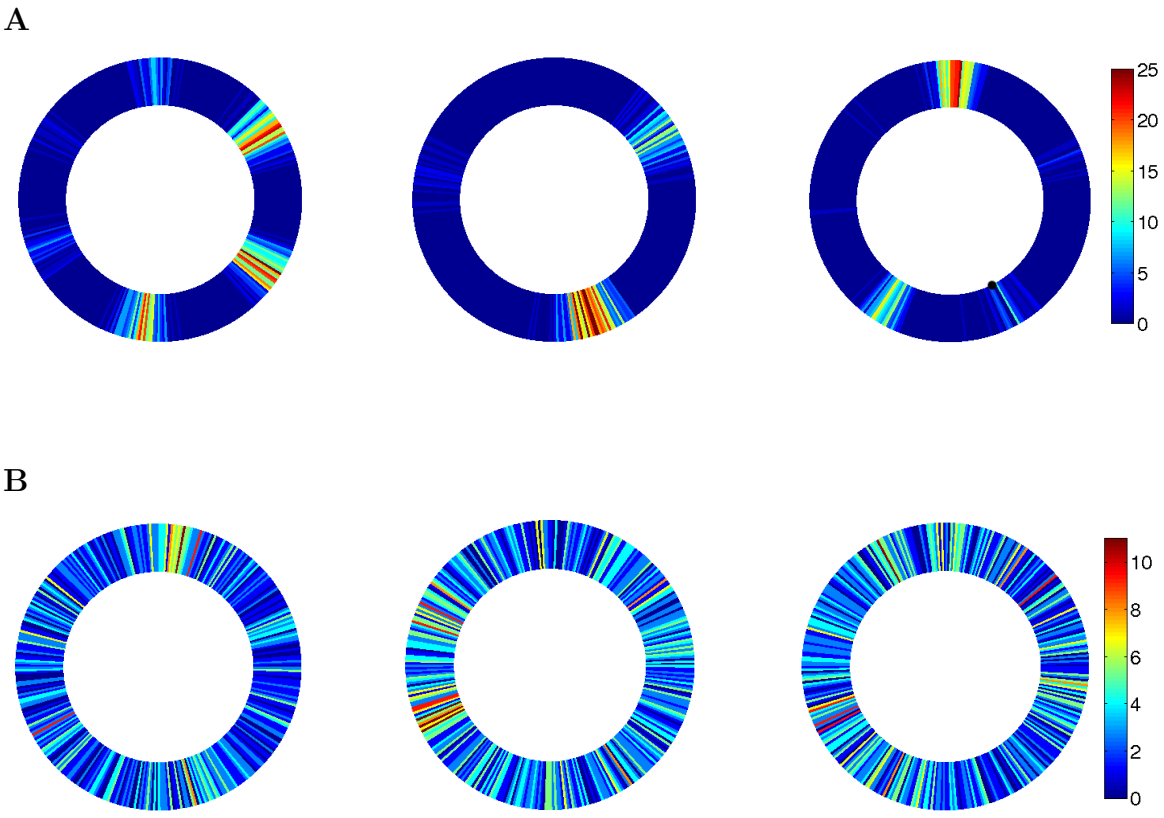


Figure 4.6: Firing rates of three representative MEC **(A)** and LEC **(B)** cells as the simulated rat is constrained to the track. Firing fields of MEC cells may appear irregular due to undersampling, and the scale of the fields varies. Firing fields of LEC cells are very noisy and exhibit a relatively weak spatial tuning. Spike rates were averaged over 15 laps for each bin of width 1°.

4.2.3 CA3 Cells

The method of generating spikes for CA3 cells is fundamentally different than that used for MEC and LEC cells as CA3 cells integrate synaptic input. There are several well-established models of the neuron which vary both in accuracy and complexity, as described in §2.1. To capture the network dynamics among large networks of cells, I choose the passive isopotential cell model due to its low computational cost for each neuron. In the future, this model could be combined with the model reduction technique described in §2.2.4 to retain the spatial specificity of inputs.

CA3 cells receive both excitatory and inhibitory synaptic inputs, as demonstrated in the circuit diagram of Figure 4.1. The distinguishing factor between the two inputs is the reversal potential associated with each synapse. For the excitatory synapse, the reversal potential V_E is typically around 0 mV, and thus an active synapse tends to depolarize a cell. On the other hand, the inhibitory reversal potential is approximately the resting potential of the cell, causing inhibitory currents to drive the potential back to rest.

Since the passive model cannot generate an action potential, I use the integrate and fire (IAF) model to approximate the spike times of the cell (Dayan and Abbot [13]). According to IAF, the cell spikes when its potential reaches a set threshold value, V_{th}. The passive model is accurate for small inputs and determines the subthreshold potential, where for cell i,

$$C_m \frac{d\phi_i}{dt}(t) = -g_L(\phi_i(t) - V_L) - g_i^E(t)(\phi_i(t) - V_E) - g_i^I(t)(\phi_i(t) - V_I)$$

for $\phi_i(t) < V_{th}$. In the absence of driving current, the cell's voltage decays to its resting potential, $\overline{\phi} = V_L$, at a rate of $\tau_m = C_m/g_L$.

The excitatory and inhibitory synaptic conductances, g^E and g^I, depend on input spike trains from excitatory and inhibitory presynaptic neurons, respectively. According to Figure 4.1, the conductances for excitatory cell i are governed by

$$\tau_E \frac{dg_i^E}{dt}(t) = -g_i^E + \sum_{j=1}^{N^{\mathrm{M}}} W_{ij}^0(t) S_j^{\mathrm{M}}(t) + \sum_{j=1}^{N^{\mathrm{L}}} W_{ij}^1(t) S_j^{\mathrm{L}}(t) + \sum_{j=1}^{N^{\mathrm{CE}}} W_{ij}^2(t) S_j^{\mathrm{CE}}(t),$$

$$\tau_I \frac{dg_i^I}{dt}(t) = -g_i^I + \sum_{j=1}^{N^{\mathrm{CI}}} W_{ij}^3 S_j^{\mathrm{CI}}(t) + \sum_{j=1}^{N^{\mathrm{CE}}} W_{ij}^5 S_j^{\mathrm{CE}}(t),$$

where N^{M}, N^{L}, N^{CE}, and N^{CI} denote the number of MEC, LEC, CA3 excitatory, and CA3 inhibitory cells, respectively; S_j^{M}, S_j^{L}, S_j^{CE}, and S_j^{CI} denote the spike train of cell j within each respective group; and W_{ij} denotes the strength of the synapse from presynaptic cell j to postsynaptic cell i. The spike train of cell j is defined as

$$S_j(t) \equiv \sum_{k=1}^{\sharp \mathbf{T}_j} \delta(t - \mathbf{T}_j(k)), \tag{4.5}$$

where $\mathbf{T}_j$ is a vector containing the times at which cell j spikes, $\sharp \mathbf{T}_j$ is the number of spikes contained in $\mathbf{T}_j$, and δ denotes the Dirac-delta function.

The conductances for inhibitory CA3 cell i are similarly governed by

$$\tau_E \frac{dg_i^E}{dt}(t) = -g_i^E(t) + \sum_{j=1}^{N^{\mathrm{CE}}} W_{ij}^4 S_j^{\mathrm{CE}}(t) \quad \text{and} \quad g_i^I = 0.$$

When $\phi_i(t) \geq V_{th}$, cell i generates a spike, and the current time t is added to the vector $\mathbf{T}_i$. The voltage is then reset to its reset value and held constant until the cell's refractory period has transpired. Table B.2 provides the parameters used for the IAF model.

4.3 Place Field Dynamics during Training

In the double rotation experiment, rats were trained over several days before recordings were taken so that the standard cue configuration would be a learned environment. Likewise, I simulate thirty training laps before examining place field activity in the standard and mismatch sessions. Understanding the role of inputs and plasticity mechanisms before the cue rotations is essential to understanding the place cell responses in the mismatch session.

4.3.1 Place Field Formation

Place cells are stimulated by a random subset of MEC and LEC cells, each of which has multiple firing fields. As a result, place cells would also have multiple firing fields were the weights fixed. Theoretical studies have shown that there exists a set of weights such that synaptic input from multiple grid cells creates a localized place field (Solstad and Moser [83] and Franzius [16]) and that synaptic plasticity can drive the weights to such a distribution (Savelli and Knierim [78]). Place fields are generated in the model through applying the rate-based plasticity rule of Savelli and Knierim, described in §1.2.1. As shown in Figure 4.7, the learning rule is robust enough that grid cells beget place cell activity despite noise from the LEC. Furthermore, the learning rule allows place cells to adapt to a modified environment, as described in Chapter 5.

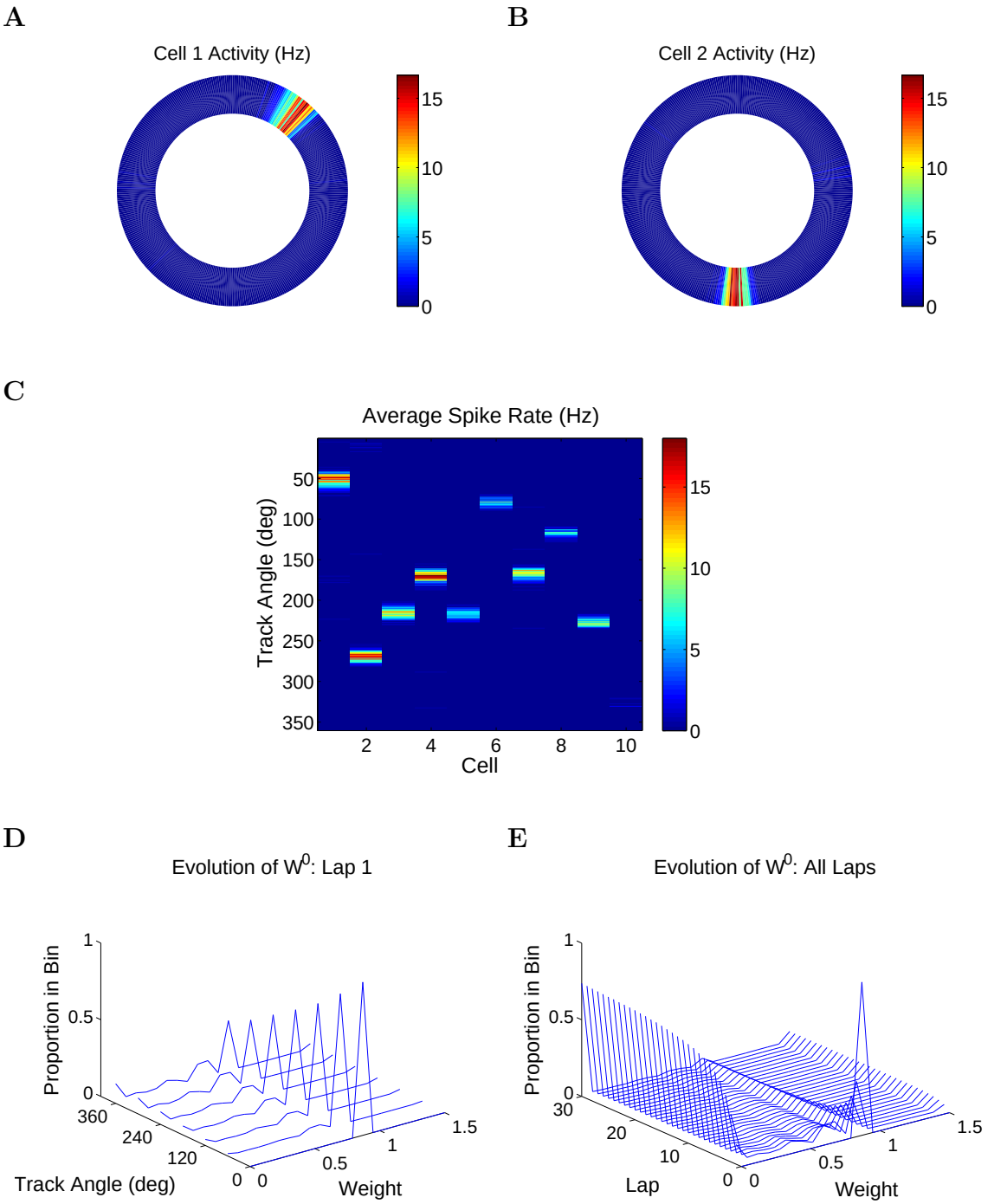


Figure 4.7: Place field formation. **A-B.** Average spike rate (Hz) for cells 1 and 2 in **(C)** at each degree on the track. **C.** Average spike rate for ten place cells, where the angle of the rat's position is plotted on the y-axis. **D-E.** Evolution of weights. Rate-based plasticity polarizes the weights from the MEC to the CA3. Weights onto silent cells remain fixed at their initial value of 0.9.

According to Eq. 1.1, the weight from MEC cell j to excitatory CA3 cell i is governed by

$$\frac{dW_{ij}^{0}}{dt}(t) = k(R_j(t) - \bar{R})R_i(t).$$

The weight change is gated by the postsynaptic spike rate, R_i, and the direction of weight change is determined by the presynaptic spike rate, R_j. As a result, when a place cell fires, the weights from all grid cells which contributed to its firing increase, while the weights from all other grid cells decrease. This causes a bipolarization in the resulting weight distribution, as illustrated in Figure 4.7, in which the place cell is increasingly driven by a subset of grid cells with one location in common among their firing fields.

$S_j(t)$, the discrete spike train given by Eq. 4.5, is converted to an instantaneous spike rate via

$$R_j(t) = \beta \int_0^\infty \exp\left(\frac{s-t}{\alpha}\right) S_j(s)ds = \beta \sum_{k=1}^{\sharp \mathbf{T}_j} \exp\left(\frac{\mathbf{T}_j(k)-t}{\alpha}\right).$$

The parameter β is set such that if the rat were to remain fixed at one of its grid points, the instantaneous spike rate would be f_{max}, the same maximal rate used to generate the Poisson spike train in Eq. 4.2. The decay constant α captures the traces of the pre- and postsynaptic firing activity. Although the learning rule is a high-level model, the presynaptic trace is inspired by the NMDA-dependent calcium influx, which has a decay constant between 100 and 1000 ms. For all results presented in this thesis, $\alpha = 100$ ms in agreement with (Savelli and Knierim [78]).

I characterize each place cell by the center of mass (COM) of its place field. For

a given cell j, let $R_j(d)$ denote the average spike rate of cell j at the rat's position, $(r\cos(d), r\sin(d))$, where r is the radius of the track and d is given in degrees. The track is discretized into bins of width 1°, and the average rate is given by the quotient of the number of spikes and the time spent within the given bin during all laps of the session. The COM of cell j is then given by

$$c_j \equiv \frac{\sum_{d=a_j}^{b_j} dR_j(d)}{\sum_{d=a_j}^{b_j} R_j(d)}. \tag{4.6}$$

To account for noise due to a few isolated spikes away from the place field, the COM is computed over the cell's place field, not over the entire track, where the place field is defined to be the interval over which the cell has its greatest activity. The upper boundary of the place field is given by

$$b_j \equiv \min\{x : x > M_j \text{ and } R_j(x+i) < 0.001 \text{ for } i = 1, \ldots, 10\},$$

where M_j is the angle at which cell j has its peak activity. The lower boundary, a_j, is defined analogously.

Factors Affecting Place Field Formation The rate-based plasticity model is effective only if the learning rate, specified by k in Eq. 1.1, is sufficiently large. This fast learning rate is in agreement with the rapid formation of place fields seen experimentally, and it is necessary in the model to resolve competition among different locations on the track. If the learning rate is fast, then when the CA3 cell spikes, the weights from all grid cells with a high spike rate are sufficiently increased for the cell to have a much higher probability of spiking at that same location during

subsequent laps. Perhaps more importantly, weights from grid cells with a low spike rate are sufficiently decreased for the cell to have a small probability of spiking at other locations on the track. If the learning rate is slow, the weights oscillate back and forth, and the cell forms multiple firing fields.

In addition to a fast learning rate, the variety in base values and the upper weight bound also significantly influence the formation of place fields. Rate-based plasticity isolates the location at which the place cell is driven by a large number of grid cells. The success of the plasticity model requires this group of grid cells to have only one firing location in common. If the group shares multiple firing locations along the track, the place cell has no mechanism for choosing one location over another, and it will form multiple firing fields. A variety of base values for the grid cells ensures that multiple grid cells generally do not have multiple firing fields in common.

If the upper bound for the weight from grid cells is too small, CA3 cells never receive enough input from the MEC to spike consistently. Place cells in this situation form a weak place field over which the cell has a very low spike rate. If the upper bound is too large, a single grid cell has too great of an impact on the CA3 cell. Because grid cells spike at several locations, this causes the CA3 cell to spike at several locations as well. The success of the rate-based plasticity model in forming place fields requires the upper bound to be small enough such that a CA3 cell spikes consistently only when it receives coincident input spikes from a subgroup of MEC cells.

Non-specific feedback inhibition plays the important role of evenly distributing place fields throughout the track, as demonstrated in Figure 4.8. Without this inhibition, most place fields would form in the initial segment of the track due to the fast learning rule, but feedback inhibition limits the amount of activity at any one location on the track.

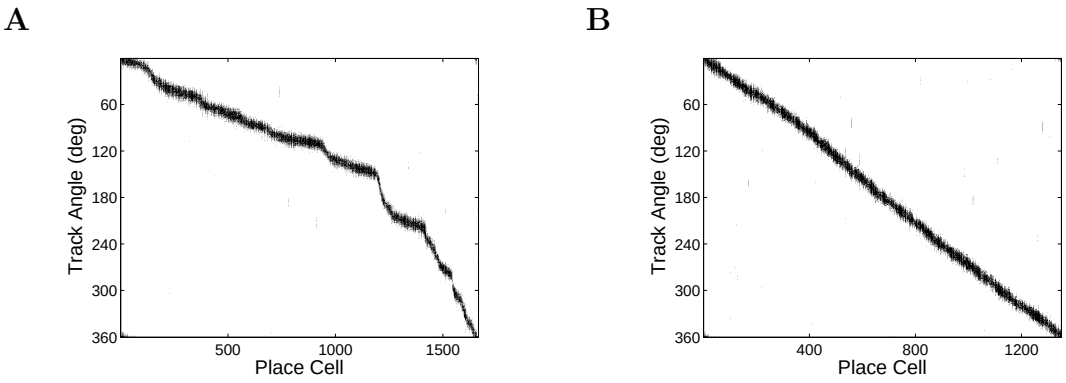


Figure 4.8: Place cell activity without **(A)** and with **(B)** nonspecific feedback inhibition. Place cells are sorted according to the center of mass (COM) of their place fields, and shaded regions correspond to locations on the track at which the given cell has an average spike rate of at least 2 Hz. Nonspecific feedback inhibition causes place fields to spread evenly throughout the track.

Plasticity of Connections from the LEC To model the double rotation experiment, I assume there is a weak coupling between a given place cell and an LEC cell whose firing rate is slightly higher during its place field. As place fields form dynamically during the training session, this weak coupling is achieved through the same plasticity rule used to generate the strong coupling between MEC cells and place cells with correlated activity. The learning rule is robust to the high noise level in LEC cells given a relatively slow learning rate (between 10% and 40% of the learning rate for MEC connections).

Noise from the LEC has little effect on the transformation from grid fields to place fields as long as the place cell activity driven by this noise is significantly weaker than its activity during its place field. As LEC connections become stronger, place fields become increasingly unstable, and multiple fields begin to form. As described in §4.4, however, this instability can be controlled by targeted inhibition.

4.3.2 Backward Shift

Place fields shift in the counter-clockwise direction as the rat moves clockwise around the track (see §1.3.2), possibly biasing place cells for the local cues that are also rotated counter-clockwise. The place cells of the model also show this backward shift in their firing fields as STDP changes the recurrent weights among place cells.

Backward Shift in a Simplified Ring To develop intuition, consider a simple ring of 120 place cells in which each cell i connects to cells $i-1$ and $i+1$, the weights being controlled by STDP. Cell i is driven by an excitatory current injection from an external source (presumably the MEC) when the angle of the rat's position is within the interval $[3(i-1), 3i)^\circ$, where the rat spends 50 ms within each 1° of the track. Figure 4.9 shows the connections and backward shift seen in this simplified setting.

As the rat traverses the track, cell i consistently spikes just prior to cell $i+1$. STDP then drives the weights to a feedforward structure in which the weight from cell i to cell $i+1$ is strong while the weight from cell $i+1$ to cell i is weak. As a result, each cell receives a larger stimulus from the place cell network just prior to its

own place field, causing the cell to spike sooner as its potential is closer to threshold at the onset of its external stimulus.

