Synaptic plasticity models in brain development and neurodevelopmental disorders



Stephanie Yayoi Teramoto Kimura

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2019

To Megumi Kimura and Tomás Kimura.

I would trade all the work in this thesis for one more day with you, but since that is not possible, may your names be remembered in the Bodleian Libraries for a very long time.

Acknowledgements

Getting my DPhil has been the most challenging thing I have ever done. That just shows how lucky and privileged my life has been. I could not have arrived to where I am today without my family. There were many moments in the last few years that I was ready to give up, but their words of encouragement, wisdom, and love kept me going. To my ojiichan and obaachan *de tía Nori*, thank you for always reminding me of taking care of my health. To my ojiichan and obaachan *de la casa*, thank you for teaching me how to fight and never give up. I wish I had your strength and courage. To my dad for teaching me to always smile, and my mom for never doubting me (and also for her occasional healthy doses of realism). You are the craziest parents I know (who else flies from Mexico to England for just a weekend just to visit their daughter?), and I could have never asked for better parents. To my uncles, aunts, and cousins for loving me, supporting me, and pampering me even though I live so far away. Even though we sometimes only see each other for a single meal a year, I carry you in my heart and I look forward to those brief two hours when I get to see you again. To Coco for keeping my family company while I was in Oxford doing my DPhil.

When I signed up for a DPhil, I had no clue what I was doing, and am grateful to have had Professor Tim Vogels and Professor Simon Butt as my supervisors. I came to Oxford wanting to work at the intersection of theory and experiments, and that is exactly what you both enabled me to do. Thank you for helping me become a better scientist, and trusting me to develop my own projects. A moment that changed my DPhil completely was when I met Dr. Laura Cancedda at the Vogels lab retreat. I am so grateful to Dr. Laura Cancedda for taking a chance and collaborating with me and having endless patience and curiosity with modelling. I am indebted to you and to Dr. Salvatore Magara for your willingness to invite a random Computational Neuroscience student and trust her to take your data and ideas and build models.

One of the great parts of being in two labs during my DPhil was that I felt like I had two families. Thank you to my two academic families for making

Oxford feel like home.

Being part of the Vogels lab was never boring. In my time as a student, we had everything from lunch and coffee breaks, to broken bones and wifi-less scientific retreats. I have learned so much from everyone in the lab. A special thank you to Dr. Everton Agnes, who was my desk neighbour throughout my PhD and has been a constant source of wisdom, constructive criticism, gaucho sayings, and laughter. It made a huge difference to have Jake Stroud, Nas Ahmad, Bill Podlaski, and Bayar Menzat to share the DPhil experience with, and who set an example for me to follow and aspire to both in their work ethic and their passion for science. Many thanks to Dr. Rui Costa, Dr. Friedemann Zenke, Dr. Victor Hernández, Dr. Georgia Christodolou, Dr. Chaitanya Chintaluri, and Panos Panagiotis for always being willing to discuss science with me and supporting me when I needed it.

In the Butt lab, I am indebted to Cristiana Vagnoni and Jacqui Stacey for always finding time in their busy experiment schedules to have lunch with me and teach me about laser scanning photostimulation, lasers, and microscopes (plus the occasional 'girl talk'). To Liad Baruchin, I am grateful for always telling me things were going to be okay and for convincing me to sign up to that 10K race. Many thanks to Dr. Mari Panniello, Ana Barros, Sev Limal, Dr. Rob Lees, Jimmy Rowland, Filippo Ghezzi, Andrew Shelton, Sara Ratti, Dr. Adam Packer, and Dr. Michael Kohl for all the good times in the shared lab space and for graciously inviting me along to all science-related, social events and coffee in the Bukopa space.

I would have never been able to thrive in Oxford without the help of Sinead Wright, Fiona Woods, Sarah Noujaim, and Stacey McGowen, who were always happy to answer my questions and guide me through the paperwork needed to smoothly meet all the requirements to graduate. Dr. Debbie Clarke and Professor Andy King were also incredibly supportive during my time as a student from the MSc all the way through to the end of my DPhil. Furthermore, I could have never even dreamed of doing my DPhil in Oxford without the support of the Wellcome Trust, the Clarendon Fund, and the Foley-Béjar Scholarship.

In addition to a stimulating intellectual Neuroscience community, I could never have survived without my friends and mentors. Thank you to Renato Carregha, Luis Angel Montiel, Lillian Toro, and Andi Terminel, who have known me for over a decade and still want to see me whenever I fly to the other side of the pond. To Ana Laura Soto for being that friend who

just gets me even though we live far away and see each other very rarely. To Héctor Moreno, Randi Brown, Diana Liao, Christian Rivera, Alejandro Arroyo, Jane Wang, and Gene Katsevich who have beaten the odds and come all the way to Oxford to visit me. To Thais Rocha, Karen Haidinger, and Késia Decoté for always treating me as a fellow Brazilian. To Megan, Mitch, and Laurelin who remind me of the importance of home and family. To Professor Philip Holmes, Professor Asif Ghazanfar, Dr. Phil Trinh, and Professor Daniel Takahashi for being my first mentors in science and research. To Dr. Tom Brughmans and Dr. Iza Romanowska for being such an energetic inspirational power couple. To the History of Science Museum for providing me with a space that I could go to unwind, probably learn one or two new things. Especially in the writing up period I could not have stayed sane without Dr. Katrhyn Eccles, Jamie Cameron, and Sarah Griffin. Knowing that I could come into the office and expect to see one or more new objects added to Cabinet gave me incredible joy always. To Karine Yuki, Midori Amano, Davide Folloni, Dr. Anne Pfister, Jessica Popkin, Alexa Phunn, Camargos + Gregs, Is Us Us (and all the variations), the Oxford Latin Combo, BMM friends, Andréa Melloni, Tai Ying Lee, Thomas Leissing, Ferdinando Randisi, Jéssica Marcomini, Iberê Souza, Gustavo Quino, Pablo Astudillo, Ga Ferreira, Ivan Ferreira, Jessica Schulte, Julie Sanchez, Silvana Alberti, and all the people who have brought joy to my life but I have somehow forgotten, for being a part of my life during these last four years and being a source of peace and true happiness in the last few years. Finally, a huge thank you to Chico Camargo for helping me become more me than I ever thought I could be, for all the conversations about anything and everything, and for staying at my side during the happiest and the toughest moments in my DPhil.

Abstract

Mathematical and computational models are powerful tools that can help us gain a deeper understanding about the world around us. In Neuroscience, however, there is still debate as to what the role of models is in research. In this thesis we build models of synaptic plasticity of varying degrees of complexity, and at different levels of abstraction in order to generate experimentally testable hypotheses and propose new ways of thinking about synaptic plasticity in development and in neurodevelopmental disorders. We applied our models to two datasets.

First, we studied a transient translaminar cortical circuit that connects layer 5b (L5b) interneurons with L4 excitatory spiny stellate neurons (SSN), or the L5b-L4 loop, that is present in the mouse barrel cortex during the first two postnatal weeks (Marques-Smith et al., 2016). We used spiking neural networks and rate models with plastic connections to explore the mechanisms that drive these large connectivity changes, as well as the possible role these may play in the developing mouse barrel cortex.

Second, we analysed extracellular plasticity data recorded from Ts65Dn animals, a mouse model for Down Syndrome. We used three abstract models to formalise and test the hypothesis that synapses in Ts65Dn animals are more saturated, *i.e.* the baseline synaptic weights are closer to their maximum. First, we used a theory of synaptic plasticity that explicitly fit the baseline and the maximum synaptic weight, then we fit a descriptive phenomenological model, and finally, we adapted the Moran process, a model from evolutionary biology, to describe extracellular plasticity experiments.

The methodological differences between how experimental and theoretical work is carried out makes it challenging to integrate modelling into an experimental framework, and vice versa. However, through these five modelling examples, we argue that a successful model is not one that can capture all the features of the data, but rather one that can address the questions and hypotheses that cannot be studied with experiments alone.

Statement of Authorship

The text and figures presented in this thesis represent my own work, but all the experimental data presented was collected by my experimental collaborators. From the Butt lab, I used results collected by Dr. André Marques Smith, Dr. Jacqui Stacey, and Dr. Liad Baruchin. From the Cancedda lab, I used data collected by Dr. Salvatore Magara. I also received a lot of helpful guidance and advice from Prof. Tim Vogels and the Vogels lab on the best modelling approaches, guidance on how to interpret the experiments from members in the Butt lab and the Cancedda lab, and was introduced to the Moran process used in Chapter 8 by Dr. Chico Camargo.

Funding

I was funded by the Wellcome Trust (grant code: 105251/Z/14/Z), the Clarendon Fund, and the Foley-Béjar Scholarship.

Contents

1 Introduction **1**
- 1.1 Models in Neuroscience . . . 3
- 1.2 Models of synaptic plasticity . . . 4
- 1.3 Thesis overview . . . 8
  - 1.3.1 Part I: Setting up cortical circuits . . . 8
  - 1.3.2 Part II: Down Syndrome and synaptic plasticity . . . 9

I A transient translaminar circuit in developing mouse barrel cortex **11**

2 Inhibitory transient circuits in brain development **13**
- 2.1 Transient circuits in the brain . . . 13
- 2.2 Transient GABAergic circuits . . . 14
- 2.3 A modality-specific activity-dependent GABAergic circuit . . . 16
- 2.4 Part I outline . . . 18

3 Plasticity Mechanisms driving Transient Circuits **21**
- 3.1 Introduction . . . 21
- 3.2 Methods . . . 22
  - 3.2.0.1 *K*-means clustering . . . 22
  - 3.2.0.2 Principal components analysis (PCA) . . . 22
  - 3.2.1 Implementation . . . 23
- 3.3 Results . . . 23
  - 3.3.1 No parameters found to reproduce connectivity changes in the data . . . 23
  - 3.3.2 Stable weight configurations are determined by small number of parameters . . . 26
  - 3.3.3 Connectivity changes are reproduced in a network with tuned plasticity rules . . . 31
- 3.4 Discussion . . . 33

4 The functional role of transient circuits **39**
- 4.1 Introduction . . . 39
- 4.2 Methods . . . 40
  - 4.2.1 Network model . . . 41
  - 4.2.2 Input streams . . . 42
  - 4.2.3 Tuning of the input excitatory and inhibitory populations . . . 43
    - 4.2.3.1 Unspecific tuning . . . 44

4.2.3.2 Group-specific, stream-unspecific tuning 45
4.2.3.3 Gaussian 45
4.2.3.4 One-to-one 46
4.2.4 Weight initialisation 46
4.2.5 Plasticity rules 46
4.2.6 Modelling the formation of barrels 47
4.3 Results 47
4.3.1 Unspecific inhibition 47
4.3.2 Group inhibition 49
4.3.3 Gaussian inhibition 52
4.3.4 Edge cases 53
4.3.5 Random 54
4.3.6 Forming a barrel field 57
4.4 Discussion 58

