A Spiking Circuit Model of Sequence Learning

by

Kevin D. Himberger

A thesis submitted in conformity with the requirements for the degree of Masters of Arts

Psychology

University of Toronto

© Copyright by Kevin D. Himberger 2015
A Spiking Circuit Model of Sequence Learning

Kevin D. Himberger

Masters of Arts
Psychology
University of Toronto
2015

Abstract

The ability to integrate information over time is fundamental to cognition. Any sensory experience, from object recognition to speech comprehension, utilizes this ability. Contrary to traditional models in which memory and perception are supported by separate circuits, recent work has demonstrated perceptual circuits can integrate information over time, as their response to an input at one moment may depend on what happened seconds earlier. We implemented a spiking model utilizing spike-timing dependent synaptic plasticity in a localized visual cortical circuit to investigate how sensory circuits are able to integrate information over time. We demonstrate that a lengthening circuit memory develops as a result of repeated exposure to a stimulus and that repeated exposure also leads to stronger synapses between neurons processing the input as well as associating sequential stimuli.
Acknowledgments

I have been privileged to pursue varied interests and careers throughout my life. The primary enablers of this privilege have been my parents, Alberta and David Himberger. I may not have become a baseball player, or geologist, or alpaca owner, or professional poker player, but without them I would probably be a homeless vagabond working odd jobs in an underdeveloped country. They supported my decision to eschew an established career and this is the first result from that change.

Another crucial person in the creation of this work and who has enabled me to pursue a career in neuroscience is my advisor, Dr. Christopher J. Honey. His patience, availability, and comprehension of terse, vague statements are greatly appreciated. Without his editorial abilities, this thesis may have remained a Joycean stream of tangentially related nouns.

I appreciate the helpful feedback from and support of the most adored woman in the world, Anna Rose Miller, and hope she continues to provide assistance from a closer locale in the future.

I’d also like to thank my committee members, Dr. Dirk Bernhardt-Walther and Dr. Taufik Valiante. Their feedback and advice is critical to my continued growth and without them the readership for my thesis could have been counted on a single hand.

Thank you to each of the advisors who helped determine my career change: Jake Palmer, Dr. Clark Jeffries, Dr. Mohammad Peyravian, Dr. John Sakon, Dr. Kerry Jordan, Dr. Kelly Giovanello, Dr. Lila Davachi, LaKissa Bright, Rhonda Hicks, and Paul Himberger.

I appreciate my lab mates, Dr. Kathrin Müsch and Dr. David Groppe, helping me better understand the field of neuroscience and the idiosyncrasies of Matlab.

Thank you to everyone who has supported and assisted in my transition from neuroscience neophyte to slightly smarter student!
# Table of Contents

Acknowledgments .......................................................................................................................... iii  
List of Tables ................................................................................................................................ vii  
List of Figures .............................................................................................................................. viii  

1 Motivation and Hypotheses ........................................................................................................ 1  

2 Background and Relation to Existing Work .............................................................................. 4  
2.1 Empirical Findings in the Visual System ............................................................................ 4  
2.2 Memory & Perception ......................................................................................................... 6  
2.3 Spatiotemporal Hierarchies ................................................................................................. 9  
2.4 Spike Timing ..................................................................................................................... 11  
2.5 Synaptic Plasticity Basics ................................................................................................. 12  
2.6 Statistical Learning ........................................................................................................... 12  
2.7 Computational Models of Brain and Mind ....................................................................... 14  

3 General Methods ...................................................................................................................... 15  
3.1 Human Primary Visual Cortex (V1) ................................................................................. 17  
3.2 Modeling a V1 cortical circuit in silico ............................................................................ 17  
3.3 Izhikevich Neurons and Neuronal Dynamics ................................................................... 23  
3.4 Background Activity ......................................................................................................... 24  
3.5 Synaptic Plasticity ............................................................................................................. 25  
3.5.1 Spike-Timing Dependent Plasticity (STDP) ......................................................... 25  
3.5.2 Heterosynaptic plasticity ................................................................................. 27  
3.6 Stabilization Period ........................................................................................................... 28  

4 Computational Experiments ..................................................................................................... 30  
4.1 Experiment 1: Effects of Repeating a Single Stimulus ..................................................... 31
4.3.3.2 Training Stimulation ........................................................................ 51
4.3.3.3 Anatomical Changes from Training ........................................ 51
4.3.3.4 Stimulation Procedure for Experiment 3 .................................... 54
4.3.4 Functional Changes Resulting from Training ..................................... 55
4.3.5 Discussion for Experiment 3 ............................................................. 57

5 General Discussion .................................................................................... 57
6 Future Directions and Model Predictions .................................................. 59
References .................................................................................................... 61
List of Tables

Table 2.2.1: Memory & Perception Definitions ................................................................. 8
Table 3.2.1: Parameters Used For Anatomy ................................................................. 20
Table 3.2.2: Synaptic Weights at Initialization .............................................................. 22
Table 3.3.1: Parameters for Izhikevich Dynamics ......................................................... 24
Table 3.4.1: Poisson Parameters .................................................................................. 26
Table 3.5.1.1: STDP Parameters .................................................................................. 27
Table 3.5.2.1: Heterosynaptic Parameters ................................................................... 29
Table 4.1.3.4.1: Mean Excitatory-Excitatory (E-E) Synaptic Weights after Training .... 36
Table 4.1.3.5.1: Training Regime and Test Case Pairing Possibilities ......................... 37
Table 4.1.4.1: Average Firing Rates for Excitatory Neurons in the 40ms Post-Stimulus across all neurons (Hz) .............................................................. 38
Table 4.1.4.2: Mean Firing Rates (Hz) for Training Regimes Presented Stimulus A (40ms Post-presentation) ........................................................... 40
Table 4.1.4.3: Mean Firing Rates (Hz) for Training Regimes Presented Stimulus B (40ms Post-presentation) ........................................................... 40
Table 4.2.3.2.1: Parameters used to calculate firing rate ............................................. 45
Table 4.3.3.3.1: Average E-E Synaptic Weights ......................................................... 54
# List of Figures

Figure 1.1: Theoretical Construct of Lengthening the Window of Temporal Integration .............. 2

Figure 2.3.1: Rough Schematic of Hierarchical Visual Processing Underlying Object Recognition ................................................................................................................................. 10

Figure 2.6.1: An example of statistical learning ............................................................................ 15

Figure 3.1: Experimental Design ................................................................................................ 17

Figure 3.2.1: Functional Clusters and Neuron Types within the modeled Cortical Circuit ....... 19

Figure 3.2.2: Axonal delays between neurons ............................................................................ 22

Figure 3.2.3: Synaptic Connections and Weights after Initialization ........................................ 23

Figure 3.2.4: Synaptic Weights Distribution after Initialization ............................................... 24

Figure 3.5.1.1: STDP Explanation ............................................................................................ 27

Figure 3.6.1: Synaptic Weight Matrix after Stabilization ......................................................... 30

Figure 3.6.2: Synaptic Weight Distribution after Stabilization ............................................... 31

Figure 4.1.3.2.1: Example Spike Raster and Firing Rate for Visually Stimulated Training Regime ........................................................................................................................................................................ 34

Figure 4.1.3.3.1: Example Spike Raster and Firing Rate for Unstimulated Training Regime ..... 34

Figure 4.1.3.4.1: Distribution of Excitatory-Excitatory (E-E) Synaptic Weights ......................... 35

Figure 4.1.3.4.2: Synaptic Weight Change from Baseline to Post-Training ............................... 36

Figure 4.1.4.1: Average firing rates for 200ms following presentation of stimulus A ................. 38

Figure 4.1.4.2: Average firing rates for 200ms following omission of stimulus at stimulation onset .............................................................................................................................................. 39

Figure 4.1.4.3: Average firing rates for 200ms following presentation of Stimulus B ............. 40

Figure 4.1.4.4: Synaptic Weight Changes when for each training regime ............................... 41

Figure 4.2.3.3.1.1: Possible 3-motif connections ....................................................................... 46

Figure 4.2.4.1.1: Classification Accuracy and Associated Firing Rates ................................. 47
Figure 4.2.4.1.2: Classifier Accuracy using average cluster firing rates ........................................48
Figure 4.2.4.2.1: Changes in 3-neuron Motifs & Cluster Coefficients for Trained Regimes .....49
Figure 4.3.3.3.1: Synaptic Weight Distribution for E-E Synapses ........................................52
Figure 4.3.3.3.2: Synaptic Weight Change for Clusters A & B ........................................53
Figure 4.3.3.3.3: Differences strong synapses between clusters A&B .....................................54
Figure 4.3.3.3.4: Motifs and Clustering Coefficients for the sequence training regimes (w>0.33) ..................................................................................................................................................55
Figure 4.3.3.4.1: Spike Raster of a Random Tested Sequence ........................................56
Figure 4.3.4.1: Average Firing Rates for the Network trained on ABCD ..........................57
Figure 4.3.4.2: Average Firing Rates for the Network trained on Random Sequences .........57
Figure 4.3.4.3: Average Firing Rate by Time Bin of Network trained with Repeated ABCD Sequence ..................................................................................................................................................57
1 Motivation and Hypotheses

Have you ever glimpsed something moving behind some bushes and then gradually, as you gather more information, realized what it was? How was your brain able to combine the sequence of visual data into the eventual recognition of an object?

Over the course of a few seconds, you integrated the initial and subsequent visual features from discrete and overlapping epochs of milliseconds (VanRullen and Koch, 2003) into a single conceptual object. The most parsimonious explanation would be that each batch of sensory input is processed the same way from one second to the next, yet our current, “online”, processing of sensory inputs is often shaped by sensory experiences over previous minutes, seconds, and even milliseconds of time (Hasson et al., 2008; Lerner et al., 2011).

Traditionally, the ability for prior information to affect processing in the present moment has been assumed to rely on specialized “memory systems”, but more recent evidence suggests that temporal integration occurs in “perceptual” systems of the brain as well (Gavornik and Bear, 2014; Hasson et al., 2015). The capacity for temporal integration is an unexpected property in primary sensory cortices because they are typically characterized as feedforward systems in which incoming information is ephemeral and must be quickly processed to avoid decay or interference (Christiansen and Chater, 2015).

Given the evidence for perceptual circuits integrating information over time (Hasson et al. 2015; Gavornik & Bear, 2014, as elaborated in Section 2), we asked, “How do cortical circuits acquire the ability to integrate information over time?”

A particular neural circuit is able to integrate information over time when exposure to an earlier stimulus affects the processing of a subsequent stimulus. Importantly, the ability to integrate information over time appears to be much greater for sequences of information that are familiar and inter-related. Thus, we propose to test a theory (Figure 1.1) in which cortical circuits learn to integrate information over longer timescales (during online processing) via repeated prior exposure to stimuli in a form of perceptual or statistical learning. Specifically, we suggest a model in which:
(1) The local cortical circuit is repeatedly exposed to the same stimulus, resulting in an increased neuronal response for that stimulus.

(2) The increased amplitude of the neuronal response also lengthens the trace of reverberating activity in the circuit, and thus the duration of circuit “memory” for the stimulus.

(3) The lengthening circuit “memory” for the stimulus enables the formation of new associative links between that stimulus and other, subsequently presented stimuli.

Thus, over time, an individual stimulus that occurs frequently will produce elevated responses (greater and more lasting firing). Similarly, over time, we should observe greater associative responses from sequences that are presented frequently (e.g. in the sequence of letters “jinx” versus “xnij”).

Figure 1.1: Theoretical Construct of Lengthening the Window of Temporal Integration

(Single Stimulus Exposure) Repeated exposure to a single stimulus leads to an increased response to that stimulus and an increase in the resonant information of that stimulus in local cortical circuits (i.e. process memory)

(Sequence Exposure) A repeated sequence of stimuli with constant timing between elements results in increased responses to each of the stimulus elements individually. However, the elongation of the process memory resulting from more opportunities to learn associations will bias the subsequent sequence presentations with gradually increasing opportunities to learn associations between sequence elements in a recurrent fashion.
The increased response suggested in (1) above is not especially controversial. For over a century, we have known that repeated exposure to perceptual information results in behavioral changes (Whipple, 1910). This experience-dependent learning depends on changes in cortical assemblies via plasticity at the level of synapses and circuits (Medini, 2014). We propose that one fundamental mechanism driving the change in response to repeatedly exposed stimuli is a selective increase in synaptic strength from stimulus-driven neurons.

The second point of our schematic model requires that the increase in the magnitude of stimulus response is accompanied by an increased duration in the stimulus response. This response contains a circuit memory, or “process memory”, in that the neuronal ensemble activated contains some information about the previously presented stimulus. We suggest that the persistence of the stimulus-related information (encoded in subsequent dynamics) will increase as the firing response to the stimulus is lengthened.