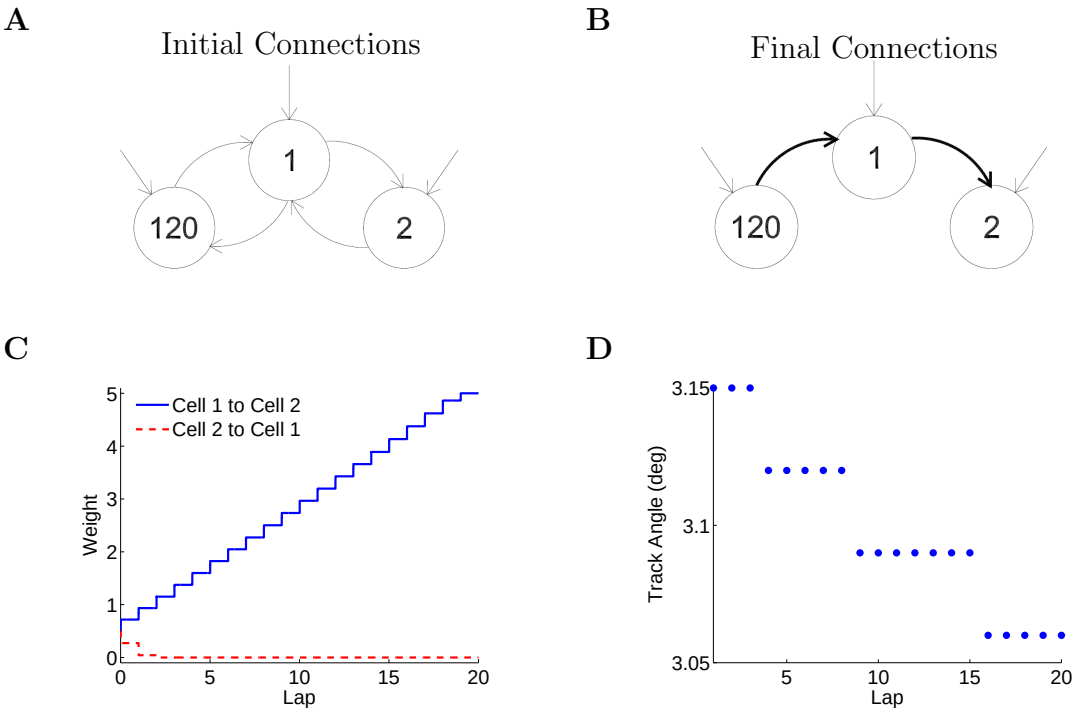


Figure 4.9: Simplified example of the backward shift. This simple network consists of a ring of 120 place cells governed by IAF (§4.2.3), each receiving an external current injection within its distinct place field. **A.** Initial connections. Each cell connects to its two neighbors and receives an external stimulus. **B.** Connections after 20 laps. STDP annihilates the feedback weights, while the feedforward weights are increased to their upper bounds. **C.** Example of the weight changes for each lap. **D.** Backward shift in cell 2. The angle at which cell 2 emits its first spike is plotted for each lap. Each place cell shifts backward due to increased input from the place cell preceding it. Place fields stabilize when the weights reach their bounds.

Backward Shift in the Full Model The full model is considerably more complex than the simplified ring setting described above. As grid cells form multiple firing fields and LEC cells have a high noise level, place cells receive synaptic input throughout the track. Furthermore, weights from external inputs (MEC and LEC

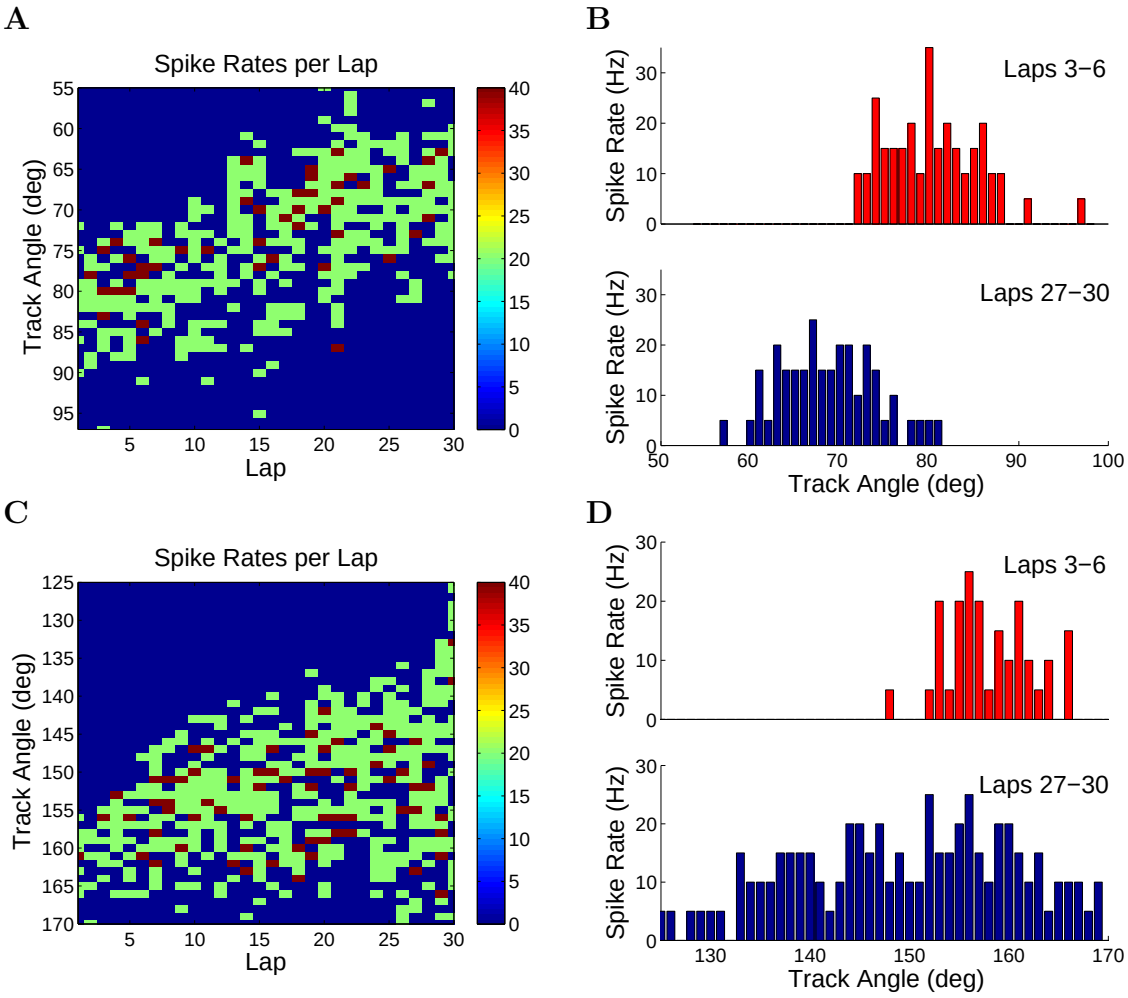


Figure 4.10: Examples of the backward shift of individual place fields. **A.** Spike rate per degree averaged over each lap. **B.** Spike rates averaged over laps 3-6 (top) and laps 27-30 (bottom). The cell was silent for angles not shown. The COM for this cell shifted from 80.3° to 68.7°, giving it a shift of −11.6°. **C-D.** Same as **(A-B)**, but for a second cell whose place field grows wider throughout the training session. The COM for this cell shifted from 158.4° to 149.8°, giving it a shift of −8.6°.

cells) are plastic as well as recurrent weights, affecting the dynamics as fields begin to shift. The recurrent weights are also more complex. Far from a ring structure, place cells synapse onto a random subset of place cells regardless of the proximity in their place fields. Perhaps the biggest challenge to the backward shifting is in the overlap of place fields, resulting in the loss of monotonicity in the changing weights.

Despite this complexity, a backward shift is consistently seen in the model given STDP among recurrent weights. Figure 4.10 illustrates two types of backward shifting seen in individual place cells. A backward shift in the place cell population is dependent on the maximum strength of connections among place cells, as demonstrated in Figure 4.11. Without potentiation of recurrent weights, the population shows no shifting in the place fields. Given a symmetric Hebbian plasticity rule rather than the asymmetric STDP rule, fields tends to shift forward, in the same direction as the rat's movement (data not shown).

4.4 Emergence of a Cell Assembly

According to the theory introduced by Donald Hebb, an ensemble of place cells representing a given environment forms a cell assembly in which the repetitive sequence of activated place cells is stored in the recurrent connections among place cells (Hebb [32]). A cell assembly forms dynamically in the model during the training session, and it adapts to the partially remapped place cell representation of the mismatch session.

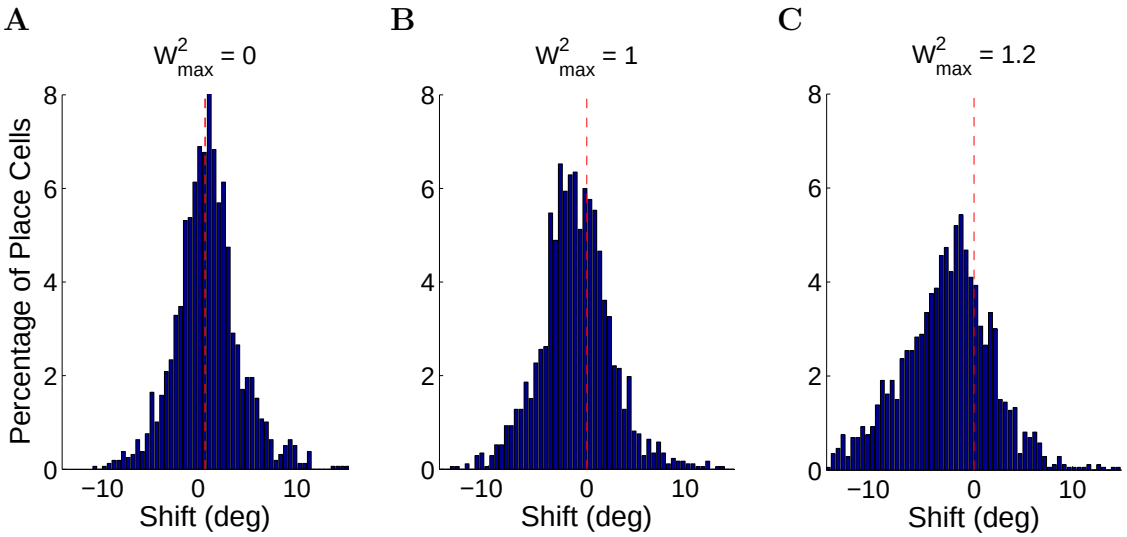


Figure 4.11: Shift in the population of place cells as a function of the maximum weight among place cells. **A.** No STDP is applied, and fields show no net shift. **B-C.** When STDP is applied, fields shift backward according to the strength of the recurrent connections.

As described in §4.1, STDP applied to W^2 encodes the sequence in which place cells are activated, where an active place cell excites other place cells in the cell assembly that are subsequently activated along the learned route. Rate-based plasticity applied to W^5 enforces stability on the cell assembly, causing an active place cell to effectively inhibit place cells with uncorrelated activity. This enforces stability within the cell assembly as the assembly suppresses any activity that disagrees with the learned sequence of activation. In addition, the plasticity rule is gated by the presynaptic spike rate to enforce stability on the place cells included in the cell assembly, where the assembly inhibits all silent cells to create a stable, sparse representation of the environment among the place cell population. When place fields disappear in the mismatch session, new place cells can become active due to the release in inhibition, forming a new cell assembly of the modified environment. I refer to the inhibitory

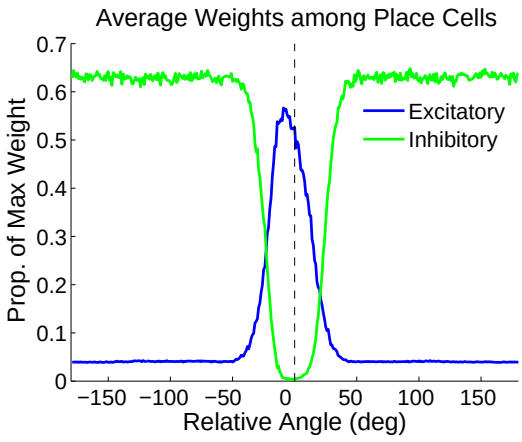


Figure 4.12: Structure of recurrent connections among place cells according to Eq. 4.7. Neighboring place cells excite one another, and place cells with non-overlapping fields inhibit one another. The asymmetry in STDP shifts the excitatory weights such that the average place cell receives its largest recurrent excitation from place cells with COMs just prior to its own.

facet of the cell assembly as *targeted inhibition.*

To determine the emerging structure among recurrent weights, I compute the average weight between two place cells as a function of the difference in their place field COMs, given by Eq. 4.6. For a given place cell k, define the set $N_k(\Delta d)$ by

$$N_k(\Delta d) \equiv \{j : c_j - c_k \in [\Delta d - 1,\ \Delta d)\},$$

where c_j and c_k denote the place field COM for cell j and cell k, respectively. The average excitatory weight among the network of active place cells at relative angle Δd is then given by

$$w^2_{avg}(\Delta d) \equiv \left(\sum_{k=1}^{N_{plc}} \sum_{j \in N_k(\Delta d)} W^2_{jk} \right) \Big/ \left(\sum_{k=1}^{N_{plc}} |N_k(\Delta d)| \right), \tag{4.7}$$

where N_{plc} is the number of active place cells, $|N_k(\Delta d)|$ is the number of cells in the set $N_k(\Delta d)$, and W^2_{jk} is the excitatory weight from place cell j to place cell k. The average inhibitory weight is defined analogously. Figure 4.12 shows a representative example of the emergent structure of both W^2 and W^5.

4.4.1 Relative Drive of each Subregion

To determine the respective roles of MEC, LEC, and CA3 cells in governing place cell activity, I evaluate the effective input, or drive, of each subgroup as a function of the track angle relative to the place field COM. Let c_k denote the place field COM for a given place cell k, and define the MEC input into cell k at relative angle Δd by

$$I_k^{\mathrm{M}}(\Delta d) \equiv \sum_{j=1}^{N^{\mathrm{M}}} W_{jk}^0 R_j^{\mathrm{M}}(c_k + \Delta d), \tag{4.8}$$

where N^{M} is the number of MEC cells, W_{jk}^0 is the weight from MEC cell j to place cell k at the end of the given session, and $R_j^{\mathrm{M}}(x)$ is the average spike rate of MEC cell j at track angle x. The average MEC input is then given by

$$I_{avg}^{\mathrm{M}}(\Delta d) \equiv (1/N_{plc}) \sum_{k=1}^{N_{plc}} I_k^{\mathrm{M}}(\Delta d). \tag{4.9}$$

LEC and CA3 input are defined analogously.

As shown in the representative example of Figure 4.13, MEC and LEC input signals are greatest at the COM of each place field, while CA3 input is greatest just prior to the COM. This indicates that the recurrent place cell network plays the primary role in causing place fields to shift backward. The input curves also illustrate the degree of spatial selectivity in the MEC and LEC signals through the difference between the peak and baseline of each input curve. As the MEC input is greater than twice the LEC input at its peak, MEC cells exert far greater control on place cell activity than do LEC cells.

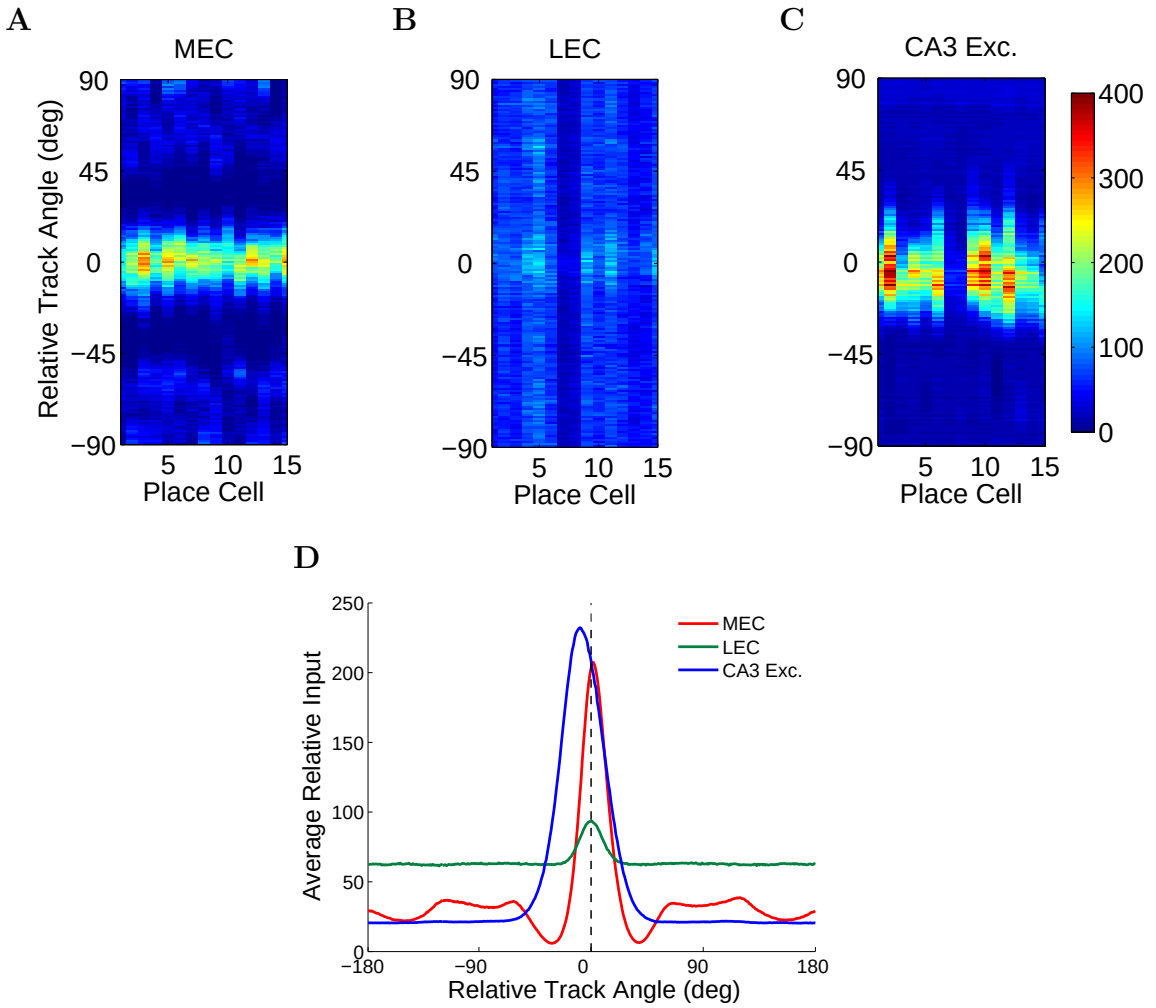


Figure 4.13: Relative input from each subregion. **A-C.** MEC, LEC, and CA3 input (Eq. 4.8) into fifteen place cells as a function of the relative angle between the rat's position and the place field COM. **D.** Average input (Eq. 4.9) from each subgroup into the population of 1200 active place cells, again as a function of the track angle relative to place field COMs. The CA3 input curve is shifted to the left due to the asymmetry in the STDP learning rule, causing place fields to shift backward as they receive input prior to their COMs. MEC cells have a much greater spatial selectivity than do LEC cells, as reflected in the difference between the peak and baseline of the input curves. The agreement between the peak of both curves and the place field COMs indicates that the rule is effective in coupling MEC and LEC cells to place cells with correlated activity. The oscillations in the MEC input curve reflect the periodicity of grid cell firing patterns, while the relatively high baseline in the LEC input curve reflects the high noise level in LEC cells.

Chapter 5

Partial Remapping in the Double Rotation Experiment

Given the strong spatial signal of MEC cells for distal cues, one would expect place fields to rotate with distal cues in the double rotation experiment. The preference for local cues observed among CA3 cells then leads to a search for mechanisms and behavioral paradigms that could give the LEC signal sufficient drive for place fields to rotate with local cues. I use two complementary approaches to investigate place field responses to the cue rotations. I first simulate morphologically accurate pyramidal cells in the CA3 with appropriate ionic currents using the reduced quasi-active model of §3.2. The activity of these reduced cells in a simplified setting of the double rotation experiment is presented §5.1. In §5.2, I use the network model to examine how place cells resolve the competition between MEC and LEC input signals. The

approaches are complementary in that the reduced model incorporates the dendritic computations performed by each cell within a simplified network setting, while the network model incorporates the dynamics among large networks of simplified cells. In the future these two models could be combined to examine the dendritic computations performed by each cell within a large dynamic network.

5.1 Simulations of Reduced Place Cells

To limit the computational expense of simulating large networks of cells, place cells in the network model are passive and isopotential. That is, active ionic currents are neglected, and all synapses are considered to be at the soma. However, the specific distribution of synapses within the dendritic tree affects activity at the soma, as described in §1.4.1. I now simulate a small network of reduced place cells to demonstrate three ways in which neighboring place cells, which form synapses in dendritic regions close to the soma, affect the way in which a place integrates conflicting input signals from the entorhinal cortex arriving in the distal apical tuft.

5.1.1 CA3 Model

All simulations presented in this section are of a CA3 pyramidal cell in the rat with morphology data obtained from neuromorpho.org, cell l48b (Turner et al. [91]). The cell is illustrated in Figure 5.1. The axial current and the ionic currents included are adapted from a computational model specific to CA3 pyramidal cells (Migliore

et al. [59]). The model relies on data and computational studies of CA3 cells from several sources and has been adapted for current studies of CA3 pyramidal cells, including (Hemond et al. [34], Baker et al. [5]).

I linearize the ionic currents that are nonzero in the subthreshold regime to derive the quasi-active model. These currents consist of delayed-rectifier, A-type, and noninactivating potassium currents with a total of five gating variables. All channels are uniformly distributed throughout the dendrites, and the leakage conductance is set to $1/15\ mS/cm^2$. All other parameters are set according to (Migliore et al. [59]).

Distribution of Synaptic Inputs The cell bodies of CA3 pyramidal cells are located in the pyramidal cell layer, or the *stratum pyramidale* (s.p.). Apical dendrites extend perpendicularly from the soma toward the center of the hippocampus, traversing three strata, while the shorter basal dendrites extend in the opposite direction through a single stratum. Dendrites within each stratum contain synapses from distinct subregions. Afferents from the dentate gyrus (DG) synapse onto either the cell body or the proximal apical dendrites in *stratum lucidum* (s.l.). Neighboring pyramidal cells in the CA3 network synapse onto either the basal dendrites in *stratum oriens* (s.o.) or the apical dendrites in *stratum radiatum* (s.r.). Finally, afferents from the entorhinal cortex synapse onto the distal apical tuft in *stratum lacunosum-moleculare* (s.l.-m.), where synapses from the LEC are located in the outer third of the stratum while MEC synapses are located in a narrow band within the middle third (Johnston and Amaral [40], Witter [95]). Figure 5.1 shows the dendritic regions contained in

each stratum as well as the distribution of synapses onto the CA3 pyramidal cell modeled in this work.