II Synaptic plasticity in Down Syndrome 61

5 When development goes wrong: A mouse model for Down Syndrome 63
5.1 What is Down Syndrome? 63
5.2 Mouse models of Down Syndrome 64
5.2.1 Altered plasticity and inhibition in Ts65Dn mice 65
5.3 Experimental results 67
5.4 Outline of Part II 70

6 Applying statLTSP to extracellular LTP data 73
6.1 Introduction 73
6.2 Methods 74
6.2.1 statLTSP 74
6.2.2 Fitting statTLSP to data 78
6.3 Results 78
6.3.1 Obtaining model parameters from the extracellular data . . . 78
6.3.2 Fitting statLTSP to the data 80
6.3.3 Fitting statLTSP to multiple induction data 83
6.3.4 Using mean and variance to test the predictions of statLTSP . 85
6.3.5 Can statLTSP be used on extracellular recordings? 87
6.3.6 How does the system behave for $N > 1$? 92
6.3.7 The effect of $N > 1$ on the predicted directions 93
6.4 Discussion 94

7 Phenomenological model 97
7.1 Methods 98
7.1.1 Fitting the phenomenological model 98
7.1.2 Decision trees 99
7.2 Results 100
7.2.1 Parameter fit distributions are similar across genotypes 100
7.2.2 The data confirm the patterns of the phenomenological model parameters 101

7.2.3 The transient response to a plasticity protocol is most predictive of genotype . . . 102
7.2.4 The transient response to a plasticity protocol is not predictive of stable level of plasticity . . . 104
7.3 Discussion . . . 105

8 Using a Moran Process for LTP and LTD **107**
8.1 Methods . . . 108
8.1.1 The Moran process . . . 108
8.1.2 Neutral Moran Process . . . 109
8.1.3 Neutral Moran Process with selection . . . 110
8.1.4 Moran process with mutations . . . 111
8.1.5 The Moran process as a model of plasticity . . . 112
8.1.6 Simulating the Moran process . . . 117
8.1.7 Fitting the parameters of the Moran process to data . . . 119
8.2 Results . . . 120
8.2.1 The Moran Process fits experimental data qualitatively . . . 120
8.2.2 Evaluating the performance of the model . . . 125
8.2.3 Saturation determines stable LTP levels . . . 125
8.3 Discussion . . . 128

9 Conclusion **133**
9.1 Summary and Discussion . . . 134
9.2 Future Directions . . . 138
9.3 On making useful models and on making models useful . . . 140

A Table of Abbreviations **143**

B Stationary points of the original statLTSP formulation **145**

Bibliography **148**

Chapter 1

Introduction

> "*A man sets out to draw the world. As the years go by, he peoples a space with images of provinces, kingdoms, mountains, bays, ships, islands, fishes, rooms, instruments, stars, horses, and individuals. A short time before he dies, he discovers that the patient labirynth of lines traces the lineaments of his own face*"
> -Jorge Luis Borges, The Aleph and other stories

Using mathematical models to understand the world around us is a tale as old as time. One of the earliest mathematical models were built and applied in the context of calendars and timekeeping. Ancient Egyptians and Ancient Mayans independently found ways to predict eclipses, which were regarded as dangerous events that should be prepared for (Filar, 2002; Thompson, 1974). Predicting eclipses required a deep understanding of astronomy coupled with mathematical derivations. Furthermore, physical evidence of the existence of astronomical models remain in the form of scientific instruments. For example, the astrolabe is a scientific object that was used to make several astronomical measurements, and would have required a model of the way stars move in relation to the Earth. In particular, their model assumed all stars lay on a sphere around our planet, and then projected this sphere onto a flat disc that then became the face of the astrolabe (Morrison, 2007). Although the 'model' underpinning the astrolabe was geocentric, which we now know is incorrect, astrolabes enabled people to tell time from the height of the stars

in the sky.

More recent mathematical models that continue to inspire those who work in this field today include those proposed by Sir Isaac Newton and Albert Einstein. In their cases, the models they proposed did not take a physical form like the astrolabe, and yet, in a few lines of equations, they provided deep insights into the way the world works. For instance, even though all humans have a lot of personal experience with the way objects move, Newton's three laws of motion summarises the principles governing any individual instance of motion. One of the strengths of models is that they can move beyond what has ever been observed and make predictions, and can venture into areas that are not accessible experimentally. This is the case for Einstein's gravitational waves, which were predicted by Einstein in 1916, but were only directly observed a century later in 2016 (Abbott et al., 2016). As such, mathematical models are compelling tools in science because they can help simplify complex phenomena into a set of principles and make predictions that can inspire future experiments.

Nonetheless, the utility of mathematical or computational models in the life sciences is still a subject of debate (Ghani, 2016). It seems that part of the challenge of using mathematical models in biology is that there is a big gap in the language used by biologists and that used by modellers. Beyond the language gap, one of the challenges that models face in Neuroscience is to reduce the complexity of the brain whilst still producing relevant insights for experimentalists. Navigating the trade-off between precision and simplicity is illustrated by Ask & Reza (2016), who argue that the limited applicability of models in Neuroscience is due to the fact that models often ignore important variables or oversimplify the complexity of biological processes, but recognise that including too much detail also reduces the explanatory value of a model. Given the diversity of questions and data collection methods in Neuroscience it is clear that there will never be a one-fits-all model, especially since models can be used for a variety of purposes, from understanding the mechanisms and principles producing the data patterns observed, to generating

hypotheses and integrating knowledge (Brodland, 2015). Nonetheless, it remains crucial to understand what makes models useful to be able to integrate theoretical and experimental approaches in research.

1.1 Models in Neuroscience

One of the most important models in Computational Neuroscience is the Hodgkin-Huxley model (Hodgkin & Huxley, 1952). This study used a series of experimental findings on the way charged ions flow through the membrane of the squid giant axon at rest and during an action potential to propose a mathematical model of an action potential. Hodgkin & Huxley (1952) used voltage-clamp recordings to see the permeability of the axon to sodium and potassium, which were presented to the axon in different concentrations. These responses were then fit using ordinary differential equations, and when the effects of sodium and potassium were put together, they could effectively capture the shape of an action potential. This is an example of a simple model that was built based on very detailed biophysical observations of a particular system, and provided insights into how action potentials were produced. This model remains quite close to the data and its results are easy to relate to the biological system.

On the other hand, there are theoretical models that are inspired by data, but their goal is to understand fundamental principles of how the brain processes information. For example, Olshausen & Field (1996) investigated whether the responses of neurons in primary visual cortex (V1) could result from the brain using a sparse code to process images. Sparse coding and related theories like predictive coding, continue to influence the way that experimentalists and theorists think of the brain (Friston & Kiebel, 2009; Bastos et al., 2012; Whittington & Bogacz, 2017).

Models that remain faithful to the experiments and data that inspired them like the Hodgkin-Huxley model, as well as models that abstract away from data to propose principles governing biological processes in the brain like Olshausen & Field

(1996) are both valuable types of models in Neuroscience. However, the way each model interacts with and contributes to experimental work is different. It is easier to use models that remain close to the data to propose future experiments and iterate between doing experiments and updating the model. On the other hand, abstract models that care more about fundamental computations in the brain provide a framework to think about what the brain is doing, and that sets the intellectual foundation to develop new models and interpret experimental results.

A framework that captures the proximity of models to experiments comes from David Marr, who proposed that there were three different levels from which one could discuss computation in the brain (Stevens, 2012; Marr & Poggio, 1976). These levels were the following: (i) computational, or the goal of a computation and why that makes sense, (ii) algorithmic, or the algorithm that can transform the input to the appropriate output, and (iii) implementation, or the way in which this algorithm can be achieved given the particular constraints of the machinery available to the brain (McClamrock, 1991). These levels are clearly related to each other, but the type of questions that are asked at each level is quite different. In the course of this thesis we will explore models of synaptic plasticity from all three levels of abstraction.

1.2 Models of synaptic plasticity

Neurons are often modelled as analogue processing units that process, transform, and propagate information in the form of electrochemical signals. Due to differences in the inner and outer concentration of charged ions like sodium (Na^+), potassium (K^+), and chloride (Cl^-), there is an electric potential across the membrane of the neuron, which at rest is around -60mV (Hille, 2001). Inputs to the neuron result in changes in the membrane potential. When the membrane potential exceeds a certain voltage, the neuron fires an action potential due to the presence of voltage-gated ion channel that open at certain values of the membrane potential (Hodgkin & Huxley, 1952). An action potential results in the release of vesicles containing

neurotransmitters that can then bind to receptors that lead to an electrical deflection in the membrane potential of downstream neurons (Purves et al., 2014). The connections between neurons are called synapses, the neuron that releases neurotransmitters is the presynaptic neuron, and the neuron that receives the signal is the postsynaptic neuron. Information is propagated through these all-or-nothing events, and when they are not spiking, neurons are integrating inputs that they receive. There is a different class of connections known as electrical synapses or gap junctions that directly connect the cytoplasm of two neurons, and as a result, even subthreshold changes in membrane potential are transmitted and shared between neurons (Purves et al., 2014). However, in this thesis, we focus solely on chemical synapses.

Since communication between neurons happens through changes in the membrane potential, we can characterise the connection between neurons with the magnitude of the deflection and whether the deflection was depolarising or whether it was hyperpolarising. The main depolarising neurotransmitter in the mammalian brain is glutamate as it opens AMPA and NMDA receptors, which allow cations or positive ions to flow into the cell (Kandell et al., 2000). The main inhibitory neurotransmitter in the mammalian brain is GABA as it binds to $GABA_A$ receptors that are permeable to chloride and $GABA_B$ that are linked to opening potassium channels through second messengers (Kandell et al., 2000). Any given neuron will release a single class of neurotransmitters, so excitatory neurons are glutamatergic as their effect brings the postsynaptic neuron closer to firing an action potential, and inhibitory neurons are GABAergic as their effect brings postsynaptic neurons further away from firing an action potential. The magnitude of the postsynaptic membrane potential deflection in response to a presynaptic spike depends on various factors like the number of neurotransmitters released into the synaptic cleft and the number of receptors that are in the postsynaptic membrane (Costa et al., 2017). In most models, however, the connection strength is represented by a single parameter that is called the weight, w, of the connection (Gerstner et al., 2009). However,

in more complex models, the weight parameter is broken down into more detailed descriptions that, for example, distinguish between the presynaptic and postsynaptic contributions to the magnitude of the response (Larkman et al., 1992; Costa et al., 2017).

