

BOSTON UNIVERSITY  
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**MODELING BIOPHYSICAL AND NEURAL CIRCUIT BASES FOR CORE  
COGNITIVE ABILITIES EVIDENT IN NEUROIMAGING PATTERNS:  
HIPPOCAMPAL MISMATCH, MISMATCH NEGATIVITY, REPETITION  
POSITIVITY, AND ALPHA SUPPRESSION OF DISTRACTORS**

by

**STEFAN BERTEAU**

B.S., The American University, 2002

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

2017

© 2017  
STEFAN BERTEAU  
All rights reserved

Approved by

First Reader

---

Daniel Bullock, Ph.D.  
Professor of Psychological and Brain Sciences

Second Reader

---

Ennio Mingolla, Ph.D.  
Professor, Department of Communications Sciences and Disorders  
Northeastern University

Third Reader

---

Paul Miller, Ph.D.  
Associate Professor of Biology  
Brandeis University

## **DEDICATION**

I would like to dedicate this work to my parents, who raised me to value the search for knowledge and understanding, and who supported me in every way a child could hope for over this long journey.

## **ACKNOWLEDGMENTS**

I would like to acknowledge the help of my advisors, Daniel Bullock, Ennio Mingolla, Robert Sekuler, and Paul Miller. Their advice and assistance through this project have been invaluable. I would also like to particularly acknowledge the contribution of Dr. Xiao-ping Lu for initially suggesting that the M-current might be what I was looking for in terms of dynamically labeled predictive coding. Finally, a special thanks to Mr. Kuykendall, my fifth grade science teacher and the first person to call me Dr. Berteau.

STEFAN BERTEAU

Boston University Graduate School of Arts & Sciences, 2017

Major Professor: Daniel Bullock, Professor of Psychological and Brain Sciences

## ABSTRACT

This dissertation develops computational models to address outstanding problems in the domain of expectation-related cognitive processes and their neuroimaging markers in functional MRI or EEG. The new models reveal a way to unite diverse phenomena within a common framework focused on dynamic neural encoding shifts, which can arise from robust interactive effects of M-currents and chloride currents in pyramidal neurons. By specifying efficient, biologically realistic circuits that achieve predictive coding (e.g., Friston, 2005), these models bridge among neuronal biophysics, systems neuroscience, and theories of cognition.

Chapter one surveys data types and neural processes to be examined, and outlines the Dynamically Labeled Predictive Coding (DLPC) framework developed during the research. Chapter two models hippocampal prediction and mismatch, using the DLPC framework. Chapter three presents extensions to the model that allow its application for modeling neocortical EEG genesis. Simulations of this extended model illustrate how dynamic encoding shifts can produce Mismatch Negativity (MMN) phenomena, including pharmacological effects on MMN reported for humans or animals.

Chapters four and five describe new modeling studies of possible neural bases for alpha-induced information suppression, a phenomenon associated with active ignoring of stimuli. Two models explore the hypothesis that in simple rate-based circuits, information

suppression might be a robust effect of neural saturation states arising near peaks of resonant alpha oscillations. A new proposal is also introduced for how the basal ganglia may control onset and offset of alpha-induced information suppression. Although these rate models could reproduce many experimental findings, they fell short of reproducing a key electrophysiological finding: phase-dependent reduction in spiking activity correlated with power in the alpha frequency band.

Therefore, chapter five also specifies how a DLPC model, adapted from the neocortical model developed in chapter three, can provide an expectation-based model of alpha-induced information suppression that exhibits phase-dependent spike reduction during alpha-band oscillations. The model thus can explain experimental findings that were not reproduced by the rate models. The final chapter summarizes main theses, results, and basic research implications, then suggests future directions, including expanded models of neocortical mismatch, applications to artificial neural networks, and the introduction of reward circuitry.

## TABLE OF CONTENTS

DEDICATION.....	iv
ACKNOWLEDGMENTS .....	v
ABSTRACT.....	vi
TABLE OF CONTENTS.....	viii
LIST OF TABLES .....	xiv
LIST OF FIGURES .....	xv
LIST OF ABBREVIATIONS.....	xxv
CHAPTER ONE: INTRODUCTION.....	1
The Predictive Coding Hypothesis .....	1
KCNQ and the M-current .....	2
Phasic-Tonic Mode Switching.....	4
Labile Encoding .....	5
Rate Based Encoding .....	6
Labeled Line Encoding .....	7
Dynamically Labeled Predictive Coding .....	9
CHAPTER TWO: HIPPOCAMPAL MISMATCH DETECTION .....	14
The Functional Anatomy of CA1 as a Comparator.....	14
Existing CA1 Comparator Models.....	16

<b>A Biologically Detailed Comparator that Uses Pyramidal Mode Switching .....</b>	<b>17</b>
<b>Materials and Methods.....</b>	<b>22</b>
<i>Proposed Model .....</i>	23
<i>Architecture .....</i>	23
<i>Network Connectivity.....</i>	23
<i>Membrane and Synapse Equations .....</i>	24
<i>Input .....</i>	29
<i>Output.....</i>	30
<i>Pharmacological Manipulations .....</i>	30
<i>Data Analysis .....</i>	30
<b>Results.....</b>	<b>31</b>
<i>Investigation 1: Match/Mismatch .....</i>	31
<i>Investigation 2: Pharmacological Manipulations .....</i>	33
<b>Discussion.....</b>	<b>36</b>
<b>CHAPTER THREE: A BIOLOGICALLY REALISTIC SINGLE-CELL</b>	
<b>COMPARATOR MODEL OF THE MISMATCH NEGATIVITY .....</b>	
<i>Abstract.....</i>	38
<i>Introduction.....</i>	40
<i>The MMN .....</i>	41
<i>A Controversy.....</i>	44
<i>Constraints on MMN Models .....</i>	45

Existing Models .....	49
Limitations of Existing Models .....	52
Proposed New Model.....	56
Materials and Methods.....	60
Proposed Model .....	62
Architecture.....	62
Network Connectivity.....	62
Neuron and Synapse Models .....	63
Input .....	64
Output .....	65
Data Analysis .....	65
Results.....	69
Investigation 1: Match/Mismatch .....	69
Investigation 2: Repetition Positivity.....	69
Investigation 3: Pharmacological Manipulations.....	71
Investigation 4: Robustness to Noise.....	78
Investigation 3: Confusability Testing.....	81
Discussion .....	82
Limitations of the New Model .....	83
Prediction and Observation.....	83
Future Directions .....	84
Appendix .....	85

Elements of the Model .....	85
Membrane Equations .....	85
Parameters.....	88
Simulating electrical field effects generated by pyramidal neuron dynamics .....	89
<b>CHAPTER FOUR: A MECHANISTIC MODEL OF ALPHA-INDUCED INFORMATION SUPPRESSION .....</b>	<b>98</b>
Significance Statement.....	98
Abstract.....	98
Introduction.....	100
Materials and Methods.....	102
Model .....	104
Network Architecture.....	104
Network Connectivity.....	107
Neuron and Synapse Models .....	108
Input (Stimuli).....	111
Output (Data Recorded).....	113
Analysis.....	113
Results.....	116
Investigation 1: Mutual Information, Single Thalamo-Cortical Loop.....	116
Investigation 2: Discrimination Threshold in a Single Thalamo-Cortical Loop ....	120
Investigation 3: Mutual Information between Two Thalamo-Cortical Loops.....	122
Discussion.....	123

<b>CHAPTER FIVE: TOWARD A MORE REALISTIC MODEL OF ALPHA-INDUCED INFORMATION SUPPRESSION .....</b>	<b>128</b>
Introduction.....	128
Basal Ganglia Control of Thalamo-Cortical Alpha .....	129
Methods.....	134
Model .....	136
Network Architecture.....	136
Network Connectivity .....	138
Neuron and Synapse Models .....	139
Input (Stimuli).....	142
Output (Data Recorded).....	144
Analysis.....	144
Results.....	146
Investigation 1: Mutual Information .....	146
Investigation 2: Discrimination Threshold .....	149
Discussion .....	151
Limitations of the current model.....	151
Methods.....	156
Model Overview .....	156
The Predictive Coding Hypothesis .....	156
KCNQ and the M-current .....	157
Phasic-Tonic Mode Switching.....	158

Network Architecture.....	160
Network Connectivity.....	160
Neuron and Synapse Models .....	161
Input.....	161
Output .....	162
Results.....	162
Discussion.....	165
<b>CHAPTER SIX: DISCUSSION AND FUTURE DIRECTIONS .....</b>	<b>167</b>
Expectations, Surprise, Resolution, and Reward: Sustainable Novelty .....	168
Down the Garden Path .....	168
<b>BIBLIOGRAPHY .....</b>	<b>170</b>
<b>CURRICULUM VITAE .....</b>	<b>184</b>
Stefan Berteau .....	184
Education .....	184
Publications, Presentations, and Patents .....	185
Research Experience.....	187

## LIST OF TABLES

Table 2.1: Model Parameters .....	27
Table 3.1: Summary of pharmacological manipulations of the MMN. Particularly relevant effects are the dose-dependent reduction in MMN amplitude under NMDA blockade and the reduction in MMN amplitude in Alzheimer's populations under a cholinesterase inhibitor.....	46
Table 4.1: Model weights and conduction delays, with evidence for each type of connection .....	108
Table 4.2: Population parameters .....	110

## LIST OF FIGURES

Figure 1.1: Simulated voltage clamps demonstrating a voltage-dependent time constant for the KCNQ channels that mediate M-currents. Abruptly instating and clamping new transmembrane voltages (blue: -50 mV; green: -40 mV; red: -30 mV; light blue: -20 mV) between 200 ms and 700 ms show the dependence of the KCNQ channel's time constant on voltage. Details of the mathematical simulation can be found in chapter 2. ....	4
Figure 1.2: An example of CA1 pyramidal neurons switching between tonic and phasic modes of firing under different injected currents (y-axes) and shunting conductances (the two columns). Increased shunting produced a change from tonic to phasic firing, and increased the amount of stimulation (pA) required to produce even phasic firing.....	5
Figure 1.3: Circuitry mediating cortical predictions.....	11
Figure 2.1: Hippocampal mismatch circuitry, adapted from Lisman, (1999). Recurrent connectivity between the dentate gyrus and CA3 helps to create representations of current inputs that are sensitive to the recent stream of inputs. The predictions represented by synaptic weights between CA3 and CA1 can therefore be highly context dependent. ....	16
Figure 2.2: A) A standard predictive coding comparator, where a stimulus elicits predictions which inhibit a population, reducing the frequency of its output when the predicted stimulus occurs. B) A standard habituative comparator, which fires at a reduced rate due to afferent habituation when a recently experienced stimulus is	

repeated. C) A mode-shifting comparator, wherein a stimulus elicits predictions which prime a population, placing it into a high-conductance state and causing a shift from tonic to phasic firing. .... 21

Figure 3.1: A cartoon example of a mismatch negativity. The dashed line represents the average electric field generated in response to oddball stimuli, the solid line represents the average electric field generated in response to standards, and the difference is shaded grey. Negative voltages are plotted above the zero voltage line in this plot, in accord with the tradition used in earlier MMN reporting. In this plot, with tick marks representing 10 ms intervals, the MMN emerges around 80 ms, and lasts until 150 ms, post-event..... 42

Figure 3.2: Three examples of stimulus sequences from the oddball paradigm. Top: a series of ascending digits ending with a letter, this breaks both the pattern of digits and the pattern of ascending order. Middle: a series of ascending musical tones, followed by a larger descending step. This breaks both the pattern of ascending order and the pattern of consistent step size. Bottom: A series of colored squares which abruptly change color from green to red. This breaks the pattern of consistent color. .... 43

Figure 3.4: The network architecture of Wacongne et al. (2012), designed for a simulated experiment with two sounds. It consists of an abstracted higher-level memory trace, and then two cortical subpopulations for each layer (layer 2/3 and layer 4), as well as two subpopulations of thalamic input. Adapted from Wacongne et al. (2012) ..... 49

Figure 3.5: May and Tiitinen’s adaptation model. In the adaptation model, the standards and deviants activate overlapping neural populations. The repetitive standard leads to cells tuned to the standard to become adapted. When the deviant is presented, nonadapted “fresh afferents” contribute to an enhanced response. Also shown is the beginning of a stimulus sequence, four standards (S) followed by a deviant (D), and the event-related responses produced by the neural population pictured at the top.

Adapted from May & Tiitinen, (2010) ..... 51

Figure 3.6: The architecture of a traditional memory-based conception of the MMN, with separate modules performing the feature detection and the comparison and MMN generation. Adapted from May & Tiitinen, (2010b)..... 53

Figure 3.9: Baseline mismatch negativity. A) A rapid tonic firing from the unprimed comparator. B) A brief phasic burst from the primed comparator. C) The negative deflection of an unprimed comparator. D) The positive and slight negative deflection of the primed comparator. E) The difference between the two electric fields, showing a negative deflection, or MMN. Negative is plotted downward in the electric field panels. ..... 69

Figure 4.10: The simulated repetition positivity scales with priming strength. As the lambda was reduced from 100% of investigation 1 (darkest blue line) to 33% (lightest blue), the amplitude of the simulated repetition positivity dropped. Negative is plotted downward. ..... 70

Figure 3.11: The model under an NMDAR blockade, as under PCP. A) The unprimed comparator voltage trace shows extremely rapid tonic firing. B) The primed

comparator voltage trace shows fitful firing throughout the stimulus, failing to achieve a full mode shift. C) The deflection in the electric field for the unprimed comparator is reduced. D) There is a similar reduction in the negative deflection of the primed comparator's field. E) The MMN is reduced compared to baseline (c.f. investigation 1). Negative is plotted downward in the electric field panels. .... 71

Figure 3.12: The model under reduced KCNQ conductance, as would be produced by increased ACh. A) A rapid tonic firing from the unprimed comparator. B) A similar rapid tonic firing from the primed comparator. C) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the primed comparator. E) The negative deflection of an unprimed comparator. D) The positive and slight negative deflection of the primed comparator. E) The difference between the two electric fields, showing a negative deflection, or MMN. Negative is plotted downward in the electric field panels. Note that while the neural voltage trace in panel B is clearly distinguishable from the corresponding trace in the baseline figure, the MMN remains undiminished. This is due to the electrical field being derived from the difference in apical and basal transmembrane currents, unlike the basal voltage trace (which determines spike output). A more complete explanation is found below..... 72

Figure 3.13: The model under increased KCNQ conductance, as would be produced by retigabine. A) A single phasic onset spike from the unprimed comparator. B) A similar phasic onset firing from the primed comparator. C) The transmembrane

(external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the primed comparator. E) The negative deflection of an unprimed comparator. F) The positive and negative deflection of the primed comparator. G) The difference between the two electric fields, showing a diminished negative deflection, or MMN. Negative is plotted downward in the electric field panels. Note that in this case the MMN is extremely diminished, in accordance with the failure of the spiking output of the comparator to distinguish between primed and unprimed conditions..... 74

Figure 3.13: The model under both an NMDA antagonist and a KCNQ agonist. A) A rapid tonic firing from the unprimed comparator. B) A similar rapid tonic firing from the primed comparator. C) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the primed comparator. E) The slight negative deflection of an unprimed comparator. F) The slight positive and negative deflection of the primed comparator. G) The difference between the two electric fields, showing an extremely diminished negative deflection, or MMN. Negative is plotted downward in the electric field panels. .... 76

Figure 3.14: Classification performance as a function of noise. As injected white noise increases, the rate of true positives (green) falls and that of false negatives (cyan) rises. True negatives (red) and false positives (blue) remain unchanged. In this

figure the true positive rate is the proportion of expected stimuli which evoke phasic firing, which would correctly be decoded to indicate an expectation was matched.

The false negative rate is the proportion of expected stimuli which evoke tonic firing, as though they were unexpected. Likewise, the true negative rate is the proportion of unexpected stimuli which correctly evoke tonic firing, and the false positive rate is the proportion of unexpected stimuli which evoke phasic firing. .... 80

Figure 4.1: Our model circuitry. Two brain regions (the thalamus and cortex) are modeled, each with populations of excitatory and inhibitory neurons. Arrows represent excitatory connections, circles inhibitory..... 105

Figure 4.2: The expanded model used in investigation three. Two copies of the first model were connected via a unidirectional excitatory connection between the cortical excitatory populations. The “signal” entering the second population is considered noise, for purposes of mutual information calculation. .... 106

Figure 4.3: Response curves and saturation of different model neuron types. A-D) f-I curves for the four neural populations, showing firing rate as a function of input current. E) Oscillation amplitude and cortical response as a function of frequency. The colored background indicates the steady state cortical excitatory response amplitude when it is driven by input of a given frequency (x-axis) and given peak amplitude (y-axis). The actual peak amplitude of thalamic drive as a function of frequency is represented by the white line..... 111

Figure 4.4: Simulation Results demonstrating A) the Ornstein-Uhlenbeck-generated input signal, B-C) the thalamic activity at roughly 10 and 6 Hz, respectively, D-E) the

cortical activity driven by a combination of the input signal and the thalamic activity, and F) the significant difference in mutual information between the 6.1 Hz and 10 Hz conditions. .... 119

Figure 4.5: Particular properties of the model system. A) The thalamic frequency saturated as frontal input to the relay nucleus cells increased. B) The mutual information varied with frequency, increasing in the theta range and falling to very low levels below theta and in alpha. .... 119

Figure 4.7: The results of investigation 3, using two coupled thalamo-cortical loops. A) The mutual information by relative phase angle at roughly 10 Hz, demonstrating the lack of a preferred phase in the alpha band, with all MI values below 0.023. B) The mutual information by relative phase angle at roughly 6 Hz, demonstrating the strongly preferred phase. C) Mean mutual information is reduced by a lack of coherence in the oscillations, and the reduction is much greater (-0.022) in theta, where there is more best-phase MI to lose than in alpha (-0.002). D) The preferred phase from B can be altered by increasing the conduction delay from 12ms to 60ms between the two regions..... 123

Figure 5.1: The pathways by which learned excitation of D2-MSPNs in the striatum can control alpha generation. Blue arrows represent excitatory connections, red circles represent inhibitory connections, and green bars represent ACh-releasing connections. Thicker lines denote stronger connections. A comparison of the left panel (before learning) with the right (after learning) shows that the learned cortical excitations of D2-MSPNs in striatum has the effect of changing three inputs to TC

(thalamo-cortical) neurons: enhanced inhibition from TRN (thalamic reticular nucleus) and GPi (internal/medial globus pallidus) and reduced cholinergic input from nucleus basalis. D2-MSPNs are striatal medium spiny projection neurons that express D2-type dopamine receptors..... 132

Figure 5.2: Our model circuitry. Two brain regions (the thalamus and cortex) are modeled, each with populations of excitatory and inhibitory neurons. Blue arrows represent excitatory connections, red circles represent inhibitory connections, and green triangles represent gap junctions. The basal ganglia inputs shown are from the GPi. Also, modeled, but not shown are two further inputs from the basal ganglia/forebrain: an inhibitory input from GPe to TRN, and a cholinergic input from NBM to TRN cells. .... 137

Figure 5.3: Simulation Results demonstrating A) the Ornstein-Uhlenbeck-generated input signal, B-C) the thalamic activity at roughly 10 and 6 Hz, respectively. D-E) the cortical activity driven by a combination of the input signal and the thalamic activity, and F) the significant difference in mutual information between the 6.1 Hz and 10 Hz conditions. .... 149

Figure 5.5: Macaque medial prefrontal cortex neural firing rate (extracellularly recorded) normalized to average pretrial firing rate as a function of alpha power. Adapted from Haegens, Nácher, Luna, Romo, & Jensen, (2011)..... 153

Figure 5.6: Macaque medial prefrontal cortex neural firing rate (extracellularly recorded) normalized to average pretrial firing rate as a function of binned alpha phase. Adapted from Haegens, Nácher, Luna, Romo, & Jensen, (2011). .... 154

Figure 5.7: Simulated voltage clamps demonstrating a voltage-dependent time constant for the KCNQ channels that mediate M-currents. Abruptly instating and clamping new transmembrane voltages (blue: -50 mV; green: -40 mV; red: -30 mV; light blue: -20 mV) between 200 ms 700 ms show the dependence of the KCNQ channel's time constant on voltage. Details of the mathematical simulation can be found in chapter 2 .....	158
Figure 5.8: And example of CA1 pyramidal neurons switching between tonic and phasic modes of firing under different injected currents (y-axes) and shunting conductances (the two columns). Increased shunting produced a change from tonic to phasic firing, and increased the amount of stimulation (pA) required to produce even phasic firing.....	159
Figure 5.9: A simplified subset of S. Jones et al., 2009, provided for reference to the inputs to L2/3, specifically apical and feedforward inhibitory inputs from higher order cortex (in the MMN model in Chapter Four) and non-specific thalamic neurons seen in C, and the sensory input to basal dendrites from Layer 4, as seen in B.....	161
Figure 5.10: Upper Left) the ongoing alpha drive, in the form of an oscillatory change in the priming lambda. Lower Left) The cell membrane voltage trace, showing tonic firing when stimulus input arrives in the trough of alpha. Upper Right) The same ongoing alpha drive, reproduced for ease of comparison. Lower Right) The cell membrane voltage trace, showing phasic firing when stimulus input arrives at the peak of alpha.....	164



## LIST OF ABBREVIATIONS

$\mu$ A .....	microAmps
3D .....	Three Dimensional
5HT .....	Serotonin
ACh .....	Acetylcholine
AChN .....	Acetylcholinergic Neuron
AD .....	Alzheimer's Disease
AHP .....	Afterhyperpolarization
AMPA .....	$\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ART .....	Adaptive Resonance Theory
AUC .....	Area Under Curve
BG .....	Basal Ganglia
CA1 .....	Cornu Ammonis area one
CA3 .....	Cornu Ammonis area three
Cl <sup>-</sup> .....	Negative Chloride Ion
cm .....	centimeter
CRC .....	Color-Rule-Coding
D .....	Deviant
D2-MSPN .....	Medium Spiny Projection Neuron Expressing Type Two Dopamine Receptors
DG .....	Dentate Gyrus
DLPC .....	Dynamically Labeled Predictive Coding

DLPFC	.....	Dorso-Lateral Pre-Frontal Cortex
EC2	.....	Entorhinal Cortex Layer 2
EC3	.....	Entorhinal Cortex Layer 3
EEG	.....	electroencephalogram
ERP	.....	Event-Related Potential
FEF	.....	Frontal Eye Field
f-I	.....	Firing Rate as a Function of Input
F-M	.....	Frontal-Midline
fMRI	.....	Functional Magnetic Resonance Imaging
Fz	.....	Frontal Zero, the electrode
..	placed frontally and on the midline in the 10-20 placement system	
GABA	.....	$\gamma$ -Aminobutyric acid
GPe	.....	The Exterior Portion of the Globus Pallidus
GPi	.....	The Interior Portion of the Globus Pallidus
HTC	.....	High-Threshold Thalamocortical Cell
Hz	.....	Hertz
ISI	.....	Inter-Stimulus Interval
K	.....	Potassium
KCNQ	.....	Potassium Channel, Voltage-Gated, KQT-like subfamily
k $\Omega$ cm	.....	kilo-Ohm centimeter
L2/3	.....	Neocortical Layers Two and Three
L4	.....	Neocortical Layer Four

LFP .....	Local Field Potential
LFP .....	Local Field Potential
LGN .....	Lateral Geniculate Nucleus
LIP .....	Lateral Intraparietal Cortex
M <sub>1</sub> .....	Muscarinic Type One Receptor
mA .....	milliAmps
mAChR .....	Muscarinic Acetylcholine Receptor
mAChRs .....	muscarinic acetylcholine receptors
MEG .....	Magnetoencephalogram
Mg <sup>2+</sup> .....	Positive Magnesium Ion
mGLuR .....	Metabotropic Glutamate Receptor
MI .....	Mutual Information
MMN .....	Mismatch Negativity
MPC.....	Medial Prefrontal Cortex
ms .....	milliseconds
mV .....	millivolts
N <sub>I</sub> .....	The Negative ERP Component found at 100ms post-stimulus
N <sub>100</sub> .....	The Negative ERP Component found at 100ms post-stimulus (alternate name)
NA .....	Noradrenaline
Na .....	Sodium
NBM .....	Nucleus Basalis Magnocellularis
NE .....	Norepinephrine

NMDA	.....	N-Methyl-D-aspartate
NMDAR	.....	N-Methyl-D-aspartate Receptor
ORC	.....	Orientation-Rule-Coding
P30	.....	Positive ERP Component found at 30 ms post-stimulus
P3a	.....	Positive ERP Component found at 300 ms post-stimulus, part A
P50	.....	Positive ERP Component found at 50 ms post-stimulus
PCP	.....	Phencyclidine
PTB	.....	Hierarchical predictive processing theory of the brain
REM	.....	Rapid Eye Movement
ROC	.....	Receiver Operator Characteristic
RP	.....	Repetition Positivity
S	.....	Standard
SNr	.....	Substantia Nigra Pars Reticulata
SOCRATIC	.....	Sequences of Condensed Representations, Autocorrected, Theta-Gamma Coded, in context
SSA	.....	Stimulus Specific Adaptation
TC	.....	Thalamo-Cortical
TMS	.....	Transcranial Magnetic Stimulation
TRN	.....	Thalamic Reticular Nucleus
VI	.....	Primary Visual Cortex

## **CHAPTER ONE: INTRODUCTION**

This dissertation introduces a proposed dynamic neural code-switching mechanism that is compatible with the predictive coding hypothesis (a framework typified in Friston, 2005). The physiology of this mechanism explains many experimental findings in the realm of predictions, expectations, and attention. To introduce this novel encoding mechanism, I will first provide a brief overview of: the predictive coding hypothesis; M-currents and KCNQ channels; shifts between spike firing modes in neocortical pyramidal cells; and relevant previously-proposed encoding mechanisms, including one of the earliest neural encoding proposals, labeled line encoding. Then I will explain how the physiological interaction of M-currents and Cl<sup>-</sup> currents enable a new type of labeled line encoding – one that is labile and dynamically negotiated – and situate such encoding within the framework of a modified Predictive Coding Hypothesis.

### **The Predictive Coding Hypothesis**

The predictive coding hypothesis, an approach typified by Clark (2013) and Friston (2005), is an extension of the hierarchical predictive processing theory of the brain (often abbreviated PTB), which primarily consists of the ideas that brains are effectively hierarchical prediction testers, and that brains are constantly attempting to minimize prediction errors, e.g., via learning processes that improve predictions. The extension endorsed by Friston consists primarily of defining an encoding strategy wherein the feedback from "higher" to "lower" brain areas predicts features of inputs to lower areas, and inhibits them. The resultant subtraction of predictions from inputs leaves only

unpredicted features (unpredicted “residuals” that may be regarded as error signals) to be transmitted up the hierarchy via feedforward pathways from lower to higher areas. Predictive Coding is most often implemented in a Bayesian framework with little regard to neural constraints, but neurally plausible implementations of predictive coding have recently received some attention (e.g., Wacongne, Changeux, & Dehaene, 2012).

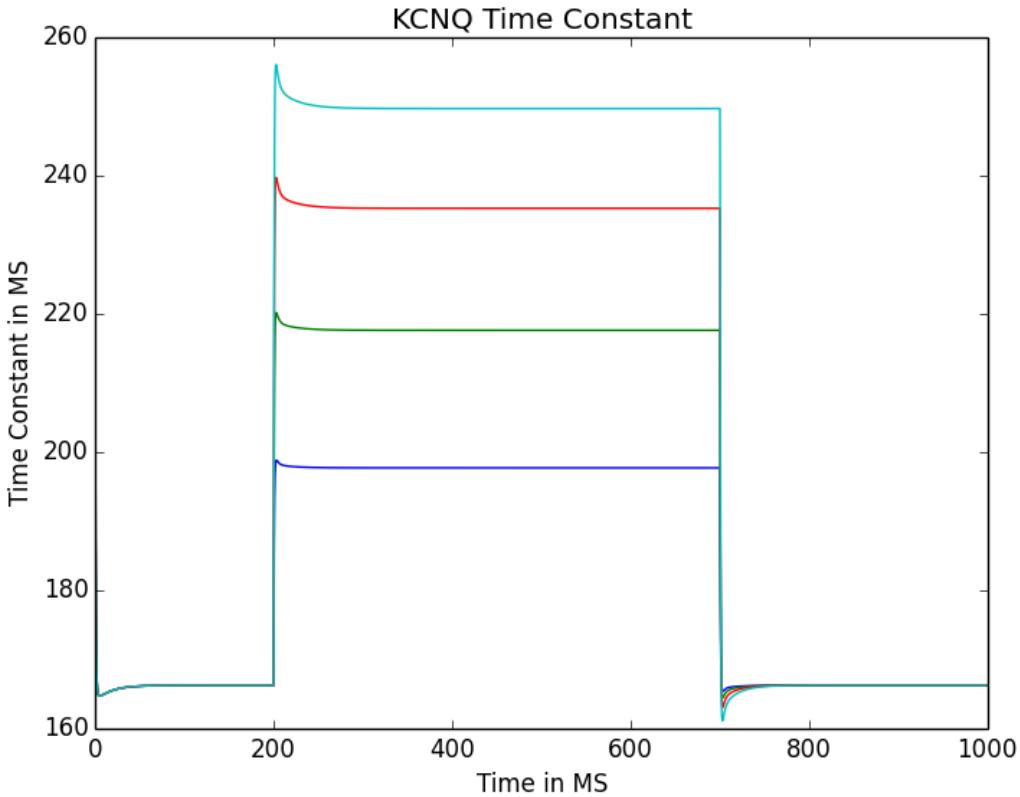
Predictive Coding proposals often come with broader claims that it offers a grand unified theory of perception and cognition arising from the interplay of top-down predictions and bottom-up error signals. I want to stress that the detailed neuronal processing framework that I am developing here is agnostic about, and not dependent on, any of these grander claims.

### **KCNQ and the M-current**

Membrane biophysicists and neurochemists have characterized the M-current as an outward potassium current via channels constituted by KCNQ proteins, often heteromeric KCNQ<sub>2</sub>/KCNQ<sub>3</sub> proteins (Fedorenko et al., 2008). Originally discovered in sympathetic ganglion neurons of bullfrogs (Brown and Adams, 1980), it is non-inactivating and voltage-dependent, being activated by depolarization of the membrane to a level greater than approximately -60mV. Because of these properties, both sub-threshold depolarizations and those large enough to generate action potentials reliably also activate the M-current. However, this is subject to muscarinic neuromodulation. Indeed, the M-current is so named because KCNQ channels are closed by activation of mAChRs (muscarinic acetylcholine receptors). Such closure effectively disables the current. The M-

current mechanism is found in neurons throughout the nervous system, including the pyramidal cells that are the principle neurons of the cerebral cortex (reviewed in Jentsch, 2000; Marrion, 1997) and the medium spiny cells that are the principle neurons of the striatum (reviewed in McCarthy et al., 2008).

Another key feature of the M-current is that the KCNQ channel has a variable time constant: the speed with which it opens and closes is dependent upon the membrane voltage. This is illustrated in Figure 1.1 by plots of the time constants resulting from four simulated voltage clamps, and as will be shown in chapters two and three this property allows a "plateau" of KCNQ conductance to develop during sustained spiking, which when high enough can end the spiking behavior entirely, transforming a tonic response into a phasic one.

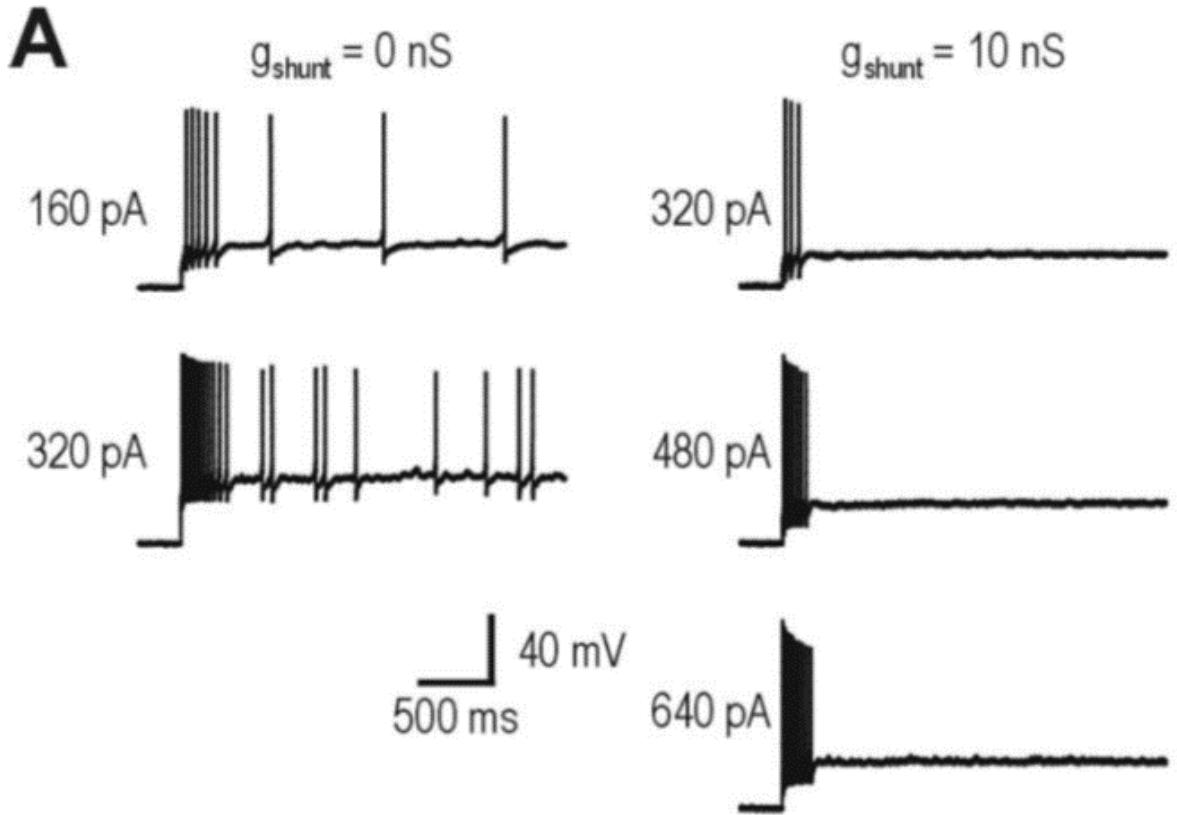


**Figure I.I:** Simulated voltage clamps demonstrating a voltage-dependent time constant for the KCNQ channels that mediate M-currents. Abruptly instating and clamping new transmembrane voltages (blue: -50 mV; green: -40 mV; red: -30 mV; light blue: -20 mV) between 200 ms and 700 ms show the dependence of the KCNQ channel's time constant on voltage. Details of the mathematical simulation can be found in chapter 2.

### Phasic-Tonic Mode Switching

The M-current is usually thought of as a rate-modulating current, which contributes to the oft-observed slow adaptation in spike rate during prolonged excitation (Madison and Nicoll, 1984). Under certain circumstances, however, it can act quite differently, as illustrated in Figure 2. Notably, Prescott et al. (2006) showed that the M-current can, via an interaction with Cl<sup>-</sup> currents, reliably cause a shift from regular (tonic) spiking to onset

(phasic) bursting. This was demonstrated both in computational models (similar in structure to our proposed model below) and in dynamic clamp recordings of pyramidal neurons from the CAI part of the hippocampus.



**Figure 1.2:** An example of CAI pyramidal neurons switching between tonic and phasic modes of firing under different injected currents (y-axes) and shunting conductances (the two columns). Increased shunting produced a change from tonic to phasic firing, and increased the amount of stimulation (pA) required to produce even phasic firing.

### Labile Encoding

As defined by Perkel & Bullock (1968), labile encoding is the use of different kinds of representations and transformations by a given individual at different times under

different circumstances. Later, the ability of neurons to switch between tonic and phasic modes is proposed as a key basis for labile encoding by neural circuits. Labile encoding has been shown to operate in the central nervous system at the neuronal level. For example, changes in encoding/representation (as determined by regression models correlating neural activity to load and posture during a loaded reaching/holding task) by motor cortical (presumed pyramidal) neurons were observed within the course of a single trial (Kurtzer, Herter, & Scott, 2005).

The phenomenon of labile encoding highlights the general information processing principle that successful encoding, and decoding for effective use, is not just about how information is sent. It's also about how it is received. A rate-based code is useless if the time constant of the next-stage integrator is too fast; and precise spike timing with single spike signaling is not helpful if the target neuron requires a rapid spike train to depolarize it enough to fire.

### Rate Based Encoding

As Perkel and Bullock wrote in 1968, frequency or rate is not by itself an adequate specification for encoding, so it is best thought of as one defining property of a class of codes. For our purposes, we will use *rate-based encoding* to refer to integration-and-reset-based encoding. Here, the rate of spikes is what matters, and it is decoded via neural integration, such that the net impact on the post-synaptic neuron is (at least for a viable range of rates) proportional to the presynaptic spike rate over a brief time period.

It is important to recognize both that some neuronal modes may lie outside the rate-based domain, and that remaining within this rate-based domain does not mean changes in encoding are not occurring. For instance, anything that alters the encoded rate without also altering the representation of the encoded feature (e.g. habituation to a constant-intensity visual edge) may require a coordinated adjustment of the decoder, and exemplify labile encoding.

### **Labeled Line Encoding**

In labeled line encoding, the presence of electrical activity in particular lines (here meaning axons, nerve fibers or groups of fibers), but not elsewhere, is what conveys most of the information to the rest of the nervous system about events that have transpired. Originally proposed by Muller (1838) in "On the specific energies of nerves", it has been observed repeatedly, such as in optic nerve fibers, where a particular fiber indicates where in the retinotopic map visual stimulation occurred. Historically, it has been assumed that the "label" for a given fiber was generally static and pre-existing (Perkel & Bullock, 1968). This assumption has caused labeled line encoding to be dismissed when it comes to regions of the central nervous system such as the cortex, because detailed representation of the vast variety of possible perceived stimuli would require an enormous number of pre-labeled lines, most of which would be quiescent at most times. For instance, if each level of pressure felt on a fingertip required an individually pre-labeled line passing through each brain region that processes pressure-related information, the necessary brain mass would rapidly become unsustainably large.

In the early stages of sensory processing in the neocortex we often face a combination of rate-based and labeled line encoding, where some aspects of a stimulus are encoded via labeled line (as in retinotopy in vision). Other aspects are represented via rate-based features, like the intensity of the oriented edge detected at a particular retinotopic location in VI.

However, imagine that labile encoding is in play, and that neural representation of say, visual feature intensity can switch between a rate-based encoding scheme and labeled line encoding via a process that falls under the rubric of "priming". Primed neurons are modulated by top-down input to act as though they are in a labeled line, hence able to represent a value simply by responding, without needing a lengthy spike train. What value is being encoded by this primed line? That is where the expectations driving the top-down priming come into play, with the priming effectively setting up, or dynamically negotiating, a labeled line specifically for the brain's predictions. Provided that expectation-based priming can be learned, this mechanism obviates the need for a massive battery of labeled lines, most of which are rarely active, and sets up short-lived labeled lines only for the most likely outcomes in a given situation. Each such dynamically labeled line can fire a single brief ("phasic") burst to confirm whatever expectation caused its priming in the first place, and convey detailed information on an exact stimulus. Neither a rate nor precise spike timing are required for effective coding, which is also highly efficient. In terms of information theory, one could say that a single bit is sufficient to confirm a complex string of expected information.

If, however, an unlabeled (i.e., unprimed) line is activated, then the neuronal system will respond using rate-based encoding, allowing detailed communication of the exact features not expected, albeit with two costs: increased communication time (due to use of rates and integration), and higher metabolic activity in support of tonic-mode firing.

In awake animals, principal neurons of the cerebral cortex are often observed to be in what is called a high conductance state (Destexhe et al., 2003). In this state, neurons display a sustained depolarized membrane potential and irregular, desynchronized firing activity. In the model presented below, the primed state mentioned in the prior paragraph is made possible by such a high conductance state.

### **Dynamically Labeled Predictive Coding**

As enabled by the M-current, labile switching between rate-based and dynamically negotiated labeled line encoding allows for a new, non-subtractive, non-Bayesian implementation of predictive coding. I will call this *Dynamically Labeled Predictive Coding* (DLPC), for ease of distinguishing it from the non-labile predictive coding implementations previously discussed by Friston and others.

DLPC is a novel reinterpretation and implementation of the predictive coding hypothesis. DLPC uses M-currents to achieve labile encoding. Notably, a pyramidal neuron's ability to switch from tonic to phasic mode provides a key basis for a circuit's ability to switch from rate-based encoding to dynamically negotiated labeled line encoding. Cholinergic signals

that are sufficient to inhibit M-currents can quickly return neurons to tonic firing mode, and thereby enable those neurons' embedding circuit(s) to return to rate-based encoding.

Predictive coding is based on the idea that codes are changed when comparison mechanisms detect differences between predicted and actual inputs. The mode shift in DLPC offers several benefits, most notably that there is no requirement that the channels for information processing be distinct from the channels constituting the comparison mechanism. This makes the DLPC circuitry more efficient than that described in other proposed neural implementations of predictive coding (e.g., Wacogne et al., 2012)..

Each area classifies patterns in its inputs from other areas, and each area tends to feature reciprocal connections feeding back to the areas that provided its inputs. These feedback connections tend to originate in the lower layers of a neocortical area and terminate in the upper layers of the areas acting as its input providers (Barbas & Rempel-Clower, 1997). Hebbian learning ensures the reinforcement of only those feedback projections connecting an active representation in one area with the representation which will be active in the other area immediately following (modulo any conduction delay, or possible delay via cerebellar mechanisms). By terminating on layer 2/3 neurons in their target area, these selected feedback connections then serve to convey a prediction about what lower-level representations can be expected to follow the activation of a higher-level representation.

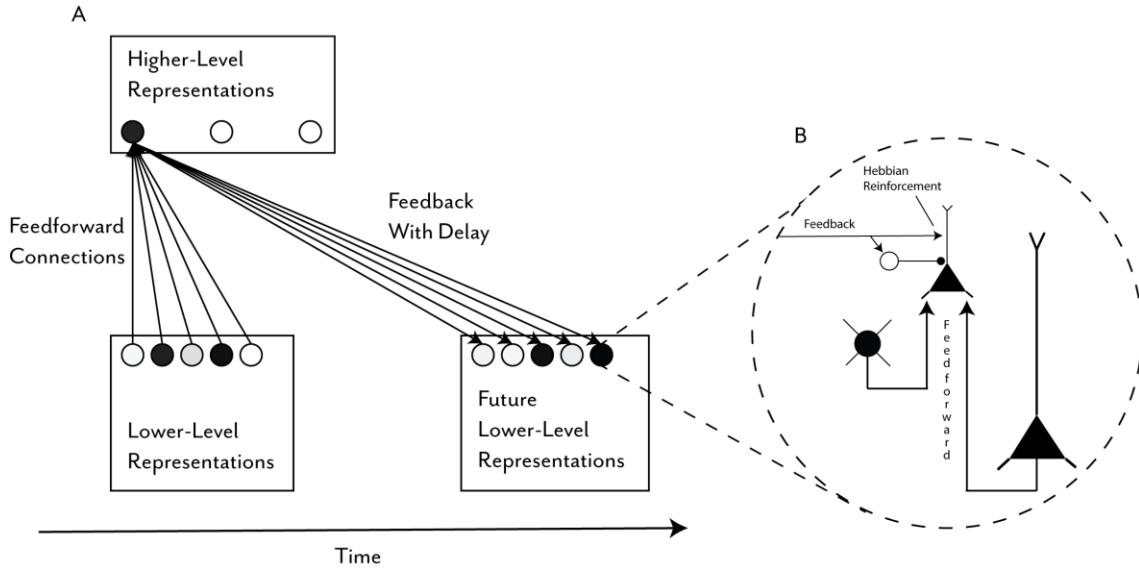


Figure 1.3: Representation activation and circuitry (inset) mediating cortical predictions. Note that a single representation being active does not necessarily correspond to a single cell being active, and many cells/columns might instantiate a particular representation via anything from vector encoding to Hebbian cell assemblies.

