University of Sheffield

Modelling the cortical representation of infrequent stimuli

Chao Han

Supervisor: Prof. Eleni Vasilaki, Dr. Hannes Saal

This dissertation is submitted for the degree of
Doctor of Philosophy

in the

Department of Computer Science

January 7, 2022

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Chao Han
June 13, 2022

Acknowledgements

In the first place, I would like to thank my primary supervisor, Prof. Eleni Vasilaki, and secondary supervisor, Dr. Hannes Saal, for their wise guidance, valuable advice, consistent support and patience throughout my whole PhD. Eleni instructed me how to become a better researcher, especially the communication and presentation skills that have been steadily improved under her guidance. Hannes delivered many helpful suggestions on my PhD project and meticulous comments on my publications. Most of all, I want to thank them for tailoring their supervision to my working style and providing a friendly academic environment.

Special thanks to my research collaborators, Dr. Wolfger von der Behrens, Prof. Giacomo Indiveri, Dr. Aditya Gilra, Dr. Natacha Vanattou-Saïfoudine, and Gwendolyn English, for their productive discussion during my visit to their labs in Zurich and virtual meetings and constructive feedback on our papers.

I would also like to thank my labmates and friends: Dr. Avgoustinos Vouros, Nada AbdelRahman, Dr. Matt Ellis, Luca Manneschi, Matthew Whelan, Zehai Tu, Dr. Yan Ge, Shuo Zhou, Xianyuan Liu, Li Zhang, Lawrence Schobs, Peizhen Bai, Chunchao Ma, Dr. Tianqi Wei, and Juan José Giraldo, who offered me great help and support during my PhD journey.

Finally, I am enormously grateful to my parents who are always mentally by my side, respect and unconditionally support each of my decisions.

Abstract

In complex natural environments, sensory systems are constantly exposed to a large stream of inputs. In order to encode the everchanging sensory surroundings efficiently, sensory systems rapidly adapt their responses to common stimuli that are parts of predictable patterns while retain their responsiveness to the novel, unexpected stimuli that usually contain behaviourally important information. Such a differential processing capacity is commonly referred to as novelty detection, whose neural signatures have been identified in human electroencephalogram (EEG) recording as mismatch negativity (MMN) and single-neuron electrophysiology as stimulus-specific adaptation (SSA). Despite the prevalence of the rare-selective responses in several sensory modalities, the intracellular and network mechanisms underlying are not yet fully understood.

In this thesis, we employ the modeling tools in the computational neuroscience community, combined with computer simulation, to investigate the neuronal network dynamics producing novelty-detecting signals in the auditory and somatosensory cortex of the rodent model. Following an earlier auditory SSA mechanism in recurrent network with synaptic depression [Yarden and Nelken, 2017], we provide further support for the SSA circuitry hypothesis by using for the first time a similar recurrent network structure yet mediated with neuronal adaptation implemented in both simulated mean-field and physical neuromorphic spiking models. Our results suggest that SSA and its pertinent properties could arise from the differential excitation of tuned neuronal populations via propagation of population activity and short-term adaptation mechanism (neuronal adaptation or synaptic depression). In addition, we adapted and expanded this cortical circuit of the auditory SSA to a multi-scale thalamocortical network model that explains and reproduces physiologically observed neuronal response patterns that demonstrate signatures of both SSA and MMN in the somatosensory cortex [Musall et al., 2017]. Specifically, our results indicate that the novelty signal arises from the complex recurrent interplay between thalamic neurons and cortical neurons in layers 4 and 6. This work therefore provides a concrete mechanism that can serve as a starting point for further investigating the neural circuit mechanisms underlying novelty detection.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

To survive is to adapt to the world we live in. Organisms adapt over disparate timescales. At timescales of generations and lifetimes, species adapt features that are advantageous for survival over the course of evolution. At slow timescales related to the development of individual organisms, animals adapt their behaviors to changes in their surroundings. For example, polar bears hibernate in a den during the snowy wintertime and revive to search for food when spring comes. At timescales of real-time interactions, we all have experienced that we become easily used to the background noises and no longer conscious of their existence when we focus on other things. Although these phenomena stem from different mechanisms, they manifest an idea in common: adaptation is changes in an organism's functions that develop over diverse timescales in response to environmental changes.

The sensory systems especially demonstrate a compelling capacity of fast adaptation to the ever-changing sensory scene occurring over timescales from milliseconds to seconds, which is omnipresent at all sensory pathways and has a profound effect on perceptual processing. One of the simplest distributions of sensory stimuli is the binary distribution, where only two possible values of a single-dimensional stimulus, such as the sound frequency or magnitude, are taken, one presented frequently and the other rarely. This stimulation paradigm has been widely used in the study of the capacity of the sensory system to highlight novel or rare stimuli, which are often associated with behaviorally important events, embedded in a sequence of different, repeated stimuli, which has less relevance. Such novelty detecting capacity is favorable for human and animal survival as the brain selectively reduces its response to the steady sensory background to save energy, but remains sensitive to rare stimuli that may signify the appearance of prey or predators.

On the level of single neurons, the novelty-detecting ability is embodied in stimulus-specific adaptation (SSA), characterized by the phenomenon that neuronal responses decrease with repetitive stimuli, but recover when a different rare stimulus is presented. SSA is a robust and widespread phenomenon found in multiple sensory modalities across several species. On the level of human EEG recording, the neural signature of novelty detection has been identified as mismatch negativity (MMN) featured by a deviance-sensitive component of event-related potentials. Despite huge similarities, SSA is not likely to be the direct neuronal correlate of MMN due to their distinct responses latencies and sensitivity to the neurotransmitter NMDA, but probably one of the upstream mechanisms giving rise to MMN.

The overall objective of this dissertation is to elucidate the intrinsic and network mecha-

nisms giving rise to cortical SSA, with an emphasis on auditory and somatosensory modalities, using computational modeling methods. Particularly, we are devoted to modeling a delayed deviant-specific response speculated as a neurophysiological origin of MMN [Musall et al., 2017], and providing a new perspective of linking the SSA evident in single-cell electrophysiology with the EEG-recorded MMN via our model.

The structure and research contributions of the dissertation are outlined as follows:

Chapter 2: "Background" starts with a brief introduction to the rapid adaptation of the sensory system and the significance of the phenomenon in the context of a sensory encoding framework called linear–nonlinear-Poisson (LNP) models. It then focuses on a special form of rapid sensory adaption named stimulus-specific adaptation (SSA) and its relevance to novelty detection. A literature survey of SSA models follows up, and the modeling of short-term synaptic plasticity, serving as a critical ingredient of SSA models, is also discussed. The chapter concludes with the mathematical formulation of the mean-field rate model and geometrical analysis of its dynamics, which will be the theoretical basis of our SSA models in different sensory modalities discussed in the following chapters.

Contributions: We extend a pre-existing derivation on the rate network model coupled with constant synapses to the case of adapting synapses. Besides, we provide an alternative geometrical interpretation for an earlier dynamical analysis on a recurrent population with depressing synapses and additional original work on the hysteretic and synchronous behaviors occurred in this network configuration. This work is instructive for understanding the different patterns of the population response to a step impulse.

Chapter 3: "A Robust Model of Auditory Stimulus-Specific Adaptation" proposes a possible mechanism for auditory SSA in a recurrent network with feedforward adaptation, which is a simplified version of a previously reported SSA model [Yarden and Nelken, 2017]. We theoretically compare responses of our mean-field network model across multiple experimental protocols, which provides practical ways of tuning the spiking model counterpart implemented on neuromorphic hardware to match the experimental recordings. We also performed a robustness test of our model against online parameter perturbations and sensitivity analysis by providing operating ranges of network parameters.

Contributions: The work identifies interactions of population activity of specifically-tuned neurons, alongside cross-population adaption (spike-frequency adaptation or short-term synaptic depression), as two essential mechanistic primitives for generating SSA and novelty detection. Theoretical predictions are provided on the model responses to various SSA paradigms. The numerical results are qualitatively consistent with the experimental observation, even in the presence of online disturbances for model parameters.

Chapter 4: "Modelling Novelty Detection in the Thalamocortical Loop" develops a thalamocortical network model that accounts for novelty detection and reproduces physiologically observed neural response patterns in the whisker-related part of the somatosensory cortex. Specifically, our results demonstrate that the novelty signal arises from the complex recurrent interplay between thalamic neurons and cortical neurons in layers 4 and 6

over multiple timescales. We also test the ability of our model to distinguish between stimuli with higher or lower predictability, making it verifiable with future experiments. Finally, we discuss several limitations of our model and potential ways to improve and expand the model's performance in future work.

Contributions: To the best of our knowledge, our model is the first computational model endeavoring to elucidate the mechanism of the novelty-specific response observed in single sensory neurons, and link it to another EEG-based novelty-detecting phenomenon called mismatch negativity (MMN), concerning their common sensitivity to presentation context and timescales of response latency. Furthermore, our model provides predictions for experimental tests of our model's capacity in detecting the unexpected in more elaborated paradigms.

Chapter 5: "Conclusions and Future Work" summarizes the key contributions of my PhD research and proposes some ideas to improve and expand my current models for future work.

The content of the thesis is primarily based on the following publications produced during my PhD journey.

Vanattou-Saïfoudine, N., **Han, C.**, Krause, R., Vasilaki, E., von der Behrens, W. and Indiveri, G., 2021. A robust model of Stimulus-Specific Adaptation validated on neuromorphic hardware. Scientific reports, 11(1), pp.1-15.

Han, C., English, G., Saal, H. P., Indiveri, G., Gilra, A, von der Behrens, W., and Vasilaki, E. (2021). Modelling novelty detection in the thalamocortical loop. *Submitted to PLOS Computational Biology.*

Chapter 2

Background

The chapter aims to provide readers with both experimental and mathematical background knowledge relevant to my PhD research on the cortical processing of novel stimuli. It begins with proposing a possible explanation for the significance of general sensory adaption from the point of the efficiency of encoding sensory stimuli using the liner-nonlinear-Poisson model in section 2.1. The following section provides a brief background of the physiological basis of electroencephalograph (EEG) signals. Section 2.3 then focuses on a particular type of sensory adaptation called stimulus-specific adaptation, and further discusses its relation to its EEG correlate termed mismatch negativity and the functional importance of both phenomena in detecting novel signals. A review of computational SSA models is presented in section 2.4, pointing out that none of the current models can explain a recent experimental finding of late-phase SSA that is one of my research topics. A phenomenological model of short-term synaptic plasticity as an essential ingredient of SSA models is introduced in section 2.5. The chapter concludes with my original work on the derivation of mean-field recurrent networks with depressing synapses and the analysis of the network dynamics in sections 2.6 and 2.7, respectively.

2.1 Rapid Sensory Adaptation

Adaptation is prominent in all sensory modalities, which occurs at multiple timescales caused by different mechanisms. Among them, rapid sensory adaptation at timescales from milliseconds to seconds is one of the well-studied cases, which is found omnipresent at all stages of sensory pathways and has an important impact on the processing of sensory stimuli both neurophysiologically and perceptually [Whitmire and Stanley, 2016; Wark et al., 2007; Series et al., 2009; Fairhall et al., 2001]. Here we try to answer the question raised naturally: what role does this rapid adaptation play in sensory processing?

In order to probe the potential purpose of this adaptive phenomenon, we need first to consider how an individual neuron encodes the stream of sensory information into its corresponding spike sequences. The linear-nonlinear-Poisson (LNP) model has been widely used as an encoding framework of single sensory neurons to capture their response properties [Schwartz et al., 2006]. As can be seen from Fig 2.1, sensory stimuli are first linearly filtered by a set of given temporal features of a neuron in question. These filtered stimuli are then summed and fed to the neuron's gain function, whose output defines the instantaneous firing

rate of the following Poisson spike generator. The spike-triggered average of white-noise stimuli is used to estimate linear filters and accordingly determine a related gain function [Schwartz et al., 2006]. The linear filtering stage can be envisaged as the temporal feature selectivity driven by synaptic inputs converging to a given neuron, and the nonlinear activation stage as intrinsic neuronal properties resulting in the spike emission given the filtered inputs. In the context of the LNP model, one can regard sensory adaptation as changes in the encoding properties of the cascade, which arise from modifications in synaptically-driven feature filters and/or neuronal gain.

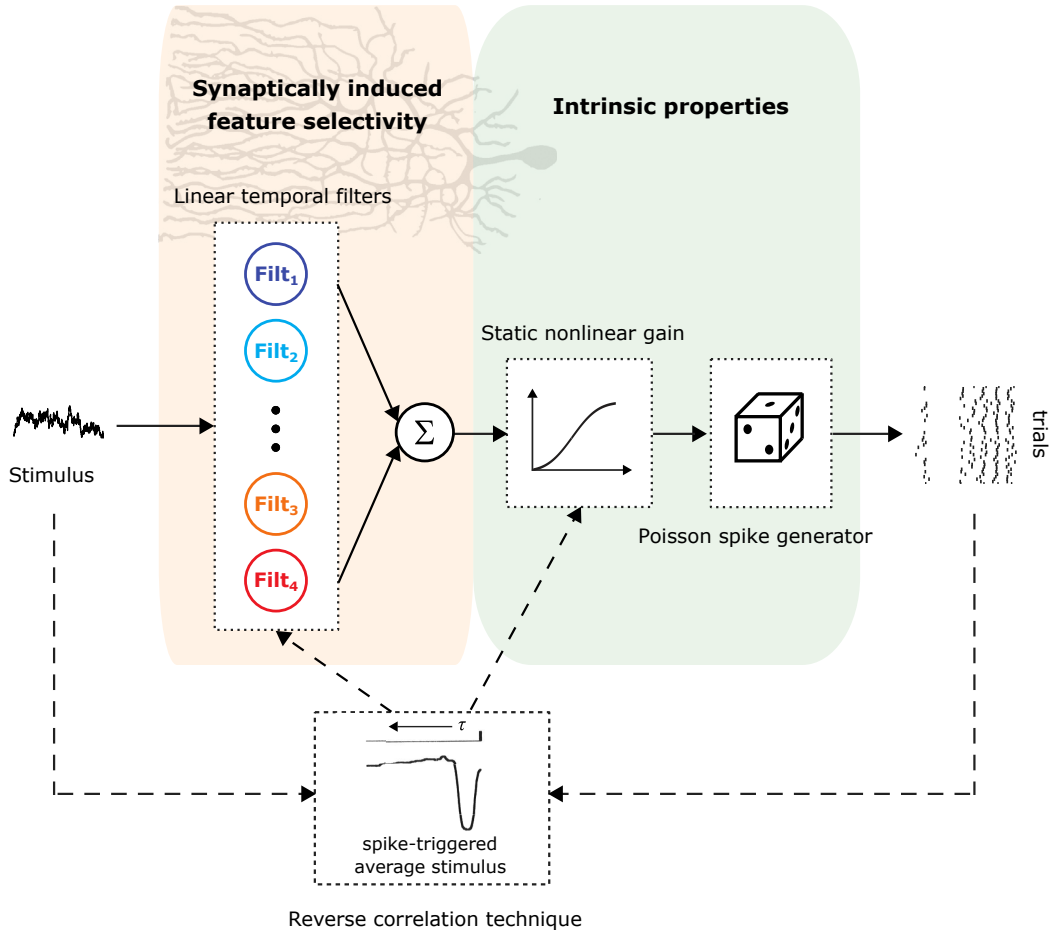

Figure 2.1: **Schema of the LNP model.** The relation between stimulus signals and neuronal spiking responses can be captured by the LNP model. The incoming stimuli are synaptically filtered by a set of pre-determined temporal features. Then aggregated filtered inputs are transformed into a Poisson spike train whose firing rate is determined by the output of the nonlinear gain function of the given neuron. The linear filters and nonlinear gain function can be computed using the reverse correlation of spikes with stimuli preceding each spike.

Barlow's efficient coding hypothesis has shed substantial light on the question about the purpose of sensory adaptation. It postulates that the sensory system always adopts the optimally efficient coding strategy to express the peripheral stimuli given a limited capacity

for information transmission [Barlow et al., 1961]. This efficient encoding strategy is expected to be achieved by single neurons to dynamically adapt its outputs to the inputs with varying statistics over relevant timescales. Theoretical analysis based on information theory indicates that immediate information transmission efficacy is maximized when a sensory neuron matches the distribution of its outputs to the local distribution of time-varying inputs [Wark et al., 2007], which has also been experimentally tested at several stages of sensory pathways across different modalities [Fairhall et al., 2001; Maravall et al., 2007, 2013]. The fast adaptation in the neuronal response can be manifested in intrinsic properties as nonlinear gain scaling [Brenner et al., 2000] and short-term synaptic plasticity as changes in features that linearly filter stimuli [Whitmire and Stanley, 2016]. Specifically, repeated stimuli can gradually increase or decrease the postsynaptic response through a facilitating or depressing synapse that acts as a high-pass or low-pass filter. The short-term synaptic adaptation and modulation of intrinsic neuronal properties can further lead to vital changes in information transmission on a network level across multiple timescales [Gjorgjieva et al., 2014].

2.2 Neural Basis of EEG

This section introduces some of the key concepts of EEG and its neurophysiological origin [Holmes and Khazipov, 2007; Tivadar and Murray, 2019], which are necessary to understand the relevance of EEG signatures of the sensory adaptation to the analogous phenomenon at the cellular level, as discussed in the next section.

The EEG is a non-invasive measure of cortical activity via multiple electrodes placed along the scalp's surface. It records the field potential fluctuations caused by summed extracellular current flows of populations of similarly oriented neurons that fire synchronously. However, the EEG electrodes could also collect electrical noises from other sources, such as muscular activity and electrical equipment. Despite its relatively low spatial resolution and signal-to-noise ratio, the EEG technique is still widely used to diagnose or treat seizures, sleep disorders and encephalopathies given its advantage of portability and millisecond-scaled temporal resolution. An event-related potential (ERP) is obtained by averaging EEG responses temporally locked to the occurrence of a specific stimulus or event. A subtype of ERP is the evoked potential (EP), which is the temporally averaged EEG responses evoked by the presentation of a specific sensory stimulus.

Neurons are electrically polarized by altering the relative concentration of intracellular ions to the extracellular milieu through the protein pumps on the membrane. Action potentials can be generated at the soma of a depolarized neuron and propagated along its axon. When an excitatory postsynaptic potential (EPSP) is induced by a presynaptic spike, an intracellular current source at the EPSP site and an adjacent current sink in the extracellular space occur together. According to Kirchhoff's current law, an intracellular source somewhere along the dendrite or at the soma of the postsynaptic neuron and its corresponding extracellular current sink must be generated concurrently. This extracellular current sink-source setup forms a current dipole that functions as a battery and gives rise to the EEG signals.

Since the pyramidal neurons in the cortex are well organized in a columnar manner, they have homogenous spatial orientations and become active simultaneously. These essential features make the pyramidal neurons the main contributors to the EEG response. Although EEG cannot directly detect activity from subcortical structures such as the thalamus and

brainstem, these deep sources have a significant influence on surface EEG activity. For instance, a large range of rhythmic EEG activity could result from the thalamocortical system. The mechanisms underlying some of these synchronous EEG patterns have been elucidated, such as the sleep spindle oscillation, while many others still are unclear.

2.3 Stimulus-specific Adaptation (SSA), Mismatch Negativity (MMN) and Novelty Detection

Sensory cortices are adept at identifying regularities and patterns obscured in the ever-changing input stream to dynamically generate predictions for the forthcoming stimulus. Their responses are then updated according to differences between the anticipated sensory inputs and the actual ones [Khouri and Nelken, 2015]. Mismatch negativity (MMN), observed in human EEG recordings [Fitzgerald and Todd, 2020], and stimulus-specific adaptation (SSA), evident in single-cell recordings in animals [Nelken, 2014], are two prominent instances of this type of short-term plasticity acting over timescales of seconds to minutes. For example, when consecutively presented with an identical stimulus over a short period, the cortical sensory neuron will decline/adapt its response and predict/expect another same one for the next stimulus. If the next stimulus violates the expectation, a larger/unadapted response will be evoked than by the expected stimulus. Therefore, both SSA and MMN conceptually represent a form of novelty detection: unexpected or infrequent stimuli are made more salient by eliciting stronger responses than expected or frequent ones. By virtue of such differential sensory processing, organisms are enabled to separate behaviorally important information from the continuous stream of sensory information.

The phenomenon of attenuated neuronal responses to common (standard) stimuli without generalizing to other, rare (deviant) stimuli has been a long-standing topic of interest, especially in the auditory system, which is referred to as SSA [Ulanovsky et al., 2003; Katz et al., 2006; Taaseh et al., 2011; Hershenhoren et al., 2014a; Musall et al., 2017]. It is typically tested using the oddball paradigm that contains a repetitive standard sequence occasionally interrupted by a deviant [Näätänen et al., 1978]. The roles of two single-dimension stimulus features as the deviant and standard are swapped in order to disambiguate the effect of response bias towards a specific stimulus feature from the occurrence probabilities on SSA. Consequently, SSA is quantified by the SSA index (SI) [Ulanovsky et al., 2003] defined below:

$$\mathrm{SI} = \frac{(\mathrm{d}(x_1) + \mathrm{d}(x_2)) - (\mathrm{s}(x_1) + \mathrm{s}(x_2))}{\mathrm{d}(x_1) + \mathrm{d}(x_2) + \mathrm{s}(x_1) + \mathrm{s}(x_2)} \tag{2.1}$$

where $\mathrm{s}(x_i)$ and $\mathrm{d}(x_i)$, $i \in \{1, 2\}$, describe neuronal responses evoked by stimuli feature x_i presented as standard and deviant, respectively.

The above index can be modified to measure SSA induced by each stimulus feature separately:

$$\mathrm{SI}(x_i) = \frac{\mathrm{d}(x_i) - \mathrm{s}(x_i)}{\mathrm{d}(x_i) + \mathrm{s}(x_i)} \tag{2.2}$$

According to the formula, SI is normalized between -1 and 1 and its positive value indicates the existence of SSA, where stronger SSA effects leads to a closer value of the index to 1 (refer to Fig 2.2).

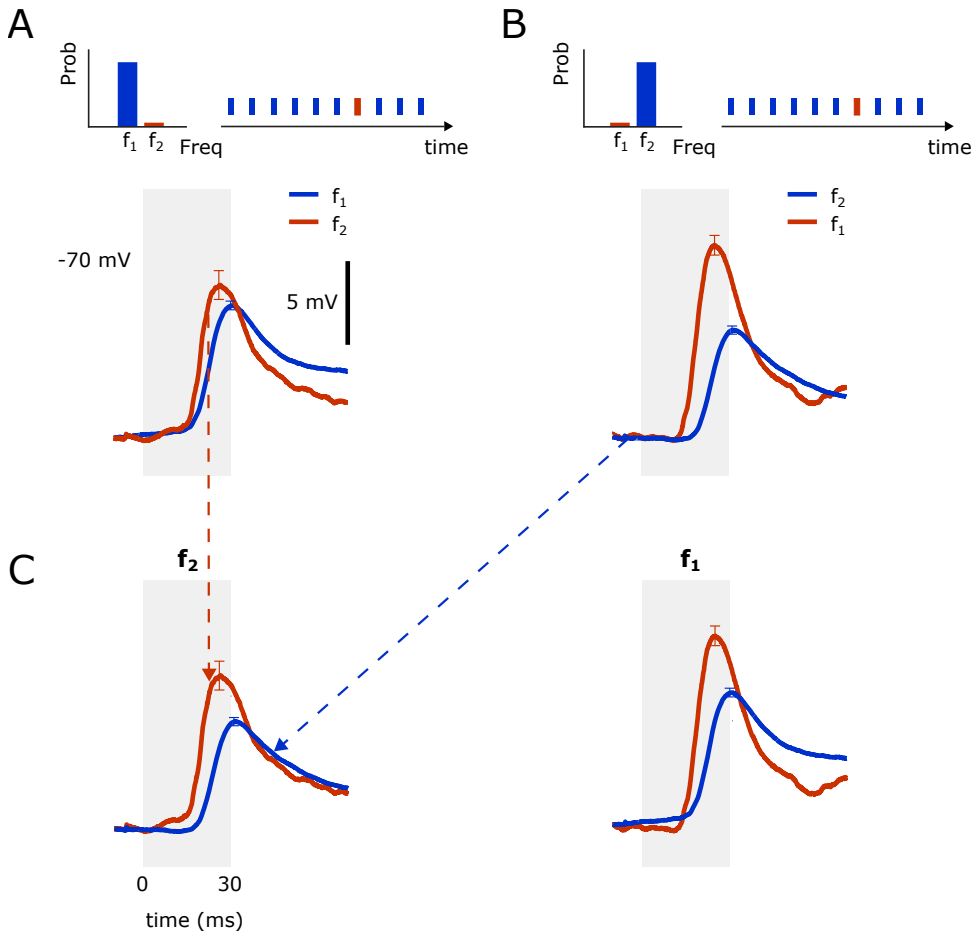

Figure 2.2: **Auditory SSA tested in the oddball paradigm. (A)** A frequency oddball sequence consists of the standard frequency f_1 of a tone (95% probability of occurrence) and deviant frequency f_2 of another tone (5% probability), which is reflected by tones of f_2 interspersed among a series of repeated tones of f_1 in the temporal presentation. The average membrane potential responses of an auditory neuron to the standard (blue) is smaller than the deviant (red). **(B)** The same convention as (A) except that identities of f_1 and f_2 as standard and deviant are swapped. Similarly, neuronal responses evoked by the standard are weaker than the deviant. Recording data are replotted from Hershenhoren et al. [2014a]. **(C)** Same responses in (A) and (B) are categorized by tone frequency. As defined in Eq 2.2, the strength of SSA elicited by each of the two frequencies is scaled by the difference between responses to the frequency as deviant and as standard. The difference is positive for both f_1 and f_2, which signifies that SSA exists at both frequency sites.

SSA is a robust and widespread phenomenon found in multiple sensory modalities. Ulanovsky et al. [2003] first observed this phenomenon in the auditory cortex, and following this pioneering work, SSA has also been identified at subcortical stages of the auditory pathway such as the thalamus [Bauerle et al., 2011; Antunes and Malmierca, 2014, 2011] and the inferior colliculus [Malmierca et al., 2009, 2015], as well as in other sensory systems such

as the somatosensory cortex [Katz et al., 2006; Musall et al., 2017] and the visual system [Reches and Gutfreund, 2008; Reches et al., 2010]. Furthermore, SSA was demonstrated in the auditory cortex of anaesthetized [Taaseh et al., 2011; Hershenhoren et al., 2014a], awake [von der Behrens et al., 2009] and freely moving rats [Polterovich et al., 2018], indicating that this phenomenon is not significantly affected by brain state and likely a hardwired component of sensory processing.

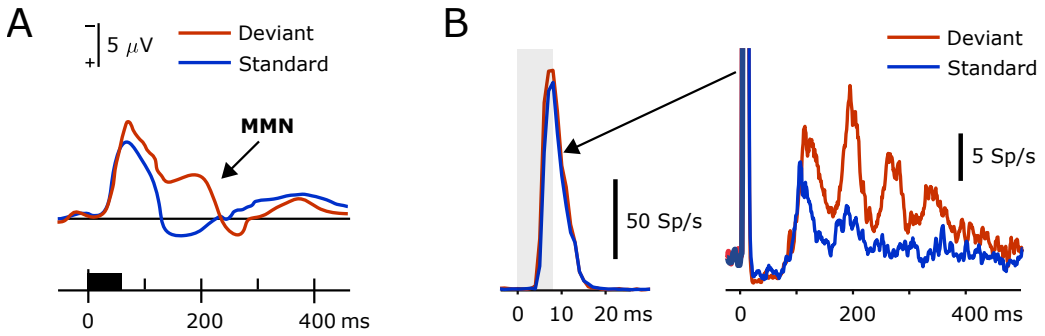

Figure 2.3: **Human auditory MMN and late SSA in whisker-related somatosensory area of the rat model. (A)** Average auditory event-related potentials in human EEG recordings in response to the deviant and standard in a frequency oddball paradigm. An additional negative component is evoked exclusively during the presentation of deviant stimuli. The duration of stimuli is indicated as a black bar along the time axis. Recording data are replotted from Sams et al. [1985]. **(B)** Average firing rate of neurons in the layer 4 of S1 in response to a whisker oddball paradigm. A secondary oscillatory deviant response with comparable latency with MMN (right panel) is observed following a transient synchronized activity (left panel). Gray shading marks the stimulation duration. Data are replotted from Musall et al. [2017].