The strength of the connections between neurons are known to change throughout life at different timescales. Short-term plasticity transiently changes the efficacy of a synapse depending on presynaptic spiking activity, and happens in the order of hundreds to thousands of milliseconds (Zucker & Regehr, 2002). However, lasting changes in the synaptic weight are also possible, and these are thought to be the changes that are most important to experience-dependent changes, learning, and memory in the brain (Kandell et al., 2000). Donald Hebb is famous for postulating that synaptic plasticity existed and was important for memory and learning based on theory (Hebb, 1949). It was only much later that these changes could be measured and confirmed experimentally, and it was shown that synapses could undergo long-term potentiation (LTP) and long-term depression (LTD) (Bliss & Lømo, 1973; Debanne et al., 1998). It is impossible to describe all models of synaptic plasticity, but we will briefly introduce different classes of models and examples that will be most relevant to the scope of the models used in this thesis. Synaptic plasticity models can broadly be subdivided into phenomenological models, biophysical models, and optimal models of plasticity.

Phenomenological models capture the basic behaviour of a system without incorporating any of the chemical and biophysical components that give rise to the phenomenon. The first experiments that confirmed the existence of long-term plasticity in the brain used tetanic stimulation protocols to induce these changes (Bliss & Lømo, 1973). As a result, the simplest phenomenological model of plasticity uses firing rates as the main contributor to the changes in synaptic strength. If a presynaptic neuron fires at a rate x, and the postsynaptic neuron fires at a rate y, then the change in the firing rate is given by a function F that depends on x, y, and the

current synaptic weight w and can be written as,

$$\frac{dw(t)}{dt} = F(x, y, w). \tag{1.1}$$

The dynamics of the weight changes will be largely determined by the particular form of F that is chosen. While rate-based models can capture changes that are seen in the brain, experiments have also shown that the relative timing of spikes are important (Bi & Poo, 1998; Vogels et al., 2013). There are models that explicitly take into account the importance of spike times in the resulting changes in the synaptic weight, and these are called spike-timing dependent plasticity (STDP) rules.

Biophysical models of synaptic plasticity explicitly model components that are known through experiments to be relevant to the magnitude of the change that is induced. For example, it thought that the intracellular calcium concentration is a key component of plasticity (Stevens et al., 1994). A hypothesis as to how calcium could mediate different changes in the synapse is that moderate intracellular calcium concentrations results in decreases in the connection weight while high levels of intracellular calcium leads to increases in the connection strength through calcium's ability to activate CaMKII (Lisman, 1989; Stevens et al., 1994). Models that explicitly model the changes in intracellular calcium concentration, use a biophysical quantity to explain the changes in the weights of connections (Shouval et al., 2002; Graupner & Brunel, 2012). Finally, there is a separate class of models that study synaptic plasticity from the perspective of the computational problem that synaptic plasticity is trying to solve. Some of these models look at the way information coding may be optimised in the brain (Olshausen & Field, 1996; Friston & Kiebel, 2009; Bastos et al., 2012; Singer et al., 2018). Other synaptic plasticity rules maximise information transmission (Toyoizumi et al., 2005, 2007), maximise expected reward (Xie & Seung, 2004), represent slowly-varying features of stimuli (Sprekeler et al., 2007), and proposes nonlinearity in Hebbian learning as the fundamental component for receptive field formation (Brito & Gerstner, 2016). Finally, one of these models is statLTSP, which proposes that synapses undergoing LTP

or LTD are aiming to become more reliable and reach a particular target strength (Costa et al., 2017). As such, this model captures the changes observed in the data through a computational principle rather than by using a descriptive variable that captures the behaviour directly like in phenomenological models or incorporating variables that are biologically relevant like in biophysical models.

1.3 Thesis overview

In spite of the diversity of computational models that exist, it still remains challenging to successfully establish collaborations between experimental groups and theoretical groups. In this thesis we will discuss the results of two collaborations through five models. The first half of my thesis is the product of a collaboration between the Vogels and Butt groups and looks at the way a particular set of connections in mouse barrel cortex changes over the course of the first two postnatal weeks. The second half of my thesis was done in collaboration between the Vogels and Cancedda groups, where we looked at how learning is affected in Down Syndrome.

1.3.1 Part I: Setting up cortical circuits

One of the most popular metaphors of the brain in the last century is that the brain is like a computer. This notion has inspired many papers and reflections on the function of a brain. Whether or not computers are a useful metaphor to think about the brain, they are definitely not suitable to describing brain development. Whenever one makes any changes to a computer, like adding RAM memory or an extra hard drive, the first step is to turn the computer off. In contrast, brains cannot turn themselves off, and still, the brain undergoes radical changes throughout development. A more appropriate metaphor for brain development might be the construction of a building. In parallel with the final construction, there are temporary scaffolds that are put into place to enable builders to safely put down the foundations of the

building. In the end, these scaffolds are taken away, and all that remains is the building.

In practice, identifying the elements of the developmental scaffolds needed to construct a brain is challenging because there are many moving pieces that seem to be part of a perfectly orchestrated performance. The first half of the thesis focuses on transient GABAergic cortical circuits in the normally developing brain (Butt et al., 2017). In particular, we focus on a specific circuit, referred to as the L5b-L4 loop, that is found in the mouse primary somatosensory cortex also known as the barrel cortex. This circuit is composed of connections between layer 5b (L5b) somatostatin-positive (SST+) interneurons with L4 excitatory neurons for a brief period in the first two postnatal weeks (Anastasiades et al., 2016; Marques-Smith et al., 2016). While *in vitro* experiments can provide a snapshot of the connectivity of neurons at different times during development, due to the acute nature of these experiments, they provide little insight into the dynamics of the connectivity changes. We proposed to use modelling to explore the activity-dependent mechanisms that drive the changes in the L5b-L4 loop, and explore the possible functional roles that this circuit may play during development.

1.3.2 Part II: Down Syndrome and synaptic plasticity

The second part looks at the consequences of changes in the GABAergic system in a neurodevelopmental disorder on synaptic plasticity in adult animals. We analyse and model plasticity data collected from the Ts65Dn mouse, which is the most common mouse model of Down Syndrome (DS). We use three models to formalise and test hypotheses proposed by the Cancedda lab to explain the origin of the differences in plasticity in Ts65Dn animals and wildtype (WT) littermates.

We shift the focus of our models away from the implementation of particular connection changes whose function we do not understand, to studying changes in connections that are tied directly to learning and memory in the brain. This enables us to not only frame the problem at the level of implementation (*i.e.* what are the

changes in the Ts65Dn animals that lead to impairments in learning?), but also at an algorithmic (*i.e.* how does a particular sequence of stimuli result in a specific change in the response magnitude?) and a computational level (*i.e.* what is the goal of a synaptic weight change in the context of learning, and given such a principle how does this computation shed light on the learning impairment observed in DS?).

While an understanding of the mechanisms by which changes in the GABAergic connections in DS affect synaptic plasticity is needed in order to develop pharmacological treatments, we chose to approach this data from computational and algorithmic levels instead. A big factor in this modelling choice was the large number of parameters needed to build spiking or rate networks. The majority of papers that exist use parameter values that are taken from measurements of "normal" adult rodent brains. These parameters, however, may not be the same in Ts65Dn mouse brains due to the complex set of changes in the brain observed in Down Syndrome. To increase the interpretability of our results, we opted for abstract models that may allow us to build a framework that can be used to think about and guide future experiments on synaptic plasticity in Down Syndrome, even if they do not provide direct evidence of the role of GABA in DS.

Part I

A transient translaminar circuit in developing mouse barrel cortex

Chapter 2

Inhibitory transient circuits in brain development

A wide range of genetic and activity-driven mechanisms are active at different moments of development in order to build the complex networks of neurons that are part of the brain (Polleux, 2005; Markram et al., 2004; Khazipov & Luhmann, 2006). Interfering with either the genetic components or sensory-driven activity can have long-term consequences, leading to changes in the connectivity of the brain and can ultimately impact brain function, and behaviour (Rubenstein & Merzenich, 2003). In the first half of this thesis we will focus on a particular circuit that assembles and disassembles within the first two postnatal weeks in the mouse whisker barrel cortex. Although this circuit has been characterised in detail through experiments and genetic manipulations (Marques-Smith et al., 2016), the experiments could not fully address the questions of what mechanisms were driving the structural connectivity changes, nor their role during development.

2.1 Transient circuits in the brain

During brain development, overproduction of connections and neurons that are subsequently refined and pruned in later stages of development is necessary for the proper wiring of the brain (Riccomagno & Kolodkin, 2015). Both excessive and

insufficient pruning and refinement are thought to underpin diseases like autism and schizophrenia (Neniskyte & Gross, 2017; Thomas et al., 2015). Activity helps sculpt the network of connections both for small-scale refinements like pruning a single synapse, as well as for large-scale changes like pruning axon collaterals or cell death (Vanderhaeghen & Cheng, 2010; Cohen-Cory et al., 2010).

In cortex, one of the most well-studied transient neural populations is the subplate (SP). SP contains some of the earliest neurons in cortex, plays an important role in setting up the correct wiring of cortical circuits, and is the first cortical population to make connections with thalamus and amplify thalamic inputs into cortex during development (Kanold, 2009). Although SP was thought to disappear after development, it is evident that some non-pyramidal subplate cells form part of Layer 6b in the adult mouse (Marx et al., 2015; Luhmann et al., 2018). It has been proposed that SP plays an important role in setting up the thalamocortical inputs onto cortex, and ablation of SP in primary visual cortex (V1) interferes with the formation of ocular dominance columns and the GABAergic transition from depolarising to inhibitory (Kanold & Shatz, 2006). In barrel cortex, subplate neurons also receive thalamic input early, and GFP labelled subplate neurites are shown to extend towards the cortical plate early, by P4-P6 they cluster within the barrels, and this pattern changes again by P10, when the neurites move towards the space between the barrels called the barrel septa (Pinon et al., 2009). Although SP is the most well-known and well-described transient population in cortical development, we focus on a more recently discovered GABAergic transient circuit in barrel cortex.

2.2 Transient GABAergic circuits

GABAergic interneurons play a variety of roles throughout development (Butt et al., 2017; Ben-Ari et al., 2007; Akerman & Cline, 2007). GABA is thought to be depolarising early in development due to the expression of a sodium-potassium-chloride transporter called NKCC1 that increases the internal chloride concentration (Kaila

et al., 2014) As NKCC1 creates a larger concentration of intracellular chloride than extracellular chloride, when GABA opens $GABA_A$ receptors, chloride ions leave the cells leading to a depolarising deflection in the membrane potential. Recent experiments confirmed the developmental changes in intracellular chloride by visualising chloride concentrations in cells *in vivo* (Sulis Sato et al., 2017). There has been a lot of debate about the effect of depolarising GABA on developing circuits, and there is evidence that while GABA can be excitatory in *in vitro* experiments, in *in vivo* experiments GABA still effects a net inhibitory effect on network activity (Kirmse et al., 2015; Valeeva et al., 2016).