The third and final requirement is that directed associative learning occurs between elements of the sequence. Consider the simple sequence, A->B, in which stimulus A is followed by stimulus B after 50 milliseconds. The proposed mechanism relies on potentiation of outgoing synapses from neurons most immediately responsive to A, resulting in a larger population response amongst the assembly of downstream neurons, and thus, a longer duration of the population response to stimulus A. As the duration of the response to stimulus A increases, this creates more opportunities for an overlap in firing between neurons responsive to A and neurons responsive to B. The increased overlap in firing between these neuronal ensembles allows for selective strengthening of the synapses from “A” neurons to “B” neurons.

Sequence learning requires the formation of links that are directed. Exposure to A->B should produce different synaptic changes and behavioral effects from exposure to B->A. Thus a desirable property of a mechanism that builds associations between neuronal ensembles is that the associations should only form in the temporal direction of the presented stimuli. Luckily there exists an asymmetric synaptic plasticity mechanism – spike-timing dependent plasticity (STDP) -- that modifies synaptic strength depending upon the relative timing of pre and post-synaptic spikes. We will employ STDP in our model of sequence learning.

If repeatedly presenting visual stimuli to a circuit leads to changes of the kind described in our schematic model (Points 1-3 above, and Figure 1.1), we still need to explain the effect on
online processing. Thus our second research question is: “if circuit responses to a repeated stimulus change as suggested above, how does this support the ability of cortical circuits to integrate information from that stimulus over time?”

To begin to answer this question, if indirectly, we propose that the persistence of information in a circuit during online processing of a sequence should vary as a result of the amount of previous learning of the particular stimulus sequence being processed. This information persistence, or process memory, can result in a circuit that produces a larger and more sustained response to one ordered sequence (e.g. “ABCD”) as compared with its reverse (e.g. “DCBA”), even in cases when the circuit has been exposed to the individual stimulus elements an equal number of times.

In this manuscript, we formalize and test the schematic model above by implementing it as a computational model of a small visual cortical circuit. In this way, we investigate both (i) how cortical circuits acquire the ability to integrate information over time and (ii) how the responses of a circuit change as the “process memory” lengthens for the elements of the learned sequence. We expect that, given repeated exposure to a stimulus or sequence:
(1) More neurons will fire from stimulus (sequence) presentation compared to baseline
(2) The local cortical circuit will be able to assess stimulus (sequence element) presentation at a point further into the future
(3) A repeated sequence of stimuli will induce a larger neuronal response in the trained condition compared to the reversed sequence

These experiments can inform our understanding of how cortical circuits are able to integrate information over time for specific stimuli, and how they acquire the ability to do so.

2 Background and Relation to Existing Work

2.1 Empirical Findings in the Visual System

Visual perceptual learning is often investigated using oriented and localized patches of contrast called “Gabor patches”, which elicit robust responses in early visual cortices in most
mammalian brains (Gabor, 1946; Daugman, 1985). A Gabor patch example can be found in Gavornik & Bear (2014), Figure 1b. These learning studies regularly provide evidence for experience-dependent plasticity (Polat and Sagi, 1994). In the case of repeated exposures to the oriented Gabor stimulus, an increased neuronal response is typically observed for subsequent presentations of the learned stimulus. For example, one study suggested that adult mice have increased neural responses to horizontal lines versus vertical lines because of their visual experiences in the world (Frenkel et al., 2006).

Building on these studies of how early visual circuits respond to repeated exposures to the same orientations, more recent work has probed how neural circuits are reshaped by repeated sequences of orientations (Gavornik and Bear, 2014). Gavornik & Bear repeatedly presented mice with a sequence of four Gabor patches (total sequence length: 600ms), and measured the local field potential response in primary visual cortex (V1) of these mice. After training, they observed increased response amplitudes to the learned sequence compared to a reverse ordering of the same stimulus sequence. By contrast, mice trained with random non-repeating sequential combinations of the same four Gabor patches did not show a preferential response to the experimental repeated sequence compared to the reverse. The Gavornik & Bear (2014) results can be found in Figures 1c & 1e of that paper. In this way, the results of Gavornik & Bear (2014, hereafter “GB14”) extend the classical perceptual learning paradigm from learning about features in space to learning about features in space and time.

The observation of selective responses in primary visual cortex (V1) to these 600 ms sequences is surprising if V1 is considered as a kind of feedforward, specific feature detector. Although studies have shown reinstatement of prior activity patterns in sensory regions (Polyn, 2005), low-level regions like V1 are still commonly treated as transformational relays of thalamic input to higher order cortical regions. As Gavornik and Bear point out, if primary visual cortex can retain “memory” of sequences, then the ability to respond selectively to the temporal structure of input may be a ubiquitous property of sensory circuits, rather than a process that is relegated to dedicated “memory regions”.
2.2 Memory & Perception

How does learning affect memory and perception? In the GB14 study, neuronal responses to a previously presented sequence are increased, while responses to the same stimuli with a different ordinal sequence or timing are decreased. This implicit recognition of previously seen elements and their specific timing belies simple perception of sensory information in the subjective present.

Memory is the manifestation of learning. In the GB14 experimental condition, “learning” is the process of sensitization to the sequence of Gabor patches and “memory” is the increased neural response for that condition versus random sequences. Given the differential neural responses, we consider the mouse to have learned something about these oriented line segments and implicitly expressed “memory”. However, classic definitions of memory deal more with overt behavior as opposed to simple awareness of external information in the “subjective present” (i.e. perception) (Fairhall et al., 2014). Even further removed from the classic concept of memory are differential neuronal responses to trained vs untrained stimuli, which do not manifest in behavior, despite the fact that these responses result from learning. Although it has long been appreciated that perceptual learning is possible (Woodworth and Thorndike, 1901; Foster, 1911), the location of the underlying memory has remained unclear (Schoups et al., 1995). As neural recording techniques have improved, demonstrable behavior has shifted from observable effects to the human eye (e.g. declarative memory, response time) (Gibson, 1953), to measuring brain waves (e.g. using electrocorticography) and neuronal spiking (e.g. using multi-electrode arrays), thus further blurring the line between memory and perception.

Inherent in the classic psychological distinction between memory and perception is the implication that mnemonic and sensory process are associated with distinct neural processes, and perhaps distinct brain regions. For example, working memory theories have evolved over the past 40 years to postulate the existence of numerous distinct memory buffers (Baddeley, 2012), with neural substrates that are separate from the regions involved in “primary” or perceptual processing. This neural separation has grown out of the separation of memory and perceptual processes within psychological models of memory, but it is unclear whether it is a natural decomposition of the underlying neural circuits. Indeed, there have been may theoretic proposals over the decades in which sensory information is integrated “online” (i.e.
continuously) into existing cortical memories, (Hayek, 1952; Fuster, 1997); the most recent of these proposals is the “Process Memory” framework (Hasson et al., 2015).

The Process Memory framework postulates that some aspects of memory are intrinsic to virtually all perceptual processing, and that temporal integration occurs continuously over many timescales at different levels of the cortical hierarchy. Common cortical mechanisms underlie the processing of sensory information and allow the possibility that although some information is processed in conjunction with executive and attentional systems, those systems are not required for all processing (c.f. Just and Carpenter, 1992).

Another confusing aspect of “memory” are potentially overlapping terms: “short-term memory” (STM), “working memory” (WM), “long-term memory” (LTM), “sensory memory”, “perception”, and now “process memory” (Baddeley, 2012; Fuster, 1997; Cowan, 2008; Jonides et al., 2008; Hasson et al., 2015). We define these terms in Table 2.2.1 by specifying whether the memory is “persistent” (i.e. whether a neural representation of the memory is persistently accessible in the future) and the duration of the memory.

Table 2.2.1: Memory & Perception Definitions

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Persistent?</th>
<th>Duration of Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term Memory</td>
<td>Recollection of information presented in a temporally proximate instance</td>
<td>No</td>
<td>Seconds - Minutes</td>
</tr>
<tr>
<td>Working Memory</td>
<td>Mental manipulation of information that is either currently presented or intrinsically queried (Baddeley, 2012)</td>
<td>No</td>
<td>Seconds - Minutes</td>
</tr>
<tr>
<td>Long-term Memory</td>
<td>Recollection of information presented in the subjective “past”</td>
<td>Yes</td>
<td>Minutes - Years</td>
</tr>
<tr>
<td>Sensory Memory</td>
<td>Recollection of information presented via a specific modality (can span STM, WM, &amp; LTM) (Harris et al., 2002; Weinberger, 2004). Can result from perceptual learning (Pearson and Brascamp, 2008).</td>
<td>Yes/No</td>
<td>Seconds - Years</td>
</tr>
<tr>
<td>Perception</td>
<td>Online awareness of environmental information in the subjective present</td>
<td>No</td>
<td>Milliseconds - Seconds</td>
</tr>
<tr>
<td>Process Memory</td>
<td>Systems-level description of neuronal activation in response to stimuli</td>
<td>No</td>
<td>Milliseconds - Minutes</td>
</tr>
</tbody>
</table>

While the distinctions in Table 2.2.1 can be useful, these distinctions tend to be slightly arbitrary in that LTM consists of items that can be remembered after a threshold amount of time and the others typically cannot. Process memory is fundamentally different from the other terms in that it is a systems level description that implies inherent information in the circuit that cannot
be explicitly recollected or manipulated in working memory. Process memory is akin to a form of implicit long-term memory, but differs in that it is a response to stimuli and cannot be internally queried. All of the other memory terms can be considered a continuum of cortical consolidation, whereas process memory is the output of a circuit that has experienced cortical consolidation.

The psychological notions of “sensory memory” and “short-term memory” can both arise as a result of “process memory” within sensory circuits. The ability for a circuit to contain information about recent events, especially since it is heavily disruptable by new information or time, is a good match for the classic definition of short-term memory (Atkinson and Shiffrin, 1968). Perceptual learning appears to occur in a modality-specific hierarchy (Harrison and Tong, 2009), and sensory memory can be stored solely in primary sensory cortices. Sensory memories have been shown in auditory primary cortex (Alain et al., 1998; Weinberger, 2004), somatosensory primary cortex (Harris et al., 2002), and visual primary cortex in mice (Frenkel et al., 2006; Gavornik and Bear, 2014). In Alain et al, hippocampal lesions did not affect performance, supporting the theory that sensory memory can be independent of hippocampal activity. Additionally, many have shown that regions used for processing sensory input are the same as those used in memory encoding and retrieval (Fuster, 1997; Harris et al., 2001; James and Gauthier, 2003). All of these studies support the theory that sensory processing should not be conceptually separated from cortical memory formation.

Classical frameworks for memory propose that associative learning happens in the hippocampus, motor movements are learned by a basal ganglian-cerebellar partnership, and episodic memories are recalled via reinstatement of prior distributed cortical states. In this classical framework, modality-specific sensory processing is an input stage for the processing hierarchy and does not possess a memory capacity of its own (McClelland and O’Reilly, 1995; Serre et al., 2007; Jonides et al., 2008; Baddeley, 2012). However, recent studies have demonstrated that sequence learning and prediction are possible even pre-cortically, allowing that the same type of mechanism could occur in the “higher order” sensory cortices (Schwartz et al., 2007).
2.3 Spatiotemporal Hierarchies

Sensory processing appears to occur over time in an anatomical hierarchy specific to each modality. For example, in most vertebrate brains, visual features are processed within a hierarchy of ascending feature selectivity and complexity, ranging from basic features such as orientation in V1 to object representation in Inferior Temporal cortex (Hubel and Wiesel, 1962; Ewert, 1989; Goodale, 2011) (Figure 2.3.1 shows an abstracted example). Anatomical hierarchies have also been shown in imaging studies in the integration of cross-modal information in “higher order” regions (e.g. audiovisual processing in Superior Temporal Gyrus and Middle Temporal Gyrus) (Zatorre et al., 2007; Alpert et al., 2008; Honey et al., 2012).

The Process Memory framework proposes that shorter integration intervals (30-250ms) capture low-level (“less specific”) visual information (i.e. general environmental statistics) and longer integration intervals (500ms+) are associated with processing “more specific” visual information (i.e. slowly-evolving, less transient environmental objects or concepts). Within the Process Model framework, multimodal and amodal regions at the apex of the cortical hierarchy may exhibit temporal integration on the timescale of many tens of seconds, or even minutes (Chen et al., 2015).