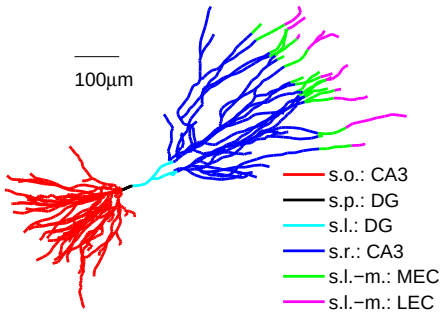


Figure 5.1: Dendritic regions of a CA3 pyramidal neuron. The neuron is modeled through a spatial discretization into compartments of length 0.2 μm, each containing a synapse. Synapses are distributed according to the subregion containing the presynaptic cells, where DG cells form synapses very close to the cell body, CA3 cells synapse onto basal dendrites and proximal apical dendrites, and MEC and LEC cells synapse onto the distal apical tuft, LEC inputs arriving furthest from the soma. Abbreviations indicate the stratum of each dendritic region.

For all simulations presented here, synaptic conductances are given by the alpha function of Eq. 2.1, where $\tau_\alpha = 2\ ms$. CA3 cells integrate synaptic inputs according to the reduced quasi-active model described in §3.2, but a threshold is now enforced on the soma potential such that spikes are generated at the soma when the potential crosses the threshold. According to the integrate and fire model (IAF), all state variables are then held at rest through a refractory period of 3 ms.

Figure 5.2 demonstrates the relative drive of inputs arriving from each subregion as well as the ability of the reduced model to approximate spike times. Figure 5.3 compares a simulation of the quasi-active model to that of the corresponding passive model, demonstrating the accuracy of the reduced system.

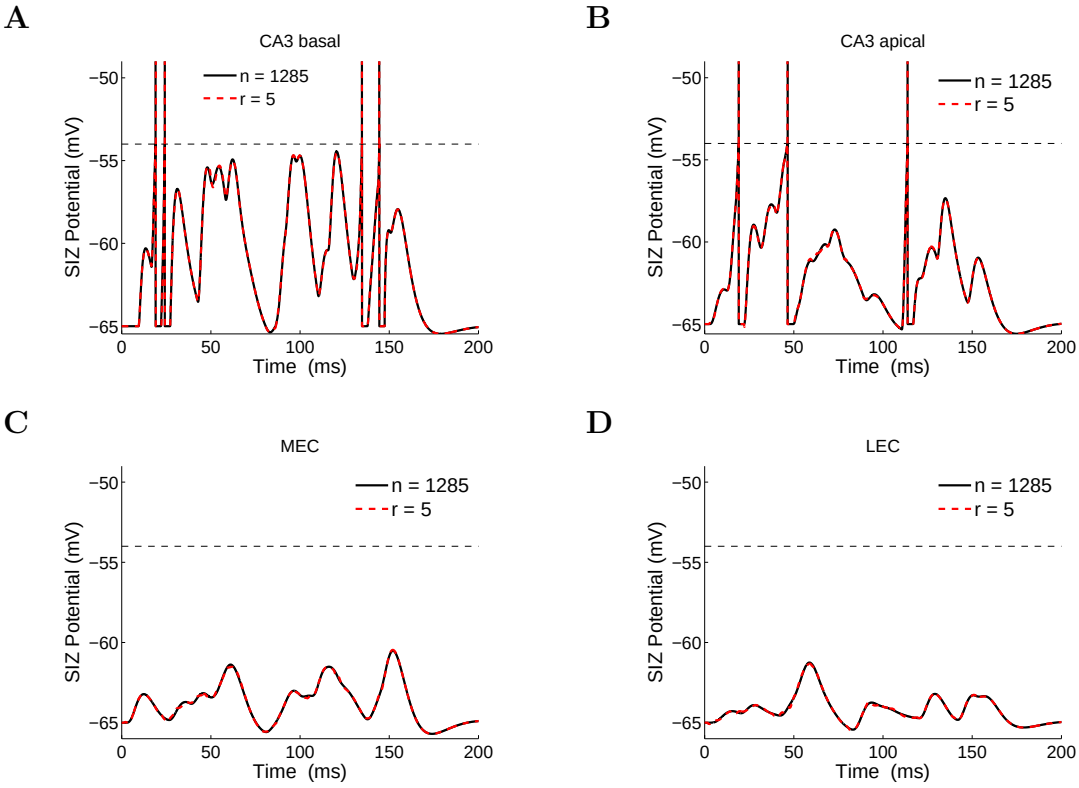


Figure 5.2: Relative response to synaptic inputs from CA3 excitatory cells that synapse onto basal dendrites **(A)** and apical dendrites **(B)**, from MEC cells **(C)**, and from LEC cells **(D)**. For each plot twenty cells within the given subregion were given spike times between 0 and 150 ms at a frequency of 10 Hz, and synapses were selected randomly according to Figure 5.1. The peak conductance for all synapses was 20 mS/cm^2. The reduced model of size 0.4% of the full system size accurately captures the subthreshold potential and spike times given synaptic inputs from any subregion, where CA3 inputs result in a greater depolarization than do comparable inputs from the MEC and LEC.

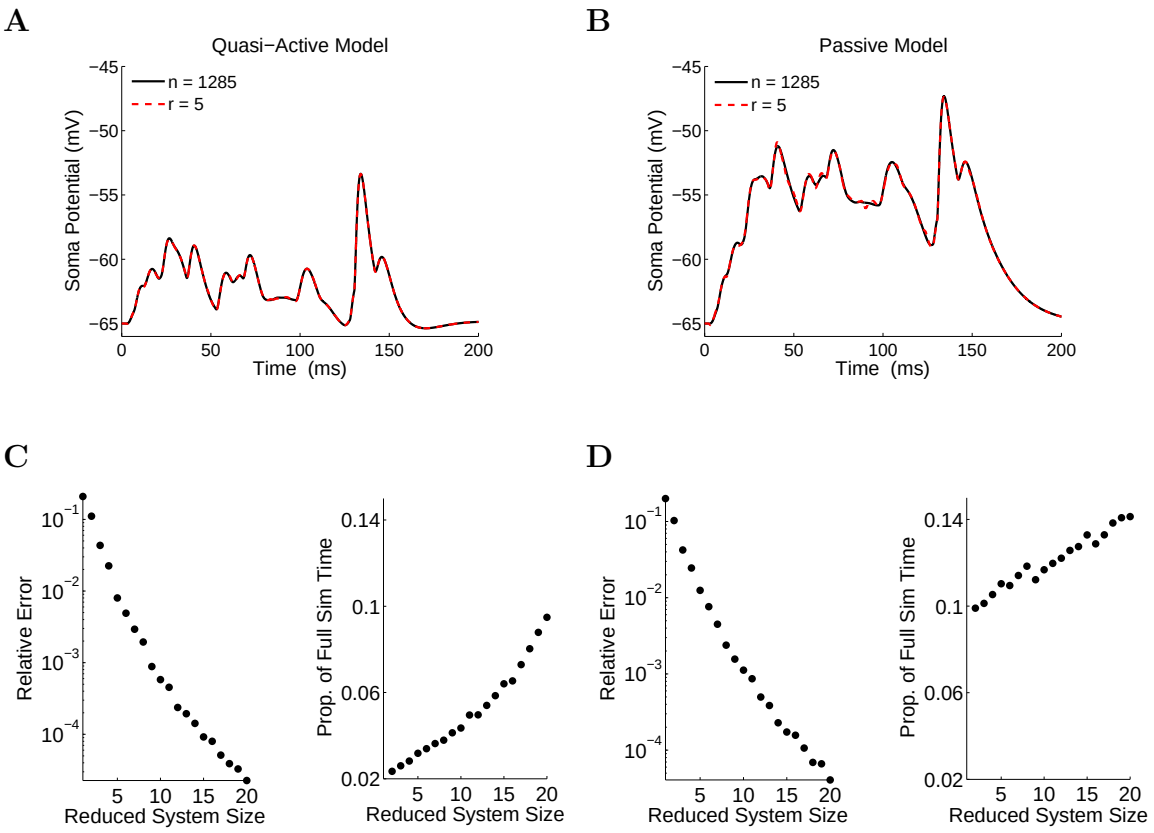


Figure 5.3: Effect of quasi-active currents. **A-B.** Comparison between somatic potentials given the quasi-active model **(A)** and associated passive model **(B)** in which all active currents are ignored. The quasi-active potassium currents act to hyperpolarize the cell. Ten MEC, ten LEC, and twenty CA3 cells were given random spike times between 0 and 150 ms at a frequency of 10 Hz each, where ten CA3 cells synapsed onto basal dendrites and ten synapsed onto apical dendrites. Identical stimuli were used for both the passive and quasi-active simulations. The soma potential for the full system with 1258 compartments (solid black) is compared to the potential for the reduced system with 5 compartments (dashed red). **C.** Relative 2-norm error in the soma potential (left) and simulation times for each reduced system proportional to that of the full system (right) given the quasi-active model. Since the quasi-active model tracks five gating variables and the potential, the number of state variables is given by the number of compartments multiplied by six. **D.** Same as **(C)**, but for the passive model. The simulation time for the full quasi-active and passive models are 22.19 and 4.29 *sec*, respectively.

5.1.2 Simplified Setting

I simulate a small network of reduced place cells in a simplified setting of the double rotation experiment. Each place cell is governed by IAF with a subthreshold potential determined by the reduced quasi-active model, as described in §5.1.1.

Figure 5.4 shows a schematic of this simplified setting. Assume a rat moves clockwise around a circular track of radius r with distinct local and distal cues, where the rat's position, $[r\cos(\theta), r\sin(\theta)]$, is characterized by the track angle, θ, and the rat takes 50 ms to pass each degree on the track. One group of LEC cells and two groups of MEC cells, denoted MEC_0 and MEC_1, provide synaptic input into each place cell. Each input group consists of forty cells spiking at a constant rate of 10 Hz when the rat is within its respective active region of width 30°. Synapses for each place cell are selected randomly according to Figure 5.1.

In the standard session, the place cell has a single, localized place field of width 30° driven by LEC and MEC_1 cells with active regions corresponding to the place field. The weights, or peak synaptic conductances, for LEC cells are weaker than those for MEC_1 cells; consequently, LEC input alone cannot cause the cell to fire, while MEC input can drive a robust place field. Meanwhile, the active region of MEC_0 cells precedes the place field by 90°, but weights from MEC_0 cells are initialized to zero.

In the mismatch session, the active region of each MEC group rotates clockwise by 45° with distal cues, while the active region for LEC cells rotates counter-clockwise by 45° with local cues. The place cell must then resolve competition between a local

and distal field as the active regions of LEC and MEC_1 cells no longer coincide.

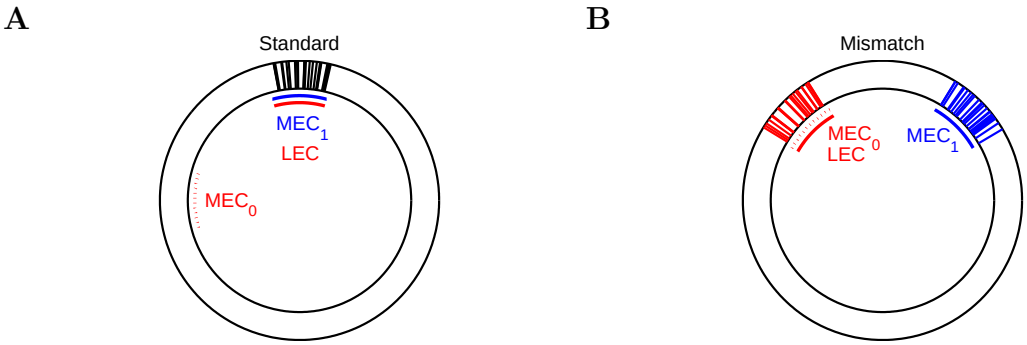


Figure 5.4: Simplified setting of the double rotation experiment. **A.** During the standard session, the place cell has a localized place field driven by weak input from LEC cells and strong input from MEC_1 cells. Black lines on the track indicate spikes of the place cell, and the bars in the interior of the track indicate active regions for each input group. The weights of MEC_0 synapses are initialized to zero, the weights of LEC synapses are set such that LEC input weakly depolarizes the cell, and weights of MEC_1 synapses are strong enough to drive the place field. **B.** During the mismatch session, MEC active regions rotate 45° clockwise with distal cues, while the LEC active region rotates −45° with local cues, forcing the place cell to choose between a local region (red) and distal region (blue).

5.1.3 Place Cell Responses to Conflicting Inputs

Given the relatively weak nature of the LEC input signal, the place cell would consistently form a distal field if the setup of Figure 5.4 were a complete picture of the dynamics involved. However, a variety of plasticity mechanisms and behavioral paradigms have been observed in related experiments, and I examine how a subset of these observed dynamics may combine to cause the counter-intuitive local rotation in CA3 place fields observed experimentally.

Assumption 1 The efficacy of synaptic inputs from the LEC increases transiently upon the rat's initial exposure to the modified environment.

Justification As described in Chapter 4, LEC cells provide contextual information to place cells as they respond to local objects. One application of this theory is that the LEC may provide the sensory feedback signal necessary for error correction in models of path integration. In a similar manner, I assume that LEC cells provide a much stronger teaching signal as the rat attempts to resolve conflicting information regarding its modified environment than when the rat is in a learned, stable environment.

Perhaps the best supporting evidence is in the differential activation of neurons along the transverse axis of the CA1 upon exposure to novel objects. Cells in the CA1 are topographically organized such that LEC cells synapse onto CA1 cells in the distal region of the transverse axis (closer to the subiculum), and MEC cell synapse onto CA1 cells in the proximal region (closer to the CA3). When the rat is initially exposed to novel objects, an increased activity level is observed among CA1 cells primarily along the distal half of the transverse axis (Ito and Schuman [38]).

I assume that, similar to CA1 cells, CA3 cells are also more driven by LEC inputs as the rat pays attention to local cues, which can be interpreted as local objects. This increased efficacy could be due to neuromodulators, synchrony among the field potentials in the LEC and CA3 subregions, or an increased activity level in LEC cells. As the first two mechanisms are beyond the scope of the model, I increase the spike rate of LEC cells transiently at the onset of the mismatch session.

Assumption 2 MEC cells are dynamically coupled to place cells with activity correlated to their own activity, and the connections from grid cells to place cells are quickly updated when place fields remap.

Justification There is no apparent structure in the place cell representation across environments as seemingly random subsets of place cells represent locations within each environment. Grid cells, however, provide a global reference frame as grid cell activation is consistent regardless of the environment. Assuming MEC cells provide the spatial tuning for place cells, connections from the stable spatial map of grid cells to the remapped place cell representation must be dynamic. Furthermore, learning must be fast as place fields are apparent upon the initial exposure to a new environment.

In the network model grid cells adapt to changes in place cell activity in the mismatch session according to the same plasticity rule used to generate place fields in the training session. This fast plasticity is also implemented in the simulations of reduced place cells to update the weights from the MEC_0 and MEC_1 input groups.

Implications I simulate four place cells without recurrent connections to investigate the effect of Assumptions 1 and 2. In the standard session, each cell has a place field of width 30°, where the place fields between cell i and cell $i + 1$ are offset by 5°. In the mismatch session, each cell must choose between a local region over which LEC and MEC_0 cells are active and a distal region over which MEC_1 cells are active. Results are presented in Figure 5.5.

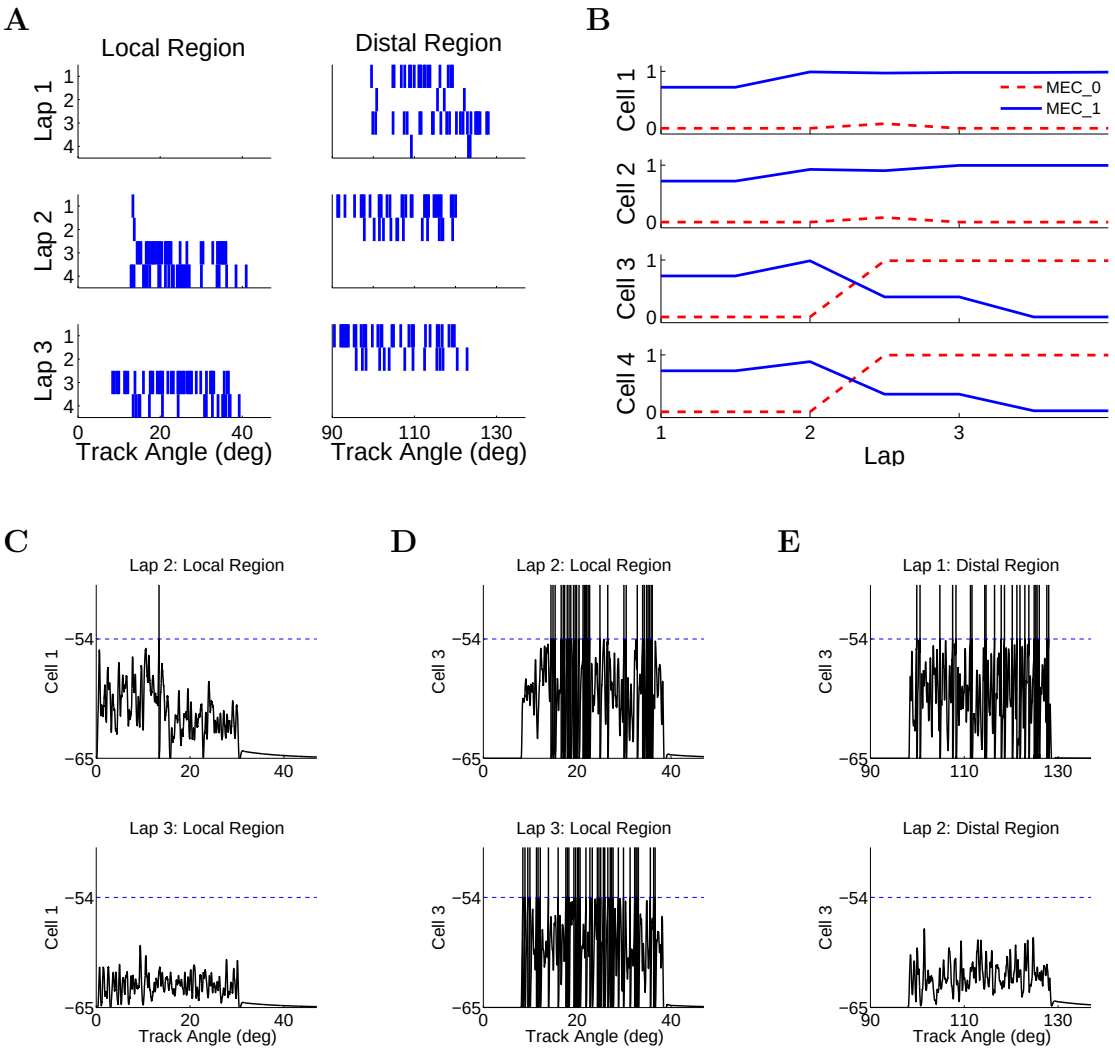


Figure 5.5: Transient increase in LEC activity among isolated cells. Each of four place cells were stimulated by LEC and MEC_0 cells within its local region and by MEC_1 cells within its distal region, where regions for neighboring place cells were offset by 5°. The distal region was simulated first to give no bias to the local region. The spike rate of LEC cells increased by a factor of three during the first 15° of each place cell's local region during lap 2. **A.** Spikes of each place cell within their respective local and distal regions. Variation in synaptic locations causes half the cells to rotate with local cues and half to rotate with distal cues. **B.** Average weights from MEC_0 cells (dashed red) and from MEC_1 cells (solid blue). **C.** Somatic potential of cell 1 during its local region. The cell spiked only once within its local region as the increased LEC activity was insufficient to drive the potential above threshold. When the LEC spike rate was reset to 10 Hz, the cell is weakly depolarized. **D.** Somatic potential of cell 3 during its local region. Increased LEC activity causes multiple spikes in cell 3; consequently, weights from MEC_0 cells increase, resulting in a place field seen in the local region during lap 3 despite the weak LEC drive. **E.** Somatic potential of cell 3 during its distal region. Weights from MEC_1 cells decrease due to the burst of spikes in the local region, silencing the activity in the distal field during subsequent laps.

If LEC cells had a constant spike rate, then all place cells would form distal fields since MEC_1 weights are strong, MEC_0 weights are zero, and LEC weights are too weak to drive the potential to threshold. Two of the four cells simulated, however, form a local place field when the LEC spike rate is increased from 10 Hz to 30 Hz during the initial 15° of each cell's local region. The increased LEC activity causes cells 1 and 2 to spike only once during the local region; consequently, the cells form distal place fields. On the other hand, cells 3 and 4 have a burst of activity during the initial half of their local regions. As the plasticity rule is postsynaptically gated, this burst causes weights from the active MEC_0 cells to increase while weights from the silent MEC_1 cells decrease, resulting in local place fields.

It is important to understand the respective roles of LEC and MEC cells in this simplified setting. MEC cells provide the primary spatial drive throughout the simulations, and LEC cells have a strong impact only during the first 15° of the rat's initial pass through the local region. When coupled with the fast learning rule, the burst of spikes in place cells due to increased LEC activity resets the weights such that the place field is controlled by MEC_0 cells instead of by the MEC_1 cells it followed in the standard session. Thus, place fields rotate with LEC cells and the local cues despite being driven by MEC cells rotating with distal cues.

The simulation also demonstrates one effect of spatially distributed synapses. Each of the four place cells received statistically identical synaptic input from each subregion, but half the place cells formed a local field while the other half formed a

distal field. The only difference between the cells was in the synaptic locations, which were chosen randomly for each cell within the appropriate region. When the LEC rate increased, the relative locations of activated synapses could result in subthreshold activity (cells 1 and 2) or a burst of spikes (cells 3 and 4).

I use the reduced model to generate Figure 5.5 due to its accuracy and reduced simulation time. Figure 5.6 shows the agreement between the reduced system and the full system given the same experimental setup described above.