Figure 1.3 depicts how lower-level representations drive a higher-level representation via feedforward connections. The reciprocal connections (with delay) then project to the next set of lower-level representations to be active. In the case of a match between active feedback and lower-level activity, the neural circuitry proposed in chapters 3 and 4 provides the ideal conditions for Hebbian learning to occur, reinforcing the feedback connections which have matched, and which are effectively predicting the next pattern of lower-level representations. It should be mentioned that while this is activating a single representation in the higher level, this does not necessarily correspond to a single cell (i.e. grandmother cell).

The linked hypotheses that constitute DLPC provide a framework for explaining a number of apparently paradoxical, or at least counter-intuitive, results, and suggest common

underlying mechanisms for prediction and attentional control. They also highlight ways in which rarely examined ideas about coding have led to measurement practices that may have missed key processes by which the brain transfers and learns from information.

The specific applications of this novel hypothesis will be treated in depth in the following chapters.

Chapter two models hippocampal prediction and mismatch, using the DLPC framework. Chapter three presents extensions to the model that allow its application in modeling neocortical EEG genesis. Simulations of this extended model illustrate how dynamic encoding shifts can produce Mismatch Negativity (MMN) phenomena, including pharmacological effects on MMN reported for humans or animals.

Chapters four and five describe new modeling studies of possible neural bases for alpha-induced information suppression, a phenomenon associated with active ignoring of stimuli. Two models explore the hypothesis that in simple rate-based circuits, information suppression might be a robust effect of neural saturation states arising near peaks of resonant alpha oscillations. A new proposal is also introduced for how the basal ganglia may control onset and offset of alpha-induced information suppression. Although these rate models could reproduce many experimental findings, they fell short of reproducing a key electrophysiological finding: phase-dependent reduction in spiking activity correlated with power in the alpha frequency band.

Therefore, chapter five also specifies how a DLPC model, adapted from the neocortical model developed in chapter three, can provide an expectation-based model of alpha-induced information suppression that exhibits phase-dependent spike reduction during alpha-band oscillations. The model thus can explain experimental findings that were not reproduced by the rate models.

The final chapter summarizes main theses, results, and basic research implications, then suggests future directions, including extended models of neocortical mismatch, applications to artificial neural networks, and the introduction of reward circuitry.

## CHAPTER TWO: HIPPOCAMPAL MISMATCH DETECTION

A significant body of evidence has accumulated to support the hypothesis (e.g., Hasselmo et al., 1995; Lorincz and Buzsaki, 2000; Lisman and Otmakhova, 2001; Vinogradova, 2001; Meeter et al., 2004; Lisman and Grace, 2005; Kumaran and Maguire, 2007; Duncan et al., 2009) that the hippocampus includes a comparator, i.e., a neuronal circuit that enables it to compare two signal streams, one carrying predictions/expectations about sensory inputs, the other carrying actual inputs. While these streams match, the hippocampus can continue making predictions based on what was learned from past experiences. When these streams mismatch, the hippocampus generates a response that constitutes a “prediction error signal.” Such signals serve to switch brain circuits into a learning mode that encodes the novel, unpredicted information, and stores it for use in future predictions.

### **The Functional Anatomy of CA1 as a Comparator**

Within the hippocampus, area CA1 exhibits both the anatomical connectivity and the signaling that would be expected of a comparator that serves as a mismatch detector, i.e. a comparator that sends a positive signal for a mismatch but little or no signal for a match. The anatomical connectivity and signal convergence that characterize CA1 are summarized in Figure 2.1 (adapted from Lisman & Grace, 2005). Notably, the dentate gyrus (DG) and CA3 zones of the hippocampus receive sensory streams from layer 2 of entorhinal cortex (EC2), whereas area CA1 receives sensory streams from layer 3 of entorhinal cortex (EC3). Then, sparse recoding and recurrence within DG-CA3 enable context-dependent associative retrievals of past experiences to inform predictions about upcoming sensory

input. These predictions are conveyed to CA1 via the synaptic weights of the Schaffer collaterals, i.e., those collaterals of CA3 fibers that project to CA1. These weights effectively store associations between current CA3 states and formerly present CA1 states. CA1 can therefore perform a comparison between the predictions arriving from CA3 and actual sensory inputs from EC3. Physiological studies strongly support this view. For example, in human fMRI studies, there is much greater CA1 activation during an associative mismatch than during a match (Kumaran and Maguire, 2006). Likewise, CA1 shows strong activity to unexpected inputs but habituates rapidly as inputs become predictable, but this habituation vanishes if CA3 input to CA1 via Schaffer collaterals is eliminated (Vinogradova, 2001) while EC3 input to CA1 is preserved. Such evidence (see reviews in Duncan et al. 2009, and Lisman & Grace, 2005; but also reports by Buzsáki et al. 1979, Vinogradova, 1984, Otto and Eichenbaum, 1992, Knight, 1996, Dolan and Fletcher, 1997) indicates that it is the learned predictions from CA3 that are causing a reduction in CA1 activity as the associations form and increase in certainty.

In summary, hippocampal processing in the CA1 zone is now widely considered to be exemplary of a computation that compares memory-dependent sensory predictions/expectations with actual sensory inputs. Whereas comparators can be built to respond more in the match, or more in the mismatch, the comparator in CA1 serves as a mismatch detector, in the sense that it habituates as learning-updated expectations come to match inputs, but sends a strong signal whenever inputs mismatch expectations.

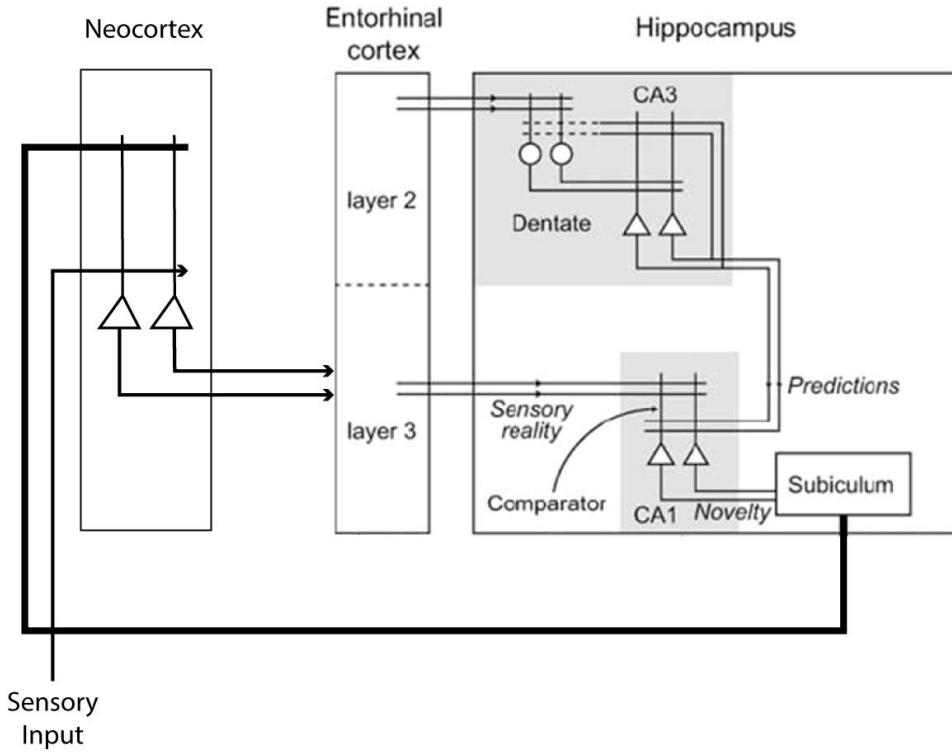


Figure 2.1: Hippocampal mismatch circuitry, adapted from Lisman, (1999). Recurrent connectivity between the dentate gyrus and CA3 helps to create representations of current inputs that are sensitive to the recent stream of inputs. The predictions represented by synaptic weights between CA3 and CA1 can therefore be highly context dependent.

### Existing CA1 Comparator Models

To date, there are no fully implemented computational models of CA1 Mismatch Detection. Katz et al. (2007) presents a model of CA1 function. Although this model does not approach the hippocampus from a framework of episodic memory and sensory mismatch – instead it deals in context and spatial information – the model can arguably be reinterpreted, without requiring any change in the simulation or their results. Effectively, CA1 place cells can be interpreted as comparing an abstract sensory representation of current location (computed by the neocortex from sensory input) with a predicted location

provided as a corollary of CA3 playback of recollected place sequences. Katz et al. note that CA3 pyramidal neurons' axon (Schaffer) collaterals target the proximal zone of CA1 pyramidal neurons' apical dendrites, whereas "perforant pathway" axons from EC3 to CA1 target the distal zones (see Figure 2.1) of CA1 pyramidal neurons' dendrites. This arrangement allows the model to perform a sort of subtractive/inhibitory mismatch detection, because the more proximal inputs (carrying predictions) can, when present, block the postsynaptic effects of the more distal inputs. However, the authors did not propose the model in Katz et al. (2007) under this interpretation, and have not to date promoted it as a model of hippocampal mismatch.

Lisman and Otmakhova (2001) presented a partial computational implementation of the SOCRATIC (Sequences of Condensed Representations, Autocorrected, Theta-Gamma Coded, In Context) model, which detailed the encoding and retrieval of sequences in the CA3 region and the role of CA1 as a decoder and mismatch detector. However, CA1 was not computationally implemented in that paper.

### **A Biologically Detailed Comparator that Uses Pyramidal Mode Switching**

In this section we specify biological bases for a detailed model of a feature-rich CA1 comparator. This model emerged from the simulation of often-overlooked features of CA1 pyramidal neurons, and is based on the ability of these features to dynamically alter neural properties that have traditionally been considered "intrinsic". A clue to the existence of dynamic properties is that so-called intrinsic properties can markedly vary between reports based on slice experiments versus *in vivo* recordings (Prescott et al.,

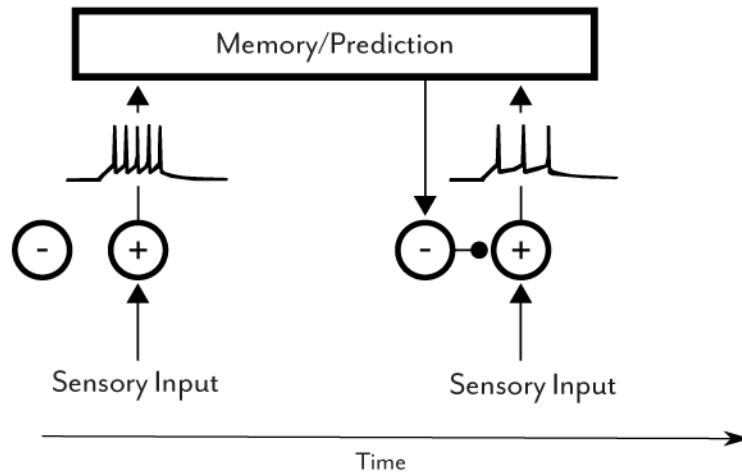
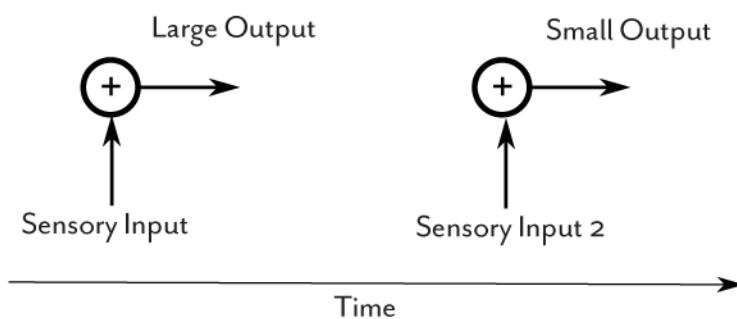
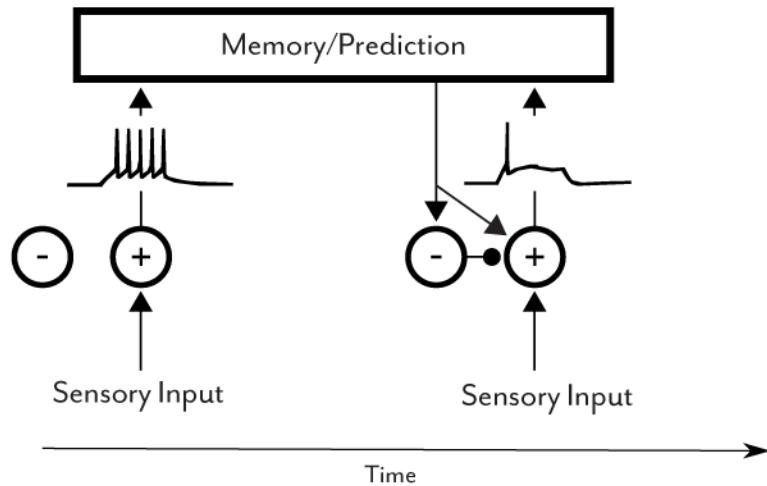
2008), especially when the *in vivo* reports come from awake, behaving, animals rather than from anesthetized animals. Whereas neurons have previously been sorted into types, e.g., excitability classes (per Hodgkin's classification in Hodgkin, 1948) based on "intrinsic properties", such properties may in fact be dynamically determined by external factors, such as afferent neural activity that is absent in slices, or during anesthesia. For example, Wolfart et al. (2005) showed that even afferent noise can determine the transfer function of thalamic neurons, a feature often used for "intrinsic" typing.

Prescott et al. (2008) called out two of these features in particular, specifically the M-current and its interaction with Cl<sup>-</sup> currents. The M-current, which is mediated by KCNQ channels, is an inward potassium current, originally discovered in sympathetic ganglion neurons of bullfrogs by Brown & Adams (1980). It is non-inactivating and voltage-dependent, being activated by depolarization of the membrane above approximately -60mV (Brown & Adams, 1980). Because of this, depolarizations large enough to generate action potentials reliably also activate the M-current. M-currents are so named because KCNQ channels are closed by activation of adjacent Muscarinic ACh receptors. M-currents are found in neurons throughout the nervous system, including pyramidal cells of the cerebral cortex and hippocampus (reviewed in Jentsch, 2000; Marrion, 1997).

The M-current is usually thought of as a rate-modulating current, contributing to slow adaptation in spike rate during prolonged excitation. Under certain circumstances, however, it can act quite differently. Prescott et al. (2006) showed that the M-current

interaction with the chloride leak current can reliably cause a shift from regular (tonic) spiking to onset (phasic) bursting. This was demonstrated both in computational models, similar in structure to our proposed models below, and in dynamic clamp recordings from CA1 hippocampal pyramidal cells. However, these properties were discussed with respect to low frequency resonances, not with respect to CA1 as a comparator.

When in the mode described (equivalently) as "regular firing" or "tonic firing", the cell responds to an input with proportionate output for as long as the input persists. When in the mode described (equivalently) as "phasic firing" or "onset burst firing", the cell responds only transiently, with a single spike or short burst of high-frequency spikes, at the onset of the input, but does not sustain above baseline firing, and may even become quiescent, for the remaining duration of the input. The new model developed and simulated here is based on the hypothesis that the ability of single pyramidal neurons to switch modes enables them to operate as mismatch-detecting comparators. As shown in Figure 2.2, this allows for the construction of comparators using more efficient circuitry than in prior proposals.

**a****b****c**

**Figure 2.2:** A) A standard predictive coding comparator, where a stimulus elicits predictions which inhibit a population, reducing the frequency of its output when the predicted stimulus occurs. B) A standard habituative comparator, which fires at a reduced rate due to afferent habituation when a recently experienced stimulus is repeated. C) A mode-shifting comparator, wherein a stimulus elicits predictions which prime a population, placing it into a high-conductance state and causing a shift from tonic to phasic firing.

This mode shift allows a neural population to dynamically switch from encoding detailed, rate-based information about an unexpected stimulus to quick, efficient phasic reporting of the onset of expected stimuli, and back again. Predicted stimuli do not require detailed information transmission, since a single “byte” of confirmation can tell the rest of the brain to act as though its expectation were confirmed for all practical purposes.

We will show below that mode switching in the model depends on intact NMDA currents. This allows the model to be used to explain data that indicate that NMDA blockade disrupts mismatch computations in some common experimental paradigms (Farley, Quirk, Doherty, & Christian, 2010; Kreitschmann-Andermahr et al., 2001; Leung, Croft, Baldeweg, & Nathan, 2007). Furthermore, given the ubiquity of M-currents and KCNQ channels in CNS neurons, it can be expected that aspects of the model should be applicable to non-hippocampal comparators and even beyond mismatch detection per se. Recently, dysfunctions of KCNQ channels, which come in several variants, have been implicated in transient drug-induced psychoses as well as in chronic syndromes including schizophrenia (e.g., Soty et al. 2009), depression (e.g., Friedman et al., 2016), and epilepsy (e.g., Gunthorpe et al, 2012). In addition, novel pharmaceuticals that act via KCNQ channels are being explored for possible therapeutic values. Therefore, we report several simulations below that further illustrate the new comparator model’s operation, and illustrate its compatibility with recent reports regarding drug effects on KCNQ functions.

## Materials and Methods

The proposed model's biophysical and neuroanatomical details, including strongly nonlinear interactions between the excitation-driven habituation and inhibition-driven subtraction, allow it to address limitations found in the alternative comparator models, with a much greater economy of circuitry, as noted in the caption of Figure 2.2. The model was constructed by combining the range of receptors and channels included in prior models of pyramidal cells (Prescott et al., 2006) and synapses (e.g. Brunel & Wang, 2001). However, the model also includes further receptors and interactions needed to realistically model ACh-dependent M-currents, which play a fundamental role in our treatment of cortical comparison processes. In particular, a voltage-dependent M-current was added, based on the mathematical modeling in Mainen & Sejnowski (1996). We use simulations to assess and document the model's explanatory power vis-à-vis existing data from CAI mismatch experiments and current theories of CAI comparator function (such as Lisman's SOCRATIC model).

Our simulations can be grouped into two major investigations. The first demonstrates the ability of a simulated CAI pyramidal cell to differentiate between expected and unexpected input. The second is a pharmacological investigation designed to offer predictions regarding learning and disordered processing associated with psychosis. In this case, we simulate the effects of NMDA and KCNQ blockade on the neuron, then demonstrate the ability of the model to capture the effect of a KCNQ agonist, such as retigabine, following

NMDA blockade. In the model, it can rescue the cell's ability to mode switch, and thus distinguish matches from mismatches.

### *Proposed Model*

In documenting our proposed model here, we employ the standards recommended by Nordlie, Gewaltig, & Plesser (2009), which proposed that the specification of computational neural models be broken down into the architecture, network connectivity, neural and synapse models used, and the input to and measured output from the simulated system.

### *Architecture*

The architecture is extremely simple. A single neuron representing a CA1 pyramidal cell performing sensory comparison is the core of the model, as it is a single-neuron comparator.

### *Network Connectivity*

While there are a high number of synaptic connections in our model, they form a very simple pattern. The sensory comparator is innervated by 600 synapses: 200 driving synapses (presumed to be from layer 3 entorhinal cortex (EC3) pyramidal neurons), 200 basal priming synapses (presumed to be from CA3 neurons), and 200 basal inhibitory priming synapses (presumed to be from priming-driven feed-forward inhibitory interneurons within CA1). All synapses have weights that are randomly selected using a Gaussian distribution with  $\mu$  of 1.0 and  $\sigma$  of 0.1. The weights are of the form  $w \cdot \text{gauss}(\mu, \sigma)$ , where the value of  $w$  is 16 for driving synapses, 12 for excitatory priming

synapses, and 8 for inhibitory priming synapses. Whether these synapses were active or not was directly controlled in the simulation: neither EC3 pyramids, CA3 pyramids, nor CA1 inhibitory interneurons, as such, were simulated.

### *Membrane and Synapse Equations*

Each comparator neuron is made up of two, coupled, point models, one representing the apical portion of the cell, and the other the basal. The membrane voltage  $V$  of each point model is conductance-based, following the general convention of

$$C \frac{dV}{dt} = I_1 + I_2 + \dots + I_n$$

where  $I$  is a current of the form  $g_{max} \cdot g_{proportion} \cdot (V_E - V_m)$ , where  $g_{max}$  is the maximum conductance of the channels carrying that current,  $g_{proportion}$  is the proportion (from 0 to 1) of the max conductance currently engaged (or the proportion of membrane channels currently open and non-inactivated), and  $V_E - V_m$  is the driving voltage differential between the Nernst equilibrium of that particular current's ion ( $V_E$ ) and the current membrane voltage ( $V_m$ ). These values, and their published sources, are specified for each channel type in Table 2.1.

In our model, the complete membrane equation is

$$C \frac{dV}{dt} = I_{Na} + I_{AMPA} + I_{NMDA} + I_{GABA} + I_K + I_M + I_{Cl^-}$$

The currents involved are as follows, each with their own  $g_{proportion}$  function:

- NMDA (Na)

$$g_{proportion} = \frac{g_{open}}{1 + [Mg^{2+}] \exp \frac{-0.062V_m}{3.57}}$$

$$\frac{dg_{open}}{dt} = \sum_{j=1}^{C_e} (w_j \cdot s_j^{NMDA})$$

$$\frac{ds_j^{NMDA}}{dt} = -\frac{s_j^{NMDA}}{\tau_{NMDA}^{decay}} + \alpha x_j (1 - s_j^{NMDA})$$

$$\frac{dx_j}{dt} = -\frac{x_j}{\tau_{NMDA}^{rise}} + \sum_k \delta(t - t_j^k)$$

- AMPA (Na)

$$\frac{dg_{proportion}}{dt} = \sum_{j=1}^{C_e} (w_j \cdot s_j^{AMPA})$$

$$\frac{ds_j^{AMPA}}{dt} = -\frac{s_j^{AMPA}}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k)$$

- GABA (Cl)

$$\frac{dg_{proportion}}{dt} = \sum_{j=1}^{C_i} (w_j \cdot s_j^{GABA})$$

$$\frac{ds_j^{GABA}}{dt} = -\frac{s_j^{GABA}}{\tau_{GABA}} + \sum_k \delta(t - t_j^k)$$

- Inactivating Voltage-Based Na

$$g_{proportion} = 0.5(1 + \tanh((V_m - V_1)/V_2)) \cdot (1 - h)$$

$$\alpha_h = 0.07 \exp\left(\frac{V_m}{20}\right)$$

$$\beta_n = \frac{1}{\exp\left(\frac{V_m + 30}{10}\right) + 1}$$

$$\frac{dh}{dt} = \alpha_h(1 - h) - \beta_n h$$

- Potassium

$$\frac{dg_{proportion}}{dt} = \phi(w_{inf} - g_{percent}) / \tau_w$$

$$w_{inf} = 0.5(1 + \tanh((V_m - \beta_w)/\gamma_w))$$

$$\tau_w = \frac{1}{\cosh\left(\frac{V_m - \beta_w}{2 \cdot \gamma_w}\right)}$$

- Cl Leak Current

$$g_{proportion} = 1$$

- M-current, from (Mainen & Sejnowski, 1996)

$$\frac{dg_{proportion}}{dt} = \frac{\frac{1}{1 + exp(\frac{\beta_z - V_m}{\gamma_z}) - g_{percent}}}{tau_z}$$

$$tau_z = \frac{tau_{z_{peak}}}{3.3exp(\frac{V_m - \beta_z}{20}) + exp(\frac{-\beta_z}{20})}$$

$$tau_{z_{peak}} = \frac{1000}{2.e^{(\frac{T-36}{10})}}$$

And the parameters used in our baseline simulations are as follows:

Table 2.1: Model Parameters

Parameter	Value	Citation
$C$	$2 \frac{\mu F}{cm^2}$	(Prescott, Ratté, De Koninck, & Sejnowski, 2006)
$g_{Na}^{max}$	$20 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$V_{Na}$	$50mV$	(Prescott et al., 2006)
$V_1$	$-1.2mV$	(Prescott et al., 2006)
$V_2$	$23mV$	(Prescott et al., 2006)
$w$	$28$	(Fernandez, Mehaffey, Turner, & Fernando, 2005)
$V_C$	$-64mV$	(Fernandez et al., 2005)
$A$	$232$	(Fernandez et al., 2005)
$g_K^{max}$	$20 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$V_K$	$-100mV$	(Prescott et al., 2006)
$\phi$	$0.15$	(Prescott et al., 2006)
$\beta_w$	$-2mV$	(Prescott et al., 2006)
$\gamma_w$	$21mV$	(Prescott et al., 2006)
$g_M^{max}$	$4 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$g_{AHP}^{max}$	$1 \frac{mS}{cm^2}$	(Prescott et al., 2006)

$\beta_z$	-35mV	(Prescott et al., 2006)
$\gamma_z$	10mV	(Yamada, Koch, & Adams, 1989)
$T$	36°C	(Yamada et al., 1989)
$\alpha_{AHP}$	0.005	(Prescott et al., 2006)
$\beta_{AHP}$	0mV	(Prescott et al., 2006)
$\gamma_{AHP}$	5mV	(Prescott et al., 2006)
$V_{shunt}$	-70mV	(Prescott et al., 2006)
$g_{shunt}$	$2.0 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$Mg^{2+}$	$1e-3$	(Brunel & Wang, 2001)
$g_{AMPA}^{max}$	$7.5e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$g_{GABA}^{max}$	$7.5e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$g_{NMDA}^{max}$	$2e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$\tau_{AMPA}$	2ms	(Brunel & Wang, 2001)
$\tau_{GABA}$	10ms	(Brunel & Wang, 2001)
$\tau_{NMDA}^{rise}$	2ms	(Brunel & Wang, 2001)
$\tau_{NMDA}^{decay}$	100ms	(Brunel & Wang, 2001)
$\alpha_{NMDA}$	0.5ms	(Brunel & Wang, 2001)

The bridge current between the apical and basal portions of the neuron is computed based on their relative voltages, the axial resistance of the apical dendrite, and the distance between the apex and the base:

$$I_{bridge} = \frac{V_{apex} - V_{base}}{dr_{axial}}$$

In the above equation  $d$  represents the distance between the apex and base (i.e. the height of the pyramidal cell) in cm,  $V$  is the voltage at one of the two points, and  $r_{axial}$  is the axial resistance. In our model CA1 pyramidal cells,  $d = 0.0003$  and  $r_{axial} = 200\Omega cm$ .

### *Input*

Each of the hundreds of input synapses is driven by a Poisson process whose lambda variable changes over time, going from 0.0001 at rest to 0.01 during activation. Activation lambda values and onset/offset timing are specific to each population of inputs, and to whether or not we were simulating an expected (primed) stimulus, or an unexpected (unprimed) stimulus.

Regardless of expectations, driving inputs (presumed to be from EC3) are activated at 350 ms into the simulation, with lambda switching to a value of 0.01 for a duration of 350 ms (stimulus duration), whereupon the driving inputs returned to the resting lambda. During primed stimulation, the excitatory priming inputs are activated at 150 ms, switching to a lambda value of 0.01 for a duration of 550 ms, returning to a resting lambda (0.0001) at the same time as the driving inputs.

During primed stimulation, the inhibitory (I) priming inputs are activated at 175 ms, also at a lambda value of 0.01, and return to a resting lambda at the same time as the excitatory (E) priming inputs. During unprimed stimulation, the priming inputs (both E and I) will not activate at all, remaining at rest lambda throughout the entire simulation.

### *Output*

The membrane voltage and extracellular current were measured for analysis at both the apex and the base of all neurons, along with the net KCNQ-mediated M-current.

### *Pharmacological Manipulations*

Each of our pharmacological manipulations was set to produce either a 25% increase (for agonists) or a 25% decrease (for antagonists) of the maximum value of an affected channel's conductance. To examine the impact of various pharmacological agents, the basic match/mismatch simulations were performed under the following conditions:

- NMDA activity decrease (e.g. via PCP; (Anis, Berry, Burton, & Lodge, 1983))
- KCNQ activity decrease (e.g. via high doses of a Cholinesterase inhibitor; (Riekkinen et al., 1997))
- KCNQ activity increase (e.g. via retigabine; (Soty et al., 2009))
- NMDA activity decrease and KCNQ activity increase combined

### *Data Analysis*

Spike detection was performed on the simulated membrane voltage data by recording, as a spike time, each crossing from below -10 mV to above 10 mV.

Each neuron's sequence of spikes was then classified as phasic or tonic by dividing the time up into two segments. The first, the onset segment, consisted of the first 100ms after the onset of driving (not priming) synaptic activity. The second, sustained segment consisted of the time between the end of the onset segment and the offset of driving activity. The rate

in spikes/second was calculated for each segment, and if the onset rate was at least three times the sustained rate, as well as the sustained rate being below 5 spikes/second, the neuron's firing pattern was classified as phasic. Otherwise, the neuron's response was classified as tonic.

## Results

### Investigation 1: Match/Mismatch

The results confirmed our prediction that a realistic model of CA1 neurons can perform comparison, and mismatch detection, between a priming input (presumed to be from CA3) and a driving input (presumed to be from EC3). When the two inputs coincided, a mode switch was observed and the neural response became phasic, as seen in Figure 2.3. This results from the M-current in the unprimed case (as seen in Figure 2.3c's red trace) having a lower magnitude of deviation from baseline (zero) as indicated by a higher plateau than that of the primed case (Figure 2.3c's blue trace, which shows a larger deviation below baseline). This unprimed plateau is not negative enough to prevent continuous spiking, and thus allows the tonic response to unprimed stimulation. The primed plateau is sufficiently negative to prevent sustained spiking; however, a high-frequency burst of spikes is seen before the plateau is reached.

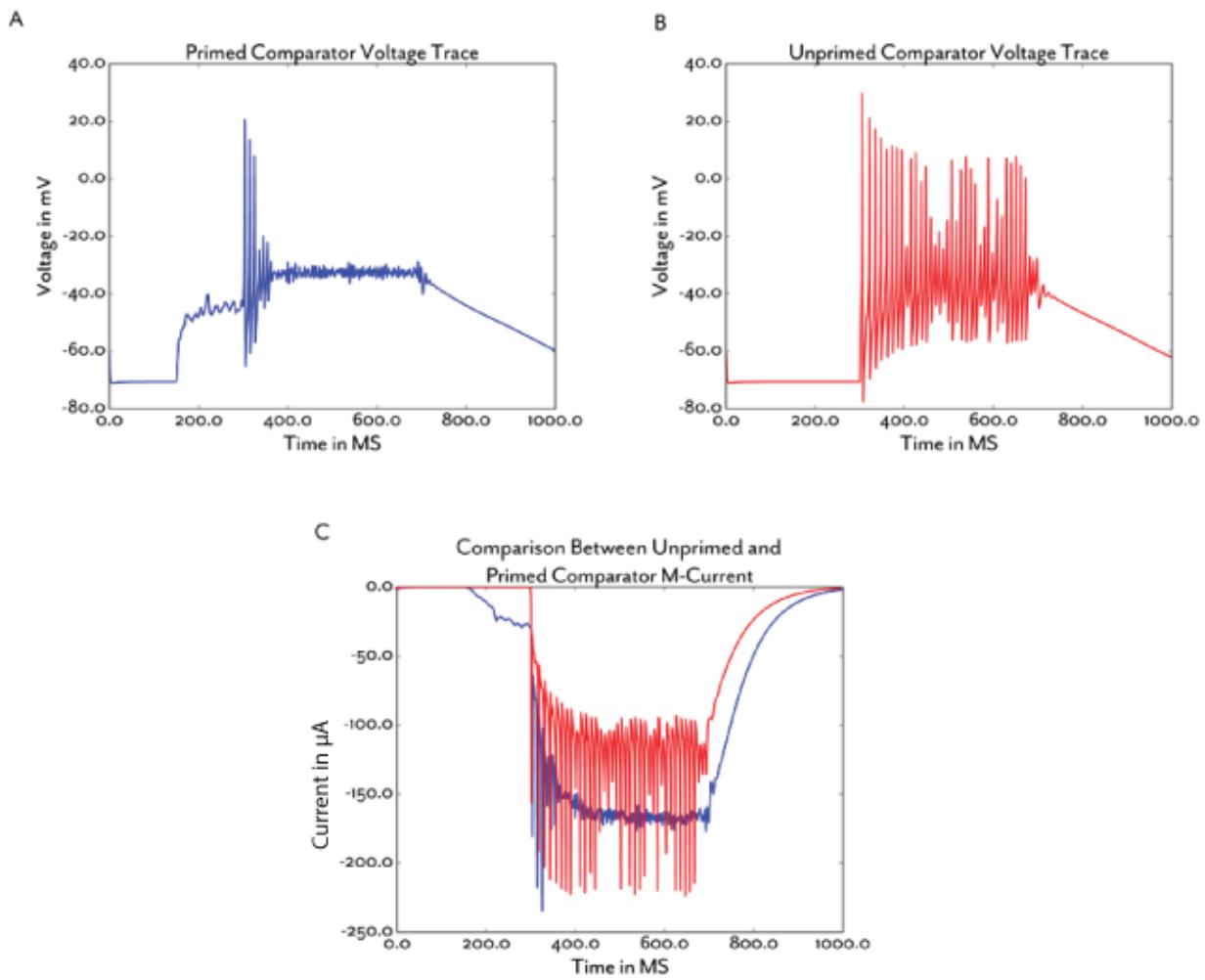
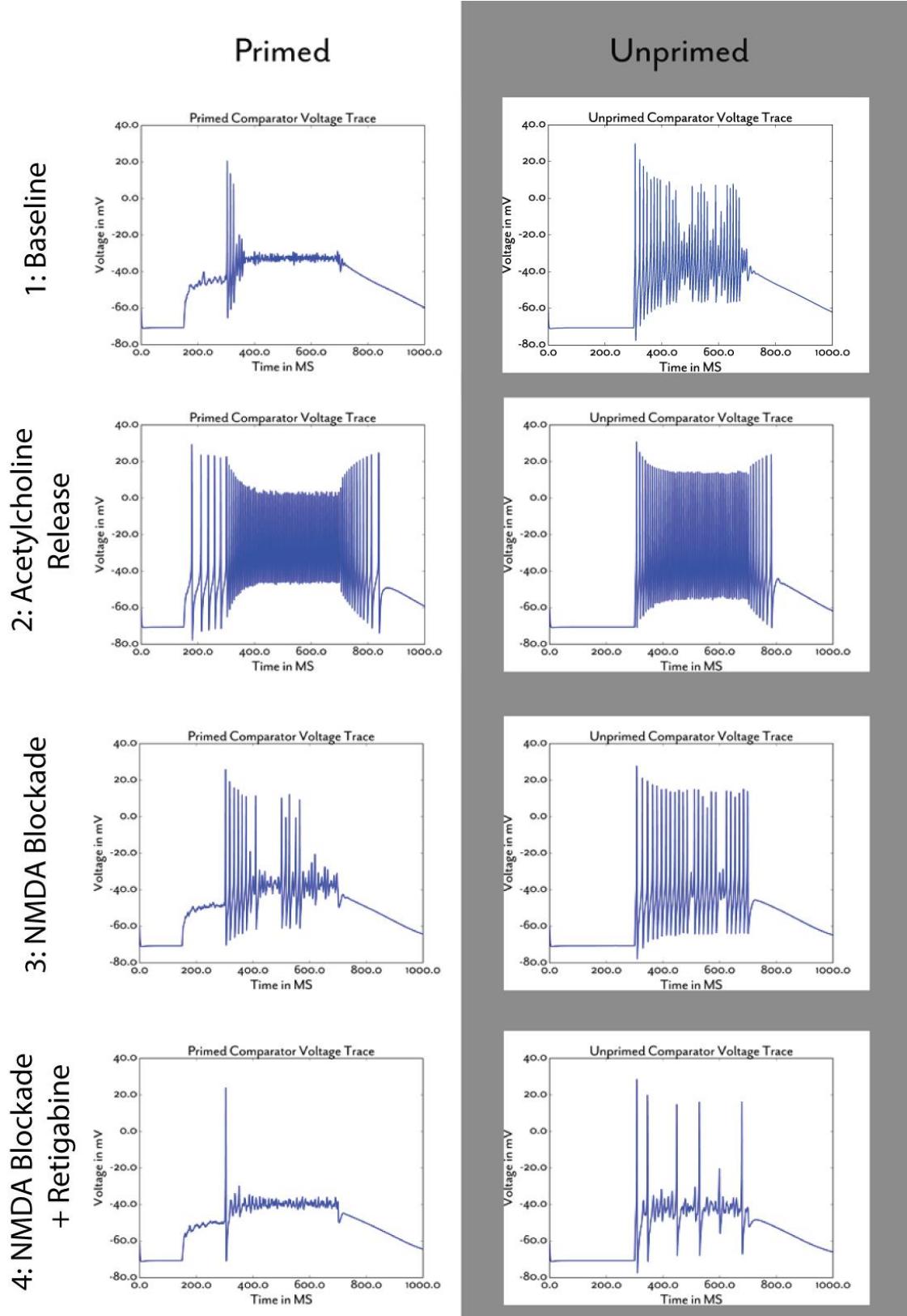


Figure 2.3: A and B) Simulated voltage trace (in millivolts membrane potential) of primed (blue) and unprimed (red) conditions. C) Net M-current (in micro-amps) observed in primed (blue) and unprimed (red) conditions. Panel C clearly shows a greater sustained amplitude of M-current in the primed condition, which prevents the simulated neuron from firing.

## Investigation 2: Pharmacological Manipulations

### NMDAR Blockade

As can be seen by comparing Figure 2.4, row 3 with row 1, reducing NMDA max conductance by 25% prevents the mode switch to a phasic response. This mode switch is required for the system to reveal that it has learned an expectation and has used it via priming to compute a match. NMDA blockade would prevent match computation in any circuit using a comparator as modeled here, so it would appear in an experiment that all stimuli are treated as mismatches. This might lead an experimenter to infer that this is because NMDA blockade blocks the learning of any new expectations. Though such might also be pertinent *in vivo*, note that the result shown here is entirely due to an effect of NMDA currents on neuronal responses to weighted synaptic inputs that would be sufficient to serve the priming function (as in row 1, left column) in the absence of the NMDA blockade. Hence the effect depicted does not depend on well-known roles of NMDA currents in plasticity mediated by long-term synaptic strength changes. Rather, the mechanism by which partial NMDA blockade prevents the mode-switch is a reduction in sodium conductance that would otherwise be produced by excitation of NMDARs on the modeled CA1 pyramidal cell.



**Figure 2.4: Primed and unprimed comparator responses to pharmacological manipulation.** The panels on the left represent simulated membrane voltage (in mV) responses to primed (aka expected) stimuli, and the panels on the right represent the corresponding unprimed responses. From top to bottom, the rows represent responses under a baseline condition, under 25% reduction in KCNQ conductivity (simulating elevated ACh), under 25% reduction in NMDA conductivity (simulating an NMDA antagonist like PCP), and under a combination of 25% reduction in NMDA conductivity and a 25% increase in KCNQ conductivity (simulating the ability of retigabine to reduce the symptoms of NMDA blockades like PCP).

### ACh Agonists

Increasing ACh sufficiently to reduce the M-current's maximum conductance by 25% is another way to prevent the mode switch to a phasic response, as can be seen by comparing Figure 2.4, row 2 with row 1. Like the results under NMDA blockade, the tonic spiking in both primed and unprimed conditions indicates that a real neural circuit described by our comparator model would not exhibit match-detection, thus treating all stimuli as mismatches. The spike rate is drastically increased in our simulation, but a more reasonable spike rate might be achieved with more nuanced modulation of maximum KCNQ conductance. Again, this effect is achieved with no change in synaptic weights, and thus represents a direct effect of the simulated ACh boost, not an ACh-dependent learning effect.

### NMDA antagonism combined with retigabine rescues comparator function

In line with the findings of Soty et al. (2009), simulation of a retigabine effect taking the form of a 25% increase in maximum KCNQ conductance rescues the comparator functionality from a 25% NMDA antagonism, such that the mode switch is once again observed under coincidental priming inputs (from CA3) and driving inputs (from EC3).

This would allow for restoration of CAI discrimination between expected and unexpected experiences, and thus may point to one mechanism by which retigabine has anti-psychotic effects.

## Discussion

These results indicate that a realistic CAI neuron model can simulate a number of phenomena known from the literature on hippocampal function, and specifically data that provide evidence that CAI serves as a mismatch detector. The mode switch allows a brief phasic response to an expected stimulus to be supplanted by a sustained tonic response to an unexpected (novel) one. We hypothesize that this bears a direct relation to reports (Kumaran and Maguire, 2006) that the same CAI region exhibits both small CAI fMRI signals to expected events and huge fMRI signals to unexpected events. Both phasic and tonic modes of CAI pyramids were also simulated in Prescott et al. (2006), and in Prescott et al. (2008), but they were not discussed in the context of mismatch detection. In the current, more complete, model, additions allowed simulation of a number of additional effects, notably pharmacological effects that can eliminate either the match (phasic mode) response or the mismatch (tonic mode) response. These simulations revealed that dysfunctions can arise from either too much or too little ACh stimulation of muscarinic receptors that control KCNQ channels. The simulations also revealed that a dysfunction caused by administration of an NMDA antagonist could be rescued by simultaneous administration of a KCNQ channel agonist, such as retigabine. Finally, simulations revealed that a primed CAI neuron's response to a driving input has a lower latency (is faster) than an unprimed CAI neuron's response to the same driving input. This conforms

to the general observation that priming confers the advantage of faster responses to expected events.

*Limitations of the model.* A truly complete model, and understanding, of CAI pyramidal neurons remains a task for the future. For example, Prescott and Sejnowski (2008) notes that M-currents and AHP currents impact the two modes differently, offering important benefits that would improve reliability of each mode by itself, but might interfere with the other mode. For instance, the M-current can provide for precise onset responses below a certain mode-switch input level, and beyond that level (when intensity overrides the expectations), whereas AHP introduces some reduction in reliability at a given firing rate, but also increases the range of inputs over which reliability remained high (Prescott and Sejnowski, 2008). Some of these effects, and new ones, have been further explored in subsequent models, but, again, without regard to the issue of mismatch detection.

*Potential applications of the comparator model.* It is important to note that this model is above all a model of mode switching by a pyramidal neuron. Therefore, the results may be generalizable to many pyramidal neurons that exist outside CAI, e.g. in various parts of neocortex. Indeed, although CAI mismatch signals are unmistakable in fMRI studies, evidence has long existed from other types of studies, notably EEG/ERP studies, that many areas of cortex react much differently to matches than to mismatches of expectations. Indeed, one of the best studied phenomena in cognitive neuroscience is the mismatch negativity, i.e. the MMN, which emerges when averaged ERPs to expected stimuli are subtracted from averaged ERPs to stimuli that violate expectations. The present results may

help explain such phenomena, either because hippocampal predictions are sent to neocortical areas (not just to CA1), or because a priming relationship exists between any higher order cortical area and the lower order cortical areas to which it projects (Barbas and Rempel-Clower, 1997; Markov et al., 2014), and because mismatch detection is needed to guide learning in all such instances, at least at some phases of development. Indeed, such an assumption is fundamental in predictive coding models of the cerebral cortex (see recent reviews in Shipp, 2016 and Spratling, 2017). In such contexts, as earlier noted in Figure 2.3, our model offers a more efficient neural architecture than alternatives (e.g., Wacongne et al. 2012).

## **CHAPTER THREE: A BIOLOGICALLY REALISTIC SINGLE-CELL COMPARATOR MODEL OF THE MISMATCH NEGATIVITY**

### **Abstract**

Leading models of the well-established ERP phenomenon known as mismatch negativity (MMN) have advanced two partially conflicting interpretations: predictive coding and afferent habituation. Wacongne et al. (2012) uses top-down inhibition, based on a hierarchy of predictions, to reduce the response to sensory inputs, whereas May & Tiitinen (2010) relies on afferent habituation to reduce the driving input to neurons. Despite the clear differences between these approaches, both of the alternative processes are subtractive,

in that they feature an approximately linear reduction of activity in response to graded effects of previous events (either via predictions or via habituation).