MMN, like SSA, is usually studied by applying the oddball paradigm and manifested in human EEG recordings as an additional negative deviation elicited by a deviant stimulus that breaks the regularity established by repeated presentations of the standard [Näätänen et al., 1978] (see Fig 2.3A). MMN can appear in any sensory modality and has been well-studied in the auditory Näätänen et al. [2007] and visual areas [Pazo-Alvarez et al., 2003]. Both cortical SSA and MMN demonstrate a "true" deviance-detecting property, where the deviant response is dependent on its presentation context and beyond mere sensory input depression by applying several control paradigms [Jacobsen and Schröger, 2001; Taaseh et al., 2011; Musall et al., 2017]. In addition to SSA, animal models also show MMN-like responses that are homologous with human MMN in terms of their electrophysiological, pharmacological and functional properties [Harms et al., 2016; Shiramatsu and Takahashi, 2021]. Despite significant similarities, SSA is unlikely to be the direct neuronal substrate of MMN, but maybe one of the several mechanisms leading to the generation of MMN [Nelken and Ulanovsky, 2007]. While MMN has been pharmacologically demonstrated to be NMDA-receptor dependent [Umbricht et al., 2000], SSA seems to be not affected by NMDA-receptor antagonists [Farley et al., 2010]. Besides, the response latency of SSA (peaks <10 ms after stimuli onset) is substantially shorter than MMN (peaks ∼150-200 ms post-stimulus). Instead, SSA temporally matches better another event-related potential called mid-latency response (MLR) that peaks around

20 ms after stimuli onset [Shiramatsu and Takahashi, 2021]. MLR observed in humans and animals also demonstrates true deviance detection [Grimm et al., 2011] and its insensitivity to NMDA-receptor antagonists [Shiramatsu et al., 2013]. Thus, it seems more plausible to postulate that SSA is the potential neuronal correlate of mid-latency response. Moreover, Musall et al. [2017] recently found a late deviant-sensitive sensory response in the primary somatosensory cortex (S1) occurring hundreds of milliseconds post-stimulus, similar to MMN latency (see Fig 2.3B). The late sensory response has also been tested to be context-specific. These evidence strongly implied that the late response possibly lies in the physiological origin of MMN.

2.4 A review of SSA models

Over the past decade, a large number of computational models have been proposed to account for cortical SSA, with a focus on the auditory pathway. These models can be grouped into two main classes:

1) The first class of SSA models is based on a feedforward network with synaptic depression [Taaseh et al., 2011; Mill et al., 2011b,a, 2012; Hershenhoren et al., 2014b] which, unlike general neuronal fatigue, is stimulus-specific. Thus reduction in synaptic strength only occurs at those synapses mediating frequently presented stimulus features, leading to the specifically attenuated postsynaptic inputs to the neuron in question and correspondingly the generation of SSA. This class of feedforward adaptation models can be summarized in the adaptation in narrowly-tuned modules (ANTM) model, which functions as linear temporal filters within the context of LNP neuron introduced in 2.1. As illustrated in Fig 2.4, a downstream neuron integrates feedforward excitatory inputs from populations of presynaptic neurons tuned for specific acoustic frequencies. The feedforward synapses show activity-dependent adaptation. SSA in the postsynaptic neuron arises from the differential adapted drives from populations of narrowly-tuned presynaptic neurons. The ANTM model can reproduce several properties of SSA observed in the auditory cortex, such as the dependence of SSA on oddball protocol parameters, from the perspective of network-level adaptation. Indeed, as prominent synaptic depression is found in the thalamocortical pathway, cortical SSA emerges probably due to the differentially adapted thalamocortical inputs [Taaseh et al., 2011; Mill et al., 2011b].

In the ANTM model, the standard's regularity cannot affect the deviant response because the responses to each stimulus feature solely depend on the amount of adaptation load accumulated in its input channel. Hence the model fails to exhibit true deviance detection in which the presentation context formed by the standards also contributes to the deviant response. Nevertheless, this flaw of the ANTM model can be potentially addressed through some modifications, such as employing the model in a multi-layer configuration [Mill et al., 2012] or assuming the width of the tuning curve of the input channel is history-dependent [Nelken, 2014].

2) On top of the ANTM structure that explains SSA solely by presynaptic mechanisms of an LNP neuron, the contributions of network mechanisms to cortical SSA have also been considered in the form of recurrent networks [Wacongne et al., 2012; Wang and Knösche, 2013; May et al., 2015; Yarden and Nelken, 2017; Kudela et al., 2018]. The group of models has more complex network connectivity, and combined with synaptic plasticity, it usually produces rich

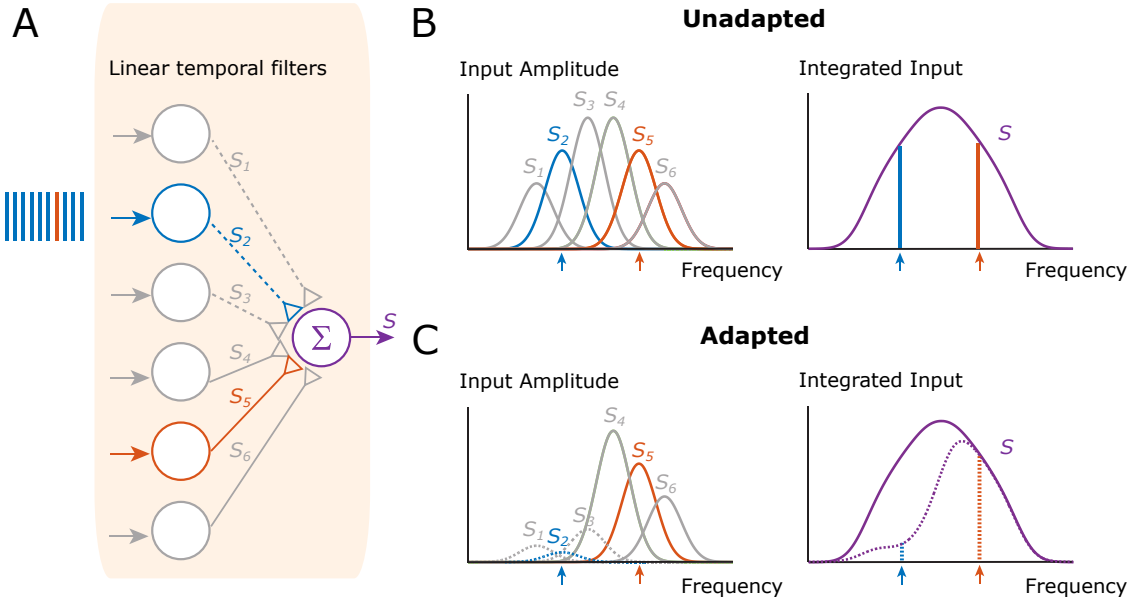

Figure 2.4: **Adaptation in the narrowly tuned modules (ANTM) model. (A)** A widely-tuned auditory neuron integrates inputs from narrowly-tuned presynaptic populations organized along a frequency axis. The tone sequence is composed of one off-peak frequency of the downstream neuron that frequently occurs (the standard in blue) and another off-peak frequency that occasionally occurs (the deviant in red). The repeated standard stimuli will gradually weaken its corresponding feedforward synapses (dashed line). Stimulus-specific inputs are dynamically filtered by the feedforward depressing synapses and aggregated to a downstream neuron, which can be envisaged as the feature selectivity stage of the LNP framework. **(B)** Before the adaptation happens, the integrated inputs produced by standard and deviant stimuli are equal (solid purple curve). **(C)** The amplitude of adapted input tuned to standard stimuli attenuated significantly (dashed frequency tuning curves) while that tuned to deviant stimuli is almost unaffected, which distorts the original symmetric frequency tuning of the downstream neuron (dashed purple curve). Figure is adapted from Whitmire and Stanley [2016].

intracortical behaviors. Predictive coding based neuronal network of the auditory cortex was studied by Wacongne et al. [2012] to account for deviance detection in the context of MMN. Yarden and Nelken [2017] proposed a generic SSA mechanism mediated via the propagation of synchronous population activity across a local neural circuit. Recently a laminar network of morphologically plausible, multi-compartmental neuron units was designed to capture SSA in the form of local field potentials [Kudela et al., 2018].

However, there is still no computational model accounting for the biphasic sensory responses in S1 that can be either sensitive to the rarity and presentation context of stimuli or not, depending on the response latency [Musall et al., 2017]. In particular, the latency of the secondary cortical response cannot be solely accounted for by the dynamics of a recurrent network involving short-term synaptic depression that is commonly regarded as an indispensable constituent of SSA models. Thus the latency is possibly produced by some

subcortical source(s) other than the intra-cortical dynamics. Inspired by the idea that SSA arises from the interaction between differentially adapted populations of neurons tuned for specific feature stimuli [Yarden and Nelken, 2017], we proposed a multi-scale recurrent network with synaptic depression and excitatory/inhibitory imbalance to explain how novelty detection can arise in the whisker-related part of the somatosensory thalamocortical loop and reproduce physiologically observed neural response patterns in S1 (see Chapter 4 for details).

2.5 Modeling Short-term Synaptic Plasticity

One of the prominent signatures of neocortical and thalamocortical synapses is their rapid changes in connection strength induced by sustained stimulation, known as short-term synaptic plasticity [Thomson and Deuchars, 1994; Gil et al., 1999; Katz et al., 2006]. The dynamics of synaptic plasticity depend on the types of neurons connected. In general, connections between excitatory pyramidal neurons exhibit synaptic depression, characterized by the reduction of postsynaptic responses to a repetitive presynaptic spike train (see Figure 2.5A). In contrast, couplings involving inhibitory interneurons often display synaptic facilitation, which has the opposite effect described by the increase of postsynaptic responses to the same spike train [Markram et al., 1998; Gupta et al., 2000] (see Figure 2.5B). It has been widely assumed that short-term synaptic depression is a necessary ingredient of SSA models, either in feedforward (such as ANTM model) or recurrent architectures [Nelken, 2014; Khouri and Nelken, 2015], as reviewed in the previous section.

Here, we will introduce a phenomenological model of short-term synaptic plasticity proposed by Tsodyks and Markram (also called TM model) [Tsodyks and Markram, 1997], which can qualitatively reproduce the experimentally-recorded kinetic data of short-term plasticity. The TM model can also be used to derive mean-field dynamics of a neural network coupled by adapting synapses (see an example in next section), which is a useful computational tool to investigate the local network dynamics of the cortex.

2.5.1 Modelling Synaptic Depression

The TM model assumes that each presynaptic site reserves an absolute amount of synaptic resources, and the amplitude of postsynaptic responses depends on the fraction of presynaptic resources in an active state. The dynamics of conversion between different states of resources are described by the following equations:

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{z}{\tau_{\mathrm{rec}}} - ux^{-}(t^f)\sum_f \delta(t-t^f) \\
\frac{\mathrm{d}y}{\mathrm{d}t} &= -\frac{y}{\tau_{\mathrm{in}}} + ux^{-}(t^f)\sum_f \delta(t-t^f) \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= \frac{y}{\tau_{\mathrm{in}}} - \frac{z}{\tau_{\mathrm{rec}}}
\end{aligned} \tag{2.3}$$

where the fraction of synaptic resources in the recovered, active, and inactive states is denoted by x, y and z, respectively. The Dirac delta function $\delta(t)$ represents a presynaptic spike. A presynaptic spike elicited at time t^f instantaneously transforms a rate u of resources in a recovered state into an active state that will trigger the opening of postsynaptic ion channels.

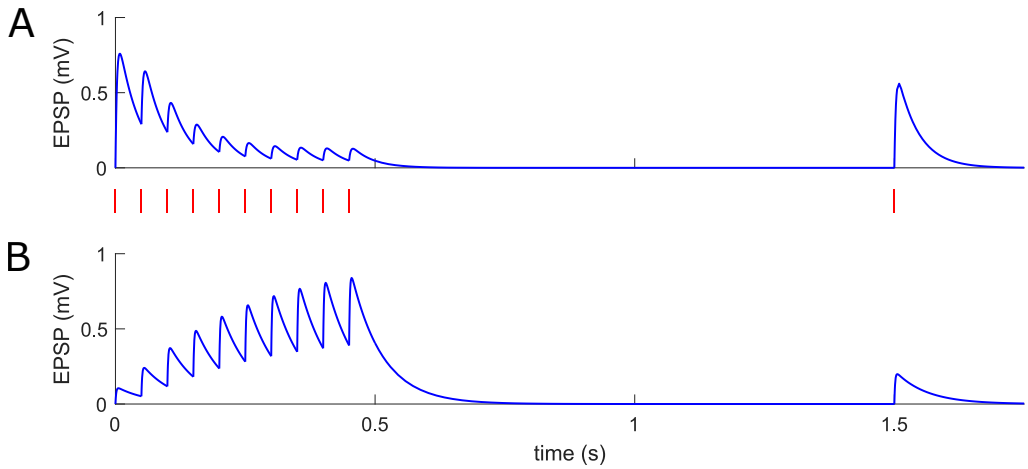

Figure 2.5: **Short-term synaptic plasticity. (A)** At a depressing synapse, attenuated excitatory postsynaptic potentials (EPSP) are evoked by a presynaptic spike train at 20 Hz. The amplitude of the postsynaptic response quickly reaches a plateau value following the first several presynaptic spikes. The synapse largely recovers after a 1-second silence of simulation. The spike train is marked under the panel. **(B)** Driven by the same presynaptic spike train as (A), the facilitating synapse demonstrates an opposite effect. Both depression and facilitation are stimulated using the TM model described in Eq 2.11. The postsynaptic potential is modeled by passive membranes ($\tau_{\rm m}{\rm d}u/{\rm d}t = -u + RI_{\rm syn}(t)$) with input resistance R and membrane time constant $\tau_{\rm m}$. Parameters: (A) $J = 250$ pA, $U = 0.5$, $\tau_{\rm rec} = 800$ ms, $\tau_{\rm in} = 3$ ms, $R = 100$ MΩ, $\tau_{\rm m} = 40$ ms; (B) $J = 1540$ pA, $U = 0.03$, $\tau_{\rm rec} = 130$ ms, $\tau_{\rm facil} = 530$ ms, $\tau_{\rm in} = 1.5$ ms, $R = 100$ MΩ, $\tau_{\rm m} = 60$ ms.

The active resources y immediately become inactive with a time constant $\tau_{\rm in}$ and subsequently recover with a time constant $\tau_{\rm rec}$ (see Fig 2.6). $x^-(t^f)$ denotes the fraction of recovered synaptic resources right before the presynaptic firing time t^f.

Here the postsynaptic current is modeled as directly current-sourced rather than conductance-based

$$I_{\rm syn}(t) = Jy(t) \tag{2.4}$$

where J is the synaptic efficacy expressed as the maximum capacity of synaptic transmission when all the resources are active.

By taking into account the fact $\tau_{\rm in} \ll \tau_{\rm rec}$ and the firing rate of the presynaptic neuron is usually much lower than $1/\tau_{\rm in}$, as well as the normalization condition $x + y + z = 1$, Eqs 2.3 can be approximated by the following equation involving x only:

$$\frac{{\rm d}x}{{\rm d}t} = \frac{1-x}{\tau_{\rm rec}} - ux^-(t^f)\sum_f \delta(t - t^f) \tag{2.5}$$

By integrating on both sides of Eq 2.5, one can obtain the iterative expressions for the recovered resources x in response to the presynaptic spike train

$$x_{n+1} = x_n(1-u)e^{-\Delta t_n/\tau_{\rm rec}} + 1 - e^{-\Delta t_n/\tau_{\rm rec}} \tag{2.6}$$

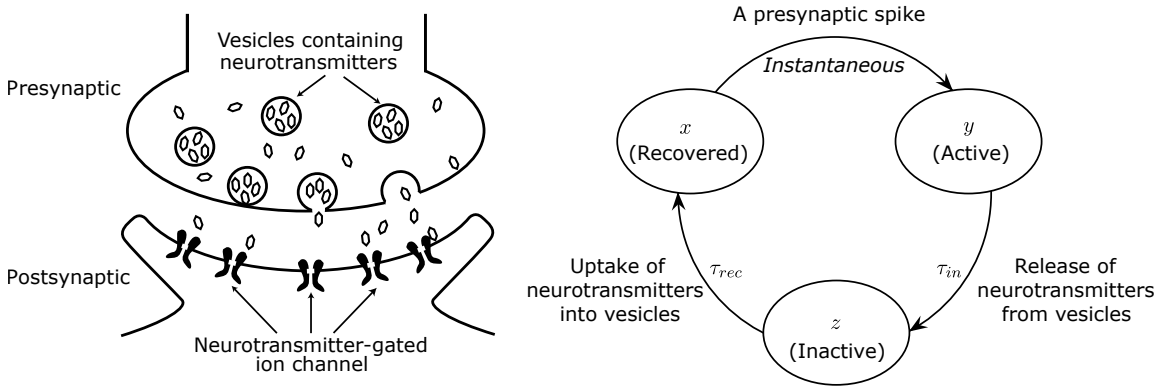

Figure 2.6: **Schematic illustration of activity-dependent synaptic transmission.** Upon the arrival of a presynaptic spike, synaptic vesicles quickly release their filled neurotransmitters into the synaptic cleft with the time constant τ_{in} . These neurotransmitters are then captured by receptors on the postsynaptic membrane and correspondingly lead to the opening of transmitter-gated ion channels. The depleted vesicles gradually replenish themselves with neurotransmitters with the time constant τ_{rec} and become available again for utilization by the next presynaptic spike. The left plot is modified from Gerstner et al. [2014]

where x_n represents the value of x immediately before the nth presynaptic spike. Δt_n is the time interval between the n-th and $(n+1)$-th presynaptic sikes. Correspondingly, the iterative form for the postsynaptic current is as follow

$$I_{\text{syn},n+1} = (I_{\text{syn},n} + Jux_n)e^{-\Delta t_n/\tau_{\text{in}}} \tag{2.7}$$

Particularly, if the presynaptic neuron fires at a constant rate f ($\Delta t_n \equiv 1/f$), the recovered resource will reach a steady-state value

$$x_{ss} = \frac{1 - e^{-1/(f\tau_{\text{rec}})}}{1 - (1-u)e^{-1/(f\tau_{\text{rec}})}} \tag{2.8}$$

as well as the postsynaptic current

$$I_{\text{syn},ss} = \frac{Jux_{ss}e^{-1/(f\tau_{\text{in}})}}{1 - e^{-1/(f\tau_{\text{in}})}} \tag{2.9}$$

2.5.2 Modeling Synaptic Facilitation

Since the amplitude of instant jump of postsynaptic current elicited by a presynaptic spike is proportional to the product of u and x, to model the facilitating effect, one can assume that u is not a constant rather being elevated by each spike and decaying between spikes. This assumption reflects an immediate increase in calcium concentration in the presynaptic terminal caused by an incoming spike, which leads to the release of neurotransmitters, and slow dissipation of calcium ions afterward [Bertram et al., 1996; Mallart and Martin, 1967; Zengel and Magleby, 1982]. The corresponding kinetics equation is therefore expressed as

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{U-u}{\tau_{\text{facil}}} + U[1-u^-(t^f)]\sum_f \delta(t-t^f) \tag{2.10}$$

where u is the fraction of opened calcium channels, U denotes the steady-state value of u, and τ_{facil} is the time constant for the close of calcium channels following their instant opening by incoming spikes.

Similarly, the above differential equation can be converted into an iterative form for the value of u right before the arrival of n-th presynaptic spike, which, along with the iterative equation of x defined in Eq 2.11, give rise to the corresponding expression for the postsynaptic current

$$\begin{aligned} u_{n+1} &= u_n(1-U)e^{-\Delta t_n/\tau_{\text{facil}}} + U \\ x_{n+1} &= x_n(1-u_n)e^{-\Delta t_n/\tau_{\text{rec}}} + 1 - e^{-\Delta t_n/\tau_{\text{rec}}} \\ I_{\text{syn},n+1} &= (I_{\text{syn},n} + Ju_nx_n)e^{-\Delta t_n/\tau_{\text{in}}} \end{aligned} \tag{2.11}$$

Driven by a regular presynatic spike train at the frequency of f ($\Delta t_n \equiv 1/f$), u can reach a plateau value

$$u_{ss} = \frac{U}{1-(1-U)e^{-1/(f\tau_{\text{facil}})}} \tag{2.12}$$

When $\tau_{\text{facil}} \to 0$, u_{ss} becomes equivalent to U and therefore the full model is reduced to a model of synaptic depression with $u = U$

It should be noted that mechanisms of synaptic facilitation and depression can exist concurrently. Different combination of parameters τ_{facil}, τ_{rec}, and U in Eq 2.11 could give rise to a wide range of short-term synaptic dynamics from strong facilitation ($\tau_{\text{facil}} \gg \tau_{\text{rec}}$ and relatively small values of U) to dominating depression ($\tau_{\text{facil}} \ll \tau_{\text{rec}}$ and large values of U) (see Fig 2.7).

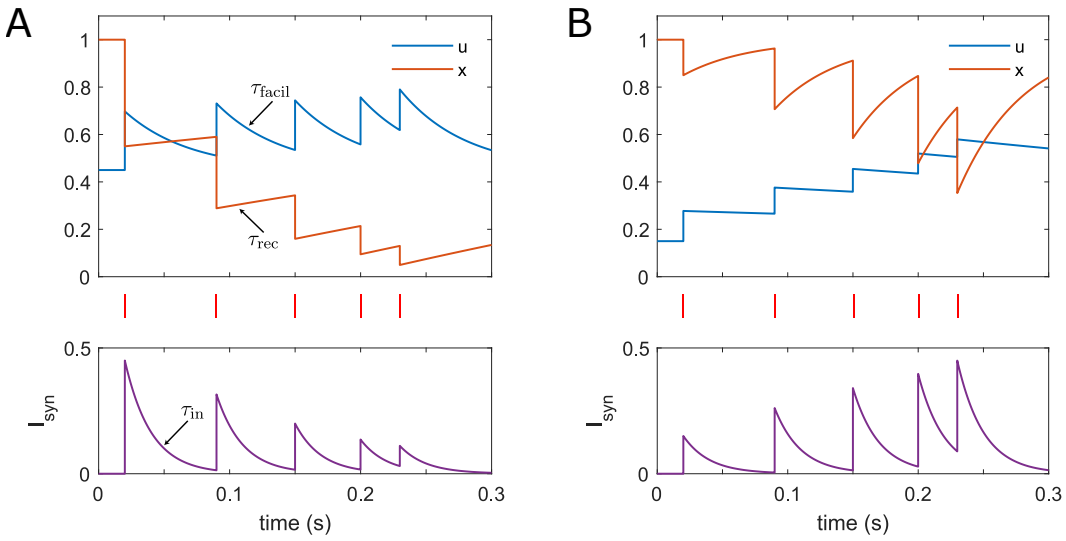

Figure 2.7: **Synaptic depression vs. facilitation. (A)** The evolution of the utilization rate u, recovered resources x as well as postsynaptic currents I_{syn} of a depression-dominated synapse in response to a presynaptic spike sequence (depicted under the top panel). Parameters: $J = 1$, $U = 0.45$, $\tau_{\text{rec}} = 750$ ms, $\tau_{\text{facil}} = 50$ ms, $\tau_{\text{in}} = 20$ ms. **(B)** A facilitation-dominated synapse driven by the same spike sequence as (A). Same parameters used as (A) except for $U = 0.15$, $\tau_{\text{rec}} = 50$ ms, $\tau_{\text{facil}} = 750$ ms.

2.6 From Spiking to Rate Network

In this section, we derive the rate equations for a network of interacting populations mediated by depressing synapses under the assumption that each population is large and homogeneous, using standard techniques [Gerstner et al., 2014; Vasilaki and Giugliano, 2012, 2014; Esposito et al., 2015]. The work is an extension of a previous work tackling populations connected with fixed connection strength [Gerstner et al., 2014] to the scenario of adapting synapses. By homogenous, we mean that (i) all neurons have the same properties; (ii) each neuron receives identical external input; and (iii) the coupling strength between any pair of neurons within the population is statistically uniform. The general conclusions drawn here will be applied in the subsequent two chapters to formalize mean-field cortical networks composed of multiple coupled cortical columns.

We define the mean firing rate as the population activity $A(t)$ averaged over a pool of homogenous neurons, as did in Gerstner et al. [2014]:

$$A(t) \equiv \lim_{\Delta t \to 0} \frac{n_{\mathrm{act}}(t; t+\Delta t)}{N \Delta t} = \frac{1}{N} \sum_{j=1}^{N} \sum_{f} \delta(t - t_j^f) \tag{2.13}$$

where N is the number of neurons within the population, $n_{\mathrm{act}}(t; t+\Delta t)$ denotes the number of spikes of all neurons in the population, which appear betweem time t and $t+\Delta t$. Since Δt is a very short time interval, each neuron fires at most one spike during the interval and correspondingly $n_{\mathrm{act}}(t; t+\Delta t)$ also represents the number of neurons that fire between time t and $t+\Delta t$. $\delta(t)$ is the Dirac delta function and t_j^f denotes the firing time of the j-th neuron in the population. This notion of the rate as population activity averaged over many neurons will be widely used in the following derivation.

The dynamics of short-term depression for the connection between pre- and post-synaptic neuron pair j, i is given by the phenomenological model [Tsodyks et al., 1998; Hennig, 2013]:

$$\frac{\mathrm{d}x_{ij}(t)}{\mathrm{d}t} = \frac{1 - x_{ij}(t)}{\tau_{\mathrm{rec}}} - U x_{ij}(t) \sum_{f} \delta(t - t_j^f) \tag{2.14}$$

where each pre-synaptic spike emitted at t_j^f utilizes a certain rate U of the available synaptic resource x_{ij}that is a fraction of full resources. The depleted available resource then returns to its baseline value with a time constant τ_{rec}. For the current-based type of post-synaptic input, which is independent of neuronal membrane potential, the increment in the amplitude of post-synaptic current $I_i(t)$ triggered by a pre-synaptic spike arriving at t_j^f is given by

$$\Delta I_i(t_j^f) = J U x_{ij}^{-}(t_j^f) \tag{2.15}$$

where J is the absolute synaptic efficacy and $x_{ij}^{-}(t_j^f)$ denotes the value of x_{ij} immediately before the arrival of the spike.

We define $x_i(t) \equiv 1/N \sum_{j=1}^{N} x_{ij}$ as the mean available resource in a bunch of synapses from a population of pre-synaptic neurons j to a post-synaptic neuron i. By taking the average over the pool of pre-synaptic neurons on both sides of Eq 2.14, we get

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \frac{1 - x_i(t)}{\tau_{\mathrm{rec}}} - \frac{1}{N} \sum_{j=1}^{N} U x_{ij}(t) \sum_{f} \delta(t - t_j^f) \tag{2.16}$$

The second term on the right-hand side of Eq 2.16 can be further transformed, assuming uncorrelated Poisson spiking from a large population of pre-synaptic neurons, to

$$\begin{aligned}\frac{U}{N}\sum_{j=1}^{N}\sum_{f}x_{ij}(t)\delta(t-t_j^f) &= U\lim_{\Delta t\to 0}\frac{\sum_{j\in\Omega}x_{ij}(t)}{N\Delta t}\\ &\approx U\lim_{\Delta t\to 0}\frac{x_i(t)n_{\mathrm{act}}(t;t+\Delta t)}{N\Delta t} \quad = Ux_i(t)A(t)\end{aligned} \tag{2.17}$$

where N is the number of neurons in the pre-synaptic population. Ω denotes the subpopulation of pre-synaptic neurons that fire between time t and $t+\Delta t$, and $n_{\mathrm{act}}(t;t+\Delta t)$ is the size of the subpopulation. The approximation becomes an equality in the limit of $N\to\infty$. $A(t)$ represents the population activity of pre-synaptic neurons as defined in Eq 2.13.

Using Eq 2.17, Eq 2.16 can be reformulated as

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \frac{1-x_i(t)}{\tau_{\mathrm{rec}}} - Ux_i(t)A(t) \tag{2.18}$$

We find that the dynamics of the mean synaptic resource $x_i(t)$ is independent of the neuronal identity and therefore is the same for all neurons in the population.

In addition, we assume that each pre-synaptic spike emitted at t_j^f triggers an instantaneous post-synaptic current with the temporal profile $\alpha(t-t_j^f)$ and the efficacy for each synaptic connection in an all-to-all coupled population of N neurons is scaled as J_0/N. The input current to a neuron i in the population is generated by all spikes of all neurons within the same population,

$$\begin{aligned}I_i(t) &= \sum_{j=1}^{N}\sum_{f}\frac{J_0}{N}Ux_{ij}^{-}(t_j^f)\alpha(t-t_j^f)\\ &= J_0U\sum_{j=1}^{N}\sum_{f}\int_0^\infty\frac{1}{N}x_{ij}(t-s)\alpha(s)\delta(t-t_j^f-s)\mathrm{d}s\\ &= J_0U\int_0^\infty\alpha(s)\frac{1}{N}\sum_{j=1}^{N}\sum_{f}x_{ij}(t-s)\delta(t-t_j^f-s)\mathrm{d}s\end{aligned} \tag{2.19}$$

Making use of Eq 2.17 with t replaced by $t-s$ to substitute the average quantity on the right-hand side of Eq 2.19, we obtain

$$I(t) = J_0U\int_0^\infty\alpha(s)x(t-s)A(t-s)\mathrm{d}s \tag{2.20}$$

We dropped the neuronal index i since the input current at time t depends on past mean synaptic resource and population activity and is identical for all neurons.

Now we generalize the arguments from a single fully-connected pool to multiple interacting pools. Under the assumption that neurons are homogenous in each population, the activity of neurons in population k is

$$A_k(t) = \frac{1}{N_k}\sum_{j\in\Gamma_k}\sum_{f}\delta(t-t_j^f) \tag{2.21}$$

where N_k is the size of population k and Γ_k represents the set of neurons that belongs to population k.