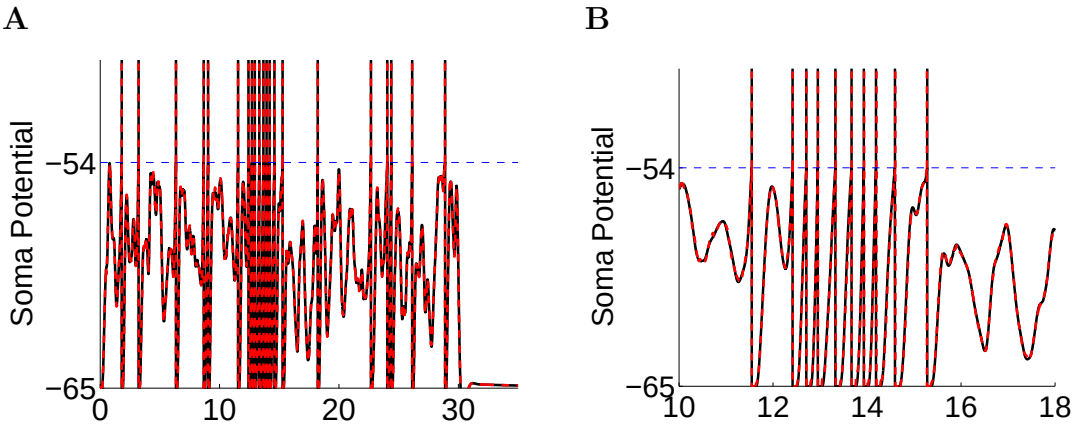


Figure 5.6: Comparison between the full system with 1285 compartments (solid black) and reduced system with 5 compartments (dashed red) for a single place cell given the experimental setup of Figure 5.5. The cell forms a local field as increased LEC activity effectively resets the weights from MEC cells. **A.** The reduced system accurately captures all dynamics of Figure 5.5, for which only reduced cells were simulated. **B.** Burst of spikes for the place cell occurring between 10° and 18° of **(A)**.

Assumption 3: Place cells with correlated activity excite one another (see §4.4).

Implications In the network model, structure emerges in the recurrent connections among place cells due to STDP, where place cells are the most strongly driven by cells with place fields immediately preceding their own (Figure 4.12). I examine the

effect of this emergent structure in the simplified setting by explicitly enforcing a feedforward structure on the recurrent weights. I simulate eight place cells such that during the standard session, each cell has a place field of width 30°, place fields of neighboring cells are offset by 5°, and each cell i forms twenty synapses onto cell $i+1$. Multiple synapses are used to depolarize a wider area of dendrites, where each place cell represents a group of twenty place cells responsive to the same location on the track. Synaptic locations are uniform across all cells. Results are shown in Figure 5.7.

The first effect of recurrent connections is an increased coherence in the place field rotations. If the connections are strong enough, which is expected since the synapses are proximal to the soma, then they can carry the activity of the first few cells across the network. To demonstrate this effect, I increase the LEC spike rate to 30 Hz within the interval $[0, 20]^\circ$ during lap 1 of the simulation. Cells 1 through 4 are thus stimulated by increased LEC activity for 20°, 15°, 10°, and 5°, respectively, during their local regions, but cells 5 through 8 never see an increase in LEC activity. As shown in Figure 5.7, cells 1 through 4 form local fields while cells 5 through 8 form distal fields if the recurrent connections are weak. When the weights are increased, the burst induced by increased LEC activity propagates through all cells, causing all eight cells to form local fields.

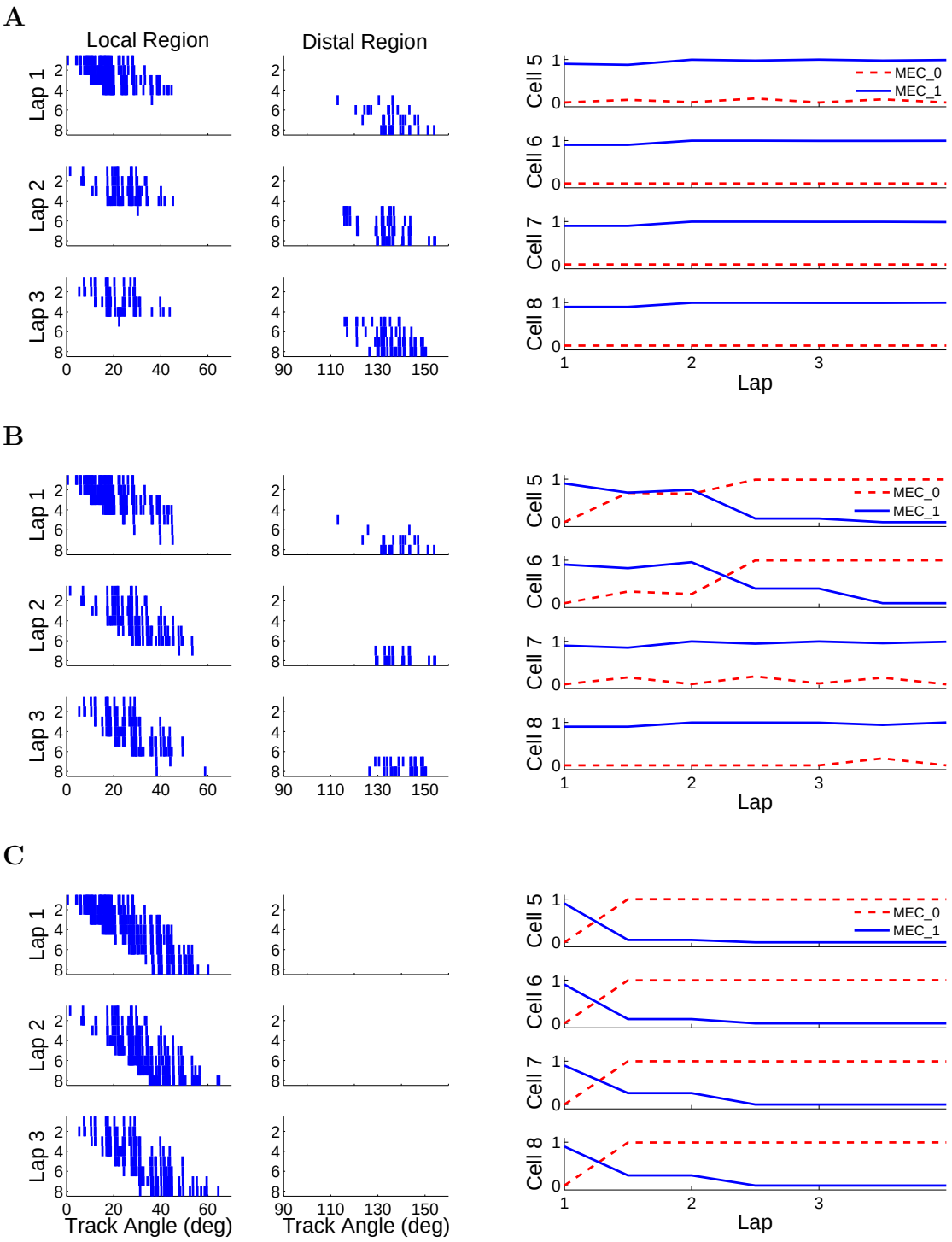


Figure 5.7: Effect of feedforward excitation. Eight place cells are arranged in a feedforward structure such that each cell i synapses onto cell $i+1$, where their place fields are offset by 5°. The firing rate of LEC cells increases from 10 Hz to 30 Hz during the first 20° of lap 1, differentially activating cells 1 through 4 while only weakly depolarizing cells 5 through 8. **A.** When the weight of recurrent connections is set to $\overline{g} = 1.5\ mS/cm^2$, cells 1 through 4 form local fields while cells 5 through 8 form distal fields. **B.** When $\overline{g} = 1.75\ mS/cm^2$, cells 5 and 6 are recruited to form local fields. **C.** When $\overline{g} = 2\ mS/cm^2$, all eight cells form local fields. The synaptic locations and spike times are identical for all three simulations.

Assumption 4 Place cells with uncorrelated activity inhibit one another (see §4.4).

Implications As demonstrated in Figure 5.5, local and distal regions are in competition for the place field of an individual cell when cues are rotated. The targeted inhibition of the network model, which is based on Assumption 4, creates a second mode of competition in which two place cells compete for a single location on the track as they inhibit one another. Similar to recurrent excitation among place cells, targeted inhibition among place cells enforces coherence in the place cell responses.

Figure 5.8 demonstrates this effect. I simulate two place cells that have learned to inhibit each other due to a 90° difference between their place field COMs in the standard session. This targeted inhibition is implemented as described in §4.1. In the mismatch session, in which the local region for cell 1 corresponds to the distal region for cell 2, LEC firing rates are increased to 30 Hz for the first 15° of each cell's local region.

For the example shown, without targeted inhibition the cells share a firing field in the mismatch session as cell 1 forms a local field and cell 2 forms a distal field. When the weights of the inhibitory synapses on each cell are nonzero, cell 1 silences cell 2 during the first lap when cell 1 has a burst of activity driven by increased LEC input. This gives MEC_0 weights onto cell 2 more time to grow while MEC_1 weights become weaker, resulting in a local field for cell 2.

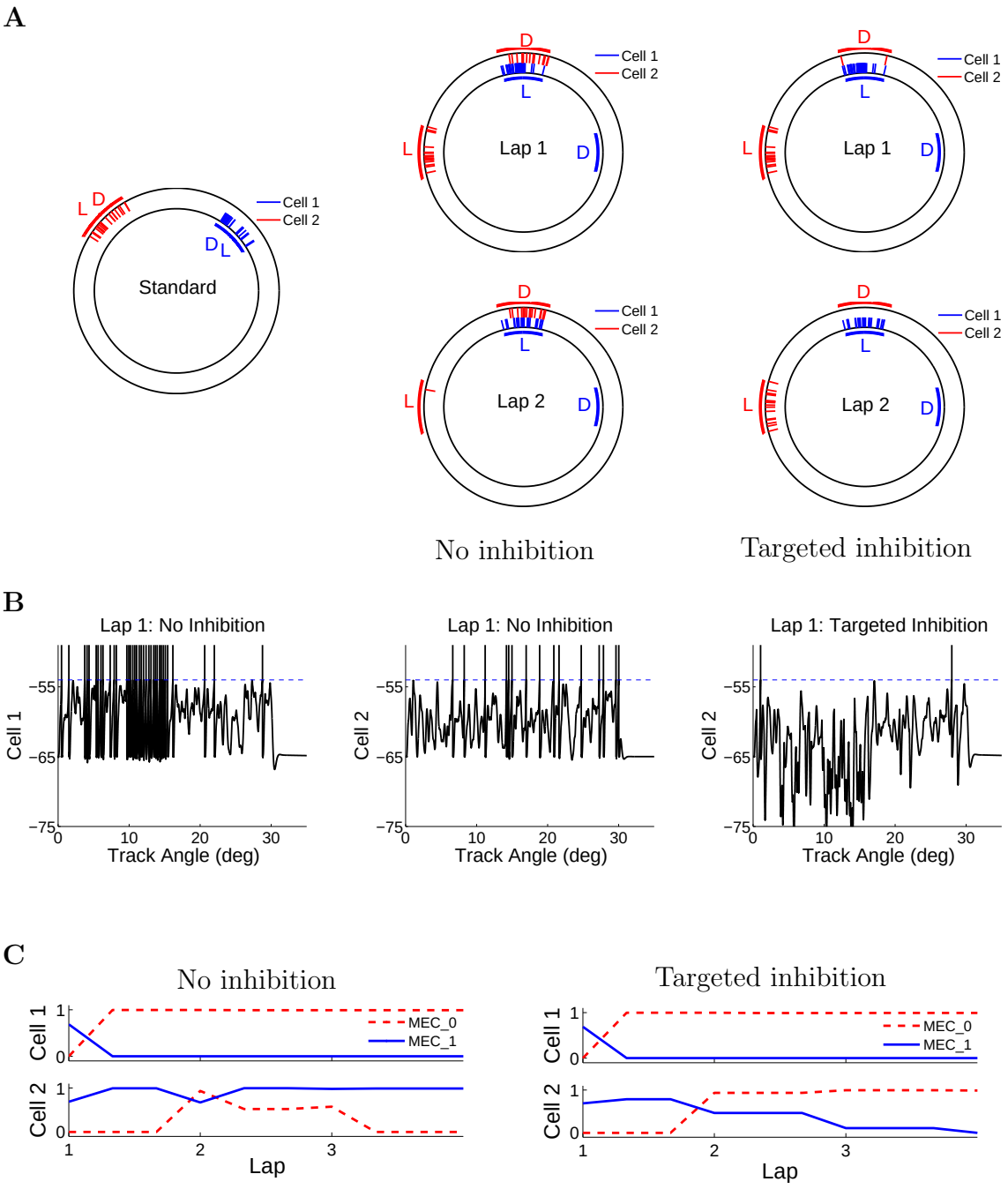


Figure 5.8: Effect of targeted inhibition. **A.** Two place cells inhibit one another as their place fields are separated by 90° in the standard session (left). Lines on the track indicate spikes, and bars indicate the local (L) and distal (D) regions. During the mismatch session, the local region of cell 1 corresponds to the distal region of cell 2, and the cells share a firing field in the absence of inhibition. Given inhibition, competition from cell 1 causes cell 2 to form a local field. The rat begins at the onset of the shared region and moves clockwise around the track. **B.** Somatic potentials. During the first lap, increased LEC activity drives a burst of activity in cell 1 (left), hyperpolarizing cell 2 within its distal region (right). **C.** Given inhibition, MEC_1 weights are held constant during the distal region, and activity during the local region ensures that cell 2 remains silent during subsequent laps.

5.2 Network Responses to Cue Rotations

The network model has the following properties after the simulated training session:

- MEC and LEC cells generate Poisson spike trains with a firing rate dependent on the position of the rat, providing a strong and weak spatial signal, respectively.
- CA3 excitatory cells have localized place fields developed dynamically and directly controlled by synaptic input from the MEC and LEC.
- Coactive place cells excite one another, where a place cell is the most strongly driven by place cells firing slightly prior to its own place field. Place fields shifting backward in the training session have largely stabilized.
- Place cells inhibit one another both through non-specific and targeted feedback inhibition, creating a sparse, stable representation of the standard cue configuration by place fields evenly distributed throughout the track.

MEC and LEC responses to the cue rotations of the mismatch session mimic experimental data, as described in §5.2.1, but CA3 cells continue to integrate synaptic input without any explicit changes due to the modified environment. As place fields are dynamically controlled by MEC and LEC cells, the model can be used to investigate the partial remapping observed in place cells as its two sources of input provide conflicting signals. I examine the effect of the assumptions made in §5.1.3 through simulations of four systems that vary in the nature of their recurrent connections among place cells after the training session.

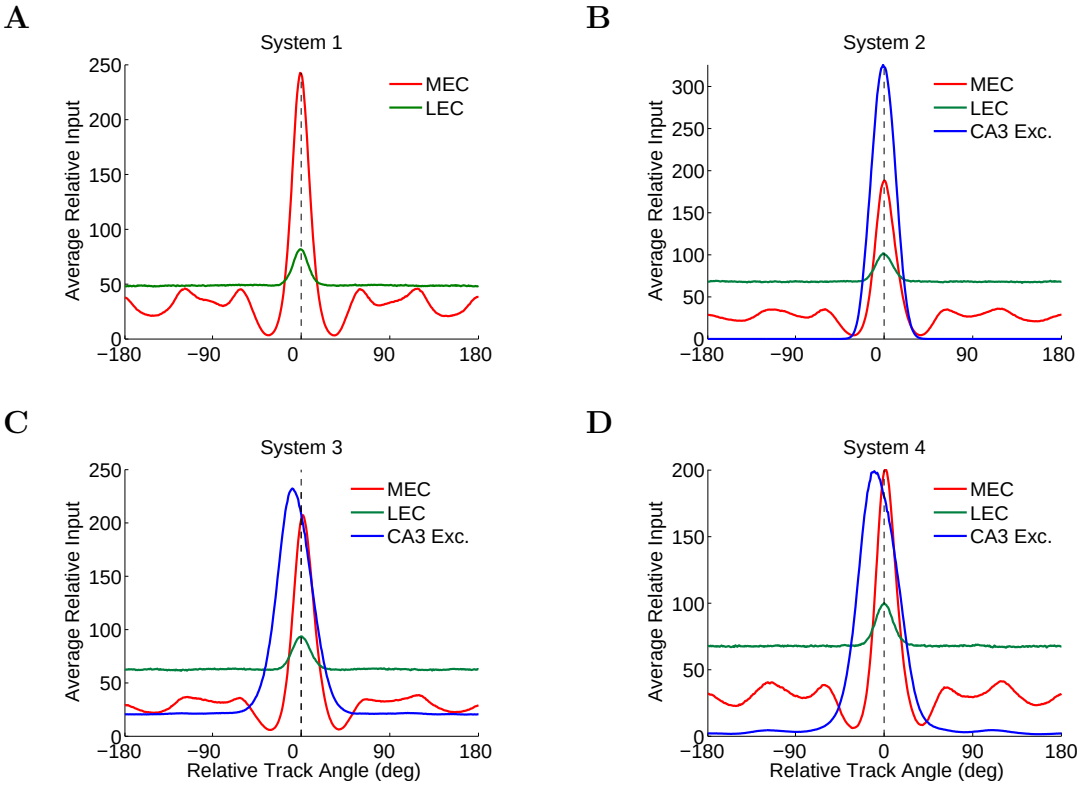


Figure 5.9: Average relative inputs for systems simulated in the standard and mismatch sessions. The input curves are explained in §4.4.1. All four systems statistically receive the same inputs from the MEC and LEC but differ in the nature of the recurrent connections among place cells. Although not shown, targeted inhibition is implemented for Systems 2-4, where the structure of inhibitory weights resembles that shown in Figure 4.12. **A.** System 1 has no recurrent connections among place cells. 60% of place cells are active in the standard session, but place fields continue to form throughout the simulation in the absense of targeted inhibition. **B.** Recurrent connections of System 2 are governed by the symmetric rate-based plasticity rule. 50% of place cells are active, and place fields show a forward shift in the training session. The system is stable during the standard session despite the relatively strong recurrent input since this input is centered over the place fields. **C.** Recurrent connections of System 3 are governed by the asymmetric STDP rule. 50% of place cells are active, and place fields show a strong backward shift in the training session. Place fields are stable in the standard session if the recurrent weights are held constant, but fields become unstable if the weights remain plastic. **D.** System 4 is identical to System 3 except in the learning rate for targeted inhibition, which is 40% faster. 35% of place cells are active, and the backward shift seen in System 3 is tempered by the fast inhibition. Unlike System 3, the system reaches a balance during the training session such that the system is stable in the standard session given the same plasticity rules applied in the training session.

The systems are summarized through the average relative inputs arriving from each subregion, shown in Figure 5.9. I compare the response of each system to the cue rotations when the LEC activity level remains fixed (§5.2.2), increases over the initial lap(s) of the mismatch session (§5.2.3), and increases periodically according to an alternating state hypothesis (§5.2.4).

5.2.1 Input Signals for Mismatched Cues

Data collected from the double rotation experiment indicate that MEC cells respond to the distal cue set (J. Neunuebel [manuscript in preparation]). As grid cells provide a global reference frame in which activation patterns among grid cells are consistent across environments, grid cells respond to cue rotations as a unit in the model.

Grid fields recorded in the experiment do not rotate perfectly with distal cues but seem to shift, causing some grid cells to represent entirely new locations on the track in the mismatch session than they represented in the standard session, as illustrated in Figure 5.10. In the model all grids rotate 45° with respect to the center of the track and shift 6 *cm* along the grid axis. The resulting correlation in the grid cell activity between the standard and mismatch sessions agrees with similar plots generated from experimental data. Figure 5.11 provides examples of individual grid cell responses as well as the correlations of the population's activity.

Correlations in LEC activity recorded during the standard and mismatch sessions show a slight preference for the local cue set. I assume that all LEC cells rotate

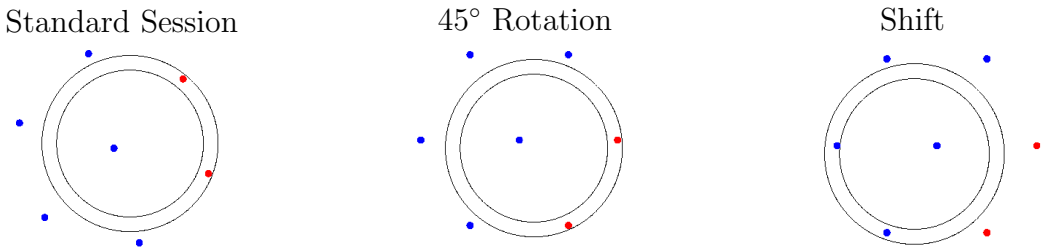


Figure 5.10: Illustration of the grid cell response to cue rotations. Grid cells appear to rotate with global cues and shift upon the cue rotations, causing some grid cells to represent entirely new locations on the track. For this illustration the grid points located on the track in the standard session (shown in red) shift away from the track, and two new firing fields form on the track.

with local cues, which means that the hot spots of Eq. 4.4 rotate $-45°$. To incorporate noise, the base probability, P_0, and the probability difference for each hot spot, $P_{D,i}$, change randomly by up to 100% of the original values. The lower and upper boundaries for each parameter are maintained in the mismatch session. Figure 5.12 provides examples of individual LEC cell responses as well as the correlations in the population's activity, which again agree with experimental data.