Figure 2.3.1: Rough Schematic of Hierarchical Visual Processing Underlying Object Recognition
These anatomical hierarchies appear to have analogous (and potentially overlapping) functional hierarchies based upon the time scale of information integration (Hasson et al., 2008). Overlapping timescales of information processing have been shown for various cognitive tasks in macaques (Murray et al., 2014), where longer timescales (200-350ms) correlate with brain regions “higher” in the cortical hierarchy and more perceptual circuits at the bottom of the cortical hierarchy have shorter timescales (50-200ms).

In language comprehension, different neural regions are activated dependent upon whether sounds, words, sentences, or paragraphs are perceived (Xu et al., 2005). The lower order region responsible for audition (primary auditory cortex; A1) has a relatively short window for integration of sounds (<1 second) as compared to the STG (which has an intermediate window for integration) or the precuneus (>30 seconds). This is further supported by comparisons between subjects showing common neural activation across participants in auditory processing involving ascending cortical regions as elements of a story unfold (Lerner et al., 2011). Figure 3 of the Lerner et al paper shows a neural map of auditory information integration over various timescales. Common cortical circuitry, agnostic to timescale, could manifest as neural activity in different neural regions at progressively longer timescales. The observed variance on information integration windows is then an emergent property of what is processed and not how it is processed.

Despite the existence of sequence learning throughout all modalities (e.g. movies, speech, somatosensory) (Conway and Christiansen, 2005), humans are better at integrating information over time for some sequences relative to others. Consider the following example; imagine trying to learn the word “engreconition”, a new five syllable English word. This is likely less difficult than learning “sweezooshnix” despite the latter word comprising fewer letters and syllables. The reason for the increased ease of learning engreconition is non-obvious, but the Process Memory framework suggests that it is easier to remember because of an implicit processing bias in the phonemic associations that permit us to learn the sequence of syllables faster than sweezooshnix. This implicit processing bias does not imply a subconscious processing of the phonemic associations because we know it’s an English word, but rather the processing itself is easier for engreconition because of the underlying neuronal scaffolding and resultant process memory. In essence, the subsequent phonemes are primed (or partially predicted) for one word, but not the other.
In this project we seek to understand the mechanisms that may operate in local cortical circuits to support the integration of information over time. If these mechanisms are general, they may be instantiated repeatedly at multiple levels of the anatomical cortical hierarchy, giving rise to a corresponding functional hierarchy of temporal integration. GB14 allows us to investigate how a distributed cortical hierarchy could function by modeling how a single level in a singular cortical hierarchy may operate (i.e. V1 in the visual modality). We propose that the implicit time receptive window (TRW) associated with neural circuits is not a region-specific process because it is mediated by a central system or necessary differences in connectivity, but rather because it follows a general cortical outline.

2.4 Spike Timing

The directionality of learning a sequence of elements is only possible when the relative timing between elements is known. It is unclear whether rate-coding models (looking at action potentials over tens or hundreds of milliseconds) permit the detailed examination of network activity that spike-coding (looking at action potentials at millisecond precision) allows. Beyond simple association, characterizing changes in neural connectivity beyond correlational measures is not possible with rate-coding. Motifs forming in the network and changing network dynamics emerging from the learning process may require detailed timing information (Sjöström et al., 2001). Previous work in spike timing suggests millisecond precision may be crucial for some cortical communication and that the fidelity with which the first spike arrives from a stimulus may bias the subsequent discrimination of what stimulus was presented (VanRullen et al., 2005). Additionally, experiments in vitro suggest that small cortical circuits are able to learn specific timing associated with stimulus presentation (Johnson et al., 2010).

A final point related to spike timing is the concept of synfire chains, a collection of neurons that fire in a stereotypical firing pattern (Abeles, 1982). This concept is related to the ability of a neural system in producing an elongation of neural activity beyond the temporal window previously associated with a specific stimulus response. The gradual elongation of memory associated with that stimulus response could be the result of specific neurons entraining downstream neurons in a building process. Building a coalition of neurons associated with a stimulus response could result from synaptic plasticity.
2.5 Synaptic Plasticity Basics

Synaptic plasticity can form and reshape pathways for neural activity propagation in cortical circuits. The reshaping of pathways is the entrainment of neuronal ensembles into associative neuronal formations, serving as the underlying mechanism of learning. There are multiple mechanisms of synaptic plasticity operating in parallel in the cerebral cortex, the most well-known of which is Hebbian synaptic plasticity (Hebb, 1949). In Hebbian plasticity, synapses between neuronal ensembles, A and B, will be potentiated when the neurons in the two ensembles fire in temporal proximity. Importantly, the connections formed via standard Hebbian plasticity are bidirectional: the synapses projecting from A to B are altered in the same way as the synapses from B to A. Thus Hebbian learning suggests a kind of symmetric (non-directional) associative learning between ensembles. Symmetric learning is broadly inconsistent with the behavioral findings (McMahon and Leopold, 2012) and biological findings (Pfister and Gerstner, 2006) of directed associations, and is also less conceptually useful for learning directed contingencies (e.g. the word part “tion” is much more often a suffix than a prefix). In order to capture the temporal statistics of environmental input, and to allow the formation of directed associations, we employ an empirically well-supported (Feldman, 2012; Markram et al., 2012) plasticity paradigm known as Spike-Timing Dependent Plasticity (STDP). In STDP, the changes in synaptic weight depend not only on whether two neuronal ensembles fire in temporal proximity, but also on which ensemble fires first. In this way, STDP enables the formation of directed association, which matches empirical observations more closely (Bi and Poo, 1998; Caporale and Dan, 2008). The mechanisms of synaptic plasticity, and the specific STDP model employed, are covered in greater detail in Section 3.5.

2.6 Statistical Learning

Statistical learning is the study of how neural circuits are reshaped by the statistics of input from the environment. The structure of sights, sounds, and movements in our world can all be statistically described, and so in a sense all learning is statistical. However, the field of statistical learning has focused primarily on learning that occurs without awareness or volition, a process exemplified in the ability of infants to acquire information like linguistic structure via passive exposure (Saffran, 2001; Wu et al., 2011). Partly because of its early connections to
children’s language learning, statistical learning has been unusually heavily focused on the
statistics of information over time. Thus, although most of the field is framed in terms of
psychological theory rather than neural mechanisms, the experiments in this field are highly
relevant to our question of how neural circuits gain the ability to integrate information over time,
and how acquired temporal associations shape our perception (Nissen and Bullemer, 1987; Shin
and Ivry, 2002; Keele et al., 2003).

Statistical learning has been observed across many non-humans, including rhesus
monkeys (Blanchard et al., 2014), songbirds (Brenowitz and Beecher, 2005), and mice
(Gavornik and Bear, 2014). Despite the variability inherent in inter-species comparisons, there
appear to be commonalities between the neural circuits utilized in statistical learning (Wilson et
al., 2013).

“Statistical Learning” and “Sequence Learning” both describe processes of generating
associative connections between temporally ordered elements. A common example is the
combination of phonemes into syllables and syllables into words (Figure 2.6.1). The
combination of phonemes and syllables is only a word when combined in a specific spatial order
and within some temporal bounds between each element. These associations can be formed
between body movements (Schubotz, 2007), musical elements (Zatorre et al., 2007), gestures
(Maurer et al., 2005), rhythms (Watanabe et al., 2007), tones (Saffran et al., 1999), tactile
pressure (Conway and Christiansen, 2005), words (in speech), and likely anything else that can
be perceived or imagined.

“Eh” — “ko” — “low” — “kay” — “shun”

Echolocation

Figure 2.6.1: An example of statistical learning

“Sequence learning” typically concerns the learning of deterministic sequences; it is
therefore a special case of “statistical learning” which allows for the learning of probabilistic
associations between elements. An example of statistical learning (i.e. non-deterministic) is
speech segmentation. This phenomena is demonstrated in infants, adults, and non-native
language speakers where highly variable phonemes must be parsed spatially and temporally into
words and sentences (Mattys et al., 1999; Kaan et al., 2007). Thus, the learning of the deterministic sequences presented during training in GB14 is an example of sequence learning. While the associations between elements are recognized as sequences most intuitively, the timing between the elements (i.e. the inter-stimulus interval, or ISI) is just as important in the formation of a sequence.

Statistical learning appears to be modality-specific (Conway and Christiansen, 2006). Since most modalities have inherent hierarchies, this statistical learning finding appears to fit well with the suggestion of information integration over different timescales across a cortical hierarchy. The apices of these hierarchies are likely responsible for cross-modal learning and there is some evidence that cross-modal statistical learning can occur (Altmann et al., 1995). Thus, a higher cortical region has the opportunity for signals to be assessed as co-occurring percepts in time. We focus within the visual modality for parsimony and the ability to leverage known anatomy and timing in the existing literature.

2.7 Computational Models of Brain and Mind

In neuroscience, computational modeling serves two broad purposes: (i) to precisely link quantitative assumptions to quantitative predictions; (ii) to provide a simulated experimental basis to explore neuroscientific concepts and processes that are not feasibly studied in biological models due to temporal, financial, or ethical constraints. However, computational models of brain function are highly variable in the degree to which they emulate the biological brain, with some aiming to model it at a biophysically realistic level while others attempt to capture more abstract but essential features of neural dynamics (Vogels et al., 2005; Gerstner et al., 2012.).

Prior computational modeling has investigated how the sequential structure of neural dynamics shapes biological function (e.g. models of synaptic change due to spike timing dependent plasticity) (Izhikevich and others, 2003; Izhikevich, 2006; Hosaka et al., 2008). At a more abstract, functional level, aspects of ordinal sequence learning have also been studied in generic coupled network models (Sussillo and Abbott, 2009).

Some of these more functional models have influenced the Process Memory framework and other theories of how temporal information is processed hierarchically. For example, Botvinick showed how the sensory and motor information in Fuster’s hierarchy (Fuster, 1997)
could be processed serially using sequential input (Botvinick, 2007), akin to our suggestion of sensory processing and memory. However, there has been little work aiming to understand sequence learning in terms of neuronal and circuit level mechanisms.

A few studies have attempted to link the temporal dependence of biological learning with the broader functional questions regarding temporal structure (Litwin-Kumar and Doiron, 2014; Veliz-Cuba et al., 2014). Each of these models, like we do, suggest that synaptic plasticity in an experience-dependent manner can accurately reproduce previously seen sequences of stimuli.

In this thesis, we aim to place the Process Memory model on a more solid mechanistic footing, by jointly considering the biological and functional aspects of learning and processing information over time. We ask: (i) how do cortical circuits acquire the ability to integrate information over time? And (ii) how do the responses of a circuit change as a sequence is learned and the temporal integration window lengthens? In the next section, we introduce the modeling framework that is used to address these questions.

3 General Methods

The Experimental Design is summarized in Figure 3.1. Network Initialization and Stabilization are the same for all experiments, but Training and Testing vary by experiment.
Each experiment measured the changes in a local cortical circuit resulting from repeated exposure to particular stimuli.

The General Methods section focuses on the first two steps of the experimental design: Initializing the Neural Network (i.e. What was modeled and How) and Stabilization of the Neural Network. The details of Training and Testing for each computational experiment, along with the individual results, are provided in dedicated subsections within Section 4.
3.1 Human Primary Visual Cortex (V1)

The experiments undertaken by Gavornik & Bear (2014) investigated the effects of repeated exposure of sequential stimuli on visually evoked responses within mouse primary visual cortex. Both mouse and human primary visual cortices comprise multiple clusters of neurons which share common functional processing. For example, spatially proximate clusters of neurons are selective for common stimulus orientations (Bosking et al., 1997). The multicolored pattern shown in Figure 1b of Bosking et al (1997) illustrates eight distinct oriented edges and the neural clusters most responsive to that orientation. We chose to model four such orientation clusters in a localized region to permit investigation of the dynamics associated with response to oriented edge stimuli.

3.2 Modeling a V1 cortical circuit in silico

A population of 2000 Izhikevich neurons was spatially organized in a ring. The ring was divided into four cortical quadrants (Figure 3.2.1), inspired by the distinct biological domains shown in Bosking et al (1997). The neurons in each quadrant are referred to as a “cluster”, and each cluster responds to direct input from external stimuli of a specific orientation (Figure 3.2.1, legend).
Figure 3.2.1: Functional Clusters and Neuron Types within the modeled Cortical Circuit

Following standard approximations for the proportions of neuronal types (Bush and Sejnowski, 1996; Hosaka et al., 2008; Izhikevich, 2006), excitatory neurons (80% of neurons, shades of orange in Fig 3.2.1) are intermixed spatially with inhibitory neurons (20% of neurons, shades of blue in Fig 3.2.1). Each cluster contains the same number of neurons ($N_{\text{per}}$) and the same proportion of inhibitory and excitatory neurons.