In addition to shaping activity in the developing brain, GABAergic interneurons are thought to play an important role in controlling plasticity mechanisms, and may even control the generation of new excitatory synapses (Oh et al., 2016). For example, in primary visual cortex, GABAergic interneurons and specifically parvalbumin-positive (PV+) interneurons are thought to control the onset and closure of critical period plasticity (Hensch et al., 1998; Hensch, 2005). Critical period plasticity refers to a defined window of time when certain connections are most plastic and sensitive to external stimuli. If during the critical period, visual experience was somehow altered, this will have lasting consequences even if normal vision is restored afterwards (Hensch, 2005). In the optic tectum of the tadpole, depolarising GABAergic signalling was shown to be required for the onset of sensory-driven spiking activity (van Rheede et al., 2015). Recent work has shown that there are transient interneuron circuits in mouse barrel cortex (Anastasiades et al., 2016) that are thought to be important for feedforward inhibitory circuits to mature (Tuncdemir et al., 2016), and to shape the responses of excitatory cells to sensory input control the size of barrels in primary mouse somatosensory cortex in the adult (Che et al., 2018). For the rest of this chapter we focus on a third transient GABAergic circuit that connects L5b SST+ interneurons and L4 SSNs (Marques-Smith et al., 2016).

2.3 A modality-specific activity-dependent GABAergic circuit

Marques-Smith et al. (2016) describe a transient circuit involving somatostatin-positive interneurons (SST+ cells) in Layer 5b (L5b) that receive thalamic input and Layer 4 (L4) spiny stellate neurons (SSNs), which are the primary recipients of thalamic input in the mature cortex even though recent studies have found thalamic inputs into other layers like L5b (Constantinople & Bruno, 2013). The transient circuit connecting these two neural populations, which we will refer to as the L5b-L4 'loop', is present from postnatal day 4 (P4) to P6, begins to disassemble between P7-P9 and is replaced by a local inhibitory circuit from P10-P15 (Figure 2.1).

The disassembly of the circuit was hypothesised to be at least partially determined by sensory activity since early transection of the infraorbital nerve (ION) in P1 pups led to a delay in the collapse of the 'loop'. Marques-Smith et al. (2016) argued that the delayed timeline was due to a decrease in activity. To test this hypothesis, similar experiments were done on the "neocortex isolé" conditional Celsr3, where thalamocortical axons fail to innervate cortex leading to an 'isolated' cortex (Zhou et al., 2008). The translaminar L5b-L4 loop was found to exist in Celsr3 mutant animals and the timeline of the disassembly of the circuit was the same as the wildtype animals (Stacey, 2018). These results seemed contradictory as the two experimental interventions were thought to decrease activity in cortex but had different effects on the timeline of the circuit. However, Frangeul et al. (2017) showed that the ION transection leads to neuronal hyperexcitability in VPM, *i.e.* that neurons are more responsive to depolarising stimulation. These results raise the possibility that ION transection leads to excessive input activity during development. On the other hand, *in vivo* recordings in the first postnatal week of cortex isolé animals show spindle burst activity that is statistically not different to activity wildtype animals (*unpublished data*). This suggests that the timeline of the L5b-L4 loop is extended by excessive activity, and that as long as the amount of network activity is simi-

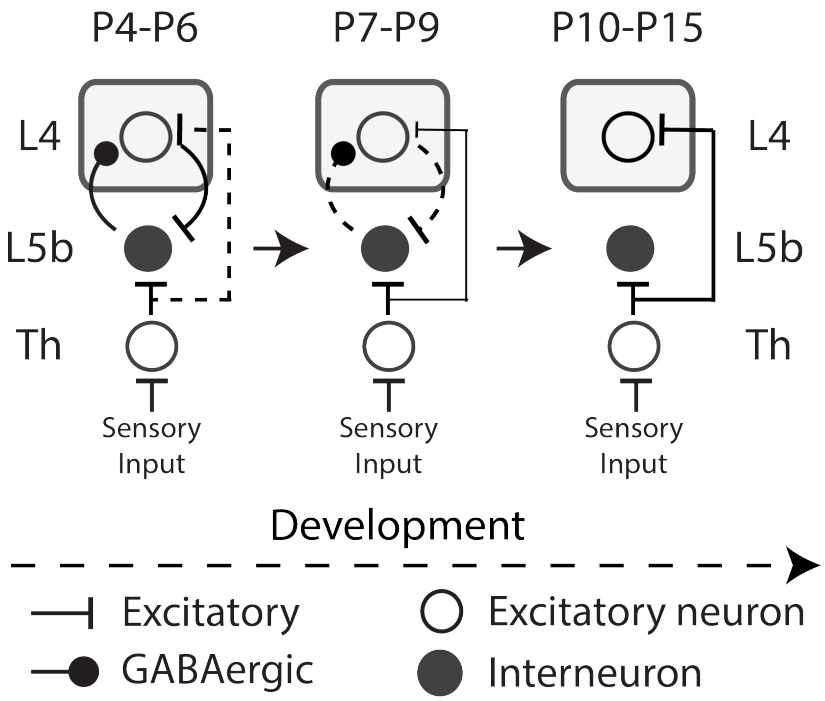


Figure 2.1: **L5b-L4 loop.** Marques-Smith et al. (2016) described a transient translaminar connection between L5b interneurons and L4 excitatory neurons that was present from postnatal day 4 (P4), and had disappeared by P10.This coincides with the period of time when thalamic inputs onto L4 have arrived and connections are refined.

lar, the disassembly of the connections will follow the same time course as those in wildtype animals.

Finally, this circuit seems to be modality specific as it was not found in visual cortex during early development (Stacey, 2018). Experiments were carried out not only in the ages when the loop was found in barrel cortex, but also in older animals, as the critical period of plasticity in visual cortex happens after the critical period of plasticity in barrel cortex (Gordon & Stryker, 1996; Fox, 1992). This suggests that the L5b-L4 loop might be important to set up connections that are relevant to processing somatosensory information specifically. One of the unique characteristics of barrel cortex is the organisation of connections into 'barrels' where thalamic afferents corresponding to a whisker will reach the centre of a barrel in L4, and neurons in L4 will orient towards the centre (Erzurumlu & Gaspar, 2012). On the other hand, primary visual cortex is organised retinotopically, but is known to have a salt-and-pepper organisation of orientation tuning in mouse, so it is not as highly ordered as

barrel cortex (Kondo et al., 2016). One of the hypotheses we will test is whether the L5b-L4 loop could play any role in refining the barrel maps during the first week.

2.4 Part I outline

In Chapter 3, a spiking network was used to study whether the disassembly of the L5b-L4 loop could be mediated by spike-timing dependent plasticity (STDP) mechanisms. The goal of this study was to disambiguate the potential role of activity in the disassembly of the cortical circuit. Throughout the simulations, we measured the average connection strengths for the three plastic connections that were the connections between thalamus (Th) and L4, the connection from L4 to L5b, and the GABAergic connection from L5b to L4 (see Figure 2.1). The spiking network was simulated many times with different combinations of parameter values, and the results were evaluated based on how close the final average strengths of each connection type were to the experimentally measured connections (Marques-Smith et al., 2016). We found that in order to capture the connectivity changes, we had to use different plasticity rules for the two excitatory connections, and that the connection strength of the GABAergic neuron for one of the rules we used was highly dependent on the amount of input activity coming from thalamus.

In Chapter 4, a model of interacting excitatory and inhibitory plasticity introduced by Clopath et al. (2016) was extended to study the conditions under which an interneuron circuit like the L5b-L4 loop could enable barrels to form correctly. The L5b-L4 loop is present towards the end of the refinement of L4 connectivity, and we hypothesised that the loop could be involved in the refinement of L4 SSN receptive fields. In our model, each whisker input was represented by a set of independent input streams. We varied the specificity of the inhibitory responses to whisker input and found a set of conditions for the excitatory neuron to develop a receptive field that were specific to one whisker. We found that in some circumstances, the inhibitory input could stabilise the response specificity of a L4 excitatory cell, and

that the feedback excitatory connection onto the inhibitory neuron could then further refine the response of the L4 cell to responding to a subset of the inputs coming from a single whisker.

Chapter 3

Plasticity Mechanisms driving Transient Circuits

3.1 Introduction

In the first postnatal week, mouse brain development depends on a combination of genetic factors like transcription factors, and activity-dependent factors such as synaptic plasticity and patterned electrical activity (Khazipov & Luhmann, 2006). In the case of the transient translaminar connections between L5b SST+ interneurons and L4 SSNs, if the disassembly of the connections were genetically encoded, we would expect the disappearance of the connections to happen regardless of the type and quantity of activity in barrel cortex. However, activity appears to play a role in the the disassembly of the loop. For example, transecting the infraorbital nerve (ION transection), which prevents normal sensory activity from reaching barrel cortex, also changes the timeline of the loop (Marques-Smith et al., 2016). In this chapter we used models of spiking networks to attempt to disambiguate the potential role as activity-dependent mechanisms in the disassembly of this circuit.

3.2 Methods

3.2.0.1 K-means clustering

After running many simulations for different parameter combinations, we wanted to understand whether there were any patterns in the parameters that gave rise to different synaptic weight dynamics in the simulations. We used K-means clustering, which is one of several methods to group data by their similarity into K clusters or groups (Bishop, 2006). Assume that we have a dataset has N D-dimensional points $\{x_1, ..., x_N\}$, and we want to partition the data into K clusters. Ideally, the points in each cluster would be closer to other points in that cluster compared to their distances from points in other clusters. This is achieved by minimising the within-cluster sum of squares also known as a distortion measure and is defined as

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||x_n - \mu_k||^2, \tag{3.1}$$

where μ_k is a D-dimensional vector that represents the centroid of each cluster k, $r_{nk} \in \{0, 1\}$ is a binary variable that is 1 when x_n is in cluster k. The algorithm iterates between assigning points to clusters by updating r_{nk} and recalculating the cluster centroids μ_k, in order to minimise J, which is the sum of within-cluster distances from points to the centroid of their cluster.