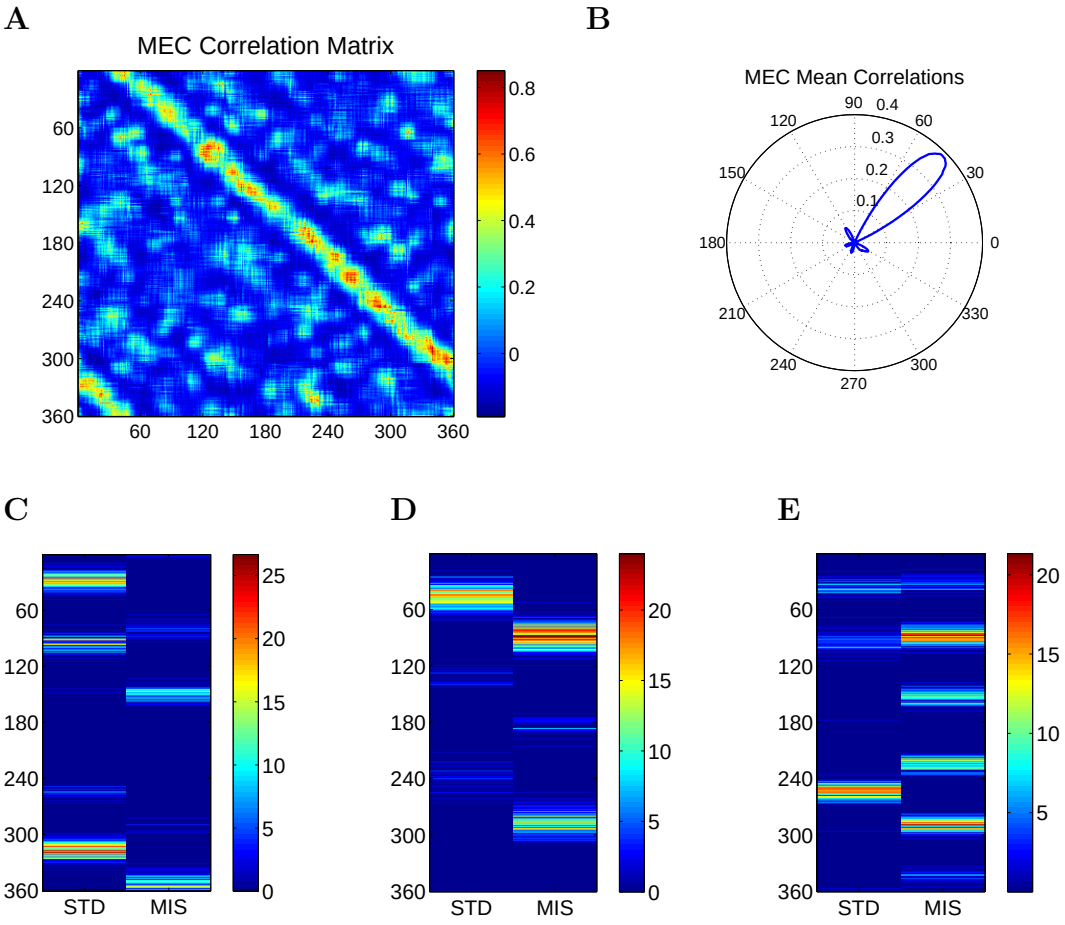


Figure 5.11: MEC response to cue rotations. All grids rotate 45° with the distal cue set and shift 6 *cm*. **A.** Correlation matrix for the firing rates of one hundred MEC cells in the standard and mismatch sessions. The firing rate of each cell was averaged across bins of width 1° of the track for all laps in the given session. **B.** Mean of each diagonal in **(A)**. MEC cells show a clear preference for the distal cues. **C-E.** Representative examples of individual MEC cell responses to the cue rotations. Some cells show a rate remapping **(C-D)**, while others show a global remapping **(E)** due to the grid shift.

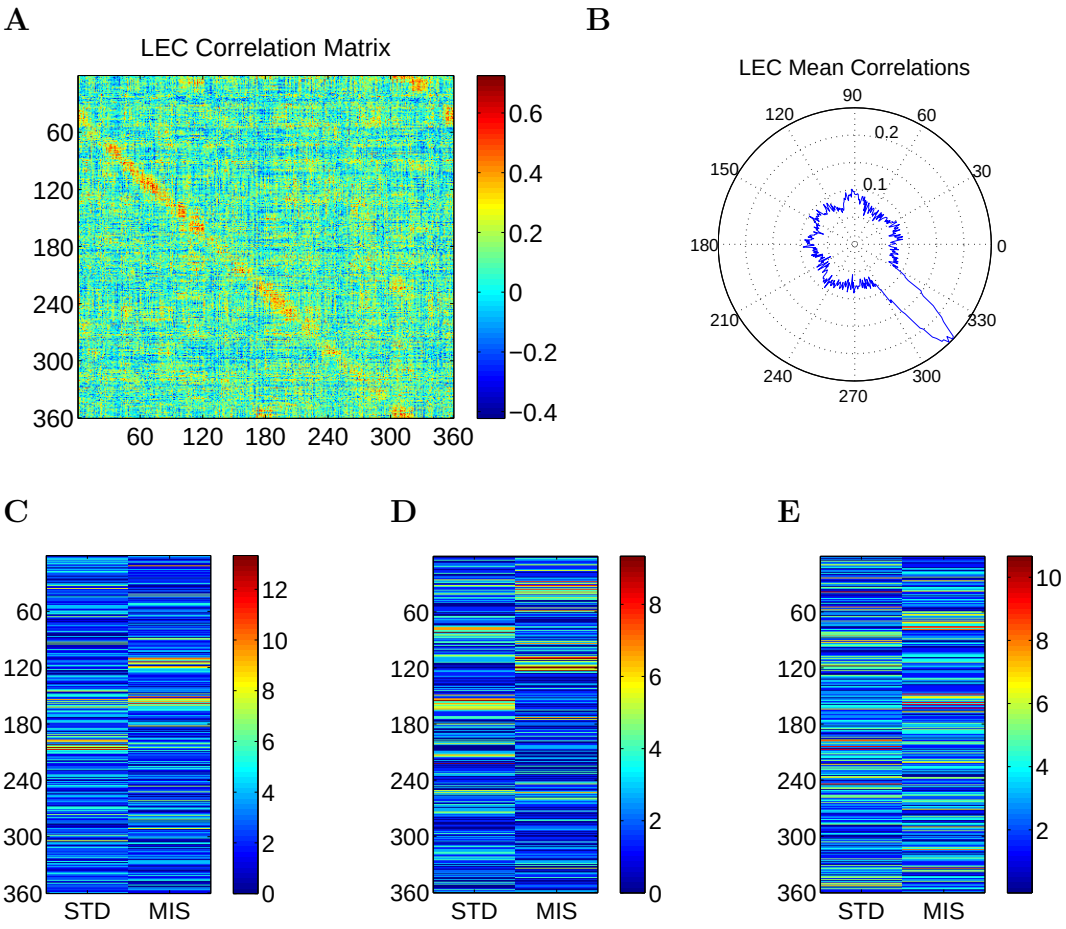


Figure 5.12: LEC response to cue rotations. The hot spots of all firing fields rotate $-45°$ with the local cue set, while the base probability and probability differences, P_0 and $P_{D,i}$, are modified by as much as 100%. Fifty LEC cells were simulated to compare the results with experimental data from 50 active LEC cells. **A.** Correlation matrix for the firing rates of all cells in the standard and mismatch sessions. **B.** Mean of each diagonal in **(A)**. LEC cells show a clear preference for the local cues, although their signal is much more noisy than is the corresponding signal from the MEC. **C-E.** Representative examples of individual LEC cell responses to the cue rotations.

5.2.2 Place Cell Responses

As was done in the experiment, I simulate fifteen laps for the standard session followed by fifteen laps for the mismatch session. Between sessions MEC and LEC firing fields are adjusted according to §5.2.1 to model the cue rotations. For these simulations no additional changes are made to the LEC activity. Figures 5.13-5.15 show the results for Systems 1, 2, and 4, respectively, where the results for Systems 3 and 4 are similar. The dominant response in every simulation is a rotation with distal cues, which is expected given the relative strength of the MEC signal in terms of both its peak input into place cells and its spatial selectivity.

System 1, which has no recurrent connections among place cells, shows the weakest response for distal cues as only 50% of place fields rotate distally. Most place cells remap, where place fields either rotate randomly, disappear, or appear. The weak LEC signal also drives 6% of place cells to rotate with local cues.

Fields rotate much more coherently with distal cues given recurrent excitation and targeted inhibition among place cells. In Systems 2-4, almost 90% of place fields rotate distally. Even more telling, no fields rotate with local cues or to a random location. When the recurrent weights are symmetric due to rate-based plasticity, place fields show a strong forward shift similar to that seen in the training session, on average bypassing the distal cues by almost 20°. On the other hand, the asymmetric recurrent weights result in rotations centered at the distal cues.

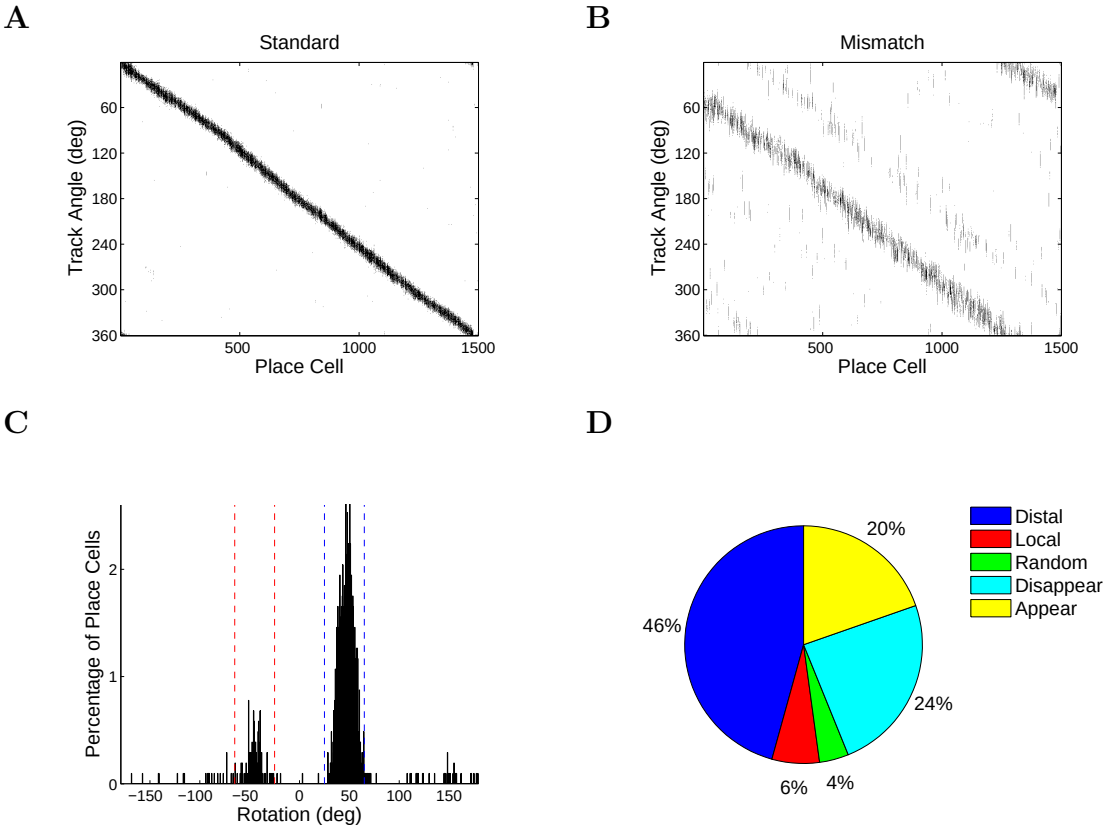


Figure 5.13: Mixed response in System 1. Without recurrent connections, half the fields rotate with distal cues while most other fields remap. **A.** Place cell activity during the standard session. Shaded regions correspond to locations on the track at which the given place cell has a spike rate greater than 2 Hz. Place cells are sorted according to their place field COMs. **B.** Place fields during the mismatch session, where place cells are ordered as in **(A)**. The band of activity below the diagonal corresponds to place fields that rotated with distal cues, while the band above the diagonal corresponds to fields that rotated with local cues. **C.** Histogram of rotations among the population, where a rotation is defined to be the difference between the place field COM in the standard and mismatch sessions. **D.** Responses of all place cells. *Distal* and *local* refer to place field rotations within the dotted blue and red lines of **(C)**, respectively, and all other rotations are classified as *random*. *Disappear* and *appear* refer to place cells that are active in only one session.

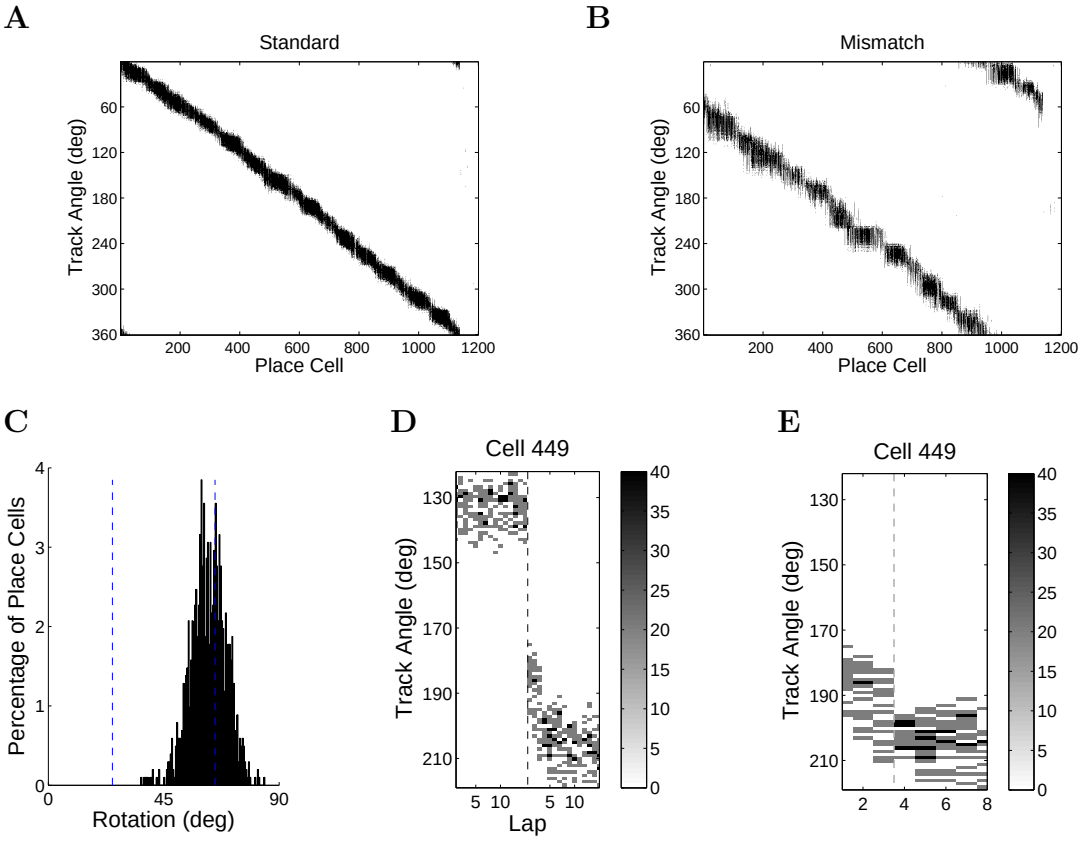


Figure 5.14: Over-rotation in System 2, for which recurrent weights are symmetric. **A-B.** Roughly 90% of place fields rotate with distal cues, although the representation of the track is not as smoooth. **C.** Place fields over-rotate, on average bypassing the distal cues by almost 20°. There were no rotations for the angles not shown. **D.** Average spike rate of a representative cell computed for each lap. The dotted line separates the standard and mismatch sessions, and the colorbar indicates the spike rate (Hz). **E.** First eight laps of the mismatch session shown in **(D)**. The forward shift occurs during the initial few laps of the mismatch session, and the activity during the first lap is centered at the distal cue.

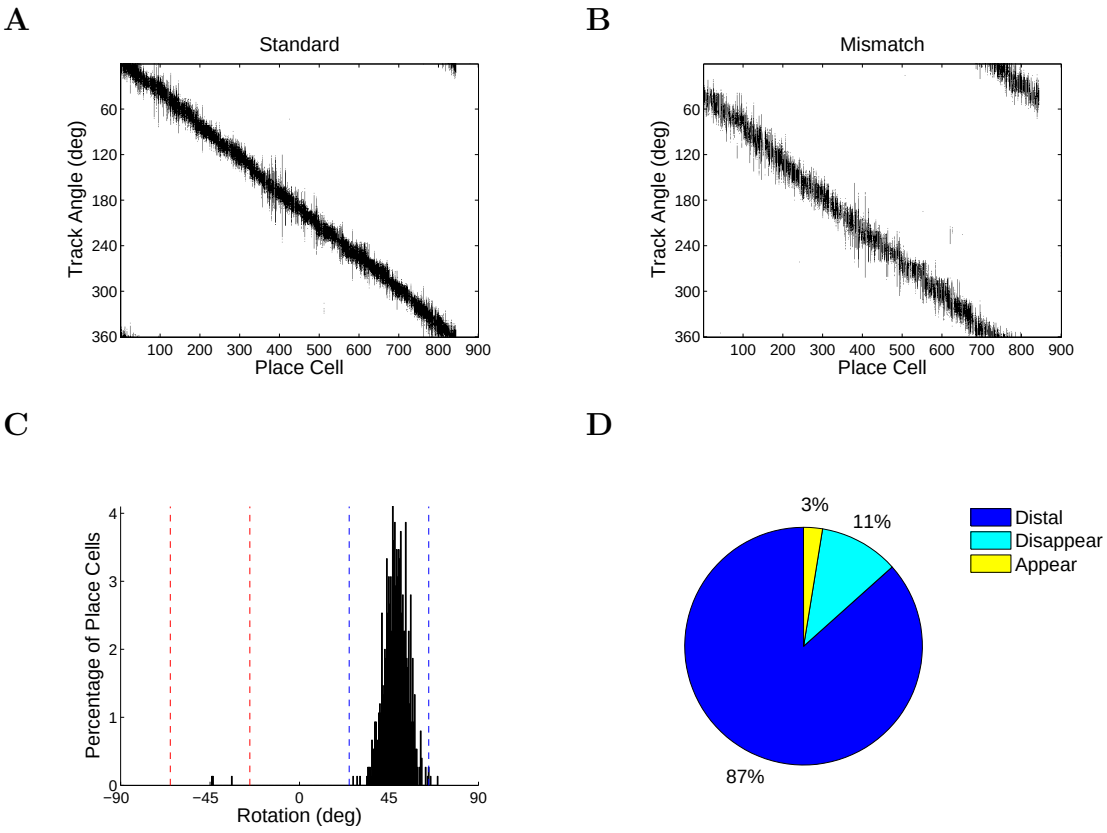


Figure 5.15: Centered distal rotation in System 4, in which recurrent weights are asymmetric. **A-B.** The band of place cell activity rotates with distal cues with very little distortion. **C.** Place field rotations are centered at the distal cues. **D.** Recurrent connections cause the place cells to rotate as a unit with distal cues.

5.2.3 Effect of an Increased LEC Spike Rate

I next apply Assumption 1 of §5.1.3 by increasing the LEC activity level during the initial *response laps*, which consist of a set number of laps immediately following the cue rotations. Response laps are not part of the double rotation experiment but allow one to examine the activity given heightened LEC input separately from the activity during the mismatch session, which consists of fifteen laps over which the LEC activity is returned to its normal level.

For these simulations I increase the LEC activity level by multiplying the maximum frequency, f_{max}, by a factor of three, where f_{max} is used to generate spike trains as described in Algorithm 3 of §4.2.1. This effectively increases the LEC spike rate evenly throughout the track, giving LEC cells no increase in their spatial selectivity. Figure 5.16-5.18 show the results for Systems 1, 3, and 4, respectively. The results for Systems 2 and 3 are similar, but a forward shift is again seen in System 2. Figure 5.19 compares the simulation of System 4 to an identical simulation for which all weights are held constant after training, isolating the role of plasticity mechanisms.

Without recurrent connections, place cells exhibit a mixed response as approximately the same number of place fields rotate with distal and local cues while most cells remap. A mixed response is seen regardless of the number of response laps simulated, where the LEC activity was increased for five response laps for the results presented in Figure 5.16. This mixed response resembles the response observed in CA1 place cells, which do not have recurrent connections.

Given recurrent connections, increasing the LEC activity by a factor of three consistently causes most fields to rotate with local cues. This local rotation depends on the extent to which LEC activity increases. If the rate were increased by a factor of two, distal rotations would be dominant. Likewise, the local rotation becomes stronger the more response laps are simulated, as demonstrated in Figures 5.17 and 5.18.

Due to the instability of System 3, excitatory recurrent weights among place cells are held constant during the standard session but are again governed by STDP during the response and mismatch sessions. The brain has homeostatic mechanisms that could reduce the amount of plasticity within a learned environment, allowing the system to then adapt to the modified environment when cues are rotated. The plasticity mechanisms are unchanged throughout the simulation for System 4 as the fields are stable during the standard session. The responses of Systems 3 and 4 are very similar, but System 4 resolves the competition between local and distal fields more quickly.