We use spatially homogeneous functional clusters with the belief that there is some spatial clustering in feature selectivity in most cortical regions, and thus any results are potentially generalizable to the cortex at large. Additionally, the modularity makes it easier to implement functionally specific lateral interactions, e.g. cross-orientation suppression.

The random two-dimensional coordinates of neurons within each cluster were assigned according to the following formulas:

$$x = cc - r_o \cdot \cos(o \cdot (C-1) + \delta_{\text{spread}}) \cdot \delta_{\text{depth}}$$  \hspace{1cm} \text{Equation 3.2.1}

$$y = cc - r_o \cdot \sin(o \cdot (C-1) + \delta_{\text{spread}}) \cdot \delta_{\text{depth}}$$  \hspace{1cm} \text{Equation 3.2.2}

where $x$ is the 2D coordinate on the x-axis, $y$ is the 2D coordinate on the y-axis, $cc$ is the center of the ring, $r_o$ is the radius of the outer edge of the ring, $o$ is the number of degrees in a cluster, $C$
is the ordinal number of the cluster to which the neuron belongs, $\delta_{\text{spread}}$ is the lateral spatial spread permissible for neurons within a cluster, $\delta_{\text{depth}}$ is the maximum permissible distance from $cc$. The variables $cc$, $r_o$, and $o$ are constants, while $\delta_{\text{spread}}$ and $\delta_{\text{depth}}$ are random variables which were drawn uniformly from a predefined range of permissible values (see Table 3.2.1).

**Table 3.2.1: Parameters Used For Anatomy**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Number of Neurons</td>
<td>2000</td>
</tr>
<tr>
<td>$N_{\text{exc}}$</td>
<td>Number of Excitatory Neurons</td>
<td>1600</td>
</tr>
<tr>
<td>$N_{\text{inh}}$</td>
<td>Number of Inhibitory Neurons</td>
<td>400</td>
</tr>
<tr>
<td>$N_{\text{clust}}$</td>
<td>Number of Clusters</td>
<td>4</td>
</tr>
<tr>
<td>$N_{\text{per}}$</td>
<td>Number of Neurons per Cluster</td>
<td>500</td>
</tr>
<tr>
<td>$cc$</td>
<td>Circle Center ($N/2,N/2$)</td>
<td>(1000,1000)</td>
</tr>
<tr>
<td>$r_o$</td>
<td>Outer Radius ($N/4$)</td>
<td>500</td>
</tr>
<tr>
<td>$r_i$</td>
<td>Inner Radius</td>
<td>0.2</td>
</tr>
<tr>
<td>$o$</td>
<td>Degrees per Cluster ($360/N_{\text{clust}}$)</td>
<td>90</td>
</tr>
<tr>
<td>$C$</td>
<td>Cluster being calculated</td>
<td>$[1, N_{\text{clust}}]$</td>
</tr>
<tr>
<td>$\delta_{\text{spread}}$</td>
<td>Randomization Term for spread within cluster</td>
<td>$[0, o]$</td>
</tr>
<tr>
<td>$\delta_{\text{depth}}$</td>
<td>Randomization Term for distance between the outer and inner circles</td>
<td>$[r_i, r_o]$</td>
</tr>
<tr>
<td>$k_{\text{exc_intra}}$</td>
<td>Outbound synapses from excitatory neurons to intracluster neurons</td>
<td>110</td>
</tr>
<tr>
<td>$k_{\text{exc_extra}}$</td>
<td>Outbound synapses from excitatory neurons to extraccluster neurons (average synapses to each extraccluster)</td>
<td>30</td>
</tr>
<tr>
<td>$k_{\text{inh_intra}}$</td>
<td>Outbound synapses from inhibitory neurons to intracluster neurons</td>
<td>92</td>
</tr>
<tr>
<td>$k_{\text{inh_extra}}$</td>
<td>Outbound synapses from inhibitory neurons to extraccluster excitatory neurons (average synapses to each extraccluster)</td>
<td>36</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>Euclidean distance between two neurons</td>
<td>Variable</td>
</tr>
<tr>
<td>$\overrightarrow{\text{dist}}$</td>
<td>Vector of the sum of all distance values from each neuron</td>
<td>Variable</td>
</tr>
<tr>
<td>$\overrightarrow{\text{prox}}$</td>
<td>Proximity factor used to normalize distances between two neurons</td>
<td>Variable</td>
</tr>
<tr>
<td>$\overrightarrow{\text{prox_max}}$</td>
<td>Maximum proximity factor in the network between any two neurons</td>
<td>31.5</td>
</tr>
<tr>
<td>$\overrightarrow{\text{prox}}$</td>
<td>Vector of the sum of all proximity values from each neuron</td>
<td>Variable</td>
</tr>
<tr>
<td>$\text{prob}_\text{prox}$</td>
<td>Probabilities of synapse creation from a single neuron based upon normalization</td>
<td>Variable</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>Delay between two neurons</td>
<td>$[1, d_{\text{max}}]$</td>
</tr>
<tr>
<td>$d_{\text{max}}$</td>
<td>Maximum axonal delay</td>
<td>10ms</td>
</tr>
</tbody>
</table>
To reflect the fact that neurons with similar functional preferences are biased towards connecting to other neurons sharing those functional preferences (Ko et al., 2011), a probabilistic calculation was used to determine whether any two neurons had a synaptic connection based upon proximity. The probability of any two neurons (denoted as i and j) creating a synapse is a function of their normalized Euclidean distance:

\[
\begin{align*}
\text{dist}_{ij} & = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \\
\text{prox}_{ij} & = \frac{1}{\sqrt{\text{dist}_{ij}}} \\
\text{prox} & = \sum \text{prox}_i \\
\text{prob}_{\text{prox}} & = \frac{\text{prox}_{ij}}{\text{prox}}
\end{align*}
\]

Equation 3.2.3
Equation 3.2.4
Equation 3.2.5
Equation 3.2.6

The number of outgoing synapses was capped by a neuron-specific threshold which varied according to the type of sending neuron (excitatory vs inhibitory) and the synaptic target (intracluster or extracluster). “Intracluster” neurons are neurons that share a functional preference (relative to the sending neuron) and “extracluster” are neurons that do not. Synaptic connections favored intracluster vs extracluster targets, but the proportions varied dependent upon sending neuronal type (as shown in Table 3.2.1) (Kisvárday et al., 1997).

After it was determined which pairs of neurons, \(ij\), were synaptically connected, each synapse was associated with an axonal conduction delay (\(d_{ij}\)). The axonal conduction delays were defined to be integers that were linearly proportional to the proximity factor (\(\text{prox}_{ij}\)) between each pair of neurons, according to the following formula:

\[
d_{ij} = \frac{\text{prox}_{ij}}{\text{prox}_{\text{max}}} \times d_{\text{min}}
\]

Equation 3.2.7

where \(\text{prox}_{ij}\) is the proximity factor based upon Euclidean distance between the neurons, \(\text{prox}_{\text{max}}\) is the largest proximity factor between any pair of neurons, and \(d_{\text{max}}\) is a scaling factor in milliseconds which specifies the maximal delay. An example conduction delay matrix is shown in Figure 3.2.2, below.
Synaptic weights \( (w_{i,j}) \), which represent the strength of the connection from neuron \( i \) to \( j \), were initialized with values dependent upon the types (Excitatory/Inhibitory; E/I) of sending and receiving neurons. Thus, fixed synaptic weights were assigned for E->E, E->I, I->E, I->I connections (Table 3.2.2). All synaptic weights were restricted to the range \([-1, 1]\) at all times. An example of a synaptic connectivity matrix for a network at initialization is depicted in Figure 3.2.3. The corresponding distribution of weights across all synapses is shown in Figure 3.2.4.

**Table 3.2.2: Synaptic Weights at Initialization**

<table>
<thead>
<tr>
<th>Type of Synaptic Connection</th>
<th>Synaptic Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory to Excitatory (E-E) Synapse</td>
<td>0.15</td>
</tr>
<tr>
<td>Excitatory to Inhibitory (E-I) Synapse</td>
<td>0.25</td>
</tr>
<tr>
<td>Inhibitory to Excitatory (I-E) Synapse</td>
<td>-0.4</td>
</tr>
<tr>
<td>Inhibitory to Inhibitory (I-I) Synapse</td>
<td>-0.4</td>
</tr>
</tbody>
</table>
Figure 3.2.3: Synaptic Connections and Weights after Initialization

Figure 3.2.4: Synaptic Weights Distribution after Initialization
3.3 Izhikevich Neurons and Neuronal Dynamics

The initialization of the neural network (Figure 3.2.4) establishes a spatially embedded and interconnected system of neurons of varying input selectivity. We now turn to consider the dynamics of each neuron as a function of its synaptic input. As noted in Section 3.2, there are 2000 Izhikevich neurons used in our cortical circuit. Izhikevich neurons are a modified form of quadratic integrate-and-fire neuron whose state evolves according to the following two-dimensional differential equations (Izhikevich and others, 2003):

\[ v' = 0.04v^2 + 5v + 140 - u + l \quad \text{Equation 3.3.1} \]
\[ u' = a(bv - u) \quad \text{Equation 3.3.2} \]
\[ \text{If } v \geq 30 \text{mV; then } \begin{cases} v & \leftarrow c \\ u & \leftarrow u + mrm \end{cases} \quad \text{Equation 3.3.3} \]

where \( v \) is the membrane potential (microvolts), \( u \) tracks membrane recovery, \( I \) is the current received by a neuron (microamperes/cm\(^2\)), \( a \) is the gain control determining membrane recovery rate of change, \( b \) is a modifier of the membrane potential in calculating the rate of change in membrane recovery, \( c \) is the membrane potential that a neuron returns to once it has fired an action potential (microvolts), \( mrm \) is the amount of change in the membrane recovery variable after an action potential (Table 3.3.1).

**Table 3.3.1: Parameters for Izhikevich Dynamics**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variable Description</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v )</td>
<td>Membrane Potential</td>
<td>Variable</td>
<td>mV</td>
</tr>
<tr>
<td>( u )</td>
<td>Membrane Recovery</td>
<td>Variable</td>
<td>-</td>
</tr>
<tr>
<td>( l )</td>
<td>Current</td>
<td>Variable</td>
<td>( \mu A ) ( / ) ( cm^2 )</td>
</tr>
<tr>
<td>( a )</td>
<td>Gain Control on Membrane Recovery</td>
<td>(excitatory) 0.02</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(inhibitory) 0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b )</td>
<td>Membrane Potential Modifier for Calculation of Membrane Recovery</td>
<td>0.21</td>
<td>-</td>
</tr>
<tr>
<td>( c )</td>
<td>Membrane Potential Reset Value</td>
<td>-65</td>
<td>mV</td>
</tr>
<tr>
<td>( mrm )</td>
<td>Post-Action Potential Membrane Recovery Modifier (( d ) in Izhikevich equations)</td>
<td>(excitatory) 4 (inhibitory) 2</td>
<td>-</td>
</tr>
</tbody>
</table>
When a neuron’s membrane potential rises above 30mV, that neuron generates an action potential, and its value is reset to -65mV. When a neuron spikes (i.e. generates an action potential), it passes current to its synaptically-connected recipient neurons (Figure 3.2.3). If a presynaptic neuron, i, spikes at time $t_i$, then the current delivered to the postsynaptic neuron, j, is updated as follows:

$$I_j(t_i+d_{i,j}) = I_j(t_i+d_{i,j}) + w_{i,j} \ast \alpha_w$$

Equation 3.3.4

where $I_j$ is the current received by neuron $j$, $d_{i,j}$ is the axonal conduction delay between neurons $i$ and $j$, $w_{i,j}$ is the synaptic weight between those neurons, and $\alpha_w$ is a scaling factor for the synaptic weight to add current.

### 3.4 Background Activity

The background Poisson firing input to the model and the “visual” stimulus exposures are considered to be “exogenous stimulation” in our paradigm, while the action potentials that result indirectly from prior neuronal activity in the model (via Equation 3.3.4) are considered “endogenous stimulation”. Our model does not include any neurons that fire in the absence of exogenous stimulation. Thus, we require exogenous stimulation to our visual cortical circuit in order to measure neural responses in our experiments, and to differentiate neural responses across experimental conditions.