3.2.0.2 Principal components analysis (PCA)

Principal components analysis is a technique that is used for a wide variety of applications, and can be defined as an orthogonal projection of data onto a lower dimensional linear space, in a way that maximises the variance of the projected data. As a result, the first principal component (PC1) is the component that captures the largest amount of variance in the data when the data is projected onto it. The second principal component (PC2) captures the largest amount of the remaining variance, under the constraint that it has to be orthogonal to PC1. Typically, most of the variance will be captured by the first few PCs and therefore, PCA is used to project the

data into a lower-dimensional space that is easier to study and visualise.

3.2.1 Implementation

All the spiking neuron simulations were written in C. Variables were updated in the simulation using a forward Euler method with timestep of 0.1ms, but mean weights and mean rates for the three connection types and neuron types were calculated at every second. All plotting and analysis were then done using Python.

3.3 Results

3.3.1 No parameters found to reproduce connectivity changes in the data

We simulated the network for a total of 648 parameter combinations for 100 minutes of simulated time, and compared the final connectivity of the network to the desired connectivity where the thalamic input onto L4 strengthened and the other two connections weakened. All the simulations ended far from the desired connectivity (Figure 3.1A). The simulation with the greatest error was one where the mean of all connection types was at the maximum strength, and would mean that the transient circuit did not disassemble (Figure 3.1B). On the other hand, the simulation whose final connectivity was closest to the desired connectivity was also not a good fit to the data, as it had not reached its steady state, and both excitatory connections were continuing to strengthen (Figure 3.1C).

In Figure 3.1 we plotted the weights for the trajectories that had the lowest and the largest error, but in order to get a more comprehensive understanding of the model's behaviour, we used K-means clustering. The elbow method was used to choose $K = 4$ clusters, because the decrease in the distortion or sum of within-cluster squared distances was much smaller after the fourth cluster and can be seen graphically as a kink or 'elbow' in the graph (Figure 3.2). The four clusters were

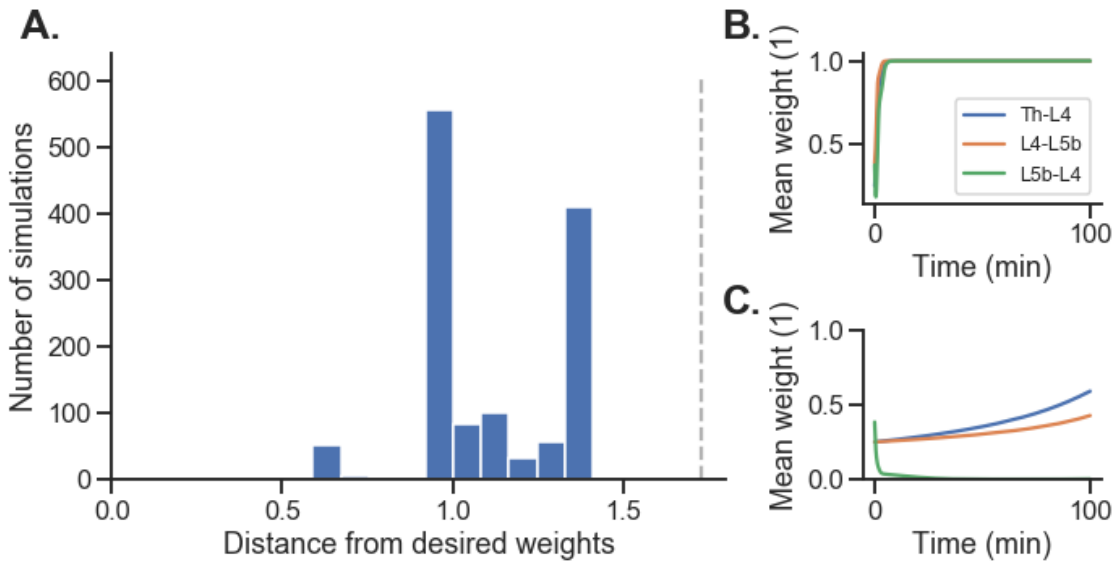


Figure 3.1: **No parameter combination tested reproduced experimental data with a single inhibitory and excitatory STDP rule. A:** The Euclidean distance from the desired mean weight configuration, and the simulated end mean weight configuration. All of the weights are expressed as a fraction of the maximum weight, so all possible weight combinations are constrained to a unit cube so the maximum Euclidean distance is $\sqrt{3} \approx 1.73$ marked with the vertical gray dotted bar. **B:** The mean weights of the thalamic connections onto L4 excitatory cells (blue), the L4 SSN connections onto L5b SST+ interneurons (orange), and the L5b SST+ interneuron input onto L4 SSNs (green) for the simulation whose final connectivity was the furthest away from the desired configuration. **C:** The trajectory of the three mean weights over time for the simulation that was closest to the desired configuration.

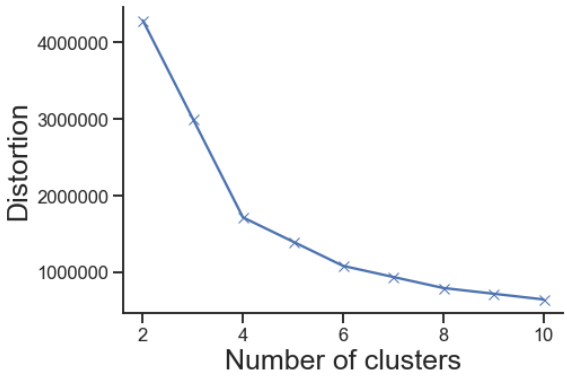


Figure 3.2: **Determining number of clusters using the 'elbow method'.** Simulated trajectories were clustered using k-means. The sum of within-cluster squared distances for different numbers of clusters. As numbers of clusters are increased, this value will decrease, but there is a change in the rate at which it decreases at $k = 4$. This means that adding a fourth cluster reduces the sum of squared distances more than adding any additional ($k > 4$) clusters.

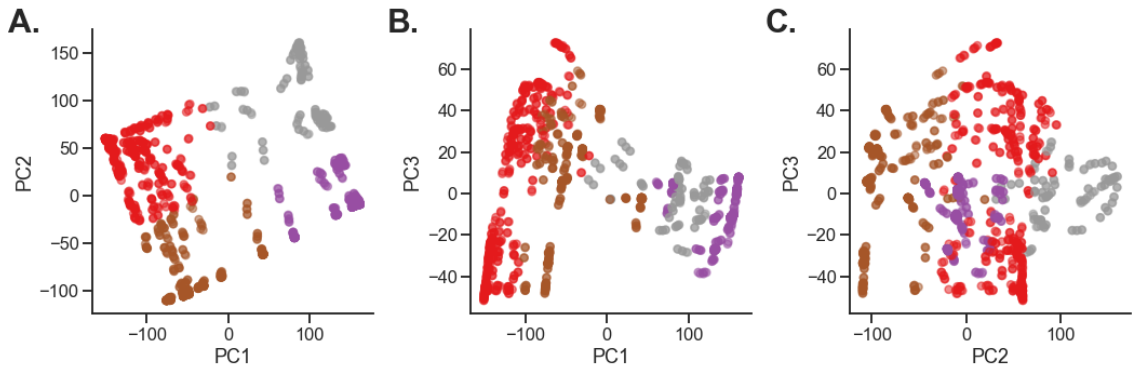


Figure 3.3: **Visualising the trajectory clusters using PCA.** A full set of trajectories contained 1800 points (600 points for each connection types), so to visualise the four clusters we first used PCA to project the data to a lower-dimensional space. (**A**:) Each point represents a single simulation, coloured by its respective cluster plotted on the first and second principal components (PCs). The same data are plotted on (**B**:) the first and third PCs and on (**C**:) the second and third PCs.

visualised by projecting the data to a lower dimensional space using PCA. The first principal component (PC1) captured 68% of the variance, the first two PCs captured 91% of the variance, and the top three PCs captured 95% of the variance. We projected the data onto 2-dimensional axes with combinations of the first three PCs (Figure 3.3). We found that the data did not separate into distinct clusters but rather seemed to lie on a surface as seen from the rectangular arrangement of points in the projection onto the first two PCs (Figure 3.3A). However, this surface was probably curved as seen from the projection onto the first and third PCs (Figure 3.3B). Although the points were mostly distinguishable into clusters when projected onto the first principal component, the projection onto the second and third PCs showed a larger overlap of all the clusters. When we plotted the points in a 3D scatterplot, we observed that the data seemed to fall on a curved 2D surface embedded in a 3D space (Figure 3.4). These results suggested that the data could be projected onto a roughly rectangular 2D hyperplane embedded in a 3D space, and that the clusters corresponded to the four quadrants of this rectangular manifold.

In order to understand what the four quadrants corresponded to in terms of the weights, the average trajectories of all simulations in each cluster were plotted (Figure 3.5). We found that the two excitatory connections plotted in blue and or-

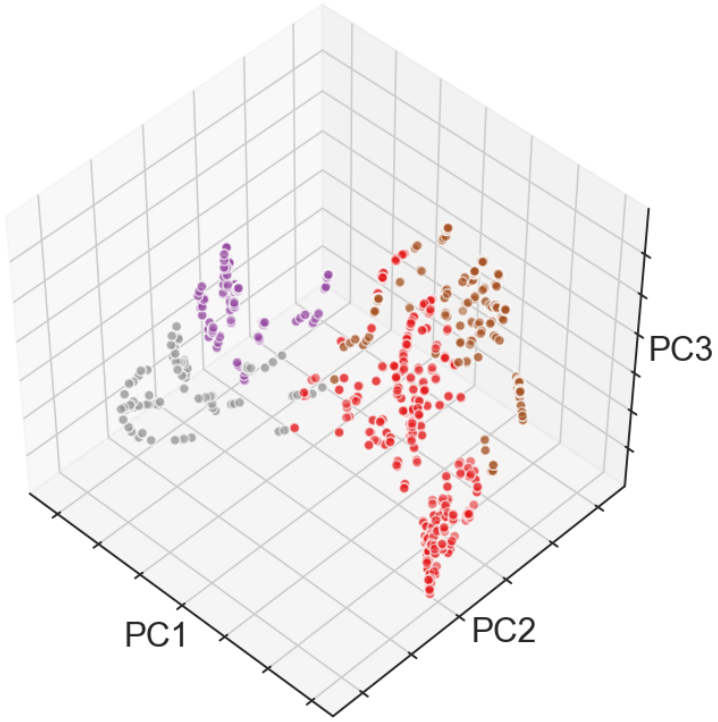


Figure 3.4: **Data projected onto three dimensions.** 3D plot showing all the data coloured by their assigned cluster projected onto the three first principal components identified through PCA.

ange strengthened or weakened together, and that out of the eight possible final connection strength combinations of the excitatory and inhibitory connections, the network tended to go towards one of following four configurations: strong excitatory and strong inhibitory connections (Figure 3.5A), weak excitatory and weak inhibitory connections (Figure 3.5B), strong excitatory and weak inhibitory connections (Figure 3.5C), and weak excitatory and strong inhibitory connections (Figure 3.5D).