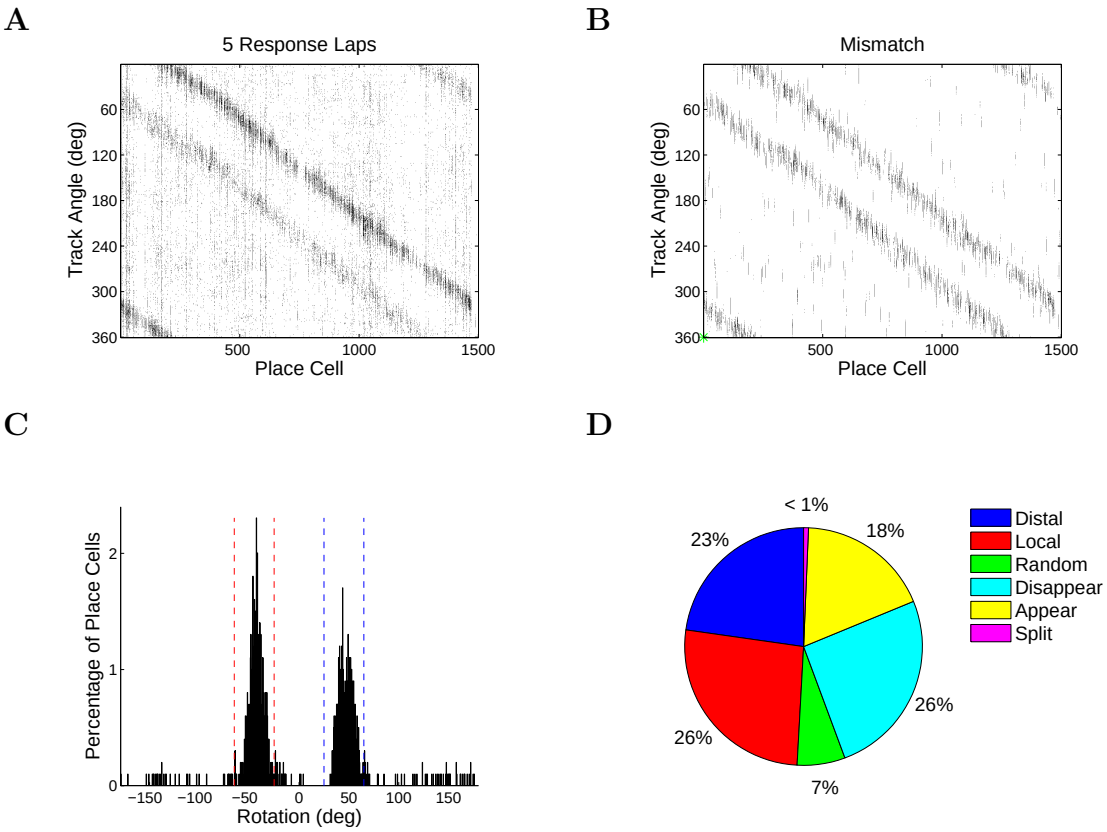


Figure 5.16: Mixed response in System 1, for which there are no recurrent connections. The LEC spike rate was increased by a factor of three for five response laps. **A.** Place cell activity during the response laps. Cells are sorted according to place field COMs in the standard session, shown in Figure 5.13A. The bands below and above the diagonal correspond to the distal and local cues, respectively. Although increased activity is seen at the local and distal cues, many cells are active throughout the track. Recurrent connections largely silence this noise, as demonstrated in Figure 5.19. **B.** Place cell activity during the mismatch session. **C.** Rotation histogram. **D.** Place cell responses. The local and distal rotations are roughly even, and most cells remap. *Split* refers to place cells that split their representation, forming more than one firing field. Most cells with split fields are responsive to both local and distal cues.

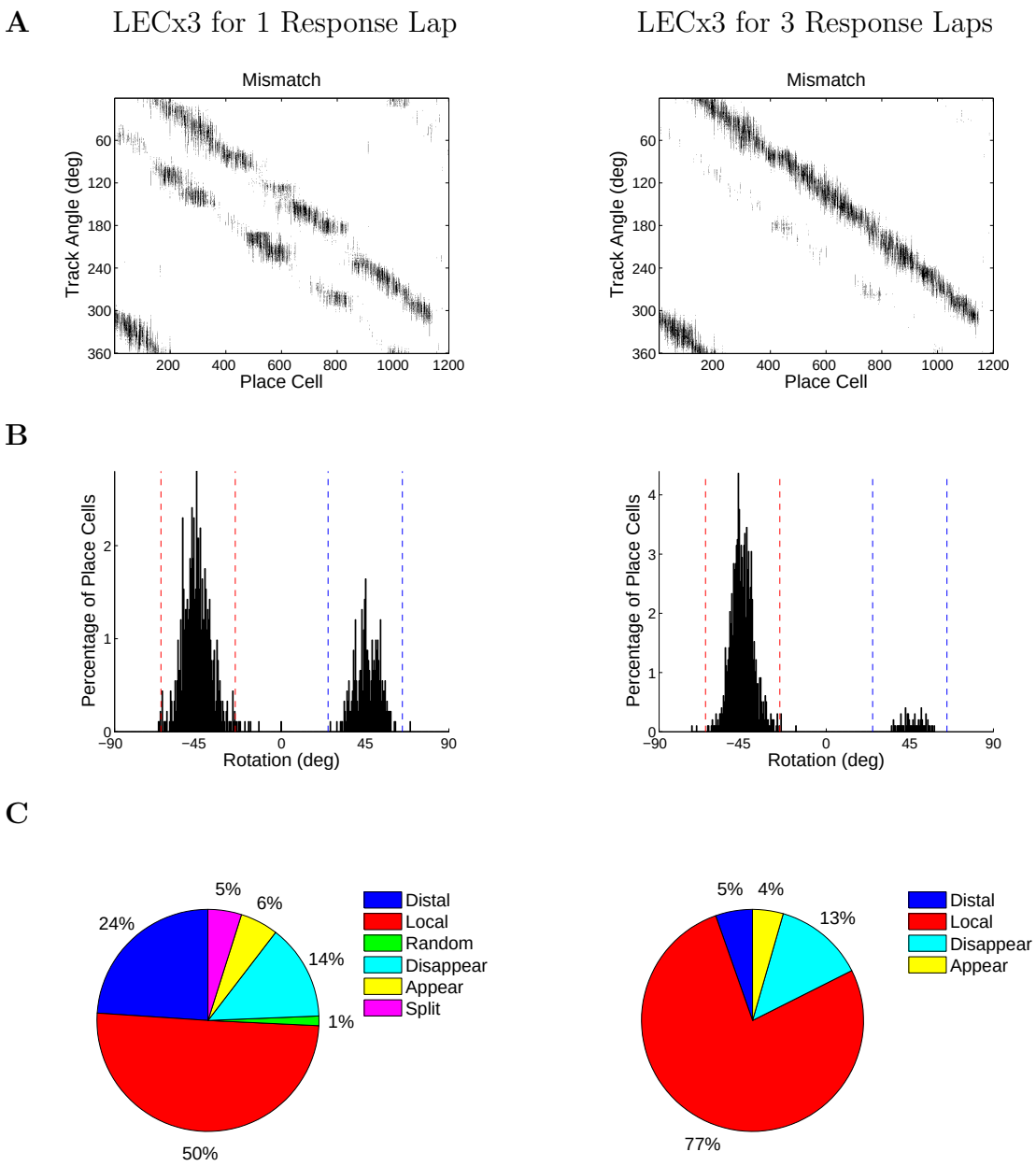


Figure 5.17: Local response in System 3. Given recurrent connections, increased LEC activity for one response lap (left) or three response laps (right) causes local rotations to be the dominant response. **A.** Place cell activity during the mismatch session. During the standard session all fields were along the diagonal. Both local and distal rotations are seen given one response lap, but the distal band is silenced during two additional laps with increased LEC input. **B.** Rotation histogram. Rotations are centered over the cues. **C.** Place cell responses. 5% of place fields split their representation between local and distal cues given a single response lap, but the competition is resolved given three response laps.

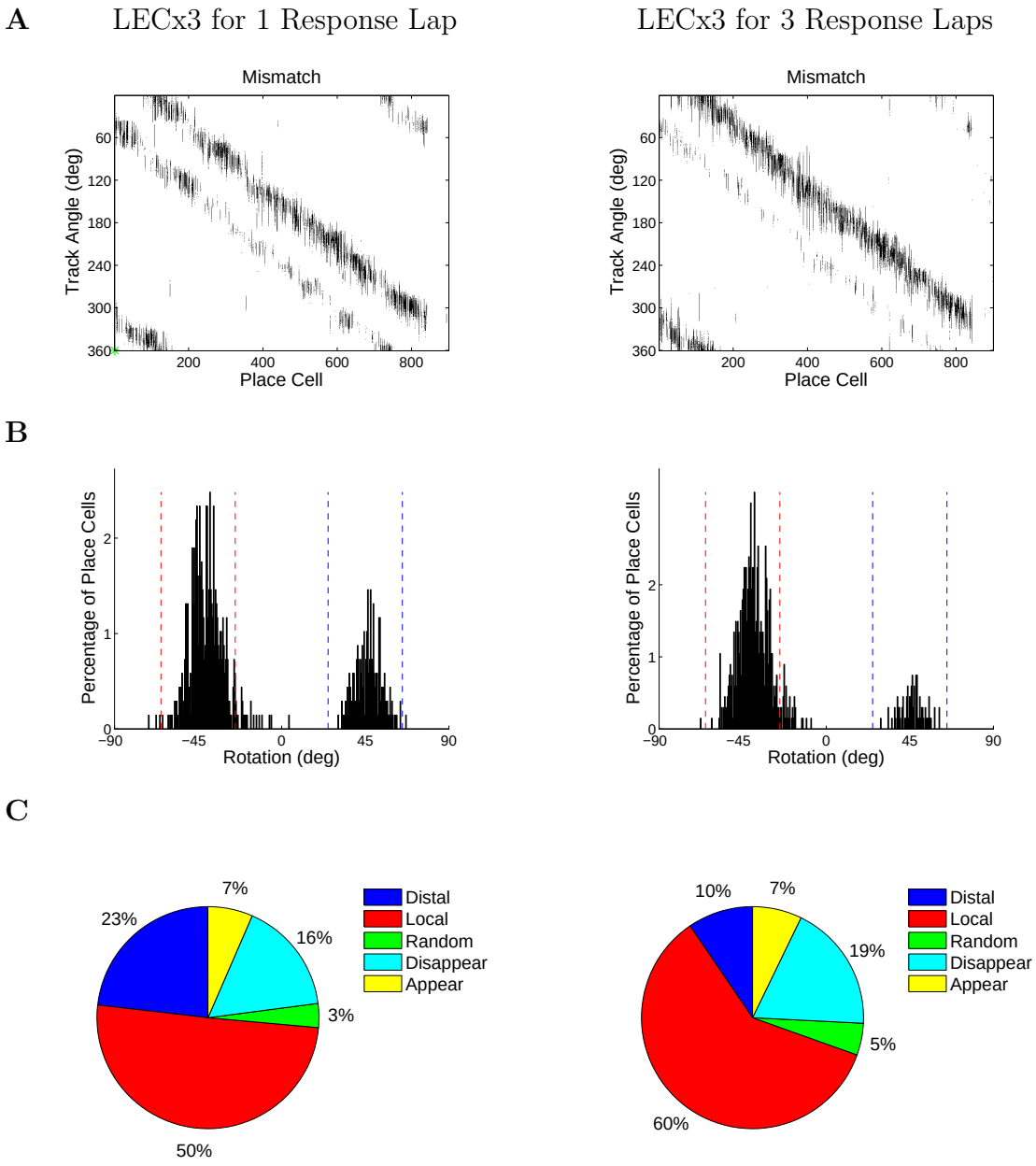


Figure 5.18: Local response in System 4. Given a faster learning rate for targeted inhibition, competition is resolved more quickly. **A.** Place cell activity during the mismatch session given increased LEC activity for one lap (left) or three laps (right). The standard session is shown in Figure 5.15A. **B.** Rotation histogram. Fields become slightly skewed compared to the response seen in System 3, shown in Figure 5.17. **C.** Place cell responses. Competition is quickly resolved as no fields split their representation, even given a single response lap.

Synaptic plasticity during the response lap is essential to the local response observed in all systems. Figure 5.19 compares the activity of System 4 during a single response lap with and without plasticity. Without plasticity, local and distal bands of activity are seen in the response lap. During the mismatch session, the local band disappears along with the transient increase in the LEC input signal, resulting in rotations with distal cues. Given plastic weights, place cell activity along the local band suppresses weights from MEC cells active along the distal band. At the same time, local fields form as weights from MEC cells active along the local band increase.

The input curves from the response lap elucidate a possible role for the asymmetry in recurrent weights. Rotated fields are centered over the local cues due to a balance between recurrent input, which peaks slightly before the cues, and recently formed MEC input, which peaks slightly after the cues, during the response lap. This balance would be lost if the recurrent weights were symmetric, resulting in a forward shift in the place fields.

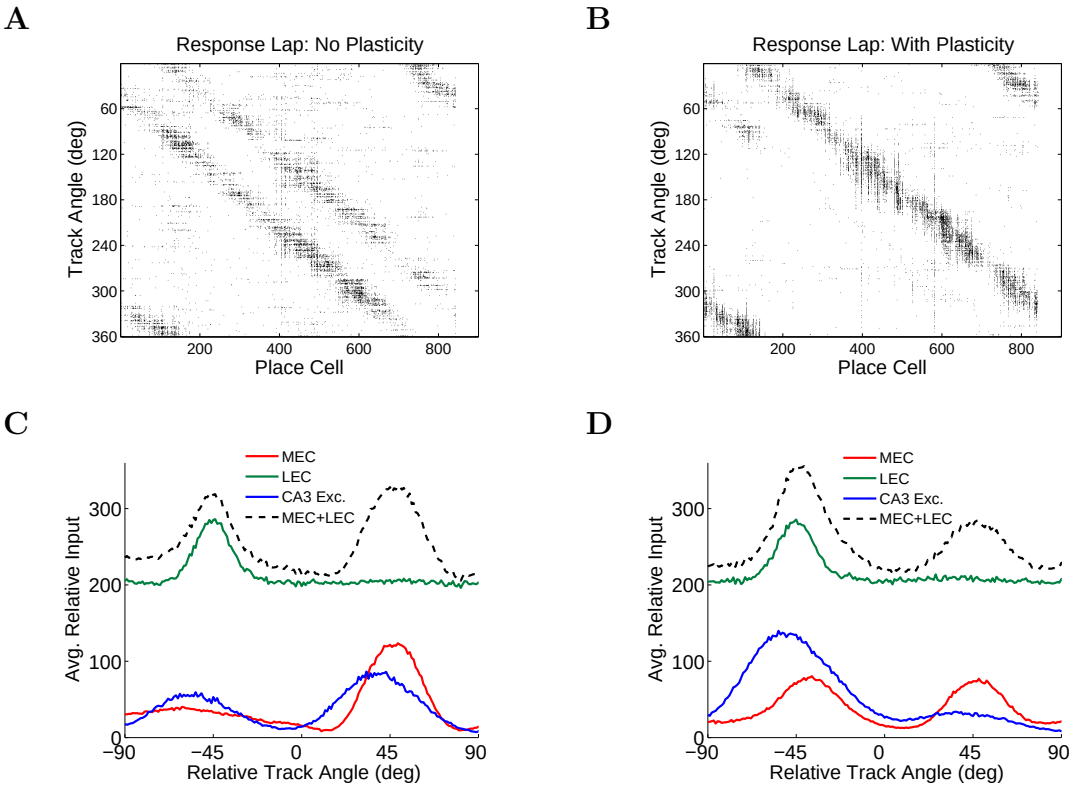


Figure 5.19: Role of synaptic plasticity. The LEC spike rate increases by a factor of three during a single response lap, and System 4 is simulated. The activity during the response lap of the simulation used to generated Figure 5.18 is compared to an identical simulation in which all weights are held constant after training. **A.** Without plasticity during the response lap, the system cannot resolve the competition between the increased LEC input and strong MEC input. When the LEC rate is restored in the mismatch session, all place fields either rotate with the distal cues or remap. **B.** Given synaptic plasticity, most place fields rotate with local cues during the response lap. The activity during the mismatch session is shown in Figure 5.18. **C-D.** Average relative inputs during the response lap, where the relative track angle is taken with respect to place field COMs in the standard session. Without plasticity, MEC cells provide input exclusively at the distal cues, causing place fields to rotate distally when the LEC rate returns to its normal level. With plasticity, place cell activity driven by increased LEC input causes a decrease in MEC weights corresponding to distal cues and an increase in MEC weights corresponding to local cues.

5.2.4 Alternating Response States

The term *transient* in Assumption 1 can be interpreted in several ways. In §5.2.3 transient means the initial lap(s) after cue rotations in their entirety. A second interpretation is that the LEC rate increases only at the beginning of the track as the rat localizes itself. In the reduced CA3 model, strong recurrent connections can carry a burst of place cell activity to other place cells not driven by increased LEC input, recruiting them for the local cues (Figure 5.7). The network model also exhibits this propagation of local activity, but I have not seen the propagation extend beyond 30° of the track. I close by examining a third interpretation based on theories of cognitive control, or the processing of multiple streams of information, a task involving the hippocampus (Wesierska et al. [93]).

Assumption 1B The transient increase in the efficacy of the LEC input signal is periodic as the network oscillates between local and distal states.

Justification During experiments in which a rat must reconcile two misaligned reference frames, distinct ensembles of coactive place cells represent each reference frame (Gothard et al. [23]), and place cell firing patterns have been observed to switch from one reference frame to another in navigational tasks (Redish et al. [72], Gothard et al. [22]). In one such set of experiments, alternating states of activity are observed in a setup similar to the double rotation experiment. In the experiment a rat is placed on a rotating track and must use both local and distal reference frames to avoid shock

zones. As the track rotates continuously, the relative positions of the two cue sets are continuously changing. The experiment indicates that a coactive ensemble of cells represents the same type of information (local or distal), and that ensembles alternate such that the reference frame represented by active place cells changes intermittently (Kelemen and Fenton [41]). In other words, the rat alternates between two behavioral states, paying attention to local cues in one state and distal cues in the other.

Cues are rotated only once in the double rotation experiment, and the place cell population remaps to form a single, stable representation of the modified environment in the cue mismatch session. However, it is possible that alternating states of activity exist in the initial laps after the cue rotations as the rat resolves the conflict in its two reference frames. I assume that the network oscillates between local and distal states in the first two laps after cue rotations but has only one state of activity during the remainder of the mismatch session.

Implications To investigate the alternating state assumption, I simulate two response laps after the standard session over which the LEC rate increases by a factor of four during local states, set to be the intervals $[20i, 20(i+1)]^\circ$ for $i = 0, \ldots, 16$. In the distal states, LEC activity is restored to its standard level.

Figures 5.20 and 5.21 show the results for simulations of Systems 3 and 4, respectively. Without recurrent connections, System 1 shows a mixed response, often with a preference for the distal cues. The results for Systems 2 are similar to results for System 3, but fields again shift forward with respect to the cues.

Place cells in Systems 3 and 4 show a preference for the local cues, where activity in the local states propagates through the distal states. In System 3 the place cell representation of the mismatch session is smooth and centered over the cues sets. In System 4 the representation is more clustered along both the local and distal bands, and more fields rotate with distal cues rather than local cues. Faster learning of targeted inhibition can explain this survival of a distal band of activity. As demonstrated in Figure 5.8, targeted inhibition places cells that were not coactive in the standard session in competition for the same location on the track, enhancing the stronger band of activity while suppressing the weaker band of activity. With a fast inhibitory learning rate, the inhibitory drive between two cells that are coactive in the response laps quickly decreases, allowing the cells to share a firing field. Indeed, simulations of System 4 show a greater local response when inhibitory weights are held constant (data not shown).

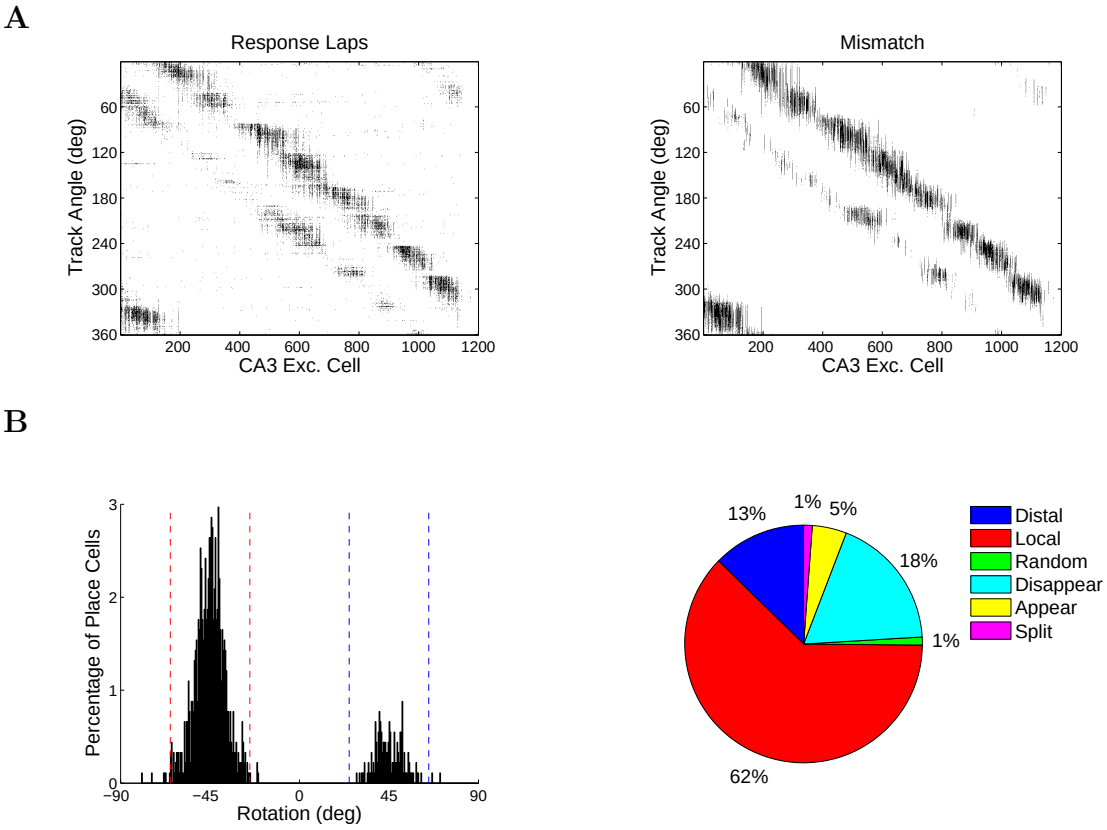


Figure 5.20: Response given System 3 to alternating states. **A-B.** Place cell activity averaged over the two response laps **(A)** or over the mismatch session **(B)**. Activity during the local states propagates through the distal states to form a continuous band of activity corresponding to local cues. **C-D.** Place cell responses. Rotations are centered over the cues, and a local rotation is the dominant response. As was done in the simulation of Figure 5.17, recurrent excitatory weights were held constant during the standard session but were again governed by STDP during the response and mismatch sessions.