To overcome limitations of prior models that rely on such subtractive mechanisms, we propose a new class of predictive coding model, based on the highly nonlinear properties of the principal neurons of the cerebral cortex, pyramidal neurons. These properties, which depend on the inclusion of realistic KCNQ channels and M-currents, provide a robust mechanism by which top-down priming creates conditions that enable single neurons to compare inputs and expectations. The same properties enable on-the-fly switching between two distinct firing modes and linked encoding schemes. One mode/scheme rapidly flags and reacts to confirmed expectations, whereas the other mode/scheme conveys more detailed information about an unexpected stimulus.

Importantly, the same neurons that perform comparisons mediate bottom-up sensory processing, so the model overcomes a key shortcoming of prior models such as Wacongne et al. (2012), which required extra, parallel circuits devoted solely to computing prediction errors. When KCNQ channels are activated by a top-down subthreshold depolarization, the induced M-current switches the state of the cell membrane. When the neuron is then driven by a sufficient bottom-up input to cause suprathreshold depolarization, the expectation-induced state enables the neuron to exhibit a phasic burst at the onset of bottom-up input, followed by depolarized quiescence. This response is radically different from the tonic firing that the neuron would exhibit in the absence of the M-current induced by top-down priming.

After we demonstrate that such model neurons can serve as comparators capable of differentiating between match and mismatch events, we then embed them in a model local circuit and show that it can produce signal deflections of the two types that underlie MMN phenomena. We also show that the new model, unlike linear-subtractive models, allows downstream neurons to classify different levels of bottom-up input regardless of the level of expectation. Because of its biophysical realism, the model can simulate known effects of pharmacological manipulations, including NMDA blockade and the abolition of the MMN in PCP-based animal models of schizophrenia. In doing so, our model becomes the first to predict the possible dissociation of the scalp-measured MMN from underlying neural comparison processes. It also makes new predictions regarding: classes of antipsychotics that can rescue the MMN; how strong modulation of muscarinic acetylcholine receptors will disrupt the MMN; and an observable dynamic shift in the encoding of information based on top-down representation of expectations.

## Introduction

To evaluate and revise its learned internal models, the brain must compare expectations to reality and generate mismatch signals if sensory input deviates from expectations. Many of these mismatch signals are associated with learning and redirections of attention. We focus on one phenomenon from studies of such attention-redirecting mismatch signals, the mismatch negativity, or MMN, and then propose a novel model of the comparison mechanism underlying MMN phenomena. In this section, we first provide background on the MMN, its origins, the experimental and pharmacological properties that will be used to

constrain the proposed model, as well as the state of existing models. Then we review the cellular mechanisms that we propose interact to produce the dynamic coding shifts that mediate distinctive processing of expected versus expectation-violating stimuli.

### *The MMN*

The MMN (see Figure 3.1) is a deflection found in a difference wave taken between two electroencephalogram (EEG) or magnetoencephalogram (MEG) measurements, and is believed to capture the processing of a cortical mismatch event (Winkler, 2007). The averaged response to many stimulus events demonstrates deflections or event related potentials (ERPs) that wax and wane over many hundreds of milliseconds after event-onset. Expected events induce different ERP waveforms than events that violate expectations. When the waveforms of expected events are subtracted from those of events which feature the same stimulus but which violate expectations, a prominent difference has been observed in EEG (Garrido et al., 2009) and MEG (Yabe et al., 1998) traces at about 100–200 ms post-event. This difference is called the MMN.

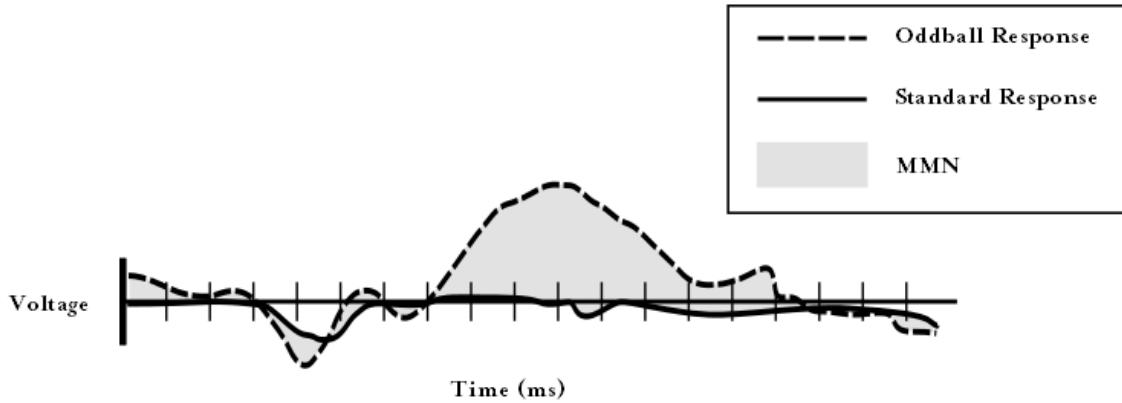
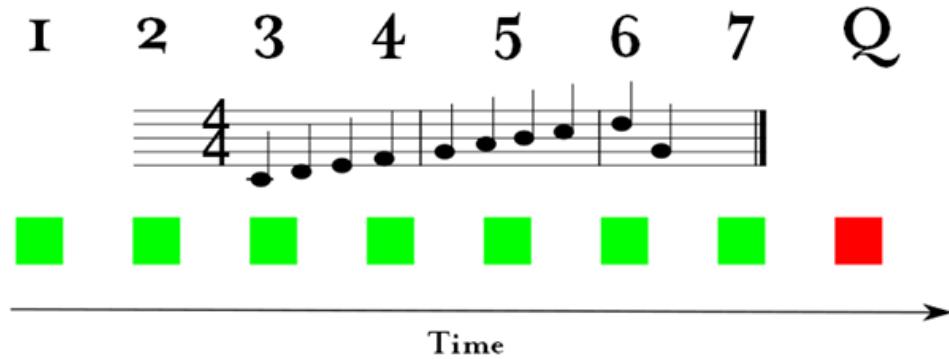


Figure 3.1: A cartoon example of a mismatch negativity. The dashed line represents the average electric field generated in response to oddball stimuli, the solid line represents the average electric field generated in response to standards, and the difference is shaded grey. Negative voltages are plotted above the zero voltage line in this plot, in accord with the tradition used in earlier MMN reporting. In this plot, with tick marks representing 10 ms intervals, the MMN emerges around 80 ms, and lasts until 150 ms, post-event.

The prototypical experimental method for eliciting an MMN is the *oddball paradigm* (Näätänen et al., 2007; Winkler, 2007). In this paradigm, illustrated in Figure 3.2, the unexpected events are violations of stimulus patterns that have been experientially established by repeated presentations. Events that exemplify the pattern are called *standards*. Those that violate it are called *deviants* or *oddballs*.



**Figure 3.2:** Three examples of stimulus sequences from the oddball paradigm. Top: a series of ascending digits ending with a letter, this breaks both the pattern of digits and the pattern of ascending order. Middle: a series of ascending musical tones, followed by a larger descending step. This breaks both the pattern of ascending order and the pattern of consistent step size. Bottom: A series of colored squares which abruptly change color from green to red. This breaks the pattern of consistent color.

In its simplest (auditory) form, the oddball paradigm is as follows:

1. Short duration tones of the standard frequency are repeatedly presented with a fixed ISI (inter-stimulus interval) between tones.
2. Much more rarely, a tone of a fixed alternate frequency is presented, after some pseudo-random number of standard tones. This alternate frequency tone is termed the oddball.
3. A difference is computed between the average oddball-induced ERP waveform and the average standard-induced ERP waveform. As shown in Figure 3.1, in this difference waveform the MMN is the prominent negative deflection (here plotted upward) whose amplitude peaks between 100–200 ms after stimulus onset.

There are many variants on the paradigm. The repeating pattern can exemplify an abstract structural rule, such as a fixed relationship between successive tones in a tone sequence (Paavilainen, Simola, Jaramillo, Näätänen, & Winkler, 2001; Tervaniemi, Saarinen, Paavilainen, Danilova, & Näätänen, 1994; Vuust et al., 2005). There are also variants in multiple modalities, such as visual (Pazo-alvarez, Cadaveira, & Amenedo, 2003), tactile (Kuchenbuch, Paraskevopoulos, Herholz, & Pantev, 2014), and olfactory (Sabri, Radnovich, Li, & Kareken, 2005). When examined before the subtraction, certain variations end up evoking a P3a instead of, or in addition to, the MMN, such as the global/local paradigm (Basirat, Dehaene, & Dehaene-Lambertz, 2014). Finally, some variants elicit a MMN via omission of the stimulus instead of via substitution of an oddball stimulus (H. C. Hughes et al., 2001). This MMN to omission will be discussed in detail in a later section.

### *A Controversy*

There are competing hypotheses regarding MMN "genesis". Predictive coding (Garrido, Kilner, Stephan, & Friston, 2009) is an evolution of the original memory trace (or "model adjustment") hypothesis first put forward in Sams, Paavilainen, Alho, & Näätänen (1985), updated in line with the hierarchical predictive coding framework (Clark, 2013; Friston, 2005). It proposes that top-down predictions come from "higher" brain regions (such as the superior temporal gyrus), and are passed to "lower" levels (such as primary auditory cortex) to "explain away," i.e., predict and filter out, predictable components of bottom-up inputs. Any unexplained input is treated as an error signal, and is propagated to a higher level, e.g. the inferior frontal gyrus.

The sensory-specific adaptation (SSA) hypothesis emerged as a direct challenge to the predictive, memory-based models. SSA is a well-established phenomenon in which a neuronal population that responds to a given stimulus exhibits a reduced response following repeated exposure to that stimulus (Figure 3, right panel). Jääskeläinen et al. (2004) argued that the MMN could be generated entirely within the temporal (sensory) cortex, due to SSA. Similarly, May & Tiitinen (2010) argued that the MMN is a result of SSA at neural sites responsible for the N100 (or N1) ERP component. If sufficient, the SSA hypothesis would explain MMN with no cognitive machinery required.

For further reviews covering the arguments being made between these two schools of thought, see May & Tiitinen (2010), and Garrido et al. (2009).

### *Constraints on MMN Models*

Any successful model of the MMN must reproduce most of its key properties, while making no predictions contrary to fact. The first set of properties we consider here depend on the stimulus parameters. First, the MMN magnitude depends on the strength of the expectation and the deviance of the unexpected events, as well as attentional and other variables (Kujala & Näätänen, 2003). The latency to the MMN's peak, within the 100-200 ms after the oddball's onset, depends on multiple factors. Tone deviants that differ from the standard in pitch have a different MMN latency than deviants that differ only in duration or intensity (Kujala & Näätänen, 2003). Furthermore, for a given standard tone, MMN latency declines as a function of the difference in Hz between the oddball and the

standard (Naatanen & Alho, 1995). The MMN peak's amplitude of a few microvolts is a linear function of the oddball's deviation from the standard (Kujala & Näätänen, 2003).

In addition to stimulus-based features, the MMN has been studied in both humans and animal models under a variety of pharmacological manipulations, as tabulated below:

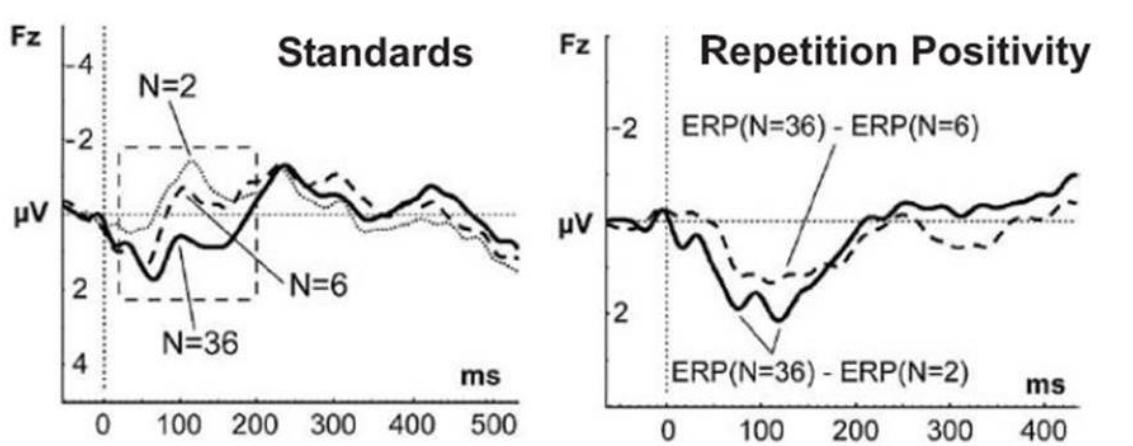
Receptor Type	Studies	Agonist MMN Results	Antagonist MMN Results
NMDA	(Ehrlichman, Maxwell, Majumdar, & Siegel, 2008; Farley et al., 2010; Heekeren et al., 2008; Javitt, Steinschneider, Schroeder, & Arezzo, 1996; Umbricht, Koller, Vollenweider, & Schmid, 2002)	No Studies	Dose-dependent reduction or abolition of the MMN
Muscarinic ACh	(Klinkenberg, Blokland, Riedel, & Sambeth, 2013; Pekkonen, Hirvonen, Jääskeläinen, Kaakkola, & Huttunen, 2001; Riekkinen et al., 1997)	Cholinesterase inhibitor reduces MMN in Alzheimer's populations, but not in non-clinical population	Mixed results: one study shows abolition, two show no effect
Nicotinic ACh	(Engeland, Mahoney, Mohr, Ilivitsky, & Knott, 2002)	Increased Amplitude	

**Table 3.1: Summary of pharmacological manipulations of the MMN.** Particularly relevant effects are the dose-dependent reduction in MMN amplitude under NMDA blockade and the reduction in MMN amplitude in Alzheimer's populations under a cholinesterase inhibitor.

NMDA antagonists, administered prior to the oddball paradigm, have been repeatedly and reliably shown to reduce and then abolish the MMN in both humans and animal models (Ehrlichman et al., 2008; Farley et al., 2010; Heekeren et al., 2008; Javitt et al., 1996; Umbricht et al., 2002). Models have begun using this as a constraint, and some existing models (May & Tiitinen, 2010; Wacongne et al., 2012) feature abolition of the MMN in response to simulated NMDA blockade. Results are mixed, however, with regard to

muscarinic ACh agonist effects on MMN. Dose-dependent reduction/abolition has so far only been found in patients with Alzheimer's disease (AD), not in young, healthy individuals. This does not necessarily imply that the effect is absent in healthy subjects. Instead, evidence from other studies suggests that the effects of elevated ACh levels would be particularly enhanced in AD patients, due to their chronically reduced levels of ACh and the correspondingly elevated ACh receptor counts (Overk et al., 2010), including type I muscarinic (M<sub>1</sub>) receptors, which control key transmembrane ionic currents, such as the M-current. If so, much larger agonist doses might produce an MMN reduction in non-AD subjects.

In addition to the MMN itself, there is a positive deflection before the negativity, the amplitude of which corresponds to the number of standard stimulus repetitions preceding the oddball. This was first analyzed by Baldeweg (2007), and named the Repetition Positivity (RP). It is shown in Figure 3.3.



**Figure 3.3:** The Repetition Positivity (RP), adapted from Baldeweg, 2007. (A) The standard ERP at Fz after 2, 6, and 36 repetitions, recorded in a roving standard experiment of Haenschel et al. (2005) during passive listening. The inset indicates the time window for early ERP components (P30 and P50) shown in (C). (B) RP is

**the difference wave between the standard ERPs after 36 and 2 (black line) and 36 and 6 (dashed line) repetitions, respectively.**

This positivity can be taken as an electrical index of the strength of expectation-based priming, and as such should be accounted for by expectation-based models of the MMN.

It should be noted that many MMN modelers emphasize the ability of their models to produce a mismatch to omission. As stated before, some variants on the experimental paradigm elicit an MMN not only to presentation of an oddball stimulus but also to omission of any stimulus at the expected time of the standard. However, evidence suggests that this "MMN to omission" obeys different laws and so may be regarded as a distinct phenomenon. In particular, it is elicited only by protocols with short ISIs. Reports of the longest ISI to produce an MMN to omission range from 140 ms (Tervaniemi et al., 1994) to 800 ms (Halgren et al., 1995), whereas the MMN to a deviant is reliable at ISIs up to 10 seconds (Bottcher-Gandor & Ullsperger, 1992). Two explanations for a response to omission have emerged (Oceák, Winkler, Sussman, & Alho, 2006). These base the MMN to omission on either loudness summation or a window of integration. Thus neither of the currently competing explanations attribute the omission effect to the same mechanism used to explain MMN responses to deviants. Therefore, while many neural mechanisms such as the gated dipole (Banquet & Grossberg, 1987) could reproduce this response to omitted stimuli, we will not be requiring that our model of MMN genesis include such mechanisms or be capable of producing an MMN to omission.

### Existing Models

Computational neural models of the MMN have a history dating back to Banquet & Grossberg (1987). As stated above, the most complete model to date under the predictive coding hypothesis is Wacongne et al. (2012).

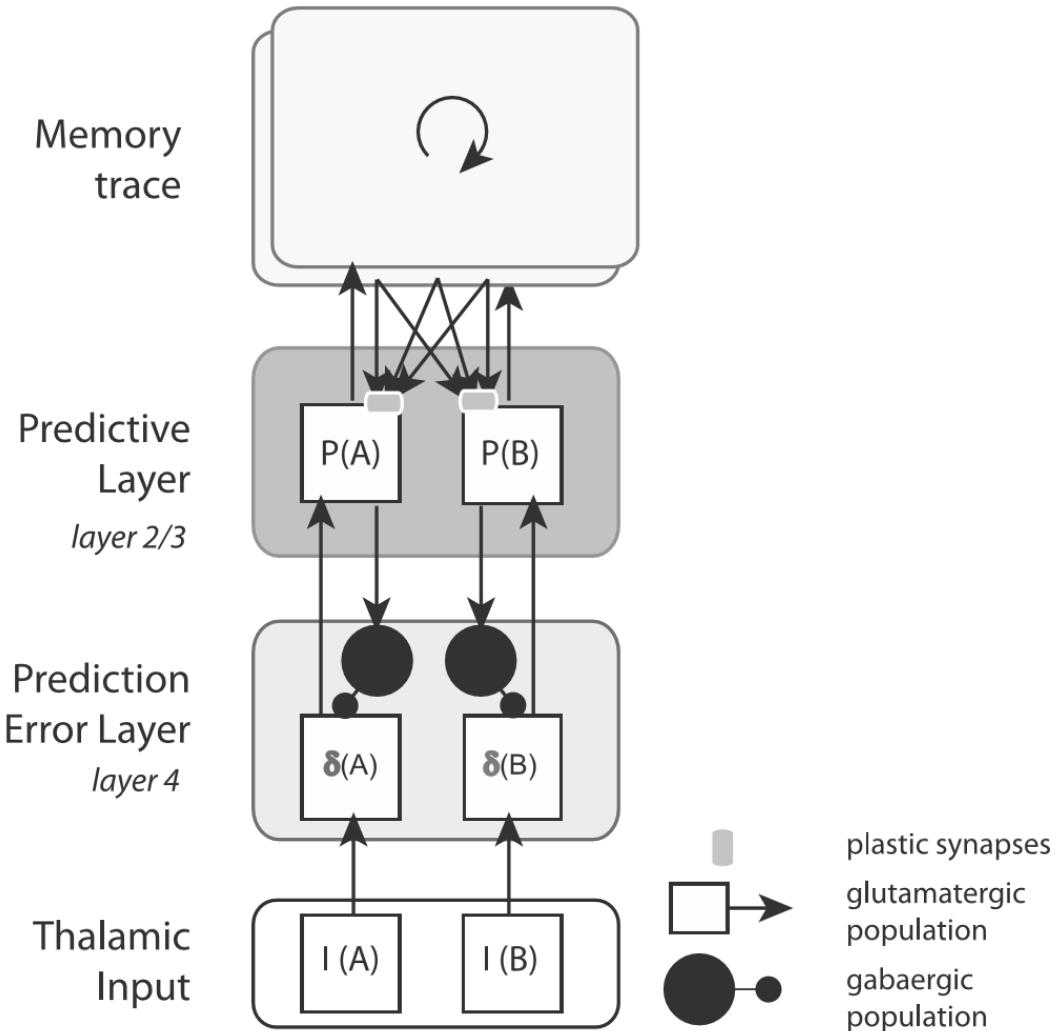


Figure 3.4: The network architecture of Wacongne et al. (2012), designed for a simulated experiment with two sounds. It consists of an abstracted higher-level memory trace, and then two cortical subpopulations for each layer (layer 2/3 and layer 4), as well as two subpopulations of thalamic input. Adapted from Wacongne et al. (2012)

Wacongne et al. (2012)'s model neurons are based on Izhikevich (2003), with synaptic

currents taken from the Brunel & Wang (2001) model. The model uses the simple

subtractive logic of most neural comparator models to generate an MMN as follows: as a stimulus sequence is presented, a memory trace learns the pair-wise associations. When the first stimulus of a pair occurs again, the memory trace activates inhibitory neurons in the second stimulus's column. These neurons inhibit the response of L4 excitatory cells to the second stimulus of the pair. If the predicted second stimulus occurs, L4 exhibits a reduced response. If a different stimulus occurs, a non-inhibited column is stimulated, producing a full-scale response. Those two responses, when compared by subtraction, produce a non-zero difference, and a resultant deflection from zero in a plotted ERP (difference) wave. This model was able to reproduce an MMN, an MMN to omission, and MMN abolition under NMDA blockade.

The model of May & Tiitinen (2010) simulates multiple areas of auditory cortex using simple non-spiking models. They use an extremely simplified model at the column level, describing an average depolarization and average firing rate for the pool of all excitatory cells in a cortical column. They do not explicitly model inhibitory neurons, but do include a general inhibition term that is based on local and neighboring column activity. They model stimulus specific adaptation (SSA) as a temporary reduction in afferent synaptic weights, as noted in Figure 3.5.

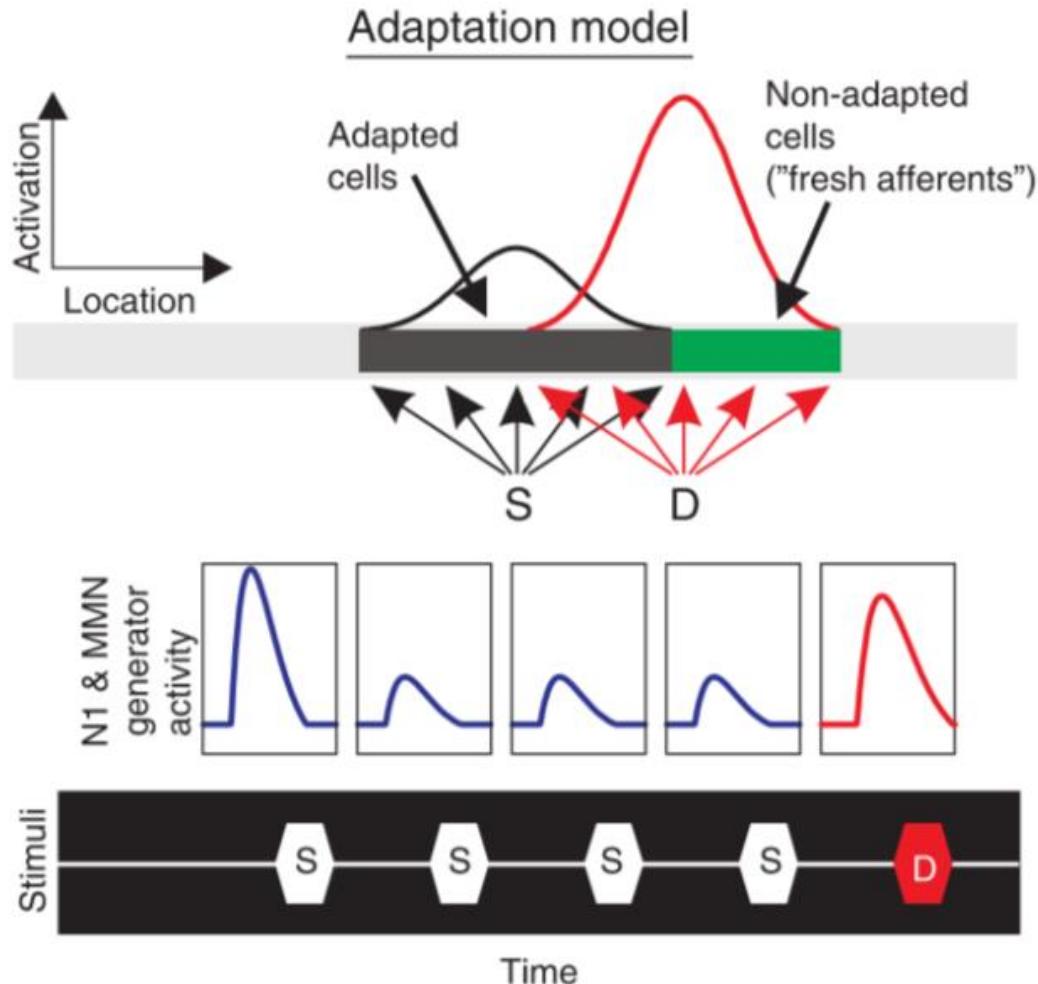


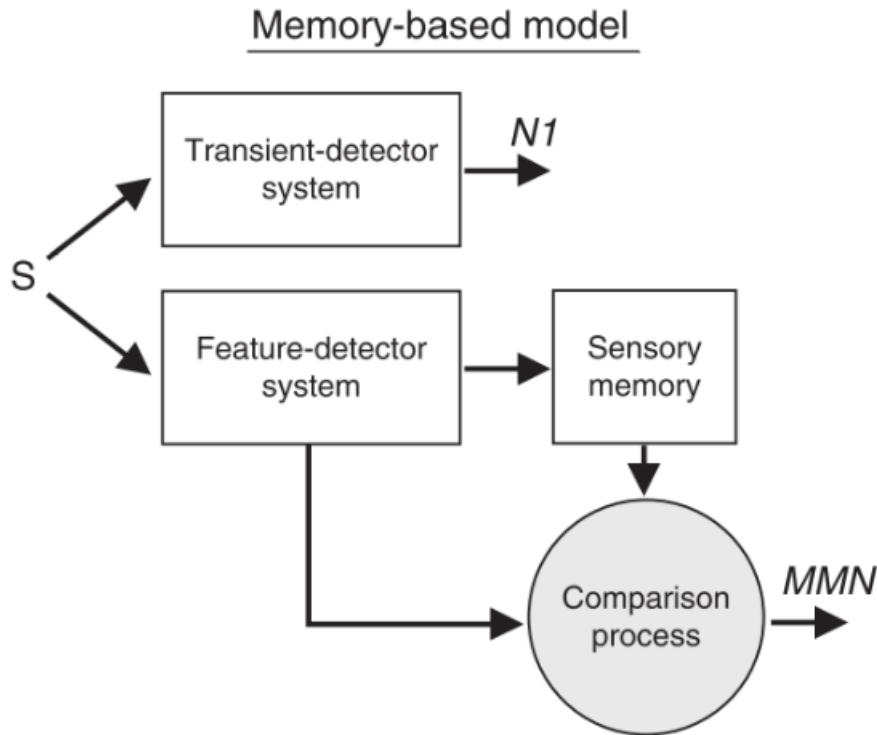
Figure 3.5: May and Tiitinen's adaptation model. In the adaptation model, the standards and deviants activate overlapping neural populations. The repetitive standard leads to cells tuned to the standard to become adapted. When the deviant is presented, nonadapted "fresh afferents" contribute to an enhanced response. Also shown is the beginning of a stimulus sequence, four standards (S) followed by a deviant (D), and the event-related responses produced by the neural population pictured at the top. Adapted from May & Tiitinen, (2010)

Lateral inhibition was simulated between core area neurons, with an onset time constant of 10ms. Lateral excitation was also included, without a divisive time constant term in the equation. This fast lateral excitation and slow lateral inhibition creates a nonlinear oscillator.

Their model successfully produces an activity pattern that mimics the MMN, which they suggest implies that MMNs can be explained by SSA. It is also capable of producing an auditory MMN to omission, which strongly depends on the properties of its nonlinear oscillator circuitry. In particular, the repeated stimulation provided by the standard starts the core oscillating with a peak response at roughly 100ms post-stimulus-latency. It is this peak, they claim, that constitutes the “MMN to omission”.

#### *Limitations of Existing Models*

Despite some successes, both these approaches to MMN modeling have notable limitations. Wacongne et al., (2012) follows in the tradition of earlier psychological models of the MMN (Naatanen, 1990, Winkler, 2007), which rely on the existence of separate circuitry dedicated to stimulus feature processing and to detecting mismatches and generating the MMN (for an overview and critique of this trend, see May & Tiitinen, (2010)), but not both.



**Figure 3.6:** The architecture of a traditional memory-based conception of the MMN, with separate modules performing the feature detection and the comparison and MMN generation. Adapted from May & Tiitinen (2010).

Requiring a separate comparison module for each feature detection system in the neocortex is extremely inefficient, however. As implemented in Wacongne et al., (2012), the comparison module contains the same number of neurons as the feature detectors. Thus, this model requires a doubling of the size of the neocortex relative to any model that uses a single circuit to perform both functions.

The model proposed by May & Tiitinen, (2010) has an equally significant limitation. It lacks all top-down feedback. This is somewhat defensible, because the whole point of their construction is to explore a way to explain the MMN with only a bottom-up process.

Nevertheless, top-down input is a highly prominent feature of the neocortex (e.g., Barbas and Rempel-Clower, 1997; Markov et al., 2014), and of thalamo-cortical circuits, and it remains to be seen whether their MMN model would survive the incorporation of realistic top-down feedback. Such feedback must eventually be incorporated in order to advance any serious model of cortical computations.

We address the limitations of both the existing predictive coding and habituative theories by introducing aspects of the May and Tiitinen model to a predictive coding framework; in particular we model mismatch computations and the MMN via adaptation in the same neurons that perform sensory feature processing. However, rather than rely on purely afferent adaptation, we introduce predictive adaptation driven by top-down feedback, in accordance with the predictive coding framework and neuroanatomical principles. This approach, detailed below, produces a model that mitigates both of the above-noted limitations of prior computational models.

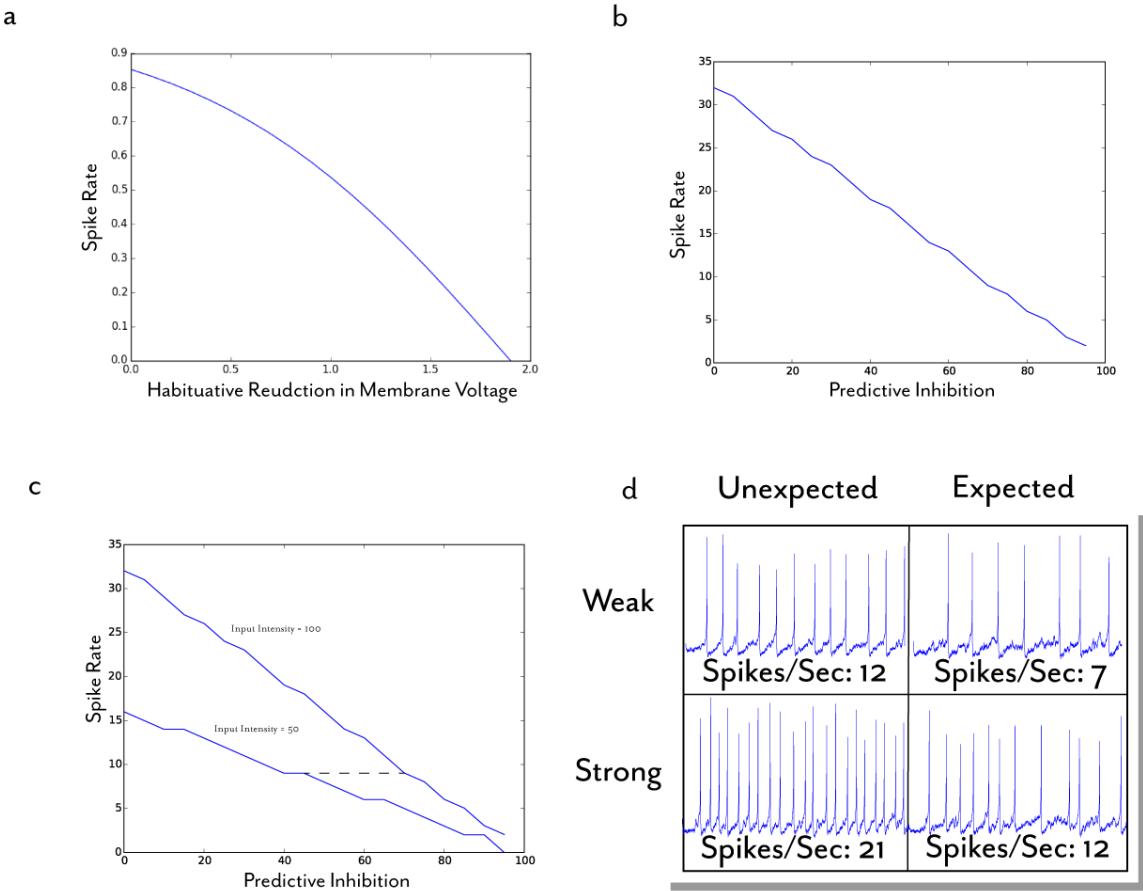
Moreover, there is one other key limitation shared by both Wacongne et al.'s and May & Tiitinen's models: both of them use subtractive comparators. Specifically, the output firing rate as a function of prediction (or of habituation) strength varies along a mostly linear downward trajectory (as demonstrated in Figure 4.7), which presents several problems for any recipient neurons attempting to decode this output.

- Any decoder will experience potential confusion between an unexpected weak stimulus and an expected strong stimulus. Recall that an oddball with a mere

change in intensity can elicit an MMN. In the extreme, both models predict that a perfectly expected stimulus will be indistinguishable from an absent stimulus, which is clearly not the case.

- Graded lowering of spike rates implies lengthening times needed by recipient cells to exceed threshold. Greater delays would be incurred during processing of better expected stimuli, which is in opposition to evidence from electrophysiological, EEG, and psychophysical studies (Noonan et al., 2016).
- The constant modification of the firing rate means that either extensive circuitry must be in place to reconstruct the original signal, or the mismatch comparison must be performed on a copy of the signal, which entails redundant circuitry.

Figure 3.7c illustrates the principle, implied above, that linear subtractive comparators suffer from an ambiguity, specifically an inability to differentiate in their output between a predicted strong input and a weak or absent input. In fact, any given rate of output from the comparator could indicate any number of different input intensities, all predicted with varying levels of success.



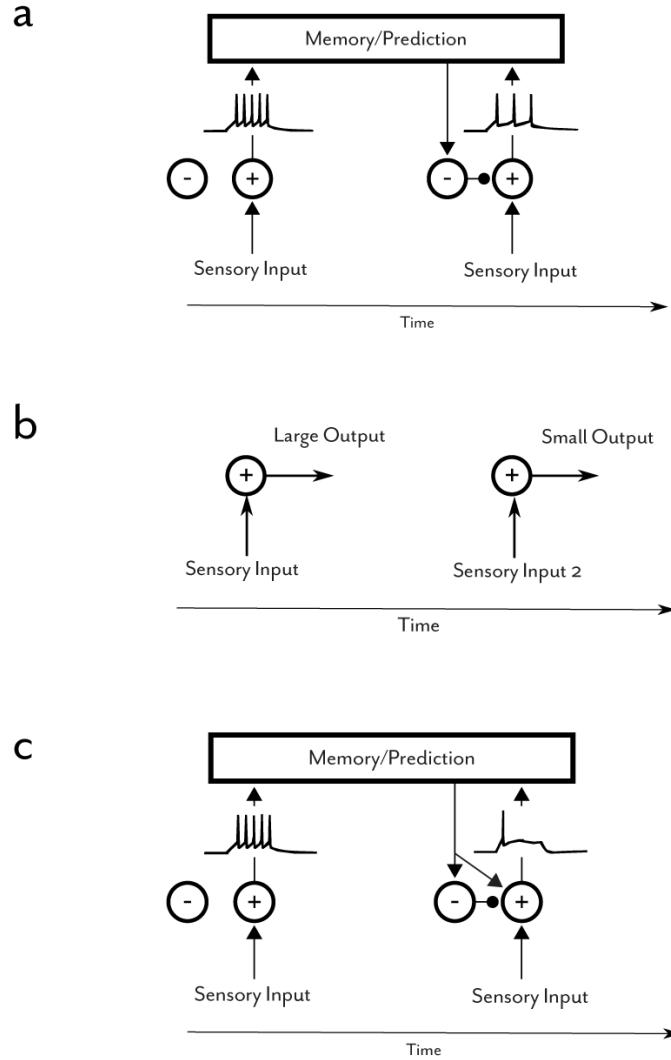
**Figure 3.7: a)** The nearly linear habituative reduction in membrane voltage of the comparator in May and Tiitinen (2010), with the output spike rate on the Y-axis and the amount of habituative inhibition on the X. As inhibition increases the spike rate is smoothly reduced. **b)** The linear response of the comparator in Wacongne et al. (2012). Output spike rate is again on the Y-axis, with the amount of inhibition (per Izhikevich (2003)'s quadratic integrate-and-fire model) coming from expectations on the X-axis. In this model the reduction in spike rate is almost perfectly linear. **c)** Confusability between a weak, unexpected input and a stronger, more expected input in the Wacongne comparator. Because the model expectations simply scale down the spike rate, a given spike rate could indicate a wide variety of stimuli under different expectations, making it difficult to decode. **d)** The same confusability illustrated with neuron voltage traces, showing how different strengths of inputs can result in the same spike rate.

### *Proposed New Model*

The above-mentioned limitations can be effectively addressed by a model of mismatch detection that is surprisingly simple at the circuit level, once we grant the criticality of accurately modeling how pyramidal dendrites process both top-down and bottom-up

signals. The model's circuit-level simplicity also critically depends on the incorporation, and functional reinterpretation, of prominent yet overlooked features of single neocortical pyramidal neurons. Notable among these features are the M-current and its interaction with Cl<sup>-</sup> currents.

The M-current, a current mediated by KCNQ channels, is an inward potassium current, originally discovered in sympathetic ganglion neurons of bullfrogs by Brown & Adams (1980). It is non-inactivating and voltage-dependent, being activated by depolarization of the membrane above approximately -60mV (Brown & Adams, 1980). Because of this, depolarizations large enough to generate action potentials reliably also activate the M-current. KCNQ channels are closed by Muscarinic ACh receptors, and are found in neurons throughout the nervous system, including pyramidal cells of the cerebral cortex (reviewed in Jentsch, 2000; Marrion, 1997). The M-current's function is usually interpreted from the perspective of rate-modulation: it contributes to slow adaptation in spike rate during prolonged excitation. However, under readily achieved physiological conditions, it can act quite differently, and serve other functions. Prescott, et al. (2006) showed that the M-current interaction with the chloride leak current can reliably cause a shift from regular (tonic) spiking to onset (phasic) bursting. This was demonstrated both in computational models, similar in structure to our proposed models below, and in dynamic clamp recordings from CA1 hippocampal pyramidal cells.



**Figure 3.8: A) A standard predictive coding comparator, where a stimulus elicits predictions which inhibit a population, reducing the frequency of its output when the predicted stimulus occurs. B) A standard habituative comparator, which fires at a reduced rate due to afferent habituation when a recently experienced stimulus is repeated. C) A mode-shifting comparator, wherein a stimulus elicits predictions which prime a population, placing it into a high-conductance state and causing a shift from tonic to phasic firing.**

In the new comparator model proposed here, this mode shift allows a neural population to dynamically switch from encoding detailed, rate-based information about an unexpected stimulus to quick, efficient phasic reporting of the onset of expected stimuli. These stimuli

will not require detailed information transmission, since a single “byte” of confirmation – a phasic, high-frequency burst followed by quiescence – can enable recipient neurons to act as though the brain’s expectation was confirmed. Thus the same circuit may suffice for purposes of feature processing and mismatch detection. This process is illustrated in figure 3.8.

Furthermore, the limitations of linear comparators are all at least partially addressed by our proposed model.

1. The nonlinearity of response means that confusability is reduced, although not eliminated.
2. Reporting matches by brief phasic bursts – just a few high-frequency spikes – allows for very fast processing of expected stimuli by recipient neurons, without incurring the delays associated with graded lowering of spike rates.
3. A perfectly expected stimulus is still reported via a phasic burst, differentiating it from a stimulus omission.
4. Because expected stimuli are not reported in a rate-based code, success does not depend on a preservation of the original sensory signal’s firing rate. Unexpected stimuli preserve the original firing rate, and induced release of ACh removes any bias from top-down priming, allowing the original, unaltered rate-based information to be passed on.

## Materials and Methods

Although the new model shares some features with the Wacongne et al. (2012) model, unlike Wacongne et al. and other predictive coding treatments, it explores the hypothesis that the comparison process does not require full parallel circuitry. Instead it uses both excitation-based and inhibition-based priming within the main stimulus processing channel, to allow synaptic inputs that mediate top-down expectations to activate traditionally "habituation" currents before neurons are driven by any event-related bottom-up ("driving") inputs. Such priming, which occurs only after expectation-development, induces a qualitative change in firing mode (Figure 4.8C). Thus, the proposed comparator reacts in qualitatively different manners to driving input, depending on whether or not it has recently received priming input. When primed, it responds to a range of driving inputs with a single spike (or transient burst) just at the onset of the sustained input. This suffices to confirm, to recipient brain areas, that expectations were met. If unprimed, the same "comparator" neuron generates a train of spikes whose frequency varies to reflect the intensity of driving input. This provides detailed encoding of stimulus features. It should be noted that our particular parameter set produces a high frequency response, which may be closer to firing saturation than would be optimal, but tests have shown that the mode switch operates under a broad range of parameters, so long as the basic ratios of priming and driving weights are maintained.

The proposed model's biophysical and neuroanatomical details, including strongly nonlinear interactions between the excitation-driven habituation and inhibition-driven subtraction, allow it to address limitations found in the alternative comparator models,

with a much greater economy of circuitry, as noted. Appendix A formulates mathematically a neuronal biophysical model that can be used to simulate these conditions. As specified there, the model incorporates the range of receptors and channels included in prior models of pyramidal cells (Prescott et al., 2006) and cortical synapses (e.g. Brunel & Wang, 2001), but also adds further receptors and interactions needed to realistically model ACh-dependent M-currents, which play a fundamental role in our treatment of cortical comparison processes. In particular, a voltage-dependent M-current was added, based on the mathematical modeling in (Mainen & Sejnowski, 1996). We use simulations to assess and document the model's explanatory power vis-à-vis all the phenomena noted and tabulated above as constraints on MMN models.

Our simulations can be grouped into five major investigations. The first demonstrates the ability of a simulated layer 2/3 cortical pyramidal cell to differentiate between expected and unexpected input. Layer 2/3 pyramidal cells were selected as our model comparator for reasons explained previously in chapter 2. The second investigation examines the impact of the strength of priming (as a placeholder for strength of expectation) on the repetition positivity. The third is a pharmacological investigation, simulating the effects of NMDA and KCNQ blockade on the neuron, along with a KCNQ agonist, such as retigabine. The fourth investigation tested robustness to the presence of noise and to changes in input weights. Finally, the fifth investigation focused on the ability of the comparator to avoid the regions of confusion found in the firing state space of linear comparators.

### *Proposed Model*

In documenting our proposed model here, we employ the standards recommended by Nordlie, Gewaltig, & Plesser, 2009, for specifying the architecture, network connectivity, neural and synapse models used, and the input to and measured output from the simulated system.

### *Architecture*

For the first four investigations, the architecture is extremely simple. A single neuron representing a layer 2/3 pyramidal cell performing sensory comparison is the core of the model, as it is a single-neuron comparator. For investigation five, we simulated two neurons, one representing a sensory-area layer 2/3 comparator, and one representing a pyramidal cell located in an association area, which receives its inputs from the sensory comparator.