We assume that each neuron i in population k receives input from all neurons j in population l with adapting coupling strength $(J_{kl}/N_l)U_{kl}x_{ij}(t)$ and the post-synaptic current of the neuron i elicted by a spike of a presynaptic neuron j has the time course $\alpha_{kl}(t)$. Here the synaptic efficacy J_{kl}, utilization rate U_{kl} and post-synaptic current $\alpha_{kl}(t)$ depend on the type of synaptic connection from a neuron in population l to a neuron in population k rather than the neuronal identity. The input current to a neuron i in pool k is induced by all spikes of all neurons in the network of pools,

$$\begin{aligned} I_{i,k}(t) &= \sum_l \sum_{j\in\Gamma_l} \sum_f \frac{J_{kl}}{N_l} U_{kl} x_{ij}^-(t_j^f)\alpha_{kl}(t-t_j^f) \\ &= \sum_l J_{kl}U_{kl} \int_0^\infty \alpha_{kl}(s)\frac{1}{N_l}\sum_{j\in\Gamma}\sum_f x_{ij}(t-s)\delta(t-t_j^f-s)\mathrm{d}s \\ &= \sum_l J_{kl}U_{kl} \int_0^\infty \alpha_{kl}(s)x_i(t-s)A_l(t-s)\mathrm{d}s \end{aligned} \tag{2.22}$$

Correspondingly the dynamics of mean synaptic resource $x_i(t)$ is generalized as

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \frac{1-x_i(t)}{\tau_{\mathrm{rec},kl}} - U_{kl}x_i(t)A_l(t) \tag{2.23}$$

By assuming that U_{kl} and $\tau_{\mathrm{rec},kl}$ are homogenous across different pools, we drop the subscript kl and reformulate the above equation as

$$\frac{\mathrm{d}x_l(t)}{\mathrm{d}t} = \frac{1-x_l(t)}{\tau_{\mathrm{rec}}} - Ux_l(t)A_l(t) \tag{2.24}$$

Here we change the index of mean synaptic resource from post-synaptic neuron i in population k to pre-synaptic population l since the evolution of mean synaptic resource is governed by the activity of pre-synaptic population and is the same for all neurons within post-synaptic populations.

In this case, the input current,

$$I_k(t) = \sum_l J_{kl}U \int_0^\infty \alpha_{kl}(s)x_l(t-s)A_l(t-s)\mathrm{d}s \tag{2.25}$$

which is independent of the neuronal index i but the post-synaptic population index k.

Finally, we characterize the dynamics of the input current $h_k(t)$ of population k with the differential equation of passive membrane and employ for each population the rate model $A_l(t) = F_l(h_l(t))$ in which F_l is the stationary gain function of neurons in population l. The input current $h_k(t)$ takes into account both synaptic coupling $I_k(t)$ and external drive $I_k^{\mathrm{ext}}(t)$,

$$\tau_{\mathrm{m}}\frac{\mathrm{d}h_k(t)}{\mathrm{d}t} = -h_k(t) + \sum_l J_{kl}U\int_0^\infty \alpha_{kl}(s)x_l(t-s)F_l(h_l(t-s))\mathrm{d}s + I_k^{\mathrm{ext}}(t) \tag{2.26}$$

where all neurons in the network have the same membrane time constant τ_{m}.

Particularly, in the case of instantaneous post-synaptic current pulse $\alpha_{kl}(s) = \delta(s)$, Eq 2.26 can be reduced to a first-order differential equation

$$\tau_{\mathrm{m}} \frac{\mathrm{d}h_k(t)}{\mathrm{d}t} = -h_k(t) + \sum_l J_{kl} U x_l(t) F_l(h_l(t)) + I_k^{\mathrm{ext}}(t) \tag{2.27}$$

2.7 Phase Plane Analysis of a Recurrent Population with Depressing Synapses

Following the derivation of a mean-field rate model in the previous section, here we present a systematic analysis of the dynamics of a recurrent, excitatory population with short-term synaptic depression. This work re-expresses a prior stability analysis of the type of network under different parameter settings [Vasilaki and Giugliano, 2014] in the context of phase portraits, and extends the original work by further considering hysteretic and synchronous behaviors of the network, respectively led by Hopf bifurcation by a specific internal parameter and switch of external input. This network structure and its corresponding dynamics will later be employed as an elementary component of the somatosensory SSA model discussed in Chapter 4.

We geometrically analyze the model dynamics on the phase plane under different parameter regimes. We identify the model can show either bistability or monostability. The hysteresis behavior in the former case will result in sustained neuronal activity even after eliminating the external stimulus. This phenomenon is thought of as the neuronal signature of working memory that enables humans and animals to keep track of a temporary sensory stimulus for a few seconds after its disappearance. Furthermore, we examine the transient model dynamics in response to the switch of external input given a specific set of network parameters. On the onset of external stimuli, the model can exhibit a transient, sharp increase in population activity, reflecting almost synchronous firings of a group of neurons within a short time interval. The coincident population activity rapidly exhausts the synaptic resources within the population, which leads to the immediate termination of the concurrent activity and prevents its subsequent generation.

The rate dynamics of a homogenous population fully connected by depressing synapses is given by the following differential equations:

$$\begin{cases} \tau_{\mathrm{m}} \dot{h} = -h + JUxE + I_{\mathrm{ext}} \\ \dot{x} = (1-x)/\tau_{\mathrm{rec}} - UxE \end{cases} \tag{2.28}$$

where E is the excitatory population activity defined by the following gain function of the mean-field input current h:

$$E = [\alpha(h-\theta)]_+ = \max\left(\alpha(h-\theta), 0\right) \tag{2.29}$$

Since each neuron in the population has the same dynamics by assuming that the population is fully-connected, large and homogenous, the collective dynamics of the population are described as if a single neuron behaves. The summed mean-field input current h consists of the self-recurrent feedback mediated by depressing synapses $JUxE$ and the external input I_{ext}.

In order to investigate the dependency of the network dynamics on parameters, we first need to determine the nullclines of the dynamical system (2.28, 2.29). By setting $\dot{h} = 0$ in Eq 2.28, we get the following algebraic equation:

$$-h + JUx[\alpha(h-\theta)]_+ + I_{\text{ext}} = 0 \tag{2.30}$$

which has the solution

$$h(x) = \begin{cases} (\theta - I_{\text{ext}}) / (\alpha JUx - 1) + \theta & h \geq \theta \\ I_{\text{ext}} & h < \theta \end{cases} \quad (h\text{-nullcline}) \tag{2.31}$$

The h-nullcline consists of a constant function $h(x) = I_{\text{ext}}$ for $h < \theta$ and one branch of the hyperbola $h(x) = (\theta - I_{\text{ext}}) / [\alpha JU(x - 1/\alpha JU)] + \theta$ for $h \geq \theta$ with vertical asymptote $x = 1/\alpha JU$ and horizontal asymptote $h(x) = \theta$, as depicted in Fig 2.8.

Similarly, given $\dot{x} = 0$, we obtain the equation

$$(1-x)/\tau_{\text{rec}} - Ux[\alpha(h-\theta)]_+ = 0 \tag{2.32}$$

which gives the x-nullcline

$$\begin{cases} h(x) = 1/\alpha\tau_{\text{rec}}Ux - 1/\alpha\tau_{\text{rec}}U + \theta & h \geq \theta \\ x = 1 & h < \theta \end{cases} \quad (x\text{-nullcline}) \tag{2.33}$$

The x-nullcline is composed of a vertical line $x = 1$ for $h < \theta$ and the right branch of the hyperbola $h(x) = 1/\alpha\tau_{\text{rec}}Ux - 1/\alpha\tau_{\text{rec}}U + \theta$ for $h \geq \theta$ with vertical asymptote $x = 0$ and horizontal asymptote $h(x) = \theta - 1/\alpha\tau_{\text{rec}}U$. The vertical line and the hyperbola join at the point $(x, h) = (1, \theta)$ (see Fig 2.8).

When $I_{\text{ext}} \leq \theta$, there is always an intersection between h- and x-nullcline at $(x, h) = (1, I_{\text{ext}})$. However, there could be zero, one or two intersections between the hyperbola parts of the two nullclines in the domain of $h > \theta$, depending on the parameters (see Fig 2.8A, B and C). When $I_{\text{ext}} > \theta$, the scenario is greatly simplified with only one intersection at $h > \theta$ for arbitrary choices of the rest parameters (see Fig 2.8D).

We observe that there exists a critical value for network parameters where two intersections coalesce into a single one at $x = x_{\text{c}}$ and $h = h_{\text{c}} > \theta$ when $I_{\text{ext}} \leq \theta$. By definition, h-nullcline is tangent to x-nullcline at $x = x_{\text{c}}$, namely

$$\begin{cases} h^{h-\text{null}}(x_{\text{c}}) = h^{x-\text{null}}(x_{\text{c}}) \\ \dot{h}^{h-\text{null}}(x_{\text{c}}) = \dot{h}^{x-\text{null}}(x_{\text{c}}) \end{cases} \tag{2.34}$$

which can be further expressed in an explicit form as

$$\begin{cases} (\theta - I_{\text{ext}}) / (\alpha JUx_{\text{c}} - 1) = 1/\alpha\tau_{\text{rec}}Ux_{\text{c}} - 1/\alpha\tau_{\text{rec}}U \\ J(\theta - I_{\text{ext}}) / (\alpha JUx_{\text{c}} - 1)^2 = 1/\left[\tau_{\text{rec}}(\alpha Ux_{\text{c}})^2\right] \end{cases} \tag{2.35}$$

The above equations have the solution

$$x_{\text{c}} = 1 - \alpha U\sqrt{J\tau_{\text{rec}}(\theta - I_{\text{ext}})} \tag{2.36}$$

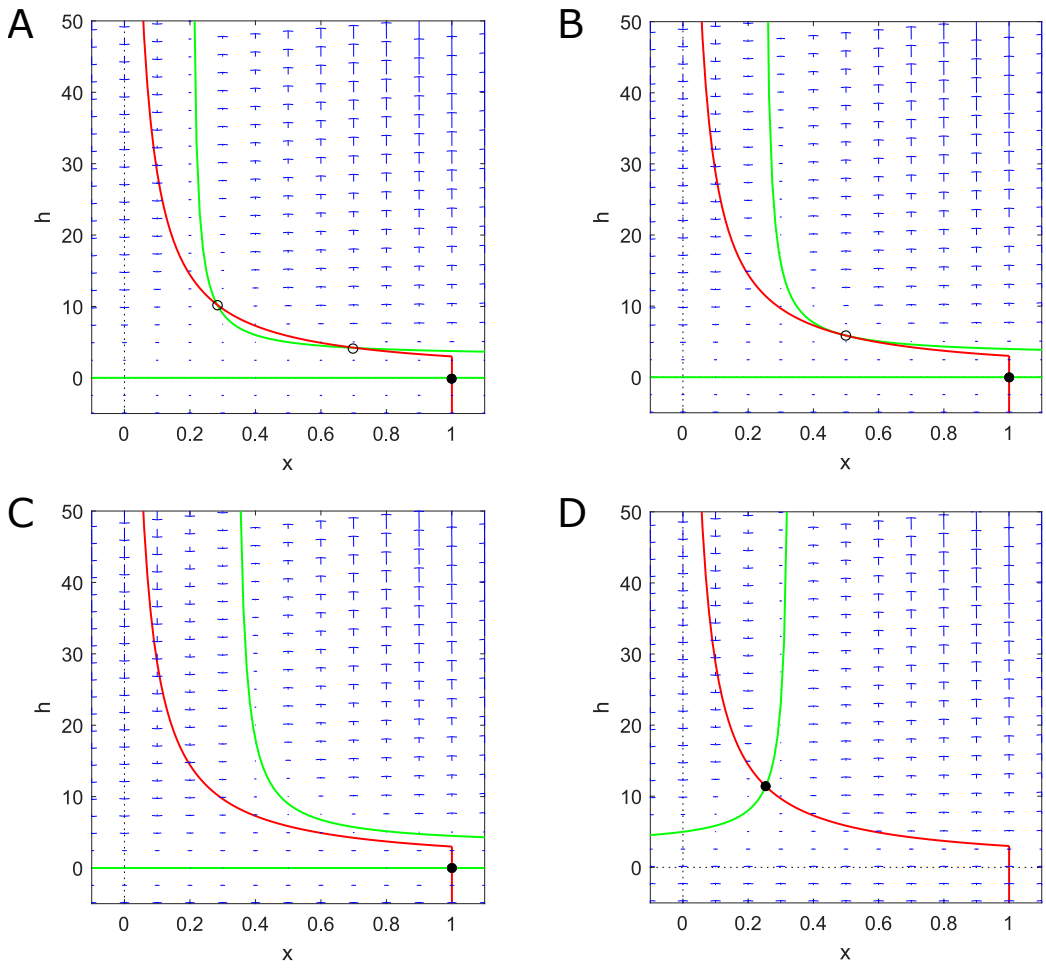

Figure 2.8: **Phase portraits of the dynamical system (2.28) for different J and I_{ext}.** The h- and x-nullclines can intersect three times (**A**: J=10, I_{ext}=0), twice (**B**: J=J_{c}=8.199, I_{ext}=0) or once (**C**: J=6, I_{ext}=0). They can also have only one intersection in another case where $I_{\text{ext}} \geq \theta$ (**D**: J=6, I_{ext}=5). Red and green curves respectively represent the x- and h-nullclines. The stable and unstable equilibria are denoted by filled and open circles, respectively. The rest parameters: $\tau_{\text{m}} = 0.001$ s, $U = 0.5$, $\tau_{\text{rec}} = 0.7$ s, $\theta = 3$, $\alpha = 1$.

Substituting x_{c} into the second equation of Eqs 2.35, one can obtain the critical value of the synaptic efficacy J related to the single intersection scenario (Fig 2.8B)

$$J_{\text{c}} = \left(\sqrt{\tau_{\text{rec}}\left(\theta - I_{\text{ext}}\right)} + 1/\sqrt{\alpha U}\right)^2 \tag{2.37}$$

We then consider the stability of the equilibrium points. For the convenience of analysis, we compactly rewrite the dynamical system (2.28) in the following form

$$\begin{cases} \dot{x} = G_1(x,h) = (1-x)/\tau_{\text{rec}} - UxE \\ \dot{h} = G_2(x,h) = \left(-h + JUxE + I_{\text{ext}}\right)/\tau_{\text{m}} \end{cases} \tag{2.38}$$

By definition, the Jacobian matrix L of the above system at the equilibrium $(x_{\text{eq}}, h_{\text{eq}})$ is

$$L\left(x_{\text{eq}}, h_{\text{eq}}\right) = \left(\begin{array}{cc} \partial G_1/\partial x & \partial G_1/\partial h \\ \partial G_2/\partial x & \partial G_2/\partial h \end{array} \right)\Bigg|_{(x_{\text{eq}}, h_{\text{eq}})} \tag{2.39}$$

To be specific,

$$L\left(x_{\text{eq}}, h_{\text{eq}}\right) = \left(\begin{array}{cc} -1/\tau_{\text{rec}} - \alpha U\left(h_{\text{eq}} - \theta\right) & -\alpha U x_{\text{eq}} \\ \alpha JU\left(h_{\text{eq}} - \theta\right)/\tau_{\text{m}} & \left(-1 + \alpha JU x_{\text{eq}}\right)/\tau_{\text{m}} \end{array} \right) \quad \text{For } h_{\text{eq}} > \theta \tag{2.40}$$

and

$$L\left(x_{\text{eq}}, h_{\text{eq}}\right) = \left(\begin{array}{cc} -1/\tau_{\text{rec}} & 0 \\ 0 & -1/\tau_{\text{m}} \end{array} \right) \quad \text{For } h_{\text{eq}} \leq \theta \tag{2.41}$$

When $h_{\text{eq}} \leq \theta$, since the eigenvalues $\lambda_{1,2}$ of $L\left(x_{\text{eq}}, h_{\text{eq}}\right)$ are always negative real numbers($\lambda_1 = -1/\tau_{\text{m}},\ \lambda_2 = -1/\tau_{\text{rec}}$), the equilibrium $(x_{\text{eq}}, h_{\text{eq}})$ is a stable node (the equilibrium at $h_{\text{eq}} = I_{\text{ext}}$ when $I_{\text{ext}} \leq \theta$, i.e. the rightmost intersection in Fig 2.8A, B and C).

We notice that in the case of Fig 2.8A, the leftmost intersection can transit from an unstable to stable focus when τ_{m} increases. That is to say, this equilibrium undergoes a Hopf bifurcation, which occurs when the eigenvalues of the equilibrium become pure imaginary numbers. We numerically compute the bifurcation value $\tau_{\text{m}}^* = 0.082$. The middle intersection is always an unstable saddle as τ_{m} increases in the same range. Hence, when $\tau_{\text{m}} > \tau_{\text{m}}^*$, the model can show bistability in absence of external input. After a long presence of external input $I_{\text{ext}} > \theta$, the system states will inhabit the single global attractor. However, the system states fail to return to their previous rightmost equilibrium when the external input I_{ext} is removed. The hysteresis behavior can cause persistent neuronal activity even after the removal of external stimuli, which is postulated to be the neuronal correlate of working memory [Barak and Tsodyks, 2007a].

Finally, we investigate the dynamic relation between synaptic resource x and summed input current h when the external input I_{ext} exists or not, respectively. When the external current I_{ext} is stepped from 0 to a non-zero value, say $I_{\text{ext}} = 5$, the phase portrait changes instantaneously, while the old state of the system does not have time to catch up. Hence we need to consider the network dynamics when $I_{\text{ext}} = 5$ with initial values (x_0, h_0) set to the state of the system right before the step current. We can see from Fig 2.10A that a set of trajectories with different initial points all make an anticlockwise excursion before finally approaching the attractor. In the temporal domain, the excursion represents a transient increase of the mean synaptic input h and correspondingly the population activity E. The ensemble phenomenon is termed a population spike featured by nearly coincident firings of a pool of neurons. The size of the excursion depends on the initial value of the mean synaptic resource x_0. The closer the initial synaptic resource x_0 is to its full state 1, the larger the population spike (excursion of h) is triggered. On the other hand, once the step current is removed, the states of the system are immediately drawn to the h-nullcline $x = 0$ and then travel along it to the global attractor with a time constant τ_{rec} (see Fig 2.10B). It is worth noting that for our SSA model in Chapter 4, we choose parameters in the regime showing single equilibrium in the absence of the external input, as illustrated in Fig 2.8C and 2.10B.

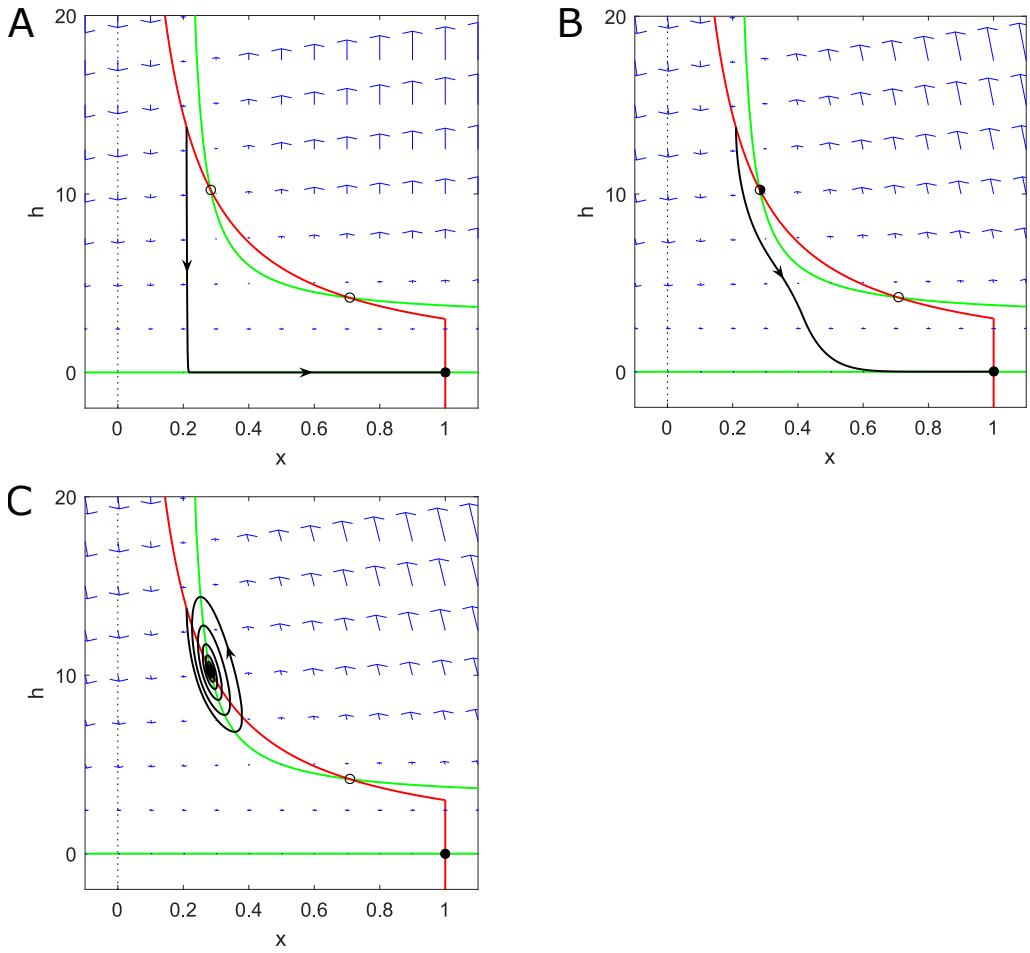

Figure 2.9: **Hopf bifurcation in the dynamical system (2.28).** The leftmost equilibrium gains stability as the bifurcation parameter τ_{m} increases (**A**: τ_{m}=0.001 s, **B**: τ_{m}=τ_{m}^{*}=0.082 s, **C**: τ_{m}=0.1 s) Red and green curves respectively represent the x- and h-nullclines. Stable, unstable equilibria as well as equilibria undergoing a bifurcation are denoted by filled, open and half-filled circles, respectively. The rest parameters: $J = 10$, $U = 0.5$, $\tau_{\mathrm{rec}} = 0.7$ s, $\theta = 3$, $\alpha = 1$, $I_{\mathrm{ext}} = 0$.

2.8 Summary

In this chapter, we introduced the organism's ability to detect novel sensory signals and accordingly presented the concepts of stimulus-specific adaptation (SSA) and mismatch negativity (MMN) as two relevant neural signatures verified in the oddball protocol. Following these experimental findings, we reviewed the current SSA models and concluded that none of them could account for a deviant-specific response with MMN-like latency observed in the somatosensory cortex [Musall et al., 2017]. Therefore, one of the goals of the thesis is to develop a mechanistic model to explain this particular response pattern that may be the electrophysiological origin of the mismatch response in EEG. The mean-field model derivation and its related phase plane analysis were made to form the theoretical foundation of our

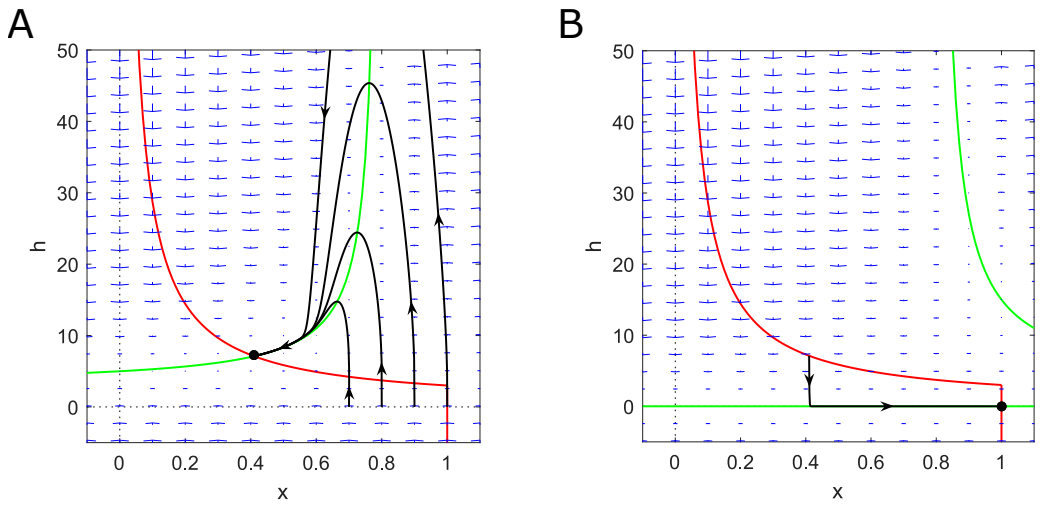

Figure 2.10: **Representative trajectories of the system (2.28) when I_{ext} is on and off, respectively. (A)** Phase portrait for the step input current $I_{\text{ext}} > \theta$ ($I_{\text{ext}} = 5$). Trajectories with ascending initial synaptic resource $x_0 = 0.7, 0.8, 0.9, 1$ but the same initial input current $h_0 = 0$ demonstrate increasing amplitudes of a detour before reaching the equilibrium. **(B)** Phase portrait for the absent external input ($I_{\text{ext}} = 0$). A trajectory starting from the prior steady states under external stimuli terminates at a global attractor. The remaining parameters are: $J = 2.5$, $\tau_{\text{m}} = 0.001$ s, $U = 0.5$, $\tau_{\text{rec}} = 0.7$ s, $\theta = 3$, $\alpha = 1$.

SSA model of the auditory cortex in chapter 3 and the somatosensory cortex in chapter 4. The following chapter will discuss our simplified implementation of a promising network architecture producing auditory SSA [Yarden and Nelken, 2017]. This SSA structure will be further adapted and extended to reproduce and elucidate the late SSA response within the context of a somatosensory thalamocortical loop (chapter 4).

Chapter 3

A Robust Model of Auditory Stimulus-specific Adaptation

This chapter is mainly established on my co-authored published material [Vanattou-Saïfoudine et al., 2021], which studies auditory SSA using both a theoretical mean-field model and its physical equivalent model implemented on a neuromorphic hardware circuit. In 2017, Yarden and Nelken [2017] proposed a novel mechanism for SSA involving both a recurrent network and local synaptic depression. Based on their SSA network architecture, our work provides further support for the hypothesis of SSA as a consequence of the propagation of synchronous firing events across the network of populations of tuned sensory neurons [Yarden and Nelken, 2017], using for the first time both a mean-field model and a neuromorphic implementation. As the neuromorphic hardware used in this study does not possess short-term synaptic depression, we employed a spike-frequency adaptation mechanism to simulate synaptic depression by introducing an adaptive input layer. Spike-frequency adaptation is a phenomenon in which a neuron shows a reduction in its firing rate in response to a constant stimulus [Benda and Herz, 2003]. Furthermore, it is worth noting that the adaptation feature used in our network is an interesting tool to investigate SSA as past research has demonstrated that the cortical adaptation recovery time scale to tones was consistent with the cortical synapse recovery scale in synaptic depression [Ulanovsky et al., 2004, 2003; May and Tiitinen, 2010]. We identity through this work that the propagation of population activity of neurons tuned for specific stimuli features, along with the activity-dependent fast adaptation mechanism (spike-frequency adaptation or synaptic depression) that modulates the extent of activity propagation, are two indispensable components to reproduce SSA and its related properties fully.

The following content shows my primary contributions to the joint work in developing the minimal mean-field model, which is summarized in three folds: (1) I performed transient dynamics and stability analysis of the mean-field populations to find the parameters qualitatively approximating the experimental recordings; (2) presented a theoretical analysis of the full mean-field network behaviors under different protocols and practical ways of tuning its equivalent spiking version of the network model implemented on neuromorphic hardware; (3) performed robustness test of the model against online parameter perturbations and sensitivity analysis by supplying the parameter ranges.

3.1 Results

The present work proposed a hierarchical network of the primary auditory cortex (A1), which is a simplified version of the network presented in Yarden and Nelken [2017], to explain the dynamics of auditory SSA and reproduce its related experimental findings. The network model is implemented in mean-field rate populations under certain assumptions, which allows us to treat part of the network analytically and derive the necessary conditions to demonstrate SSA and its relevant properties under this architecture. The theoretical results derived from the mean-field model also provided insights for tuning its neuromorphic spiking counterpart to reproduce previously obtained SSA experimental results [Taaseh et al., 2011]. The network presented holds special features to observe SSA discussed in [Yarden and Nelken, 2017], such as recurrent connectivity, lateral propagation of network activity and a short-term-depression-like mechanism.