Given the spatial scale of our model, we do not explicitly model other cortical and subcortical sources of input to our patch of visual cortex. Instead we use exogenous “noise” from external regions to both support the maintenance of inhibitory and excitatory activation (Bulsara and Schmera, 1993; Moss, 2004) and ensure the model is not quiescent in the absence of stimulus (Lestienne, 2001; Kenet et al., 2003). Thus we introduce Poisson-based “background firing”, as a surrogate for these cumulative sources of input. Specifically, we provided Poisson current at rate, $R_{poi}$, to each neuron. Whenever a current injection event was generated by the

<table>
<thead>
<tr>
<th>$t_i$</th>
<th>Time Variable</th>
<th>Synaptic Weight</th>
<th>Variable</th>
<th>ms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Synaptic Weight</td>
<td>[-1, 1]</td>
<td>-</td>
</tr>
<tr>
<td>$w_{i,j}$</td>
<td>Synaptic weight modifier (for current)</td>
<td>7.8</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Poisson process for a neuron, input current to that neuron at that millisecond was increased by a constant, $\alpha_b$:

$$I_j(t) = I_j(t) + \alpha_b$$ \hspace{1cm} \textbf{Equation 3.4.1}

This non-specific stimulation produced an average firing rate of $\sim 1.25$Hz per neuron in the absence of any other stimulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{poi}$</td>
<td>Poisson Firing Rate</td>
<td>2Hz</td>
</tr>
<tr>
<td>$\alpha_b$</td>
<td>Current Addition from Background Firing</td>
<td>12mA</td>
</tr>
</tbody>
</table>

### 3.5 Synaptic Plasticity

#### 3.5.1 Spike-Timing Dependent Plasticity (STDP)

The neurons in the network were connected by plastic synapses, which change their weight as a function of spiking activity (Abbott and Nelson, 2000). The specific form of synaptic plasticity that we chose to utilize is known as Spike-Timing Dependent Plasticity (STDP) (Bi and Poo, 1998; Feldman, 2012; Izhikevich, 2006; Song et al., 2000). STDP is a mechanism by which synapses are potentiated or depotentiated dependent upon the relative timing of action potentials between two connected neurons. Figure 3.5.1.1 illustrates the valence of synaptic potentiation as a function of the relative spike timing between when neurons $i$ and $j$ fire action potentials. The synaptic weight change ($\Delta w$, y-axis) is shown as a function of the relative timing difference in action potentials ($\Delta t$, x-axis), following the model described in Equation 3.5.1.1.
Figure 3.5.1.1: STDP Explanation  

If neuron i fires prior to neuron j (i+, shown with orange line segment), then the synapse from i to j is potentiated (Equations 3.5.1.1-3.5.3; Table 3.5.1.1). Conversely, if i fires after j (i-, shown with blue line segment), then the synapse from j to i is reduced (Hosaka et al., 2008).

\[
\Delta t = t_j + d_{i,j} - t_i
\]  

Equation 3.5.1.1

\[
f_+ (w) = (1 - w)^{\mu} \quad \text{and} \quad f_- (w) = \alpha w^{\mu}
\]  

Equation 3.5.1.2

\[
\Delta w(\Delta t) = \begin{cases} 
-\lambda f_+ (w) * \exp \left(-\frac{\Delta t}{\tau_p}\right) & (\Delta t \geq 0) \\
\lambda f_- (w) * \exp \left(-\frac{\Delta t}{\tau_d}\right) & (\Delta t < 0)
\end{cases}
\]

Equation 3.5.1.3

where \( f \) is the magnitude of synaptic weight change, \( \mu \) is a modifier controlling whether additive or multiplicative STDP is used, \( \alpha \) is the gain control on the magnitude of synaptic strength decrease, \( \lambda \) is the gain control on \( f \), and \( \tau_p \) & \( \tau_d \) are the maximum interval (milliseconds) of inspection for synaptic weight updates (for potentiation and depotentation, respectively). All variables described are constant, except \( f \).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>Time of Action Potential</td>
<td>Variable</td>
<td>ms</td>
</tr>
</tbody>
</table>
Gain control on magnitude of synaptic weight change

\[ f \]

Modifier determining STDP type (multiplicative or additive)

\[ \mu \]

Gain control on influence of current synaptic weight on synaptic weight decrease

\[ \alpha \]

Gain control on \( f \)

\[ \lambda \]

Interval of inspection for potentiation

\[ \tau_p \]

Interval of inspection for depotentiation

\[ \tau_d \]

Plasticity was restricted to excitatory to excitatory (E-E) synapses. Therefore, all other synapses (E-I, I-I, I-E) remained constant (as defined in Table 3.2.2). Although inhibitory plasticity exists in visual cortex (Maffei et al., 2006; Gandhi et al., 2008; Vogels et al., 2013), we have limited plasticity to E-E synapses for the sake of parsimony, and because the kinds of plasticity operating at E-I, I-E, and I-I synapses is less well understood empirically. Exclusion of inhibitory plasticity has been shown in previous work to not significantly affect empirical modeling (Bush and Sejnowski, 1996; Izhikevich, 2006; Hosaka et al., 2008).

### 3.5.2 Heterosynaptic plasticity

STDP is the most frequent method of synaptic weight change in the model, but it is not the only mechanism of synaptic plasticity in our model. In order to prevent the known issue of runaway growth of synaptic strength arising from STDP, we incorporate heterosynaptic plasticity, a mechanism that has been shown empirically to support network homeostasis (Chistiakova et al., 2014; Lee et al., 2012) and which has attractive computational properties in the setting of sequential learning (Fiete et al., 2010).

Heterosynaptic plasticity occurs when, for a single neuron, changes at one of its synapses lead directly to changes at its other synapses. For example, if strengthening and maintaining synapses in the dendritic arbor of a neuron draws upon a limited metabolic resource, then long-term potentiation (LTP) or long-term depression (LTD) at one synapse in the dendritic arbor, may lead to LTD or LTP, respectively, at other synapses. In this way, the total incoming synaptic weight may be conserved even as the relative contribution of synapses changes (Royer and Paré, 2003). Figure 2c of Chistiakova et al (2014) illustrates an induction method for Heterosynaptic LTD from an in vitro experiment.

The “Mexican Hat” (two-colored circles) indicates the regions of potentiation (in red) and depotentiation (in green), while the red lightning indicates the induction zone. This empirical
finding is interesting because of the lack of direct stimulation to regions undergoing plasticity, suggesting that STDP is insufficient as a singular mechanism to model synaptic plasticity. We model this form of plasticity by taking the total incoming synaptic strength from all excitatory neurons to each excitatory neuron and compare that summated value against a threshold, \( W_{\text{thres}} \). If the threshold is exceeded, then all synaptic weights for that excitatory neuron are modified using divisive normalization:

\[
W_{\text{col}} = \sum w_{ij} \quad \text{(where i & j are excitatory neurons)} \quad \text{Equation 3.5.2.1}
\]

\[
\text{If } W_{\text{col}} > W_{\text{thres}}, \quad w_{ij} = w_{ij} / W_{\text{thres}} \quad \text{Equation 3.5.2.2}
\]

where \( W_{\text{col}} \) is the sum of all synaptic weights coming into a neuron.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_{\text{col}} )</td>
<td>Summation of excitatory to excitatory synaptic weights to a specific excitatory neuron</td>
<td>Variable</td>
</tr>
<tr>
<td>( W_{\text{thres}} )</td>
<td>Maximum value in vector of all ( W_{\text{col}} ) values</td>
<td>(~33)</td>
</tr>
</tbody>
</table>

Thus the metabolic cost of maintaining a dendritic tree is held relatively stable given potentiation pressure resulting from STDP.

### 3.6 Stabilization Period

The preceding subsections define the selected aspects of V1 connectivity that are emulated by the model, as well as how synapses are plastic. Before we can begin to conduct computational experiments using this system, however, it is important to allow the model to establish synaptic variability from this initialized state until it reaches a dynamical equilibrium. This will provide a common basis from which we can compare subsequent experimentally-induced changes in the network structure and dynamics.

To allow the model to reach an equilibrium state before experimental training regimes were applied, we simulated 1.8 million consecutive time steps (30 virtual minutes) in a regime in which neurons received only spontaneous Poisson background activity as exogenous input. The synaptic weights in the intracluster regions both potentiate and depotentiate (as shown in Figure 3.6.1, c.f. 3.2.3). These changes in E-E synaptic weights are evident in the distribution changes shown in Figure 3.6.2 (c.f. Figure 3.2.4):
Figure 3.6.1: Synaptic Weight Matrix after Stabilization

Figure 3.6.2: Synaptic Weight Distribution after Stabilization
This distribution of synaptic weights was used as the starting point for each experiment.

4 Computational Experiments

Following the creation of the baseline network (Figure 3.6.1), we studied the learning properties of the network, and the persistence of information in process memory, in a series of three computational experiments. These experiments were inspired by the Gavornik & Bear findings discussed in Section II and they test the three hypotheses posited in the Introduction. We hypothesize that, after repeated exposure to a stimulus (or sequence) during training,

1. More neurons will become responsive to the training-exposed stimulus (or sequence) compared to baseline;
2. The local cortical circuit will be encode information about stimulus (or sequence element) presentation farther into the post-stimulus period;
3. A sequence of stimuli (ABCD) will induce a larger neuronal response than the reverse sequence (DCBA) only when networks trained on ABCD. The advantage of ABCD over DCBA will not be present for networks trained on a non-repeating random sequence.

Experiment 1 investigated whether repeated exposure of the network to a specific stimulus (“A”) elicited an enhanced response to that particular stimulus, as was found in GB14. Experiment 2 investigated whether repeated exposure of the local circuit to a particular stimulus would lengthen the process memory for that stimulus within the local circuit. Finally, Experiment 3 investigated whether repeated exposure to a particular sequence of stimuli (ABCD) would lead to an enhanced sequence-specific response. The visual stimuli presented to in training are referred to by letter, consistent with the cluster labels in Figure 3.2.1.

In all experiments, we compared the circuit changes elicited by three different training regimes (“experimental conditions”):

1) Baseline Network: No stimulus or sequence presented, no plasticity
2) Spontaneous Activity: No stimulus or sequence presented, with plasticity enabled
3) Repeated Exposure: A (Exp. 1&2) or ABCD (Exp. 3) presented, with plasticity enabled
When plasticity is enabled, the excitatory-excitatory (E-E) synapses were mutable according to the rules described in Section 3. Stimulus presentation indicates excitation of a subset of the network (in addition to the Poisson background activity).

4.1 Experiment 1: Effects of Repeating a Single Stimulus

4.1.1 Overview

In primary visual cortex, repeated exposure to a stimulus results in an increased neural response (Frenkel et al., 2006). Frenkel et al called this experience-dependent plasticity stimulus-specific response potentiation (SSRP). SSRP was induced in head-restrained, awake mice by presenting a visual stimulus 100-400 times to left and right eyes in random alternation. We employ an analogous design in our computational experiments in order to investigate the development of response potentiation in our model.

4.1.2 Hypothesis

The expectation of Experiment 1 was that the SSRP effect (increases in response to a repeatedly presented stimulus) would be obtained only in the training regime which exposed the circuit to the repeated stimulus during training and which allowed synaptic plasticity during training. Systems without plasticity cannot learn and thus cannot be affected by the presentation of stimuli. We hypothesized that the response potentiation would arise from synaptic potentiation among neurons initially responsive to stimulus A (i.e. in cluster A).

4.1.3 Methods

4.1.3.1 Training Overview

Following the stabilization of the network described in Section 3.7, we exposed the network to three distinct training regimes and compared the anatomical and functional changes elicited by training. The training regimes were:

1) Baseline (No change from Stabilization period) – No Plasticity
2) No Stimulus Presentation – Plasticity Enabled
3) Repeated Presentation of Stimulus A – Plasticity Enabled

The model was simulated with a 1 ms step size for 200,000 time steps (200 simulated seconds).
4.1.3.2 Training with Exposure to Visual Stimulus

To emulate the geniculocortical volley that arrives in V1 following the visual presentation of a Gabor patch (Stimulus A) in the receptive field of those neurons, we activate an ensemble of 13 excitatory neurons within the cluster of neurons (Cluster A) that is selective for that stimulus orientation. The same set of neurons within cluster A was stimulated at each visual stimulus exposure. The incoming volley of spikes (“stimulation”) onset occurred exactly 200ms into each second of model simulation (i.e. at 0.2 seconds, 1.2 seconds, 2.2 seconds, etc…), for the duration of the visual training period. Figure 4.1.3.2.1 shows a spike raster for a representative 1-second window within the training period.