### *Network Connectivity*

While there are a high number of connections, they form a very simple pattern. The sensory comparator is innervated by 600 synapses: 200 basal driving synapses (presumed to be from excitatory layer IV neurons, which were not simulated), 200 apical priming synapses (presumed to be from higher cortical regions), and 200 basal inhibitory priming synapses (presumed to be from priming-driven feed-forward cortical interneurons). All synapses have weights that are randomly selected using a Gaussian distribution with  $\mu$  of 1.0 and  $\sigma$  of 0.1. The weights are of the form  $w \cdot \text{gauss}(\mu, \sigma)$ , where the value of  $w$  is 16

for driving synapses, 12 for excitatory priming synapses, and 8 for inhibitory priming synapses.

### *Neuron and Synapse Models*

#### Neuron Model

Each comparator neuron is made up of two point models, one representing the apical portion of the cell, and the other the basal. The membrane of each point model is conductance-based, following the general convention of

$$C \frac{dV}{dt} = I_1 + I_2 + \dots + I_n$$

where  $I$  is a current of the form  $g_{max} \cdot g_{proportion} \cdot (V_m - V_E)$ , where  $g_{max}$  is the maximum conductance of the channels carrying that current,  $g_{proportion}$  is the proportion (from 0 to 1) of the max conductance currently engaged (or the proportion of membrane channels currently open and non-inactivated), and  $V_m - V_E$  is the driving voltage differential between the current membrane voltage ( $V_m$ ) and the Nernst equilibrium of that particular current's ion ( $V_E$ ).

In our model, the complete membrane equation is

$$C \frac{dV}{dt} = I_{Na} + I_{AMPA} + I_{NMDA} + I_{GABA} + I_K + I_M + I_{Cl^-}$$

with  $I_{Na} + I_{AMPA} + I_{NMDA}$  representing the summed synaptic currents. The remaining, non-synaptic currents are detailed in the appendix, along with the parameters used.

### Synapse Model

Synaptic connections addressed by fibers coming in from outside the scope of the model are driven by Poisson point-process spike generators associated with the assumed fibers. The Poisson function behind each one varies as a function of time, with an onset and offset time. The details of this change over time are provided in the "Input" section, below. When a synapse is hit by one of these Poisson-generated spikes it activates postsynaptic conductances in accordance with either the AMPA and NMDA or the GABA membrane conductance equations from part two. The effect depends on whether or not the synapse has a positive (excitatory, and therefore AMPA and NMDA) or negative (inhibitory, and therefore GABA) weight. The synapse equations are also detailed in the appendix.

The equations put forward by Brunel & Wang (2001) were used to specify each particular synaptic conductance. The precise equations used are described in the appendix.

### *Input*

Each of the hundreds of input synapses is driven by a Poisson process whose lambda variable changes over time, going from 0.0001 at rest to a given value during activation. Activation lambda values and onset/offset timing are specific to each population of inputs, and to whether or not we were simulating an expected (primed) stimulus, or an unexpected (unprimed) stimulus.

Regardless of expectations, driving inputs are activated at 350ms into the simulation, with lambda switching to a value of 0.01 for a duration of 350ms (stimulus duration), whereupon the driving inputs returned to the resting lambda. During primed stimulation, the excitatory priming inputs are activated at 150 ms, switching to a lambda value of 0.01 for a duration of 550ms, returning to a resting lambda (0.0001) at the same time as the driving inputs.

During primed stimulation, the inhibitory priming inputs are activated at 175 ms at a lambda value of 0.01, and also return to a resting lambda at the same time as the driving inputs. During unprimed stimulation, the priming inputs (both excitatory and inhibitory) will not activate at all, remaining at rest lambda throughout the entire simulation.

### *Output*

The membrane voltage and extracellular current were measured for analysis at both the apex and the base of all neurons, along with the net KCNQ-mediated M-current. In addition, for each neuron we calculated the quasi-static electric field using the methods provided in the Appendix.

### *Data Analysis*

This field was then low-pass filtered with a cutoff of 14Hz to remove individual spike artifacts which do not come through in scalp EEG. The resulting field trace represents the first physically accurate modeling of the electrical fields in any model of the mismatch

negativity. Spike detection was performed on the simulated membrane voltage data by recording as a spike's time each crossing from below -10 mV to above 10 mV.

Each neuron's sequence of spike times was then classified as phasic or tonic by dividing the time up into two segments. The first, the onset segment, consisted of the first 100ms after the onset of driving (not priming) synaptic activity. The second, sustained segment consisted of the time between the end of the onset segment and the offset of driving activity. The rate in spikes/second was calculated for each segment, and if the onset rate was at least three times the sustained rate, as well as the sustained rate being below 5 spikes/second, the neuron's firing pattern was classified as phasic. Otherwise, the neuron was classified as tonic.

#### Investigation 1: Match/Mismatch

The first investigation simply involved simulating a primed and an unprimed condition as detailed above in order to demonstrate that the primed condition produced phasic firing while the unprimed condition produced tonic firing. No further analysis was performed in this baseline investigation.

#### Investigation 2: Repetition Positivity

As mentioned previously, the amplitude of the repetition positivity has been found to vary with the strength of expectations (Baldeweg 2007). To determine whether our model could reproduce this effect, a "match" condition was simulated three times, once as in investigation 1, once with the lambda for both excitatory and inhibitory priming synapses

reduced by 33% (simulating a weaker expectation), and once with it reduced by 66% (for an even weaker expectation). Per Baldeweg (2007), we then plotted the simulated electric field responses to just the standards, to look for a reduction in amplitude as the priming signal became weaker.

#### Investigation 3: Pharmacological Manipulations

Each of our pharmacological manipulations was set at either a 25% increase (for agonists) or a 25% decrease (for antagonists) of the maximum channel conductance value. To examine the impact of various pharmacological agents, the basic match/mismatch simulations were performed under the following conditions:

- NMDA activity decrease (e.g. via PCP; (Anis, Berry, Burton, & Lodge, 1983))
- KCNQ activity decrease (e.g. via high doses of a Cholinesterase inhibitor; (Riekkinen et al., 1997))
- KCNQ activity increase (e.g. via retigabine; (Soty et al., 2009))
- NMDA activity decrease and KCNQ activity increase combined

#### Investigation 4: Robustness To Noise

For this investigation, a single comparator neuron was simulated with injected white noise currents added to the apex and the base. The simulation was repeated 20 times each (10 primed and 10 unprimed) at 30 different levels of noise amplitude. For each simulation, the response of the neuron to driving input was classified as either phasic or tonic, per the ratio of the onset spike rate (the spike rate during the 100ms following the onset of driving

input) to the sustained spike rate (the rate during the remaining 250ms of the driving input). If this ratio was higher than 3:1 and the sustained spike rate was lower than 5 then the neuron was classified as phasic. This was then treated as a binary classifier, with true positives occurring when a neuron was primed and responded in a phasic manner, false negatives when a neuron was primed and responded in a tonic manner, false positives when a neuron was unprimed but responded in a phasic manner, and true negatives when a neuron was unprimed and responded in a tonic manner. These four values were then plotted for each noise level simulated.

#### Investigation 5: Confusability Testing

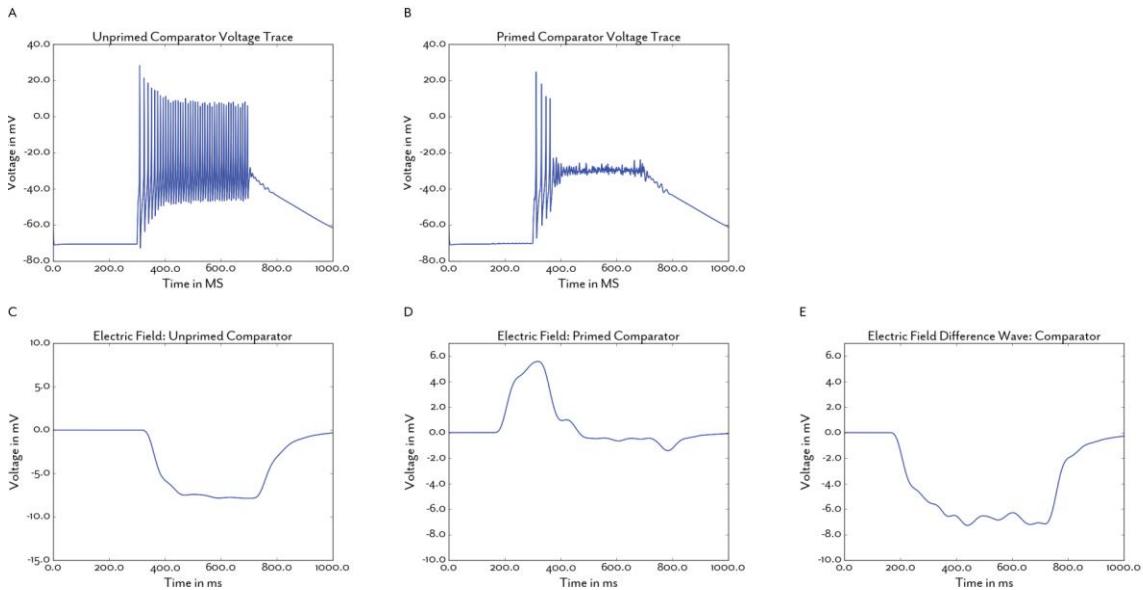
We defined a discriminability metric as the difference between sustained firing rates (as defined above in the Robustness Evaluation section) of two conditions. This is useful in that it means that two perfectly phasic responses (with a sustained spike rate of zero) are highly confusable (i.e. difficult to discriminate), a phasic and tonic response are easy to discriminate (i.e. this metric will have a high value), and two tonic responses can vary in this metric based on their relative firing rates.

Our confusability testing involved the simulation of a single comparator neuron over the space of three different driving input weights and nine different predictive priming weight sets (since priming involves both an excitatory and an inhibitory connection, both weights must be scaled by a single scalar value). The discriminability metric was evaluated for each simulation and its value was plotted along the predictive priming scalars for each of the three driving inputs.

## Results

### *Investigation 1: Match/Mismatch*

Our simulations performed as expected, and the results confirmed our predictions. A mode switch was indeed observed between the primed and unprimed conditions, as seen in figure 3.9A.



**Figure 3.9: Baseline mismatch negativity.** A) A rapid tonic firing from the unprimed comparator. B) A brief phasic burst from the primed comparator. C) The negative deflection of an unprimed comparator. D) The positive and slight negative deflection of the primed comparator. Note that the priming deflection begins before the spiking in B, due to the subthreshold influence of priming synaptic currents. E) The difference between the two electric fields, showing a negative deflection, or MMN. Negative is plotted downward in the electric field panels.

### *Investigation 2: Repetition Positivity*

The electric fields, plotted in Figure 3.10, showed a strong positivity, with a corresponding reduction in “match” as the expectational priming is reduced. This finding makes our

model of the MMN the first to produce a realistic repetition positivity, which is reliably found in the standard trials of experimental MMN data.

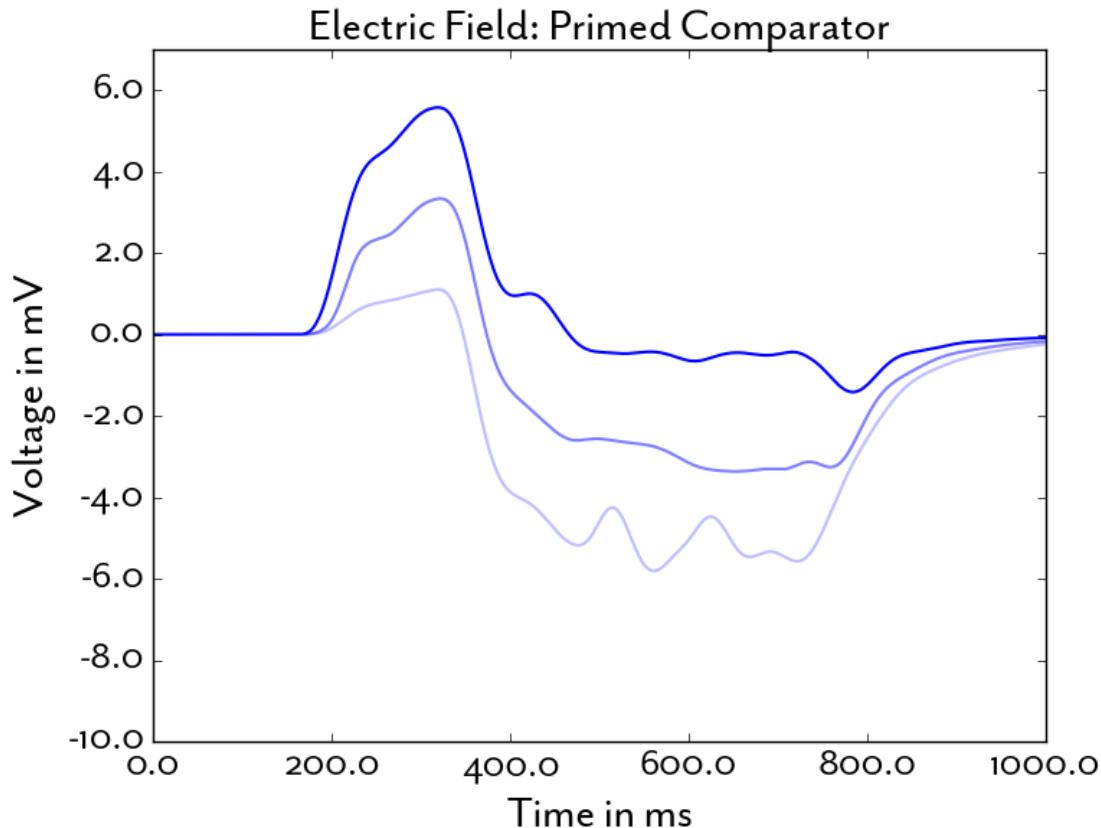
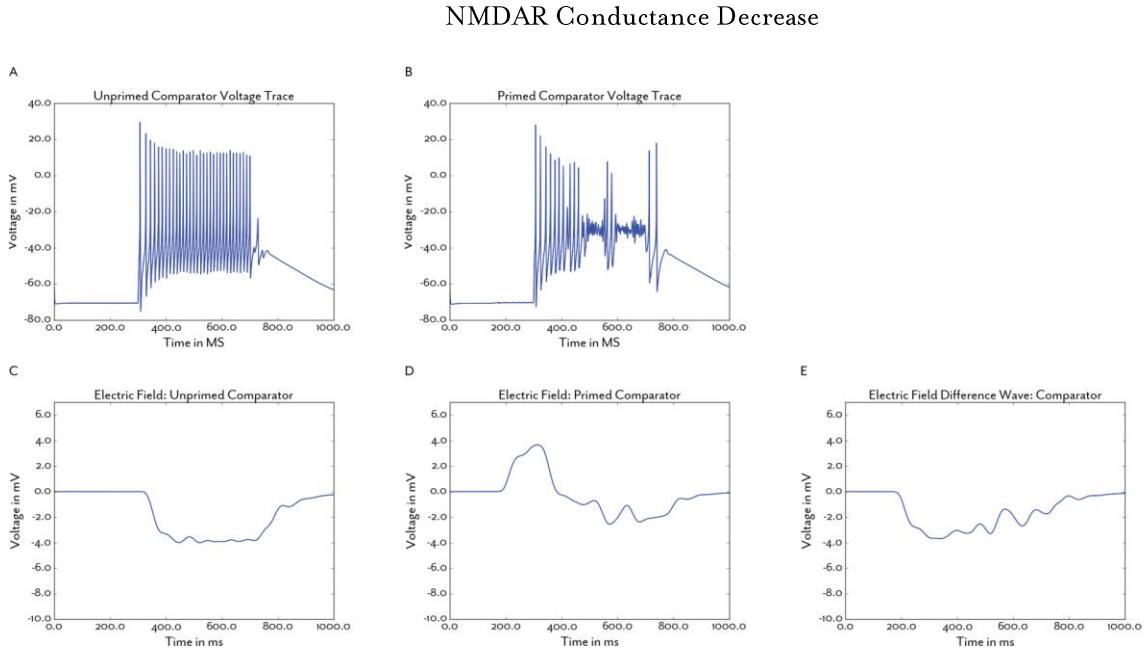


Figure 3.10: The simulated repetition positivity scales with priming strength. As the lambda was reduced from 100% of investigation 1 (darkest blue line) to 33% (lightest blue), the amplitude of the simulated repetition positivity dropped. Negative is plotted downward.

### Investigation 3: Pharmacological Manipulations



**Figure 3.II:** The model under an NMDAR blockade, as under PCP. A) The unprimed comparator voltage trace shows extremely rapid tonic firing. B) The primed comparator voltage trace shows fitsful firing throughout the stimulus, failing to achieve a full mode shift. C) The deflection in the electric field for the unprimed comparator is reduced. D) There is a similar reduction in the negative deflection of the primed comparator's field. E) The MMN is reduced compared to baseline (c.f. investigation 1). Negative is plotted downward in the electric field panels.

As shown in Figure 3.II, reducing NMDA max conductance by 25% prevents the mode switch to a phasic response. Under the view that processing based on standard formation relies on this mode switch, this blockade would prevent such processing from occurring in a model based on our comparator, causing all stimuli to be treated as mismatches. This is an effect of NMDA currents on neuronal responses to un-modified synaptic weights. Hence it does not depend on well-known roles of NMDA currents in plasticity mediated by long-term synaptic strength changes. Instead, the mechanism by which partial NMDA blockade prevents the mode-switch is a reduction in excitatory sodium conductance.

## KCNQ Conductance Decrease

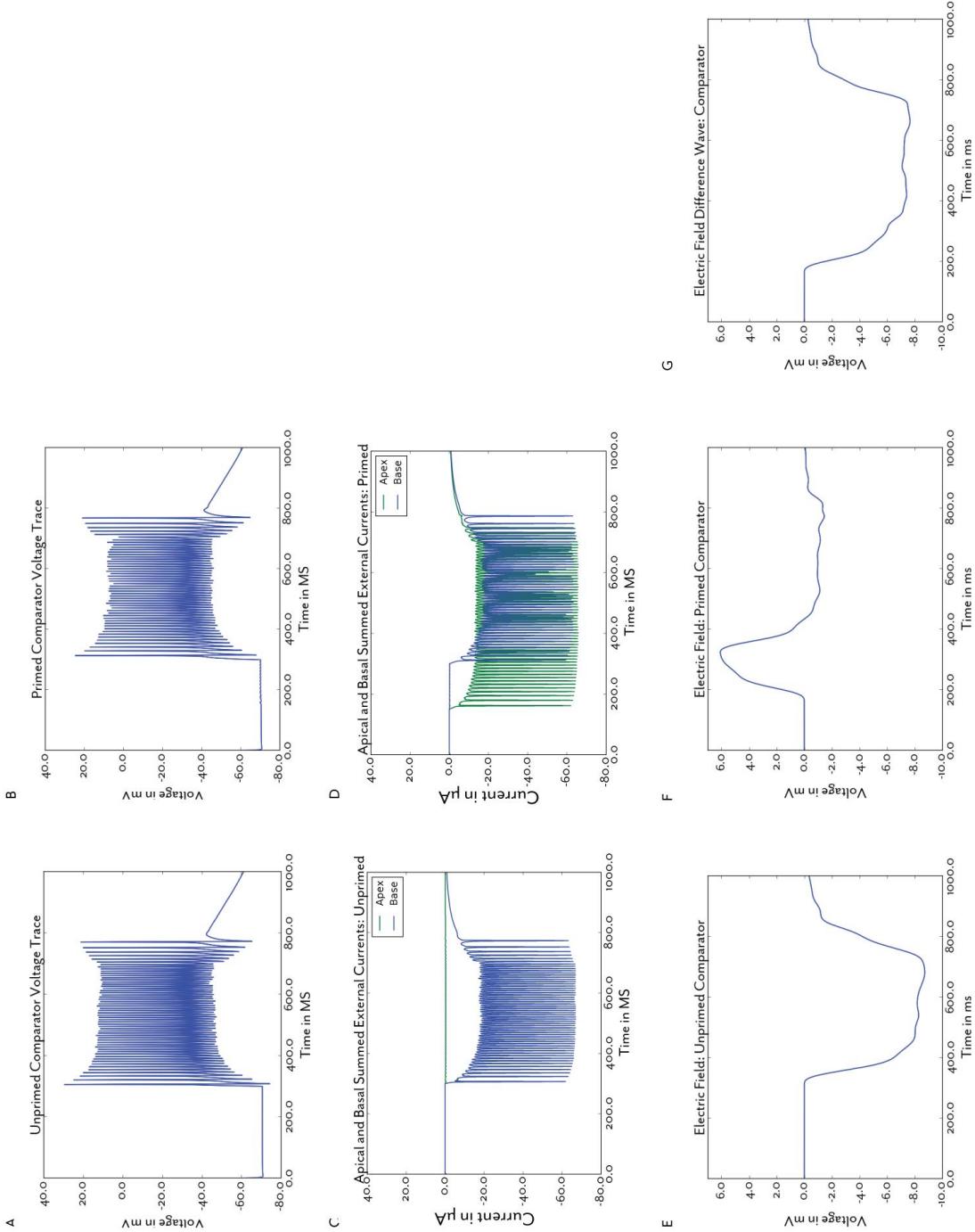


Figure 3.12: The model under reduced KCNQ conductance, as would be produced by increased ACh. A) A rapid tonic firing from the unprimed comparator. B) A similar rapid tonic firing from the primed comparator. C) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the

primed comparator. E) The negative deflection of an unprimed comparator. D) The positive and slight negative deflection of the primed comparator. E) The difference between the two electric fields, showing a negative deflection, or MMN. Negative is plotted downward in the electric field panels. Note that while the neural voltage trace in panel B is clearly distinguishable from the corresponding trace in the baseline figure, the MMN remains undiminished. This is due to the electrical field being derived from the difference in apical and basal transmembrane currents, unlike the basal voltage trace (which determines spike output). A more complete explanation is found below.

Increasing ACh sufficiently to reduce the M-current's maximum conductance by 25% also prevents the mode switch to a phasic response (Figure 3.12, panel B). Similar to the results under NMDA blockade, these results indicate that a circuit consistent with our comparator would be rendered incapable of processing based on standard formation, instead treating all stimuli as mismatches, during epochs with sufficiently elevated ACh.

How can this happen? The key factor is that the simulated electric field (defined in the appendix) is defined by the relative transmembrane current flow in the apical verses basal regions of the cell (see figure 4.12, panels C and D), whereas the spike output is determined by the transmembrane currents in the basal section combined with the axial current from the apical dendrite. This means under certain conditions, the electrical field can be modulated independently of the spike output, which is seen in the figure above. This unaltered MMN fits the findings on cholinesterase inhibitors in healthy subjects, while challenging any assumption that the lack of modulation in the MMN means ACh is unimportant to mismatch detection. In doing so, our model also becomes the first theory of MMN generation to propose that the difference wave ERP signature is dissociable from the spike-communicated results of the underlying comparison/mismatch detection process.

## KCNQ Conductance Increase

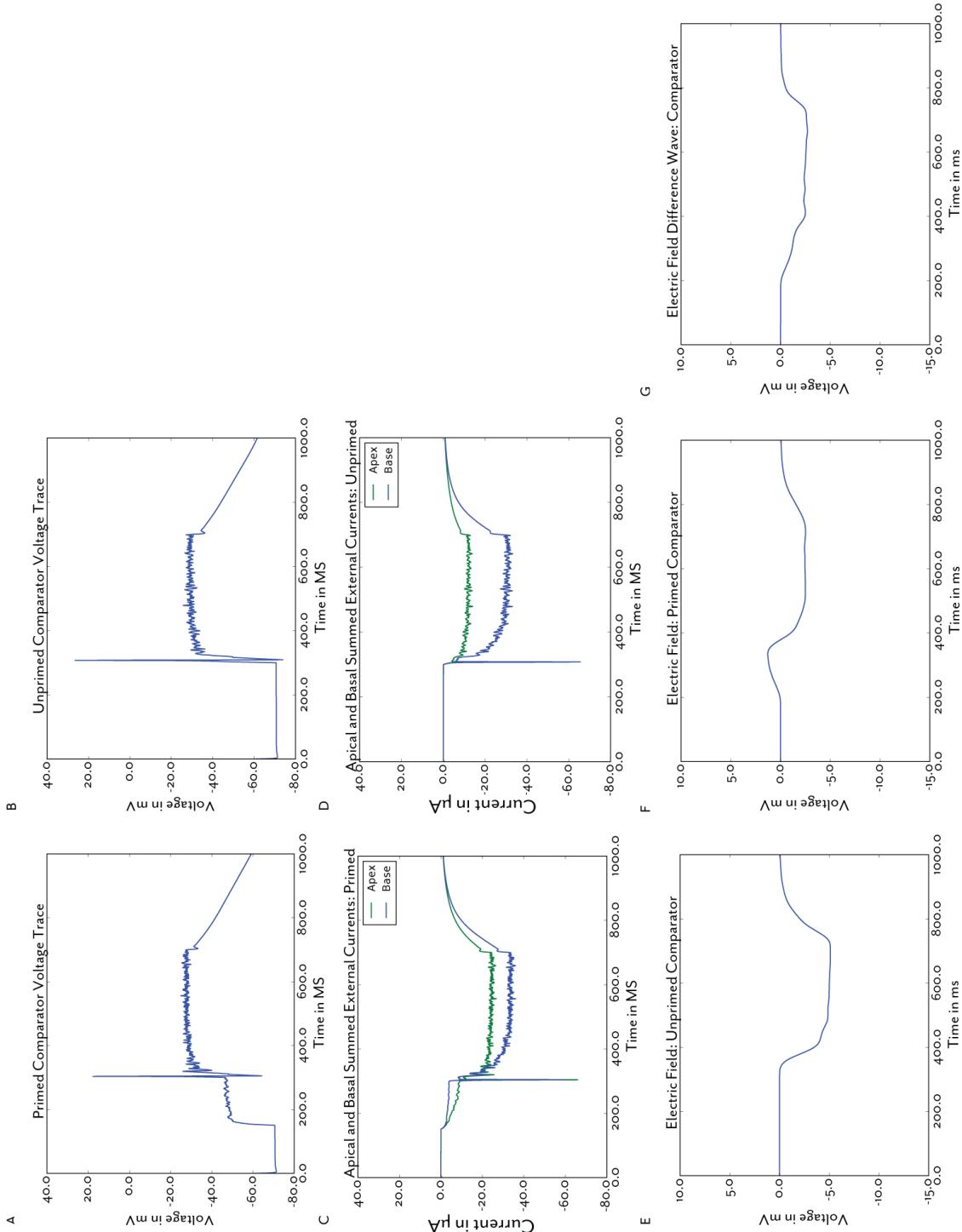


Figure 3.13: The model under increased KCNQ conductance, as would be produced by retigabine. A) A single phasic onset spike from the unprimed comparator. B) A similar phasic onset firing from the primed comparator.

C) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the primed comparator. E) The negative deflection of an unprimed comparator. F) The positive and negative deflection of the primed comparator. G) The difference between the two electric fields, showing a diminished negative deflection, or MMN. Negative is plotted downward in the electric field panels. Note that in this case the MMN is extremely diminished, in accordance with the failure of the spiking output of the comparator to distinguish between primed and unprimed conditions.

KCNQ activity increase: In a shift from the other manipulations, KCNQ agonism (reducing KCNQ max conductance by 25%) prevents the mode switch to a tonic response. This suggests that a circuit consistent with our model comparator would be incapable of mismatch, treating unexpected inputs as though they were predicted and primed. Unlike the previous increased ACh condition (figure 4.12), this condition shows a corresponding reduction of the MMN, to near extinction. To our knowledge, no studies have been performed on the impact of retigabine or other KCNQ agonists on the mismatch negativity, as well as none on the behavioral markers of mismatch detection (such as reaction time), making this a novel testable prediction. Our model predicts that differences in behavioral markers between standards and oddballs will be reduced, along with a reduced mismatch negativity.

## NMDA Conductance Decrease and KCNQ Conductance Increase

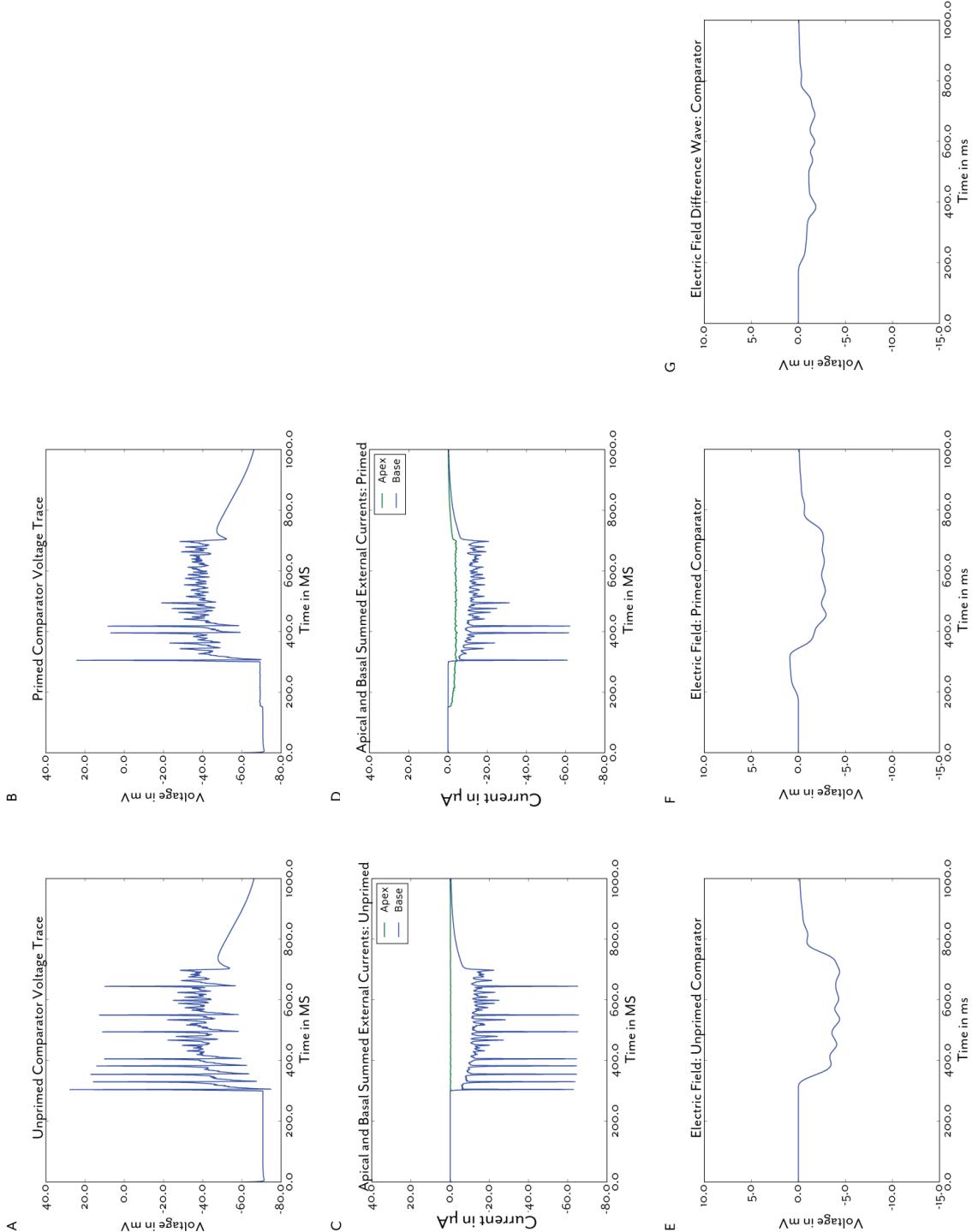


Figure 3.13: The model under both an NMDA antagonist and a KCNQ agonist. A) A rapid tonic firing from the unprimed comparator. B) A similar rapid tonic firing from the primed comparator. C) The transmembrane

(external) currents of the apical (green) and basal (blue) sections of the unprimed comparator. D) The transmembrane (external) currents of the apical (green) and basal (blue) sections of the primed comparator. E) The slight negative deflection of an unprimed comparator. F) The slight positive and negative deflection of the primed comparator. G) The difference between the two electric fields, showing an extremely diminished negative deflection, or MMN. Negative is plotted downward in the electric field panels.

In line with the findings of Sotty et al. (2009), simulating the joint effects of a 25% reduction in NMDA currents and of a 25% increase in maximum KCNQ conductance shows how a KCNQ agonist such as retigabine rescues the basic comparator functionality compromised by deficient NMDA currents. Figure 3.13 shows that the mode switch is once again observed under coincidental priming and driving inputs. This would allow for neocortical discrimination between expected and unexpected experiences, and, combined with the hippocampal retigabine findings from chapter two suggests a mechanism of action for retigabine's antipsychotic effects.

Also notable is that the simulated MMN did not fully recover under retigabine; it remained reduced, as in the NMDA blockade investigation. This extends our novel testable prediction, allowing us to predict that schizophrenics (or subjects on PCP) will benefit behaviorally and cognitively from the restored mismatch detection (and possibly even reduced distractibility, if the KCNQ alpha hypothesis put forward in chapter five is correct), but will still display a reduced MMN consistent with untreated schizophrenics. This is non-obvious, and to our knowledge this has not been tested.

*Investigation 4: Robustness to Noise*

The model turned out to be extremely robust to the presence of noise in its inputs, with the white noise parameter able to increase to more than 1000 mA before a significant falloff in classification sensitivity occurred. Specificity remained extremely high regardless of noise levels, as sufficient injected noise prevented the mode shift to phasic firing, which in turn prevented classification of a stimulus as a “positive”, or expected stimulus.

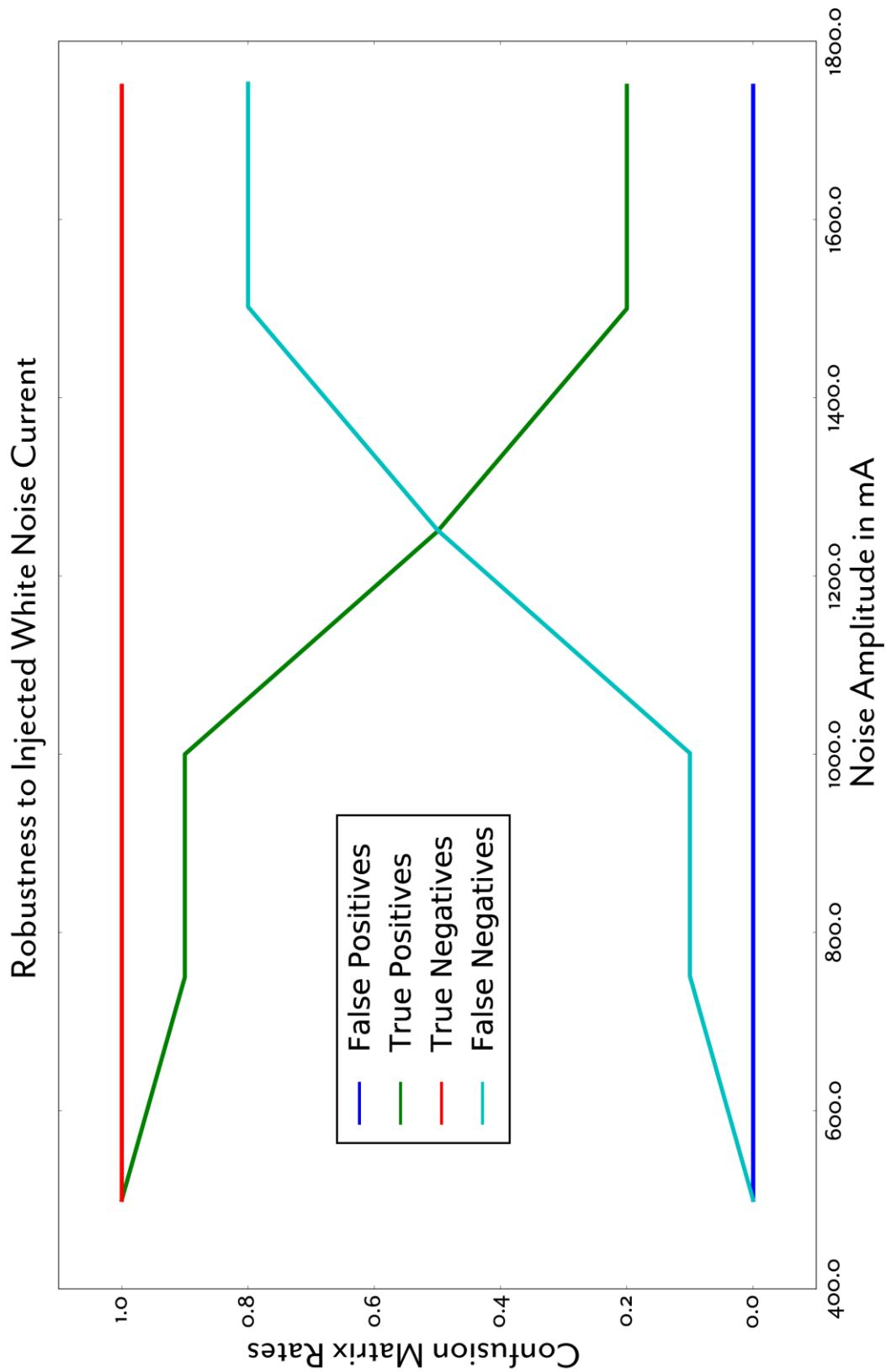
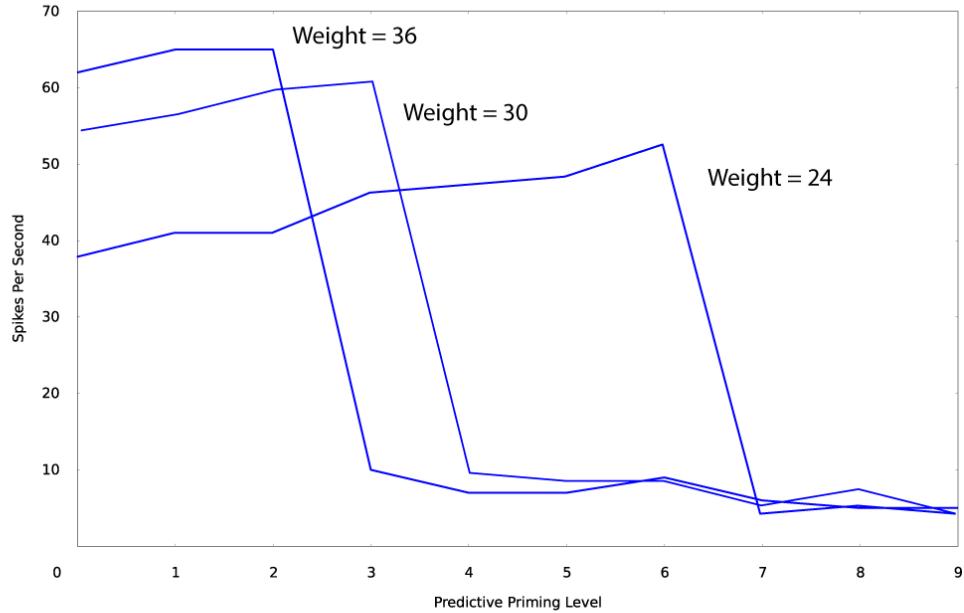


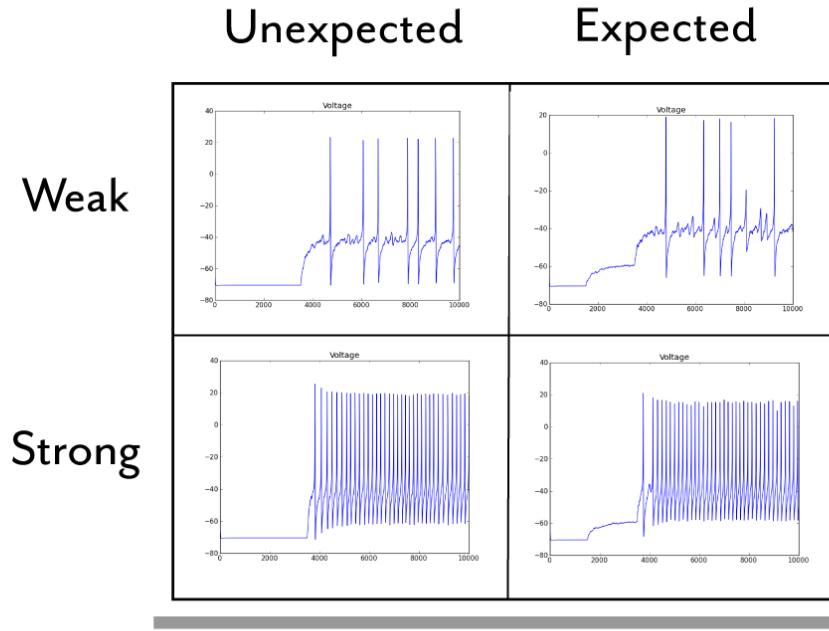
Figure 3.14: Classification performance as a function of noise. As injected white noise increases, the rate of true positives (green) falls and that of false negatives (cyan) rises. True negatives (red) and false positives (blue) remain unchanged. In this figure the true positive rate is the proportion of expected stimuli which evoke phasic firing, which would correctly be decoded to indicate an expectation was matched. The false negative rate is the proportion of expected stimuli which evoke tonic firing, as though they were unexpected. Likewise, the true negative rate is the proportion of unexpected stimuli which correctly evoke tonic firing, and the false positive rate is the proportion of unexpected stimuli which evoke phasic firing.

*Investigation 3: Confusability Testing*

a



b



Our proposed comparator with 25% M-current reduction due to ACh.

**Figure 3.15:** The robustness of our proposed model to confusability. A) The output spikes per second of our model under different levels of expectational priming input for three different stimulus intensities (represented by driving inputs with weights 24, 30, and 36, respectively). The highly nonlinear response (compared with other models in figure 4.7) helps to keep the three stimulus intensities distinct. B) Furthermore, upon mismatch,

ACh is released (Ranganath and Rainer, 2003), which drastically reduces the effect of expectational priming (here simulated as a 25% reduction in M-current conductance), allowing a target neuron to decode the intensity of a stimulus without bias due to expectations.

## Discussion

Our comparator robustly tags primed versus unprimed inputs by switching between phasic and tonic firing modes. This behavior makes it a good candidate for inclusion in future expectation/prediction-based models of the MMN. Unlike other comparators, it generates a faster response along with less activity on a primed “match”, consistent with behavioral patterns observed in studies such as (Noonan et al., 2016). Beyond regular mismatch generation, the comparator’s tagging is appropriately disrupted by NMDA blockade. Realistic membrane currents also allow our comparator to predict MMN disruption in a healthy population under high doses of KCNQ agonists but not antagonists.

The proposed model bears some similarity to, but is at the same time remarkably different, from one previous MMN model in particular. That model was based on an application of Adaptive Resonance Theory (ART) to the MMN (Banquet & Grossberg, 1987). In that rate-based formulation, matches produced increased and sustained activity even after asymptotic learning of top-down expectations, whereas in our model, such matches are indicated by brief bursts that are followed by drastically reduced spiking while the input remains on. This is quite an important difference, since “resonance in the match” was a defining feature of ART. Each model features a reset mechanism for mismatch, and a “goodness of fit” threshold wherein no match is reported if none of the predictions are matched closely enough by the input.

*Limitations of the New Model*

Of course, this model only addresses one of the metabotropic roles of ACh, so there is much work still to be done investigating the ionotropic, and other metabotropic, impacts of ACh. As always, it remains to be seen whether computational principles derived from a partial model, which simplifies dendritic morphology and lacks some known conductances, and others yet to be discovered, will survive greater realism.

*Prediction and Observation*

There exists a recent trend in neuroscience to think of the brain as a prediction machine, applying hierarchical sets of rules to predict the next input from the outside world. Other neuroscientists think of the brain as an observing and categorizing machine, taking in input, filtering it, and classifying it.