3.3.2 Stable weight configurations are determined by small number of parameters

If the thalamic input onto L4 had strengthened in the second cluster (Figure 3.5 B), or if the excitatory connection from L4 to L5b had weakened in the third cluster (Figure 3.5 C), then the final connectivity of the network would have matched the exper-

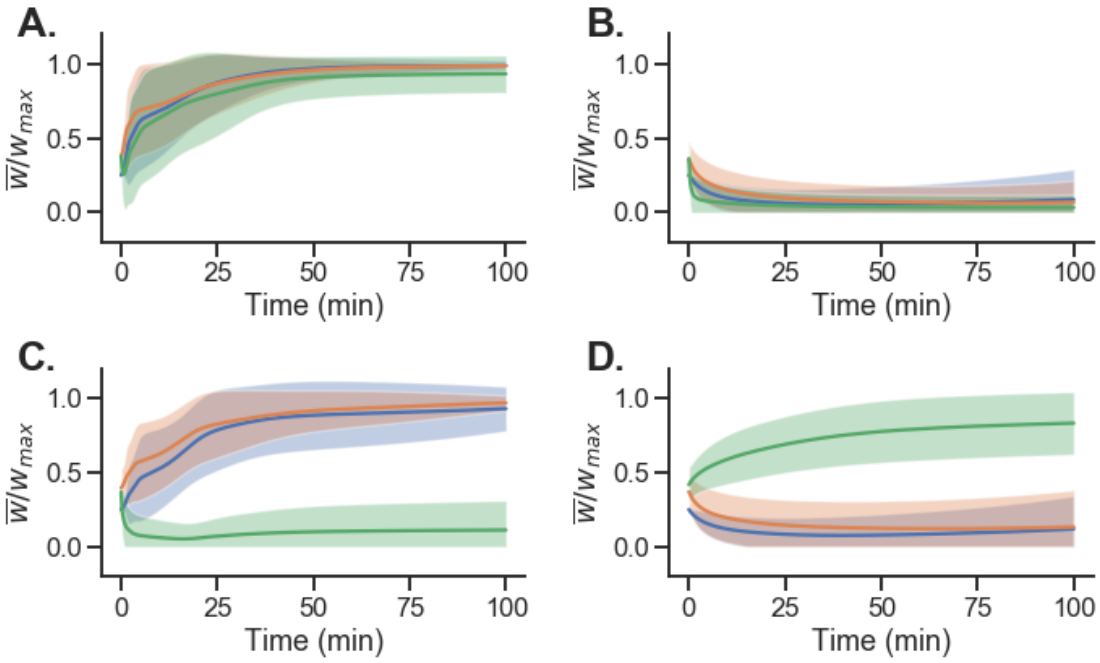


Figure 3.5: **Mean trajectories in each cluster.** The trajectories in each cluster were averaged, and the mean trajectory for each connection type is shown (Th-L4 in blue, L4-L5b in orange, and L5b-L4 in green). The shaded region is the standard deviation of each weight over the simulated time. The two excitatory connections strengthened or weakened together, and the clusters can approximately be described as combinations of strong and weak excitatory and inhibitory connections. Clusters had trajectories that displayed the following connection strengths at the end of a simulation: (**A**:) strong excitatory and strong inhibitory connections, (**B**:) weak excitatory and weak inhibitory connections, (**C**:) strong excitatory and weak inhibitory connections, and (**D**:) with weak excitatory and strong inhibitory connections.

imental data. In order to understand how to change the network to match the experimental data, we first explored the effect of individual parameters on the changes in the synaptic weights in our network. The parameter search was performed on five parameters that were the input rates, the reversal potential of GABA (E_{GABA}), a parameter setting the bias towards LTP or LTD for the excitatory STDP (eSTDP) parameter (β_{ex}), the type of inhibitory rule that was used (asymmetric (AH) or symmetric (SH)), and the inhibitory STDP (iSTDP) parameter, (ρ_0 or β_{inh}). In the case of the symmetric inhibitory rule, the parameter was the target postsynaptic firing rate (ρ_0), and for the asymmetric rule the parameter was similar to that of the eSTDP rule and set the bias towards LTP or LTD (β_{inh}).

In Figure 3.6 the effect of each parameter (rows) on the weight of each connection type (columns) was plotted by looking the distribution of final weights of a connection for each value of that parameter in a violin plot. Violin plots are similar to box plots, but the width of the envelope shows the probability density of the data at each value. We chose this visualisation method, because in most simulations the average weight for each connection type was either 0 or 1, and box plots would not be informative. We were interested in assessing the parameter sensitivity of each connection, or in other words, how much changing the value of the parameter changed the distribution of final weights for that connection type. Visually, this meant that when a connection was sensitive to a parameter, the violin plots would change for different values of that parameter (*e.g.*, Figure 3.6Ci).

All three connections were sensitive to the input firing rate (Figure 3.6A). For both excitatory connections, there was a subset of simulations with an input rate of 5Hz that had an intermediate final weight as there was an additional 'bump' in the violin plot. This bump corresponded to the simulations that had not reached a steady state like the example trajectory in Figure 3.1C. A higher proportion of weights strengthened for higher firing rates, but there was little difference between 10Hz and 15Hz (Figure 3.6Ai,ii). There was a similar trend for the inhibitory connection, as most inhibitory connections weakened for 5Hz and a larger proportion of connections strengthened as the firing rates increased (Figure 3.6Aiii). The distributions for E_{GABA} were the same for all three weights suggesting that the reversal potential did not have an effect on the final weight distribution of any connection type (Figure 3.6B). On the other hand, the parameter controlling the eSTDP had an strong effect on the two excitatory connections, where a value of 0.5 and 1 led to the strengthening of most connections, and a value of 1.5 led to both types of excitatory connections to weaken (Figure 3.6Ci,ii). Probably due to the strong effect of the eSTDP parameter on the excitatory connection weights, β_{ex} also had an effect on the distribution of connection weights for the inhibitory connection, where most inhibitory connections weakened for a value of 1.5 (Figure 3.6Ciii).

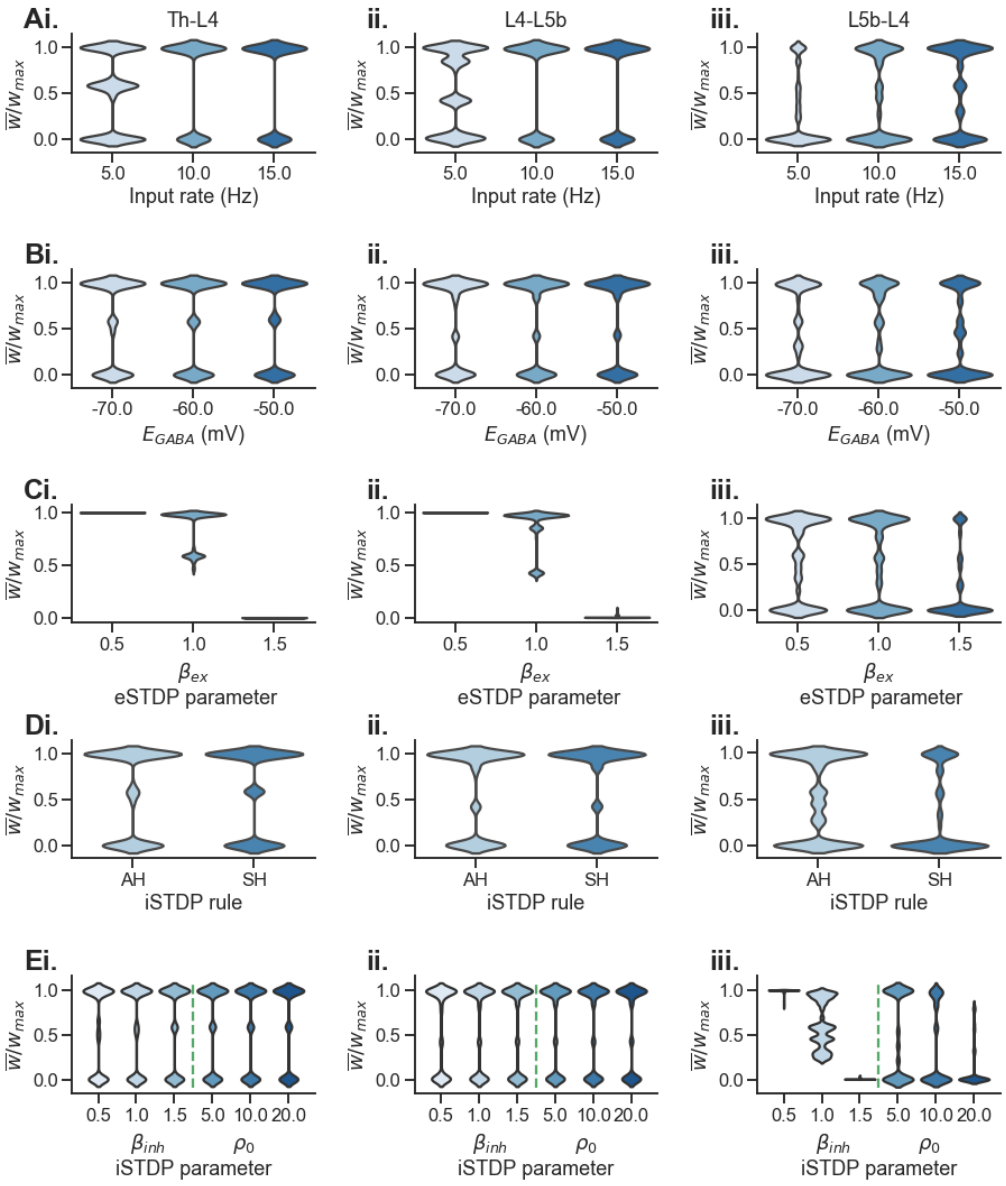


Figure 3.6: **Few parameters explain the final weights.** The importance of (**A:**) The activity of the input thalamic population affects the proportion of weights that strengthen for all connection types. (**B:**) The distribution of strong and weak weights for all three connection types is the same regardless of the reversal potential of $GABA_A$ (**C:**) The parameter controlling the excitatory plasticity affects the distribution of weak and strong connections for all three connection types, but has a stronger effect on the two excitatory connections (**Ci,ii**). (**D:**) The type of iSTDP affects only the distribution of connection weights of the inhbitory connection (**Diii**). (**E:**) The parameter controlling iSTDP affects only the distribution of inhibitory connection weights for both types of inhibitory rules. The first three distributions are the parameters for the Antihebbian rule, and the latter three distributions are the parameters for the symmetric stabilising rule.