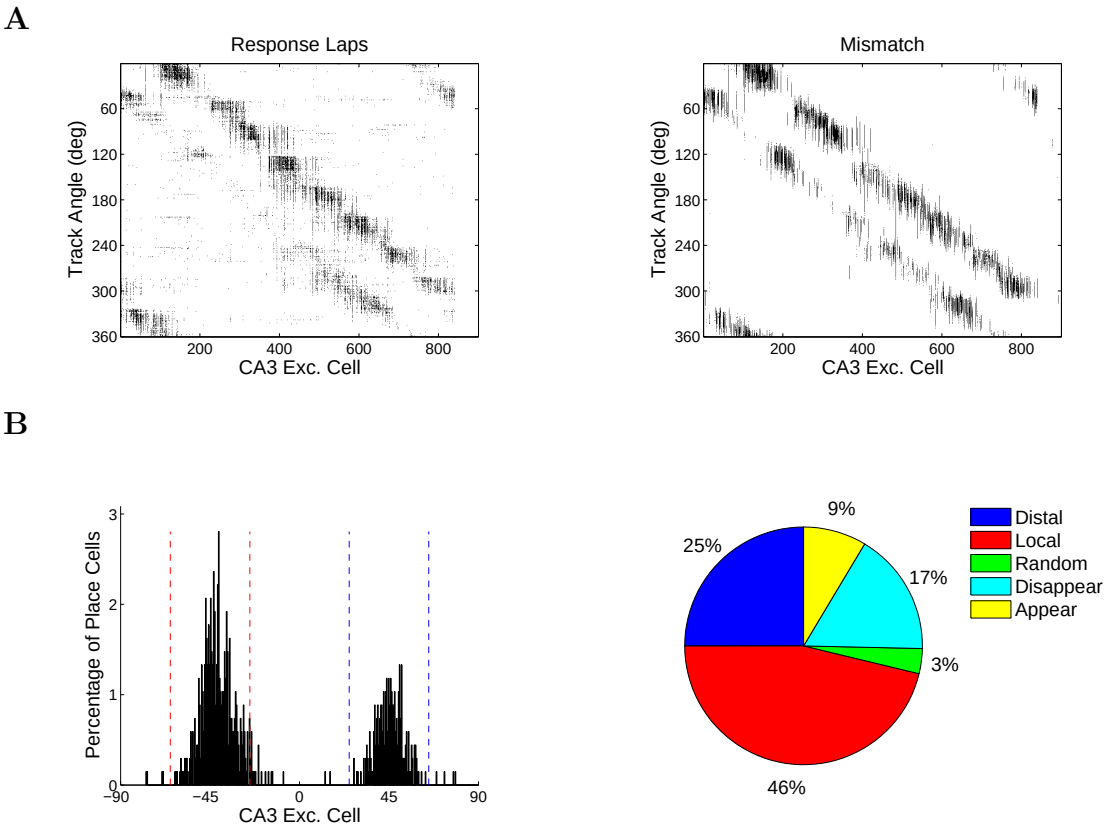


Figure 5.21: Response of System 4 given alternating states. A faster learning rate for targeted inhibition causes more place fields to rotate with distal cues. **A-B.** Place cell activity averaged over the two response laps **(A)** or over the mismatch session **(B)**. Activity during the local states propagates through the distal states, but the place fields are more clustered compared to Figure 5.20, for which the inhibitory learning rate was slower. **C-D.** Place cell responses. Local rotations show a slight forward shift past the local cues. Compared to Figure 5.20, more fields rotate with distal cues rather than local cues. This could be caused by faster inhibitory learning as the inhibition between two cells competing for one region on the track quickly decreases.

Figure 5.22 shows the dynamics of System 3 during the response laps. The system requires two response laps to resolve the competition between local and distal cues as both a local and distal band of activity are seen in the first response lap, but only the local band survives in the second response lap. This results in a local response if two response laps are simulated but in a mixed or distal response if only one response lap is simulated.

Individual cells rotating with local cues differ in their activity during the response laps. As shown in Figure 5.22C, cells 602 and 645 have two active regions corresponding to local and distal cues in the first response lap but only one active region in the second response lap. The local activity in cell 645 is much wider in the response laps than in the mismatch session due to heightened activity propagating down the place cell network. Cell 710 directly rotates with the local cues.

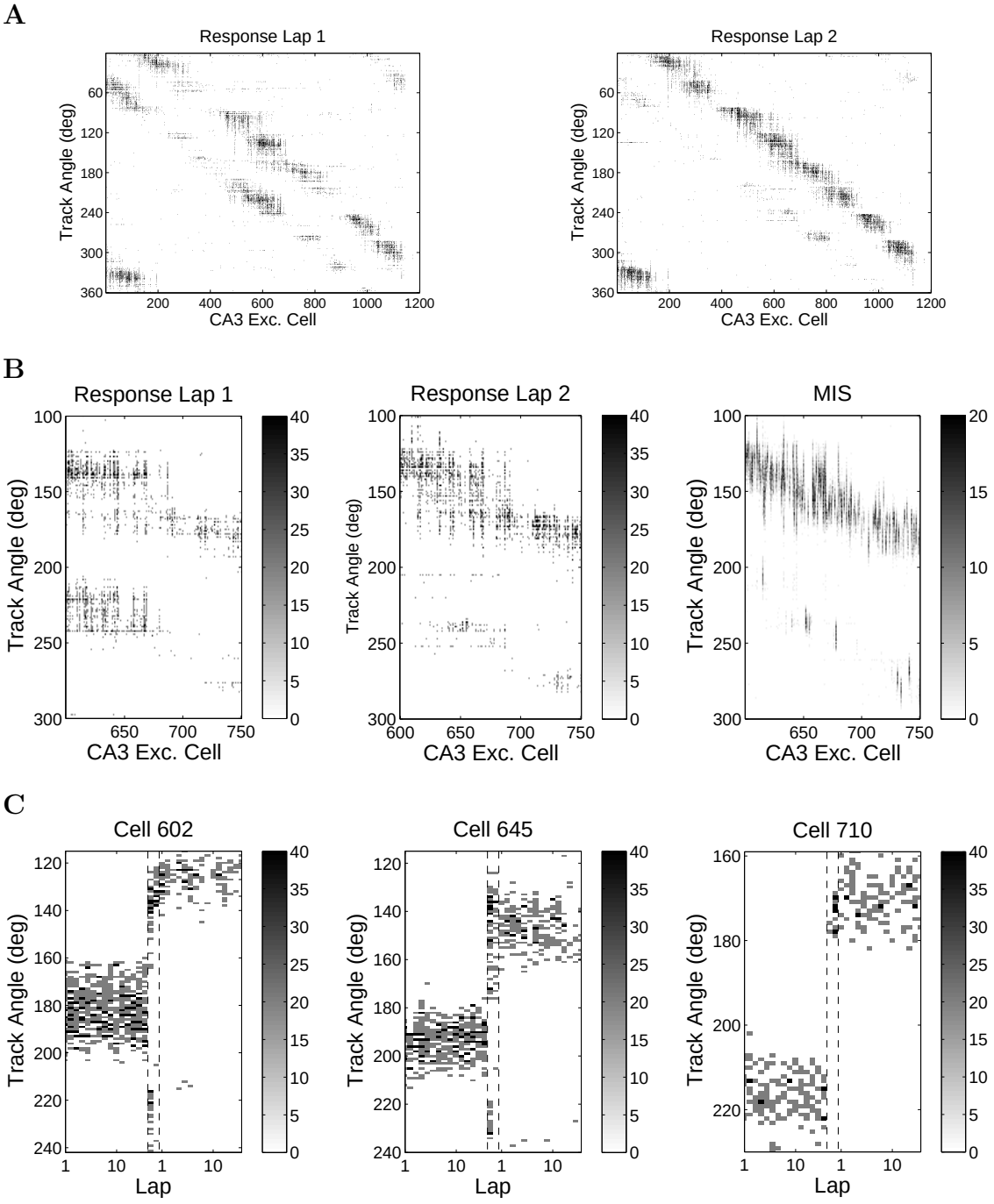


Figure 5.22: Dynamics of System 3 during the response laps. **A.** Place cell activity averaged over the first response lap (left) or second response lap (right). The system alternated between local and distal states identically in both laps. Competition between local and distal bands of activity is resolved in the second response lap as the local band is enhanced and the distal band largely disappears. **B.** Activity of cells 600-750 averaged over the first response lap (left), second response lap (center) and fifteen mismatch laps (right). **C.** Spike rates averaged over each lap for three cells. Some cells require two response laps to resolve competition between local and distal fields (cells 602 and 645), and some cells immediately rotate with local cues (cell 710). Dashed lines indicate the response laps.

Chapter 6

Conclusion

The hippocampus contains a spatial, context-dependent map of the external environment through the activity of place cells driven by synaptic input carrying two distinct streams of information. Grid cells in the MEC provide an absolute spatial reference frame through their hexagonal firing patterns, while cells in the LEC provide sensory feedback regarding local objects or cues. For this thesis I developed computational models to investigate how place cells in the CA3 resolve conflicting information from these two inputs signals. I first examined how synaptic inputs are integrated in dendrites by constructing reduced models of individual neurons that reproduce the subthreshold potential while preserving neuronal structure and the spatial component of the input signal. I then examined the dynamics among a network of interconnected cells in the context of the double rotation experiment. Simulations from both models demonstrate that plasticity mechanisms, behavioral paradigms, and cell assembly dy-

namics can combine to enable the weak LEC input signal to reset place fields largely controlled by the MEC.

Both models provide insight into the processes underlying the counter-intuitive results of the double rotation experiment. The central question I address is why place cells in the CA3 coherently rotate with the weak signal from the LEC for local cues over the much stronger spatial signal from the MEC for distal cues. The simulations point to distinct roles for the MEC, LEC, and CA3 inputs into a place cell in determining its response to cue rotations.

MEC vs. LEC Cells in the MEC and LEC provide conflicting input signals into place cells as the LEC signal rotates with local cues while the MEC signal rotates with distal cues. At first glance, this seems to indicate that MEC and LEC input signals are antagonistic, competing with each other for control of the place fields. The simulations tell a different story. On their own, LEC cells are incapable of causing place fields to rotate with local cues, regardless of the increase in the strength of their input signal, due to the lack of spatial selectivity. Rather, the LEC input signal causes place fields to rotate locally only when it is combined with input from the MEC.

The process can be summarized by considering the LEC signal as a teaching signal, where its strongest influence is seen during the critical initial moments after the cue rotations in which it causes place cells to have a slightly higher firing rate at the local cues. The long-term effect depends on how well weights from MEC cells to place cells

have learned. If the weights are governed by a fast Hebbian plasticity rule gated by postsynaptic activity, then the transient LEC teaching signal effectively resets the weights from MEC cells, causing place fields driven by MEC input to rotate with local cues.

The models predict a transient increase in the efficacy of the LEC input signal upon cue rotations. This would be difficult to test directly as CA3 neurons receive afferents from both the LEC and MEC, and it is very difficult to obtain recordings from distal CA3 dendrites. If the increased efficacy were due to an increase in the LEC spike rate, then one could simply examine the LEC activity during the first few laps of the mismatch session. However, other mechanisms in the brain could be responsible, such as neuromodulators or synchrony in the field potential oscillations in the CA3 and LEC subregions.

It would be interesting to test if there is a differential activation of CA1 place cells along its transverse axis. The simulations show a response similar to that observed experimentally in the CA1 given the same process described above in the absense of associational connections. Thus, the models predict that after cue rotations, place cells in the distal region of the transverse axis would have a transient increase in their activity relative to the activity of place cells in the proximal region.

Cell Assembly Dynamics I refer to a cell assembly as an emsemble of place cells representing a given environment such that cells with correlated activity excite one another through associational connections while cells with uncorrelated activity

inhibit one another through a process I refer to as *targeted feedback inhibition.* In the network model, synaptic plasticity causes a cell assembly to form dynamically during the training session and to adapt to the partial remapping seen in the mismatch session.

Before cue rotations targeted inhibition enforces stability on the cell assembly in two ways. First, it causes the system to be robust to noise from the LEC. In the simulations presented in this thesis, the baseline for the LEC input was relatively small compared to MEC input. Given targeted inhibiton, this baseline could be increased without distorting place fields as the assembly suppresses any place cell activity that disagrees with the learned sequence of activation. Second, it enforces stability on the assembly as a whole as place cells that are silent in the first few laps remain silent upon subsequent laps. When place fields disappear in the mismatch session, new place cells can join the cell assembly due to the release of inhibition, allowing the system to maintain its overall activity level.

The simulations also point to important roles for both recurrent excitation and targeted inhibition as facilitators when the cues are rotated. Their primary effect is to enhance differential activity. When the LEC rate increases during the response laps, it causes heightened place cell activity throughout the track with only a slight preference for the local cues. Recurrent excitation then enhances the stronger signal for local cues while targeted inhibition suppresses the weaker activity away from the local cues. The cell assembly thus facilitates the LEC role as a teaching signal.

As was true in the standard session, the cell assembly causes place cells to respond more coherently to cue rotations. In a similar manner, recurrent excitation can facilitate the propagation of a local signal across distal states in the alternating state hypothesis. This propagation is dependent on the spatial signal from the MEC and thus on fast Hebbian plasticity. Without input from the MEC, recurrent excitation strengthens the activity in the local state caused by increased LEC input. However, the LEC signal cannot nudge the band of activity down the track due to its weak spatial tuning, resulting in a strong cluster of activity in the local state with little propagation into distal states. On the other hand, if MEC input increases during the local state, then its spatial signal can move the strong cluster of activity along the track.

The models predict that recurrent excitation and targeted inhibition enhance differential activity, resulting in a more coherent response to cue rotations. This prediction agrees with the differences observed in the CA1 and CA3 responses to cue rotations. The model similarly predicts a differential degree of coherence in the CA3 response along the septotemporal axis due to the shift in the density of associational connections (Witter and Amaral [94]).

Future Directions The two computational models presented in this thesis are complementary in that the reduced model incorporates the dendritic computations performed by each cell within a simplified network setting, while the network model incorporates the dynamics among large networks of simplified cells. In the future the

models can be further developed and applied to related studies of spatial memory. In this thesis reduced quasi-active neurons were connected into networks through the integrate and fire model. This could also be done by coupling an active soma with a reduced quasi-active tree to reproduce the bursting patterns of CA3 neurons. The ultimate goal in developing reduced models of the neuron is to reduce the full active system while preserving the underlying neuronal structure. Such a reduced model would be invaluable in studying phenomena such as coincidence detection between entorhinal cortical and hippocampal synaptic inputs.

The network model in its current state can be used to investigate a variety of hypotheses concerning the double rotation experiment. It could also be easily adapted to investigate more general tasks of spatial memory. However, the model currently neglects the spatial component in the input signal as cells are isopotential, and the reduced and network models could be combined to examine dendritic computations performed by each individual neuron within a large dynamic network.

Appendix A

Volterra Series

A.1 Construction

The Volterra series is a useful representation of nonlinear systems, but thorough accounts of its derivation and convergence are sparse in the literature. In this appendix we derive the Volterra series for the passive cable given synaptic input and establish its convergence by relating it to the well–studied Picard iterates. For an alternative approach to the Volterra series, see §5.3 of (Mohler [60]).

Consider the full model for the passive cable driven by monosynaptic input,

$$v'(t) = Av(t) + bg(t) + Bv(t)g(t), \quad v(0) = 0,$$

where by Eq. 3.31, $A = -C^{-1}G$, $b = C^{-1}Ee_p$, and $B = -C^{-1}N$. Define $f(t) \equiv \mathrm{e}^{-At}v(t)$. Then,

$$f'(t) = \mathrm{e}^{-At}\big[bg(t) + B\mathrm{e}^{At}f(t)g(t)\big], \quad f(0) = 0. \tag{A.1}$$

If $g(t)$ were transient, the product $f(t)g(t)$ would be relatively small, motivating the initial guess

$$f^1(t) = \int_0^t \mathrm{e}^{-As} bg(s)ds.$$

Eq. A.1 is solved by the fixed point problem

$$f(t) = f^1(t) + \int_0^t K(s)f(s)ds,$$

where

$$K(t) = \mathrm{e}^{-At} B\mathrm{e}^{At} g(t).$$

Each iterate is used to generate the next guess for f via

$$f^k(t) = f^1(t) + \int_0^t K(s)f^{k-1}(s)ds.$$

This method is known as Picard iterations or the method of successive approximations, and as both f^k and K are continuous, $f^k(t)$ converges to $f(t)$, as shown in §1.10 of [9]. We construct the Volterra series by defining $v^1(t) \equiv \mathrm{e}^{At} f^1(t)$ and $v^k(t) \equiv \mathrm{e}^{At}(f^k - f^{k-1})(t)$ for $k > 1$. Its convergence follows via

$$\sum_{k=1}^{\infty} v^k(t) = \mathrm{e}^{At}\big(f^1 + \sum_{k=1}^{\infty}(f^{k+1} - f^k)\big) = \mathrm{e}^{At} f(t) = v(t).$$

Unwrapping each term reveals

$$v^1(t) = \mathrm{e}^{At} f^1(t) = \int_0^t \mathrm{e}^{A(t-s)} bg(s)ds, \tag{A.2}$$

implying that $(v^1)'(t) = Av^1(t) + bg(t)$, or

$$C(v^1)'(t) + Gv^1(t) = Ee_p g(t), \quad v^1(0) = 0. \tag{A.3}$$

Similarly, for $k \geq 2$,

$$v^k(t) = \mathrm{e}^{At} \int_0^t K(s) e^{-As} v^{k-1}(s) ds = \int_0^t \mathrm{e}^{A(t-s)} B g(s) v^{k-1}(s) ds, \tag{A.4}$$

implying that $(v^k)'(t) = Av^k(t) + Bg(t)v^{k-1}(t)$, or

$$C(v^k)'(t) + Gv^k(t) = -g(t)Nv^{k-1}(t), \quad v^k(0) = 0. \tag{A.5}$$

The series easily generalizes for polysynaptic input by replacing the input of Eq. A.3 with $\sum_{j=1}^m E_{p_j} e_{p_j} g_{p_j}(t)$ and of Eq. A.5 with $-\sum_{j=1}^m g_{p_j}(t) N^{(p_j)} v^{k-1}(t)$.

A.2 Transfer Functions

When the cable is driven by injected input currents, we construct a reduced system to match the leading moments of the transfer functions for the full and reduced systems. The same procedure can lead to an appropriate reducer for the cable driven by synaptic input by considering the transfer functions for each term in the Volterra series, where each transfer function maps the synaptic conductance g to $y^k = e_{siz}^T v^k$. In this appendix we derive the transfer functions for each Volterra term given monosynaptic input. The transfer functions can be generalized for polysynaptic input.