We propose that each area of the brain can act in either of those roles, and in fact they are interconnected and emerge from a single neural substrate. In order to generate a framework for thinking about the brain in this manner, we consider the constraints for each role found in the literature, and the way it interacts with the other role. We then examine fundamental units such as neurons and glial cells, and the properties they have which might cause these different roles to emerge.

We view prediction as a common cortical trait which primes a given area of the brain in anticipation of future inputs. New inputs can either be directly relayed from sensory areas, from subcortical areas of the brain such as the amygdala, or relayed from other cortical areas. EEG, single-cell recording, and behavioral experiments have all examined the difference between expected input and unexpected input, and found several constraints. This prediction role appears to emerge from the M-current, which helps convert a neuron from a rate-based coding mechanism to a spike timing-based mechanism. This current, which is not in most existing neural models, appears to be governed by the level of acetylcholine (ACh) in a given region of the brain. Thus, in our model, a low level of ACh is proposed to bias the brain towards predictions and generalizing new inputs into already existing categories, and a high level discounts the effect of the predictions and causes each part of the brain to simply observe its inputs as though they were novel, seeking out new patterns. We propose that this is a major component of attention, and is the reason for ACh's association with attentional focus. In the high-ACh state, we shift roles from prediction to observation.

#### *Future Directions*

In addition, studies have suggested important interactions between other pharmaceuticals and the MMN, including Norepinephrine (NE). Investigation of this interaction in the proposed model will require its extension, but they remain feasible.

## Appendix

### *Elements of the Model*

This neural comparator model currently consists of two point (single compartment) models with realistic synapses, bridged by a resistor. Each point model consists of all of the membrane equations in section 2, and the bridge current is computed based on their relative voltages, the axial resistance of the apical dendrite, and the distance between the apex and the base.

### *Membrane Equations*

The membrane voltage model is conductance-based, following the general convention of

$$C \frac{dv}{dt} = I_{current_1} + I_{current_2} + \dots + I_{current_n} \quad (1.1)$$

where  $I_{current_n}$  is a current of the form  $g_{current}^{max} \cdot g_{percent} \cdot (V_E - V_m)$ , where  $g_{current}^{max}$  is the maximum conductance of the channels carrying that particular current,  $g_{percent}$  is the percent of the max conductance currently achieved due to open, non-inactivated channels, and  $(V_E - V_m)$  is the driving voltage differential, between the Nernst equilibrium of that particular ion ( $V_E$ ) and the current membrane voltage ( $V_m$ ).

$$C \frac{dv}{dt} = I_{Na} + I_{AMPA} + I_{NMDA} + I_{GABA} + I_K + I_M + I_{CaK}$$

In total, the equations cover

The currents involved are as follows, each with their own  $g_{percent}$  function:

- NMDA (Na)

$$g_{percent} = \frac{g_{open}}{(1 + [Mg^{2+}] \exp(-0.062V_m) / 3.57)}$$

$$\frac{dg_{open}}{dt} = \sum_{j=1}^{C_e} (w_j \cdot s_j^{NMDA})$$

$$\frac{ds_j^{NMDA}}{dt} = -\frac{s_j^{NMDA}}{\tau_{NMDA}^{decay}} + \alpha x_j (1 - s_j^{NMDA})$$

$$\frac{dx_j}{dt} = -\frac{x_j}{\tau_{NMDA}^{rise}} + \sum_k \delta(t - t_j^k)$$

- AMPA (Na)

$$\frac{dg_{percent}}{dt} = \sum_{j=1}^{C_e} (w_j \cdot s_j^{AMPA})$$

$$\frac{ds_j^{AMPA}}{dt} = -\frac{s_j^{AMPA}}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k)$$

- GABA (Cl)

$$\frac{dg_{percent}}{dt} = \sum_{j=1}^{C_i} (w_j \cdot s_j^{GABA})$$

$$\frac{ds_j^{GABA}}{dt} = -\frac{s_j^{GABA}}{\tau_{GABA}} + \sum_k \delta(t - t_j^k)$$

- Inactivating Voltage-Based Na

$$g_{percent} = 0.5(1 + \tanh((V_m - V_1)/V_2)) \cdot (1 - h)$$

$$\alpha_h = 0.07 \exp\left(\frac{V_m}{20}\right)$$

$$\beta_n = \frac{1}{\exp\left(\frac{V_m + 30}{10}\right) + 1}$$

$$\frac{dh}{dt} = \alpha_h(1 - h) - \beta_n h$$

- Potassium

$$\frac{dg_{percent}}{dt} = \phi(w_{inf} - g_{percent}) / \tau_w$$

$$w_{inf} = 0.5(1 + \tanh((V_m - \beta_w)/\gamma_w))$$

$$\tau_w = \frac{1}{\cosh\left(\frac{V_m - \beta_w}{2 \cdot \gamma_w}\right)}$$

- Cl Leak Current

$$g_{percent} = \Gamma_{Cl}$$

- M-current, from (Mainen & Sejnowski, 1996)

$$\frac{dg_{percent}}{dt} = \frac{\frac{1}{\tau_z} - g_{percent}}{1 + exp(\frac{\beta_z - V_m}{\gamma_z}) - g_{percent}}$$

$$\tau_z = \frac{\tau_{z_{peak}}}{3.3exp(\frac{V_m - \beta_z}{20}) + exp(\frac{-\beta_z}{20})}$$

$$\tau_{z_{peak}} = \frac{1000}{2.e^{(\frac{T-36}{10})}}$$

### Parameters

Parameter	Value	Citation
$C$	$2 \frac{\mu F}{cm^2}$	(Prescott et al., 2006)
$g_{Na}^{max}$	$20 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$V_{Na}$	$50mV$	(Prescott et al., 2006)
$V_1$	$-1.2mV$	(Prescott et al., 2006)
$V_2$	$23mV$	(Prescott et al., 2006)
$w$	28	(Fernandez et al., 2005)
$V_C$	$-64mV$	(Fernandez et al., 2005)
$A$	232	(Fernandez et al., 2005)
$g_K^{max}$	$20 \frac{mS}{cm^2}$	(Prescott et al., 2006a)
$V_K$	$-100mV$	(Prescott et al., 2006)
$\phi$	0.15	(Prescott et al., 2006)
$\beta_w$	$-2mV$	(Prescott et al., 2006)
$\gamma_w$	$21mV$	(Prescott et al., 2006)
$g_M^{max}$	$4 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$g_{AHP}^{max}$	$1 \frac{mS}{cm^2}$	(Prescott et al., 2006)

$\beta_z$	-35mV	(Prescott et al., 2006)
$\gamma_z$	10mV	(Yamada et al., 1989)
$T$	36°C	(Yamada et al., 1989)
$\alpha_{AHP}$	0.005	(Prescott et al., 2006)
$\beta_{AHP}$	0mV	(Prescott et al., 2006)
$\gamma_{AHP}$	5mV	(Prescott et al., 2006)
$V_{shunt}$	-70mV	(Prescott et al., 2006)
$g_{shunt}$	$2.0 \frac{mS}{cm^2}$	(Prescott et al., 2006)
$Mg^{2+}$	$1e-3$	(Brunel & Wang, 2001)
$g_{AMPA}^{max}$	$7.5e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$g_{GABA}^{max}$	$7.5e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$g_{NMDA}^{max}$	$2e-3 \frac{mS}{cm^2}$	(Brunel & Wang, 2001)
$\tau_{AMPA}$	2ms	(Brunel & Wang, 2001)
$\tau_{GABA}$	10ms	(Brunel & Wang, 2001)
$\tau_{NMDA}^{rise}$	2ms	(Brunel & Wang, 2001)
$\tau_{NMDA}^{decay}$	100ms	(Brunel & Wang, 2001)
$\alpha_{NMDA}$	0.5ms	(Brunel & Wang, 2001)

### *Simulating electrical field effects generated by pyramidal neuron dynamics*

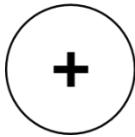
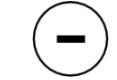
A key portion of the literature on event related potentials (ERPs) pertains to the study of mismatch signals generated within the neocortex, such as the *mismatch negativity*, hereafter the **MMN** (Näätänen et al., 2007; Winkler, 2007), which follows sensory inputs that are rare enough to mismatch current expectations for sensory input. Most studies of the MMN have been conducted via EEG. In order to fit our comparator model to the pertinent

experimental findings, we must first expand the cellular model to simulate the quasi-static electrical fields being measured in EEG studies.

The term "dipole" is frequently used to refer to the sources of electrical activity being measured in EEG. What this normally refers to is the second (dipole) term of the multipole expansion of a series of current sources and sinks, describing the component of the total field which is equivalent to one generated by exactly one source and one sink, separated in space. This component is the focus of our EEG analysis because it is the one which contributes the most to the measured fields at the distances encountered in scalp recordings.

The simplest model of a neuron capable of generating a realistic dipole field will consist of two points, separated in space. It should be noted that this is far simpler computationally than a compartmental model (even one with only two compartments), since there is no integration of space required. In this section, we develop a two-point model of a cortical pyramidal cell.

Rall (1962) plotted the dipole fields of more detailed compartmentally modeled pyramidal neurons, and demonstrated that one of the poles in the dipole term localizes to the apical section of the pyramidal cell, and the other one localizes to near the soma. In keeping with this, one point in the neuron developed here will represent the electrical activity of the apical portion of the cell and the other will represent the soma and basal dendrites.



**Figure 3.16:** Two point neurons, one representing the basal pole and one the apical. In this particular instance, the apical is depolarizing, which will produce a current sink, indicated by the negative sign, at the apical pole.

Each of these two points in our model is governed by a set of equations. All that is required to be able to complete the prediction of scalp voltage is an equation which offers  $dv/dt$  for each point. This, combined with the positions of each point within the 3D space of a cortical region (which will later be mapped to a source voxel), is sufficient to precisely calculate the dipole field. However, accurate and useful predictions will rely heavily on accurate representation of the change in voltage at each point. In addition, defining the transmembrane currents in the membrane equation is helpful in simplifying later steps, although these can in theory be derived from  $dv/dt$  and a cell's capacitance, if necessary. In the interest of demonstrating best practices, we here use a generic conductance based model. The membrane equations for each point are defined in terms of three currents, specifically an excitatory current, a rectifying current, and a shunting current:

$$\frac{dv}{dt} = I_{excite} + I_{rectify} + I_{shunting}$$

Having established the voltage change over time at each point on its own, we must now modify the equations to include a current flowing between the points, otherwise our dipole effectively remains two neurons instead of representing the two regions of a single neuron. Adding this current to the above equation, we arrive at:

$$\frac{dv_i}{dt} = I_{excite_i} + I_{rectify_i} + I_{shunting_i} \pm I_{ij}$$

where  $i$  is the point in question, and  $j$  is the opposing point of the dipole. The current  $I_{ij}$  will be added if  $i$  is the second point, and subtracted if it is the first point. Although that is just an artifact of the definition of  $I_{ij}$  below, and could validly be reversed if the current was defined in terms of the second point, rather than the first.

As mentioned above, without some sort of bridge these are effectively just two point neurons. What is needed to make them into a single dipole neuron is the axial current which connects the apical point to the basal point. This is electrically equivalent to bridging the two points with a resistor.

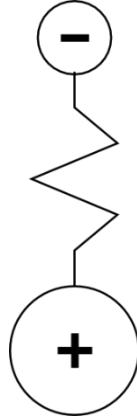


Figure 3.17: A dipole model neuron bridged with a resistor, apically depolarizing as in Figure 3.16.

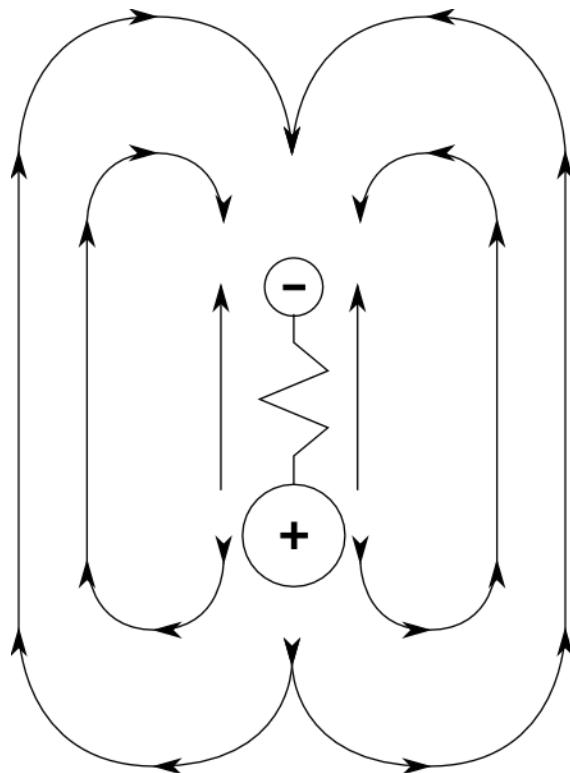
For two points separated by a distance  $d$  in meters, the specific axial resistance per unit length ( $\hat{\rho}_i$ ) is constant regardless of position (. Since resistance sums, the resistance for a cell with axial length  $\Delta x$  is  $\hat{\rho}_i \cdot \Delta x$ . For layer 2/3 neocortical pyramidal cells, the total volume resistance can be assumed to be  $100k\Omega cm$  at an apical dendrite length of 400 micrometers, per Larkman et al. (1992). This gives us  $250,000\Omega / \text{cm}$ . Therefore for a 300 micrometer apical dendrite, the axial current between location  $x$  and location  $(x + \Delta x)$  is then the voltage difference times the inverse axial resistance of the neurite connecting the two points, per the following equation from Niebur (2008).

$$I(x + \Delta x) = \frac{v(x) - v(x + \Delta x)}{\hat{\rho}_i \cdot \Delta x}$$

Substituting  $I_{ij} = I(x + \Delta x)$ ,  $v(x) = V_i$ ,  $v(x + \Delta x) = V_j$ , and the distance  $\Delta x = d$ , the current flowing into the basal point  $j$  will be:

$$I_{ij} = \frac{V_i - V_j}{\hat{\rho}_i d}$$

Having defined the primary (intracellular) current, we must now define the secondary (extracellular) currents. These are called secondary because they are indirectly evoked outside the cell (in the intracellular medium) by the primary currents. When a transmembrane current sends a flow of ions either into or out of the cell, the extracellular region experiences a negative or positive current, depending on the charge of the ions and the direction of the flow. Conservation of charge dictates that other currents flow either toward or away from this extracellular region, preventing a local buildup of charge in this space. As a result, large curls of extracellular current are evoked by the transmembrane currents. These secondary currents are the currents which drive the quasi-static electric field being measured in EEG.



**Figure 3.18:** The extracellular currents (pictured as arrows) due to the source sink formed as a result of the previously pictured apical depolarization.

The secondary current we are defining is driven by the relative difference in transmembrane current between the two points. Thanks to having employed a conductance-based model, we can now simply remove the axial current term from each of the two point membrane equations to get their respective transmembrane currents:

$$I_{transmembrane_i} = I_{excite_i} + I_{rectify_i} + I_{shunting_i}$$

Unless the two points have exactly equal  $I_{transmembrane}$  values, one point will act as a current source, and the other will act as a sink. The amplitude of this current dipole source/sink ( $s$ ), once again defined relative to the first point, will be equal to the difference between this equation evaluated for each of the two points.

$$s = (I_{transmembrane_i} - I_{transmembrane_j})$$

Alternatively, if we did not have a conductance-based model, because micro-faradays, micro-volts, and micro-amps are in play,  $dv/dt = I$ , so the  $dv/dt$  of whatever membrane equation was used can be substituted, with  $I_{ij}$  subtracted from it to produce the transmembrane current.

Now, having defined the extracellular currents, we are in a position to be able to calculate the quasi-static electric field.

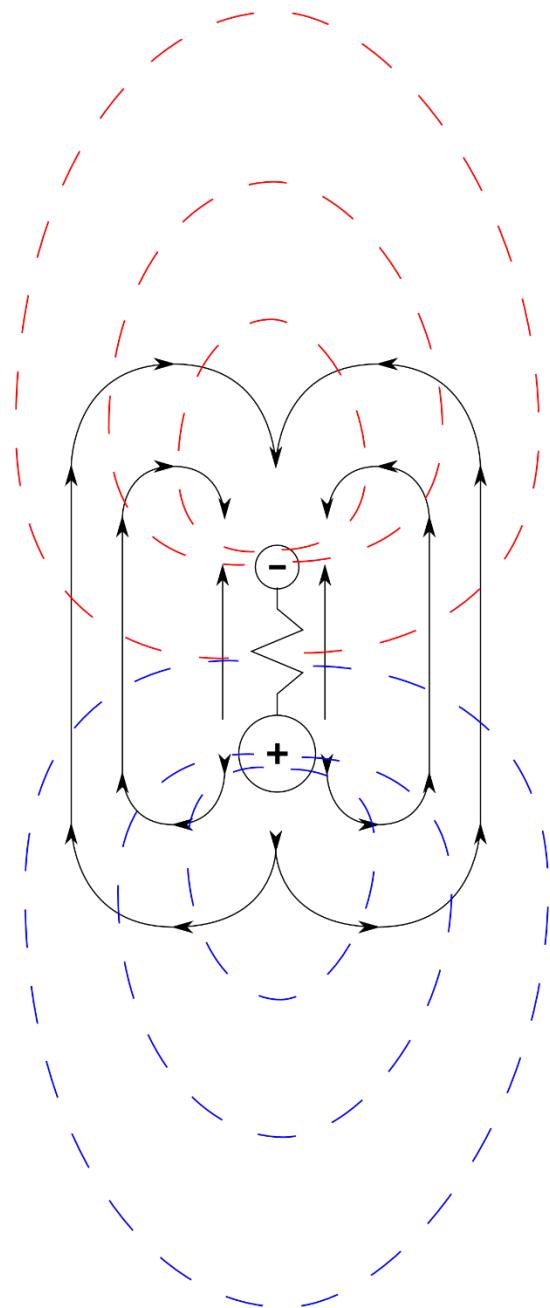


Figure 3.19: Equipotential lines for the quasi-static electric field produced by the extracellular currents which result from the apical depolarization. Red indicates positive voltage, and blue negative.

Full spherical coordinates are not necessary, because a dipole field is rotationally symmetrical, which allows us to express the voltage measured by an electrode in the field in

terms of polar coordinates  $r$  and  $\theta$ . Per equation 3.7 of Nunez and Srinivassen (2006), for two point currents of equal and opposite magnitude ( $s$ ), separated in space by a distance of  $d$  meters, the equation is as follows:

$$\Phi(r, \theta) \equiv \frac{sd \cos(\theta)}{4\pi\epsilon_0 r^2}$$

With  $\theta$  being an angle measured clockwise from the "top" of the polar coordinates, and  $r$  being the radial distance from the center of the two-point system.  $\epsilon_0$  is a constant known as the permittivity of empty space, and is here assigned the value  $8.854187817 \times 10^{-12} F \cdot m^{-1}$ .

## CHAPTER FOUR: A MECHANISTIC MODEL OF ALPHA-INDUCED INFORMATION SUPPRESSION

### **Significance Statement**

In EEG, assorted functions have been ascribed to oscillations depending on their frequency and location, with a common suggestion that coherence between oscillators in two different brain regions indicates communication between those regions. However, alpha (8-14 Hz) oscillations increase in amplitude when information transfer is inhibited, raising the possibility that coherent alpha oscillations suppress communication. In this paper we propose a model of such alpha suppression. We show via computer simulations that when circuits are driven to oscillate near their 8-14Hz resonances neural activity can be unresponsive to stimuli, instead dominated by large amplitude oscillations. Such saturating oscillations suppress communication as measured by mutual information between a signal and downstream activity, offering the first successful mechanistic model of alpha suppression.

### **Abstract**

Suppression of information transfer by alpha (8-14 Hz) oscillations, a phenomenon known as "alpha induced information suppression", has been demonstrated in both

correlative and causal experiments. However, despite extensive experimental work, an explanation for alpha suppression in terms of cortical activity and neural responses has yet to be established. Here we propose such an explanation based on the behavior of simulated models of coupled resonant neural circuits. When constant drive causes the thalamic resonant frequency to match the 10 Hz resonance of the cortical oscillator, the amplitude of oscillations in the coupled thalamo-cortical circuit increases. This increase causes the cortex to transition from occasional input-correlated, information-rich bursts of activity into periodic, uninformative suprathreshold activity. Alternating saturation and inhibition of the high-amplitude oscillations, as in the Inhibition Timing Hypothesis (Jensen & Mazaheri, 2010; Klimesch, Sauseng, & Hanslmayr, 2007), reduces the impact of incoming signals on cortical excitatory population activity. The ensuing reduction in communication, measured by mutual information between sensory input and cortical response, when cortical oscillations are in the alpha range, agrees with results of transcranial magnetic stimulation. The significant phase-dependence of mutual information when oscillations are in the low alpha/high theta range, replicates EEG and behavioral results from Busch, Dubois, & VanRullen, (2009). Finally, we address an implicit assumption underlying many functional network studies, that coherence implies communication between cortical regions. We find that relative phase, as well as coherence, determines communication between two modeled cortical regions that are oscillating in the theta range, whereas high-amplitude alpha oscillations are reliably detrimental to communication.

## Introduction

Intentional suppression of task-irrelevant stimuli during sustained attention is associated with increased oscillations in the 8–14 Hz cortical alpha band (Freunberger, Fellinger, Sauseng, Gruber, & Klimesch, 2009; Payne, Guillory, & Sekuler, 2013; Worden, Foxe, Wang, & Simpson, 2000). It is thought that such attentional processes entail not only a decrease in alpha activity over regions of active encoding, but also an increase in regions whose function needs to be reduced. In this view, alpha oscillations reflect an active inhibitory mechanism (Foxe & Snyder, 2011; Klimesch et al., 2007).

In 2007, Klimesch et al. proposed the Inhibition Timing Hypothesis, a modification of the alpha suppression hypothesis that includes specific phases at which alpha would enhance, rather than suppress, information processing. In accordance with this hypothesis, the particular phase of stimulus onset was found to be important in the reliability of target detection (Busch et al., 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009) and the reaction time in perceptual discrimination tasks (Vanrullen, Busch, Drewes, & Dubois, 2011).

Romei and colleagues subjected the alpha suppression hypothesis to a direct test, using transcranial magnetic stimulation (TMS) to induce pre-stimulus oscillations in the right or left parietal-occipital cortex. By varying the frequency of oscillations during a near-threshold target detection task, they found that target detection was impaired by oscillations in the alpha band, but not by oscillations in other frequency ranges (Romei, Gross, &

Thut, 2010). This provided evidence of a causal link between alpha oscillations and information suppression.

The most commonly discussed mechanism of alpha information suppression is pulsed inhibition (Jensen & Mazaheri, 2010). This takes evidence implicating GABAergic interneurons in alpha production (S. R. Jones, Pinto, Kaper, & Kopell, 2000; Lorincz, Kékesi, Juhász, Crunelli, & Hughes, 2009) and uses it to explain the apparently phase-dependent nature of alpha information suppression (Busch et al., 2009; Mathewson et al., 2009; Vanrullen et al., 2011). In this paper, we offer a mechanistic computational model consistent with this mechanism.

Our mechanistic model is motivated also by theoretical studies on coupled nonlinear oscillators. The effect of any interaction between coupled oscillators depends strongly on their relative phase, due to phase resetting (Rinzel & Ermentrout, 1998), making communication between oscillators phase-dependent. Moreover, for nonlinear oscillators, responses to perturbations—such as ones arising from interaction with another oscillator or from sensory input—typically depend on the amplitude of oscillation, which itself depends on the oscillation frequency. In particular, as thalamo-cortical neural circuits possess a resonance in the alpha range (Gutfreund, Yarom, & Segev, 1995; Herrmann, 2001) and neural responses to input require a balance of excitation and inhibition (Bell, Mainen, & Sejnowski, 1994), we expect that the neural response to input will be increasingly impacted by saturation/inhibition when the neural circuit is driven to oscillate in the alpha range.

Finally, a growing number of recent studies have attempted to define coherence-based networks of “functional connectivity” in the brain. These studies have identified coherence networks in the alpha, beta, and gamma bands (Hipp, Engel, & Siegel, 2011; Palva, Monto, Kulashekhar, & Palva, 2010). In this paper, we extend our model to include two cortical regions, in order to examine an implicit assumption underlying such studies, that greater oscillatory coherence implies greater information transfer.

## **Materials and Methods**

We propose a mechanism of alpha information suppression that is based on alpha-band resonance of the thalamo-cortical circuit (Herrmann, 2001; VanRullen & Macdonald, 2012). To simulate the thalamo-cortical resonance, we used a simple model excitatory-inhibitory feedback circuit, which oscillates with a natural frequency of 10 Hz. The model thalamus receives variable tonic drive that we consider as frontal input from higher order cortical areas (per Guillory and Sherman, 2002). Nonlinearities in the system cause the oscillation frequency of the thalamic circuit to vary with the tonic external drive. We use this result to adjust thalamic oscillation frequency across its observed range from 2–13 Hz (S. W. Hughes & Crunelli, 2005). In our model, this range emerges from a rapid exponential rise in the amount of frontal input required to shift the frequency. The coupling of thalamic and cortical oscillators in a loop makes the frequency of cortical oscillations also depend on the tonic drive to the thalamus. The coupled circuit possesses a resonance, which appears when the tonic drive to the thalamus causes the thalamic circuit’s oscillatory frequency to match the 10 Hz natural frequency of the cortical circuit. We investigate the consequence of such an alpha resonance by simulating three tasks, which we

summarize below. Details of the model's implementation and the task-specific inputs follow the brief description of each task.

Our first investigation tested how the mutual information between a cortical input signal and the cortical activity depended on the frequency of thalamo-cortical oscillations. A continuously varying signal was sent to the cortex while the thalamus was driven with different levels of frontal input. The frontal input to the thalamus was systematically varied in order to move the thalamus through its natural range of bursting oscillations (2-13 Hz, per Hughes & Crunelli, (2005)). At each level of input (and therefore frequency), mutual information was evaluated between the signal and the cortical spike rate.

In the second investigation, we assessed how cortical responses depended on the phase of an input, aiming to reproduce the experiment reported in Busch et al., (2009). Our model was simulated with continuously varying noise input to the cortex, while sufficient frontal input to the thalamus drove the system into the high theta range. In addition to the varying noise, a 12 ms pulse was injected at various phases of the thalamo-cortical oscillation. This pulse simulated the presentation of a brief threshold-level visual stimulus. The model's response to the stimulus pulse was determined by a simple threshold classifier (see Analysis), and behavioral statistics were calculated as in Busch et al.'s original work.

The third investigation examined the role of coherence and the relative phase of oscillations in communication between two connected instances of our model thalamo-cortical loop. This arrangement was meant to represent two separate but connected

thalamo-cortical regions that processed incoming information. In both the alpha and theta frequency bands, the model thalamic nuclei were driven so as to oscillate either coherently or incoherently (at different frequencies within the band). In the coherent condition, the thalamic oscillations were simulated with varying phase delays relative to each other, allowing us to examine the impact of these two parameters on communication of a continuously varying signal between two oscillating cortical circuits. Finally, the conduction delay between the two model systems was varied, and the optimum phase for communication was found for both short and long delays.

## Model

In describing our model, we follow the standards proposed by (Nordlie et al., 2009), documenting the network architecture, the connectivity, the neuron and synapse models used, the input (stimuli), and the output (data recorded). Our model was implemented in two forms: one for investigations one and two, and the other in an expanded form for investigation three.

### *Network Architecture*

The basic modeled network (Figure 4.1) consists of one thalamo-cortical loop, with the thalamus and cortex each containing two populations of neurons.

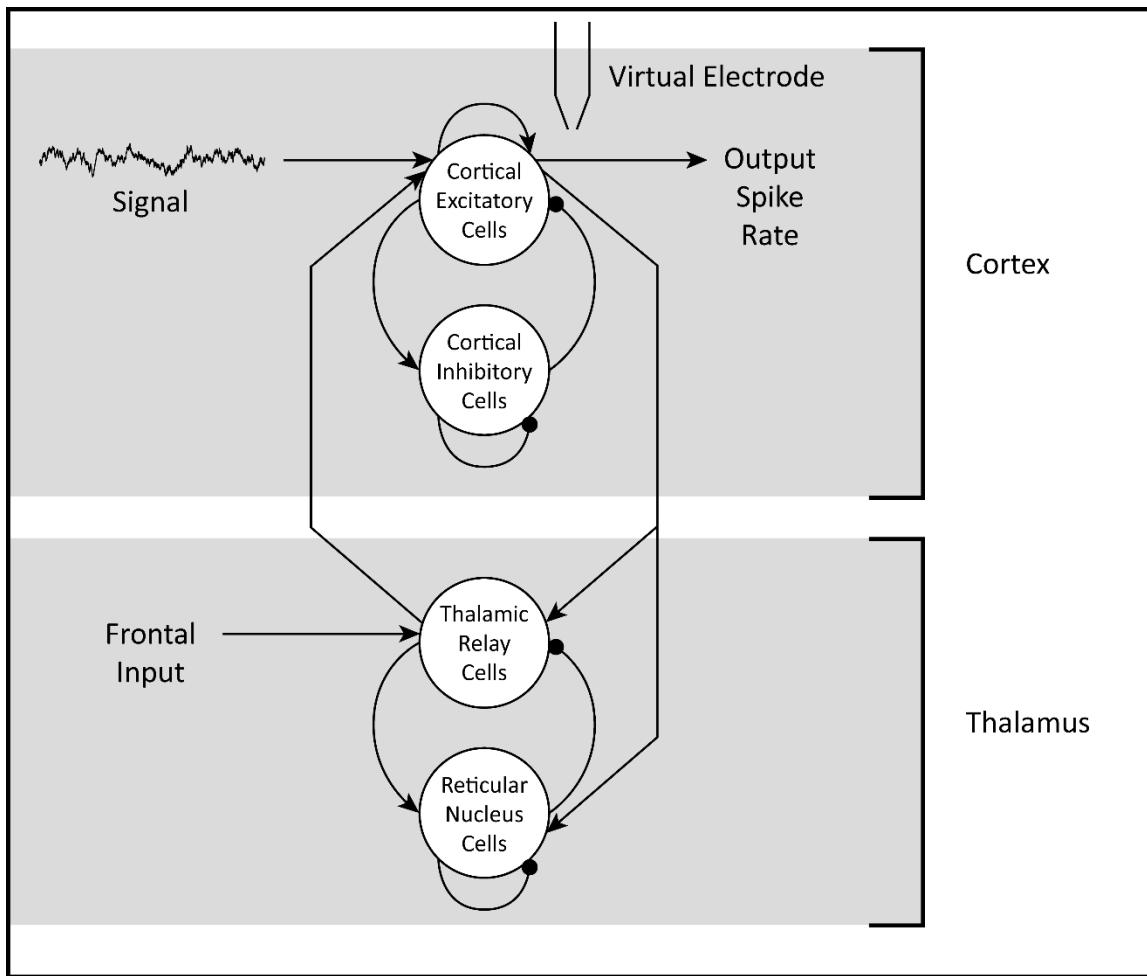


Figure 4.1: Our model circuitry. Two brain regions (the thalamus and cortex) are modeled, each with populations of excitatory and inhibitory neurons. Arrows represent excitatory connections, circles inhibitory.

The model thalamus consists of an excitatory thalamic relay nucleus (a population of thalamic relay cells) and its corresponding section of inhibitory thalamic reticular nucleus (TRN, consisting of a population of reticular nucleus cells). The model cortical area contains a population of excitatory projection neurons and a population of inhibitory interneurons.

For our third investigation of this model (investigation descriptions to follow), we expanded the model by duplicating the network architecture in full, creating two distinct thalamo-cortical loops, to be linked via a cortico-cortical connection from the excitatory projection neurons in Area I to the equivalent population in Area II. (Figure 4.2) Feed-forward inhibition was omitted from the model, because the inhibitory cells that are targets of long range cortical connections do not generally have reciprocal connections with the excitatory neurons we are modeling (Apicella, Wickersham, Seung, & Shepherd, 2012).

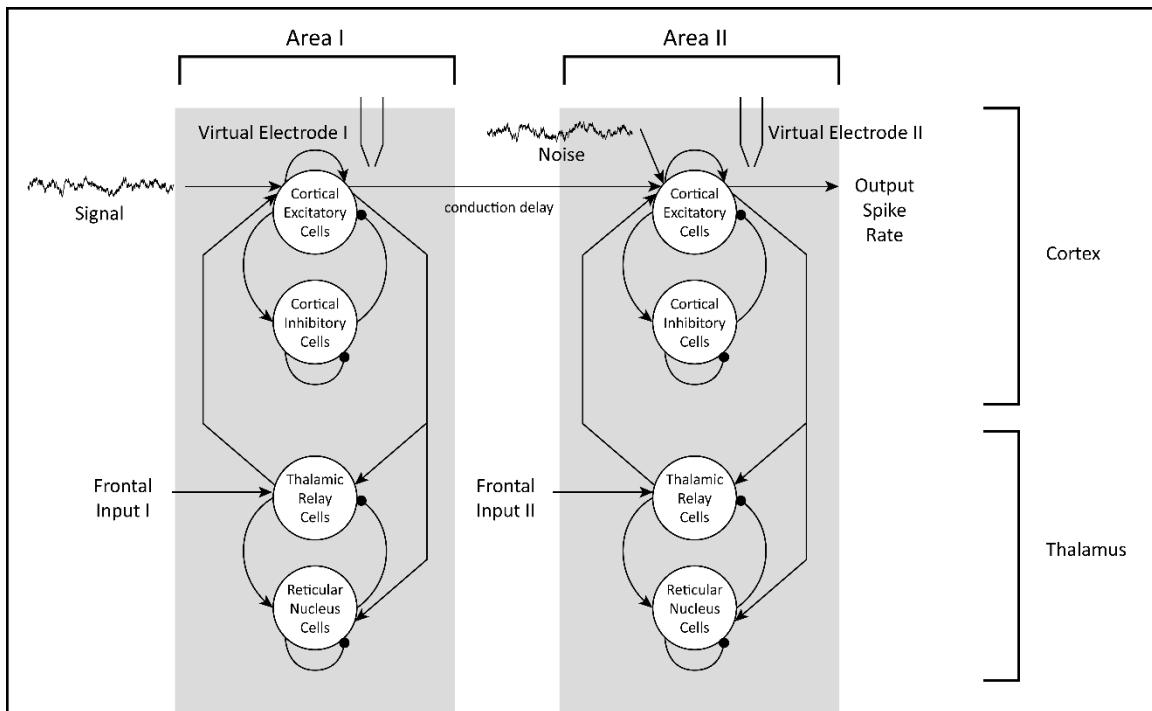


Figure 4.2: The expanded model used in investigation three. Two copies of the first model were connected via a unidirectional excitatory connection between the cortical excitatory populations. The "signal" entering the second population is considered noise, for purposes of mutual information calculation.

### *Network Connectivity*

This model defines synapses between neurons via a pair of parameters: a weight  $W_{ij}$ , and a conduction delay  $d_{ij}$ . These are enumerated for each simulated neural group in Table X.

Connection	Weight $W_{ij}$	Conduction Delay $d_{ij}$	Sources
Thalamic projection neurons to TRN interneurons	12	0.25 ms	Jones, 2002
Thalamic projection neurons to cortical excitatory neurons	3.5	4 ms	Sherman and Guillery, 2004
The thalamic projection neurons' calcium current self-excitation	1.0	0 ms	Huguenard, 1996
TRN interneurons to thalamic projection neurons	-7.0	0.25 ms	Sherman and Guillery, 2004
TRN interneurons to TRN interneurons	-1.75	0.25 ms	Destexhe et al., 1998
Excitatory cortical to cortical inhibitory	5.0	0.25 ms	Destexhe et al., 1998

Excitatory cortical to excitatory cortical	8.0	0.25 ms	Destexhe et al., 1998
Excitatory cortical to thalamic reticular nucleus	2.75	4 ms	Jones, 2002
Excitatory cortical to thalamic projection neurons	2.0	4 ms	Destexhe et al., 1998
Inhibitory cortical to excitatory cortical	-7	0.25 ms	Destexhe et al., 1998
Inhibitory cortical to inhibitory cortical	-0.75	0.25 ms	Destexhe et al., 1998
Inter-area cortical excitatory to cortical excitatory	1.0	12 ms or 60 ms	Apicella et al., 2012

Table 4.1: Model weights and conduction delays, with evidence for each type of connection.

### *Neuron and Synapse Models*

Our model is constructed out of rate-based populations of neurons whose responses track a sigmoidal firing rate vs. input ( $f$ - $I$ ) curve and are described by a simple differential equation. Synapses are current-based. With the exception of the thalamic excitatory population, the equations for each neural population  $i$  in the model are as follows:

$$I_i = \sum_j r_j(t - d_{ij}) W_{ij}$$

$$r_i^\infty = r_i^0 + \frac{(r_i^{max} - r_i^0)}{(1 + e^{(-I_i - I_i^{threshold})/I_i^{sigma}})}$$

$$\frac{dr_i}{dt} = \frac{(r_i^\infty - r_i)}{\tau_i},$$

where  $I$  is the summed current-based input to the population across all inputs  $i$ ,  $r_j(t - d_{ij})$  is the firing rate of each input source at a time delay of  $d_{ij}$ , and  $W_{ij}$  is the weight of input  $i$  to population  $j$ . In the second equation,  $r_i^\infty$  is the steady-state population activity rate if inputs were held constant,  $r_i^0$  the minimum possible, and  $r_i^{max}$  the maximum possible firing rate.  $I_i^{threshold}$  represents a soft threshold (the current required to reach the midpoint of the f-I curve), and  $I_i^{sigma}$  is inversely proportional to the slope of the sigmoidal f-I curve. Finally,  $r_i$  represents the instantaneous firing rate of the population, which chases  $r_i^\infty$  at a rate determined by the time constant  $\tau_i$ . These values are set for the different simulated neural groups as listed in Table 2.

Population $i$	$r_i^0$	$r_i^{max}$	$I_i^{threshold}$	$I_i^{sigma}$
Thalamic Excitatory	0	70	50	10
Thalamic Inhibitory	0.025	200	100	10
Cortical Excitatory	0	70	95	25

Cortical Inhibitory	5	200	175	15
---------------------	---	-----	-----	----

Table 4.2: Population parameters

The thalamic excitatory population uses the following equations, modified to account for the self-excitation of the calcium currents found in projection neurons:

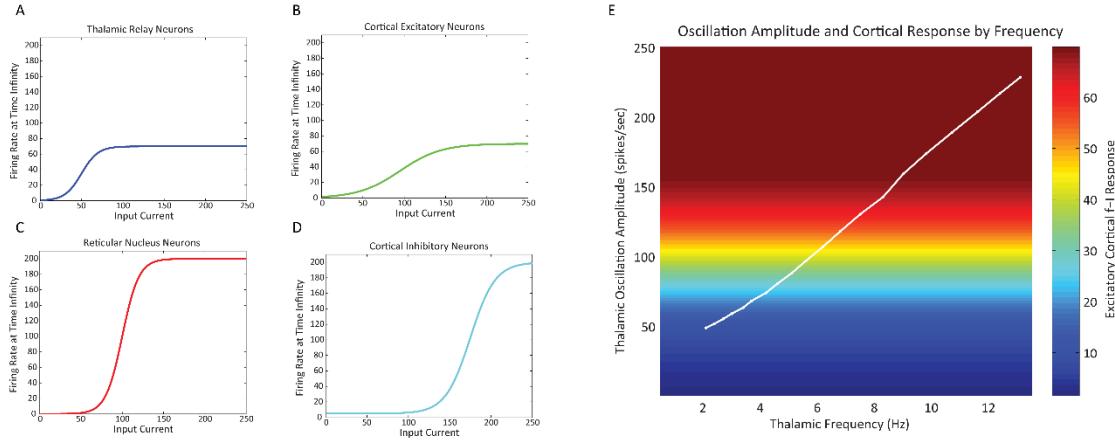
$$I_i^{in} = \sum_j a_j(t - d_{ij}) W_{ij}$$

$$\tau_i^{ca} \cdot \frac{dI_i}{dt} = -I_i + I_i^{in} + W_i^{ca} I_i$$

$$r_i^\infty = r_i^0 + \frac{(r_i^{max} - r_i^0)}{(1 + e^{(-I_i - I_i^{threshold})/I_i^{sigma}})}$$

$$\frac{dr_i}{dt} = \frac{(r_i^\infty - r_i)}{\tau_i}$$

The terms are defined as in the other populations above, except that  $I_i$  now represents the combination of the synaptic input current summed with the calcium current.  $I_i^{in}$  represents the summed synaptic input current,  $W_i^{ca}$  represents the weight of the self-exciting thalamic calcium current, and  $\tau_i^{ca}$  is the time constant of the thalamic calcium current. For thalamic excitatory populations,  $W_i^{ca} = 1$ , and  $\tau_i^{ca} = 0.015$ . The combination of the above equations and the parameter values found in Table 2 produces the following f-I curves for the 4 types of neuron involved in our simulation, as shown in Figure 4.3.



**Figure 4.3: Response curves and saturation of different model neuron types.** A–D) f–I curves for the four neural populations, showing firing rate as a function of input current. E) Oscillation amplitude and cortical response as a function of frequency. The colored background indicates the steady state cortical excitatory response amplitude when it is driven by input of a given frequency (x-axis) and given peak amplitude (y-axis). The actual peak amplitude of thalamic drive as a function of frequency is represented by the white line.

### *Input (Stimuli)*

The model system has two input sources. The first is a signal input to the cortical pyramidal population. This represents information coming either from sensory sources or other brain regions. The second source of input is frontal drive to the thalamic projection nucleus. The magnitude of this latter input largely determines the frequency and amplitude of the thalamo-cortical oscillations. We use different relative onset times of the frontal drive to the two thalamic regions to alter the relative phase of the oscillators in the third investigation (below).

For the first investigation, the signal consisted of a pseudo-random continuous current input, generated by an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930), and varying from roughly -50 to 50 mA. The specific equation used was as follows:

$$dx_i = \theta(\mu - x_i)dt + \sigma dW_{ij}$$

The parameters of the process were  $\theta = 10$ ,  $\mu = 0$ , and  $\sigma = 90$ , and the final result was scaled by a factor of 50 to achieve the desired current range. The frontal drive in this investigation was a constant level of current input over each simulation, with the amplitude varying systematically between simulations from 45 to 330, with the onset delay of the frontal input being held at zero.

The second investigation used an identical pseudo-random continuous input as noise, summed with a single 25 amplitude square "pulse" 6 ms in duration. The pulse was delivered at varying phases of the thalamo-cortical oscillation. For this investigation, frontal drive was fixed at 170 mA, which produced oscillations in the high-theta range. The onset delay of the frontal input was once again held at zero.

In the third investigation, two pseudo-random continuous inputs were generated for each simulation. One served as the signal input for the first system, and the second one served as "irrelevant" or "noise" signal input to the second system, representing input it was receiving from other areas beyond the scope of the model. The parameters for these two Ornstein-Uhlenbeck processes were the same as above. The two frontal drives in this investigation were varied between 140 (low theta), 170 mA (high theta), 200 mA (low alpha), and 300 mA (high alpha), and for each of these input amplitudes, the onset delay of frontal input to the second system was varied in 36 steps between the duration of one and two cycles of the oscillation (with one cycle being the minimum in order to allow the

system to relax from its initial conditions). This provided 36 relative phases separated by 10 degrees.

#### *Output (Data Recorded)*

Spiking rate data was recorded from the excitatory population of cortical projection neurons, as well as from a simulated virtual electrode recording the local field potential (LFP) generated by the summed synaptic inputs to these same neurons. In the third investigation, there were two virtual electrodes simulated, one for each of the thalamo-cortical systems involved. Each electrode was simulated through the simple method of summing excitatory and inhibitory current inputs to each cell population, and multiplying by -1 so that excitatory inputs produce a negative deflection in simulated voltage.