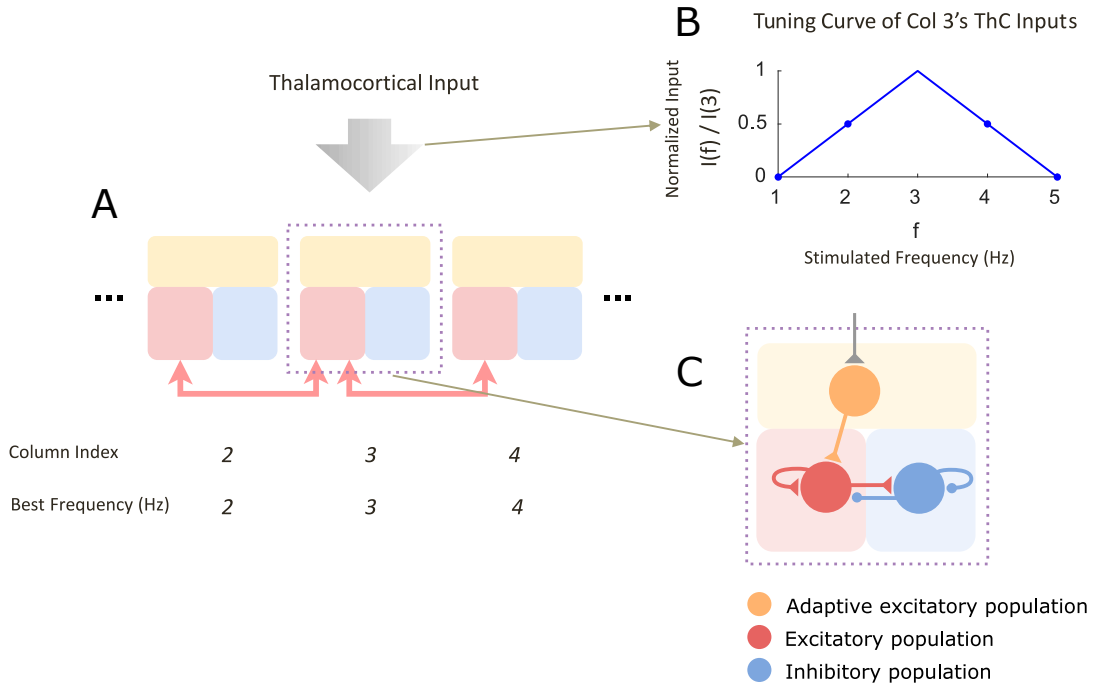

Figure 3.1: **Network architecture. (A)** The network is composed of five cortical columns that are arranged along the frequency axis by their best frequencies. Each column is constructed in three blocks: the adaptive layer (pale yellow) receives thalamocortical sensory inputs and sends its feedforwards outputs to the excitatory (pale red) population that is coupled with the inhibitory (pale blue) population. The excitatory group is also laterally connected to its nearest excitatory counterparts, illustrated by arrows. **(B)** The thalamocortical input to each column (for example, column 3) linearly declines as the stimulated frequency is away from the column's best frequency, which forms the tuning curve of the column's thalamocortical input. **(C)** To be detailed, within every column, the feedforward adaptive population excites a recurrent excitatory population that is regulated by another recurrent inhibitory population. The homogenous excitatory and inhibitory populations are fully connected.

3.1.1 Model Architecture

As illustrated in Fig 3.1A, the primary auditory cortex is modeled as coupled cortical columns, each of which consists of two cascaded layers: the adaptive layer where neurons demonstrate spike-frequency adaptation and the recurrent excitatory-inhibitory layer. Columns are unidimensionally organized by the most responsive frequency of each column, called the column's best frequency. The adaptive population in each cortical column is evoked by feedforward thalamocortical inputs dependent on the frequency difference between the best frequency of the column and the frequency of the stimulated tone. Typically, the farther the stimulated frequency is from the best frequency of a column, the weaker the thalamocortical input the column will receive (see Fig 3.1B). The adaptive population subsequently activates the recurrent excitatory group regulated by the recurrent inhibitory group (Fig 3.1C). Furthermore, each excitatory population feeds its output to its two lateral excitatory populations, leading to a lateral propagation of population activities across the network.

3.1.2 Deviance Sensitivity Assessed in Different SSA Protocols

SSA emerges in the network in response to a "traditional" oddball paradigm (OD) that consists of two stimuli, standard and deviant, which differ in one specific stimulus feature (such as frequency or amplitude of tone stimuli) activating different populations of tuned neurons. In this study, the standard stimulus is presented with a higher probability than the deviant stimulus. In order to control for potential asymmetries in processing the two stimuli (e.g., being more sensitive to one tone frequency than the other), a second oddball paradigm is presented where the occurrence probabilities of the two stimuli are switched and the stimulus that was earlier the standard is now the deviant. Such a flip-flop paradigm allows one to compare the responses to an identical stimulus, respectively presented as the standard and deviant.

To test SSA in the model, we stimulated the network by an oddball sequence composed of tones of 2 Hz and 4 Hz as the standard and deviant stimuli respectively, and recorded the activity in the excitatory population of the middle column (column 3) to verify the existence of SSA. The standard stimulus evokes more adapted activities in the standard column than the deviant does in its column due to the repeated stimulation (see Fig 3.2A). Therefore, the less adapted input from the deviant adaptive population evokes more effective propagation of activities across the excitatory network than the standard input. Correspondingly, in the recording column, stronger responses are triggered by the deviant than the standard, bringing about the generation of SSA (see Fig 3.2B).

It is worth mentioning that a phenomenon of fast and synchronized firings of a pool of neurons has been widely observed in the auditory cortex [Luczak et al., 2009, 2013; Sakata and Harris, 2009], under the name of population spike. The synchronous activity enables its lateral propagation across multiple cortical columns, which is associated with SSA generation [Yarden and Nelken, 2017]. From a modeling point of view, prolonged activation could also implement the lateral propagation of activity by increasing the excitatory population's firing rate and then spreading to its neighboring populations via inter-column connections, similarly as population spikes do. However, to match the population bursts recorded in the primary auditory cortex (A1) [Luczak et al., 2009, 2013; Sakata and Harris, 2009], we modeled responses of our network model to sensory stimulation as a transient synchronization in population activity (see Fig

3.2B, bottom panel).

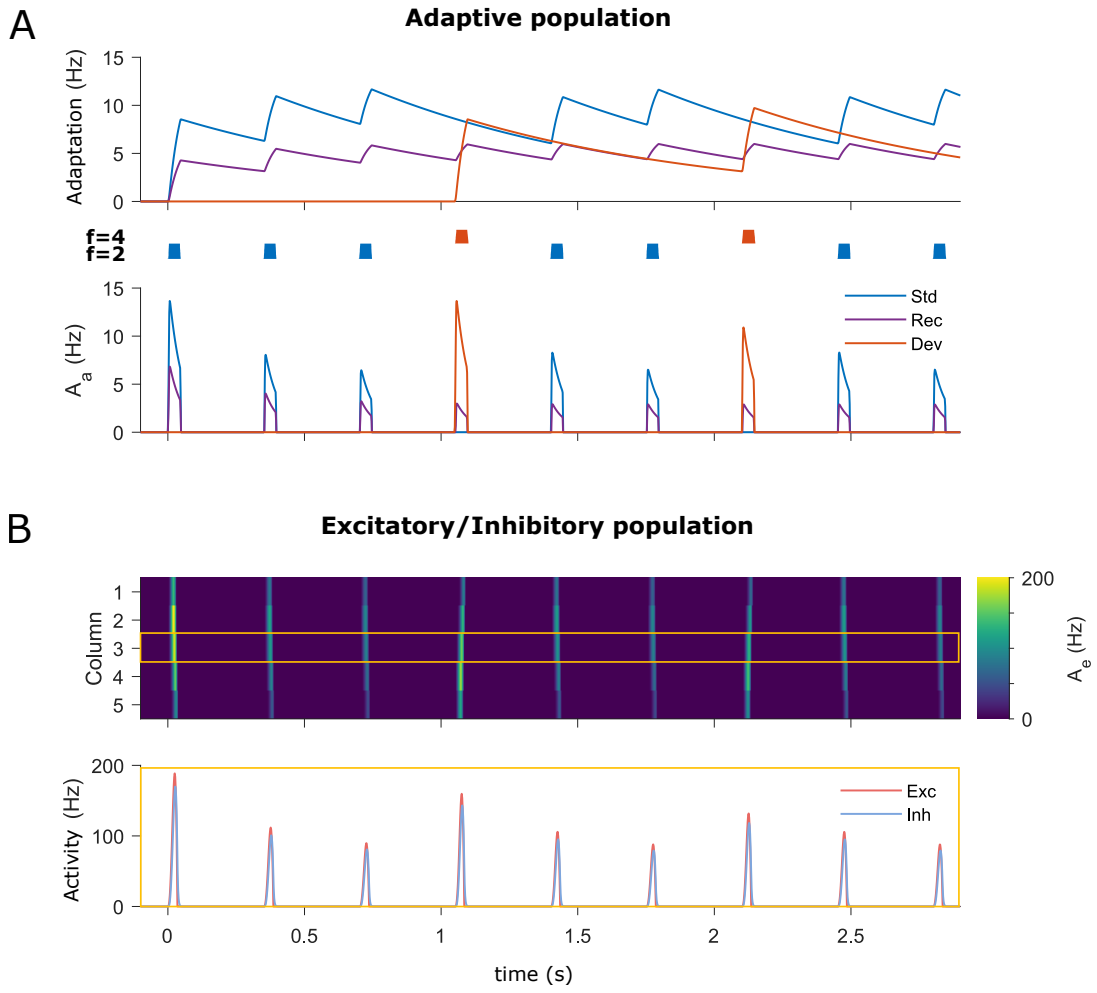

Figure 3.2: **SSA arising from the differential adaptation of populations of frequency-tuned neurons. (A)** The amounts of adaptation a accumulate immediately upon the presentation of stimuli and are gradually released during the inter-stimulus interval. Adaptation in the standard column (blue) is often more substantial than the deviant column (red). Yet, adaptation induced in the recording column (purple) is independent of the identity of stimuli after the first several ones due to the symmetry of the thalamocortical input around the stimulated frequency (top panel). Accordingly, the adaptive population activity A_a in the standard column is often weaker than the deviant. The population activity in the recording column reaches a steady state after the first several stimuli (bottom panel). The oddball stimuli sequence is shown under the top panel. **(B)** Excitatory population activity A_e in column 4 initiated by the deviant usually propagates more significantly into the recording column midway than by the standard-triggered activity in column 2 (top). Correspondingly, excitatory (pale red) and inhibitory (pale blue) populations in the recording column often respond stronger to the deviant compared to the standard stimuli (bottom). The time axis is vertically aligned for all panels.

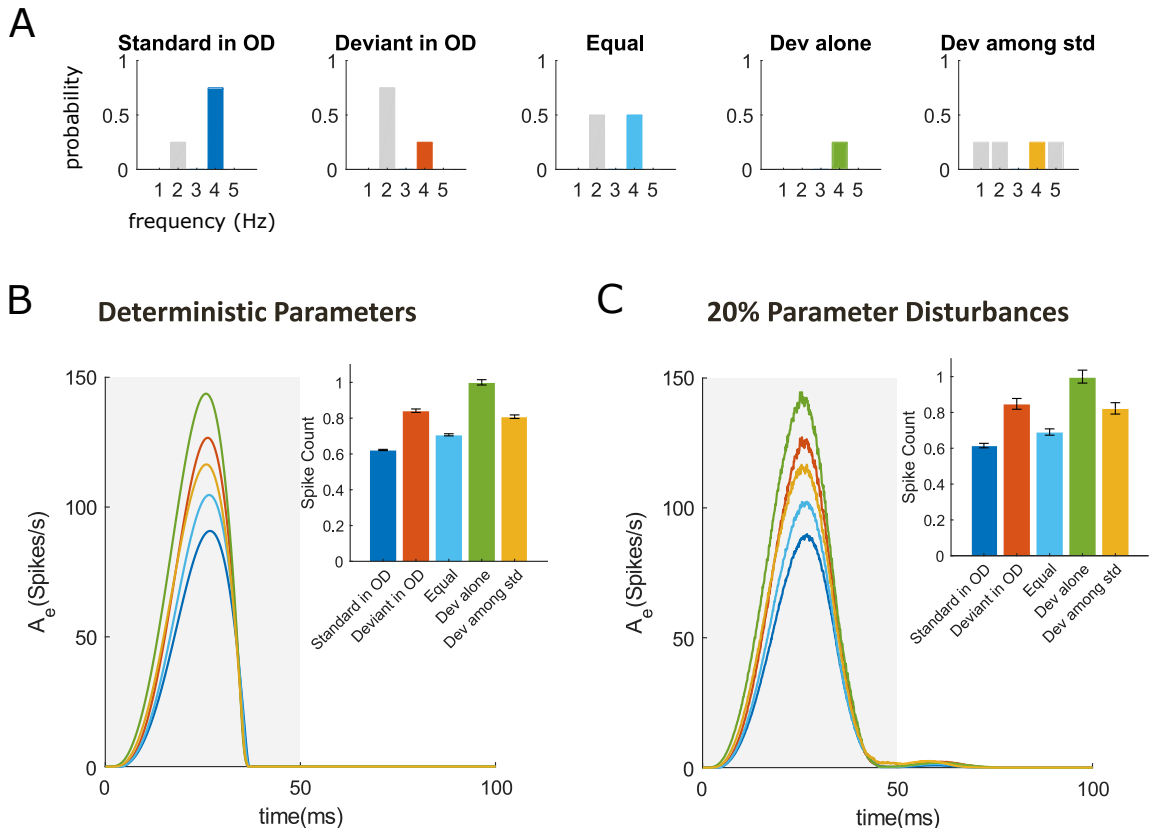

Figure 3.3: **Average responses of the model under five protocols. (A)** To investigate the deviant response characteristics, five control oddball (OD) protocols are used: "Standard in OD" in which a stimulus is presented to column 4 as a standard with a 75% probability of appearance (blue), "Deviant in OD" in which the same stimulus applied to column 4 is used as a deviant with a 25% probability of appearance (red), the "Equal" paradigm where two stimuli are presented 50% of the time each to column 2 and 4 respectively (cyan), "Dev alone" (green) where the stimulus is presented to column 4 alone as a deviant against silence and the "Dev among Std" (yellow) in which four different stimuli are presented 25% of the time each to column 1, 2, 4 and 5, respectively. **(B)** The population activity averaged over stimuli in the recording column evoked by tones of 4 Hz embedded in five different protocols. The gray shade marks the stimulus duration. Inset: the average spike counts with the standard error of the mean, normalized to the Deviant alone condition. Data are average over responses of recording column to auditory stimulation of 4 Hz in each protocol consisting of 800 stimuli in total. In all simulation protocols, network parameters are chosen from Table 3.1, input amplitude $A = 15$ spikes/s, stimulus duration = 50 ms and inter-stimulus interval (ISI) = 300 ms (stimulus offset to onset). **(C)** The same setup was employed as (B), except that 20% real-time parameter disturbances are introduced to parameters in Table 3.1, excluding number of columns M, width of tuning curve λ and membrane time constants τ, τ_e, τ_i that are taken from literature [Yarden and Nelken, 2017].

Besides the "traditional" OD protocol, three more protocols were tested on our network (the "Equal", "Dev alone", and "Dev among std"), which we also employed to test the neuronal

sensitivity to contextual regularity called true deviant detection (TDD) on the model (Fig 3.3A). The trends, across the different protocols, are qualitatively in line with the experimental findings in rat A1 [Taaseh et al., 2011] as well as spiking responses in the neuromorphic model, executed under biologically-realistic network noises [Vanattou-Saïfoudine et al., 2021]. (Fig 3.3B). For the "traditional" and equal oddballs, when the deviant column is stimulated, its excitatory populations receive less adapted inputs due to the rarity of the presentation of this stimulus. Consequently, more robust activity is propagated into the recording column midway, leading to the strongest response observed in the "Deviant in OD" protocol, followed by the "Equal" response and finally, the lowest response in the "Standard in OD". Besides, the "Dev alone" protocol evokes the largest response among all conditions as there is no cross-frequency adaptation caused by other stimuli. At last, the condition for the network demonstrating TDD is given in Eq 3.17 (refer to subsection 3.2.2 for details). Both positive SI and CSI values (SI = 0.1495, CSI = 0.0203) indicate the presence of SSA and TDD in our simulated network model (refer to 3.2.3 for the definitions of SI and CSI).

3.1.3 Robustness to Network Parameters

To test the robustness of the mean-field model, except for the network structure parameters, the adaptation and membrane time constants that are based on the literature[Yarden and Nelken, 2017], we simultaneously varied the rest free parameters in Table 3.1 by adding 20% online perturbations of their optimal values, leading to the fluctuation of actual parameters uniformly bound between 1-20% and 1+20% of their optimal values. We then averaged the responses of column 3 to tone of frequency 4 Hz within each protocol as done for the deterministic parameter circumstance. As a result, one can see that the parameter perturbation only increase the response variations and has a negligible effect on the SSA and IDD according to the indices calculated from mean responses (SI = 0.1592, CSI = 0.0151) (see Fig 3.3C). This demonstrates the robustness of our model against online parameter disturbances.

In addition, we did the sensitivity analysis of the free network parameters by varying each of them individually (with all the others fixed in their optimal values) and exploring the operating range of the parameter where SSA and TDD are both valid, based on positive metrics values. From parameter ranges reported in Table 3.1, one can see that the SSA and TDD of our model are relatively sensitive to the excitatory connection weights and then the inhibitory connection weights, but much less susceptible to the adaptation connection weight and adaptation scaling factor. The results meet our expectation since one critical element of SSA in our model is the extent of the spread of excitatory activities across columns that is mainly restricted by the strength of intra-column recurrent excitatory and inter-column connections.

3.2 Materials and Methods

We reduced the activity of each cortical column to the minimal mean-field rate representation, similar to Vasilaki and Giugliano [2014]; Esposito et al. [2015], to analyze its behavior. The stability analysis of the coupled excitatory and inhibitory populations was done in order to choose the parameter set leading to the biologically plausible network behavior. To be specific, we emulated the neuronal population bursts in response to transient stimuli recorded in A1

[Luczak et al., 2009, 2013; Sakata and Harris, 2009] by producing the oscillatory behavior near the fixed point of the excitatory-inhibitory network. The full mean-field network behaviors under five different protocols were then analyzed to conclude the consistency of the model prediction with the experimental observations [Taaseh et al., 2011]. A necessary condition for true deviance detection generation in the model is given in the end.

The architecture used in our mean-field model is closely related to the one proposed in Yarden and Nelken [2017] of A1. In this study, we used a network comprising five interconnected cortical columns. The thalamocortical inputs to each column linearly decrease along with the distance between the frequency of the stimulated tone and the column's best frequency on the frequency gradient of the auditory cortex. Specifically, every column is modeled by homogeneous excitatory and inhibitory populations that receive cross-adapted inputs from the adaptive layer when stimulated by a tone featured by a certain frequency, as shown in Fig 3.1A.

3.2.1 Transient Dynamics Analysis

In a similar vein as the derivation of a mean-field network in section 2.6, we can treat any neuronal population in our model as an arbitrary neuron in the pool, under the assumption that the population is fully-connected, large and homogenous. On the level of the collective activity of a neuronal population, the mean-field rate model is theoretically equivalent to a homogeneous group of leaky integrate-and-fire neurons [Gerstner et al., 2014].

The mean-field dynamics of an adaptive population is defined below, which a simplified version of the adaptation model proposed in [Benda and Herz, 2003]:

$$\begin{cases} \tau \dot{h}_a = -h_a + I_{ext} \\ \tau_a \dot{a} = -a + cA_a \end{cases} \tag{3.1}$$

where a represents the amount of adaptation accumulated by neurons, and A_a is the population activity defined by a threshold-linear activation function of the mean synaptic input h_a:

$$A_a = [h_a - a]_+ = \max(h_a - a, 0) \tag{3.2}$$

If the population fires at a constant rate A_a, the adaptation a will be gradually accumulated, leading to an effective decrease in the input current h_a and accordingly the population activity A_a.

Suppose that the synaptic current kinetics is much faster than the adaptation kinetics, namely, $\tau \ll \tau_a$, then Eqs 3.1 can be reduced to one equation by treating the synaptic current h_a as instantaneously reaching its asymptotic value I_{ext}:

$$\tau_a \dot{a} = -a + c(I_{ext} - a) \tag{3.3}$$

The solution for zero initial state $a(0) = 0$ is

$$a(t) = \frac{c}{1+c} I_{ext} \left(1 - e^{-\frac{1+c}{\tau_a} t}\right) \tag{3.4}$$

This means that if the external input I_{ext} never stopped, the adaptive threshold a would exponentially approach to its asymptotic value $cI_{ext}/(1+c)$ with time constant $\tau_a/(1+c)$

and exponentially decay to 0 with time constant τ when later external input was removed (Supplementary data S1a, top). Correspondingly, since $A_a(t) = I_{ext} - a(t)$ when external input is on, starting from I_{ext}, the population activity A_a will exponentially decay to its asymptotic value $I_{ext}/(1+c)$ with $\tau_a/(1+c)$. A_a will instantaneously drop to zero with no external input (see Fig 3.2A).

In addition, the network model for interacting excitatory and inhibitory populations is adapted from the well-known Wilson and Cowan's model [Wilson and Cowan, 1972], which is a typical neural circuit for generating oscillatory behaviors due to exhibitory-inhibitory imbalance. A similar dynamical analysis for Wilson and Cowan's model can be found in the textbook by Dayan and Abbott [2001]. Here we use the type of oscillator network to resemble the cortical response to transient sensory stimulation. The dynamics of interacting mean-field excitatory and inhibitory populations is described as

$$\begin{cases} \tau_e \dot{h}_e = -h_e + w_{ei} A_i + w_{ee} A_e + w_a A_a \\ \tau_i \dot{h}_i = -h_i + w_{ie} A_e + w_{ii} A_i \end{cases} \tag{3.5}$$

Similarly, the excitatory or inhibitory population activity A_e or A_i are defined as the rectified linear activation function of its mean synaptic inputs $A_e = [h_e]_+$ or $A_i = [h_i]_+$. w_{ei}, w_{ee}, w_{ie}, w_{ii} and w_a are the inhibitory-to-excitatory, excitatory-to-excitatory, excitatory-to-inhibitory, inhibitory-to-inhibitory and adaptive-to-excitatory connections, respectively.

To analyze the stability of the fixed point of the excitatory-inhibitory network, we rewrite the system (3.5) in the standard form by assuming the synaptic currents h_e and h_i are non-negative

$$\begin{cases} \dot{h}_e = G_1(h_e, h_i) = \left[(w_{ee} - 1)h_e + w_{ei} h_i + w_a A_a\right]/\tau_e \\ \dot{h}_i = G_2(h_e, h_i) = \left[(w_{ii} - 1)h_i + w_{ie} h_e\right]/\tau_i \end{cases} \tag{3.6}$$

By definition, the Jacobian matrix J of above system at the fixed point (h_{e0}, h_{i0}) is

$$J(h_{e0}, h_{i0}) = \left(\begin{array}{cc} \partial G_1/\partial h_e & \partial G_1/\partial h_i \\ \partial G_2/\partial h_e & \partial G_2/\partial h_i \end{array} \right)\Bigg|_{(h_{e0}, h_{i0})} = \left(\begin{array}{cc} (w_{\mathrm{ee}} - 1)/\tau_{\mathrm{e}} & w_{\mathrm{ei}}/\tau_{\mathrm{e}} \\ w_{\mathrm{ie}}/\tau_{\mathrm{i}} & (w_{\mathrm{ii}} - 1)/\tau_{\mathrm{i}} \end{array} \right) \tag{3.7}$$

which is independent of the values of fixed point and external input $w_a A_a$.

The stability of the fixed point is determined by the sign of the real parts of the eigenvalues of its Jacobian matrix J. The pair of eigenvalues $\lambda_{1,2}$ is given by

$$\lambda_{1,2} = \frac{1}{2}\left(\frac{w_{\mathrm{ee}} - 1}{\tau_{\mathrm{e}}} + \frac{w_{\mathrm{ii}} - 1}{\tau_{\mathrm{i}}} \pm \sqrt{\left(\frac{w_{\mathrm{ee}} - 1}{\tau_{\mathrm{e}}} - \frac{w_{\mathrm{ii}} - 1}{\tau_{\mathrm{i}}} \right)^2 + \frac{4 w_{\mathrm{ei}} w_{\mathrm{ie}}}{\tau_{\mathrm{e}} \tau_{\mathrm{i}}}} \right) \tag{3.8}$$

Assuming that $\tau_e = \tau_i = \hat{\tau}$, the above expression can be further simplified as

$$\lambda_{1,2} = \frac{1}{2\hat{\tau}}\left(w_{\mathrm{ee}} + w_{\mathrm{ii}} - 2 \pm \sqrt{(w_{\mathrm{ee}} - w_{\mathrm{ii}})^2 + 4 w_{\mathrm{ei}} w_{\mathrm{ie}}} \right) \tag{3.9}$$

It turns out that the eigenvalues are a complex conjugate pair with positive real parts following the optimal parameter values given in Table 3.1. Therefore the equilibrium (h_{i0}, h_{e0}) is an unstable focus, which means that the phase-plane trajectory spirals away from the equilibrium until it joins a stable limit cycle induced by the rectification nonlinearity. The corresponding behaviours of h_e and h_i settle into periodic oscillations (see Fig 3.4A). In this case, the population exhibits a synchronized activity in response to a momentary stimulus, which is qualitatively similar to the sensory responses recorded in the auditory cortex (see Fig 3.4B).

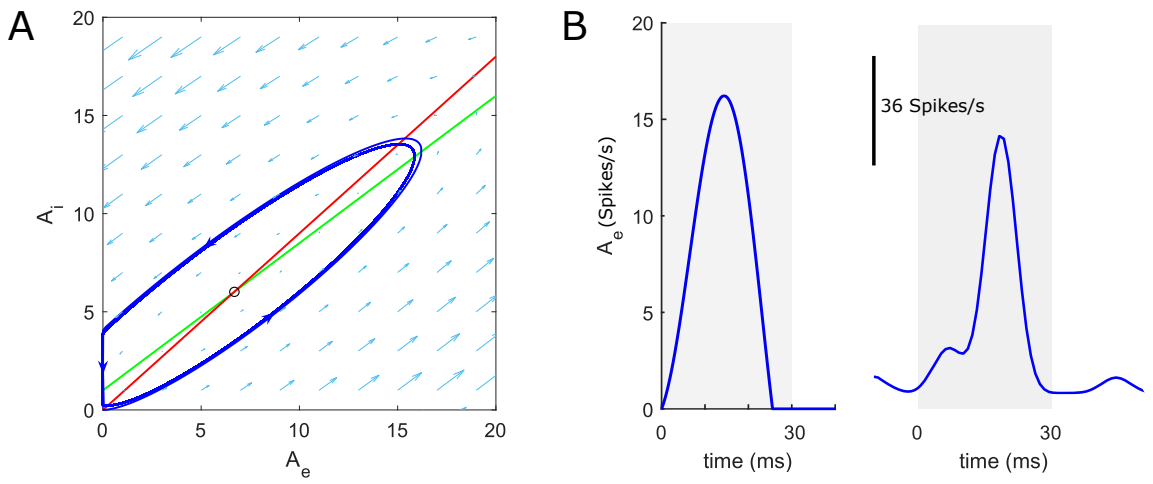

Figure 3.4: **Trajectory of the excitatory-inhibitory model (3.5) when $w_a A_a = 3$. (A)** The intersection of the h_e- (red line) and h_i-nullclines (green line) is an unstable focus (open cirlce) that is enclosed by a stable rectified limit cycle. The phase-portrait trajectory initiated at (h_e, h_i)=(0,0) immediately enters the limit cycle and thereafter oscillates periodically. **(B)** Left panel: population activity of the excitatory population in response to a momentary stimulus whose duration is marked in grey. Right panel: for comparison, a population burst recorded in A1 evoked by a tone stimulus with the same duration. Experimental recordings are reproduced from Hershenhoren et al. [2014a]. Network parameters are chosen from the optimal values of the corresponding parameters in Table 3.1.

3.2.2 SSA in a Recurrent Network with Spike-frequency Adapted Input

The whole network of cortical columns is described by the following rate equations:

$$\tau \dot{h}_a^Q = -h_a^Q + \sum_{f=1}^{M} A\xi_f T_f^Q \tag{3.10}$$

$$\tau_a \dot{a}^Q = -a^Q + c[h_a^Q - a^Q]_+ \tag{3.11}$$

$$\tau_e \dot{h}_e^Q = -h_e^Q + \sum_{R=-1}^{1} w_{ee}^{|R|}[h_e^{Q+R}]_+ + w_{ei}[h_i^Q]_+ + w_a[h_a^Q - a^Q]_+ \tag{3.12}$$

$$\tau_i \dot{h}_i^Q = -h_i^Q + w_{ie}[h_e^Q]_+ + w_{ii}[h_i^Q]_+, \; Q = 1, 2, \cdots, M \tag{3.13}$$

where Eqs 3.10 and 3.11 describe dynamics of network of adaptive populations, which are expanded from single adaptive population dynamics defined by Eq 3.1. In a like manner, Eqs 3.12 and 3.13 are adapted from Eq 3.5, respectively representing the dynamical behaviour of excitatory and inhibitory populations within interconnected columns.

Every column Q is tonotopically organized along the frequency channel axis ordered from 1 to M by its preferred frequency $f = Q$ (best frequency). The temporal profile of external

stimulation to certain frequency channel f is given by firing rate $A\xi_f(t)$, where A and $\xi_f(t)$ respectively denote the maximum amplitude and normalized temporal envelope (a fraction between 0 and 1) of the stimulus. In the simulation, $\xi_f(t)$ is chosen as a trapezoid pulse (50 ms duration with 5 ms onset/offset linear ramps).