Assume $g(t < 0) = 0$ and define the kernel

$$h_1(t) \equiv \begin{cases} \mathrm{e}^{At} b & \text{if } t \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $A = -C^{-1}G$ and $b = C^{-1}Ee_p$. Eq. A.2 can be written as

$$v^1(t) = \int_{-\infty}^{\infty} h_1(\tau)g(t-\tau)d\tau = (h_1 \star g)(t). \tag{A.6}$$

As expected, v^1 depends linearly on g, and by the convolution theorem,

$$\mathcal{L}v^1(s) = \widetilde{H}_1(s)\mathcal{L}g(s), \quad \text{where} \quad \widetilde{H}_1 = \mathcal{L}h_1$$

and $\mathcal{L}$ denotes the Laplace transform. The transfer function $\widetilde{H}_1$ is then given by

$$\widetilde{H}_1(s) = \int_0^{\infty} \mathrm{e}^{(A-sI)t}b \; dt = (G+sC)^{-1}Ee_p.$$

Since the passive system is stable, $\widetilde{H}_1$ is well-defined if $\mathrm{Re}(s) \geq 0$ and provides the mapping

$$\mathcal{L}y^1(s) = H_1(s)\mathcal{L}g(s), \quad \text{where} \quad H_1 = e_{siz}^T\widetilde{H}_1. \tag{A.7}$$

We next strive to write the second Volterra term as a convolution, which would allow us to easily compute its transfer function. By Eq. A.4, v^2 depends on the product of v^1 and g, both of which depend on g. We thus make the educated guess that v^2 can be written as

$$v^2(t) = \iint_{\Re^2} h_2(\tau_1,\tau_2)g(t-\tau_1)g(t-\tau_2)d\tau_1 d\tau_2 \tag{A.8}$$

and solve for the kernel h_2. As Eq. A.8 is not quite a convolution, we pause to show the advantage of its form. Define $\overline{g}(t_1,t_2) \equiv g(t_1)g(t_2)$, and define

$$\overline{v}(t_1,t_2) \equiv (h_2 \star \overline{g})(t_1,t_2).$$

Then, $v^2(t) = \overline{v}(t,t)$, and by the convolution theorem,

$$\mathcal{L}\overline{v}(s_1,s_2) = \widetilde{H}_2(s_1,s_2)\mathcal{L}\overline{g}(s_1,s_2) = \widetilde{H}_2(s_1,s_2)\mathcal{L}g(s_1)\mathcal{L}g(s_2),$$

where $\widetilde{H}_2 = \mathcal{L}h_2$. Therefore,

$$v^2(t) = \mathcal{L}^{-1}\big(\widetilde{H}_2(s_1,s_2)\mathcal{L}g(s_1)\mathcal{L}g(s_2)\big)(t,t), \tag{A.9}$$

and the two-dimensional transfer function $\widetilde{H}_2$ does indeed map the input g to the output v^2 in the frequency domain. We now return to the computation of the kernel h_2 and corresponding transfer function. By Eq. A.4,

$$v^2(t) = \int_0^\infty \mathrm{e}^{A\tau_1} B v^1(t-\tau_1) g(t-\tau_1) d\tau_1,$$

where $B = -C^{-1}N$. By Eq. A.6,

$$v^1(t-\tau_1) = \int_0^\infty h_1(s) g(t-\tau_1-s) ds = \int_0^\infty h_1(\tau_2-\tau_1) g(t-\tau_2) d\tau_2,$$

given the change of variables $\tau_2 = \tau_1 + s$. Eq. A.8 is then obtained by defining the kernel

$$h_2(t_1,t_2) \equiv \begin{cases} \mathrm{e}^{At_1} B \mathrm{e}^{A(t_2-t_1)} b & \text{if } 0 \le t_1 \le t_2, \\ 0 & \text{otherwise.} \end{cases}$$

The transfer function $\widetilde{H}_2 = \mathcal{L}h_2$ is given by

$$\begin{aligned} \widetilde{H}_2(s_1,s_2) &= \int_0^\infty \int_0^\infty \mathrm{e}^{-s_1 t_1 - s_2 t_2} h_2(t_1,t_2) dt_1 dt_2 \\ &= \int_0^\infty \mathrm{e}^{(A-s_1 I)t_1} B \mathcal{I}(t_1;s_2) dt_1 \end{aligned}$$

where

$$\begin{aligned}\mathcal{I}(t_1; s_2) &= \int_0^\infty \mathrm{e}^{-s_2 t_2} h_1(t_2 - t_1) dt_2 \\ &= \int_0^\infty \mathrm{e}^{-s_2(t_1+\tau)} h_1(\tau) d\tau = \mathrm{e}^{-s_2 t_1} \widetilde{H}_1(s_2).\end{aligned}$$

Hence, if $\mathrm{Re}(s_1 + s_2) \geq 0$, then

$$\begin{aligned}\widetilde{H}_2(s_1, s_2) &= \int_0^\infty \mathrm{e}^{(A-(s_1+s_2)I)t_1} B \widetilde{H}_1(s_2) dt_1, \\ &= -(G + (s_1 + s_2)C)^{-1} N (G + s_2 C)^{-1} E e_p.\end{aligned}$$

Finally, $H_2 \equiv e_{siz}^T \widetilde{H}_2$ combined with Eq. A.9 leads to

$$y^2(t) = \mathcal{L}^{-1}\big(H_2(s_1, s_2)\mathcal{L}g(s_1)\mathcal{L}g(s_2)\big)(t, t). \tag{A.10}$$

To simplify notation, the transfer function can be written in its regular form, defined in §2.3 of (Rugh [77]) such that $H_2(s_1, s_2) = H_2^{reg}(s_1 + s_2, s_2)$, or

$$H_2^{reg}(s_1, s_2) = -e_{siz}^T (G + s_1 C)^{-1} N (G + s_2 C)^{-1} E e_p.$$

In a similar manner, one can iteratively compute the kernels and corresponding transfer functions for each Volterra term. As each Volterra term has an increasingly nonlinear dependence on g, it can be written as

$$v^k(t) = (h_k \star \overline{g})(t, \cdots, t),$$

where $\overline{g}(t_1, \cdots, t_k) \equiv \prod_{j=1}^k g(t_j)$. Given $\overline{v}(t_1, \cdots, t_k) \equiv (h_k \star \overline{g})(t_1, \cdots, t_k)$,

$$\mathcal{L}\overline{v}(s_1, \cdots, s_k) = \widetilde{H}_k(s_1, \cdots, s_k)\mathcal{L}g(s_1) \cdots \mathcal{L}g(s_k),$$

and $v^k(t) = \overline{v}(t, \cdots, t)$. The transfer function for the output y^k is then given by

$$H_k^{reg}(s_1, \cdots, s_k) = (-1)^{k+1} e_{siz}^T \prod_{j=1}^{k-1} ((G + s_j C)^{-1} N)(G + s_k C)^{-1} E e_p,$$

$$H_k(s_1, \cdots, s_k) = H_k^{reg}(s_1 + \ldots + s_k,\ s_2 + \ldots + s_k, \ldots, s_k).$$

Appendix B

Network Model Parameters

Grid Cells

	Value	Description
θ	0 – $\pi/3$ rad	tilt
b	28 – 60 cm	base
r	0 – 33 cm	offset magnitude
ϕ	0 – 2π rad	offset direction
γ	0.018	spread of grid
t_{ref}	3 ms	refractory period
f_{max}	20 Hz	max spike rate

LEC Cells

	Value	Description
P_0	0.05 – 0.15	base probability
N_{hot}	3	number of hot spots
γ	4	spread of fields
$P_{D,i}$	0 – 0.3	prob. diff. for field i

Table B.1: Entorhinal Cortex Parameters.

	Value	Description
C_m	1 $\mu F/cm^2$	membrane capacitance
g_L	0.3 mS/cm^2	leakage conductance
V_L	-70 mV	leakage reversal potential
V_I	-70 mV	inh. reversal potential
V_E	0 mV	exc. reversal potential
V_{reset}	-65 mV	reset potential
t_{ref}	3 ms	refractory period
τ_E	5 ms	exc. decay constant
τ_I	5 ms	inh. decay constant

Table B.2: CA3 Parameters.

MEC to CA3 Excitatory Cells		
k :	50 ms	learning rate
β	10 Hz	rate constant
τ_{Ca}	100 ms	calcium decay rate
$\bar{R}$	5 Hz	threshold rate
LEC to CA3 Excitatory Cells		
k :	15 ms	learning rate
β	10 Hz	rate constant
τ_{Ca}	100 ms	calcium decay rate
$\bar{R}$	2.5 Hz	threshold rate
Targeted Feedback Inhibition		
k :	50 – 70 ms	learning rate
β	10 Hz	rate constant
τ_{Ca}	100 ms	calcium decay rate
$\bar{R}$	0.5 Hz	threshold rate
Associational Connections		
A_+	0.38	Maximum LTP
A_-	0.4	Maximum LTD
τ_+	20 ms	LTP decay constant
τ_+	20 ms	LTD decay constant

Table B.3: Plasticity parameters.

W^0		
ρ :	0.1	density of connections
W_{max}	1.5	maximum weight
W_{init}	0.8	initial weight
W^1		
ρ :	0.2	density of connections
W_{max}	1.5	maximum weight
W_{init}	0	initial weight
W^2		
ρ :	0.5	density of connections
W_{max}	1	maximum weight
W_{init}	0	initial weight
W^3, W^4		
ρ :	0.4	density of connections
W_{max}	3	maximum weight
W^5		
ρ :	0.5	density of connections
W_{max}	6	maximum weight
W_{init}	0	initial weight

Table B.4: Weight matrix parameters.

Bibliography

[1] D.G. Amaral and M.P. Witter. The three-dimensional organization of the hippocampal formation: A preview of anatomical data. *Neuroscience*, 31:571–591, 1989.

[2] A.C. Antoulas, D.C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *American Mathematical Society Publications*, 280:193–219, 2001.

[3] G.A. Ascoli, L. Hunter, J.L. Krichmar, J.L. Olds, and S.L. Senft. Computational neuroanatomy of the hippocampus, April 2009. http://www.krasnow.gmu.edu/L-Neuron/index.html.

[4] Z. Bai and D. Skoogh. A projection method for model reduction of bilinear dynamical systems. *Linear Algebra and its Applications*, 415:406–425, 2006.

[5] J. Baker, T. Perez-Rosello, M. Migliore, g. Barrionuevo, and G.A. Ascoli. A computer model of unitary responses from assocational/commissural and perforant path synapses in hippocampal CA3 pyramidal cells. *Journal of Computational Neuroscience*, 31:137–158, 2011.

[6] G. Bi and M. Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, 18:10464–10472, 1998.

[7] K.I. Blum and L.F. Abbott. A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, 8:85–93, 1996.

[8] L.J. Borg-Graham. Modelling the non-linear conductances of excitable membranes. In J. Chad and H. Wheal, editors, *Cellular Neurobiology: A Practical Approach*, pages 247–275. IRL Press at Oxford University Press, New York, 1991.

[9] M. Braun. *Differential Equations and Their Applications*. Springer-Verlag, New York, 1975.

[10] V.H. Brun, T. Solstad, K.B. Kjelstrup, M. Fyhn, M.P. Witter, E.I. Moser, and M.B. Moser. Progressive increase in grid scale from dorsal to ventral Medial Entorhinal Cortex. *Hippocampus*, 18:1200–1212, 2008.

[11] P.C. Bush and T.J. Sejnowski. Reduced compartmental models of neocortical pyramidal cells. *Journal of Neuroscience Methods*, 46:159–166, 1993.

[12] S. Chaturantabut and D. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32:2737–2764, 2010.

[13] P. Dayan and L. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, 2001.

[14] J. Dudman, D. Tsay, and S. Siegelbaum. A role for synaptic inputs at distal dendrites: instructive signals for hippocampal long-term plasticity. *Neuron*, 56:866–879, 2007.

[15] A.D. Ekstrom, J. Meltzer, and B.L. McNaughton. NMDA receptor antagonism blocks experience-dependent expansion of hippocampal place fields. *Neuron*, 31:631–638, 2001.

[16] M. Franzius, R. Vollgraf, and L. Wiskott. From grids to places. *Journal of Computational Neuroscience*, 22:297–299, 2007.

[17] R.W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *Journal of Computational and Applied Mathematics*, 123:395–421, 2000.

[18] R.W. Freund. The sprim algorithm for structure-preserving order reduction of general rcl circuits. In P. Benner, M. Hinze, and E.J.W. ter Maten, editors, *Model Reduction for Circuit Simulation*, pages 25–52. Springer, New York, 2011.

[19] R.C. Froemke and Y. Dan. Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*, 416:433–438, 2002.

[20] F. Gabbiani and S. Cox. *Mathematics for Neuroscientists*. Elsevier Academic Press, Boston, 2010.

[21] W. Gerstner and W. Kistler. *Spiking Neuron Models*. Cambridge University Press, New York, NY, 2002.

[22] K.M. Gothard, W.E. Skaggs, and B.L. McNaughton. Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *Journal of Neuroscience*, 16:8027–8040, 1996.

[23] K.M. Gothard, W.E. Skaggs, K.M. Moore, and B.L. McNaughton. Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. *Journal of Neuroscience*, 16:823–835, 1996.

[24] E.J. Grimme. *Krylov Projection Methods for Model Reduction*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 1997.

[25] C. Gu. Qlmor: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 30:1307–1320, 2011.

[26] S. Gugercin, A. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30:609–638, 2008.

[27] T. Hafting, M. Fyhn, S. Molden, M. Moser, and E. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436:801–806, 2005.

[28] E.L. Hargreaves, G. Rao, I. Lee, and J.J. Knierim. Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science*, 308:1792–1794, 2005.

[29] M. Hasselmo. Grid cell mechanisms and function: contributions of entorhinal persistent spiking and phase resetting. *Hippocampus*, 18:1213–1229, 2008.

[30] H. Hayashi and Y. Nonaka. Cooperation and competition between lateral and medial perforant path synapses in the dentate gyrus. *Neural Networks*, 24:233–246, 2010.

[31] S. Healy. *Spatial Representation in Animals*. Oxford University Press, 1998.

[32] D.O. Hebb. *The Organization of Behavior; A Neuropsychological Theory*. Wiley, New York, NY, 1949.

[33] K. Hedrick and S.J. Cox. Structure-preserving model reduction of passive and quasi-active neurons. *Journal of Computational Neuroscience*, under revision, 2012.

[34] P. Hemond, D. Epstein, A. Boley, M. Migliore, G.A. Ascoli, and D.B. Jaffe. Distinct classes of pyramidal cells exhibit mutually exclusive firing patterns in hippocampal area CA3b. *Hippocampus*, 18:411–424, 2008.

[35] A. Heynen, W. Abraham, and M. Bear. Bidirectional modification of CA1 synapses in the adult hippocampus in vivo. *Nature*, 381:163–166, 1996.

[36] M. Hines. Efficient computation of branched nerve equations. *International Journal of Bio-Medical Computing*, 15:69–76, 1984.

[37] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.

[38] H.T. Ito and E.M. Schuman. Functional division of hippocampal area CA1 via modulatory gating of entorhinal cortical inputs. *Hippocampus*, 22:372–387, 2012.

[39] T Jarsky, A Roxin, W Kath, and N Spruston. Conditional dendritic spike propagation following distal synaptic activation of hippocampal CA1 pyramidal neurons. *Nature Neuroscience*, 8:1667–1676, 2005.

[40] D. Johnston and D.G. Amaral. Hippocampus. In G. Shepherd, editor, *The Synaptic Organization of the Brain*, chapter 10, pages 417–458. Oxford University Press, New York, 1998.

[41] E. Kelemen and A. Fenton. Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. *PLOS Biology*, 8:e10000403. doi:10.1371/journal.pbio.10000403, 2010.

[42] A. Kellems, S. Chaturantabut, D. Sorensen, and S. Cox. Morphologically accurate reduced order modeling of spiking neurons. *Journal of Computational Neuroscience*, 28:477–494, 2010.

[43] A. Kellems, D. Roos, N. Xiao, and S. Cox. Low-dimensional, morphologically accurate models of subthreshold membrane potential. *Journal of Computational Neuroscience*, 27:161–176, 2009.

[44] J Knierim. Dynamics interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. *The Journal of Neuroscience*, 22:6254–6264, 2002.

[45] J. Knierim. Hippocampal place cells: parallel input streams, subregional processing, and implications for episodic memory. *Hippocampus*, 16:755–764, 2006.

[46] C. Koch. *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, New York, 1999.

[47] H. Krapp and F. Gabbiani. Spatial distribution of inputs and local receptive field properties of a wide-field, looming sensitive neuron. *Journal of Neurophyiology*, 93:2240–2253, 2005.

[48] I. Lee, G. Rao, and J. Knierim. A double dissociation between hippocampal subfields: Differential time course of CA3 and CA1 place cells for processing changed environments. *Neuron*, 42:803–815, 2004.

[49] I. Lee, D. Yoganarashimha, G. Rao, and J. Knierim. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature*, 430:456–459, 2004.

[50] S. Leutgeb, J. Leutgeb, M. Moser, and E. Moser. Place cells, spatial maps and the population code for memory. *Current Opinion in Neurobiology*, 15:738–746, 2005.

[51] P. Li and L.T. Pileggi. Compact reduced-order modeling of weakly nonlinear analog and rf circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 23:184–203, 2005.

[52] R.C. Li and Z. Bai. Structure-preserving model reduction using a Krylov subspace projection formulation. *Communications in Mathematical Sciences*, 3:179–199, 2005.

[53] Y.C. Liang, H.P. Lee, S.P. Lim, W.Z. Lin, K.H. Lee, and C.G. Wu. Proper orthogonal decomposition and its applications-part i: Theory. *Journal of Sound and Vibration*, 252:527–544, 2002.

[54] Y. Lin, L. Bao, and Y. Wei. Order reduction of bilinear MIMO dynamical systems using new block Krylov subspaces. *Computers and Mathematics with Applications*, 58:1093–1102, 2009.

[55] J.E. Lisman. Role of the dual entorhinal inputs to hippocampus: a hypothesis based on cue/action (non-self/self) couplets. *Progress in Brain Research*, 163:615–625, 2007.

[56] H. Markrum, J. Lübke, M. Frotscher, and B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275:213–215, 1997.

[57] C.O. Martinez, V.H. Do, and B.E. Derrick. Endogenous opioid peptides contribute to associative LTP in the hippocampal CA3 region. *Neurobiology of Learning and Memory*, 96:207–217, 2011.

[58] M. Mehta, C. Barnes, and B. McNaughton. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94:8918–8921, 1997.

[59] M. Migliore, E.P. Cook, D.B. Jaffe, D.A. Turner, and D. Johnston. Computer simulations of morphologically reconstructed CA3 hippocampal neurons. *Journal of Neurophysiology*, 73:1157–1168, 1995.

[60] R. Mohler. *Nonlinear Systems: Applications to Bilinear Control*. Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[61] J. Monaco, J. Knierim, and K. Zhang. Sensory feedback, error correction, and remapping in a multiple oscillator model of place-cell activity. *Frontiers in Computational Neuroscience*, 5, 2011.

[62] R.G.M. Morris. Synaptic plasticity and learning: Selective impairment of learning in rats and blockade of long-term potentiation in vivo by the NMDA receptor antagonist AP5. *Journal of Neuroscience*, 9:3049–3057, 1989.

[63] A. Odabasioglu, M. Celik, and L.T. Pileggi. PRIMA: Passive reduced-order interconnect macromodeling algorithm. *IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems*, 17:645–654, 1998.

[64] J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34:171–175, 1971.

[65] J. O'Keefe and L. Nadel. *The Hippocampus as a Cognitive Map*. Oxford, Clarendon, 1978.

[66] M. O'Shea and C.H. Rowell. The neuronal basis of a sensory analyser, the acridid movement detector sysem. II. response decrement, convergence, and the nature of excitatory afferents to the fan-like dendrites of the LGMD. *Journal of Experimental Biology*, 65:289–308, 1976.

[67] J. Phillips. Projection frameworks for model reduction of weakly nonlinear systems. *Proceedings of DAC 2000*, pages 184–189, 2000.

[68] J. Phillips. Projection-based approaches for model reduction of weakly nonlinear, time-varying systems. *IEEE Trans. Computer-Aided Design Integrated Circuits Systems*, 22:171–187, 2003.

[69] P. Pinsky and J. Rinzel. Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *Journal of Computational Neuroscience*, 1:39–60, 1994.

[70] R. Poznanski. A generalized tapering equivalent cable model for dendritic neurons. *Bulletin of Mathematical Biology*, 53:457–467, 1991.

[71] W. Rall. Branching dendritic trees and motoneuron membrane resistivity. *Experimental Neurology*, 1:491–527, 1959.

[72] A.D. Redish, E.S. Rosenzweig, J.D. Bohanick, B.L. McNaughton, and C.A. Barnes. Dynamics of hippocampal ensemble activity realignment: time versus space. *Journal of Neuroscience*, 20:9298–9309, 2000.

[73] R.W. Rodieck, N.Y.S. Kiang, and G.L. Gerstein. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophysical Journal*, 2:351–368, 1962.

[74] E.T. Rolls and A. Trevew. *Neural Networks and Brain Function*. Oxford University Press, Oxford, New York, 1998.

[75] J. Roychowdhury. Reduced-order modeling of time-varying systems. *IEEE Trans. Circuits Syst. II: Analog Digital Signal Process.*, 46:1273–1288, 1999.

[76] J. Rubin, D. Lee, and H. Sompolinsky. Equilibrium properties of temporally asymmetric hebbian plasticity. *The American Physical Society*, 86:364–367, 2001.

[77] W.J. Rugh. *Nonlinear System Theory*. The Johns Hopkins University Press, Baltimore, 1981.

[78] F. Savelli and J.J. Knierim. Hebbian analysis of the transformation of medial entorhinal grid-cell inputs to hippocampal place fields. *Journal of Neurophysiology*, 103:3167–3183, 2010.

[79] A.K. Schierwagen. A non-uniform equivalent cable model of membrane voltage changes in a passive dendritic tree. *Frontiers in Neuroscience*, 1:19–42, 1989.

[80] W.B. Scoville and B. Milner. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry*, 20:11–21, 1957.

[81] HZ Shouval, MF Bear, and LN Cooper. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences*, 99:10831–10836, 2002.

[82] I.G. Silkis. Involvement of the trisynaptic hippocampal pathway in generating neural representations of object-place associations (an analytical review). *Neuroscience and Behavioral Physiology*, 41:117–129, 2011.

[83] T. Solstad, E. Moser, and G. Einevoll. From grid cells to place cells: a mathematical model. *Hippocampus*, 16:1021–1031, 2006.

[84] S. Song, K. Miller, and L.F. Abbott. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926, 2000.

[85] N Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9:206–221, 2008.

[86] L.R. Squire, C.E.L. Stark, and R.E. Clark. The medial temporal lobe. *Annual Review of Neuroscience*, 27:279–306, 2004.

[87] E. Teng and L.R. Squire. Memory for places learned long ago is intact after hippocampal damage. *Nature*, 400:675–677, 1999.

[88] R.D. Traub, R. Miler, R.U. Muller, and A.I. Gulyás. Functional organization of the hipocampal CA3 region: implications for epilepsy, brain waves and spatial behavior. *Network: Computation in Neural Systems*, 3:465–488, 1992.

[89] R.D. Traub, K.S. Wong, R. Miles, and H. Michelson. A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *Journal of Neurophysiology*, 66:635–650, 1991.

[90] L. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[91] D.A. Turner, X.G. Li, G.K. Pyapali, A. Ylinen, and G. Buzsaki. Morphometric and electrical properties of reconstructed hippocampal CA3 neurons recorded in vivo. *Journal of Computational Neuroscience*, 356:580–594, 1995.

[92] C.E. Villemagne and R.E. Skelton. Model reduction using a projection formulation. *International Journal of Control*, 46:2141–2169, 1987.

[93] M. Wesierska, C. Dockery, and D. Fenton. Beyond memory, navigation, and inhibition: behavioral evidence for hippocampus-dependent cognitive coordination in the rat. *Journal of Neuroscience*, 25:2413–2419, 2005.

[94] M. Witter and D. Amaral. Hippocampal formation. In G. Paxinos, editor, *The Rat Nervous System*, chapter 21, pages 647–657. Elsevier Academic Press, San Diego, CA, 2004.

[95] M.P. Witter. Organization of the Entorhinal-Hippocampal system: A review of current anatomical data. *Hippocampus*, 3:33–44, 1993.

[96] M.P. Witter and E.I. Moser. Spatial representation and the architecture of the entorhinal cortex. *Trends in Neuroscience*, 29:671–678, 2006.

[97] B. Yan and P. Li. Reduced order modeling of passive and quasi-active dendrites for nervous system simulation. *Journal of Computational Neuroscience*, 31:247–271, 2011.

[98] X. Yu, J.J. Knierim, I. Lee, and H.Z. Shouval. Simulating place field dynamics using spike timing-dependent plasticity. *Neurocomputing*, 69:1253–1259, 2006.

[99] X. Yu, H.Z. Shouval, and J.J. Knierim. A biophysical model of synaptic plasticity and metaplasticity can account for the dynamics of the backward shift of the hippocampal place fields. *Journal of Neurophysiology*, 100:983–992, 2008.