#### **Analysis**

Data analysis was conducted in Matlab 2012a (The MathWorks, Natick, MA). For all three investigations, we used a fast Fourier transform to analyze the cortical excitatory population's local field potential (*i.e.* summed synaptic input) to determine oscillatory power over a time-frequency spectrum. We also classified the thalamic oscillations based on mean distance between peaks.

For the first investigation, at each 10 second step of frontal drive to the thalamic nucleus, the impact of the signal input on the spiking rate output data was computed via mutual

information (Dayan & Abbott, 2001). Mutual information (MI) allows us to determine how much variability in the receiver's response is due to the activity of the sender.

Mathematically, MI is equal to the receiver's full response entropy minus its noise entropy.

$$I_m = \int ds \int dr p[s]p[r|s] \log_2 \left( \frac{p[r|s]}{p[r]} \right)$$

MI can be thought of as a measure of how well variations in the sender are correlated with variations in the receiver, beyond the ability of the receiver to predict its own variability

We evaluated mutual information by binning our signals into discrete sender and receiver states, using a number of bins determined by Scotts Normal Reference Rule (Scott, 1979), a rule originally designed for optimizing bin count in histograms. It suggests bins of width

$\frac{3.5\hat{\sigma}}{n^{1/3}}$ , where  $\hat{\sigma}$  is the standard deviation of the sample, and  $n$  is the number of data points

in the sample. We then approximated the above integrals by sums over bins  $r_i$  and  $s_j$ , with

$p([r_i | s_j])$  being estimated by  $N_{ij}$ , which is the number of occurrences of rate  $i$  given

stimulus  $j$ , divided by the number of occurrences of stimulus  $j$ . This evaluation was

repeated over a range of time delays up to 40 ms, and we selected the delay that yielded the largest MI value.

We also corrected for the number of samples, since a less than infinite number of samples can cause bias in the mutual information estimation. This was corrected for by determining the number of elements in the response space  $\bar{R}$  and the number of data points  $n$ , then

subtracting  $\frac{\bar{R}-1}{2n \log(2)}$ . (Panzeri, Senatore, Montemurro, & Petersen, 2007).

The simulation was repeated 10 times under new random seeds. A Lilliefors test showed that the mutual information measurements during alpha oscillations were not normally distributed. We therefore ran a nonparametric Wilcoxon rank sum test in order to check for a significant difference in means between mutual information measured during alpha and theta oscillations.

We also defined a responsiveness metric, designed to capture how well the oscillating cortical neurons responded to changes in input. It is equal to the average of the variance in extracted peak heights across all runs, divided by the variance in the input signal. As with mutual information, we investigated the normalcy of our responsiveness results. A Lilliefors test showed that the alpha responsiveness measurements were not normally distributed. We therefore again ran a nonparametric Wilcoxon rank sum test. For the second investigation, mutual information was not calculated, and instead a threshold classifier sorted trials into hits, where activity during the cycle when the stimulus pulse occurred surpassed or equaled the threshold level of 40 Hz; and misses, where activity during the cycle did not reach 40 Hz. Hit and miss rates were calculated for each phase angle of stimulus presentation. False positives obviously could not be sorted based on stimulus phase, since there was no stimulus during these trials. False alarm rates were, however, calculated for all non-target trials.

In the third investigation, mutual information was once again calculated, as above, but between the input signal to a first cortical oscillator and the second cortical oscillator's

excitatory firing rate trace. This was repeated for each relative phase, at each frequency or pair of frequencies, and for each conduction delay.

## Results

### *Investigation 1: Mutual Information, Single Thalamo–Cortical Loop*

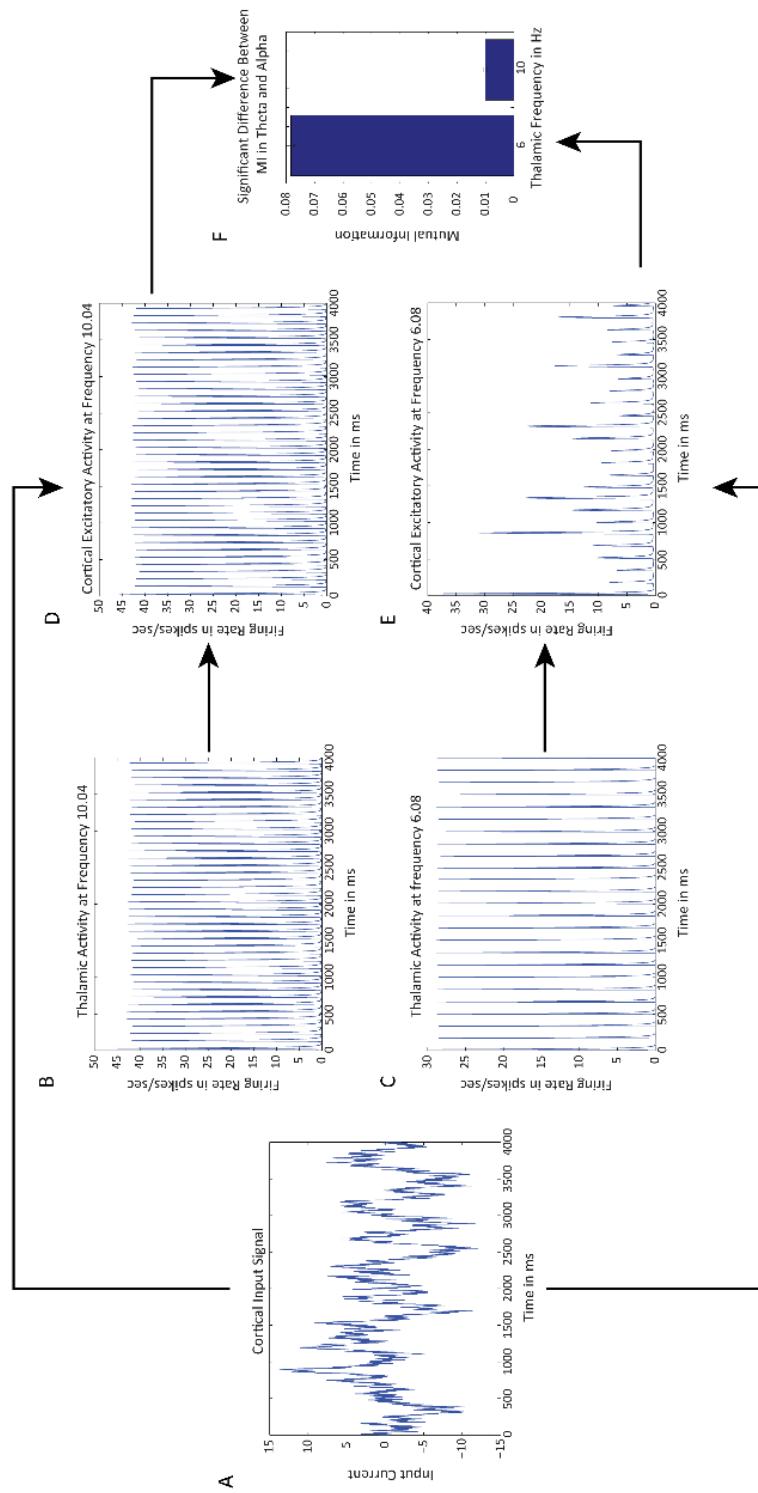
The results from the first investigation show a clear frequency-dependent reduction of mutual information when the thalamo-cortical circuit oscillates in the alpha range (10Hz). When the amplitude of periodic drive to the cortical pyramidal cells increases, the cortical activity transitions from a sequence of bursts correlated with its input signal into a regular oscillatory mode, drastically reducing the available information about the signal.

The rank-sum test rejected the null hypothesis of equal medians ( $p = 6.80e^{-8}$ ), demonstrating a significant difference between mutual information in the 6.1 Hz group (representing oscillations at theta frequencies) and the 10 Hz group of trials. Our responsiveness metric had a mean of 2.4 across the ten runs of the simulation driven at alpha frequencies, and a mean of 4.3 across the ten runs driven in theta. The rank-sum test run on responsiveness metrics also rejected the null hypothesis of equal medians ( $p = 6.80e^{-8}$ ).

Additionally, the results demonstrate the behavior we had hypothesized at each stage of the process. Higher frontal input to the thalamus produces a higher frequency oscillation in the thalamus as well as greater thalamic peak amplitude. This oscillatory thalamic behavior, when combined with the Ornstein-Uhlenbeck input signal, causes a switch from mostly subthreshold (but occasionally suprathreshold) cortical oscillations, whose amplitude is

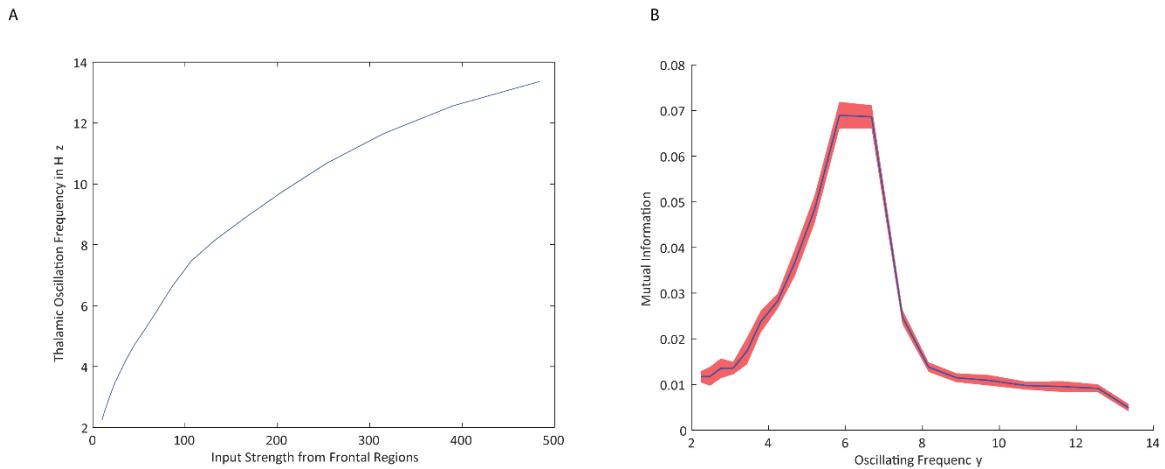
largely determined by the amplitude of the signal, to saturated suprathreshold cortical oscillations, which are largely invariant with regard to the signal amplitude. The switch to saturated suprathreshold cortical oscillations leads to a significantly lower mutual information.

The frequency of oscillation in our model thalamo-cortical loop depends on the level of frontal input (Fig. 4.5A). The dependence is weak—only slightly stronger than logarithmic (note the log-scale of the abscissa in Fig. 4.5A)—so as the frontal input varies in strength across two orders of magnitude, the range of bursting frequencies produced in our model matches the range observed in rat LGN slices, where increased depolarization leads to increasing frequencies between 2 and 13 Hz (S. W. Hughes & Crunelli, 2005).



**Figure 4.4:** Simulation Results demonstrating A) the Ornstein–Uhlenbeck-generated input signal, B–C) the thalamic activity at roughly 10 and 6 Hz, respectively, D–E) the cortical activity driven by a combination of the input signal and the thalamic activity, and F) the significant difference in mutual information between the 6.1 Hz and 10 Hz conditions.

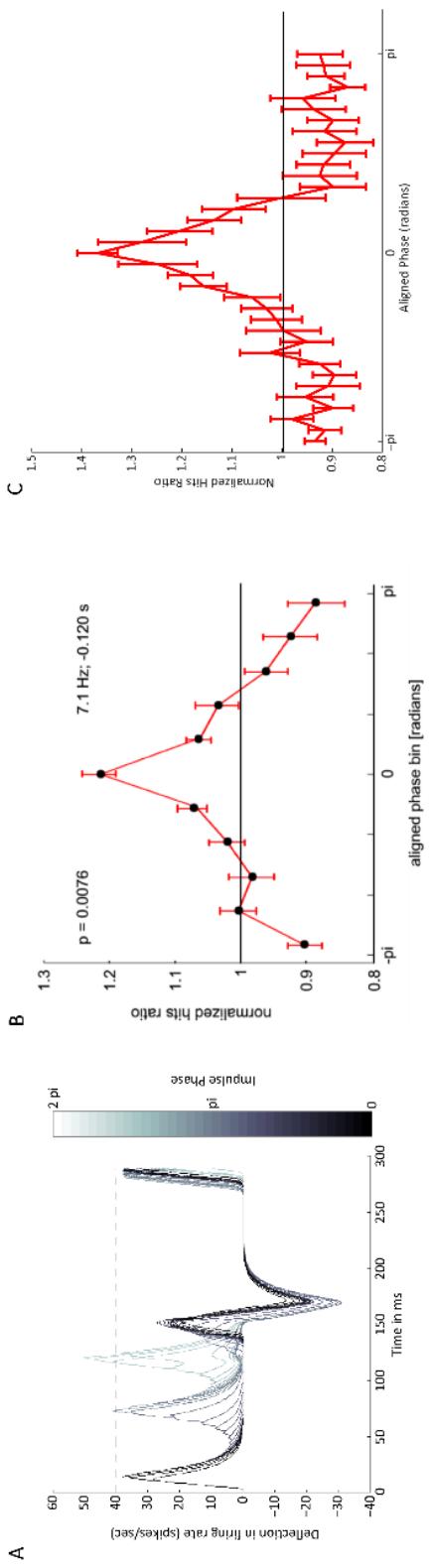
In addition, the mutual information varies non-monotonically, in an inverted-U, as a function of the oscillation frequency (Fig. 4.5B) and thus as a function of the level of frontal input. Communication (as measured by MI) is low when the thalamic oscillations are below the theta range, with an amplitude so low as to render the cortex nearly unexcitable by the signal. As it enters the theta range, thalamic amplitude is in a state of balance near the firing threshold, allowing for improved communication between signal and cortex. And finally, across the full alpha range the thalamic amplitude is too high and the entire cortical response is functionally saturated, drastically reducing MI once again.



**Figure 4.5:** Particular properties of the model system. A) The thalamic frequency saturated as frontal input to the relay nucleus cells increased. B) The mutual information varied with frequency, increasing in the theta range and falling to very low levels below theta and in alpha.

*Investigation 2: Discrimination Threshold in a Single Thalamo-Cortical Loop*

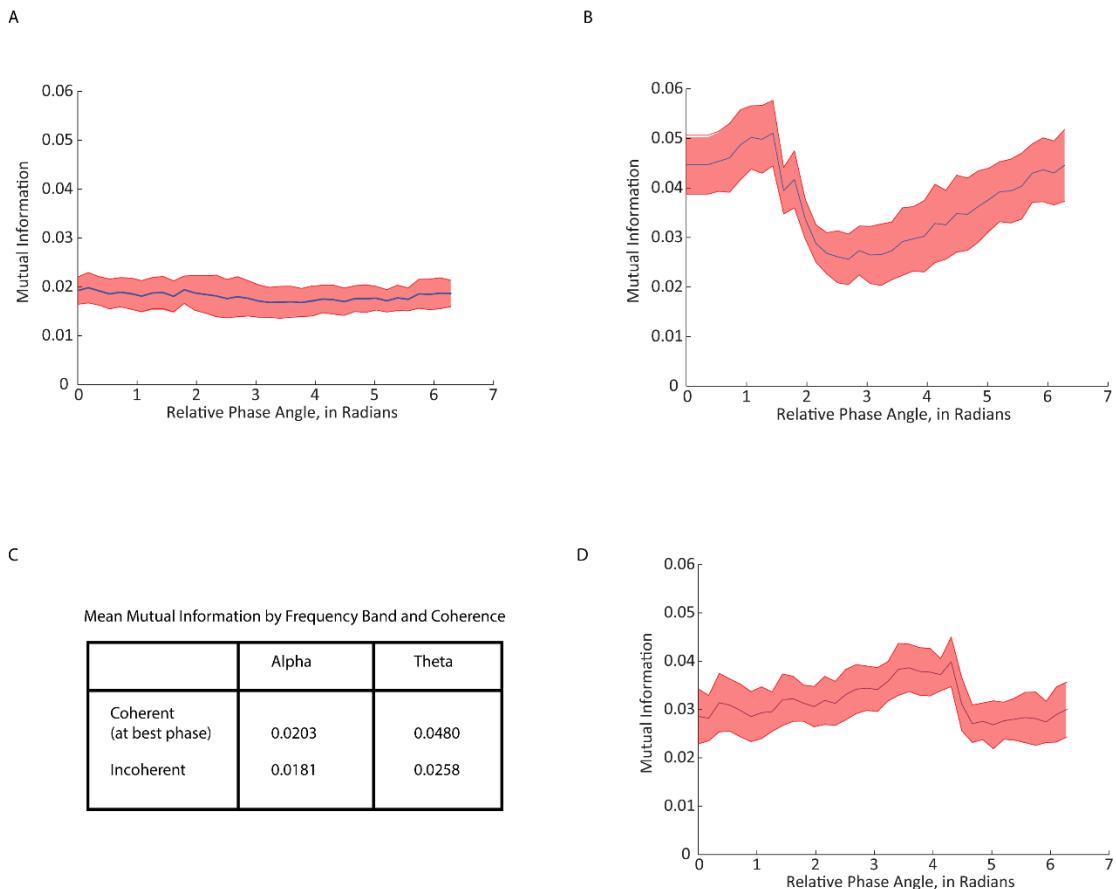
As shown in Figure 4.6, results of our second investigation were in close agreement with those reported by N. A. Busch, Dubois, & Vanrullen, (2009). For example, Fig. 6B shows that the deflection in cortical activity produced by a single stimulus pulse is strongly dependent on the phase of stimulus presentation with respect to the ongoing cortical oscillation. As can be seen in Figure 6a, when multiple phases of impulse were applied using the same (frozen) background noise, both the initial and later oscillatory impact of each stimulus impulse varied with the phase, some of them falling below the threshold of the classifier (40 spikes/sec), and others surpassing the threshold. Also, the threshold classifier demonstrated extremely strong phase dependence. When the hit rate (rate of trials classified "stimulus present" out of all trials with a stimulus impulse) is plotted as a function of phase as in Figure 4.6B, the preference for an impulse just preceding the onset of the cycle (a peak in the oscillation) becomes clearly visible.



**Figure 4.6: Deflection and hit rate by phase.** A) Cortical deflection from baseline during a cycle, as a function of impulse phase. B) Experimental normalized hits ratio by aligned stimulus onset phase (reproduced from Busch et al., 2009). C) Model normalized hits ratio by aligned stimulus onset phase.

### *Investigation 3: Mutual Information between Two Thalamo–Cortical Loops*

In our third investigation, simulation of two coupled thalamo–cortical oscillators offered unique insights into the interpretation of coherence-based functional network analysis of EEG. We first attempted to establish the preferred relative phases between two thalamo–cortical oscillators, as measured by MI. The model did not exhibit a significantly preferred phase for the alpha band, as shown in Figure 4.7A, but Figure 4.7B shows that the model exhibited a preferred phase in theta.



**Figure 4.7:** The results of investigation 3, using two coupled thalamo-cortical loops. A) The mutual information by relative phase angle at roughly 10 Hz, demonstrating the lack of a preferred phase in the alpha band, with all MI values below 0.023. B) The mutual information by relative phase angle at roughly 6 Hz, demonstrating the strongly preferred phase. C) Mean mutual information is reduced by a lack of coherence in the oscillations, and the reduction is much greater (-0.022) in theta, where there is more best-phase MI to lose than in alpha (-0.002). D) The preferred phase from B can be altered by increasing the conduction delay from 12ms to 60ms between the two regions.

We then measured the impact of coherence vs. incoherence between the two oscillators.

Coherent mutual information was measured at the optimum phase found in the first analysis, whether or not it was significant. The incoherent case, obviously, did not have a consistent phase. As expected, we observed a significant reduction in mutual information in both the alpha and theta bands, suggesting that coherence is necessary for communication. However, the reduction in MI was much higher in theta. This is not because MI dropped lower in theta, but rather because when coherence was in place theta exhibited superior communication. Finally, we examined the impact of conduction delay on the preferred relative phase of coherent oscillations. Figure 4.7D shows that there is an obvious dependence of preferred phase on conduction delay, with a shift from 1 radian to 4.5 radians being observed when the conduction delay changed from 12 ms to 60 ms.

## Discussion

Our three modeling investigations produced the first full computational account of alpha information suppression. Other computational models of alpha have focused on the mechanism of generation (S. R. Jones et al., 2009), but to date none have examined communication (by any metric), or demonstrated frequency-dependent suppression of communication specifically within the alpha band. By producing a computational model, our investigation has allowed us to concretely test the feasibility of proposed mechanisms of

alpha suppression, as well as to provide novel, testable predictions. These include the proposition that alpha impacts cortico-cortical communication, and not just the thalamo-cortical communication, which is what has been experimentally examined to date. Our model also suggests that alpha oscillations should produce an oscillation not just in the LFP/scalp field, but also in the spike rate. Theta oscillations, on the other hand, should be predominantly subthreshold, simply controlling the timing of spikes.

Our model is based on a cortical resonance in the alpha range, which combines with nonlinearities in neural firing-rate response curves to produce two effects. First, as frontal input to the thalamo-cortical loop increases, the oscillation frequency increases. As the oscillation frequency approaches the alpha band, the amplitude of the oscillation—and thus measured power—increases. Second, because neural firing rates saturate, when the oscillations are high amplitude, the ability to change activity in response to inputs is diminished—that is, when the oscillations take up the complete dynamic range of the neurons, there is no further possibility to change the range of firing rates in response to inputs. By contrast, when oscillations are weaker, even subthreshold, transient inputs can produce a large response—beyond that observed in the absence of input—and thus information is conveyed.

The high-amplitude alpha oscillations in our model comprise saturating excitation with intervening periods of high inhibition, which similarly prevent input-dependent responses, in accordance with the pulsed inhibition hypothesis (Jensen & Mazaheri, 2010). While the pulsed inhibition hypothesis does not itself envision all of the details and features

of our model, the model does provide all the features proposed by pulsed inhibition. Particularly, it produces a series of pulses in the activity traces of cortical inhibitory interneurons. In addition, it displays a phase preference for communication when operating in the high theta range, as observed by Busch et al., (2009). (Mathewson et al., 2009) observed the same type of phase preference at 10 Hz. However, they measured alpha power and phase only at 10 Hz, and theta power and phase were not examined. This leaves open the possibility that the peak oscillatory power in this target detection task (which requires focused attention) was in fact still in high theta. What they described as an alpha-phase dependence may have been equally well or better described as a theta-phase effect, which is observed in our model as phase-dependent changes in communication, which would cause phase-dependent changes in performance during a recognition task.

A potential objection to our model is the validity of our assumption of a single mechanism, based on a thalamo-cortical loop, which produces either alpha or theta oscillations. The notion that theta might be thalamically driven is literally as old as the term "theta" itself. The name was chosen to represent the thalamus, which was supposed to be the origin of the newly discovered band of low-frequency oscillations. In time, theta oscillations were discovered in the hippocampus, and in recent years these have garnered much attention from researchers. However, a cortical theta does exist as well and is distinct from hippocampal theta. Frontal-midline (F-M) theta is different from hippocampal theta, usually oscillates around 6 Hz, and is the most representative form of neocortical theta (Mitchell, McNaughton, Flanagan, & Kirk, 2008). There is evidence that F-M theta is not generated hippocampally (Mitchell et al., 2008). Furthermore, there is evidence that the

thalamus is involved in F-M theta, in a synchronized (entrained) manner (Sarnthein, Morel, Von Stein, & Jeanmonod, 2005). Ishii et al., (1999) suggested that the thalamus might be driving F-M theta. Even if F-M theta is influenced by hippocampal theta, Kirk & Mackay, (2003) proposes that the influence might be via the anterior thalamic complex. Finally, the mechanism for generation of thalamic bursts in the theta range seems to be the same as the mechanism for the generation of thalamic bursts in the alpha range (S. W. Hughes & Crunelli, 2005). Taken together with the strong evidence that cortical alpha oscillations are driven by a thalamo-cortical loop (Contreras & Steriade, 1995; S. W. Hughes & Crunelli, 2005; Suffczynski, Kalitzin, Pfurtscheller, & Lopes da Silva, 2001), these results suggest that at least one type of cortical alpha and theta share a common generative mechanism.

Although coherent oscillations between two brain regions are sometimes interpreted as indicative of communication between the two regions, it is important to note that oscillations can also strongly inhibit the transfer of information between the two regions. In particular, when considering mutual information between coupled oscillating neural circuits, the strongest predictor of the firing rate of a receiving population is the relative phase within a cycle. If the phase relationship is not conducive to communication, inhibition will dominate in the receiver at the moment of greatest input from the sender, so that the sender's signal is unable to effectively excite the receiver.

At best, one could consider "packets" of signal, one per cycle, and assess how the peak firing rate of a receiver depends on the signal during the previous cycle. If the "packet" is

sent at just the right time—such as if the two oscillators are not only coherent, but have an optimal phase relationship—the amplitude of signal from a sender could have a reliable effect on the following peak firing rate of the receiver. The optimal phase difference would depend on the conduction delay between the two regions as well as the impulse response curves of the oscillator, so may be hard to establish in practice. Moreover, if the receiver's peak firing rate is already near saturation, its value is weakly correlated with any incoming signal, so communication by this method remains weak. Information could also be transferred via a phase shift produced in the receiver's oscillation, but it is unclear how a neural circuit could extract such information. These factors contribute in our model to reduced information transfer when the thalamo-cortical loop undergoes strong, albeit coherent, oscillations.

## CHAPTER FIVE: TOWARD A MORE REALISTIC MODEL OF ALPHA-INDUCED INFORMATION SUPPRESSION

### Introduction

Following the initial research on alpha-induced information suppression (not to be confused with alpha suppression, which is the suppression of alpha oscillations, as in (Lima, Singer, & Neuenschwander, 2011) that was reported in chapter five, opportunities for extending and improving the model were identified. Many of these opportunities focused on incorporating additional biological realism into the model, e.g. a realistic T Current in the thalamocortical projection neurons. We reworked our neural model to a voltage-based rate model capable of incorporating specific currents, and specified more realistic circuitry in our model thalamus (including ACh-inactivated gap junctions in the thalamic reticular nucleus). This new realism allowed us to address additional issues, such as the control mechanism which initiates alpha oscillations in the thalamo-cortical loop. We also improved upon the investigations in chapter four, offering more specific requirements and more realistic classification systems when replicating Busch et al., (2009). As a result of this process, however, we realized some potential limitations of the saturation model, and began a preliminary exploration of an alternate model of alpha-induced information suppression, based on the predictive coding framework and the same interaction of M-currents and Chloride leak currents which provided the mode shift in earlier chapters.

*Basal Ganglia Control of Thalamo-Cortical Alpha*

There are many computational models of the thalamo-cortical system involved in the control and initiation of alpha oscillations. The model of Destexhe et al. (1998) treated spindling, which is a phenomenon observable during non-REM sleep that involves waxing and waning of an oscillation at alpha frequency. Several subsequent alpha-generator models extended and modified the Destexhe model, including a series of related efforts by McCarthy et al. (2008), Ching et al. (2010), Vijayan & Kopell (2012), and Vijayan et al. (2013). The last of these reports combined modeling ideas from the other reports to propose that an anteriorization of alpha that occurs during propofol-induced unconsciousness involves a shift of the source for elevated alpha power in scalp EEG, from an occipital-thalamic generator, which depends on HTCs, i.e., specialized high-threshold bursting thalamo-cortical neurons, to a frontal-thalamic generator, which does not depend on such HTCs. The proposal of distinct frontal and occipital models with opposite responses to propofol was defended on the basis that HTCs have only been reported in the vision- and occipital cortex-associated geniculate nucleus of the thalamus. Unfortunately, the Vijayan et al. (2013) report omitted a cortical component for the occipital generator, even though such a component was included in Vijayan & Kopell (2012). Therefore it was unclear whether propofol's disconnection of the HTC generator might actually leave a residual occipital-thalamic circuit that would, like the frontal one (that lacked HTCs) become a stronger alpha generator under propofol. In addition, although the experimental data on propofol dosage increases show progressive failure to process stimulus input, there was never any simulation of the decline in stimulus processing, even at the thalamic level. Moreover, none of these reports addressed either of two issues that are central concerns of

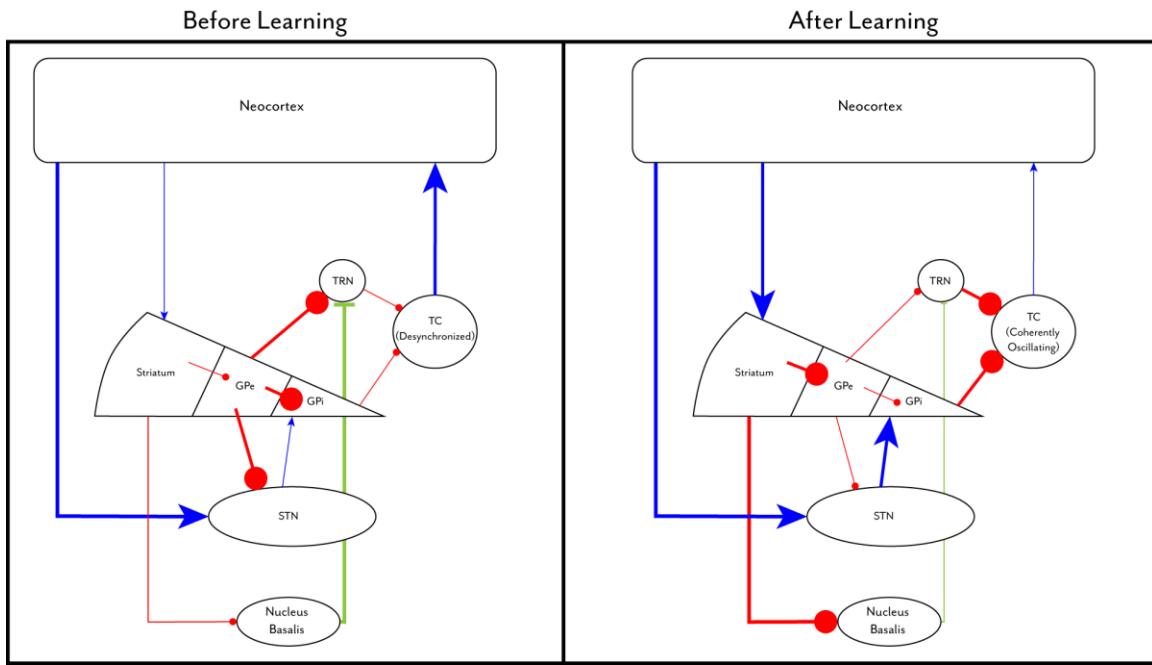
this report: First, how is it possible for a subject to learn to increase alpha for the purpose of ignoring distractors, and by what circuit is this control over alpha effected in a wide-awake, eyes-open subject? Second, what is the effect of synchronous high amplitude alpha on cortico-cortical information transmission? The only remarks found in Vijayan et al. (2013) about normal control of alpha is a pointer to the fact that the occipital generator depends on activations of mAChRs and/or mGLuRs. But no learning process that could mediate strategic deployment of alpha is described or simulated. And the only remarks in these papers about reduced information transfer pertain to a difficulty (presumed but not simulated) for trans-thalamic transmission of sensory information when the system is in alpha-mode.

In contrast, a recent paper by Hwang et al. (2014) does speak to the control issue, at least with respect to frontal-parietal alpha. They describe the need for alpha-induced information suppression in a sensory-motor context, namely that of the anti-saccade task in humans. In this task, subjects must suppress their tendency to make a saccade to the locus of a stimulus (and instead saccade to an imagined, mirror-image, location), and Hwang et al. found (using MEG) that successful suppression trials showed enhanced beta-power in DLPFC, enhanced alpha power in FEF, and a non-significant tendency toward enhanced alpha in LIP, a parietal area that relays sensory inputs to FEF. Furthermore, there was evidence that the DLPFC changes were causal to the FEF changes. They also cite primate electrophysiological studies (review in Munoz & Everling, 2004) indicating enhanced pre-cue inhibition, and depressed cue-induced activations in FEF on successful anti-saccade

trials, relative to error trials. Thus there is evidence from this paradigm of coexistent enhanced alpha and depressed stimulus-induced activations in FEF.

In their discussion, Hwang, Ghuman, Manoach, Jones, & Luna (2014) consider two (possibly non-exclusive) paths by which DLPFC might control FEF alpha: via a cortico-cortico link involving supragranular drive, or via a cortico-basal-ganglia(BG)-thalamus circuit, which may also provide the supragranular drive believed to be important for widespread cortical alpha synchrony. With respect to the latter, Hwang et al. note that beta is often a signature of cortico-striatal linkage, and there are experimental and computational models (see Munoz & Everling, 2004; Brown et al., 2004; Aron, 2011) in which the BG mediates learned suppression of cue reactivity in FEF. Our hypothesis is that learning occurs in the cortical-striatal projection on error trials, on which reactivity to distractor cues interferes with accurate performance. Error trials lead to non-reward, and depressed post-response dopamine release. Across several such trials, this causes a significant, task-dependent potentiation of cortico-striatal synapses onto D<sub>2</sub>-MSPNs, the neurons that give rise to the indirect pathway through the BG, which is depicted in Figure 5.1. On future trials in the same task, enhanced D<sub>2</sub>-MSPN activity inhibits GPe. This will have effects on the thalamus via two branches. First the reduced GPe output will mono-synaptically disinhibit TRN, which will then increasingly inhibit TC (thalamo-cortical) cells. Second, reduced GPe output disinhibits STN and GPi/SNr. The disinhibition of STN increases excitation of GPi/SNr. The resultant greater activation of GPi/SNr also increases inhibition of TC cells. These effects oppose tonic, specific-signal-driven, i.e. "relay", firing of TC cells, and thus reduce excitation of cortex. Equally important, the

much deeper inhibition of TC cells serves as a powerful inducer of thalamic alpha. In biophysically detailed models, this occurs because thalamic alpha strongly depends on intrinsic (non-synaptic) hyperpolarization-activated currents that are normally inactive due to tonic partial depolarization during the wake state. Enhanced inhibition of TC neurons by TRN and GPi/SNr, conditional upon learned activation of the BG indirect pathway, is therefore a prime candidate for reinforcement-guided recruitment of frontal, and perhaps frontal-parietal, alpha in the antisaccade task.



**Figure 5.1: The pathways by which learned excitation of D2-MSPNs in the striatum can control alpha generation.** Blue arrows represent excitatory connections, red circles represent inhibitory connections, and green bars represent ACh-releasing connections. Thicker lines denote stronger connections. A comparison of the left panel (before learning) with the right (after learning) shows that the learned cortical excitations of D2-MSPNs in striatum has the effect of changing three inputs to TC (thalamo-cortical) neurons: enhanced inhibition from TRN (thalamic reticular nucleus) and GPi (internal/medial globus pallidus) and reduced cholinergic input from nucleus basalis. D2-MSPNs are striatal medium spiny projection neurons that express D2-type dopamine receptors.

The model schematized in Figure 5.1 is consistent with, and provides a mechanistic explanation of, several other key results, including the earliest report of learning-based enhancement of alpha. As reviewed in Sterman & Egner (2006), alpha power enhancement was discovered as a reliable “side effect” of learned suppression of a previously rewarded cup-press response for food in cats. This fits well the emergence of alpha in anti-saccade tasks, in which subjects must learn to suppress one of the most highly rewarded responses made by sighted mammals. Another fit is provided by the paradigm used in Buschman et al. (2012). They trained monkeys in a rule-switching task, which required monkeys to periodically switch between using orientation or color as the decision-making criterion. Both orientation-rule-coding (ORC) ensembles and color-rule-coding (CRC) ensembles in DLPFC showed beta-band coherence when their respective rule was in force. However, when the monkey had to suppress application of the prepotent rule – orientation – the ORC ensemble oscillated at alpha frequency whereas the CRC ensemble exhibited beta synchrony. The reverse was not true: there was no alpha oscillation in the CRC ensemble when the ORC ensemble was oscillating in beta. Once again, frontal alpha appeared when the prepotent strategy had to be suppressed, but not otherwise. In this case, monkeys had made many errors by attending to orientation when color attention was needed. By our hypothesis, this would have led to strong potentiation of cortico-striatal synapses that recruit indirect pathway suppression of orientation processing.

Also shown in Figure 5.1 is yet another path by which BG indirect path activation could significantly boost the ability to instate alpha mode during waking. Although the Vijayan & Kopell (2012) model of occipital-geniculate alpha assumes that activation of mAChRs can

promote a specialized type of alpha in which HTC neurons are active but TRN neurons quiescent, there are reasons to doubt the generality of such a mode. First, the mechanism hypothesized above for control of frontal alpha utilizes hyperpolarizing inhibition of thalamus by combined GABAergic inputs from GPi/SNr and TRN. Second, there are several published reports indicating that wake-active neuromodulators including ACh, NA, and 5HT have anti-oscillatory effects. This is pertinent because a major source of ACh, the nucleus basalis magnocellularis (NBM) in the basal forebrain, receives its major GABAergic input from D<sub>2</sub>-MSPNs of ventral striatum (Sarter et al., 2006). Thus, learned recruitment of the indirect pathway's ventral striatal D<sub>2</sub>-MSPNs would inhibit AChNs in NBM and thereby reduce release of the anti-oscillatory neuromodulator ACh in the cortico-TRN-thalamic circuit. Such recruitment of D<sub>2</sub>-MSPNs in ventral striatum is consistent with current interpretations of distributed contributions to response promotion and demotion (Gruber & McDonald, 2012). Overall, Figure 5.1 summarizes how learned recruitment of striatal D<sub>2</sub>-MSPNs can promote alpha via three synergistic paths: two that enhance inhibition of thalamus, and one that reduces ACh release in TRN.

## Methods

Similarly to Chapter four, we propose a mechanism of alpha-induced information suppression that is based on alpha-band resonance of the thalamo-cortical circuit (Herrmann, 2001; VanRullen & Macdonald, 2012). We did not simulate the cortico-striatal learning controlling alpha genesis, however, because to simulate key effects of learned activation of the basal ganglia's indirect pathway, it sufficed to modify three parameters of the thalamus-TRN-cortex circuit:

1. Reduce ACh concentration in TRN.
2. Increase inhibition from GPi to the thalamic relay nucleus.
3. Decreased inhibition from GPe to TRN.

The oscillation amplitude and frequency varies with these parameters. We used this to adjust thalamic oscillation frequency across its observed range from 2–13 Hz (S. W. Hughes & Crunelli, 2005).

The coupling of thalamic and cortical oscillators in a loop makes the frequency of cortical oscillations depend on the frequency of the thalamus, and therefore indirectly on the activation of the basal ganglia's indirect pathway. The coupled circuit possesses a resonance, which appears when the tonic drive to the thalamus causes the thalamic circuit's frequency to match the 10 Hz natural frequency of the cortical circuit. We investigate the consequence of such an alpha resonance by simulating three tasks, which we summarize below. Details of the model's implementation and the task-specific inputs follow the brief description of each task.

Our first investigation tested how the mutual information between a cortical input signal and the cortical activity depended on the frequency of thalamo-cortical oscillations. A continuously varying signal was sent to the cortex while the thalamus was controlled via different levels of indirect pathway activation. This was manipulated to place the thalamus into both alpha (10 Hz) and theta (6 Hz) oscillations. At each level of input (and therefore frequency), mutual information was evaluated between the signal and the cortical spike rate.

In the second investigation, we assessed how cortical responses depended on the phase of an input, aiming to reproduce the experiment reported in Busch et al., (2009). Our model was simulated with continuously varying noise input to the thalamus and cortex while under three input regimes. In two, sufficient activation of the indirect pathway drove the system into theta or alpha range, and in the third no pulse of indirect pathway activation was administered, leaving the system in a non-oscillatory state. In addition to the varying noise, a 12 ms pulse was injected to the thalamic relay neurons at various phases of the thalamo-cortical oscillation. This pulse simulated the presentation of a brief threshold-level visual stimulus. The model's response to the stimulus pulse was determined by a simple threshold classifier (see Analysis), and behavioral statistics were calculated as in Busch et al.'s original work, along with ROC curves for each of the phases.

### *Model*

In describing our model, we follow the standards proposed by (Nordlie et al., 2009), documenting the network architecture, the connectivity, the neuron and synapse models used, the input (stimuli), and the output (data recorded).

### *Network Architecture*

The basic modeled network (Figure 5.2) consists of one thalamo-cortical loop, with the thalamus containing four populations of neurons, and the cortex containing two populations of neurons. The model thalamus consists of two excitatory thalamic relay nuclei (each a population of thalamic relay cells) and their corresponding section of inhibitory thalamic reticular nucleus (TRN, each consisting of a population of reticular

nucleus cells). The model cortical area contains a population of excitatory projection neurons and a population of inhibitory interneurons. Feed-forward inhibition was omitted from the model, because the inhibitory cells that are targets of long range cortical connections do not generally have recurrent connections from the excitatory neurons that we are modeling (Apicella et al., 2012).

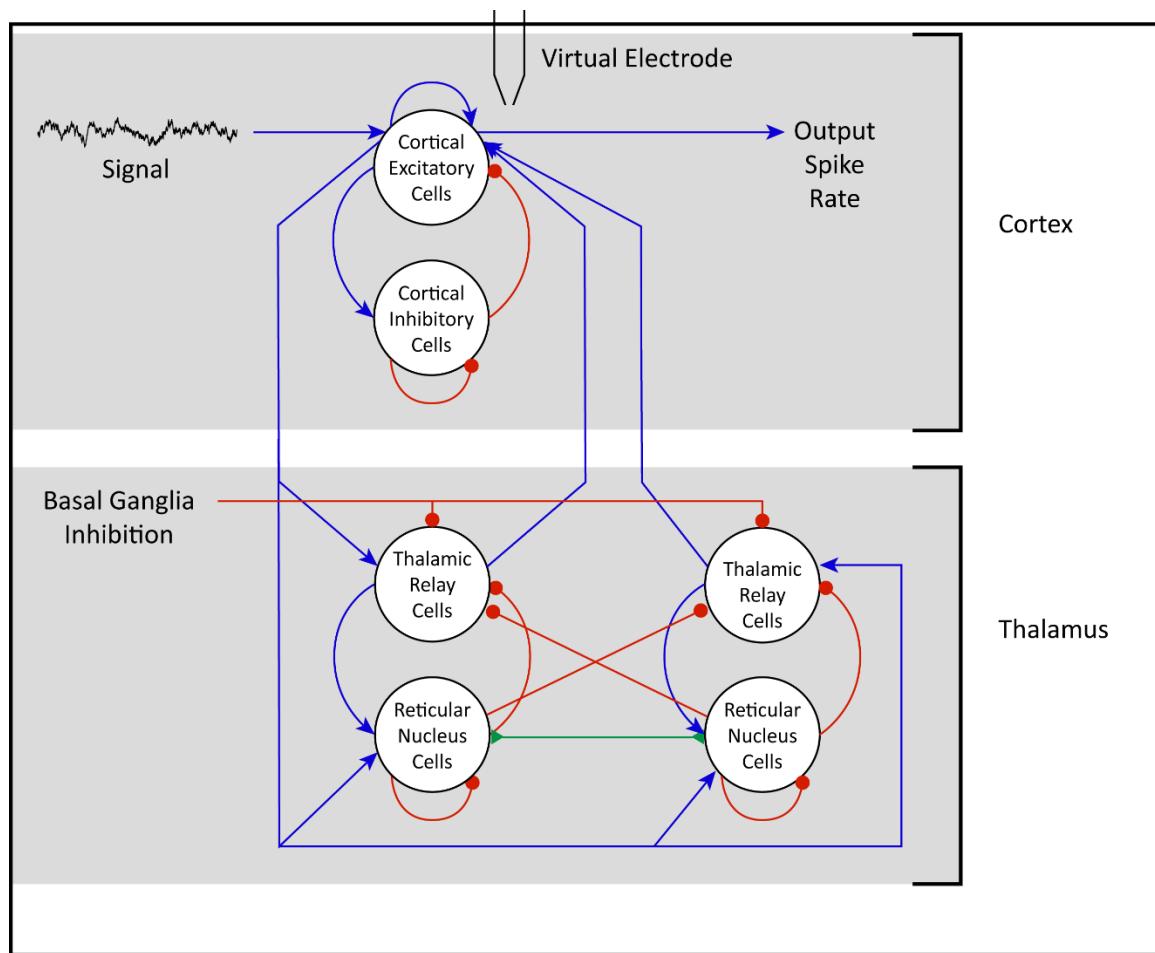


Figure 5.2: Our model circuitry. Two brain regions (the thalamus and cortex) are modeled, each with populations of excitatory and inhibitory neurons. Blue arrows represent excitatory connections, red circles represent inhibitory connections, and green triangles represent gap junctions. The basal ganglia inputs shown are from the GPe. Also, modeled, but not shown are two further inputs from the basal ganglia/forebrain: an inhibitory input from GPe to TRN, and a cholinergic input from NBM to TRN cells.