4.1.3.3 Training without Exposure to Visual Stimulus

In the “Unstimulated” regime, all neuronal firing was the direct or indirect result of Poisson background drive; there was no systematic perturbation of the spontaneous dynamics. Spike rasters of the neuronal dynamics at the time of stimulation in the stimulated regime (Figure 4.1.3.2.1) or spontaneous regime (Figure 4.1.3.3.1) are qualitatively similar. However, quantitative comparison of the firing rates (shown in insets of both figures) reveals a clear firing rate increase in response to the stimulus, selective to the stimulus-exposed training regime.
Figure 4.1.3.2.1: Example Spike Raster and Firing Rate for Visually Stimulated Training Regime

Figure 4.1.3.3.1: Example Spike Raster and Firing Rate for an Unstimulated Training Regime
4.1.3.4 Anatomical Changes Resulting from Training

Completion of training across the three regimes resulted in synaptic weight distributions similar to baseline (Figure 4.1.3.4.1A), but with an increase in strong synapses ($w > 0.4$) in the stimulus-exposed training regime (Figure 4.1.3.4.1B). This increase in synaptic strength in the stimulus-exposed regime is evident in the stimulated cluster, but not in unstimulated clusters, suggesting that learning of stimulus A has occurred, and that the learning is selective to the repeated stimulus (Figure 4.1.3.4.2).

Figure 4.1.3.4.1: Distribution of Excitatory-Excitatory (E-E) Synaptic Weights

(A) All E-E Synapses
(B) E-E Synapses with strength $> 0.4$
Repeated exposure to stimulus A led to a significant increase in the strength of excitatory synapses between neurons in cluster A ($t(70544)=-6.48$, $p<0.001$, two sample t-test comparing A-exposed and baseline regimes). Conversely, exposure to stimulus A decreased the average intracluster synaptic weights in the unstimulated clusters ($t(211520)=-2.09$, $p<0.05$, two sample t-test comparing A-exposed and baseline regimes). There were no significant changes between the baseline and unstimulated training regimes and no difference in the average synaptic weight for any training regime (Table 4.1.3.4.1).

Table 4.1.3.4.1: Mean Excitatory-Excitatory (E-E) Synaptic Weights after Training

<table>
<thead>
<tr>
<th>Cluster Type</th>
<th>Baseline</th>
<th>Training: Plasticity w/o Stimulus A</th>
<th>Training: Plasticity w/ Stimulus A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster A (Stimulated)</td>
<td>0.2081</td>
<td>0.208 n.s.</td>
<td>0.2115***</td>
</tr>
<tr>
<td>Clusters B-D (Unstimulated)</td>
<td>0.2083</td>
<td>0.2081 n.s.</td>
<td>0.2078*</td>
</tr>
<tr>
<td>All E-E</td>
<td>0.1998</td>
<td>0.1998 n.s.</td>
<td>0.1998 n.s.</td>
</tr>
</tbody>
</table>
4.1.3.5 Functional Changes Resulting from Training

In Frenkel et al (2006), stimulus-specific potentiation testing in mice was operationalized by comparing event-averaged potentials (recorded via microelectrodes) between pre-training and post-training. The equivalent metric in our model is to compare anatomical and neuronal spiking rate changes between pre-training and post-training.

Does repeated exposure to a stimulus, in the context of our synaptic plasticity model, produce an increased population response? If so, is such an increased population response specific only to the stimulus that it was repeatedly exposed to in training? To investigate these questions, we took the networks generated by each of the three training regimes, and tested their functional responses to three different input conditions: No Visual Stimulus; Single Presentation of Stimulus A; Single Presentation of Stimulus B. For each test case of visual exposure, we simulated 1500 independent trials initialized from different states of background activity. Each trial lasted for one second of model time; stimuli were delivered at 200 ms as in the training regimes; STDP was disabled for this post-training functional assessment. The combinations of visual exposure training regimes and test cases are show in Table 4.1.3.5.1.

<table>
<thead>
<tr>
<th>Training Regime</th>
<th>Test Case: Without Visual Stimulation</th>
<th>Test Case: Tested with Visual Stimulation A</th>
<th>Test Case: Tested with Visual Stimulation B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Train: Baseline Test: No Stim</td>
<td>Train: Baseline Test: Stim A</td>
<td>Train: Baseline Test: Stim B</td>
</tr>
<tr>
<td>No Visual Stimulation; Plasticity ON</td>
<td>Train: No Stim; Plasticity Test: No Stim</td>
<td>Train: No Stim; Plasticity Test: Stim A</td>
<td>Train: No Stim; Plasticity Test: Stim B</td>
</tr>
<tr>
<td>Visual Stimulation A; Plasticity ON</td>
<td>Train: Stim A; Plasticity Test: No Stim</td>
<td>Train: Stim A; Plasticity Test: Stim A</td>
<td>Train: Stim A; Plasticity Test: Stim B</td>
</tr>
</tbody>
</table>

4.1.4 Experiment 1 Results

The visually evoked responses from testing complemented the anatomical changes highlighted in section 4.1.3.4. Measuring functional responses to Stimulus A, we observed an increase in population firing rate and firing duration for the networks trained with repeated exposure to Stimulus A compared to networks trained without any stimulus (Figure 4.1.4.1).
The population response to the stimulus exhibits three peaks, because the initial stimulus input arrives simultaneously at all input neurons, and these neurons share similar axonal delays to their downstream targets within cluster A. In a more realistic model, jitter in stimulus input timing and conduction delays would avoid the overly synchronous three-phase response.

Figure 4.1.4.1: Average firing rates for 200ms following presentation of stimulus A

The mean firing rate elicited by Stimulus A was significantly larger for the stimulus-exposed training regime than the unexposed and baseline training regimes ($t(2998)=143.36$, $p<<0.001$, two sample t-test of firing rates for stimulus-exposed and spontaneous training regimes and $t(2998)=143.47$, $p<<0.001$, two sample t-test of stimulus-exposed and baseline firing rates, respectively). The baseline and spontaneous activity-trained networks produced functional responses that were not significantly different from one another (Table 4.1.4.1).

Table 4.1.4.1: Average Firing Rates for Excitatory Neurons in the 40ms Post-Stimulus across all neurons (Hz)

<table>
<thead>
<tr>
<th>Training Regime</th>
<th>Test Case: Without Visual Stimulation</th>
<th>Test Case: Tested with Visual Stimulation A</th>
<th>Test Case: Tested with Visual Stimulation B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.3777</td>
<td>1.3978</td>
<td>1.3968</td>
</tr>
<tr>
<td>No Visual Stimulation;</td>
<td>1.3764 $^{n.s.}$</td>
<td>1.4006 $^{n.s.}$</td>
<td>1.399 $^{n.s.}$</td>
</tr>
</tbody>
</table>
There were no significant differences in firing rates between the training regimes when population dynamics were measured in the absence of any stimulus (Figure 4.1.4.2) or when responses were measured to a novel (un-trained) stimulus such as Stimulus B (Figure 4.1.4.3).

**Figure 4.1.4.2**: Average firing rates for 200ms following omission of stimulus at stimulation onset
As hypothesized, the only training regime that led to a significant change in evoked response was the regime trained with stimulation in cluster A while synaptic plasticity was enabled (Figure 4.1.4.1, green line). The undifferentiated responses across training regimes for spontaneous dynamics (Figures 4.1.4.2) and for a novel stimulus (Figure 4.1.4.3) reflect the stimulus-specificity of the learning.

Somewhat surprisingly, training also increased the responsiveness of clusters B, C and D (that were not receiving any direct stimulation) upon presentation of A ($F(2,4497)=50.79$, $p<<0.001$, one-way ANOVA comparison across training regimes, Tables 4.1.4.2 & 4.1.4.3).

Table 4.1.4.2: Mean Firing Rates (Hz) for Training Regimes Presented Stimulus A (40ms Post-presentation)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Training: Plasticity w/o Stimulus A</th>
<th>Training: Plasticity w/ Stimulus A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster A (Stimulated)</td>
<td>1.3951</td>
<td>1.4122</td>
<td>7.6948***</td>
</tr>
<tr>
<td>Cluster B-D (Unstimulated)</td>
<td>1.3908</td>
<td>1.3884</td>
<td>1.4529***</td>
</tr>
</tbody>
</table>

*** $p<<0.001$

Table 4.1.4.3: Mean Firing Rates (Hz) for Training Regimes Presented Stimulus B (40ms Post-presentation)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Training: Plasticity w/o Stimulus A</th>
<th>Training: Plasticity w/ Stimulus A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster A (Stimulated)</td>
<td>1.3913</td>
<td>1.3937</td>
<td>1.3795</td>
</tr>
</tbody>
</table>
The increase in firing rates in the unstimulated clusters (B, C, D) was unexpected because the synaptic weights within those clusters decreased following training on A (Table 4.1.3.4.1). This suggests that the increase in firing was driven by connections from the stimulated cluster (A) to the unstimulated clusters. Comparing the synaptic strength between clusters from the stimulated cluster (A->B, A->C) versus to the stimulated cluster (B->A, C->A) confirms this conjecture (Figure 4.1.4.4). As a benchmark, note the relative stability of the C->D & D->C synaptic weights, as well as the synaptic weights in the stimulus omitted training regime.

Figure 4.1.4.4: Synaptic Weight Changes when for each training regime

The response to Stimulus A was dramatically increased following training on that stimulus, compared to baseline and to training without any stimulus (Figure 4.1.4.2). Stimulation using either A or B produces the expected immediate increase in firing rate at stimulation onset (200ms), and this is matched across conditions, but the duration of the response to A is extended only in the cases where testing and training used stimulus A.
4.1.5 Discussion of Experiment 1

An increased response to a repeatedly exposed stimulus could benefit an organism in multiple ways. For example, if a discrimination task is required between two stimuli, the magnitude of the cortical circuit output could indicate which stimulus was seen. Relatedly, the response magnitude increase could indicate a larger neuronal ensemble responsive to the stimulus, and thus, increased reliability in transmission of the stimulus.

The increased firing in unstimulated clusters is likely the result of the fact that no competitive stimuli were provided (i.e. clusters B, C, or D were never exposed to their preferred stimuli). In real-life, an organism would not be limited to only one type of visual exposure. However, our model predicts that if that were to occur, we would see increased responses to the stimulus in previously non-response areas (Sadato et al., 1996). In Experiment 3, the training regime is more realistic as the frequency of all stimuli is balanced, and only the sequence order is manipulated. Therefore, we anticipate that the synaptic weight changes will be specific to associations between sequence elements and not unstimulated associations.

Our stimulus-specific response potentiation suggests that experience-dependent plasticity results in a stimulus response that is not only increased in magnitude, but also in duration. Does repeated stimulus exposure also lengthen the process memory of that stimulus within the neuronal population? If so, will the memory be dependent upon the firing rate fluctuations observed in Figure 4.1.4.1?

4.2 Experiment 2: Lengthening of Process Memory

4.2.1 Overview

Having shown that the model reflects an increased neuronal response to a repeatedly presented stimulus, we asked if this repeated exposure to a stimulus lengthens the process memory of the circuit for that stimulus.

We defined the process memory of a circuit as “the amount of information about the past state of the circuit that can be inferred from the present state of the circuit” (i.e. information persistence). Thus, to operationally measure process memory, we used a classifier to decode
population states in the past given information about population states in the present. By using a classifier to decode the network history from neuronal ensembles, we can assess how process memory varies for the presented stimulus over different time intervals post-stimulation and following different training regimes. This approach allows us to test different network states over many 1000ms trials about whether there was a visual stimulus presented or omitted at time t=200ms. For each time interval (e.g. t= 280-305ms), we attempt to decode whether a stimulus was presented on that trial, based only on the spiking activity within that limited interval.

4.2.2 Hypothesis

Given that repeated stimulus exposure resulted in a relative increase in the duration of firing rates and affected unstimulated clusters, it seems probable that process memory for a stimulus is lengthened in networks which were trained with repeated exposure to that stimulus.

Intuitively, decoding from network states before 0.2s should be at chance (~50%), as the network state cannot contain any information about the not-yet-presented stimulus. At the time of stimulus presentation (200ms), we expect there to be a large population response to the visual stimulus, regardless of the training regime. Thus, around the time of stimulus presentation (200ms), decoding accuracy should be near ceiling (~100%) regardless of the training regime. However, in time intervals following the stimulus presentation on each trial, the ability to decode whether a stimulus was presented will depend on the process memory in the circuit for the presentation of the stimulus, and we expect that this process memory will vary according to the training regime.

Specifically, we hypothesize that the process memory for the stimulus will be longest in the conditions where the network was repeatedly exposed to the stimulus and synaptic plasticity was active.

The average synaptic weight increased in the training regime exposed to stimulus A in Experiment 1. This increase suggests that there is likely to be more “reverberation” in the stimulated cluster given the intracluster potentiated synapses. Therefore, we believe that neuronal ensembles have preferentially developed within the stimulated cluster compared to unstimulated clusters and these ensembles give rise to process memory.
4.2.3 Methods

4.2.3.1 Training and Testing

Training and Testing followed the same protocol used in Experiment 1. Therefore, we were able to use the previously generated test data (Methods from section 4.1.3) as the basis for the decoding analysis in Experiment 2.