T_f^Q represent the relative strength of input column Q receives from frequency channel f compared with its best frequency channel Q. The values of T_Q^f over all frequency channels form the tuning curve of column Q's thalamic inputs. Here we choose a linear tuning curve:

$$T_f^Q = \left[1 - \frac{|Q-f|}{\lambda}\right]_+ \tag{3.14}$$

where λ is the width of the tuning curve describing the level of cross-frequency thalamic inputs that cortical column Q receives (see Fig 3.1B for an example of the tuning curve of column Q=3's thalamocortical inputs).

The following demonstrates analytically the conditions at which the network is able to produce true deviance detection, i.e. response to deviant in the oddball protocol is larger than to the same deviant in 'Dev among std'. Taking inspiration from Taaseh et al. [2011], we define the accumulated adaptation load of column Q during certain protocol as

$$L(Q) = \sum_{f=1}^{M} p_f T_f^Q \tag{3.15}$$

where p_f is the probability of appearance of the tone of frequency f. Here T_f^Q can be considered as the extent of cross-frequency adaptation in column Q caused by tone stimulation of frequency f

According to the network structure, the excitatory population in the recording column is excited by feedforward adapted inputs and lateral inputs from its neighbouring columns. The steady state of excitatory synaptic inputs to the excitatory population in the recording column N on the presentation of tones of frequency $N+1$ in a protocol is formulated as

$$I_{Protocol}(N) = w_a C T_{N+1}^N B^{L(N)} + w_{ee}^1 f\left(w_a C B^{L(N+1)}\right) \tag{3.16}$$

where $B < 1$ represents the steady-state response of adaptive population in a column to tones of the column's best frequency within a sequence consisting of only tones of that frequency, normalized by the unadapted response C. Thus the more adaptation loads the adaptive population has, the weaker its responses are, and accordingly, the lower inputs to the excitatory population. $f(\cdot)$ denotes the equivalent activation function of the excitatory population mediated by the inhibitory population, which is a monotonically increasing function given the parameter values.

In all simulation protocols, $N = 3$ and $\lambda = 2$. Hence, in "Standard in OD" protocol, $L(3) = 0.25T_2^3 + 0.75T_4^3 = 0.5$ and $L(4) = 0.25T_2^4 + 0.75T_4^4 = 0.75$. Similarly, $L(3) = 0.5T_2^3 + 0.5T_4^3 = 0.5$ and $L(4) = 0.5T_2^4 + 0.5T_4^4 = 0.5$ for "Equal"; $L(3) = 0.75T_2^3 + 0.25T_4^3 = 0.5$ and $L(4) = 0.75T_2^4 + 0.25T_4^4 = 0.25$ for "Deviant in OD"; $L(3) = 0.25T_4^3 = 0.125$ and $L(4) = 0.25T_4^4 = 0.25$ for "Dev alone" and at last $L(3) = 0.25T_1^3 + 0.25T_2^3 + 0.25T_4^3 + 0.25T_5^3 = 0.25$ and $L(4) = 0.25T_1^4 + 0.25T_2^4 + 0.25T_4^4 + 0.25T_5^4 = 0.375$ for "Dev among std". Due to $B < 1$, we get $I_{Standard\ in\ OD}(3) < I_{Equal}(3) < I_{Deviant\ in\ OD}(3) < I_{Dev\ alone}(3)$.

Notation	Description	Optimal value	Operating range
M	Number of columns	5	
λ	Width of tuning curve	2	
τ_a	Time constant for adaptation process	1 s	
τ	Membrane time constant of adaptive population	1×10^{-3} s	
τ_e	Membrane time constant of excitatory population	5×10^{-3} s	
τ_i	Membrane time constant of inhibitory population	5×10^{-3} s	
w_{ee}^0	Intra-columnar excitatory-to-excitatory connection weight	3.25	[3.2, 3.5]
w_{ee}^1	Inter-columnar excitatory-to-excitatory connection weight	0.2	[0.15, 0.35]
w_{ie}	Intra-columnar excitatory-to-inhibitory connection weight	1.8	[1.75, 1.95]
w_{ei}	Intra-columnar inhibitory-to-inhibitory connection weight	-3	[-3.3, -2.85]
w_{ii}	Intra-columnar inhibitory-to-inhibitory connection weight	-1	[-1.1, -0.7]
w_a	Connection weight between adaptive and excitatory populations	0.5	(0, 2.5]
c	A constant to scale the rate of adaptive population	20	(1, 100]

Table 3.1: **Parameters in the mean-field model.** Steps of 0.05, 0.1 and 1 are used to linearly search the operating ranges of the excitatory and inhibitory connection, adaptation connection weights and adaptation scaling factor, respectively. The parameter value is valid only if both SI and CSI metrics for the model based on that parameter value are positive. Values exceeding 100 for the adaptation scaling factor have not been tested since SSA and TDD vary very slightly when increasing the parameter linearly. The protocol setup used here to compute SI and CSI is the same as that in Fig3.3.

In order to demonstrate TDD, we need to let $I_{Deviant\ in\ OD}(3) > I_{Dev\ among\ std}(3)$, that is to say, the following difference should be positive:

$$w_a C T_{N+1}^N \underbrace{\left(B^{0.5} - B^{0.25}\right)}_{<0} + w_{ee}^1 \underbrace{\left[f\left(w_a A B^{0.25}\right) - f\left(w_a A B^{0.375}\right)\right]}_{>0} > 0 \tag{3.17}$$

The first and second term in the expression is respectively negative and positive. Hence, possible ways of tuning parameter to achieve TDD are to make the second term more positive by increasing inter-columnar connection w_{ee}^1 and/or decreasing the intra-columnar adaptive-to-excitatory connection w_a.

All the model parameters are summerized in Table 3.1. The network was numerically simulated by Euler's method with a time step 1×10^{-4} s.

3.2.3 Data Analysis

The average response to a tone stimulus is evaluated as mean spike counts of the middle column (column 3) of the network triggered by all presentations of the tone. The average spike counts are calculated as the integrals of population activity with a temporal resolution of 1 ms over a window of 100 ms starting at the stimulus onset.

The strength of SSA is quantified by the index defined before in Eq 2.1, where $\mathrm{s}(x_i)$ and $\mathrm{d}(x_i)$, $i \in \{1, 2\}$, are the average responses measured as spike counts to the tone frequency x_i used as the standard and deviant stimulus respectively.

To quantify the neuronal sensitivity to a regular pattern in the presentation context, we use a context-specific index (CSI) by modifying the SSA index mentioned above accordingly. Instead of employing the deviant and standard responses, we compare the deviant responses

in the "Deviant in OD" and "Deviant among Standards" conditions.

$$\mathrm{CSI} = \frac{\mathrm{d}_{\text{Deviant in OD}}(x) - \mathrm{d}_{\text{Dev among std}}(x)}{\mathrm{d}_{\text{Deviant in OD}}(x) + \mathrm{d}_{\text{Dev among std}}(x)} \tag{3.18}$$

where $\mathrm{d}_{\text{Deviant in OD}}(x)$ and $\mathrm{d}_{\text{Dev among std}}(x)$ are the average spike counts to the same deviant frequency x in the "Deviant in OD" and "Deviant among Standards" protocols, respectively. If the value of CSI is positive, then we conclude that the model displays true deviance detection.

The code for the mean-field model can be accessed via `https://github.com/ChaoHan-UoS/AuditorySSA_Hardware`.

3.3 Summary

This chapter aims to investigate the underlying neural circuit giving rise to SSA in the auditory cortex. To this end, we recreated a simplified version of the SSA network proposed in Yarden and Nelken [2017] using a minimal mean-field rate model. The mean-field model is capable of robustly reproducing previously reported experimental observations in five SSA paradigms, even under 20% online disturbances of the optimal values of network parameters. Our collaborators at Institute of Neuroinformatics, Zurich, further expanded our mean-field rate model to a spiking version validated on neuromorphic hardware [Vanattou-Saïfoudine et al., 2021]. We concluded that interactions among differential adapted populations of tuned neurons are crucial in the development of SSA.

Chapter 4

Modelling Novelty Detection in the Thalamocortical Loop

As discussed earlier in sections 2.3 and 2.4, Musall et al. [2017] observed a deviant-specific spiking response with a comparable onset latency as MMN, in the region of the primary somatosensory cortex (S1) processing whisker information. To the best of our knowledge, the neural circuitry and mechanism giving rise to this phenomenon remain unknown.

In this chapter, we adapted and generalized the auditory SSA model developed by Yarden and Nelken [2017] to a hierarchal thalamocortical model for the whisker-related area of S1 (also known as the barrel cortex) with multiple timescales of synaptic and neuronal dynamics to capture the biphasic activity that demonstrates both signatures of SSA and MMN [Musall et al., 2017]. On the short timescale of the early response, a vital element of the transient dynamics of the cortical circuitry is the so-called population spike (PS), characterized by a burst of firings of a pool of neurons evoked by a brief stimulus. The synchronization of neuronal firings facilitate the robust propagation of population activity elicited in a cortical column into its surrounding column. On the longer timescale, we hypothesize that the late rhythmic cortical response is inherited from the thalamic spindle oscillation initiated by the interplay between the thalamic reticular (RE) and thalamocortical (TC) relay cells. The somatosensory thalamocortical network retains only those crucial neurobiological features, such as the laminar architecture and somatotopic organization of the barrel cortex, synaptic connectivity between the barrel cortex and its corresponding thalamic nuclei, as well as different synaptic receptors within thalamic nuclei, that are sufficient to elicit the late context-dependent deviant responses. Finally, we test the cortical capacity for predictive coding using more complex stimulus sequences. The results of our model indicate that the precision of expectation generated by sequence history impacts the neural response to stimuli that interrupt those expectations.

The chapter is primarily based on Han et al. [2021] that is under review while writing the dissertation.

4.1 Results

The architecture of the somatosensory thalamocortical network is organized in a loop as shown in Fig 4.1A. Based on the stereotypic somatotopic map for representing rodent whiskers, the

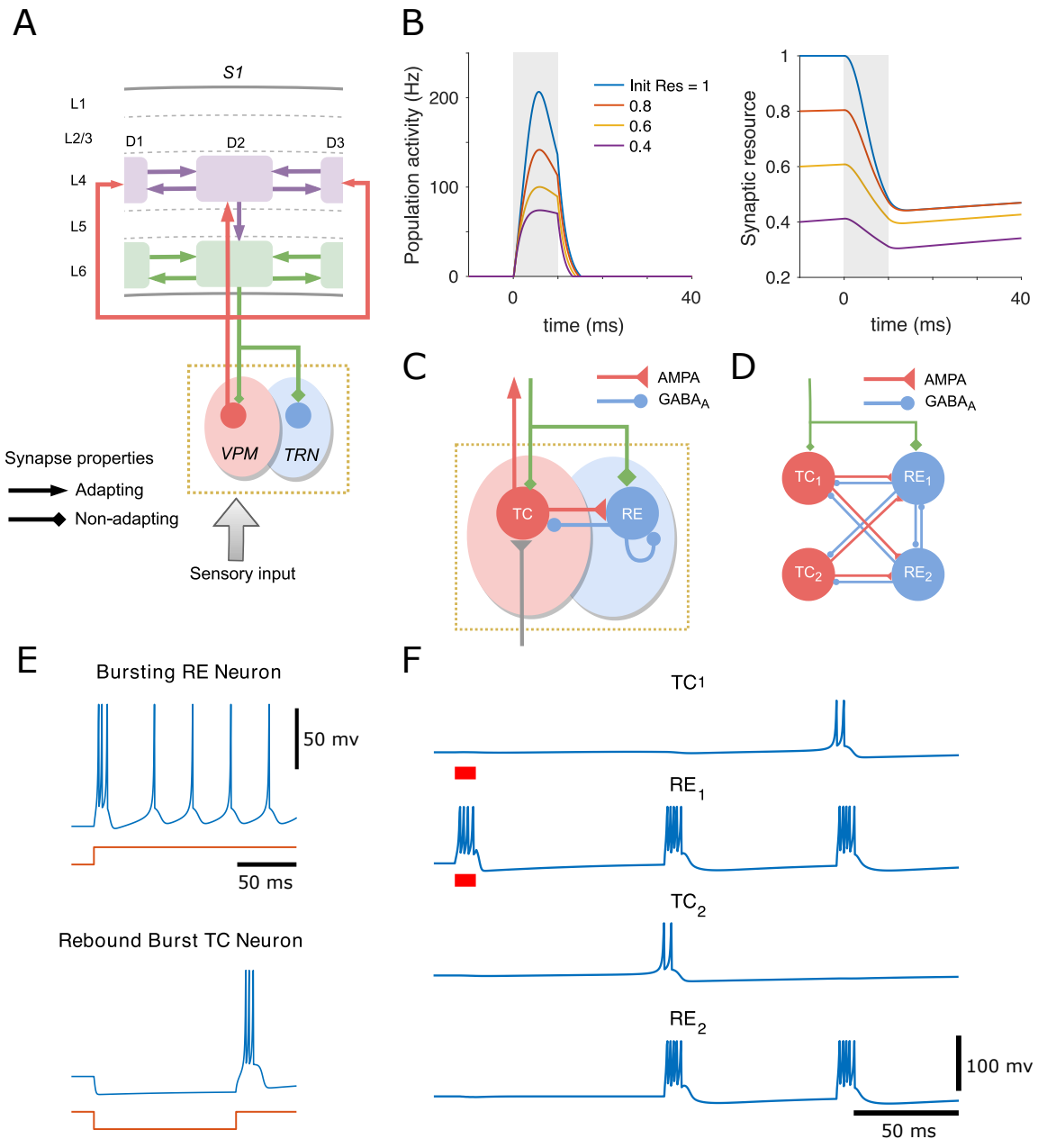

Figure 4.1: **Structure and dynamics of the thalamocortical network and minimal thalamic circuit inducing spindle oscillations. (A)** The ventral posterior medial (VPM) nucleus of the thalamus (pink shade), which is somatotopically arranged into discrete clusters called "barreloids", relays peripheral sensory information to the L4 barrels (purple), with the relative strength of connections represented by the size of the arrows. Each L4 cluster projects its output vertically to the corresponding L6 infrabarrel (green), which in turn provides feedback excitation to its somatotopic barreloid and thalamic reticular nucleus (TRN, in blue shade), with higher cortical drive on TRN than VPM (specified by the size of the diamonds). Each L4 barrel and L6 infrabarrel also connects to its neighbouring barrels. **(B)** Population spike (PS) can be generated in the cortical cluster, represented by the transient increase of

the cluster's mean firing rate (left). Mean synaptic resources of the neuronal group are depleted by the PS and later recover gradually (right). Higher levels of initial resources evoke a more substantial PS. The time 0 represents the onset of a stimulus whose duration is marked in grey shading. **(C)** Two classes of cells with different synaptic receptors are used in the thalamic circuit: thalamocortical (TC) relay cells with excitatory AMPA-mediated synapses in the VPM and reticular (RE) cells with inhibitory $GABA_A$-mediated synapses in the TRN. TC and RE neuronal populations are mutually coupled, and the RE population is also recurrently connected. The thalamus can be activated by bottom-up sensory input (grey) and top-down cortical drive (green). **(D)** A minimal thalamic network of 2 TC and 2 RE cells driven by the cortex is capable of inducing spindle oscillations. **(E)** Voltage traces of two types of thalamic cells simulated by the izhikevich model in response to a step in direct current (bottom of each panel): initial burst tonic firing in TC neuron (-70 mV at rest, hyperpolarizing pulse of 10 pA) and rebound burst in RE neuron (-62.5mV at rest, depolarizing pulse of -10 pA lasting 120 ms). **(F)** The time course of membrane potentials for the 4 thalamic cells. Cortical stimulation is marked by red bars under the traces (0.08 pA for TC_1 and 20 pA for RE_1).

barrel cortex is modeled as a grid of interconnected vertical columns, each of which primarily mediates stimuli from its principal whisker. Also, we only focus on neuronal clusters within layers 4 (barrels) and 6 (infrabarrels) of each cortical column, which respectively act as the afferent and efferent layers of the cortex from and to the thalamus [Crandall et al., 2017; Petersen, 2007; Whitmire and Stanley, 2016]. The barreloids in the ventral posterior medial (VPM) nucleus of the thalamus show an identical layout to that of the barrels. Each barreloid relays the information from its principal whisker through an ascending sensory pathway to L4 of its principal as well as surrounding barrels, with stronger coupling to the principal than the surrounding ones. We hypothesize that L2/3 and L5 contribute little to the cortical SSA but merely transmit columnar information from L4 to L6. Therefore, a direct shortcut connection from L4 to L6 within every cortical column is added, which is sufficient in the simplified model proposed here to demonstrate SSA in L6, although in reality the flow of excitation from L4 will go through L3, L2 and L5 in sequence before arriving at L6 [Armstrong-James et al., 1992; Fox, 2008; Feldmeyer, 2012]. To close the thalamocortical loop, the outputs of layer 6 are projected back to their somatotopically aligned barreloids and thalamic reticular population. Here we assume that neurons in the thalamic reticular nucleus (TRN) are also whisker-specific, but they are not necessarily arranged in spatially discrete clusters as barreloids.

We first focus only on the cortical network that is capable of generating SSA and true deviance detection by modulation of cross-barrel synchronous activities. All cortical intra- and inter-column synapses exhibit short-term depression. Only excitatory populations are considered in the cortical circuitry to simplify the mean-field analysis, and the level of excitability can be regulated by the threshold of neuronal gain function. Every layer-specific population in columns is modelled by a mean-field recurrent network coupled with depressing synapses, which can generate the so-called population spike (PS) in response to external stimulation. The PS is characterized by nearly synchronous firing of a group of neurons within a short time window [Tsodyks et al., 2000; Loebel and Tsodyks, 2002]. In the cortical mean-field rate populations we studied here, a population spike is described as a sharp increase in population activity, leading to fast depletion of mean synaptic resource that will recover

gradually after the removal of external input. Higher initial synaptic resources cause more substantial PSs to be triggered in response to the same stimulus (Fig 4.1B).

The thalamic circuitry is set up to induce spindle oscillations characterized by intermittent bursting activity at 7-14 Hz, which is postulated as the origin of the late rhythmic responses in L4 of S1 [Musall et al., 2017], since the frequency of cortical oscillation (∼10Hz) falls into the spindle range. Early studies have shown that network and intrinsic mechanisms act in combination to generate spindle oscillations [Timofeev and Bazhenov, 2005]. Two interacting populations of spiking thalamic neurons are considered to model spindle oscillation: excitatory thalamocortical (TC) neurons and inhibitory thalamic reticular (RE) neurons. As illustrated in Fig 4.1C, the TC population is excited by whisker stimulation and then provides AMPA-mediated excitatory synaptic input to the RE population and neurons in cortical L4. In turn, the RE neurons suppress the activity of the TC neurons and themselves via $GABA_A$-mediated inhibition. The top-down modulation from L6 exerts more effective excitation on RE than TC population due to more substantial corticothalamic conductance [Destexhe et al., 1998; Timofeev and Bazhenov, 2005; Petersen, 2007]. Unlike the mean-field dynamics of the cortex represented by the rate network, the intrinsic dynamics of individual spiking thalamic neuron is simulated by the izhikevich model (see Materials and methods section for model details). RE neurons exhibit transient bursting followed by tonic firing in response to depolarizing step currents, and TC relay neurons can trigger rebound spike-bursts upon release from hyperpolarization [Timofeev and Bazhenov, 2005; Izhikevich and Edelman, 2008] (Fig 4.1E).

We elucidate the mechanism of spindle oscillation triggered by the corticothalamic feedback excitation in a minimal thalamic circuit consisting of 2 RE and 2 TC cells [Destexhe et al., 1998] (Fig 4.1D). When the cortical drive is exerted on both TC_1 and RE_1 neurons, only the RE_1 neuron is able to elicit spikes due to the considerably higher cortical conductance on RE than TC neurons. The bursting activity of RE_1 cell provide inhibitory postsynaptic potentials (IPSPs) to both TC cells and then induce rebound bursting spikes in TC_2 cells when its potential is released from hyperpolarization and almost reaches its resting state. The firing of TC_2 cell reciprocally excite RE cells, leading to the next cycle of spindle oscillations (Fig 4.1F). Minute bump in the potential of TC_1 cell evoked by initial corticothalamic stimulation prevents the cell from generating rebound bursts in the first cycle, which results in the TC cells fire every two cycles while RE cells fire every cycle in the minimal circuit. With the thalamocortical pathway, spindle oscillation is synchronized over cortical areas as hypothesized to be the late oscillatory rhythms found in L4 of S1 [Musall et al., 2017].

4.1.1 Propagation of Population Activity in the Thalamocortical Loop

The PS evoked in the principal L4 barrel will propagate to the neighbouring L4 barrels via inter-column coupling, as well as to the column-aligned L6 infrabarrel via the L4-to-L6 projection [Armstrong-James et al., 1992]. The L4 responses are largely localized around the principal barrel and substantially attenuated in the surrounding barrels [Petersen and Diamond, 2000b; Lebedev, 2002] (Fig 4.2A, top panel). Responses across L4 barrels start almost simultaneously about 7 milliseconds after stimulus onset, in accordance with experimental findings [Oberlaender et al., 2012] (Fig 4.2B, top). The topographical spread of PS in L6 is similar to that in L4 but with broader temporal profiles of the PSs and slightly longer response latencies [Oberlaender et al., 2012] (Fig 4.2A and B, bottom panels). The extent of

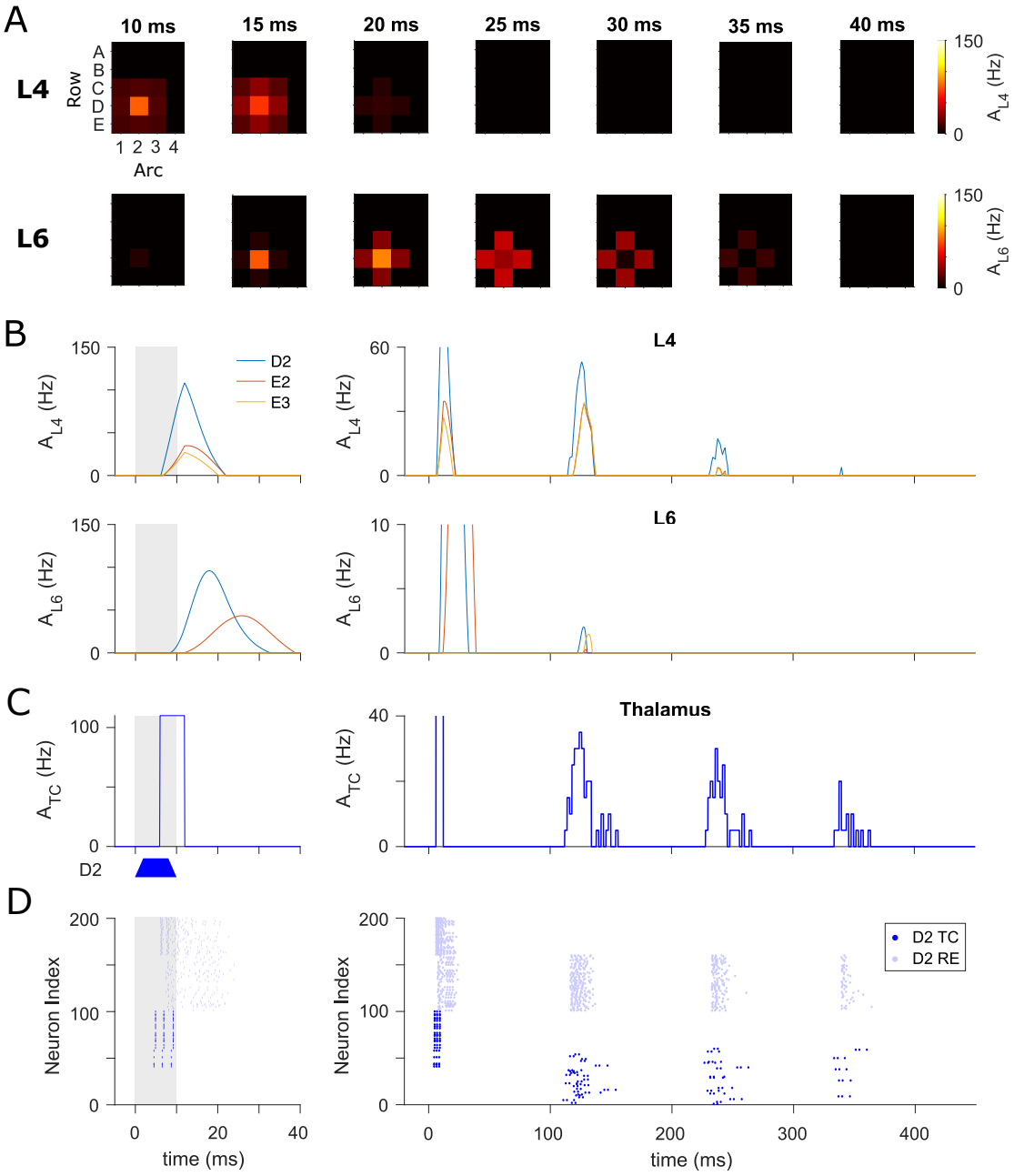

Figure 4.2: **Spatiotemporal distribution of whisker-evoked responses in cortical columns and thalamus. (A)** Top: Snapshots of cortical spatial activity patterns in L4, A_{L4}, at different times after the onset of a brief stimulation to D2 whisker. Bottom: Analogous plots for L6 populations, where population spikes occur later and over longer periods. **(B)** Temporal profiles of mean firing rate of three cortical columns in L4 (top) and L6 (bottom) respectively. Stimulus duration is indicated in grey shading. The timescale of the left panels matches that in (A) to show the behaviour of the initial population spikes evoked by sensory stimulation. Late spike-bursts in L4 driven by secondary thalamocortical input are displayed over a longer timescale (top-right), but the additional late responses can hardly spread into

L6 (bottom-right). **(C)** Population activity averaged over all 100 TC neurons in D2 barreloid, A_{TC}. The early synchronous activity is induced by the principal whisker deflection and relays the sensory information to the L4 barrels (left). The trapezoid form of stimulus is illustrated under the plot. The initial L6 population spike (bottom-left panel in (B)) projects back to the thalamus, eliciting the late oscillatory rhythm that begins after ∼110 ms of stimulation onset at an interval of ∼110 ms (right). **(D)** Spike raster for the full thalamic network of D2 barreloid (from which the population activity of TC cells shown in (C) is calculated). The circuit is composed of 100 TC cells (dark blue) in the D2 barreloid and 100 reciprocally coupled RE cells (light blue). The time axes of plots in (B), (C) and (D) are aligned.

propagation is determined by the width of the cross-whisker tuning curve of thalamocortical input, the strength of inter-column connections as well as the excitability of the neural population, where a broader tuning curve, stronger inter-population couplings and a more excitable population give rise to more extensive propagation.

In addition, a grid of barreloids in the VPM of the thalamus was simulated to convey sensory signals from individual whiskers to the cortex. The activity evoked by the principal whisker in barreloid D2 will not spread into other barreloids because no recurrent connections are made within the VPM, as observed experimentally [Brecht and Sakmann, 2002]. The following describes the information flow within the thalamocortical system in response to a D2 whisker deflection: The D2 whisker stimulation first excites the D2 barreloid through a bottom-up connection to TC neurons, and subsequently induces transient PSs in the cortical L4 populations via thalamocortical synapses. Onset PSs are also evoked in L6 through intra-cortical connections and projected back to the thalamus, where the spindle oscillations are initiated after the instantaneous onset response. As can be seen in Fig 4.2C and D, the TC cell group in the D2 barreloid exhibits a transient increase in population activity right after peripheral stimulation (early response), followed by substantial synchronized activities in the sleep spindle frequency range (roughly 9 Hz, late response). The diminishing late oscillation triggered by corticothalamic excitation occurs about 110 milliseconds after stimulus onset and lasts a couple of cycles before termination. Finally, the late thalamic activity gives rise to a secondary cortical response in L4 that only barely propagates to L6 (Fig 4.2B, right panels).

4.1.2 SSA Arises from Adapted PS to Standard Stimuli

SSA is generally tested in the oddball protocol, where repetitive deflection is applied to one whisker (called the standard), and the sequence is randomly interrupted from time to time by deflection of another whisker (called the deviant). The roles of whiskers as standard and deviant are swapped to disambiguate any response preference for individual whisker from the deflection probability on SSA [Musall et al., 2017].