### *Network Connectivity*

This model defines synapses between neurons in the two populations via a pair of parameters: a weight,  $W_i$ , and a conduction delay  $\tau_i$ . Gap junctions are defined with a weight  $W_i$  that varies between a min and max weight depending on ACh release by the projection from NBM to TRN. These are enumerated for each synapse or gap junction in Table 6.1.

Connection	Weight $W_i$	Conduction Delay $\tau_i$	Sources indicating pathway is present
Thalamic projection neurons to TRN interneurons	12	0.25 ms	Jones, 2002
Thalamic projection neurons to cortical excitatory neurons	20	4 ms	Sherman and Guillery, 2004
TRN interneurons to thalamic projection neurons	-14	0.25 ms	Sherman and Guillery, 2004
TRN interneurons to TRN interneurons	-0.2	0.25 ms	Destexhe et al., 1998
Excitatory cortical to cortical inhibitory	15.75	0.25 ms	Destexhe et al., 1998

Excitatory cortical to excitatory cortical	12.0	0.25 ms	Destexhe et al., 1998
Excitatory cortical to thalamic reticular nucleus	20	4 ms	Jones, 2002
Excitatory cortical to thalamic projection neurons	20	4 ms	Destexhe et al., 1998
Inhibitory cortical to excitatory cortical	-8	0.25 ms	Destexhe et al., 1998
Inhibitory cortical to inhibitory cortical	-0.01	0.25 ms	Destexhe et al., 1998
TRN gap junction	Min: 0 Max: 2000	N/A	

Table 6.1. Parameter values in the mathematical model.

### *Neuron and Synapse Models*

Our model is constructed out of rate-based populations of neurons whose mean voltage is simulated via a simple differential equation and translated into a firing rate via an exponential function and basic differential equation:

$$f(v) = \frac{r_{\max}}{1 + e^{((V_{half} - v)/v_{sig})}}$$

$$\frac{dr}{dt} = \frac{f(v) - r}{\tau_r}$$

Where  $r_{max}$  is the maximum firing rate  $r$ , set to 100 for cortical pyramidal and thalamic relay cells, and 200 for cortical interneurons and TRN cells.  $V_{half}$  is the voltage  $v$  at which half of the maximum firing rate is achieved,  $v_{sig}$  is the steepness of the sigmoid, and  $\tau_r$  is the time constant at which the firing rate changes.

The mean voltage changes via the dynamical system below:

$$I = C_{Ge}^2 \left( I_0 + I_{noise} + \sum_j w_j r_j \right)$$

$$\frac{dv}{dt} = \frac{I/G - v}{\tau_m}$$

Where  $I$  is the total input current to the neuron,  $I_0$  is the base external excitation (set at 100 for thalamic relay cells, varied from 5-40 for TRN cells, set at -20 for cortical interneurons, and set at 0 for cortical pyramidal cells),  $I_{noise}$  is a noise input, which in both investigations was generated via the Ornstein-Uhlenbeck process mentioned below.

The summed terms  $w_j$  and  $r_j$  represent the weight and conduction-delay-offset firing rate of synapse  $j$ , respectively.

The thalamic relay cells also include the rebound current  $I_{rbd}$  in their summation of input currents, making the equation:

$$I = C_{Ge}^2 \left( I_{rbd} + I_{e_0} + I_{noise} + \sum_j w_j r_j \right)$$

The rebound current  $I_{rbd}$  is defined by the following functions:

$$f_{m_{rbd}}(v) = \frac{1}{1 + \exp\left(\frac{m_{rbd_{half}} - v}{m_{rbd_{sig}}}\right)}$$

$$f_{h_{rbd}}(v) = \frac{1}{1 + \exp\left(\frac{v - h_{rbd_{half}}}{h_{rbd_{sig}}}\right)}$$

$$\frac{dm_{rbd}}{dt} = \frac{f_{m_{rbd}}(v) - m_{rbd}}{\tau_{m_{rbd}}}$$

$$\frac{dh_{rbd}}{dt} = \frac{f_{h_{rbd}}(v) - h_{rbd}}{\tau_{h_{rbd}}}$$

$$I_{rbd} = I_{rbd_{max}} m_{rbd} h_{rbd}$$

Where  $h_{rbd_{half}}$  and  $h_{rbd_{sig}}$  are the half-max and sigmoid steepness terms for  $h_{rbd}$ , with their  $m$  equivalents playing analogous roles for  $m_{rbd}$ . The two  $\tau$  terms are time constants, and  $I_{rbd_{max}}$  is the maximum self-excitation achievable by the rebound current.

Thalamic Reticular Nucleus neurons also included gap junction input  $I_{gap}$ , which altered their base membrane voltage equation to be:

$$I = C_{Ge}^2 \left( I_0 + I_{noise} + I_{gap} + \sum_j w_j r_j \right)$$

$I_{gap}$  was calculated as

$$\sum_i w_{gap} V_i$$

The summed terms  $w_{gap}$  and  $V_i$  represent the gap junction weight and the membrane voltage of the adjacent neuron  $I$ , respectively.

### *Input (Stimuli)*

The model system has three direct input sources, and one modulatory input. The first is a signal input to the cortical pyramidal population. This represents information coming either from sensory sources or other brain regions via cortico-cortical connections. The second source of input is from the GPi to the thalamic projection nucleus. This input is inhibitory and is increased as the indirect pathway is activated. The third direct input source is the indirect pathway's input to the TRN via the GPe, which is inhibitory and is reduced as the indirect pathway is activated, disinhibiting the TRN neurons. While these three input sources largely determine the frequency (and therefore amplitude) of thalamic oscillations, the modulatory effect of diffuse ACh release on gap junction conductivity (modified directly in our investigations, but hypothesized to be from the indirect pathway via the Nucleus Basalis) can also enable or disrupt the synchrony of oscillations, depending on its level.

For the first investigation, the signal consisted of a pseudo-random continuous current input, generated by an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930), and varying from roughly -50 to 50 mA. The specific equation used was as follows:

$$dx_i = \theta(\mu - x_i)dt + \sigma dW_{ij}$$

The parameters of the process were  $\theta = 10$ ,  $\mu = 0$ , and  $\sigma = 90$ , and the final result was scaled by a factor of 90 to achieve the desired current range. To elicit alpha oscillations, the direct inputs were set to 100 for the Thalamic Relay cells and 22 for the TRN as a baseline, with a pulse of indirect pathway activation from milliseconds 100 to 1100 changing them to -400 and 50, respectively. This pulse initiated the alpha oscillations. For theta oscillations, the same values were used except the TRN values were 35 at baseline and 63 during the pulse.

The second investigation used an identical pseudo-random continuous input as noise, summed with a single 0.5 amplitude square "pulse" 12 ms in duration. The pulse was delivered at varying phases of the thalamo-cortical oscillation. For both Alpha and Theta, we used identical baseline and pulse parameters to the first investigation. For the non-oscillating condition, the Alpha baseline parameters were used, but no pulse occurred, which prevents oscillations from beginning.

### *Output (Data Recorded)*

Spiking rate data was recorded from the excitatory population of cortical projection neurons, as well as from a simulated virtual electrode recording the local field potential (LFP) generated by the summed synaptic inputs to these same neurons. Each electrode was approximated through the simple method of summing excitatory and inhibitory current inputs to each cell population, and multiplying by -1 so that excitatory inputs produce a negative deflection in simulated voltage.

### **Analysis**

Data analysis was conducted in Matlab 2012a (The MathWorks, Natick, MA). For all three investigations, we used a fast Fourier transform to analyze the cortical excitatory population's local field potential (*i. e.* summed synaptic input) to determine oscillatory power over a time-frequency spectrum. We also classified the thalamic oscillations based on mean distance between peaks.

For the first investigation, at each 10 second step of frontal drive to the thalamic nucleus, the impact of the signal input on the spiking rate output data was computed via mutual information (Dayan and Abbott, 2001). Mutual information (MI) allows us to determine how much variability in the receiver's response is due to the activity of the sender. Mathematically, MI is equal to the receiver's full response entropy minus its noise entropy.

$$I_m = \int ds \int dr p[s]p[r|s] \log_2 \left( \frac{p[r|s]}{p[r]} \right)$$

MI can be thought of as a measure of how well variations in the sender are correlated with variations in the receiver, beyond the ability of the receiver to predict its own variability.

We evaluated mutual information by binning our signals into discrete sender and receiver states, with the number of bins determined by Scotts Normal Reference Rule (Scott, 1979), a rule originally designed for optimizing bin count in histograms. It suggests bins of width

$\frac{3.5\hat{\sigma}}{n^{1/3}}$ , where  $\hat{\sigma}$  is the standard deviation of the sample, and  $n$  is the number of data points

in the sample. We then approximated the above integrals by sums over bins  $r_i$  and  $s_j$ , with  $p([r_i | s_j])$  being estimated by  $N_{ij}$ , which is the number of occurrences of rate  $i$  given stimulus  $j$ , divided by the number of occurrences of stimulus  $j$ . This evaluation was repeated over a range of time delays up to 40 ms, and we selected the delay that yielded the highest MI value.

We also corrected for the number of samples, since a less than infinite number of samples can cause bias in the mutual information estimation. This was corrected for by determining the number of elements in the response space  $\bar{R}$  and the number of data points  $n$ , then subtracting  $\frac{\bar{R}-1}{2n \log(2)}$ , after Panzeri et al. (2007).

The simulation was repeated 10 times under new random seeds. A Lilliefors test showed that the mutual information measurements during theta oscillations were not normally distributed. We therefore ran a nonparametric Wilcoxon rank sum test in order to check

for a significant difference in medians between mutual information measured during alpha and theta oscillations.

For the second investigation, mutual information was not calculated, and instead a threshold classifier sorted trials into hits, where activity during the cycle when the stimulus pulse occurred surpassed or equaled the threshold level; and misses, where activity during the cycle did not reach the threshold level. Hit and miss rates were calculated for each phase angle of stimulus presentation. False positives obviously could not be sorted based on stimulus phase, since there was no stimulus during these trials. False alarm rates were, however, calculated for all non-target trials. Then, by varying the threshold of the classifier, receiver operator characteristic (ROC) curves were generated for each phase, and the area under each curve (AUC) was calculated.

## Results

### ***Investigation 1: Mutual Information***

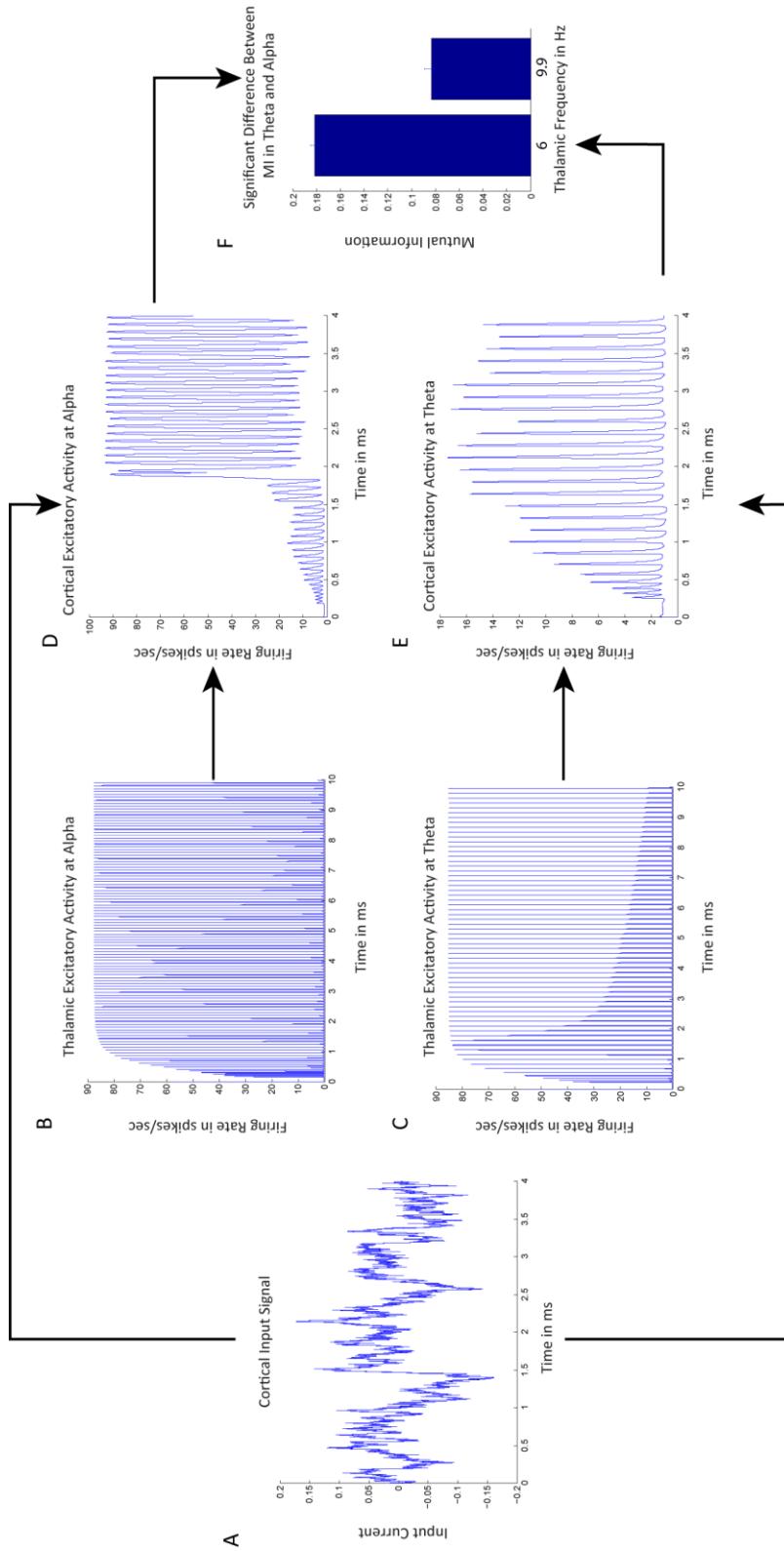
The results from the first investigation show a clear frequency-dependent reduction of mutual information when the thalamo-cortical circuit oscillates in the alpha range (10Hz). When the amplitude of periodic drive to the cortical pyramidal cells increases, the cortical activity transitions from a sequence of bursts correlated with its input signal into a regular oscillatory pattern, drastically reducing the available information about the signal.

The rank-sum test rejected the null hypothesis of equal medians ( $p = 0.0013$ ), demonstrating a significant difference between mutual information in the 6 Hz group

(representing oscillations at theta frequencies) and the 10 Hz (representing alpha) group of trials.

Additionally, the results demonstrate the behavior we had hypothesized at each stage of the process. Reduced baseline inhibition to the TRN produces a higher frequency oscillation in the thalamus as well as greater thalamic peak amplitude. This oscillatory thalamic behavior, when combined with the Ornstein-Uhlenbeck input signal, causes a switch from mostly subthreshold (but occasionally suprathreshold) cortical oscillations, whose amplitudes are largely determined by the amplitude of the signal, to saturated suprathreshold cortical oscillations, which are largely invariant with regard to the signal amplitude. The switch to saturated suprathreshold cortical oscillations leads to a significantly lower mutual information.

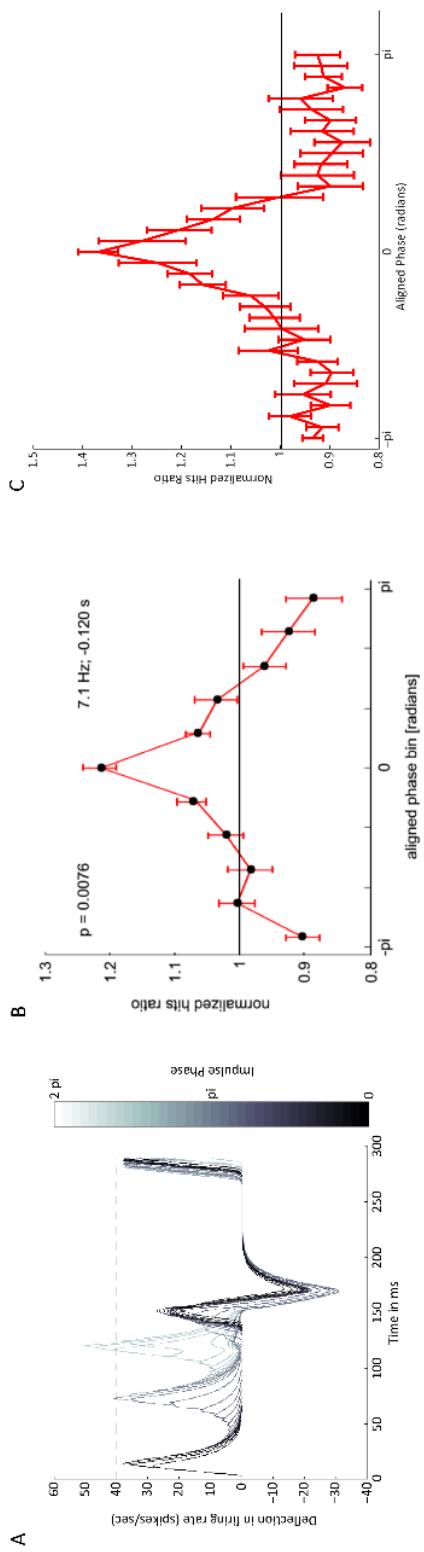
The frequency of oscillation in our model thalamo-cortical loop depends on the level of baseline inhibition to the TRN, which is governed by the indirect pathway (Fig. 6.1). Without reducing the inhibition below zero, the range of bursting frequencies produced in our model matches the range observed in rat LGN slices, where increased depolarization of relay cells leads to increasing frequencies between 2 and 13 Hz (Hughes and Crunelli, 2006).



**Figure 5.3:** Simulation Results demonstrating A) the Ornstein–Uhlenbeck-generated input signal, B–C) the thalamic activity at roughly 10 and 6 Hz, respectively. D–E) the cortical activity driven by a combination of the input signal and the thalamic activity, and F) the significant difference in mutual information between the 6.1 Hz and 10 Hz conditions.

### **Investigation 2: Discrimination Threshold**

As shown in Figure 5.4, results of our second investigation were in close agreement with those reported by N. A. Busch, Dubois, & Vanrullen, (2009). For example, Fig. 6.4B shows that the deflection in cortical activity produced by a single stimulus pulse is strongly dependent on the phase of stimulus presentation with respect to the ongoing cortical oscillation. As can be seen in Figure 5.4A, when multiple phases of impulse were applied using the same (frozen) background noise, both the initial and later oscillatory impact of each stimulus impulse varied with the phase, some of them falling below the threshold of the classifier (40 spikes/sec), and others surpassing the threshold. Secondly, the threshold classifier demonstrated extremely strong phase dependence. When the hit rate (rate of trials classified “stimulus present” out of all trials with a stimulus impulse) is plotted as a function of phase as in Figure 5.4C, the preference for an impulse just preceding the onset of the cycle (a peak in the oscillation) becomes clearly visible.



**Figure 5.4: Deflection and hit rate by phase.** A) Cortical deflection from baseline during a cycle, as a function of impulse phase. B) Experimental normalized hits ratio by aligned stimulus onset phase (reproduced from Busch et al., 2009). C) Model normalized hits ratio by aligned stimulus onset phase.

## Discussion

### *Limitations of the current model*

The introduction of a realistic mechanism through which alpha-recruitment could be controlled by a learning process extended the model to address a key problem neglected in most oscillation models: How are switches between oscillatory modes made to serve cognitive goals? However, the added realism of this extended model did not come without price. Additional parameters made tuning the model more difficult, and the redefined investigations also illuminated two other potential limitations to the model. As will be shown below, these potential limitations made it prudent to additionally explore a model that relied on subthreshold dynamics, rather than saturation. First, the saturated spiking rate in the model at peaks in the alpha cycle complicates comparisons with other experimental data. In Investigation Two, we aimed to reproduce Busch et al., (2009) by subjecting an oscillating thalamo-cortical system to an impulse (representing a roughly 12ms visual stimulus), then using a threshold classifier to determine whether a given trial was a hit, miss (false negative), false positive, or correctly identified non-stimulus trial. However, using any linear or threshold classifier, there will be interference from the increased firing rate under alpha when attempting to detect the increased firing rate due to the stimulus. Alpha can be made, with sufficient tweaking, to perform roughly at chance compared to a better-than-chance classification during theta oscillations. This at-chance performance, however, is due to the alpha oscillating classifier marking every trial as a target trial, resulting in an extremely high number of false positives, and no misses. While

the AUC can be minimized this way, this is categorically not how alpha appears to operate in human tasks, such as the one in Romei et al. (2010).

Second, and more conceptually, if saturation, as in the case of our previously presented model, involves rapid suprathreshold spiking, then we must consider the relationship between high-frequency spiking and high gamma oscillations in electric/magnetic fields. If the high-frequency spiking is synchronized, then high gamma should accompany alpha. However, some studies have demonstrated a marked reduction in gamma-band power during periods of high alpha power. This suggests that either the frequent spikes are not synchronized across a population, or that there is a reduction in the spiking activity during alpha. If the second of these is true then it would seriously call into question the saturation model. While there is no conclusive evidence, a few studies offer suggestive evidence.

Haegens, Nácher, Luna, Romo, & Jensen, (2011) observed that cortical neurons in macaques fire at a reduced mean rate during alpha oscillations during attention to a target, rather than an increased mean rate.

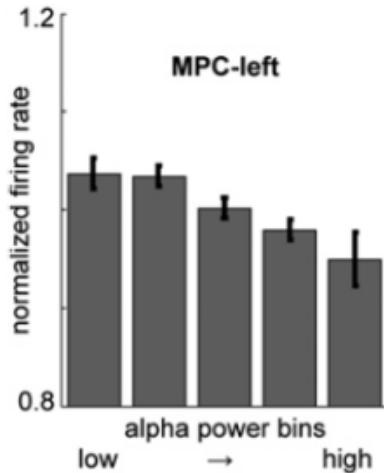


Figure 5.5: Macaque medial prefrontal cortex neural firing rate (extracellularly recorded) normalized to average pretrial firing rate as a function of alpha power. Adapted from Haegens, Nácher, Luna, Romo, & Jensen, (2011).

The same paper also observed that cortical neurons oscillating in alpha fire the most at the least depolarized moment of the alpha cycle. To the extent that the alpha peaks correspond to increased depolarizing currents in the neuron, this periodic reduction in firing rate seems to be the opposite of what is observed in our saturation model.

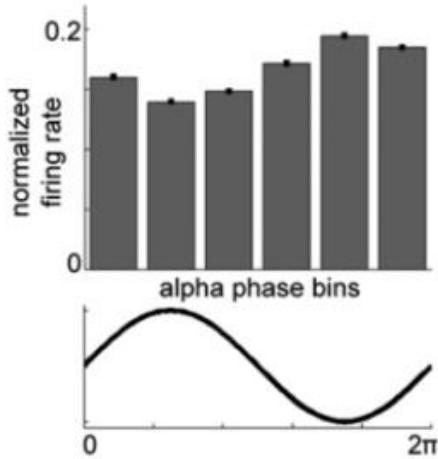


Figure 5.6: Macaque medial prefrontal cortex neural firing rate (extracellularly recorded) normalized to average pretrial firing rate as a function of binned alpha phase (derived from simultaneous LFP recordings). Adapted from Haegens, Nácher, Luna, Romo, & Jensen (2011).

This agrees with a finding of Mirpour, Bisley, & Bisley, (2013), who observed both increased alpha power and reduced spike rates for ignored distractors. In addition, Zhu et al. (2009), found that MEG deflections corresponded more closely to local field potentials than to multi-unit spiking activity, so it would seem plausible that subthreshold activity could account for observed alpha power. Unfortunately, both subthreshold activity and spike synchronization are difficult to examine in a rate-based model such as ours.

A first step within the modeling realm was then to examine possible mechanisms that could impede information transfer in a subthreshold regime. Furthermore, as earlier chapters have shown, M-currents have a potentially large impact on neocortical pyramidal spiking activity, whether in the form of a mode shift or in their more traditionally acknowledged role in spike frequency adaptation (Peters, Hu, Pongs, Storm, & Isbrandt, 2005). The saturation model is incomplete without them, and we must at some point address the

question of whether it can survive the addition of the M-current. Therefore, it was decided that any further model being explored should also include the M-current.

Furthermore, the observation by (Haegens et al., 2011) that spike rates are reduced at the peaks of alpha oscillations is reminiscent of a paradoxical observation from studies like (Kok et al., 2012), which find that (all else being equal) both reaction time and cortical activity decrease when a target is predictable (as in a cued condition).

While reaction time cannot be defined with regard to an ongoing oscillation, the reduction in spike rate for a predicted stimulus was achieved in earlier chapters via the M-current and a shift in neural encoding. This suggested that an attempt to apply these M-current-based models to alpha-induced information suppression might be fruitful. Our explanation is detailed below.

## Methods

### *Model Overview*

#### **The Predictive Coding Hypothesis**

The predictive coding hypothesis, an approach typified by Clark, (2013) and Friston, (2005), is an extension of the hierarchical predictive processing theory of the brain (often abbreviated PTB), which primarily consists of the ideas that brains are effectively hierarchical prediction testers, and that brains are constantly attempting to minimize prediction errors, e.g., via learning processes that improve predictions. The extension put forward by Friston consists primarily of defining an encoding strategy wherein the feedback from "higher" to "lower" brain areas predicts features of inputs to lower areas, and inhibits them. The resultant subtraction of predictions from inputs leaves only unpredicted features (a.k.a. error signals) to be transmitted up the hierarchy via feedforward pathways from lower to higher areas. Predictive Coding is most often implemented in a Bayesian framework with little regard to neural constraints, but neurally plausible implementations of predictive coding have recently received some attention (e.g., Wacongne, Changeux, & Dehaene, 2012).

Predictive Coding proposals often come with broader claims that it offers a grand unified theory of perception and cognition arising from the interplay of top-down predictions and bottom-up error signals. I want to stress that the detailed neuronal processing framework that I am developing here is agnostic about, and not dependent on, any of these grander claims.

## KCNQ and the M-current

Membrane biophysicists and neurochemists have characterized the M-current as an inward potassium current via channels constituted by KCNQ proteins, often heteromeric KCNQ<sub>2</sub>/KCNQ<sub>3</sub> proteins (Fedorenko et al., 2008). Originally discovered in sympathetic ganglion neurons of bullfrogs (Brown and Adams, 1980), it is non-inactivating and voltage-dependent, being activated by depolarization of the membrane to a level greater than approximately -60mV. Because of these properties, both sub-threshold depolarizations and those large enough to generate action potentials reliably also activate the M-current. However, this is subject to muscarinic neuromodulation. Indeed, the M-current is so named because KCNQ channels are closed by activation of mAChRs (muscarinic acetylcholine receptors). Such closure effectively disables the current. The M-current mechanism is found in neurons throughout the nervous system, including the pyramidal cells that are the principle neurons of the cerebral cortex (reviewed in Jentsch, 2000; Marrion, 1997) and the medium spiny cells that are the principle neurons of the striatum (reviewed in McCarthy et al., 2008).

Another key feature of the M-current is that the KCNQ channel has a labile time constant: the speed with which it opens and closes is dependent upon the membrane voltage. This is illustrated in Figure 5.7 by plots of the time constants resulting from four simulated voltage clamps.

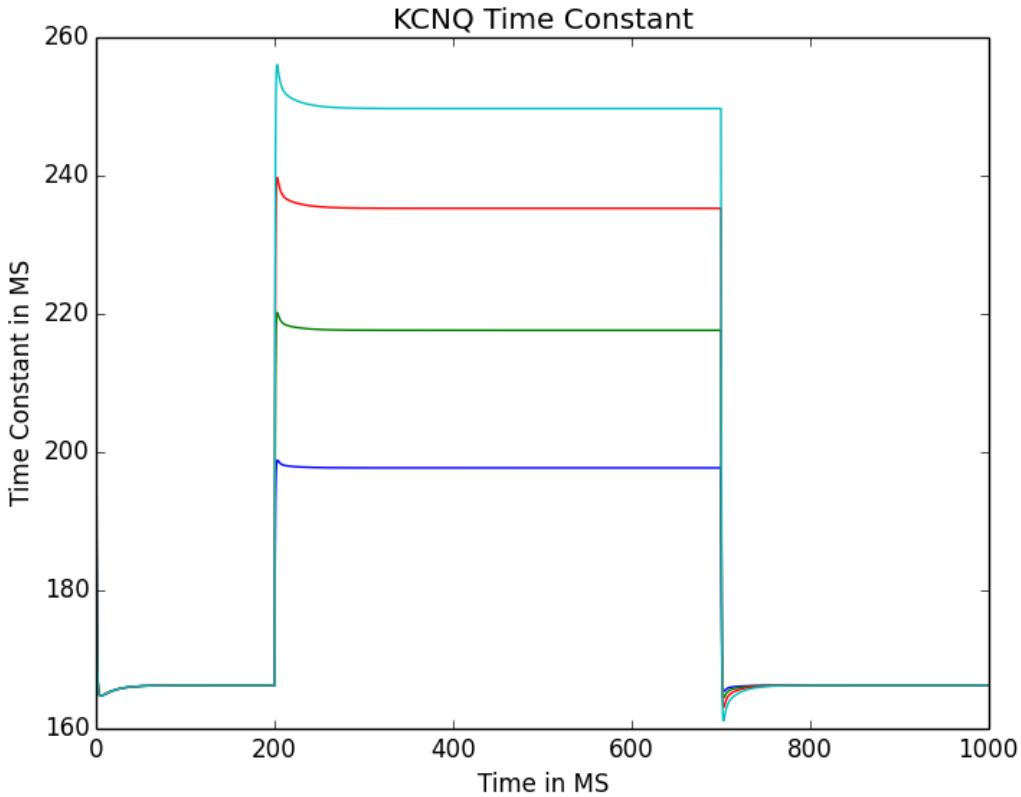


Figure 5.7: Simulated voltage clamps demonstrating a voltage-dependent time constant for the KCNQ channels that mediate M-currents. Abruptly instating and clamping new transmembrane voltages (blue: -50 mV; green: -40 mV; red: -30 mV; light blue: -20 mV) between 200 ms and 700 ms show the dependence of the KCNQ channel's time constant on voltage. Details of the mathematical simulation can be found in chapter 2.

### Phasic-Tonic Mode Switching

The M-current is usually thought of as a rate-modulating current, which contributes to the oft-observed slow adaptation in spike rate during prolonged excitation. (Kim et al, 2012) Under certain circumstances, however, it can act quite differently, as illustrated in Figure 2. Notably, Prescott et al. (2006) showed that the M-current can, via an interaction with Cl<sup>-</sup> currents, reliably cause a shift from regular (tonic) spiking to onset (phasic) bursting.

This was demonstrated both in computational models (similar in structure to our proposed model below) and in dynamic clamp recordings pyramidal neurons from the CAI part of the hippocampus.

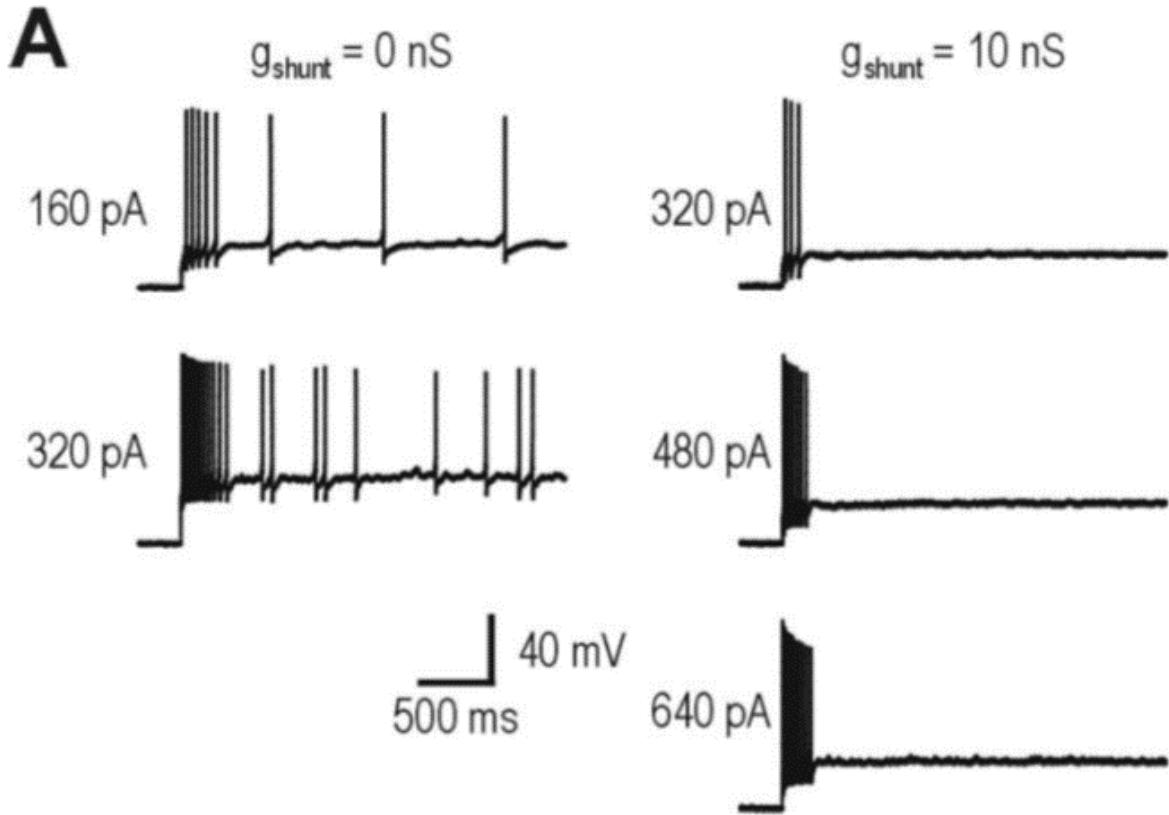


Figure 5.8: An example of CAI pyramidal neurons switching between tonic and phasic modes of firing under different injected currents (y-axes) and shunting conductances (the two columns). Increased shunting produced a change from tonic to phasic firing, and increased the amount of stimulation (pA) required to produce even phasic firing.

Under the predictive coding hypothesis, the best way to prevent something from propagating upward in the brain is to accurately predict it. In chapter three, I put forward the proposal that this is done via a top-down priming of layer 2/3 cortical pyramidal cells, causing a tonic-phasic mode shift, which in turn dynamically changes their output from a rate-based code to a dynamically negotiated labeled line scheme.

Unfortunately, when ignoring an entire modality, it is difficult to predict exactly what stimuli you will need to ignore. I propose that alpha oscillations represent a surprising way of circumventing this problem; they prime everything. When all feature representations are primed, then it very nearly doesn't matter what features are presented. They will all register as expected, no mismatch or attentional orienting response will follow, and only single spikes will ascend. If higher areas are not primed, either through top-down predictions or alpha oscillations, then the single spikes may be insufficient to drive activation, preventing further processing of the ignored stimulus.

#### *Network Architecture*

The architecture is extremely simple. A single neuron representing a layer 2/3 pyramidal cell performing sensory comparison is the core of the model, as it is a single-neuron comparator.

#### *Network Connectivity*

While there are a high number of connections, they form a very simple pattern. The sensory comparator is innervated by 600 synapses: 200 basal driving synapses (presumed to be arriving from sensory thalamus via L4, which are both beyond the scope of this model), 200 apical priming synapses (presumably from non-specific thalamic neurons, per S. R. Jones et al., 2009b), and 200 basal inhibitory priming synapses (presumed to be from priming-driven feed-forward cortical interneurons). All synapses have weights that are randomly selected using a Gaussian distribution with  $\mu$  of 1.0 and  $\sigma$  of 0.1. The weights

are of the form  $w \cdot \text{gauss}(\mu, \sigma)$ , where the value of  $w$  is 16 for driving synapses, 12 for excitatory priming synapses, and 8 for inhibitory priming synapses.

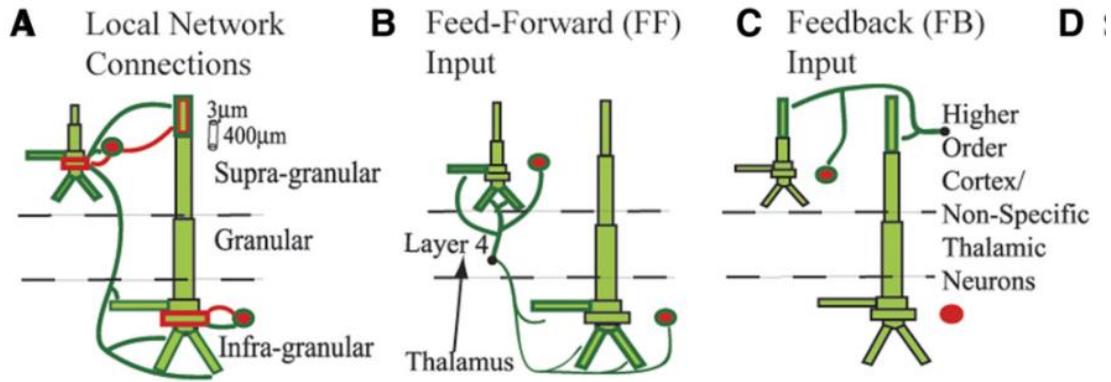


Figure 5.9: A simplified subset of S. Jones et al., 2009, provided for reference to the inputs to L2/3, specifically apical and feedforward inhibitory inputs from higher order cortex (in the MMN model in Chapter Four) and non-specific thalamic neurons seen in C, and the sensory input to basal dendrites from Layer 4, as seen in B.

There are three incoming connections in the model:

- Apical excitation (presumably from non-specific thalamic neurons)
- Basal Inhibition (presumably via L2/3 Interneurons)
- Sensory input (presumably via specific thalamus/L4)

### *Neuron and Synapse Models*

#### *Input*

Each of the hundreds of input synapses is driven by a Poisson process whose lambda parameter changes over time. In a Poisson probability distribution, the lambda parameter

represents both the mean and the variance, and a higher lambda means that the Poisson-governed event occurs more often.

Nonspecific thalamic inputs (inhibitory and excitatory) feature lambdas with a simple sine wave pattern, oscillating at 10 Hz. The sensory input synapse lambdas go from 0.0001 at rest to a given value during activation.

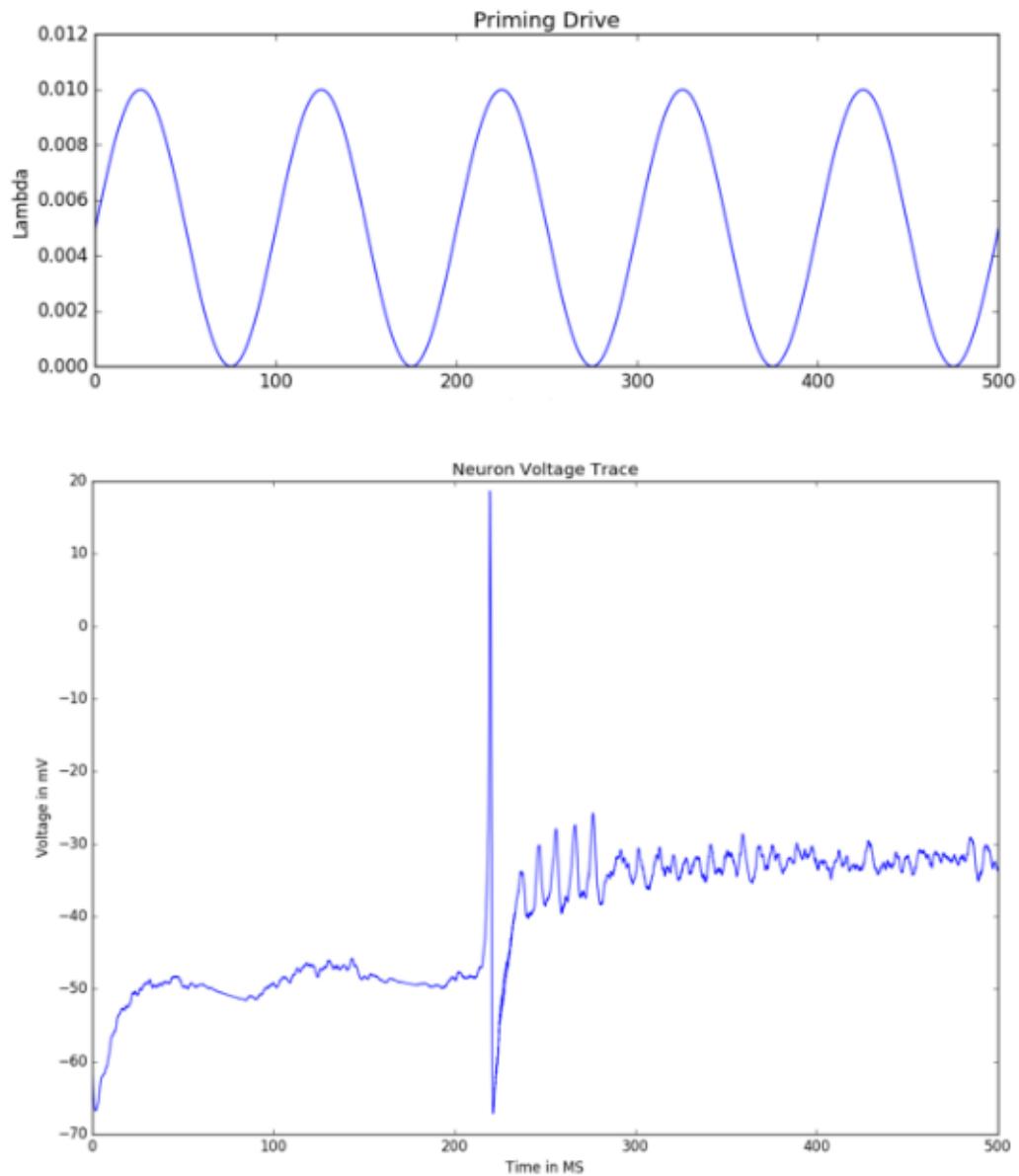
The phase of sensory input activation was varied throughout the 10 Hz cycle.

### *Output*

The membrane voltage and extracellular current were measured for analysis at both the apex and the base of the neuron, along with the net KCNQ-mediated M-current. In addition, we calculated the quasi-static electric field of the neuron using the methods provided in Chapter Three.