The last two parameters related to the inhibitory plasticity rules, and we found that unlike the eSTDP parameter, β_{ex}, the iSTDP parameters, β_{inh} or ρ_0, did not change the distributions of weights for the excitatory connections (Figure 3.6D). We found an asymmetry in the distributions of weights depending on the plasticity rule for the inhibitory connection where a larger proportion of connections that used the symmetric rule weakened whereas the asymmetric rule had slightly more connections strengthening (3.6Diii). Since the iSTDP rule did not make a difference on the distribution of weights for the two excitatory weights, it made sense that when we plotted the distributions for the individual iSTDP parameters we also did not see a difference in the weight distributions for the excitatory connections (Figure 3.6Ei,ii). For the asymmetric rule parameters shown to the left of the vertical dotted line, the pattern was the same as the eSTDP parameters for the two excitatory connections, and for the symmetric rule on the right of the dotted line, the parameter was the target output firing rate, so the higher the postsynaptic firing rate target the more likely it was for the connections to weaken to zero (Figure 3.6Eiii).

Even though connections were sensitive to the input firing rate, the eSTDP parameter, the iSTDP parameter, and the iSTDP rule, Figure 3.6 was not informative about how combinations of parameter values give rise to the different trajectories in each cluster shown in Figure 3.5. To gain a deeper understanding of how different parameter combinations shaped the weight trajectories during different simulations, the distribution of values for each of these four parameters in each cluster were plotted (Figure 3.7). The first cluster, which had trajectories where all weights were strong at the end also had higher input rates than the other clusters (Figure 3.7A). Furthermore, the two clusters that had strong excitatory connections had $\beta_{ex} = 0.5$ or $\beta_{ex} = 1.0$, whereas the clusters with weak excitatory connections had $\beta_{ex} = 1.5$ (Figure 3.7B). We found that one of the clusters (weak excitatory strong inhibitory) was comprised exclusively by simulations that used the asymmetric iSTDP rule, and had parameter values of $\beta_{inh} = 0.5$ or $\beta_{inh} = 1.0$ (Figure 3.7C,D). For simulations using the asymmetric iSTDP rule, smaller values of β_{inh}

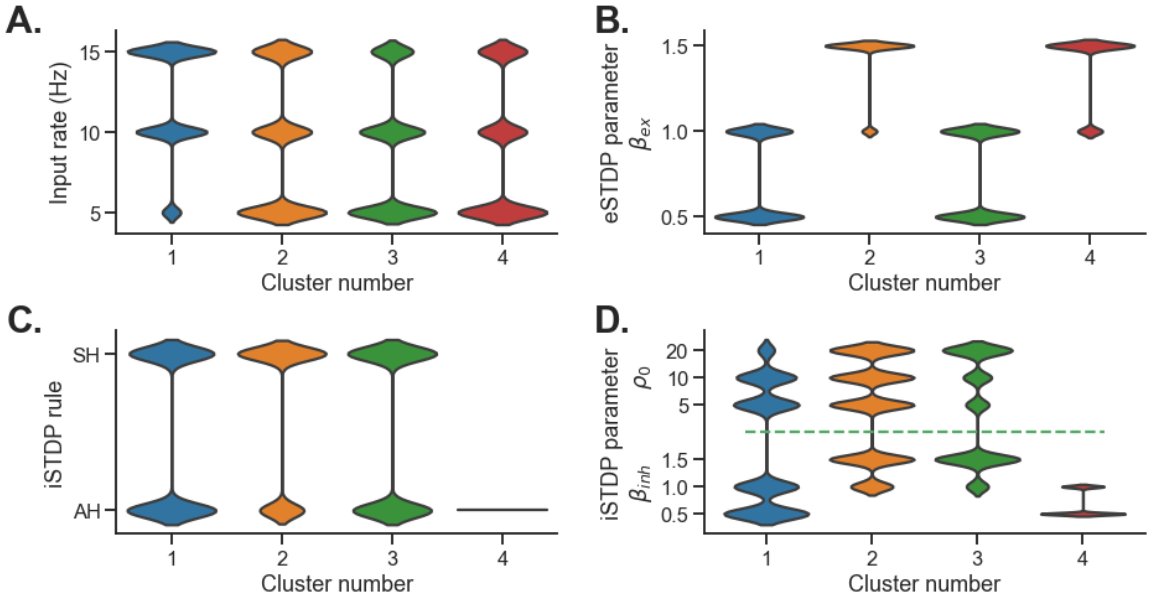


Figure 3.7: **Parameters contribute distinctly to each cluster.** (**A:**) Cluster 0, where all connections strengthened had a higher proportion of simulations with a higher input firing rate. The distributions over input rates in the other three clusters were similar to each other. (**B:**) Cluster 1 and Cluster 3 had mostly $\beta_{\text{ex}} = 1.5$, and these were the clusters where the excitatory connections weakened. This suggests that the higher β is, the weaker the excitatory connections will be. (**C:**) Cluster 0 and 2 have similar numbers of simulations with each inhibitory rule type. Cluster 0 is composed entirely by simulations with the Antihebbian inhibitory rule, and Cluster 1 has more simulations with the symmetric inhibitory rule. (**D:**) The distributions over iSTDP parameters for both inhibitory rules are shown together, and for display purposes, the y-axis was rescaled to show each parameter value equidistant.

led to a strong inhibitory final connection, and vice versa, and for the symmetric iSTDP rule, smaller values of ρ_0 made it more likely for the inhibitory connection to strengthen (Figure 3.7D).

3.3.3 Connectivity changes are reproduced in a network with tuned plasticity rules

We observed that the value of the parameter controlling the excitatory plasticity rule strongly determined whether a connection would strengthen or weaken. This would explain why the two excitatory connections strengthened together and weakened together. Originally, we had chosen to use the same parameter value for the excitatory plasticity rule for both excitatory connection types. However, the results

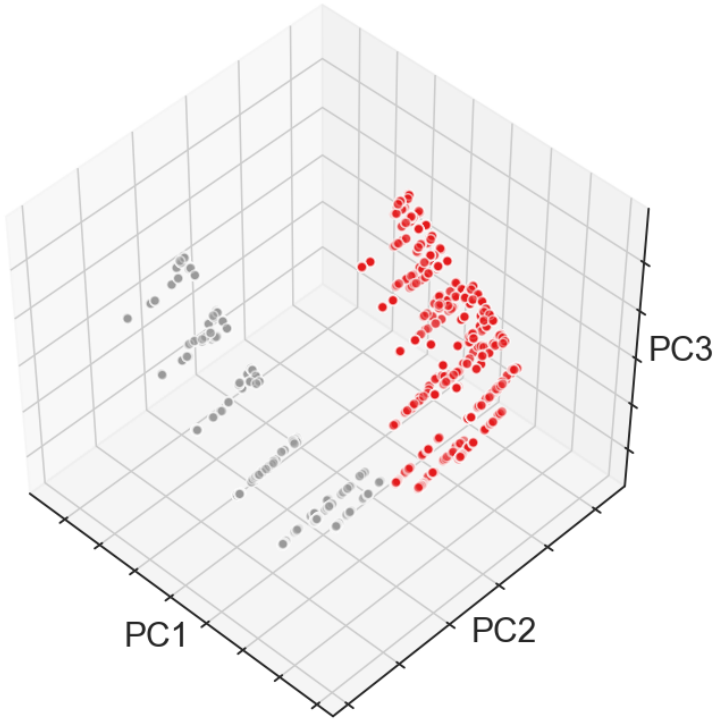


Figure 3.8: **Data projected onto three dimensions.** 3D plot showing all the data coloured by their assigned cluster projected onto the three first principal components identified through PCA.

in the previous section suggested that perhaps if we allowed the plasticity rules to be different for different connection types, then we would be able to reproduce the experimentally measured changes.

We tested this by choosing $\beta_{ex} << 1$ for the connection between the thalamic inputs and the L4 SSNs, and repeating the parameter sweep. The simulations were clustered using $K = 2$ (Figure 3.8). In one cluster, the L4-L5b connection strengthened and in the second cluster, the L4-L5b connection weakened. In both clusters, the standard deviation of the GABAergic connection was large (Figure 3.9A,B). Looking only at the simulations in cluster 2, where the excitatory component of the loop (L4-L5b) had weakened, we found that the input rate correlated with the final inhibitory connection weight, and that the input firing rate affected the outcome of the GABAergic connection but not the two excitatory connections in the system (Figure 3.9C).

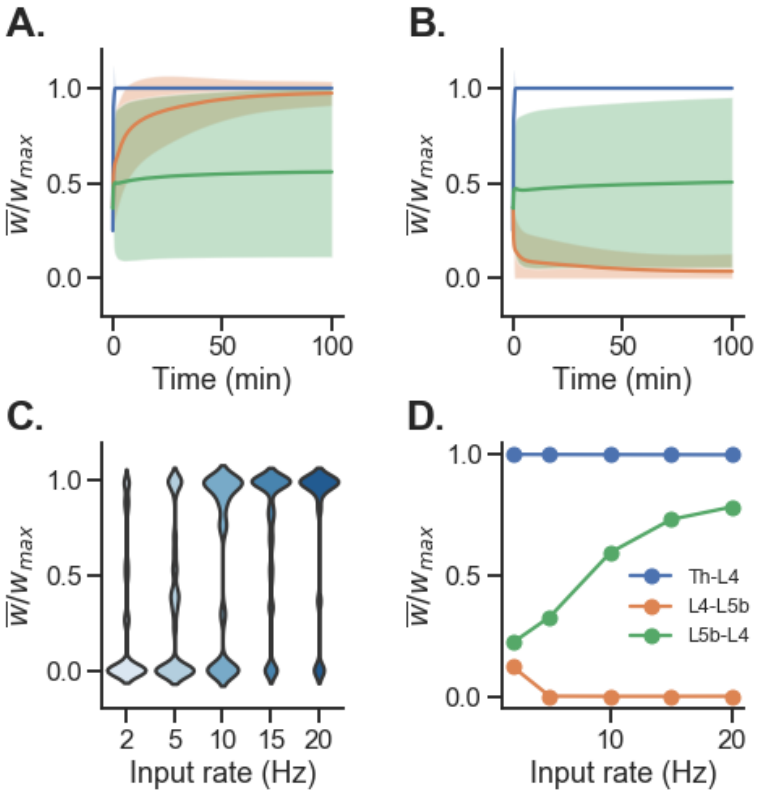


Figure 3.9: **GABAergic connection weight depends on input activity.** The trajectories in each cluster were averaged, and the mean trajectory for each connection type is shown (Th-L4 in blue, L4-L5b in orange, and L5b-L4 in green). The shaded region is the standard deviation of each weight over the simulated time. (**A:**) Cluster 1 had both excitatory connections strengthen and both strong and weak GABAergic connections. (**B:**) Cluster 2 had the L4 to L5b excitatory connection weaken, but both strong and weak GABAergic connections. (**C:**) Violin plots for the weight of the GABAergic connection for different input firing rates. (**D:**) Mean connection weights for the three connections for different input firing rates.