In accordance with experimental findings [Musall et al., 2017], the L4 network exhibits prominent SSA in the late response but hardly in the initial population spike, where a similar amplitude of activity is elicited by whisker stimulation, regardless of its identity as either standard or deviant (Fig 4.3A, middle panel). The initial PS is relatively unaffected because the depleted synaptic resources in the L4 populations can always recover almost fully before the next stimulus is presented (Fig 4.3A, bottom). In contrast, depleted synaptic resources in L6 of the standard column (D2) recover only partially during the inter-stimulus interval, due

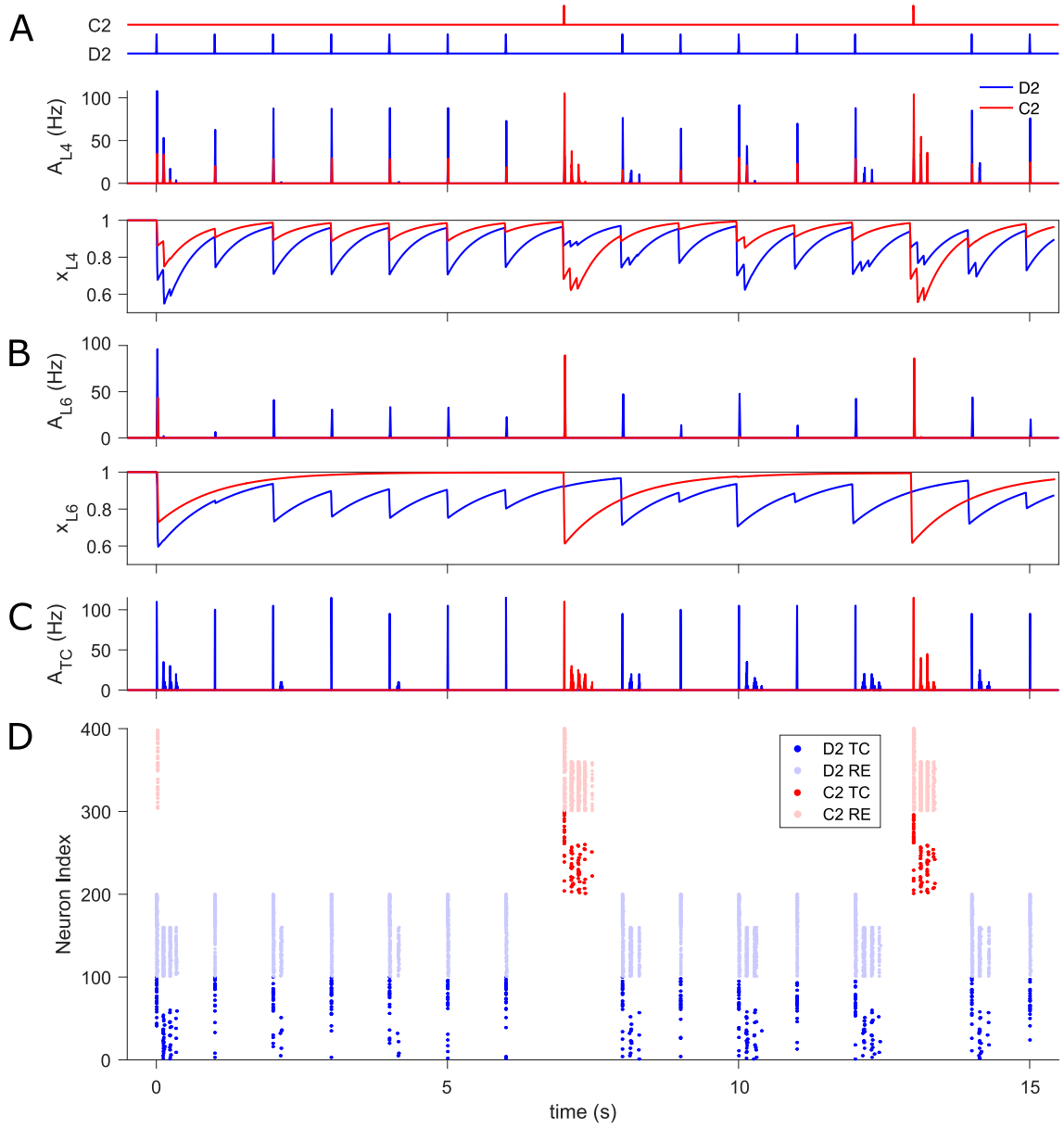

Figure 4.3: **SSA in the thalamocortical network (A)** Time course of the whisker oddball protocol (top), L4 population activity, A_{L4}, (middle), and mean synaptic resources, x_{L4}, (bottom) in standard D2 (blue traces) and deviant C2 (red) barrel. Both standard and deviant stimuli evoke comparable amplitude of PS (the first burst in response to each stimulus) due to the fast recovery of depleted resources. Conversely, late oscillations occur primarily in response to deviant stimuli. **(B)** L6 population activity, A_{L6}, (top) and mean synaptic resources, x_{L6}, (bottom) in D2 and C2 infrabarrel. The deviant stimulus generally triggers more substantial PSs than the standard because of more resource available upon presentation of the deviant than the standard. **(C)** Population activities of TC cells, A_{TC}, in D2 and C2 barreloids. The deviant corticothalamic PS is always strong enough to initiate late thalamic oscillation, yet this is not the case for the standard. **(D)** Spike raster of thalamic neurons in D2 (blue) and C2 (red) barreloids. The transient activity of C2 RE cells in response to the first stimulation

are caused by the significant cortical cross-column L6 feedback. The time axes of all plots are aligned. In the oddball paradigm, peripheral stimulus duration is 10 ms and inter-stimulus interval (ISI) is 1 s (onset-to-onset).

to the relatively long synaptic recovery time constant in this layer (Fig 4.3B, blue trace in bottom panel). The insufficient resources, along with depressed L4-to-L6 input in D2 caused by repeated standard stimuli, can only trigger inadequate PSs (Fig 4.3B, blue trace in top panel). However, as a result of the low probability of deviant stimuli, synaptic resources in L6 of the C2 cluster often reach their steady-state (Fig 4.3B, red trace in bottom panel) and accordingly evoke full PSs when the deviant stimulus finally arrives (Fig 4.3B, red trace in top panel). In other words, SSA is initiated in L6 during the early PS phase, where the cortical L6 drive on the deviant barreloid is always strong enough to elicit spindle oscillation, while it is in most cases too weak to do the same on the standard barreloid (Fig 4.3C and 4.3D). These late deviant-selective oscillations in the thalamus are then fed back to L4 populations and result in the late L4 rhythmic dynamics exhibiting SSA, consistent with the experimental results that the late cortical responses are very likely to be evoked by secondary thalamocortical inputs to L4 [Musall et al., 2017].

It is worth noting that late thalamic and synchronized cortical oscillations are also sometimes elicited in D2 by the standard stimulation, and especially the standard stimulus immediately following a deviant (Fig 4.3C, D and middle panel of A), in accordance with the sporadic standard-triggered oscillation recorded in L4 of the cortex [Musall et al., 2017]. Here these late rhythmic responses occur largely because the deviant did not deplete as many resources in L6 of D2 as another standard would have done, which occasionally allows strong enough cortical drive on the thalamus to generate a spindle oscillation during the presentation of the next standard stimulation.

4.1.3 Context-dependent Deviance Detection

It has been suggested that the larger responses to the stimuli features, when deviant than when standard, not only reflect the rarity of the deviant causing less use-dependent adaptation [Nelken, 2014], but also the sensitivity to the violation of the expectation established by the repeated standard, indicating a dynamic prediction mechanism for the next stimulus based on the short-term memory of the previous statistical sensory stream [Näätänen et al., 2001]. Such sensitivity to contextual information is termed true deviance detection/sensitivity. A common test for true deviance detection uses many-standards control protocol, where besides the standard-deviant stimuli pair in the oddball sequence, many other different stimuli are presented as the standards [Jacobsen and Schröger, 2001]. In this control, each of these distinct stimuli occurs with the same probability as the deviant stimulus, which consequently eliminates any potential expectation for the next stimulus but preserves the rarity of the deviant. If the response to the deviant stimulus in the oddball sequence exceeds the response to the same deviant one embedded in the control sequence, then the existence of true deviance detection is confirmed [Jacobsen and Schröger, 2001]. Here in the many-standards condition for the somatosensory domain, the standard stimuli are equiprobably distributed over three whiskers in a row that are adjacent to the deviant whisker, therefore each whisker is deflected with 25% probability [Musall et al., 2017] (Fig 4.4A).

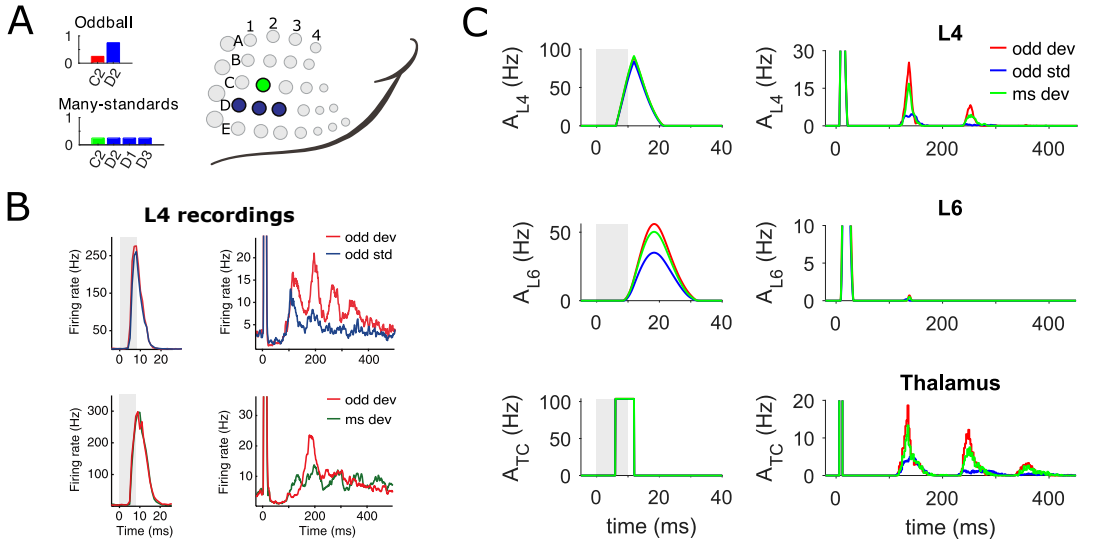

Figure 4.4: **Sensitivity to deviance in silico vs in vivo. (A)** Two protocols are used to test context-dependent deviance detection. In the whisker oddball condition, deviant stimuli are applied to the C2 whisker with 25% probability of occurrence, while standard stimuli are applied to the D2 whisker with 75% probability. In the many-standards condition, the deviant stimulus is presented with the same probability of appearance as in the oddball condition, but standard stimuli are equiprobably distributed over three whiskers D1-D3, each of which is stimulated 25% of the time. The bar height represents the probability of a specific whisker being stimulated. **(B)** A subset of multi-unit (MU) recordings in the granular layer of rat S1 shows late deviant-specific responses in the whisker oddball and many-standards protocols. Average early spike (top row, left) and late deviant-specific oscillatory responses (top row, right) are shown for standard (blue curves, activity recorded in D2) and deviant (red, activity recorded in C2) stimuli in oddball condition. The late response (bottom row) also demonstrates a distinct difference between two deviant types (both responses recorded in C2) in oddball (red curves) and many-standards protocols (green) respectively. Data are adapted from Musall et al. [2017]. **(C)** Average biphasic population activity in cortical L4 (top row), L6 (middle) and thalamus (bottom) in response to oddball standard (blue traces, recorded in D2), oddball deviant (red, recorded in C2) and many-standards deviant (green, recorded in D2) stimuli. Similar to (B), the left and right columns of panels show the early and late activities respectively across cortical layers and the thalamus. SSA and true deviance detection are initiated in the early L6 responses and subsequently enhanced in the late thalamic oscillation, which finally induces the secondary deviant-selective cortical response in L4. The third cycle of thalamic activity is too weak to produce a response in L4 since this thalamic input is below the rheobase (current threshold of gain function) of L4 neurons. Stimulation periods are marked in grey shading. Data in each protocol are averages of 120 stimulus presentations. In both stimulation paradigms, peripheral stimulus duration is 10 ms and inter-stimulus interval (ISI) is 1 s.

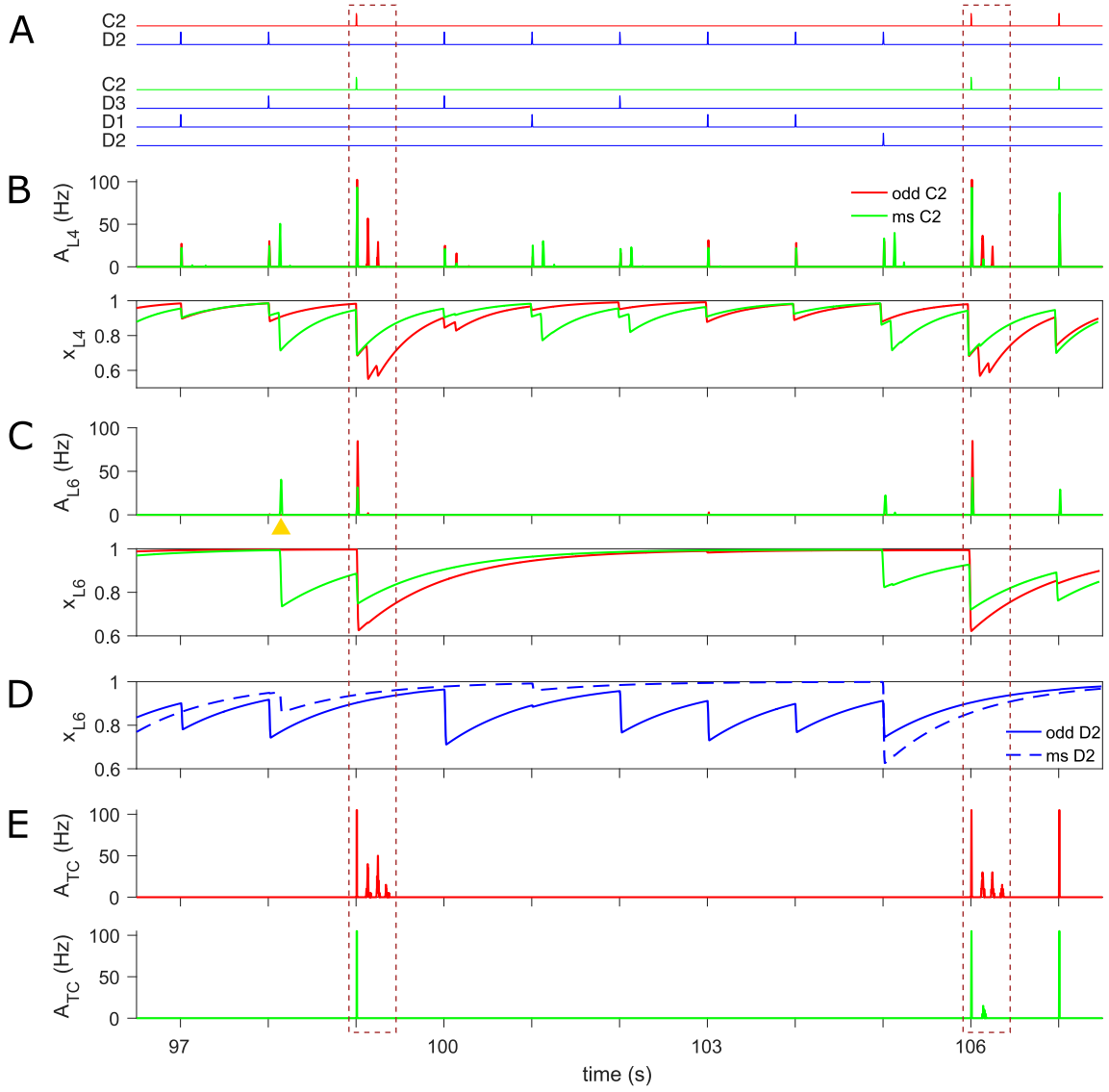

Figure 4.5: **True deviance detection in the model network. (A)** Illustration of oddball (top) and many-standards (bottom) protocols. The deviant stimulation to C2 is highlighted in red for the oddball condition and in green for the many-standards one. **(B)** Population activity (top) and synaptic resources (bottom) in L4 of C2 in response to deviant stimuli presented in oddball (red traces) and many-standards (green) condition respectively. The deviant in both protocols often evokes comparable onset responses, while a robust late oscillation is generally only induced by the oddball deviant (two typical cases are highlighted in dashed boxes). **(C)** Same as (B) but displayed for L6 of C2. Considerably larger early onset responses to the many-standards deviant compared to the oddball deviant are generated in the L6 population (see dashed boxes), because higher activity is propagated from the L6 standard column(s) into that of the deviant column in the many-standards compared to the oddball condition, and accordingly less synaptic resource is left and a weaker response is evoked upon arrival of the deviant. **(D)** Synaptic resources in L6 of D2 in response to the standard stimulus applied to D2 whisker embedded in oddball (solid traces) and many-standards (dashed) sequences.

The standard stimulus to D2 in the many-standards condition, with its lower probability of appearance than the same stimulus in the oddball condition, allows the resource in L6 of D2 to have more time to recover fully before the next standard stimulus arrives and consequently to generate an adequate PS capable of spreading into the same layer of the deviant column. **(E)** Thalamocortical (TC) population activity in C2 barreloid evoked by the deviant in the oddball (red) and many-standards (green) protocols. The corticothalamic excitation of the deviant in the many-standards condition is normally strong enough to elicit late intermittent bursting, whereas that of the deviant in the oddball condition is not (dashed boxes). The time axes are aligned for panels A-E. In both stimulation paradigms, peripheral stimulus duration is 10 ms and inter-stimulus interval (ISI) is 1 s (onset-to-onset).

Fig 4.4C illustrates the distinctive biphasic activity of C2 populations evoked by the oddball and many-standards deviant as well as of D2 populations by the oddball standard whisker deflection in the cortical L4, L6 and the thalamus. For comparison, Fig 4.4B replots from Musall et al. [2017] two different sets of cortical multi-unit (MU) recordings that respectively exhibit late SSA (top panel) and true deviance sensitivity (bottom panel) tested in the oddball and many-standards conditions. The late responses are only found in a subset of recordings mainly confined to the L4. In line with the experimental findings, the simulated L4 populations elicit nearly equally strong population spikes at short latency to both the oddball standard and deviant, as well as the many-standards deviant stimulation, while SSA and true deviance detection are only observed in the late oscillatory responses (Fig 4.4C, top row). Notably, the timescales of the L4 early and late activity as well as the amplitude of the late response in our model are quantitatively in accordance with the experimentally recorded L4 activity. Apart from the close match with the existing experimental data, our model makes further predictions of the population activity in the cortical L6 and the thalamocortical populations in both somatosensory oddball and control protocols, which, to the best of our knowledge, has not been conducted experimentally yet. We hypothesize that the late context-dependent deviant oscillation seen in L4 is reflected by the thalamic spindle oscillation that can be solely generated by the specific thalamic circuit (Fig 4.4C, right panel of bottom row). The differences in the secondary thalamic responses to each type of stimulus are driven by the early cortical PS variations in the L6 circuit, which independently detects the deviant in its own responses (Fig 4.4C, left panel of middle row).

To understand why the average early response of the L6 network for the oddball deviant is larger than for the many-standards deviant, we compared the dynamics of deviant responses in L6 across the two conditions. In the regular oddball sequence, the inadequate L6 activity in D2 evoked by the repetitive standard stimulus hardly propagates into C2, leading to negligible L6 resource depletion in C2 and little cross-whisker adaptation (refer to Fig 4.3B, bottom). In contrast, in the many-standards sequence, the standard stimuli generate substantial activity in their corresponding clusters that provide C2 with strong inter-column inputs, because of the equiprobable distribution of the standard stimuli and accordingly longer recovery time of resources between repeated stimulation. This additional lateral input generally results in higher activity in C2 evoked by the standard stimulation in the many-standards condition than the oddball condition, and consequently leads to more adaptation. To illustrate this mechanism for true deviance detection across the hierarchical network, Fig 4.5A-4.5E displays

population activities and synaptic resources for the C2 cluster in both protocols across the L4, L6, and thalamocortical populations respectively. The dashed brown frames highlight two examples for an oddball deviant triggering a significant late response in L4, while a many-standards deviant does not. In the first frame, the thalamocortical input as a deviant in both protocols triggers approximately the same onset PS in L4 of C2 (Fig 4.5B, top, first spike within the box). In contrast, less L6 activity is elicited in the many-standards than the oddball condition owing to the lower level of L6 resources reduced by the last D3 stimulation in the many-standards condition (Fig 4.5C). In consequence, only the corticothalamic feedback in the oddball protocol is large enough to trigger spindle oscillations in TC cells (Fig 4.5E), which in turn cause synchronized activity in the cortical L4 barrel (Fig 4.5B, top). In a similar vein, the second frame shows an example where resources in L6 of C2 in the many-standards condition are freshly exhausted by the preceding D2 stimulation and do not have time to recover adequately to trigger a spindle oscillation. It is also worth noting that under cross-whisker adaptation, the resources in L6 of C2 are occasionally depleted by bursting activity at long latency propagated by the first robust burst of late oscillation evoked in L4 of standard column(s) (Fig 4.5C, an instance is highlighted by golden pyramid).

4.1.4 Further Experimental Predictions

We ran another pair of novel prediction paradigms on our model with the same set of parameters used in all previous protocols to test the model's sensitivity to contextual information that includes sequential statistics. Here we used variants of the many-standards paradigm with fixed deviant position but permutated the order of the standards presentations to be either random or periodic. In the sequenced paradigm, the standard deflection was periodically applied to C2, D1 and B1 whiskers, each of which is rarely substituted for the C1 deviant in order, which is overall presented 3% of the time. Identical statistics were used in the randomized paradigm except that the three standard stimulations are randomized between the successive deviants. A schematic representation of both sequences is displayed in Fig 4.6A.

In comparison with the sequenced condition where the expectation is established by the repeated sequential standards and violated by the deviant, a precise prediction about forthcoming stimuli cannot be formed in the randomized condition. As shown in Fig 4.6B, a slightly stronger L6 response is initiated for the deviant embedded in a sequenced, expectation-forming than random stimulation train. The response discrepancy is also reflected in L4 with latency by the thalamocortical loop. We quantify this effect by computing the context-specificity index (CSI) defined as the difference between responses to sequenced and randomized deviants normalized by the sum of two responses.

Furthermore, our model predicts the dependence of the novelty-predicting effect on the inter-stimulus interval (ISI) across different cortical layers and thalamus (Fig 4.7). Due to the feature of biphasic dynamics in our network, CSI is separately computed for the early and late responses. For the early response, the novelty-predicting effect is demonstrated in the cortex over a wide range of ISI and particularly prominent in L6 that as we expected to be the novelty detector. Early novelty-predicting responses originated in L6 are subsequently revealed in the thalamic oscillation and secondary L4 responses via the L6-to-thalamus-to-L4 pathway.

To eliminate the potential effect of sensory input uncertainties on CSI, we removed the

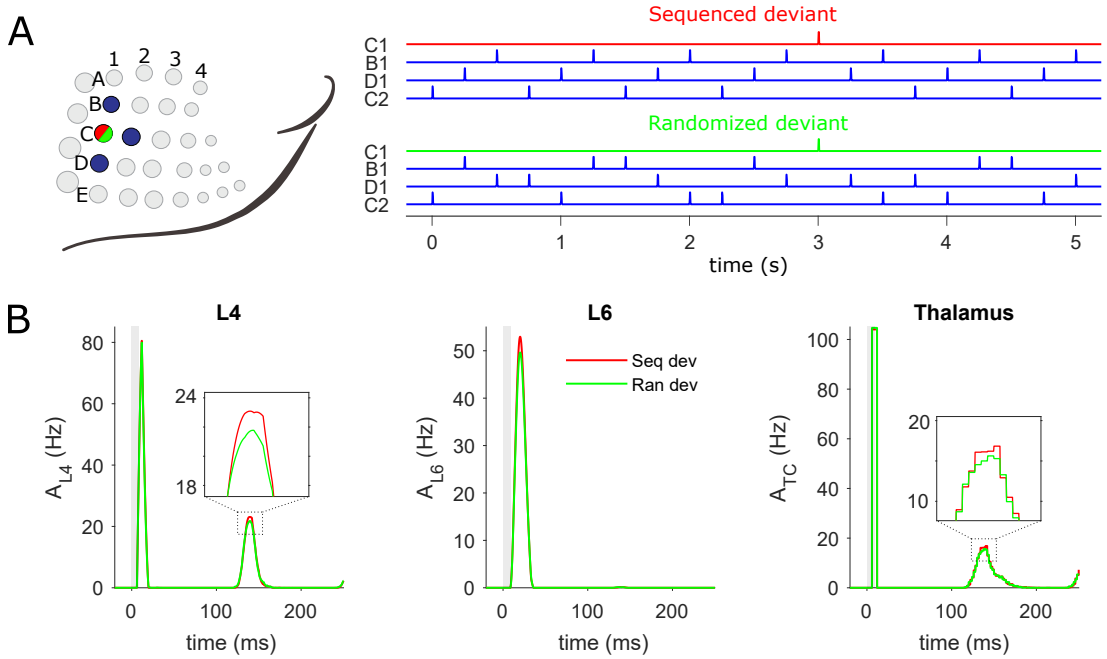

Figure 4.6: **Deviant responses of cortical and thalamic populations in sequenced and randomized paradigms. (A)** Schematic illustration of sequenced and randomized stimulation trains and the stimulated whiskers. In two sequences, the delivery of the deviant stimulus to the C1 whisker is the same, but three standard stimuli are employed either periodically (sequenced condition, top) or randomly (randomized condition, down) to C2, D1 and B1 whiskers. In the sequenced stimulus train, the deviant occasionally breaks the regularity by repeatedly taking the place of one of the standards in the order of C2, D1 and B1. In both paradigms, stimulus duration = 10 ms and inter-stimulus interval (ISI) = 250 ms (stimulus offset to onset). Deviant stimuli constitute 3% of overall stimuli (30 out of 1000 stimuli). **(B)** Average population activity across cortical L4 (left), L6 (middle) and thalamus (right) in response to the deviant in sequenced (red) and randomized (green) paradigms. A weakly differential response to sequenced and randomized deviant exclusively appears in the L6 population at short latency. Owing to the thalamocortical loop, this subtle difference is successively manifested in late thalamic and cortical L4 bursting activities, which are magnified in insets for clarity. Stimulus duration is highlighted in grey. CSI values are 0.0115, 0.0254 and -0.0061 for early L6 (computed over [0, 40] ms) and late L4 and thalamic responses (computed over [40, 250] ms), respectively.

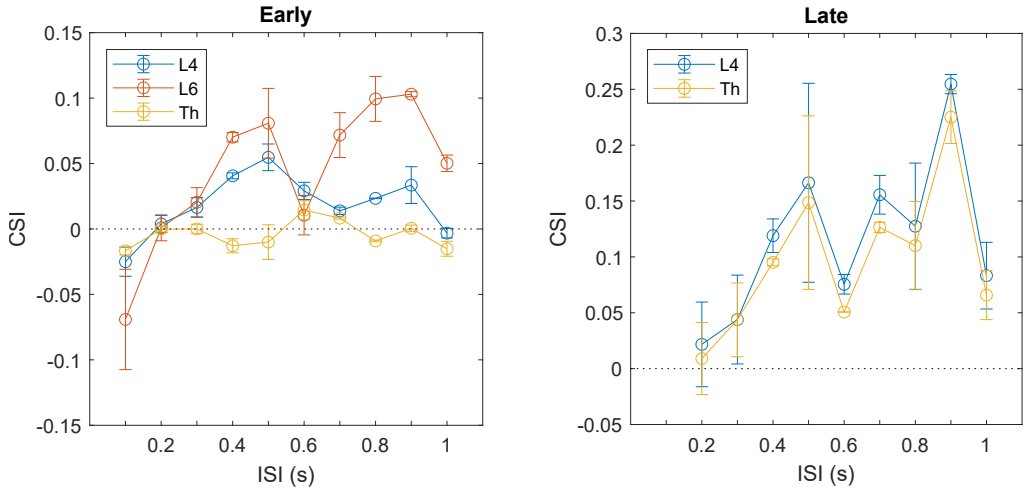

Figure 4.7: **Novelty-predicting effect as a function of inter-stimulus interval.** Early and late CSIs are respectively computed over 40 ms after stimulus onset and the remaining time before the onset of the next stimulus. Data are averaged over 5 trials of each paradigm consisting of 1000 stimuli.

stimulus-by-stimulus variation of the thalamocortical drive that was previously presented and reran both protocols on the model. In this case, each external stimulus evokes the same amplitude of thalamic responses. One can clearly see from Fig 4.8A that the novelty-predicting effect is greatly attenuated near both ends of the range allowing for it because either of the cases will lead to the steady-state of unrecovered/recovered synaptic resources in L6, which are independent of the presentation context. We also investigate the individual role of three types of deviant in the sequenced paradigm in detecting novelty across different ISIs, where the C1 deviant preceded by the B1, C2 and D1 standard is defined as type 1, 2 and 3, respectively (Fig 4.8B). The novelty-predicting effect is the most pronounced for the type 2 deviant and is not influenced much by the increase in ISI after the first several values, which indicates the capacity of novelty prediction in the cortex depends on the specific sequential context in the recent past in randomized paradigms, i.e., the precise information about the order of standard presentations over hundreds of milliseconds before the arrival of deviant.