## **Results**

Only preliminary results are available, but the subthreshold alpha drive from thalamus to cortex has been shown to achieve a mode switch, drastically reducing the number of spikes fired in response to stimulus-driven input arriving during the peaks of alpha, but reducing the spike rate much less during the troughs. (Figure 5.10)



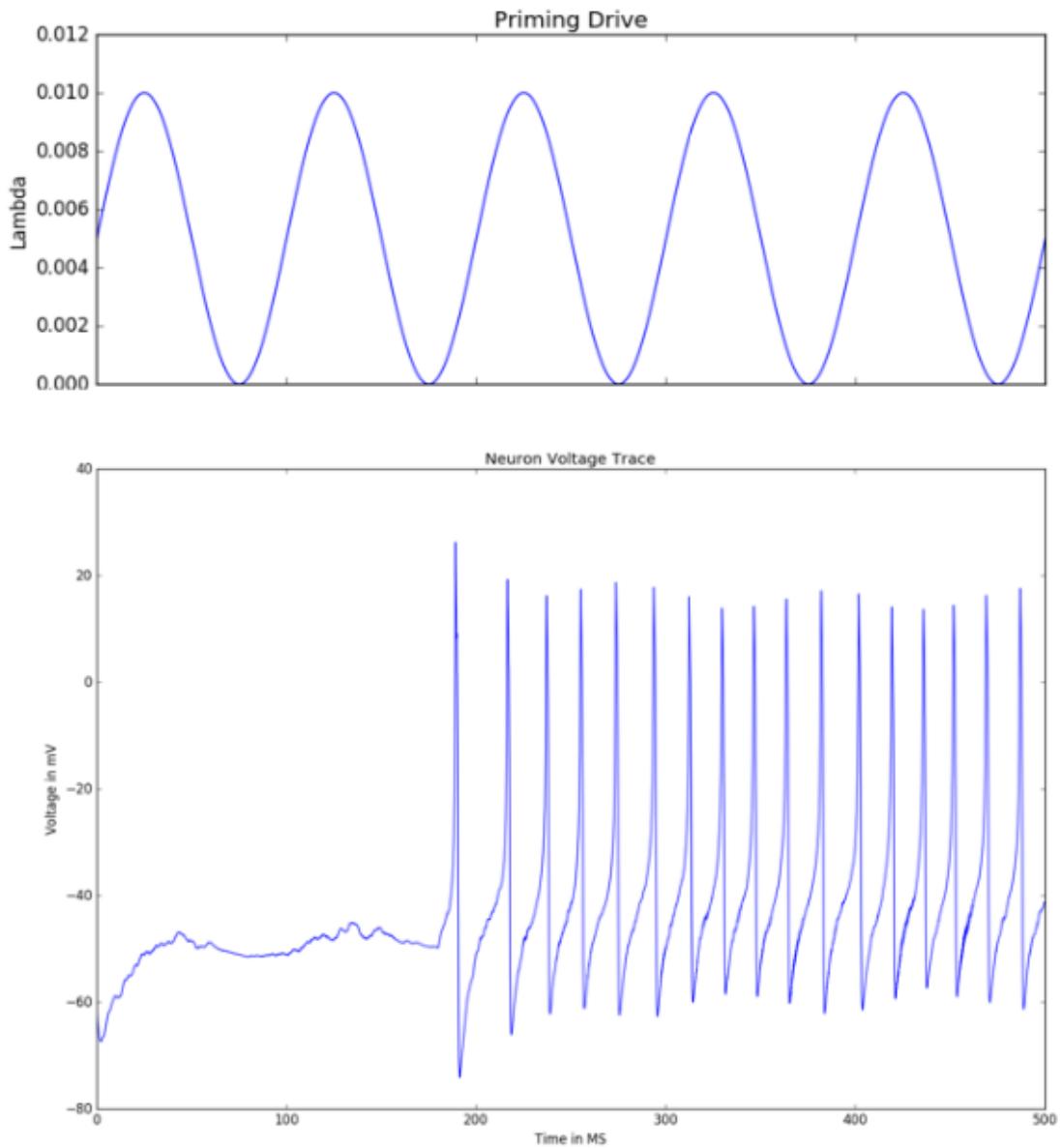


Figure 5.10: Upper Page One) the ongoing alpha drive, in the form of an oscillatory change in the priming lambda. Lower Page One) The cell membrane voltage trace, showing tonic firing when stimulus input arrives in the trough of alpha. Upper Page Two) The same ongoing alpha drive, reproduced for ease of comparison. Lower Page Two) The cell membrane voltage trace, showing phasic firing when stimulus input arrives at the peak of alpha.

These results suggest that the dynamically labeled predictive coding model of alpha proposed here is capable of reproducing the counterintuitive findings of Haegens et al. (2011), which observed not only a general reduction in firing rates as alpha power increased, but also an increase in firing during the polarized troughs of alpha.

## Discussion

The thalamic connections to the cortex in this case feature supragranular connections which drive both excitation of the apical dendrites and their feedforward inhibition. So the pulsed inhibition proposed in (Klimesch et al., 2007) is present in the new model during alpha, but such pulsed inhibition is far from the whole story. The thalamocortical system's resonance at ~10hz (Herrmann, 2001) is still relevant, in that it offers a mechanism by which this information suppression effect can be restricted to the alpha frequency range. Reducing the relative amplitude of theta and beta subthreshold oscillations can potentially prevent activation of the M-current, allowing neocortical pyramidal cells to respond to stimulus onsets with tonic firing, which generates sufficient prolonged activity to trigger an attentional orienting response. In contrast, the resonance would ensure that alpha subthreshold oscillations are of higher amplitude and therefore activate the M-current, priming the cells for a mode switch. This matches the finding of Haegens et al., (2011), by reducing the firing rate when the cortical cells are driven with suprathreshold input, such as a new stimulus onset. It also triggers the tonic-to-phasic mode shift, and so the cells oscillating in alpha will – for much of the alpha cycle, centered on the peak of the oscillation – respond only with phasic firing to stimulus onset. As described above, in a field of feature or object detectors, this effectively treats most possible inputs as expected.

This model also displays different firing rates at different phases of the alpha cycle, with the highest firing rates occurring in the trough, exactly as observed by Haegens et al. (2011)

By causing the membrane of cortical pyramidal cells in Layers 2/3 to fluctuate near, but still below the threshold, as well as non-specific thalamic drive of inhibitory neurons in the upper layers (E. G. Jones, 2001), alpha oscillations in this model result from a priming of the entire field in an oscillating manner. This suggests that diffuse thalamic projections to superficial layers – typical from GPi-inhibited thalamus but not from CBM-excited thalamus – are still extremely relevant, consistent with the hypothesis (presented earlier in this chapter) that alpha can be recruited via learned activation of the indirect pathway through the basal ganglia. This is also consistent with results indicating that the system can move very quickly into (and out of) alpha-suppressive states under learned guidance by cues that signal intervals when distractors may appear.

In conclusion, at this point there are two paths forward for the model. Either the higher areas are primed and ready to participate in the dynamically negotiated labeled line encoding scheme, in which case the information at stimulus onset that all is as expected continues up its chain in the cortical hierarchy, or else the higher areas aren't primed, in which case the single spikes coming in on what sensory cortex expects to be labeled lines will fail to elicit a suprathreshold response and the information will just die out. Either way, no attentional orienting response will be elicited, effectively allowing the owner of said alpha

oscillations to "ignore" the onset of a new stimulus for which the system has not been specifically primed.

## CHAPTER SIX: DISCUSSION AND FUTURE DIRECTIONS

This dissertation introduced the Dynamically Labeled Predictive Coding (DLPC) framework, a new proposal for the interpretation and modeling of expectation, comparison, and associative mismatch in the brain. Key features of this framework include:

- Code Switching at the level of individual neurons
- Differing representation of expected and unexpected information
- The use of dynamically established labeled line encoding
- Single-cell, non-subtractive comparison
- Efficiency in terms of spikes and downstream processing delay

This framework was then applied and validated through the construction of several models. First, a model of CAI hippocampal mismatch was developed, and shown to match experimental data. This was then expanded to model the neocortical MMN, and several non-obvious novel predictions resulted, including the potential dissociation of the recorded scalp voltage from the underlying neural mismatch signal and a novel proposed mechanism of action for the antipsychotic retigabine.

Non-DLPC models of alpha oscillations and alpha-induced information suppression were then explored, focusing on a saturation mechanism to suppress mutual information. A new proposal was also introduced for how the basal ganglia may control onset and offset of alpha-induced information suppression. Although these rate models could reproduce many experimental findings, they fell short of reproducing a key electrophysiological finding: phase-dependent reduction in spiking activity correlated with power in the alpha frequency band. To remedy this, a preliminary DLPC-based model of alpha-induced information suppression was introduced, suggesting a possible unification between mechanisms of anticipation and mechanisms of active ignoring.

In closing, the remainder of this chapter will be dedicated to a possible extension of the DLPC framework, specifically its interaction with reward circuitry and possible applications to musical neuroaesthetics.

### Expectations, Surprise, Resolution, and Reward: Sustainable Novelty

Surprise is a common element in music, and has been for centuries. There are many types of enjoyable musical surprise, from an unexpectedly loud note to a shift in instrument timbre. Most musical surprises, however, habituate rapidly, producing what might be called "novelty tracks" in the modern recording industry. An effective example of this might be 1958's, "The Chipmunk Song", which introduced America to the artificially sped-up, high-pitched voices of Alvin and the Chipmunks. The surprise was immediate and intense for listeners who had not previously encountered such production effects, but the novelty also wore off quickly.

Some musical passages, however, seem to resist habituation. A well-crafted hook or key change can often contain an element of surprise that retains its effectiveness over an adult's entire life, given that the passage or song isn't simply played on repeat for a sustained period. This suggests it is possible to shut down habituation and "sustain novelty". An explanation for exactly that sustainable novelty is what I am proposing as a future direction for my research.

### Down the Garden Path

There are key steps that seem to be common in all forms of sustainable novelty I have examined:

1. Expectations are built up
2. Associative mismatch occurs
3. A recontextualization resolves the surprise, creating a new, retroactive understanding

These three steps are proposed to interact with known brain processes in order to reliably evoke reward and a positive emotional valence. The processes involved in our proposal are:

- Musical pattern recognition, such as key and scale degree, as found in Janata et al. (2002).
- Intrinsic reward (preliminarily simplified to dopamine release) for successful prediction of recognized patterns, as shown in Satterthwaite et al. (2012).
- Habituation of reinforcer effectiveness via striatal inhibition of the ventral tegmental area (VTA), as shown in Lloyd et al. (2014).

Under this proposal, any novel music will elicit dopamine release when first heard, but as predictable patterns are recognized the striatum will inhibit the VTA, bringing it back down to baseline activity.

Sustainable novelty is proposed to be achieved first by establishing this highly predictable set of expectations exactly as above, which as above will not normally elicit reward by the time recognition is reliably correct. This is followed by step two, the associative mismatch, or

surprise. This also does not elicit reward, but what it will do is suppress striatal habituation of reinforcer effectiveness, through a cholinergic reset signal that follows mismatch (Tan & Bullock, 2008). Given the impact of ACh on DLPC comparators, this reset signal will remove expectational bias and the cue representations that could have acted via striatum to inhibit the VTA activation at the time of reward arrival. This in effect opens a short window of opportunity where a new, unexpected pattern can be recognized and will elicit increased dopamine release without inhibition on the part of the striatum.

As long as the musical piece's global statistics can continue leading the listener "down the garden path" of false expectations, then the above pattern will be robust to repeated exposures, allowing, in effect, a sustainable novelty response.

## BIBLIOGRAPHY

- Adams, P. R., & Brown, D. A. (1980). Bullfrog sympathetic neurones. *British Journal of Pharmacology*, 68, 353–355.
- Anis, N. A., Berry, S. C., Burton, N. R., & Lodge, D. (1983). The dissociative anaesthetics, ketamine and phencyclidine, selectively reduce excitation of central mammalian neurones by N-methyl-aspartate. *British Journal of Pharmacology*, 79, 565–575.
- Apicella, A. J., Wickersham, I. R., Seung, H. S., & Shepherd, G. M. G. (2012). Laminarly Orthogonal Excitation of Fast-Spiking and Low-Threshold-Spiking Interneurons in Mouse Motor Cortex. *The Journal of Neuroscience*, 32(20), 7021–7033. <http://doi.org/10.1523/JNEUROSCI.0011-12.2012>
- Aron A. R. (2011). From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biol Psychiatry*, 69(12), e55–e68.
- Baldeweg, T. (2007). ERP repetition effects and mismatch negativity generation. *Journal of Psychophysiology*, 21(3), 204–213. <http://doi.org/10.1027/0269-8803.21.34.204>
- Banquet, J., & Grossberg, S. (1987). Probing cognitive processes through the structure of event-related potentials during learning: an experimental and theoretical analysis. *Applied Optics*, 26(23), 4931–4946. <http://doi.org/10.1364/AO.26.004931>
- Barbas, H., & Rempel-Clower, N. (1997). Cortical structure predicts the pattern of corticocortical connections. *Cerebral Cortex*, 7(7), 635–646. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9373019>
- Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition*, 132(2), 137–150. <http://doi.org/10.1016/j.cognition.2014.03.013>
- Bell, A. J., Mainen, Z. F., & Sejnowski, T. J. (1994). Balancing of conductances may explain irregularity of cortical spiking. *Proceedings of the 1st Joint Symposium on Neural Computation*, 6, 1–5.
- Bottcher-Gandor, C., & Ullsperger, P. (1992). Mismatch negativity in event-related potentials to auditory stimuli as a function of varying interstimulus interval. *Psychophysiology*, 29(5), 546–550.

- Brown, D. A., & Adams, P. R. (1980). Muscarinic suppression of a novel voltage-sensitive K<sub>+</sub> current in a vertebrate neurone. *Nature*, 283(5748), 673–676.  
<http://doi.org/10.1038/283673a0>
- Brown J, Bullock D, Grossberg S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, 17, 471–510.
- Brunel, N., & Wang, X. (2001). Effects of neuromodulation in a cortical network model of object working. *Journal of Computational Neuroscience*, 11, 63–85.
- Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *The Journal of Neuroscience*, 29(24), 7869–7876.  
<http://doi.org/10.1523/JNEUROSCI.0113-09.2009>
- Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76, 838–846.
- Buzsáki, G., Grastyán, E., Tveritskaya, I. N., & Czopf, J. (1979). Hippocampal evoked potentials and EEG changes during classical conditioning in the rat. *Electroencephalography and clinical neurophysiology*, 47(1), 64–74.
- Ching S, Purdon P. L., Vijayan S, Kopell N. J., Brown E. N. (2012). A neurophysiological-metabolic model for burst suppression. *Proc Natl Acad Sci U S A*. 109(8): 3095–100.  
<http://doi.org/10.1073/pnas.1120004109>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204.  
<http://doi.org/10.1017/S0140525X12000477>
- Contreras, D., & Steriade, M. (1995). Cellular basis of EEG slow rhythms: a study of dynamic corticothalamic relationships. *The Journal of Neuroscience*, 15, 604–622.
- Dayan, P., & Abbott, L. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems (computational neuroscience)*. Cambridge, MA: MIT Press. <http://doi.org/10.1017/j.neuron.2008.10.019>
- Destexhe, A, Contreras, D., & Steriade, M. (1998). Mechanisms underlying the synchronizing action of corticothalamic feedback through inhibition of thalamic relay cells. *Journal of Neurophysiology*, 79, 999–1016.
- Destexhe, A., Rudolph, M., & Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nature Reviews. Neuroscience*, 4(9), 739–751.  
<https://doi.org/10.1038/nrn1198>
- Dolan, R. J., & Fletcher, P. C. (1997). Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature*, 388(6642), 582–585.

- Duncan, K., Curtis, C., & Davachi, L. (2009). Distinct memory signatures in the hippocampus: intentional states distinguish match and mismatch enhancement signals. *The Journal of Neuroscience*, 29(1), 131–139.
- Ehrlichman, R. S., Maxwell, C. R., Majumdar, S., & Siegel, S. J. (2008). Deviance-elicited changes in event-related potentials are attenuated by ketamine in mice. *Journal of Cognitive Neuroscience*, 20(8), 1403–1414.
- Engeland, C., Mahoney, C., Mohr, E., Ilivitsky, V., & Knott, V. J. (2002). Acute nicotine effects on auditory sensory memory in tacrine-treated and nontreated patients with Alzheimer's disease: an event-related potential study. *Pharmacology, Biochemistry and Behavior*, 72(1–2), 457–464.
- Farley, B. J., Quirk, M. C., Doherty, J. J., & Christian, E. P. (2010). Stimulus-specific adaptation in auditory cortex is an NMDA-independent process distinct from the sensory novelty encoded by the mismatch negativity. *The Journal of Neuroscience*, 30(49), 16475–16484. <http://doi.org/10.1523/JNEUROSCI.2793-10.2010>
- Fedorenko, O., Strutz-Seebohm, N., Henrion, U., Ureche, O. N., Lang, F., Seebohm, G., & Lang, U. E. (2008). A schizophrenia-linked mutation in PIP5K2A fails to activate neuronal M channels. *Psychopharmacology*, 199(1), 47–54. <http://doi.org/10.1007/s00213-008-1095-x>
- Fernandez, F. R., Mehaffey, W. H., Turner, R. W., & Fernando, R. (2005). Dendritic Na<sup>+</sup> current inactivation can increase cell excitability by delaying a somatic depolarizing afterpotential. *Journal of Neurophysiology*, 94, 3836–3848. <http://doi.org/10.1152/jn.00653.2005>
- Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, 2, 154. <http://doi.org/10.3389/fpsyg.2011.00154>
- Friedman, A. K., Juarez, B., Ku, S. M., Zhang, H., Calizo, R. C., Walsh, J. J., Chaudhury, D., Zhang, S., Hawkins, A., Dietz, D., Murrough, J. W., Ribadeneira, M., Wong, E., Neve, R., & Han, M. (2016). KCNQ channel openers reverse depressive symptoms via an active resilience mechanism. *Nature Communications*, 7, 11671.
- Freunberger, R., Fellinger, R., Sauseng, P., Gruber, W., & Klimesch, W. (2009). Dissociation between phase-locked and nonphase-locked alpha oscillations in a working memory task. *Human Brain Mapping*, 30(10), 3417–3425. <http://doi.org/10.1002/hbm.20766>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360(1456), 815–836. <http://doi.org/10.1098/rstb.2005.1622>

- Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical Neurophysiology*, 120(3), 453–463.
- Gruber A. J., McDonald R. J.. (2012). Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior. *Front Behav Neurosci*. 6:50.
- Guillery, R. W., & Sherman, S. M. (2002). Thalamic relay functions and their role in corticocortical communication: Generalizations from the visual system. *Neuron*, 33, 163–175. [http://doi.org/10.1016/S0896-6273\(01\)00582-7](http://doi.org/10.1016/S0896-6273(01)00582-7)
- Guntherope, M. J., Large, C. H., & Sankar, R. (2012). The mechanism of action of retigabine (ezogabine), a first-in-class K<sup>+</sup> channel opener for the treatment of epilepsy. *Epilepsia*, 53(3), 412–424.
- Gutfreund, Y., Yarom, Y., & Segev, I. (1995). Subthreshold oscillations and resonant frequency in guinea-pig cortical neurons: physiology and modelling. *The Journal of Physiology*, 483, 621–640.
- Hasselmo, M. E. (2005). The Role of Hippocampal Regions CA3 and CA1 in Matching Entorhinal Input With Retrieval of Associations Between Objects and Context: Theoretical Comment on Lee et al. (2005). *Behavioral Neuroscience*, 119(1), 342–345. <https://doi.org/10.1037/0735-7044.119.1.342>
- Haegens, S., Nácher, V., Luna, R., Romo, R., & Jensen, O. (2011). α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48), 19377–19382. <http://doi.org/10.1073/pnas.1117190108>
- Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Ligeois, C., Chauvel, P., & Musolino, A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. I. Superior temporal plane and parietal lobe, *Electroencephalography and clinical Neurophysiology*, 94, 191–220.
- Heekeren, K., Daumann, J., Neukirch, A., Stock, C., Kawohl, W., Norra, C., Waberski, T. D., Gouzoulis-Mayfrank, E. (2008). Mismatch negativity generation in the human 5HT2A agonist and NMDA antagonist model of psychosis. *Psychopharmacology*, 199, 77–88. <http://doi.org/10.1007/s00213-008-1129-4>
- Herrmann, C. S. (2001). Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137(3–4), 346–353. <http://doi.org/10.1007/s002210100682>

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, 69(2), 387–396.  
<http://doi.org/10.1016/j.neuron.2010.12.027>

Hodgkin, A. L. (1948). The local electric changes associated with repetitive action in a non-medullated axon. *The Journal of Physiology*, 107(2), 165–181.

Hughes, H. C., Darcey, T. M., Barkan, H. I., Williamson, P. D., Roberts, D. W., & Aslin, C. H. (2001). Responses of human auditory association cortex to the omission of an expected acoustic event. *NeuroImage*, 13(6), 1073–1089.  
<http://doi.org/10.1006/nimg.2001.0766>

Hughes, S. W., & Crunelli, V. (2005). Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *The Neuroscientist*, 11(4), 357–372.  
<http://doi.org/10.1177/1073858405277450>

Huguenard, J. R. (1996). Low-threshold calcium currents in central nervous system neurons. *Annual Review of Physiology*, 58, 329–48.  
<http://doi.org/10.1146/annurev.ph.58.030196.001553>

Hwang, K., Ghuman, A. S., Manoach, D. S., Jones, S. R., & Luna, B. (2014). Cortical Neurodynamics of Inhibitory Control. *Journal of Neuroscience*, 34(29), 9551–9561.  
<http://doi.org/10.1523/JNEUROSCI.4889-13.2014>

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., Hirabuki, N., Hiroshi, A., Taizo, K., Robinson, S. E., Takeda, M. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport*, 10(4), 675–679.  
<http://doi.org/10.1097/00001756-199903170-00003>

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, 14(6), 1569–1572.  
<http://doi.org/10.1109/TNN.2003.820440>

Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., Lin, F., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 6809–6814. <http://doi.org/10.1073/pnas.0303760101>

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *science*, 298(5601), 2167–2170.

Javitt, D. C., Steinschneider, M., Schroeder, C. E., & Arezzo, J. C. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proceedings of the National*

*Academy of Sciences of the United States of America*, 93(21), 11962–11967. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC38166/>

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, 4, 186. <http://doi.org/10.3389/fnhum.2010.00186>

Jentsch, T. J. (2000). Neuronal KCNQ potassium channels: physiology and role in disease. *Nature Reviews Neuroscience*, 1, 21–30. <http://doi.org/10.1038/35036198>

Jones, E. G. (2001). The thalamic matrix and thalamocortical synchrony. *Trends in Neurosciences*, 24(10), 595–601.

Jones, E. G. (2002). Thalamic circuitry and thalamocortical synchrony. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 357, 1659–1673. <http://doi.org/10.1098/rstb.2002.1168>

Jones, S. R., Pinto, D. J., Kaper, T. J., & Kopell, N. (2000). Alpha-frequency rhythms desynchronize over long cortical distances: a modeling study. *Journal of Computational Neuroscience*, 9, 271–291.

Jones, S. R., Pritchett, D. L., Sikora, M. A., Stufflebeam, S. M., Hämäläinen, M., & Moore, C. I. (2009). Quantitative analysis and biophysically realistic neural modeling of the MEG mu rhythm: rhythmogenesis and modulation of sensory-evoked responses. *Journal of Neurophysiology*, 102(6), 3554–3572. <http://doi.org/10.1152/jn.00535.2009>

Katz, Y., Kath, W. L., Spruston, N., & Hasselmo, M. E. (2007). Coincidence detection of place and temporal context in a network model of spiking hippocampal neurons. *PLoS Computational Biology*, 3(12), 2432–2445. <https://doi.org/10.1371/journal.pcbi.0030234>

Kim, K. S., Kobayashi, M., Takamatsu, K., & Tzingounis, A. V. (2012). Hippocalcin and KCNQ channels contribute to the kinetics of the slow afterhyperpolarization. *Biophysical journal*, 103(12), 2446–2454.

Kirk, I. J., & Mackay, J. C. (2003). The role of theta-range oscillations in synchronising and integrating activity in distributed mnemonic networks. *Cortex*, 39(4–5), 993–1008. [http://doi.org/10.1016/S0010-9452\(08\)70874-8](http://doi.org/10.1016/S0010-9452(08)70874-8)

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Research Reviews*, 53(1), 63–88. <http://doi.org/10.1016/j.brainresrev.2006.06.003>

Klinkenberg, I., Blokland, A., Riedel, W. J., & Sambeth, A. (2013). Cholinergic modulation of auditory processing, sensory gating and novelty detection in human

participants. *Psychopharmacology*, 225, 903–921. <http://doi.org/10.1007/s00213-012-2872-o>

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, 383(6597), 256.

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & De Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9), 2197–2206. <http://doi.org/10.1093/cercor/bhr310>

Kreitschmann-Andermahr, I., Rosburg, T., Demme, U., Gaser, E., Nowak, H., & Sauer, H. (2001). Effect of ketamine on the neuromagnetic mismatch field in healthy humans. *Brain Research. Cognitive Brain Research*, 12(1), 109–16. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11489614>

Kuchenbuch, A., Paraskevopoulos, E., Herholz, S. C., & Pantev, C. (2014). Audio-tactile integration and the influence of musical training. *PloS One*, 9(1), e85743. <http://doi.org/10.1371/journal.pone.0085743>

Kujala, A., & Näätänen, R. (2003). Auditory environment and change detection as indexed by the mismatch negativity (MMN). In J. Polich (Ed.), *Detection of Change: Event-Related Potential and fMRI Findings* (pp. 1–22). Norwell, MA: Kluwer Academic Publishers.

Kumaran, D., & Maguire, E. A. (2006). An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS biology*, 4(12), e424.

Kumaran, D., & Maguire, E. A. (2007). Match–mismatch processes underlie human hippocampal responses to associative novelty. *Journal of Neuroscience*, 27(32), 8517–8524.

Kurtzer, I., Herter, T. M., & Scott, S. H. (2005). Random change in cortical load representation suggests distinct control of posture and movement. *Nature Neuroscience*, 8(4), 498–504. <https://doi.org/10.1038/nn1420>

Larkman, A. U., Major, G., Stratford, K. J., & Jack, J. J. B. (1992). Dendritic morphology of pyramidal neurones of the visual cortex of the rat. IV: Electrical geometry. *Journal of Comparative Neurology*, 323(2), 137–152.

Leung, S., Croft, R. J., Baldeweg, T., & Nathan, P. J. (2007). Acute dopamine D1 and D2 receptor stimulation does not modulate mismatch negativity (MMN) in healthy human subjects. *Psychopharmacology*, 194(4), 443–451. <http://doi.org/10.1007/s00213-007-0865-1>

Lima, B., Singer, W., & Neuenschwander, S. (2011). Gamma responses correlate with temporal expectation in monkey primary visual cortex, *The Journal of Neuroscience*, 31(44), 15919–15931. <http://doi.org/10.1523/JNEUROSCI.0957-11.2011>

- Lisman, J. E., & Grace, A. a. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46(5), 703–713.  
<https://doi.org/10.1016/j.neuron.2005.05.002>
- Lisman, J. E., & Otmakhova, N. A. (2001). Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. *Hippocampus*, 11(5), 551–568.
- Lloyd, D. R., Hausknecht, K. A., & Richards, J. B. (2014). Nicotine and methamphetamine disrupt habituation of sensory reinforcer effectiveness in male rats. *Experimental and clinical psychopharmacology*, 22(2), 166.
- Lörincz, A., & Buzsáki, G. (2000). Two-phase computational model training long-term memories in the entorhinal-hippocampal region. *Annals of the New York Academy of Sciences*, 911(1), 83–III.
- Lorincz, M. L., Kékesi, K. A., Juhász, G., Crunelli, V., & Hughes, S. W. (2009). Temporal framing of thalamic relay-mode firing by phasic inhibition during the alpha rhythm. *Neuron*, 63(5), 683–696. <http://doi.org/10.1016/j.neuron.2009.08.012>
- Madison, D. V., & Nicoll, R. A. (1984). Control of the repetitive discharge of rat CA 1 pyramidal neurones in vitro. *The Journal of Physiology*, 354(1), 319–331.
- Mainen, Z. F., & Sejnowski, T. J. (1996). Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature*, 382(6589), 363–366.  
<http://doi.org/10.1038/382363a0>
- Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S. and Barone, P. (2014). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, 522(1), 225–259.
- Marrion, N. (1997). Control of M-current. *Annual Review of Physiology*, 59, 483–504.
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *The Journal of Neuroscience*, 29(9), 2725–2732. <http://doi.org/10.1523/JNEUROSCI.3963-08.2009>
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1), 66–122.  
<http://doi.org/10.1111/j.1469-8986.2009.00856.x>
- McCarthy M. M., Brown E. N., Kopell N. (2008). Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *J Neurosci*. 28(50), 13488–13504.

- Meeter, M., Murre, J. M. J., & Talamini, L. M. (2004). Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. *Hippocampus*, 14(6), 722–741.
- Mirpour, K., Bisley, J. W., & Bisley, J. W. (2013). Evidence for differential top-down and bottom-up suppression in posterior parietal cortex. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 368(1628): 20130069.
- Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta." *Progress in Neurobiology*, 86(3), 156–185. <http://doi.org/10.1016/j.pneurobio.2008.09.005>
- Müller, J.M. (1965/1838) On the specific energies of nerves. In R. Herrnstein and E. Boring (Eds.), *Sourcebook of the History of Psychology*. Cambridge, MA: Harvard University Press.
- Munoz D. P., Everling S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nat Rev Neurosci*. 5(3), 218–228.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13(2), 201–233.
- Näätänen, R., & Alho, K. (1995). Mismatch negativity—a unique measure of sensory processing in audition. *International Journal of Neuroscience*, 80(1–4), 317–337.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology*, 118(12), 2544–2590.
- Noonan, M. P., Adamian, N., Pike, X. A., Printzlau, F., Crittenden, B. M., & Stokes, M. G. (2016). Distinct mechanisms for distractor suppression and target facilitation. *Journal of Neuroscience*, 36(6), 1797–1807. <http://doi.org/10.1523/JNEUROSCI.2133-15.2016>
- Nordlie, E., Gewaltig, M., & Plesser, H. E. (2009). Towards reproducible descriptions of neuronal network models. *PLoS Computational Biology*, 5(8), e1000456. <http://doi.org/10.1371/journal.pcbi.1000456>
- Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. New York, NY: Oxford University Press.
- Oceák, A., Winkler, I., Sussman, E., & Alho, K. (2006). Loudness summation and the mismatch negativity event-related brain potential in humans. *Psychophysiology*, 43(1), 13–20. <http://doi.org/10.1111/j.1469-8986.2006.00372.x>

- Otto, T., & Eichenbaum, H. (1992). Neuronal activity in the hippocampus during delayed non-match to sample performance in rats: Evidence for hippocampal processing in recognition memory. *Hippocampus*, 2(3), 323-334.
- Overk, C. R., Felder, C. C., Tu, Y., Schober, D. A., Bales, K. R., Wuu, J., & Mufson, E. J. (2010). Cortical M1 receptor concentration increases without a concomitant change in function in Alzheimer's disease. *Journal of Chemical Neuroanatomy*, 40(1), 63-70. <http://doi.org/10.1016/j.jchemneu.2010.03.005>
- Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*, 38(2), 359-65. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11347880>
- Palva, J. M., Monto, S., Kulashekhar, S., & Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16), 7580-7585. <http://doi.org/10.1073/pnas.0913113107>
- Panzeri, S., Senatore, R., Montemurro, M. A., & Petersen, R. S. (2007). Correcting for the sampling bias problem in spike train information measures. *Journal of Neurophysiology*, 98(3), 1064-72. <http://doi.org/10.1152/jn.00559.2007>
- Payne, L., Guillory, S., & Sekuler, R. (2013). Attention-modulated alpha-band oscillations protect against intrusion of irrelevant information. *Journal of Cognitive Neuroscience*, 25(9), 1463-1476. <http://doi.org/10.1162/jocn>
- Pazo-alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology*, 63, 199-236.
- Pekkonen, E., Hirvonen, J., Jääskeläinen, I. P., Kaakkola, S., & Huttunen, J. (2001). Auditory sensory memory and the cholinergic system: implications for Alzheimer's disease. *NeuroImage*, 14(2), 376-382. <http://doi.org/10.1006/nimg.2001.0805>
- Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neurosciences Research Program Bulletin*, 6(3), 221-343
- Peters, H. C., Hu, H., Pongs, O., Storm, J. F., & Isbrandt, D. (2005). Conditional transgenic suppression of M channels in mouse brain reveals functions in neuronal excitability, resonance and behavior. *Nature Neuroscience*, 8(1), 51-60. <http://doi.org/10.1038/nn1375>
- Prescott, S. A., Ratté, S., De Koninck, Y., & Sejnowski, T. J. (2006). Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. *The Journal of Neuroscience*, 26(36), 9084-9097. <http://doi.org/10.1523/JNEUROSCI.1388-06.2006>

- Prescott, S. A., Ratté, S., De Koninck, Y., & Sejnowski, T. J. (2008). Pyramidal neurons switch from integrators in vitro to resonators under in vivo-like conditions. *Journal of neurophysiology*, 100(6), 3030–3042.
- Prescott, S. A., & Sejnowski, T. J. (2008). Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *The Journal of Neuroscience*, 28(50), 13649–13661.
- Rall, W. (1962). Electrophysiology of a dendritic neuron model. *Biophysical journal*, 2(2), 145–167.
- Riekkinen, P., Pääkkönen, A., Karhu, J., Partanen, J., Soininen, H., Laakso, M., & Riekkinen, P. (1997). THA disrupts mismatch negativity in Alzheimer disease. *Psychopharmacology*, 133(2), 203–206. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9342788>
- Rinzel, J., & Ermentrout, G. B. (1998). Analysis of neural excitability and oscillations. In C. Koch & I. Segev (Eds.), *Methods in Neuronal Modeling* (pp. 251–292). Cambridge, MA: MIT Press.
- Romei, V., Gross, J., & Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *The Journal of Neuroscience*, 30(25), 8692–8697. <http://doi.org/10.1523/JNEUROSCI.0160-10.2010>
- Sabri, M., Radnovich, A. J., Li, T. Q., & Kareken, D. A. (2005). Neural correlates of olfactory change detection. *NeuroImage*, 25, 969–974. <http://doi.org/10.1016/j.neuroimage.2004.12.033>
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials*, 62(6), 437–448. [http://doi.org/10.1016/0168-5597\(85\)90054-1](http://doi.org/10.1016/0168-5597(85)90054-1)
- Satterthwaite, T. D., Ruparel, K., Loughead, J., Elliott, M. A., Gerraty, R. T., Calkins, M. E., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2012). Being right is its own reward: Load and performance related ventral striatum activation to correct responses during a working memory task in youth. *Neuroimage*, 61(3), 723–729.
- Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and cognition*, 112, 92–97.
- Sarnthein, J., Morel, A., Von Stein, A., & Jeanmonod, D. (2005). Thalamocortical theta coherence in neurological patients at rest and during a working memory task. *International Journal of Psychophysiology*, 57(2), 87–96. <http://doi.org/10.1016/j.ijpsycho.2005.03.015>
- Sarter M, Gehring W. J., Kozak R. (2006). More attention must be paid: the neurobiology

- of attentional effort. *Brain Res Rev.* 51(2), 145–160.
- Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66(3), 605–610. <http://doi.org/10.1093/biomet/66.3.605>
- Sherman, S. M., & Guillery, R. W. (2004). Thalamus. In G. M. Shepherd (Ed.), *The Synaptic Organization of the Brain*. (5th ed., pp. 311–360). New York, NY: Oxford University Press.
- Shipp, S. (2016). Neural Elements for Predictive Coding. *Frontiers in Psychology*, 7, 1792. <http://doi.org/10.3389/fpsyg.2016.01792>
- Soty, F., Damgaard, T., Montezinho, L. P., Mørk, A., Olsen, C. K., Bundgaard, C., & Husum, H. (2009). Antipsychotic-like effect of retigabine [N-(2-Amino-4-(fluorobenzylamino)-phenyl) carbamic acid ester], a KCNQ potassium channel opener, via modulation of mesolimbic dopaminergic neurotransmission, 328(3), 951–962. <http://doi.org/10.1124/jpet.108.146944.and>
- Sterman MB, Egner T. (2006). Foundation and practice of neurofeedback for the treatment of epilepsy. *Appl Psychophysiol Biofeedback*. 31(1), 21–35.
- Suffczynski, P., Kalitzin, S., Pfurtscheller, G., & Lopes da Silva, F. H. (2001). Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention. *International Journal of Psychophysiology*, 43(1), 25–40.
- Tan, C. O., & Bullock, D. (2008). A local circuit model of learned striatal and dopamine cell responses under probabilistic schedules of reward. *The Journal of Neuroscience*, 28(40), 10062–10074. <http://doi.org/10.1523/JNEUROSCI.0259-08.2008>
- Tervaniemi, M., Saarinen, J., Paavilainen, P., Danilova, N., & Näätänen, R. (1994). Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biological Psychology*, 38(2–3), 157–167. [http://doi.org/10.1016/0301-0511\(94\)90036-1](http://doi.org/10.1016/0301-0511(94)90036-1)
- Uhlenbeck, G. E., & Ornstein, L. C. (1930). On the theory of the Brownian motion. *Physical Review*, 36, 823–841.
- Umbrecht, D., Koller, R., Vollenweider, F. X., & Schmid, L. (2002). Mismatch negativity predicts psychotic experiences induced by NMDA receptor antagonist in healthy volunteers. *Biological Psychiatry*, 51(5), 400–406. [http://doi.org/10.1016/S0006-3223\(01\)01242-2](http://doi.org/10.1016/S0006-3223(01)01242-2)
- VanRullen, R., Busch, N. A., Drewes, J., & Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Frontiers in Psychology*, 2, 60. <http://doi.org/10.3389/fpsyg.2011.00060>

- VanRullen, R., & Macdonald, J. S. P. (2012). Perceptual echoes at 10 Hz in the human brain. *Current Biology*, 22(11), 995–999. <http://doi.org/10.1016/j.cub.2012.03.050>
- Vijayan S, Ching S, Purdon P. L., Brown E. N., Kopell N. J. (2013). Thalamocortical mechanisms for the anteriorization of  $\alpha$  rhythms during propofol-induced unconsciousness. *J Neurosci*. 33(27), 11070-11075.
- Vijayan S, Kopell N. J. (2012). Thalamic model of awake alpha oscillations and implications for stimulus processing. *Proc Natl Acad Sci U S A*. 109(45), 18553-18558.
- Vinogradova, O. S. (2001). Hippocampus as comparator: Role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*, 11(5), 578–598. <https://doi.org/10.1002/hipo.1073>
- Vinogradova, O. S. (1984). Current concepts of general properties and plastic phenomena in hippocampal neurons. *Uspekhi fiziologicheskikh nauk*, 15(1), 28-54.
- Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, 24(2), 560–564. <http://doi.org/10.1016/j.neuroimage.2004.08.039>
- Wacongne, C., Changeux, J., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *The Journal of Neuroscience*, 32(11), 3665–3678. <http://doi.org/10.1523/JNEUROSCI.5003-11.2012>
- Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, 21(3-4), 147-163.
- Wolfart, J., Debay, D., Le Masson, G., Destexhe, A., & Bal, T. (2005). Synaptic background activity controls spike transfer from thalamus to cortex. *Nature neuroscience*, 8(12), 1760-1768.
- Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *The Journal of Neuroscience*, 20(RC63), 1-6.
- Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R. J., & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology*, 35(5), 615-619.
- Yamada, W., Koch, C., & Adams, P. (1989). Multiple channels and calcium dynamics. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling* (pp. 97–134). Cambridge, MA: MIT Press.



## CURRICULUM VITAE

**Stefan Berteau**

677 Beacon St, Boston MA 02215

202-213-0312 ♦ [stefan.berteau@gmail.com](mailto:stefan.berteau@gmail.com)

### **Education**

#### **Boston University, Department of Cognitive and Neural Systems (2008-2016)**

Boston, MA – PhD Candidate in Cognitive and Neural Systems

Key coursework in vision, audition, motor control, learning and reward, nonlinear mathematics, multivariate statistical analysis, developmental neurobiology, electrophysiology, and systems neuroscience.

#### **FAES Graduate School at the National Institutes of Health (2006-2007)**

Bethesda, MD – Continuing Education

Basic Human Neuroanatomy, including anatomy lab work on human brains and spinal cords

Selected to observe experimental brain surgeries from within the operating room

#### **American University, College of Arts and Sciences (1998-2002)**

Washington, DC – Bachelors of Science in Multimedia Design and Development, Computer Science Focus

Key coursework in Graphic Design, User Interface Design, Sound Design, Typography, AI Programming, 2D/3D Graphics Programming

#### **Workshops and Certifications**

NIH Human Subjects Certification

Fast Optical Imaging Summer Workshop at the Beckman Institute, University of Illinois at Urbana-Champaign

### **Publications, Presentations, and Patents**

#### **Peer-Reviewed Papers**

Berteau, S., Bullock, D., Sekuler, B., and Miller, P., "A Mechanistic Model of Alpha-Induced Information Suppression". In submission, copy provided upon request.

Berteau, S., Mingolla, E., Bullock, D., "A Biophysical Model of the Mismatch Negativity". In preparation, copy provided upon request.

Berteau, S., Mingolla, E., Bullock, D., "Predicting Scalp Voltage from Local Circuit Models". In preparation, copy provided upon request.

#### **Talks**

2015 Inter-Science of Learning Center (iSLC) Annual Meeting, "How to be a critical reader of Neuroscience Articles", co-presented with Jenessa Seymour

SecureIT 2004, and other conferences, "Spyware, Uninvited and Unwanted"

#### **Posters**

2015 Cognitive Neuroscience Society Annual Meeting, "A Mechanistic Model of Alpha-Induced Information Suppression"

2014 Computational and Systems Neuroscience (Cosyne) Annual Meeting, "Alpha coherence networks and the prerequisites for communication in simulated cortex"

2013 Cognitive Neuroscience Society Annual Meeting, "A Biophysical Perspective on Cortical Comparators in Mismatch Negativity"

2012 Society for Neuroscience Annual Meeting, "Changes in Mnemonic Activity Drive Task-Dependent Alpha Band Oscillations: A Network Model for Effective Ignoring"

2011 Society for Neuroscience Annual Meeting: "Network analysis in EEG source-space of the dynamics that support intentional ignoring of visual stimuli"

## **Patents**

Berteau, Stefan A. 2012. "System and Method for Bayesian Text Classification". U.S. Patent 8,103,110, filed May 15, 2008, and issued January 24, 2012

Berteau, Stefan A.; Grucz, Michael L.; Goldsmith, Kevin C. 2012. "System and Method for Identifying a Cookie as a Privacy Threat". U.S. Patent 8,196,176, filed April 16, 2008, and issued June 5, 2012

### **Research Experience**

#### **Brandeis University, Waltham MA**

##### ***Department of Psychology, Visiting Research Scientist (March 2011 – Present)***

Performed NSF and NIH funded research analyzing experimental data on and creating mathematical models of alpha oscillations.

Prototyped novel methods of analysis, including a method of analyzing EEG coherence networks while preserving phase information and a novel application of superparamagnetic clustering

Constructed the first computational model of the role of alpha oscillations in intentional ignoring

#### **Boston University, Boston MA**

##### ***Department of Cognitive and Neural Systems Research Assistant and Teaching Fellow (September 2008 – Present)***

Performed NSF funded research designing experiments and constructing models of cortical development, learning, expectation formation, and reward systems in the brain

Worked with complex systems consisting of dozens of interconnected differential equations

Directed and designed my own plan of research

Programmed and piloted experimental studies using eye trackers and EEG

Teaching Fellow for a graduate-level course in vision science. Led discussion sections, designed grading criteria and evaluated projects and exams

Programmed simulations in Python, C, and Matlab

**CA (Formerly Computer Associates) and HCL Technologies, Herndon, VA**

**Senior Research Engineer, Malware Threat Research (August 2004 – August 2008)**

Authored industry standards as CA's founding representative in the Anti-Spyware Coalition

Developed new applications based on expert systems and neural networks

Applied Bayesian statistics to the classification of EULAs

Designed economic models used to predict the future behavior of spyware authors

Developed internal tools using Java, Delphi, Python, and C

Researched and prototyped automated categorization and detection of software threats

Researched and prototyped analysis of complex peer-to-peer botnets via network tomography

Reverse-engineering complex malicious applications including kernel-mode driver code

Negotiated changes in software/privacy practices with high-profile companies such as Apple, Sony/BMG, Facebook, and Paypal

Addressed conferences, U.S. Congressional committees, and state legislatures as media/legislative spokesperson

Managed teams of up to five individuals to accomplish the tasks listed above