4.2.3.2 Classification Procedure

To test whether the local circuit memory could be lengthened by repeating the same stimulus, the prior network state was decoded from the present network state. To perform this decoding, we used logistic regression with a regularization technique called “Lasso” regularization to perform binary classification between two categories: (i) stimulus was presented on this trial; (ii) stimulus was not presented on this trial.

Logistic Regression is a machine learning tool which assigns a categorical output (i.e. a classification decision) from an input vector of potentially predictive variables, using a combination of a general linear model and a linking function:

\[
F(x) = \frac{1}{1-\exp[-(\beta_0 + \beta_1 x)]} \quad \text{Equation 4.2.3.2.1}
\]

\[
g(F(x)) = \ln \left( \frac{\mu(x)}{1-F(x)} \right) \quad \text{Equation 4.2.3.2.2}
\]

where \(x\) is the vector of predictor variables, \(\beta_1\) is a vector of coefficients assigned to those predictor variable, \(\beta_0\) is the intercept associated with the linear regression equation, and \(F(x)\) is the probability of the predicted non-zero category. The intent is to assign coefficients (\(\beta\)) to the potentially predictive variables which then classifies the data into categories of interest.

Since we want to predict whether “Stimulus A” or “No Stimulus” was presented during testing we use the firing rate of each neuron, \(R(t)\), from each test trial (where the tested networks share the same underlying training) as our predictor variable. Trials from each of the two test cases are randomly shuffled and combined (in equal proportions) to produce “Classifier Training Data” (75% of the total number of trials) and “Classifier Test Data” (the remaining 25%).
firing rate, \( R(t) \), is calculated for 23 different time bins of length 25ms per neuron, \( t_{bin} \), (where all bins starting \([200\leq t<500]\) have a partial overlap of 5ms; Table 4.2.3.2.1):

\[
R(t) = N_{ap}(t) \times \frac{1000}{t_{bin}}
\]

Equation 4.2.3.2.3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R(t) )</td>
<td>Neuronal firing rate within ( t_{bin} ) at time ( t )</td>
<td>Variable</td>
<td>-</td>
</tr>
<tr>
<td>( N_{ap}(t) )</td>
<td>Number of neuronal action potentials within ( t_{bin} )</td>
<td>Variable</td>
<td>-</td>
</tr>
<tr>
<td>( t_{bin} )</td>
<td>Duration of time bin</td>
<td>25</td>
<td>ms</td>
</tr>
</tbody>
</table>

where the number of action potentials for a neuron within each of the 23 time bins is \( N_{ap}(t) \).

Once the Classifier Training Data is created, then Lasso Regularization is used with “Cross Validation” on subsets of the Classifier Training Data (James et al., 2013). Lasso regularization biases the coefficients of predictor variables (\( \beta \)) towards zero to reduce the number of predictor variables ultimately used and, in conjunction with Cross Validation, we avoid “overfitting” problems associated with poor predictive ability. The vector of coefficients (\( \beta \)) reflects the relative importance of each neuron’s firing rate \( (R(t)) \) in determining whether a stimulus was presented or not within each time bin of the trial. The same coefficients are then applied to the “Classifier Test Data” (previously unseen for the coefficient calculations) to attempt to predict the proper category. Finally, for each time bin, the “Classifier Accuracy” was defined as the percentage of trials for which the classifier-predicted category matched the veridical category.

### 4.2.3.3 Measuring Anatomical Changes

Assessing whether neuronal ensembles are forming is an open problem in neuroscience (Carrillo-Reid et al., 2015). However, there are existing tools, such as the Brain Connectivity Toolbox (Rubinov and Sporns, 2010) that enable computational characterization of network properties indicative of network ensembles. We characterized changes in network structure using two network metrics: firstly, a motif analysis (i.e. counting the incidence of specific subgraphs within the anatomical network) and secondly, measurement of the clustering coefficient of neurons within the anatomical network.
4.2.3.3.1 Motif Analysis

Motifs indicate the existence of specific network connectivity patterns between three neurons. A graphical presentation of these connections are shown below (Figure 4.2.3.3.1.1 adapted from Sporns et al., 2007). The STDP rule generally encourages unidirectional synaptic strength increases with decreases to the synaptic strength in the reverse direction, therefore we expect that motifs without bidirectional connectivity like M{1-3,5,7} will be more plentiful than a motif like M6.

![Figure 4.2.3.3.1.1: Possible 3-motif connections](image)

4.2.3.3.2 Clustering Coefficients

Clustering coefficients provide a rating of the degree to which neighbors of a specific neuron connect to one another. Finding regions of high clustering coefficients suggest groups of cooperative neurons, whereas low cluster coefficients indicate potentially more chain-like (non-recurrent) behavior.

\[
\varsigma = \frac{k_n}{k_{n_{\text{max}}}}
\]

*Equation 4.2.3.3.2.1*

where \(\varsigma\) is the clustering coefficient for a neuron \(i\), \(k_n\) is the number of neighbors of \(i\) with connecting edges, and \(k_{n_{\text{max}}}\) is the maximum potential number of edges that could exist between neighbors of \(i\).
4.2.4 Results

4.2.4.1 Duration of Process Memory

As hypothesized, there was an evident increase in the process memory associated with the trained and tested pairing on Stimulus A. Classification accuracy using the firing rate from each neuron in the network is shown in Figure 4.2.4.1.1 (top).

![Classification Accuracy Comparison (Using all Neurons)](image)

![Average Firing Rate (Using all Neurons)](image)

Figure 4.2.4.1.1: Classification Accuracy and Associated Firing Rates
(top) Classification Accuracy differences between training regimes tested on Stimulus A
(bottom) Average Firing rates for training regimes tested on Stimulus A

The point at which classification accuracy for the stimulus-exposed training regime returns to chance levels aligns well with the time when the average firing rate for that training regime returns to the background firing rate (Figure 4.1.4.1.1 bottom). The classifier accuracy appeared to vary with deviation from the average background firing, remaining well above chance for the stimulus-exposed training regime compared to the other training regimes.

Since the firing rate seemed to track predictive ability, we wondered how much of the predictive power was specific to firing within the stimulated cluster. Therefore, we investigated
whether using fewer variables in a standard logistic regression (without Lasso regularization) could produce equivalent results. Figure 4.2.4.1.2 shows the classifier accuracy using only the average firing rate from each cluster as the four predictor variables.

![Classifier Accuracy using average cluster firing rates](image)

**Figure 4.2.4.1.2: Classifier Accuracy using average cluster firing rates**

The reduction in classifier accuracy shows that using only the cluster firing rates is an inferior predictor of whether a stimulus was presented or not in testing.

The accuracy of the classifier in Figure 4.2.4.1.1 matches the hypothesized output in all three ways: Classification accuracy pre-stimulus is at chance, classification accuracy at time of stimulus is at ceiling, and the model trained with plasticity and having seen the stimulus previously shows both an increased duration of memory (as shown by when the dashed green line reaches chance levels) and increased accuracy (shown as the higher measurement on the y-axis between the dashed green line and other lines). The superior predictive capabilities shown suggest that process memory in the system is more than a simple translation of firing rate in the stimulated cluster.

We conclude from the improved decoding accuracy of a stimulus-exposed circuit that memory in that local cortical circuit is lengthened by a repeated stimulus versus a transient stimulus. Process memory effects persist up to 140ms after stimulation and the memory of a system having seen a stimulus repeatedly is elongated compared to a plastic system that was not exposed to the stimulus.
4.2.4.2 Anatomical Changes

It is clear from Experiment 1 that the synaptic strength increased within the stimulated cluster and from the stimulated cluster to other clusters. Given that the STDP rule appeared to manifest on the macroscale (i.e. between clusters) we would expect the same on a microscale (i.e. a rise in motifs eschewing bidirectional connections). Additionally, if intracluster synapses are strengthening, it could suggest that clustering coefficients for the majority of neurons are increasing. This would reflect the bias in forming strong intracluster synapses versus strong synapses between clusters.

We found that both the number of motifs (especially $M\{1, 3, 5, 7\}$ as expected) and the average clustering coefficient in cluster A increased (Baseline: $\mu=0.01$, $s.d.=0.044$; Post-Training: $\mu=0.015$, $s.d.=0.029$) for the training regime pre-exposed to stimulus A (Figure 4.2.4.2.1). The average clustering coefficient in cluster A for the training regime not exposed to Stimulus A, however, decreased (Post-Training: $\mu=0.005$, $s.d.=0.021$) and motifs $M\{1, 3, 5, 7\}$ were generally stable. Note that, in all anatomical analyses, the synaptic networks were thresholded to include only synapses with a synaptic weight greater than 0.33 ($w>0.33$).

![Cluster A Motif Frequency](image1)

![Cluster A Clustering Coefficient](image2)

Figure 4.2.4.2.1: Changes in 3-neuron Motifs & Cluster Coefficients for Trained Regimes

The motif analyses suggests that the network trained with stimulus A formed more unidirectional connections, as opposed to bi-directional connections, than the network unexposed to stimulus A. Additionally, the clustering coefficient analyses suggest that exposure to stimulus A during training produced a more strong synapses within the stimulated cluster, compared to
the unstimulated network, which remained relatively stable. The increased clustering coefficient in the stimulus-exposed regime is consistent with the increase in expression of motifs 1-3, 5, 7 because propagating neuronal activity from stimulated neurons could produce strong outgoing synapses which subsequently propagate activity to their neighbors, continuing a chain of strengthening synapses. Although the spontaneous training regime showed a decrease in the average clustering coefficient, the small rise in the number of M4 and M6 motifs was unexpected. Both of these results could have resulted from random variation in intracluster dynamics, a possibility that we will test in future studies.

4.2.5 Discussion of Experiment 2

A circuit that was repeatedly exposed to a particular stimulus will subsequently, when processing that stimulus, implicitly possess information about that stimulus for a longer period of time after stimulus offset than the same circuit unexposed to the stimulus. Thus, repeated exposure to a stimulus appears to lengthen the process memory for that stimulus. The lengthening of process memory for stimulus A could not be accounted for simply by the increased mean firing rates in response to A; at least some of the process memory resides in patterns of neuronal firing.

From a modeling perspective, the formation of directed connections from the stimulated cluster, but not from unstimulated clusters, suggests that our model should demonstrate sequence learning effects when a circuit is exposed to sequences of distinct stimuli. We therefore move on to investigate sequence learning, asking: can associations between elements of a sequence be learned via repeated exposure to that sequence?

4.3 Experiment 3: Sequence Learning

4.3.1 Overview

In Experiments 1 & 2 we have shown that repeated exposure to a single stimulus results in an increased neural response (in both magnitude and duration) and that this increased neural response exists in tandem with an increase in local circuit memory. We now enter the second phase of our proposed two-phase mechanism for sequence learning and hope that these effects translate into the creation of associations between sequence elements.
4.3.2 Hypothesis

Given that repeated exposure to a single stimulus results in an increased memory of that stimulus, we hypothesized that the elongation of memory would permit additional opportunities for associations between stimulus elements to be learned. As such, we hypothesized that repeated exposure to a sequence of stimulus elements would result in stronger associations between stimulus elements only in the direction of learning compared to a random presentation of the same stimulus elements. Testing this hypothesis allows us to use GB14 style sequences of ABCD and DCBA patterns of stimulus elements, where we believe the network trained on ABCD will have an increased response to ABCD compared to an unrepeated sequence (the reverse sequence, DCBA).

4.3.3 Methods

The methodology for Experiment 3 is inspired by the GB14 experimental design shown in Figure 1.2.3. Sequence training will consist of presenting four stimulus elements where: the repeated sequence is always the same (e.g. A-B-C-D, A-B-C-D, A-B-C-D, etc.) and the random sequence changes the order of elements every presentation (e.g. D-C-B-A, A-C-D-B, C-A-D-B, etc.). The same set of neurons is activated by its specific stimulus element across training and testing, regardless of the order of presentation.

4.3.3.1 Training Overview

The training regimes were slightly modified in Experiment 3 from Experiments 1 & 2. Instead of the training regimes (Baseline, No Presentation of Stimulus A, Repeated Presentation) we used:

1) **Baseline (No change from Stabilization period) – No Plasticity**
2) **Random Sequence of Four Elements – Plasticity Allowed**
3) **Repeated Sequence of Four Elements – Plasticity Allowed**

Sequential Stimulation was a repeated pattern of stimuli every second and Random Stimulation was a randomized pattern of stimuli while preserving the relative exposure of each stimulus element within each second of training (as shown in Figure 1.2.3).
4.3.3.2 Training Stimulation

There are two major differences in training Experiment 3 compared to Experiments 1 & 2. One, instead of stimulation at a single time within each second, there were four points of stimulation spaced 25ms apart and, two, instead of stimulating only one cluster, we stimulated each of the four clusters at one of the times of stimulation. The same randomization algorithm (and same seed) used in Experiments 1 & 2 was used to choose the 13 neurons in each cluster.