4.2 Discussion

We presented a cortico-thalamo-cortical circuit of somatosensory processing with recurrent dynamics and depressing synapses that is capable of producing both SSA and the contextual sensitivity over different latencies across cortical layers and thalamus, in line with previous experimental observations [Musall et al., 2017]. The computational results suggest that the early context-specific deviant responses in S1 L6 could be a strong candidate for the electrophysiological correlates of the mid-latency response observed in EEG recordings [Shiramatsu and Takahashi, 2021; Grimm et al., 2011], because of their similar response latencies and shared sensitivity to presentation context. Besides, we show that the late context-dependent responses in S1 L4, which temporally and functionally resemble MMN, can be produced

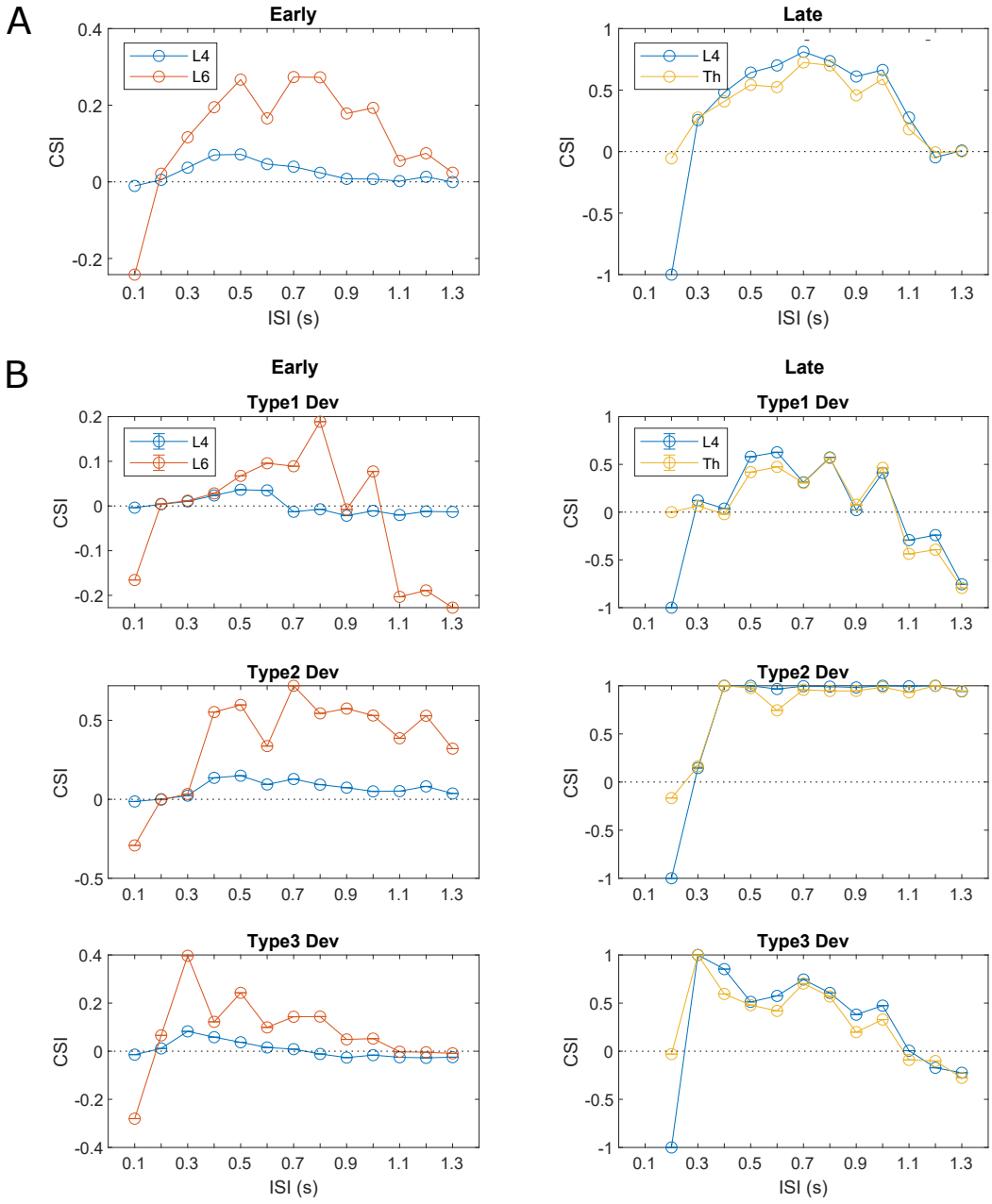

Figure 4.8: **Novelty-predicting effect as a function of inter-stimulus interval in the model with deterministic sensory inputs. (A)** Early and late CSIs are computed in the same fashion as Fig 4.7 except that data are averaged over 30 deviant out of 1000 stimuli in each paradigm. The CSI values for the early thalamic responses are ignored since they are always zero independent of ISI. **(B)** Three types of deviants show different dependence of novelty predicting capacity on ISI. Type 2 deviant has a wider ISI range allowing novelty prediction in both early and late responses than the other two types.

simply by the interaction between different cortical layers and the thalamus, without involving top-down modulation by higher brain areas.

Inspired by a previously described mechanism producing SSA through interactions among intracortical populations with different specificity for stimulus features in the auditory cortex [Yarden and Nelken, 2017], our model enables the generation of a cortical population spike upon the presentation of stimuli and its lateral propagation across different barrels to mediate SSA, consistent with the given neurobiology of the barrel cortex. The existence of whisker-evoked population spikes and their spatiotemporal distribution in S1 are inferred from voltage-sensitive dye imaging [Hubatz et al., 2020] and multielectrode array recordings [Petersen and Diamond, 2000a; Lebedev et al., 2000], where the mean firing rates show a temporary peak in the principal column upon the onset of stimulation, followed by weaker activity with short latency in the surrounding columns [Lebedev et al., 2000]. Previous modelling work has demonstrated that this population burst could be a consequence of a recurrent network with synaptic depression [Tsodyks et al., 1998, 2000; Loebel and Tsodyks, 2002; Barak and Tsodyks, 2007b]. Here, such a network is the essential ingredient of our hierarchical model to produce SSA.

Our model can be regarded as a simplified representation of the cortico-thalamo-cortical pathway of the rodent whisker system. The dynamics of the barrel cortex are described by population activity and mean synaptic resources of somatotopic neuronal clusters in a two-layered grid network. The reasons for adopting rate models for the cortical dynmaics are twofold. First, the mean-field description of population activity makes it possible to analyze the dynamics of single recurrent clusters giving rise to population spikes on the phase plane, under the mathematical abstraction that each cortical cluster is a large and homogenous network. Theoretical work has demonstrated that such a mean-field representation is equivalent to a homogeneous pool of leaky integrate-and-fire neurons in terms of its population activities [Gerstner et al., 2014]. Second, this specification allows us to focus on the collective population behaviour within the grid network, which is sufficient to generate cortical SSA and true deviance detection. SSA and true deviance sensitivity cannot be accounted for solely by the intrinsic dynamics of individual neurons, but relies to a large extent on network dynamics. It should be noted that although this mean-field network of the barrel cortex ignores some details, such as the potential contribution of additional layers, distinct neuron types and their morphology, and more realistic connectivity profiles, critical features of the barrel cortex including its somatotopic columnar arrangement and the spatiotemporal pattern of whisker-evoked responses are retained and indeed shown to be sufficient to explain the deviant selectivity of transient L6 responses.

For the thalamic contribution, following previous work by Destexhe et al. [1998], a spiking neural circuit of single-compartment thalamocortical (TC) and thalamic reticular (RE) cells are used to reproduce spindle oscillation. Synaptic currents within the network are simulated by the synaptic kinetics of AMPA and GABA receptors. In contrast to the cortical rate model, which only depends on network mechanisms, the thalamic spiking model takes into account both the intrinsic properties of single cells and their interplay to elicit synchronous oscillations. Specifically, the spiking dynamics of each thalamic neuron is simulated by an Izhikevich model that reproduces similar firing patterns to the conductance-based model used in Destexhe et al. [1998], but is computationally as efficient as the leaky integrate-and-fire model. Our modelling results implicate the slow recovery variable in these thalamocortical neurons, which may describe the inactivation of low-threshold Ca^{2+} T-currents [Izhikevich,

2007], as the main factor driving the latency of the secondary response observed in S1 L4 and therefore potentially MMN.

Finally, it is worth mentioning that the current implementation of our model still has several shortcomings. A major limitation is the sensitivity of SSA to cortical network parameters, which stems from the different layer-specific response patterns in the feedforward L4-to-L6 network. In other words, the parameter regime allowing for SSA and the true deviance sensitivity in initial L6 responses is constrained by not only the spatiotemporal pattern of whisker-evoked transient activity in the L6 network but also responses in L4 that directly delivers its output to L6. Primarily, in the L6 circuit, the PS evoked by the oddball standard should reduce with repeated stimuli while not generalize to the oddball deviant. Besides, the propagation of activity across the L6 network is supposed to be robust enough to occasionally pass the PSs triggered by standards in the many-standards condition into the deviant column, yet too weak for PSs evoked by the oddball standard to do so. Another possible shortcoming of the model is that it omitted cortical layer 2/3 and layer 5 that could modulate the laminar pathway from L4 to L6. Future work incorporating these missing layers could generate richer dynamics that expand the narrow region of SSA existence in the parameter space, leading to a robust SSA phenomenon. Lastly, it would be instructive to test the NMDA dependency of the late response in a spiking implementation of the cortical dynamics explored here to further investigate SSA's connection with MMN.

4.3 Materials and Methods

It has been observed that a cortico-thalamo-cortical loop structure exists in the rodent whisker system. The excitatory thalamocortical (TC) neurons in the ventral posterior medial (VPM) nucleus present feedforward projection to layer 4 (L4) of the cortex. On the other hand, neurons in layer 6 (L6) provide feedback excitation to TC and reticular(RE) neurons that impose reciprocal inhibition on TC cells.

4.3.1 Modeling the Barrel Cortex

Here, for the complexity needed for the work, we greatly simplified the laminar structure of the cortex by only considering two cascaded excitatory networks that model L4 and L6 respectively.

To mimic the somatotopic organization of the barrel cortex, the circuit of each layer is composed of an array of cortical columns (barrels). Each column is modelled by a fully connected excitatory population mediated by synaptic depression, with its mean-field dynamics as analyzed in the last subsection. Every column is also connected to its nearest vertical and horizontal neighbours by inter-column depressing synapses. Furthermore, the L4 columns receive thalamic inputs with cross-whisker tuning and send outputs to their aligned L6 columns.

The mean-field dynamics of cortical layer 4 network is described by the following equations

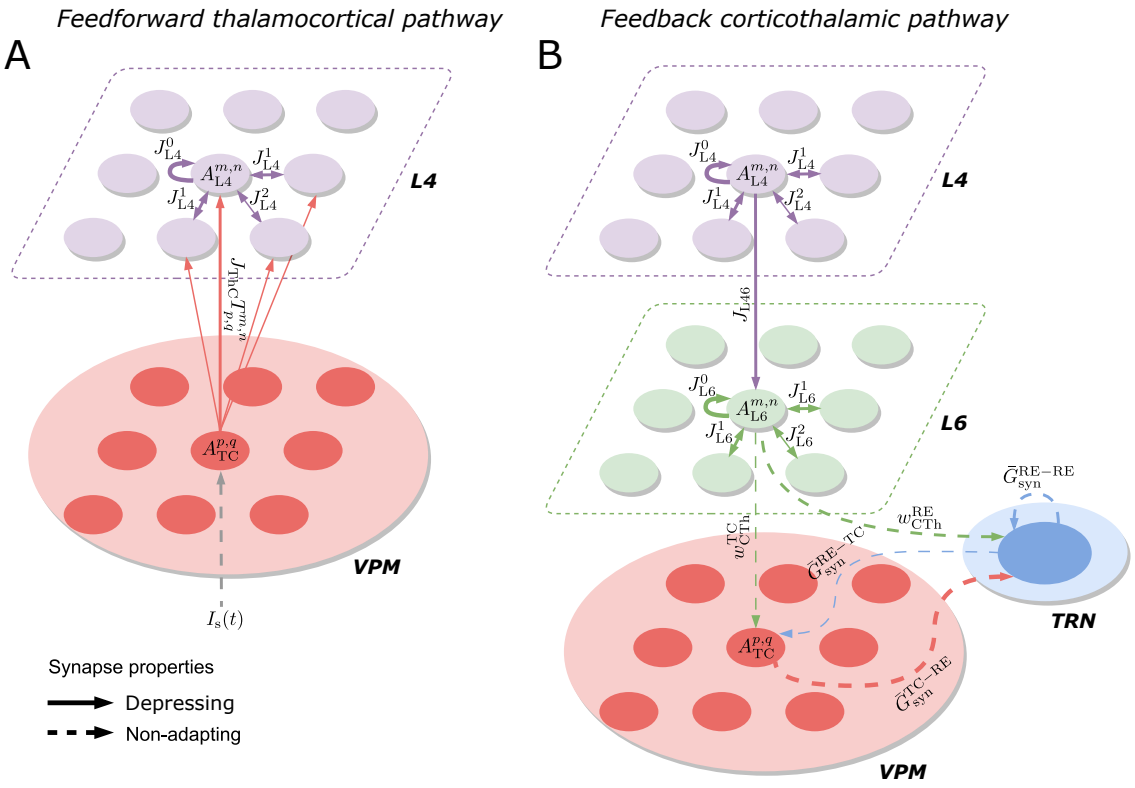

Figure 4.9: **Schematic representation of the thalamo-cortio-thalamic loop model of the barrel cortex, to illustrate the notation used in Methods. (A)** Feedforward thalamocortical pathway. **(B)** Feedback corticothalamic pathway. The width of arrows describes the relative strength of each connection.

(see Fig 4.9 for a schematic representation of the network architecture):

$$\tau_m \frac{\mathrm{d}h_{\mathrm{L4}}^{m,n}}{\mathrm{d}t} = -h_{\mathrm{L4}}^{m,n} + \sum_{k=-1}^{1}\sum_{l=-1}^{1} J_{\mathrm{L4}}^{|k|+|l|} U_{\mathrm{L4}} x_{\mathrm{L4}}^{m+k,n+l} A_{\mathrm{L4}}^{m+k,n+l} + \sum_{p=1}^{M}\sum_{q=1}^{N} J_{\mathrm{ThC}} T_{p,q}^{m,n} U_{\mathrm{ThC}} z_{p,q}^{m,n} A_{\mathrm{TC}}^{p,q} \tag{4.1}$$

$$\frac{\mathrm{d}x_{\mathrm{L4}}^{m,n}}{\mathrm{d}t} = \frac{1 - x_{\mathrm{L4}}^{m,n}}{\tau_{\mathrm{rec,L4}}} - U_{\mathrm{L4}} x_{\mathrm{L4}}^{m,n} A_{\mathrm{L4}}^{m,n} \tag{4.2}$$

$$\frac{\mathrm{d}z_{p,q}^{m,n}}{\mathrm{d}t} = \frac{1 - z_{p,q}^{m,n}}{\tau_{\mathrm{rec,ThC}}} - T_{p,q}^{m,n} U_{\mathrm{ThC}} z_{p,q}^{m,n} A_{\mathrm{TC}}^{p,q} \tag{4.3}$$

$$A_{\mathrm{L4}}^{m,n} = [\alpha_{\mathrm{L4}}(h_{\mathrm{L4}}^{m,n} - \theta_{\mathrm{L4}})]_+ \tag{4.4}$$

Similarly, the mean-field cortical layer 6 network is defined as:

$$\tau_m \frac{\mathrm{d}h_{\mathrm{L6}}^{m,n}}{\mathrm{d}t} = -h_{\mathrm{L6}}^{m,n} + \sum_{k=-1}^{1}\sum_{l=-1}^{1} J_{\mathrm{L6}}^{|k|+|l|} U_{\mathrm{L6}} x_{\mathrm{L6}}^{m+k,n+l} A_{\mathrm{L6}}^{m+k,n+l} + J_{\mathrm{L46}} U_{\mathrm{L46}} x_{\mathrm{L46}}^{m,n} A_{\mathrm{L4}}^{m,n} \tag{4.5}$$

$$\frac{\mathrm{d}x_{\mathrm{L6}}^{m,n}}{\mathrm{d}t} = \frac{1 - x_{\mathrm{L6}}^{m,n}}{\tau_{\mathrm{rec,L6}}} - U_{\mathrm{L6}} x_{\mathrm{L6}}^{m,n} A_{\mathrm{L6}}^{m,n} \tag{4.6}$$

$$\frac{\mathrm{d}x_{\mathrm{L46}}^{m,n}}{\mathrm{d}t} = \frac{1 - x_{\mathrm{L46}}^{m,n}}{\tau_{\mathrm{rec,L46}}} - U_{\mathrm{L46}} x_{\mathrm{L46}}^{m,n} A_{\mathrm{L4}}^{m,n} \tag{4.7}$$

$$A_{\mathrm{L6}}^{m,n} = [\alpha_{\mathrm{L6}}(h_{\mathrm{L6}}^{m,n} - \theta_{\mathrm{L6}})]_+ \tag{4.8}$$

The dynamics of each barrel is described by its population activity $A^{m,n}$ and mean fraction of synaptic resource available $x^{m,n}$, where m and n denote the row and arc indices of the barrel on an idealized $M \times N$ grid. Both variables are specified by subscript either L4 for the population in layer 4 or by L6 for layer 6 (which applies for all layer-specific variables). The population activity $A^{m,n}$ is defined as a threshold- linear gain function of the mean-field current $h^{m,n}$ (α and θ are respectively the slope and horizontal intercept of the gain function). The function $[\cdot]_+$ is defined as its argument when the argument is positive, otherwise zero.

The synaptic efficacy is represented by $J^{|k|+|l|}$ with superscript specifying the column-to-column distance of the connection (0, 1 and 2 respectively indicate intra-column, vertical/horizontal and diagonal inter-column connection). U denotes the utilization rate of synaptic resources. τ_m and τ_{rec} are the membrane time constant of cortical neurons and recovery time constant of synaptic resources, repectively. The dynamics of thalamocortical synapses from the whisker channel p, q to the L4 barrel m, n is characterized by the fraction of resources available $z_{p,q}^{m,n}$, parameterized by utilization rate U_{ThC} and recovery time constant $\tau_{\mathrm{rec,ThC}}$. Similarly, the L4-to-L6 depressing synapse of column m, n is described by its resources $x_{\mathrm{L46}}^{m,n}$ with utilization rate U_{L46} and recovery time constant $\tau_{\mathrm{rec,L46}}$. J_{ThC} and J_{L46} denote the efficacy of thalamocortical and laminar synapses, respectively.

$T_{p,q}^{m,n}$ represents the relative magnitude at which the barrel m, n receives thalamocotical input from whisker channel p, q compared with its primary whisker channel. The values of $T_{p,q}^{m,n}$ over all channels on the $M \times N$ grid compose the spatial tuning curve of the thalamocortical inputs to the barrel m, n. We chose a linear tuning curve in the simulation, which is defined as

$$T_{p,q}^{m,n} = [-d/\lambda + 1]_+ \tag{4.9}$$

where λ is the radius of tuning curve and d is the floored Euclidian distance between barrel m, n and whisker channel p, q on the evenly-spaced gird where the separation distance is 1

$$d = \left\lfloor \sqrt{(m-p)^2 + (n-q)^2} \right\rfloor \tag{4.10}$$

where $\lfloor \cdot \rfloor$ denotes the floor function that rounds its argument to the nearest integer less than or equal to that argument.

$A_{\mathrm{TC}}^{p,q}$ describes the thalamocortical input delivered through the whisker channel p, q. It is defined as the population activity of TC cells organized in barreloid m, n, which in practice is evaluated by Eq 2.13 with a time window $\Delta t = 2$ ms. It is worth noting that the time bin of

2 ms is only used to compute the population activity of TC cells from their spikes, and the cortical population activity is obtained by numerically solving those differential equations of the cortical dynamics with an integration time step of 0.1 ms. Specifically, the $A_{\mathrm{TC}}^{p,q}$ is updated every time window (2 ms) and held constant for 20 integration time steps (each 0.1 ms).

4.3.2 Thalamic Circuit Generating Spindle Oscillation

We assume that the late cortical rhythmic activity (roughly 10 Hz) that lies in the frequency range of the spindle oscillation (7-14 Hz) [Timofeev and Bazhenov, 2005] emerges from the thalamo-cortical system. Earlier studies have shown that spindle oscillation is initiated in the thalamus and synchronized over the cortex and its minimal neural substrate lies on the interaction between thalamic RE cells and TC relay cells [Destexhe et al., 1998; Timofeev and Bazhenov, 2005].

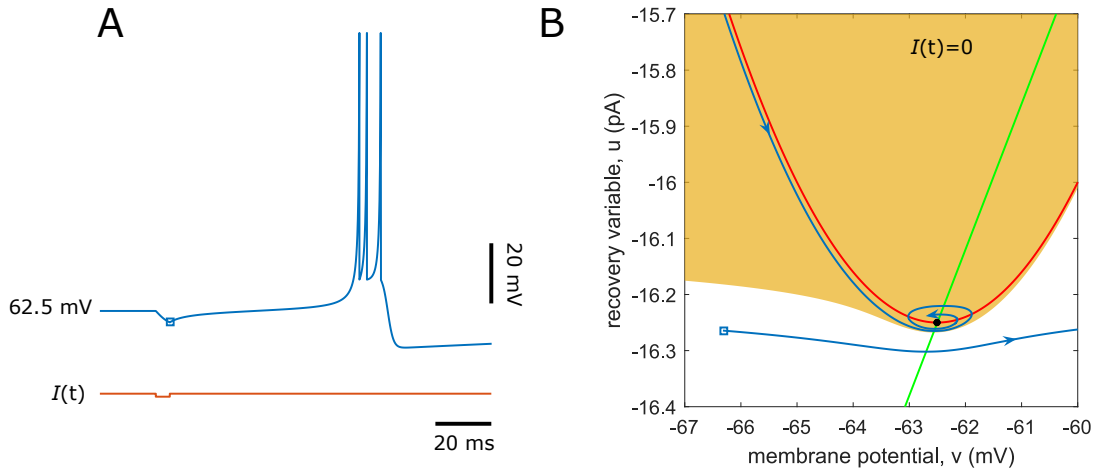

Figure 4.10: **Latency to spike bursts in a TC cell. (A)** A burst of spikes with significant latency is evoked by an inhibitory current pulse with 1 pA amplitude lasting 5 ms that approximates a transient inhibitory postsynaptic current of RE cells modulated by $\mathrm{GABA_A}$ neurotransmitters. The membrane potential value immediately after the release from inhibition is denoted in the open square. **(B)** The phase portrait of the dynamical system (4.11, 4.12) with parameters used in first subpopulation of TC cells when $I(t) = 0$. The intersection of the v-nullcline (red) and u-nullcline (green) is a stable focus, whose attraction basin is marked in orange. The trajectory initiated at the states right after the removal of inhibitory input (open square same as in (A)) slowly goes through the neighborhood of the equilibrium before accelerating towards the voltage threshold of the spike peak. After setting off several spikes, the states are reset to the attraction domain and approach along the v-nuclline to the equilibrium (resting potential).

Motivated by the anatomy of the VPM and reticular nucleus of the thalamus, we modelled each barreloid in the VPM nucleus by a cluster of 100 TC cells mediated by a pool of 100 RE cells, whose connectivity is illustrated in Fig 4.9B. The spiking dynamics of each thalamic neuron simulated by the Izhikevich model, which uses the quadratic integrated-and-fire model

for the membrane potential equation, is expressed as

$$\begin{aligned} \dot{v} &= 0.04v^2 + 5v + 140 - u + I(t) \\ \dot{u} &= a(bv - u) \end{aligned} \tag{4.11}$$

with after-spike resetting

$$\text{if } v \geq +30\text{mV}, \quad \text{then } \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \tag{4.12}$$

where the time t is in milliseconds; v and u respectively denote the membrane potential (in mV) and recovery variable that represents the difference of all influx and efflux of voltage-gated ionic currents (in pA). In particular, the population of 100 TC cells in each barreloid are divided into two subgroups. $(a, b, c, d) = (0.005, 0.26, -52, 2)$ are assigned to the first 60 TC neurons (indexed from 1 to 60) to induce rebound bursts that is an essential ingredient of spindle oscillation. The same parameter quartet except for $b = 0.25$ is assigned to the rest 40 TC neurons (indexed from 60 to 100) that only fire spike trains following stimulation from depolarized currents but fail to produce rebound spikes. Thalamic RE cells are modelled as bursting neurons by assigning $(a, b, c, d) = (0.02, 0.2, -55, 4)$.

As explained before, the latency to the secondary cortical response could be attributed to the delay of rebound bursting in TC cells elicited by inhibitory inputs from RE cells (see burt). Such a delay usually results from the slow inactivation of low-threshold Ca^{2+} T-currents in the TC neurons [Izhikevich, 2007]. In Fig 4.10, we illustrate the delay mechanism of TC cells from the perspective of dynamical systems. Significant delay to rebound bursts arises due to the slow passage of neuronal states across the vicinity of an attractor before setting off a spike. As depicted in Fig 4.10B, a brief, weak negative current pulse sets the neuronal states to a new location on the left side of the neighborhood of an attractor but outside of its attraction basin, where values of the vector field are quite small. The membrane potential increases slowly and steadily to pass that region and fire a burst of spikes before entering the attraction domain. Hence the spike burst is induced with a long latency. It is worth noting that in Fig 4.1F, TC_1 cell does not fire a spike in the first cycle as TC_2 does because the initial excitatory corticothalamic drive neutralizes the effect of inhibition from RE_1 cell and accordingly, the innate states of TC_1 cell model remain within the attraction domain after the initial bursting of RE_1 cell.

The thalamic network connectivity is illustrated in Fig 4.9B, with a connection probability of 0.6 for TC-to-RE, RE-to-TC and RE-to-RE coupling in each pair of subgroups of TC and RE cells and 0.2 for coupling between two subgroups of RE cells. $I(t)$ is the total current input to the the neuron (in pA)

$$I(t) = I_{\text{noise}}(t) - I_{\text{syn}}(t) + I_{\text{CTh}}(t) + I_{\text{s}}(t) \tag{4.13}$$

The thalamic background noise $I_{\text{noise}}(t)$ is subject to uniform distribution bound between -0.5 and 0. The conductance-based synaptic current is modelled as

$$I_{\text{syn}}(t) = g_{\text{syn}}(t)\,(v(t) - E_{\text{syn}}) \tag{4.14}$$

where E_{syn} is the synaptic reversal potential. Synaptic conductivity $g_{\text{syn}}(t)$ has a time course of exponential decay:

$$g_{\text{syn}}(t) = \sum_f \bar{g}_{\text{syn}} e^{-\left(t - t^f\right)/\tau} \Theta(t - t^f) \tag{4.15}$$

Table 4.1: Values used for the network parameters

Notation	Description	Value
M	Number of rows	5
N	Number of arcs	4
λ	Radius of tuning curve	1.6
$U_{\mathrm{L4}}/U_{\mathrm{L6}}$	Utilization rate of L4 / L6 synapses	0.5
$U_{\mathrm{ThC}}/U_{\mathrm{L46}}$	Utilization rate of thalamocortical / L4-to-L6 synapses	0.8 / 0.5
$\alpha_{\mathrm{L4}}/\alpha_{\mathrm{L6}}$	Slope of L4 / L6 gain function	1
$\theta_{\mathrm{L4}}/\theta_{\mathrm{L6}}$	Threshold of L4 / L6 gain function	5 / 3
$J^0_{\mathrm{L4}}/J^0_{\mathrm{L6}}$	Intra-column synaptic efficacy in L4 / L6	2.2 / 2.5
$J^1_{\mathrm{L4}}/J^1_{\mathrm{L6}}$	Vertical/horizontal inter-column synaptic efficacy in L4 / L6	0.05 / 0.03
$J^2_{\mathrm{L4}}/J^2_{\mathrm{L6}}$	Diagonal inter-column synaptic efficacy in L4 / L6	0.001
J_{ThC}	Synaptic efficacy of thalamocortical connection	1 / 0.05
J_{L46}	Synaptic efficacy of L4-to-L6 connection	0.24
τ_m	Membrane time constant	0.001 s
$\tau_{\mathrm{rec,L4}}/\tau_{\mathrm{rec,L6}}$	Recovery time constant of L4 / L6 synaptic resources	0.5 / 1 s
$\tau_{\mathrm{rec,ThC}}/\tau_{\mathrm{rec,L46}}$	Recovery time constant of thalamocortical / L4-to-L6 synaptic resources	0.8 / 1.2 s

where $\Theta(t)$ is the Heaviside step function and t^f represent the arrival time of pre- synaptic spikes. We simulated AMPA and $\mathrm{GABA_A}$ receptors with the decay time constant $\tau = 5$ and 6 ms; $E_{syn} = 0$ and -75 mV, respectively. The maximal conductance $\bar{g}_{\mathrm{syn}}$ for each type of connection is scaled as $\bar{g}_{\mathrm{syn}} = \bar{G}_{\mathrm{syn}}/C$, where $C = 100 \times 0.6 = 60$ is the number of randomly chosen presynaptic partners for each neuron and $\bar{G}_{\mathrm{syn}} = 2$, 0.01 and 0.2 nS for AMPA-mediated TC-to-RE, $\mathrm{GABA_A}$-mediated RE-to-TC and $\mathrm{GABA_A}$-mediated RE-to-RE synapses, respectively.