4.3.3.3 Anatomical Changes from Training

Unlike the differential growth in the right tail of the synaptic distribution in Experiment 1 (Figure 4.1.3.4.1), the synaptic distributions look very similar through the range of potential E-E synaptic weights (Figure 4.3.3.3.1).

![Figure 4.3.3.3.1: Synaptic Weight Distribution for E-E Synapses](image)

The change in intracluster synaptic weights did not appear to differ greatly between clusters. Both cluster A and B showed a subset of neurons with large increases in both training regimes (Figure 4.3.3.3.2), however, it appears the number of strong synapses (w > 0.4) was
increased more in the repeated training regime compared to the spontaneous activity training regime (Figure 4.1.3.3.3).

Figure 4.3.3.3.2: Synaptic Weight Change for Clusters A & B
Figure 4.3.3.3.3: Differences strong synapses between clusters A&B

An ANOVA between training regimes showed no significant difference in cluster A, but the overall change for each transition of the repeated sequence vs the reversed sequence and the aggregate weights were qualitatively different from baseline (Table 4.3.3.3.1).

Table 4.3.3.3.1: Average E-E Synaptic Weights

<table>
<thead>
<tr>
<th>Synaptic Group</th>
<th>Baseline</th>
<th>Trained: Random Sequences</th>
<th>Trained: Repeated Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra A</td>
<td>0.2081</td>
<td>0.2078</td>
<td>0.2084</td>
</tr>
<tr>
<td>A-&gt;B</td>
<td>0.192</td>
<td>0.1919</td>
<td>0.193</td>
</tr>
<tr>
<td>B-&gt;C</td>
<td>0.1829</td>
<td>0.181</td>
<td>0.1836</td>
</tr>
<tr>
<td>C-&gt;D</td>
<td>0.1932</td>
<td>0.1933</td>
<td>0.1953</td>
</tr>
<tr>
<td>D-&gt;C</td>
<td>0.1929</td>
<td>0.1908</td>
<td>0.1901</td>
</tr>
<tr>
<td>C-&gt;B</td>
<td>0.1832</td>
<td>0.1842</td>
<td>0.1819</td>
</tr>
<tr>
<td>B-&gt;A</td>
<td>0.1934</td>
<td>0.1932</td>
<td>0.1911</td>
</tr>
<tr>
<td>All E-E</td>
<td>0.1998</td>
<td>0.1998</td>
<td>0.1998</td>
</tr>
<tr>
<td>Total ABCD</td>
<td>0.5681</td>
<td>0.5662</td>
<td>0.5719***</td>
</tr>
<tr>
<td>Total DCBA</td>
<td>0.5695</td>
<td>0.5682</td>
<td>0.5631***</td>
</tr>
</tbody>
</table>

***p<0.001

The synaptic weights between the ABCD and DCBA connections showed inverse directions of change in the regimes trained with random versus repeated sequences. Although there was some variation, the connections in the repeated regime each increased from baseline, whereas most of the connections in the random regime decreased. Conversely, the synaptic strength of connections for the reverse sequence in the repeated sequence regime each decreased.
In aggregate there was a significant difference between the cumulative weight of the ABCD sequence between training regimes (F(2,86736)=7.38, p<0.001, one-way ANOVA) and the DCBA sequence (F(2,86418)=11.23, p<<0.001, one-way ANOVA).

**Motif and Clustering Coefficients Analyses**

Both Repeated and Random Sequence training increased the number of 3-neuron motifs from Baseline (network thresholded at w>0.33). Neither condition showed a change in the average clustering coefficient. The changes in motif distribution from cluster A->B are shown as a representative example of other associations in the repeated sequence (Figure 4.3.3.3.4).

![Motif and Clustering Coefficients Analyses](image)

**Figure 4.3.3.3.4: Motifs and Clustering Coefficients for the sequence training regimes (w>0.33)**

### 4.3.3.4 Stimulation Procedure for Experiment 3

Experiment 3 used an equivalent test case pairing to the training regimes used in Experiments 1 & 2. Instead of presenting a single stimulus (“A”) or no stimulus, we presented a repeated sequence (ABCD) or the reversed sequence (DCBA). Each sequence element had a 25ms ISI between presentations (green arrows in Figure 4.3.3.4.1). Stimulation of a sequence element (e.g. Stimulus A) involved the same neurons across training regimes and test cases to mitigate bias associated with idiosyncrasies inherent in the baseline network.
4.3.4 Functional Changes Resulting from Training

As shown by the anatomical changes in Table 4.3.3.1, associations between sequence elements are learned using STDP despite the element presentation being outside of the window of STDP mechanisms. Despite these anatomical changes, the functional results were inconclusive. The firing rates produced by presenting the sequences ABCD and DCBA to the regimes trained on ABCD (Figure 4.3.4.1) and random sequences (Figure 4.3.4.2) look similar between training regimes and test cases. When comparing the firing rates between test cases in the repeated sequence presentation regime, we saw general increases in the previously presented sequence, but not consistently so (Figure 4.3.4.3).
Figure 4.3.4.1: Average Firing Rates for the Network trained on ABCD

Figure 4.3.4.2: Average Firing Rates for the Network trained on Random Sequences

Figure 4.3.4.3: Average Firing Rate by Time Bin of Network trained with Repeated ABCD Sequence
4.3.5 Discussion for Experiment 3

Anatomical differences were evident as a result of the different training regimes. The synaptic weights between sequence elements in the repeated sequence training regime increased relative to baseline and the random sequence training condition. Additionally, the repeated sequence regime showed a decrease in strength of synapses in the reverse direction to the learned sequence (e.g. B->A).

Functionally, it appears that repeated presentation of a sequence results in an increased response to that sequence compared to baseline (data not shown), however, there is not an evident difference with the regime shown random sequences. These findings suggest that the expected changes are occurring in the synaptic connections between sequence elements, but are inconclusive in the functional response to those anatomical changes.

5 General Discussion

Our experimental findings support the process memory framework, in that a simple, constrained cortical circuit can show memory effects without requiring input from a dedicated “memory system”. These experiments demonstrate a flexible cortical mechanism, based on STDP, by which local cortical circuits can lengthen their intrinsic memory for (and response to) temporally ordered sequence elements to which they have been given repeated prior exposure.

In Experiment 1, we hypothesized that a local cortical circuit repeatedly exposed to the same stimulus would show a larger neuronal response compared to an unexposed circuit for the same stimulus. Circuit plasticity manifested as an increased firing rate and duration of the dynamics elicited by the stimulus (Figure 4.1.4.1). These functional changes appeared to result from an increase in intracluster synaptic strength within the stimulated cluster (Figure 4.1.3.4.2, Table 4.1.3.4.1).

In Experiment 2, we hypothesized that a circuit could predict stimulus presentation or omission using neuronal firing with greater accuracy when the circuit had previously seen the stimulus. Superior predictive accuracy was evident over a longer duration when the circuit had been pre-exposed to the stimulus (Figure 4.2.4.1.1). We also showed that the extended duration of process memory for the stimulus-exposed circuit could not simply be explained by the
duration of increased neuronal firing (Figure 4.2.4.1.2). The process memory elongation appeared to result from the anatomical changes discussed in Experiment 1.

Finally, in experiment 3, we hypothesized that a local cortical circuit repeatedly exposed to the same sequence would show a larger neuronal response for that sequence compared to its reversed sequence. We showed that synapses between the neurons responsive to sequence elements were potentiated in a manner that was asymmetric and selective to the order of the repeated sequence (Figure 4.3.3.3.3). Despite this unidirectional potentiation, a difference between the firing rate in the repeatedly-seen sequence and its reversed sequence was not evident (Figure 4.3.4.1).

We were able to decode whether stimulation occurred or was absent above chance levels for more than 130ms following presentation. The duration of this post-stimulus process memory extended beyond the point when firing rate dropped below baseline (38ms). This suggests that specific neuronal states contain information about the stimulus well into the future, and the persistence of this information is greater for a circuit that has been pre-exposed to the stimulus. We speculate that the lengthening of process memory for individual stimuli within a sequence enables potentiation of the synapses from the neurons responsive to a sequence element and subsequent elements of the sequence.

Repeated exposure to a stimulus may result in a more reliable neuronal representation of that stimulus. Given that increased reliability, subsequent neuronal states should form with increasing reliability. This flow of increasingly reliable neuronal states over time in response to a stimulus could carry forward into a subsequent neuronal states resulting from a subsequent sequence element. Thus repeated stimulus exposure could crystallize neuronal ensembles, where those ensembles form more reliable downstream connections, and then the synaptic connections to neurons responsive to a subsequent sequence element (which also becomes more reliable after repeated exposure) may form in the specific timing associated with the sequence.

While synaptic changes were elicited in both the single stimulus and sequence learning experiments, it remains unclear which aspects of the synaptic plasticity paradigm most influenced our results. Despite that, it appears that heterosynaptic plasticity and STDP are both critical elements for the learning demonstrated by the model.
Without heterosynaptic plasticity, synaptic strength grows until the network becomes unstable (data not shown). The synaptic stabilization shown in section 3.6 is only possible because of depotentiating pressure on synapses that would otherwise remain at a relatively constant strength. If those same synapses had not been depotentiated, then interference from their post-synaptic partners could reduce the information transmission capacity in resultant neuronal states, thus decreasing process memory.

The other obvious plasticity mechanism is STDP. With an equal time window for potentiation and depotentiation, this form of synaptic plasticity results in a synapse from one neuron to another neuron potentiating in one direction and depotentiating in the other. Without a mechanism like this, for example if the synaptic plasticity was bi-directionally potentiated with concurrent activation of neurons within the plasticity time window, the network would become over-potentiated and unstable. In addition to instability, bi-directional potentiation would result in unnecessarily potentiated synapses and produce greater competition between synapses due to heterosynaptic plasticity.

Our findings suggest that repeated exposure to specific stimuli will result in increased neural responses, which implicitly utilize ensembles of neurons to lengthen the effective memory of a stimulus and allow associations between elements to be learned over an increasing timescale.

6 Future Directions and Model Predictions

As stated in the Introduction, this examination serves as a proof-of-concept for how process memory could arise as a result of repeated exposure to stimuli. Although sequence learning is necessarily experience-dependent, learning subsequent sequence elements may vary in difficulty based upon the persistent availability of resonant information in online memory. As an example of potential memory interference, it could be that a stimulus exerts too strong a response, thus wiping the process memory for stimuli in the past. This is a potential cause of the results shown in Section 4.3.4 and will be examined in future tests of sequence learning effects on firing.
The prediction of elongating implicit process memory suggests that behavioral experiments in humans might be used to the effects of process memory on the ease of learning. Varying object complexity or sequence timing are both candidates to test process memory. As with previous implicit and statistical learning experiments, response time tasks and artificial grammar learning could be good candidates to probe the existence of process memory for previously seen stimuli (Nissen and Bullemer, 1987; Reber, 1967). The model predicts that learning a sequence within the plasticity time window should be immediately possible, whereas sequence elements with an ISI longer than the plasticity time window will take longer to learn.

We would like to investigate how this model could be used to investigate predictive coding (Clark, 2013). It is possible that given repeated presentations of stimulus A, then stimulus B, presentation of A could result in activation of neurons responsive to B, despite not presenting B. We can compare the activation of neurons responsive to B when stimulus B is presented or omitted for this paired training and compare the stimulus presentation and omission in a circuit that was not exposed to the AB pair.

Our results do not show whether sequence learning is simply multiple pairwise connections of adjacent sequence elements or if temporally non-adjacent elements are impacted by prior sequence elements. Future tests will assess whether there is greater memory in the circuit for presentation of a complete learned sequence compared to the same sequence with an initial element omitted. Our theory suggests that a system repeatedly trained on ABC will have resonant process memory in the neuronal states associated with later elements resulting in a quantitative difference between neuronal firing when BC is presented after A versus BC without A. This can be tested behaviorally with sequences of images that are repeatedly presented to human participants. One would expect higher accuracy and reduced response time to the “C” element in the sequence where A was presented first vs when A was omitted in testing.

While the experiments helped answer the research questions posed, the model produced also provides a basis for testing additional questions. For example, if, as we suggest, neuronal states are formed based upon an increase in neuronal firing, are there chains forming between neuronal states in the repeated sequence regime? We are developing methods to answer this question.
References