In addition, half of TC cells in each barreloid and their coupled RE cells were randomly picked receiving corticothalamic feedback current, which is expressed as

$$I_{\mathrm{CTh}}(t) = w_{\mathrm{CTh}} A_{\mathrm{L6}}(t) \tag{4.16}$$

where $A_{\mathrm{L6}}(t)$ denotes the population activity of L6 neurons in the somatotopic infrabarrel. We respectively assigned the corticothalamic coupling strength $w_{\mathrm{CTh}} = 0.001$ and 0.4 for TC and RE neurons.

Finally, we respectively stimulated randomly 10% and 50% of TC cells in the two complementary subgroups of each barreloid with sensory input that is described as

$$I_{\mathrm{s}}(t) = \xi(t) \cdot B \tag{4.17}$$

where B and $\xi_{p,q}(t)$ respectively represent the maximum amplitude and temporal envelope of the sensory stimulus (normalized between 0 and 1). In all simulation protocols, $B = 5$ and $\xi_{p,q}(t)$ has the profile of trapezoid pulse with 10 ms duration (2 ms onset/offset ramp time).

Forward Euler method with a time step of 0.1 ms is used to simulate the network dynamics. The values for different network parameters are listed in table 4.1. The Matlab code for running the network model is available online at `https://github.com/ChaoHan-UoS/SomatosensorySSA_Model`.

4.4 Summary

This chapter described a multi-scale recurrent network with synaptic depression to explain how a delayed deviant-detecting activity [Musall et al., 2017] can arise in the thalamocortical loop of the barrel cortex. The architecture and dynamics of the model presume that neurons in cortical layer 6 adapt, via synaptic depression, specifically to a frequently presented stimulus, resulting in reduced population activity in the corresponding cortical column when compared with the population activity evoked by a rare stimulus. This difference in population activity is then projected from the cortex to the thalamus and amplified through the interaction between neurons of the primary and reticular nuclei of the thalamus, resulting in spindle-like, rhythmic oscillations. These differentially activated thalamic oscillations are forwarded to cortical layer 4 as a late secondary response that is specific to rare stimuli that violate a particular stimulus pattern. Model results show a strong analogy between this late single neuron activity and EEG-based mismatch negativity in terms of their common sensitivity to presentation context and timescales of response latency, as observed experimentally. Our results indicate that adaptation in L6 can establish the thalamocortical dynamics that produce signatures of SSA and MMN and suggest a mechanistic model of novelty detection that could generalize to other sensory modalities.

Chapter 5

Conclusions and Future Work

This chapter recapitulates the research work and contributions delivered in the thesis. Possible ideas for future work are also discussed.

5.1 Thesis Contributions

The principal goal of the thesis is to investigate the neural substrate of the elemental phenomenon of novelty detection. To this end, we presented an exhibitory-inhibitory network moderated by spike-frequency adaptation modified from Yarden and Nelken [2017] to elucidate and replicate the auditory response trends across five SSA protocols adapted from experiments, both heuristically and theoretically (chapter 3). On top of that, we further generalized the SSA network structure of the auditory cortex [Yarden and Nelken, 2017] to the barrel cortex in the context of a thalamocortical loop to explain a deviant-specific somatosensory response with MMN-like latency observed by Musall et al. [2017] (chapter 4). Supplementary materials on the theoretical derivation and dynamical analysis for the mean-field model used in chapter 3 and 4 were provided in the last two sections of Chapter 1. We summarized the main contributions of this PhD research on SSA and deviance detection as below:

Testing an SSA hypothesis in a simplified network of the auditory cortex. In this work, we consolidated by employing a two-layered rate model of A1 the idea originated from Yarden and Nelken [2017] that SSA results from differential adaptation of populations of tuned neurons modulated by the interaction of population activity. Our rate model provided numerical comparisons among responses to different protocols, qualitatively in line with experimental results [Taaseh et al., 2011], and the analytical condition achieving the response trend to further guide the parameter tuning of its spiking counterpart. We identified through this study that, under the SSA-eliciting hypothesis, not only the synaptic depression in the original implementation but neuronal adaptation (e.g., spike-frequency adaptation) could also produce SSA and its pertinent properties. Finally, our network embodiment of the SSA hypothesis ameliorates the sensitivity of SSA to network parameters in the initial model of the hypothesis by demonstrating robustness against online parameter uncertainties and relatively wide operating ranges of free network parameters.

Modeling delayed SSA in the barrel cortex. We proposed a neural circuity mech-

anism to explain how novelty detection emerges in the whisker-related region of the somatosensory cortex (called the barrel cortex) and qualitatively reproduces its corresponding electrophysiological response patterns [Musall et al., 2017], via a somatosensory thalamocortical network with several distinct timescales of dynamics. To the best of our knowledge, our thalamocortical model is the first attempt to account for the underlying network mechanism of the late deviant-specific response recorded on the single-cell level and relate it to EEG-level MMN given their sharing response latency sensitivity to context information. Our results show that the late response in cortical L4 could be inherited from the thalamic spindle oscillation that is only evoked by the unadapted cortical drive from cortical L6 on the presentation of the deviant stimuli. Our hypothesis of late SSA can be further verified by comparing our model predictions with future experimental results.

Finally, it is worth discussing the modeling philosophy taken in the thesis. We started by converting the previous SSA model [Yarden and Nelken, 2017] to its "minimal" form in the sense that we removed the biological details that are not essential for SSA generation and only kept those indispensable elements that could elicit the qualitatively comparable deviance-detecting phenomenon to that in the full model. The minimal model could help us better understand the fundamental mechanisms behind the phenomena on the least level of complexity. Particularly, the use of mean-field representation in the minimal model facilitates the interpretation of the parameter dependence of the model dynamics via the phase plane analysis. The final phenomenological model was built upon the minimal model by augmenting it with new features to reflect the experimental data [Musall et al., 2017] while avoiding unnecessary modeling complexity. Lastly, it should be noted that there has been a long debate on the employment of reductionism vs. holism to model the brain [Mazzocchi, 2012; Pigliucci, 2014]. Prior holism efforts on large-scale spiking model of brain area [Izhikevich and Edelman, 2008; Markram et al., 2015] have gained impressive achievement on reproducing experimental observations. Still, most of them could not identify the explicit intrinsic and circuitry mechanisms giving rise to certain specific phenomena. In this thesis, we supposed that the neural network generating SSA is a non-complex system (absence of emergent behaviors) [Ladyman et al., 2013], and therefore our reductionist methodology remains viable.

5.2 Future Developments

While our hypotheses and models have explained and replicated some of the prominent SSA experimental findings in both the auditory and somatosensory cortex, there are still several open questions that need to be further investigated in the area of deviance detection. Therefore, in this section, we outlined some future work worth pursuing following the research conducted in the thesis.

One major future direction is to explore the possibility of applying a general sequence learning algorithm called hierarchical temporal memory (HTM) [Hawkins and Ahmad, 2016] to explain SSA recordings in different sensory modalities. Although current machine learning algorithms, such as the hidden Markov model (HMM) Fine et al. [1998], long short-term memory (LSTM) Hochreiter and Schmidhuber [1997] and echo state network (ESN) [Jaeger and Haas, 2004], have gained impressive performance on learning sequential data, they can only provide little insight into how the cortex deal with similar sensory stream information

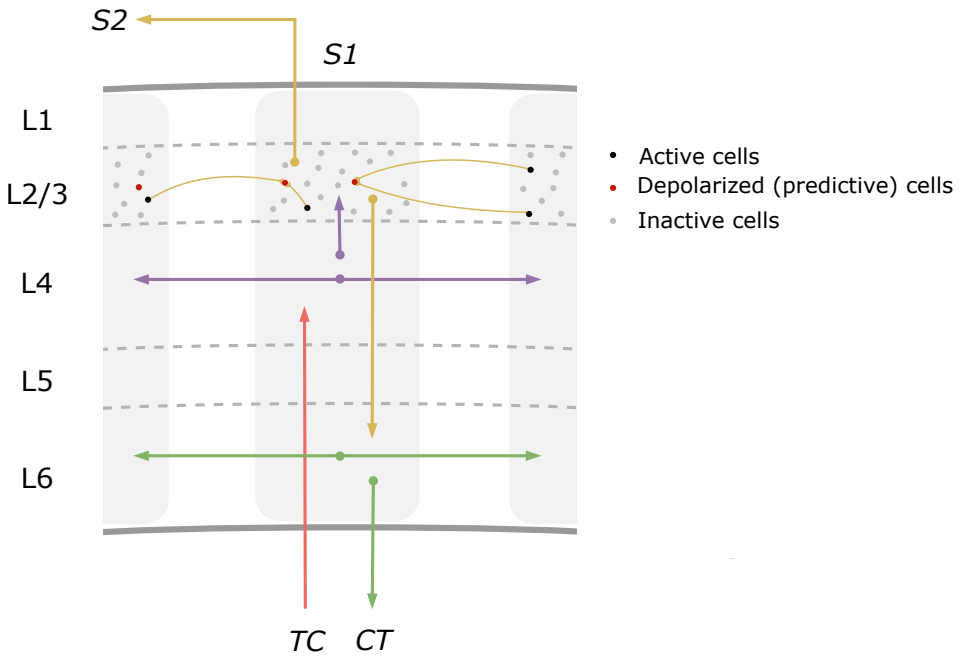

Figure 5.1: **Schema for the merging the HTM layer into our thalamocortical network.** The feedforward intra-column pathway starts from the thalamocortical (TC) cells, to L4, to L2/3 and further the high-order S2 area. Neurons in L2/3 also project back to L6 and subsequently the thalamus via the corticothalamic (CT) pathway. Intra-column connections are constructed in L2/3, L4 and L6. Cells of the HTM in L2/3 are divided into three states: active, predictive and inactive. Cells in the predictive (depolarized) state within a stimulated column will become active and reinforce the subset of intra- and inter-column synapses causing the depolarization, which forms a Hebbian-like learning rule.

due to their independence or oversimplification of neuroscience details. In contrast, HTM is an unsupervised sparse neural network greatly inspired by high-order sensory processing via laminal and columnar structures in the neocortex [Cui et al., 2016]. Since both our SSA models and the HTM model use columnar processing, we can compare and possibly combine each other to check further how closely the model predictions correlate with the SSA behavioral experiments. Specifically, we could add cortical layer 2/3 modeled by a single-layer HTM into our thalamocortical model to alleviate the sensitivity of SSA to model parameters/background noises and accordingly enlarge the difference between responses to the sequenced and randomized sequences discussed in section 4.1.4. The merged model architecture is illustrated in Fig 5.1, where the L2/3 represented by HTM receives the undifferentiated deviant inputs in both paradigms from L4 and yields robust response discrepancy independently. The context-sensitive responses of L2/3 are concurrently delivered downstream to the thalamus via L6 and upstream to the secondary somatosensory cortex (S2) in which pronounced response difference between sequenced and randomized paradigms was also found. By virtue of the thalamocortical loop, the differentiated responses could also be reflected in L4 over a long latency. The motivation for the model extension is that short-term synaptic depression and somatosensory thalamocortical loop seem not enough to show a consistent response tendency

over various inter-stimulus intervals (see Fig 4.7). Therefore some additional mechanisms need to be considered. A promising candidate could be the Hebbian-like synaptic learning rule employed in the HTM model that sparse neurons become active via a reinforced connection motif in response to anticipated stimuli while an unpredicted stimulus can activate all neurons in a column [Hawkins and Ahmad, 2016; Hawkins et al., 2017]. In addition, we can test the abilities of our simplified auditory SSA model (see chapter 3) and the HTM model to account for the sensitivity of auditory neurons to complicated statistical regularities [Yaron et al., 2012] that cannot be explained by any previous SSA models. Experimental evidence suggested a long-term dependence on the history of sensory stimuli. Thus an additional long-term synaptic effect would be plausible to be incorporated into our SSA models. We can also investigate the dependence of our SSA models on the protocol parameters (the probability of deviant appearance and inter-stimulus interval) and compare it with experimental results. Finally, it would be very instructive to study the MNDA-dependence of the early and late sensory responses via both experiments and simulation to verify our hypotheses that the early response in S1 L6 and late response in S1 L4 are respectively the physiological origins of the EEG-based MLR and MMN. A potential modeling direction would be to augment our thalamocortical network by including L2/3 modeled by the HTM model since it uses local NMDA spikes triggered on dendritic branches to activate a neuron [Cui et al., 2016]. An alternative modeling approach can also be converting the cortical rate neurons of our thalamocortical model into their spiking counterparts coupled by MNDA-mediated synapses.

5.3 Conclusions

In this thesis, we supplied further support for the SSA network mechanism by testing auditory responses to different novelty-detecting protocols in a simplified but robust version of Yarden and Nelken [2017]'s model, implemented on both a simulated mean-field minimal model and physical neuromorphic hardware. In addition, a thalamocortical network with a somatotopic arrangement was developed to explain a delayed deviant-sensitive cortical response. We concluded for the simulation results that novel signals could be initially detected in the transient L6 response to sensory stimuli as the fewer synaptic resources in the deviant column were depleted when presented in a regular than random stimuli context. The L4 delayed deviant-specific response was inherited from L6 through the thalamocortical loop, and the latency was contributed by the slow inactivation of specific ion channels in the thalamocortical neurons. Our results suggest a mechanistic model that yields signatures of both SSA and MMN, which could be expanded and generalized to explain and reproduce more experimental findings on novelty detection in future work.

Bibliography

Antunes, F. M. and Malmierca, M. S. (2011). Effect of Auditory Cortex Deactivation on Stimulus-Specific Adaptation in the Medial Geniculate Body. *Journal of Neuroscience*, 31(47):17306–17316.

Antunes, F. M. and Malmierca, M. S. (2014). An overview of stimulus-specific adaptation in the auditory thalamus. *Brain Topography*, 27(4):480–499.

Armstrong-James, M., Fox, K., and Das-Gupta, A. (1992). Flow of excitation within rat barrel cortex on striking a single vibrissa. *Journal of neurophysiology*, 68(4):1345–58.

Barak, O. and Tsodyks, M. (2007a). Persistent activity in neural networks with dynamic synapses. *PLoS Computational Biology*, 3(2):0323–0332.

Barak, O. and Tsodyks, M. (2007b). Persistent activity in neural networks with dynamic synapses. *PLoS Comput Biol*, 3(2):e35.

Barlow, H. B. et al. (1961). Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01).

Bauerle, P., von der Behrens, W., Kossl, M., and Gaese, B. H. (2011). Stimulus-Specific Adaptation in the Gerbil Primary Auditory Thalamus Is the Result of a Fast Frequency-Specific Habituation and Is Regulated by the Corticofugal System. *Journal of Neuroscience*, 31(26):9708–9722.

Benda, J. and Herz, A. V. (2003). A universal model for spike-frequency adaptation. *Neural computation*, 15(11):2523–2564.

Bertram, R., Sherman, A., and Stanley, E. F. (1996). Single-domain/bound calcium hypothesis of transmitter release and facilitation. *Journal of Neurophysiology*, 75(5):1919–1931.

Brecht, M. and Sakmann, B. (2002). Whisker maps of neuronal subclasses of the rat ventral posterior medial thalamus, identified by whole-cell voltage recording and morphological reconstruction. *The Journal of Physiology*, 538(2):495–515.

Brenner, N., Bialek, W., and de Ruyter van Steveninck, R. (2000). Adaptive Rescaling Maximizes Information Transmission. *Neuron*, 26(3):695–702.

Crandall, S. R., Patrick, S. L., Cruikshank, S. J., and Connors, B. W. (2017). Infrabarrels Are Layer 6 Circuit Modules in the Barrel Cortex that Link Long-Range Inputs and Outputs. *Cell Reports*, 21(11):3065–3078.

Cui, Y., Ahmad, S., and Hawkins, J. (2016). Continuous online sequence learning with an unsupervised neural network model. *Neural computation*, 28(11):2474–2504.

Dayan, P. and Abbott, L. F. (2001). *Theoretical neuroscience*, volume 806. Cambridge, MA: MIT Press.

Destexhe, A., Contreras, D., and Steriade, M. (1998). Mechanisms Underlying the Synchronizing Action of Corticothalamic Feedback Through Inhibition of Thalamic Relay Cells. *Journal of Neurophysiology*, 79(2):999–1016.

Esposito, U., Giugliano, M., and Vasilaki, E. (2015). Adaptation of short-term plasticity parameters via error-driven learning may explain the correlation between activity-dependent synaptic properties, connectivity motifs and target specificity. *Frontiers in computational neuroscience*, 8:175.

Fairhall, A. L., Lewen, G. D., Bialek, W., and De Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787–792.

Farley, B. J., Quirk, M. C., Doherty, J. J., and 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. *Journal of Neuroscience*, 30(49):16475–16484.

Feldmeyer, D. (2012). Excitatory neuronal connectivity in the barrel cortex. *Frontiers in neuroanatomy*, 6:24.

Fine, S., Singer, Y., and Tishby, N. (1998). The hierarchical hidden markov model: Analysis and applications. *Machine learning*, 32(1):41–62.

Fitzgerald, K. and Todd, J. (2020). Making Sense of Mismatch Negativity. *Frontiers in Psychiatry*, 11(June):1–19.

Fox, K. D. (2008). *Barrel cortex*. Cambridge University Press.

Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.

Gil, Z., Connors, B. W., and Amitai, Y. (1999). Efficacy of thalamocortical and intracortical synaptic connections: quanta, innervation, and reliability. *Neuron*, 23(2):385–397.

Gjorgjieva, J., Mease, R. a., Moody, W. J., and Fairhall, A. L. (2014). Intrinsic neuronal properties switch the mode of information. *PLoS Computational Biology*, 10(12):1–26.

Grimm, S., Escera, C., Slabu, L., and Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology*, 48(3):377–384.

Gupta, A., Wang, Y., and Markram, H. (2000). Organizing principles for a diversity of gabaergic interneurons and synapses in the neocortex. *Science*, 287(5451):273–278.

Han, C., English, G., Saal, H. P., Indiveri, G., Gilra, A., von der Behrens, W., and Vasilaki, E. (2021). Modelling novelty detection in the thalamocortical loop. *bioRxiv*.

Harms, L., Michie, P. T., and Näätänen, R. (2016). Criteria for determining whether mismatch responses exist in animal models: Focus on rodents. *Biological Psychology*, 116:28–35.

Hawkins, J. and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Frontiers in neural circuits*, 10:23.

Hawkins, J., Ahmad, S., and Cui, Y. (2017). A theory of how columns in the neocortex enable learning the structure of the world. *Frontiers in neural circuits*, 11:81.

Hennig, M. H. (2013). Theoretical models of synaptic short term plasticity. *Frontiers in Computational Neuroscience*, 7(April):1–10.

Hershenhoren, I., Taaseh, N., Antunes, F. M., and Nelken, I. (2014a). Intracellular Correlates of Stimulus-Specific Adaptation. *Journal of Neuroscience*, 34(9):3303–3319.

Hershenhoren, I., Taaseh, N., Antunes, F. M., and Nelken, I. (2014b). Intracellular correlates of stimulus-specific adaptation. *Journal of Neuroscience*, 34(9):3303–3319.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Holmes, G. L. and Khazipov, R. (2007). Basic neurophysiology and the cortical basis of eeg. *The clinical neurophysiology primer*, pages 19–33.

Hubatz, S., Hucher, G., Shulz, D. E., and Férézou, I. (2020). Spatiotemporal properties of whisker-evoked tactile responses in the mouse secondary somatosensory cortex. *Scientific reports*, 10(1):1–11.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience*. MIT press.

Izhikevich, E. M. and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9):3593–3598.

Jacobsen, T. and Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, 38(4):723–727.

Jaeger, H. and Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *science*, 304(5667):78–80.

Katz, Y., Heiss, J. E., and Lampl, I. (2006). Cross-whisker adaptation of neurons in the rat barrel cortex. *Journal of Neuroscience*, 26(51):13363–13372.

Khouri, L. and Nelken, I. (2015). Detecting the unexpected. *Current Opinion in Neurobiology*, 35:142–147.

Kudela, P., Boatman-Reich, D., Beeman, D., and Anderson, W. S. (2018). Modeling Neural Adaptation in Auditory Cortex. *Frontiers in Neural Circuits*, 12(September):1–11.

Ladyman, J., Lambert, J., and Wiesner, K. (2013). What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67.

Lebedev, M. A. (2002). Experience-dependent Plasticity of Rat Barrel Cortex: Redistribution of Activity across Barrel-columns. *Cerebral Cortex*, 10(1):23–31.

Lebedev, M. A., Mirabella, G., Erchova, I., and Diamond, M. E. (2000). Experience-dependent plasticity of rat barrel cortex: redistribution of activity across barrel-columns. *Cerebral Cortex*, 10(1):23–31.

Loebel, A. and Tsodyks, M. (2002). Computation by ensemble synchronization in recurrent networks with synaptic depression. *Journal of computational neuroscience*, 13(2):111–124.

Luczak, A., Barthó, P., and Harris, K. D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron*, 62(3):413–425.

Luczak, A., Bartho, P., and Harris, K. D. (2013). Gating of sensory input by spontaneous cortical activity. *Journal of Neuroscience*, 33(4):1684–1695.

Mallart, A. and Martin, A. (1967). Two components of facilitation at the neuromuscular junction of the frog. *The Journal of physiology*, 191(1):19P.

Malmierca, M. S., Anderson, L. A., and Antunes, F. M. (2015). The cortical modulation of stimulus-specific adaptation in the auditory midbrain and thalamus: a potential neuronal correlate for predictive coding. *Frontiers in Systems Neuroscience*, 9(March):1–14.

Malmierca, M. S., Cristaudo, S., Pérez-González, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *Journal of Neuroscience*, 29(17):5483–5493.

Maravall, M., Alenda, A., Bale, M. R., and Petersen, R. S. (2013). Transformation of adaptation and gain rescaling along the whisker sensory pathway. *PLoS ONE*, 8(12).

Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., and Diamond, M. E. (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biology*, 5(2):0323–0334.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell*, 163(2):456–492.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proceedings of the National Academy of Sciences*, 95(9):5323–5328.

May, P. J. and Tiitinen, H. (2010). Mismatch negativity (mmn), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1):66–122.

May, P. J., Westö, J., and Tiitinen, H. (2015). Computational modelling suggests that temporal integration results from synaptic adaptation in auditory cortex. *European Journal of Neuroscience*, 41(5):615–630.

Mazzocchi, F. (2012). Complexity and the reductionism–holism debate in systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 4(5):413–427.

Mill, R., Coath, M., Wennekers, T., and Denham, S. L. (2011a). A neurocomputational model of stimulus-specific adaptation to oddball and markov sequences. *PLoS computational biology*, 7(8):e1002117.

Mill, R., Coath, M., Wennekers, T., and Denham, S. L. (2012). Characterising stimulus-specific adaptation using a multi-layer field model. *Brain Research*, 1434:178–188.

Mill, R., Sheik, S., Indiveri, G., and Denham, S. L. (2011b). A model of stimulus-specific adaptation in neuromorphic analog vlsi. *IEEE transactions on biomedical circuits and systems*, 5(5):413–419.

Musall, S., Haiss, F., Weber, B., and von der Behrens, W. (2017). Deviant processing in the primary somatosensory cortex. *Cerebral Cortex*, 27(1):863–876.

Näätänen, R., Gaillard, A. W., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4):313–329.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12):2544–2590.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (2001). 'primitive intelligence'in the auditory cortex. *Trends in neurosciences*, 24(5):283–288.

Nelken, I. (2014). Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biological Cybernetics*, 108(5):655–663.

Nelken, I. and Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology*, 21(3-4):214–223.

Oberlaender, M., De Kock, C. P., Bruno, R. M., Ramirez, A., Meyer, H. S., Dercksen, V. J., Helmstaedter, M., and Sakmann, B. (2012). Cell type-specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissal cortex. *Cerebral Cortex*, 22(10):2375–2391.

Pazo-Alvarez, P., Cadaveira, F., and Amenedo, E. (2003). MMN in the visual modality: a review. *Biological psychology*, 63(3):199–236.

Petersen, C. C. (2007). The functional organization of the barrel cortex. *Neuron*, 56(2):339–355.

Petersen, R. S. and Diamond, M. E. (2000a). Spatial–temporal distribution of whisker-evoked activity in rat somatosensory cortex and the coding of stimulus location. *Journal of Neuroscience*, 20(16):6135–6143.

Petersen, R. S. and Diamond, M. E. (2000b). Spatial-temporal distribution of whisker-evoked activity in rat somatosensory cortex and the coding of stimulus location. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(16):6135–43.

Pigliucci, M. (2014). Between holism and reductionism: a philosophical primer on emergence. *Biological Journal of the Linnean Society*, 112(2):261–267.

Polterovich, A., Jankowski, M. M., and Nelken, I. (2018). Deviance sensitivity in the auditory cortex of freely moving rats. *PLOS ONE*, 13(6):e0197678.

Reches, A. and Gutfreund, Y. (2008). Stimulus-Specific Adaptations in the Gaze Control System of the Barn Owl. *Journal of Neuroscience*, 28(6):1523–1533.

Reches, A., Netser, S., and Gutfreund, Y. (2010). Interactions between Stimulus-Specific Adaptation and Visual Auditory Integration in the Forebrain of the Barn Owl. *Journal of Neuroscience*, 30(20):6991–6998.

Sakata, S. and Harris, K. D. (2009). Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. *Neuron*, 64(3):404–418.

Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 62(6):437–448.

Schwartz, O., Pillow, J. W., Rust, N. C., and Simoncelli, E. P. (2006). Spike-triggered neural characterization. *Journal of Vision*, 6(4):13.

Series, P., Stocker, A. A., and Simoncelli, E. P. (2009). Is the homunculus "aware" of sensory adaptation? *Neural computation*, 21(12):3271–3304.

Shiramatsu, T. I., Kanzaki, R., and Takahashi, H. (2013). Cortical mapping of mismatch negativity with deviance detection property in rat. *PLoS ONE*, 8(12):36–44.

Shiramatsu, T. I. and Takahashi, H. (2021). Mismatch-negativity (MMN) in animal models: Homology of human MMN? *Hearing Research*, 399:107936.

Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One*, 6(8):e23369.

Thomson, A. M. and Deuchars, J. (1994). Temporal and spatial properties of local circuits in neocortex. *Trends in neurosciences*, 17(3):119–126.

Timofeev, I. and Bazhenov, M. (2005). Mechanisms and biological role of thalamocortical oscillations. *Trends in chronobiology research*, pages 1–47.

Tivadar, R. I. and Murray, M. M. (2019). A primer on electroencephalography and event-related potentials for organizational neuroscience. *Organizational Research Methods*, 22(1):69–94.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural computation*, 10(4):821–835.

Tsodyks, M., Uziel, A., Markram, H., et al. (2000). Synchrony generation in recurrent networks with frequency-dependent synapses. *J Neurosci*, 20(1):825–835.

Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, 94(2):719–723.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience*, 24(46):10440–10453.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature neuroscience*, 6(4):391.

Umbricht, D., Schmid, L., Koller, R., Vollenweider, F. X., Hell, D., and Javitt, D. C. (2000). Ketamine-Induced Deficits in Auditory and Visual Context-Dependent Processing in Healthy Volunteers. *Archives of General Psychiatry*, 57(12):1139.

Vanattou-Saïfoudine, N., Han, C., Krause, R., Vasilaki, E., von der Behrens, W., and Indiveri, G. (2021). A robust model of Stimulus-Specific Adaptation validated on neuromorphic hardware. *Scientific Reports*, 11(1):1–15.

Vasilaki, E. and Giugliano, M. (2012). Emergence of connectivity patterns from long-term and short-term plasticities. In *International Conference on Artificial Neural Networks*, pages 193–200. Springer.

Vasilaki, E. and Giugliano, M. (2014). Emergence of connectivity motifs in networks of model neurons with short-and long-term plastic synapses. *PloS one*, 9(1):e84626.

von der Behrens, W., Bauerle, P., Kossl, M., and Gaese, B. H. (2009). Correlating Stimulus-Specific Adaptation of Cortical Neurons and Local Field Potentials in the Awake Rat. *Journal of Neuroscience*, 29(44):13837–13849.

Wacongne, C., Changeux, J.-P., and Dehaene, S. (2012). A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. *Journal of Neuroscience*, 32(11):3665–3678.

Wang, P. and Knösche, T. R. (2013). A realistic neural mass model of the cortex with laminar-specific connections and synaptic plasticity - evaluation with auditory habituation. *PloS one*, 8(10):36–39.

Wark, B., Lundstrom, B. N., and Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neurobiology*, 17(4):423–429.

Whitmire, C. J. and Stanley, G. B. (2016). Rapid Sensory Adaptation Redux: A Circuit Perspective. *Neuron*, 92(2):298–315.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24.

Yarden, T. S. and Nelken, I. (2017). Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS computational biology*, 13(3):e1005437.

Yaron, A., Hershenhoren, I., and Nelken, I. (2012). Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron*, 76(3):603–615.

Zengel, J. and Magleby, K. (1982). Augmentation and facilitation of transmitter release. a quantitative description at the frog neuromuscular junction. *The Journal of General Physiology*, 80(4):583–611.