3.4 Discussion

In this chapter we built a spiking neural network to study whether activity-dependent mechanisms, specifically STDP, could explain the disassembly of the L5b-L4 loop. We found that if we used the same excitatory STDP rule for both excitatory connections, we were unable to capture the strengthening of the thalamic afferent inputs onto L4 and the weakening of the L4 connections onto L5b. However, if we allowed the two connections to have different parameters, then we were able to find combinations of parameters that captured the connectivity changes observed

in experiments. Although we were able to reproduce the connection weights by using different plasticity parameters at each connection, this suggests that while the system is sensitive to activity levels, they require highly tuned plasticity rules. While our findings could be criticised for choosing the particular form of STDP rules that worked, it would not be surprising to find that STDP rules are tightly regulated during development, especially as multiple studies have pointed out that STDP curves change over development and are different depending on the brain region and postnatal day (Vickers et al., 2018; Jiang et al., 2007; Abraham & Bear, 1996; Itami & Kimura, 2012; Kimura & Itami, 2019). Furthermore, even though we used an asymmetric Hebbian rule for the thalamic connection onto L4, by choosing $\beta_{\text{exh}} << 1$ in the second set of simulations, we were *de facto* using an all-LTP plasticity rule which is what Kimura & Itami (2019) suggests exists in thalamocortical connections during development.

We found that the amount of input activity strongly influenced the final strength of the inhibitory connections. This is of particular relevance as it suggested that the amount and quality of the input activity into the circuit may be more important than the exact origin of the signal, or the identity of the presynaptic partners. Furthermore, it provided further evidence for the hypothesis that may reconcile the seemingly contradictory results coming from the animals with ION transections and the cortex isolé animals. Both of these experimental interventions were proposed to reduce the amount of activity in barrel cortex by physically severing the nerve connecting the whiskers to the brain or via a genetic manipulation, but the subsequent experiments in barrel cortex suggested these manipulations had different effects on connectivity. While in ION transected mice the GABAergic connection from L5b into L4 survived into P15, this connection disassembled normally in the cortex isolé animals. These results, however, support the suggestions by Frangeul et al. (2017) that in the ION transected animals, activity may in fact increase because as the input rate in the model increases we find that the GABAergic connection survives in a larger fraction of the simulations. Furthermore, even though barrel cortex in the

cortex isolé animal is not receiving somatosensory input, as long as the amount of activity is similar to that of wildtype, the model would predict that the GABAergic connection would weaken like it was observed in wildtype animals. Therefore we predicted that if *in vivo* recordings are obtained from ION transected animals, the average amount of activity should be higher than in intact mice. Likewise, it suggested that the amount and quality of the activity in cortex isolé animals should be similar to wildtype animals. Both of these are experimentally testable predictions, and preliminary work in the Butt lab suggests that the prediction is correct for cortex isolé animals.

A limitation in the interpretation of the results presented here is that we only studied a very small subset of plasticity rules that exist, and it is not straightforward to infer how our results would change had we chosen a different STDP rule. For example, if we had used a weight-dependent STDP rule, we would not have obtained a bimodal distribution of weights, so we would have had to change the way that we evaluated the performance of the simulation in capturing experimental results (van Rossum et al., 2000). In addition to the two types of rules previously discussed, there are many additional variants that were not mentioned. For example, one can include a combination of short-term plasticity, or plasticity that affects the efficacy of synaptic transmission, which can interact with STDP to form different motifs (Vasilaki & Giugliano, 2014), voltage dependence (Clopath, 2010), and other forms of non-Hebbian plasticity (Zenke et al., 2015). However, caution should be exercised when choosing more complex models for plasticity, as many of these models are based on experimental data recorded from adult animals, and they may be different during development (Itami & Kimura, 2012; Jiang et al., 2007; Vickers et al., 2018)

In addition to the limited number of plasticity rules considered, our model used two homogeneous populations of neurons to represent the L4 SSNs and the L5b SST+ interneurons. In reality, however, barrel cortex contains more neuronal subpopulations that are key to development. For example, L4 PV interneurons begin

receiving thalamic input starting P7 (Daw et al., 2007), and that when the L5b-L4 loop disassembles, feedforward inhibition via the PV interneurons provides local inhibitory control over L4 SSNs that survives into adulthood. Competition could be mediated by a combination of factors that differentiates PV interneurons from SST+ interneurons. For example, one important way in which the inhibition these interneuron subtypes differ is that SST+ interneurons target the dendrites of excitatory neurons whereas PV interneurons target the soma (Tremblay et al., 2016), and as a result a presynaptic spike from a PV interneuron may have a different effect than a presynaptic spike from a SST+ interneuron on the processing and output of excitatory cells (Miles et al., 1996).

Extending the current model one could then ask under what conditions two inhibitory populations could compete to inhibit an excitatory population. In order to incorporate the differing effects of dendrite-targeting or somatic inhibition one would need to use multicompartmental neuron models with at least one compartment for the soma and one compartment for the dendrites (Gerstner et al., 2009, Ch. 6.4.4) , but a simpler modelling approach could be to divide the inhibitory population into two groups and, using the symmetric learning rule, choose slightly different values of the target firing rate ρ_0 for each group. Letting the target postsynaptic firing rates for these populations be $\rho_0^1 < \rho_0^2$. Then, for moderate amounts of input activity the inhibitory population with target ρ_0^1 would strengthen and since the postsynaptic firing rate r would be less than ρ_0^2, then the second inhibitory population's connections would weaken. However, if there was too much excitation and the inhibitory connections had a maximum weight, then both inhibitory connections would survive as we expect to be the case for the ION transected animals.

Finally, a second extension to the model that could be explored is the effect of varying the temporal dynamics of the input. The model used constant Poisson input to the network, but activity in the developing barrel cortex is known to be different. During the first postnatal week, the brain has short bursts of oscillatory activity interleaved with long periods of silence (Khazipov et al., 2013). Further-

more, there are different types of oscillatory activity in barrel cortex (Yang et al., 2009), and part of this activity is thought to be driven by thalamic activity (Minlebaev et al., 2011). This suggested that using a time-varying input would be more biologically realistic, and due to the sensitivity of the STDP rules to timing and firing rates, it would be interesting to understand how different types of activity affect the changes in connectivity in the network.

While we cannot fully answer the question of what the underlying mechanism driving connectivity changes in developing barrel cortex is, through the work presented here, we propose that activity-dependent mechanisms like STDP are plausible candidates. Furthermore, the interaction between the STDP rules and the amount of activity in the system reconcile seemingly contradictory results in the Butt lab, and these predictions can be tested experimentally.

Chapter 4

The functional role of transient circuits

4.1 Introduction

In the previous chapter we focused on the biological mechanisms that might be driving the large changes in connectivity that were observed by Marques-Smith et al. (2016) in the L5b-L4 loop. One key limitation of the strategy employed in Chapter 3 was that it ignored the larger context of the L5b-L4 loop, disregarding other key subpopulations like PV+ interneurons, and the fact that this cortical circuit was found in barrel cortex. Recent data from the Butt lab demonstrated that regardless of the age studied, an analogous L5b-L4 loop was not found in primary visual cortex (V1) (Stacey, 2018). In this chapter we focus on studying the L5b-L4 loop in the context of the functional changes that are happening in barrel cortex from P4 to P15 and ask whether it could play a role in the refinement of thalamocortical connections onto L4.

One of the features that distinguishes barrel cortex from all other cortical areas is the presence of 'barrels', where cells in L4 organise themselves around bundles of thalamic fibers that transmit input from single whiskers (Van der Loos & Woolsey, 1973; Erzurumlu & Gaspar, 2012). The critical period for structural plasticity ends

at P4, as whisker trimming after this date does not interfere with the formation of barrels in cortex (Fox, 1992; Erzurumlu & Gaspar, 2012). Between P5 and P7, barrels begin to emerge as cellular aggregates (Rice et al., 1985), and the dendrites of barrel cells become polarised towards the centre of the barrels (Espinosa et al., 2009). As a result, in the period of time when the L5b-L4 loop is present, barrels are arriving towards the end of their critical period, and connections between thalamus and L4 are maturing. Even though the majority of events related to the structural plasticity of thalamic connections onto L4 are ending during the first half of the week when the L5b-L4 loop is present, we wondered whether the translaminar connections played a role in refining the receptive fields of L4 neurons to thalamic input, and may help enforce the activity patterns required for the dendritic polarisation that occurs between P5 and P7 (Nakazawa et al., 2018).

Modelling cannot prove that the L5b-L4 loop participates in the final stages of barrel map refinement. However, in this chapter we find the conditions under which this inhibitory circuit could contribute to the refinement of thalamocortical connections. As such, the primary goal of this chapter is exploratory and will focus on the following two relationships: (i) the specificity of the inhibitory input on the excitatory receptive field formed by a putative L4 excitatory neuron, and (ii), the role, if any, of a feedback connection from the excitatory neuron onto the inhibitory neurons. Understanding the conditions that would enable the L5b-L4 loop to participate in barrel map refinement can ultimately help us formulate future experiments that could help us test these hypotheses and understand the role of transient inhibitory circuits during brain development.

4.2 Methods

The work presented in this chapter is an extension of a model proposed by Clopath et al. (2016). This model is the only model, to our knowledge, that has explored the role of interacting excitatory and inhibitory plasticity on receptive field formation.

Our model extended the model to the developing barrel cortex by (i) introducing activity that is similar to that measured in barrel cortex during development (Mizuno et al., 2018), and (ii) looking at the effect of adding a feedback excitatory connection onto the inhibitory neuron. The methods are summarised here and any differences from the original model are described in more detail.

4.2.1 Network model

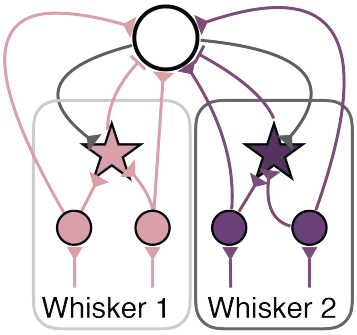


Figure 4.1: **Schematic of the network** In this figure we show a schematic of the network used. The grey boxes contain inputs from an individual whisker. Each one of these input streams was connected to inhibitory (star-shape) and an excitatory (circles) inputs onto a single output neuron. In different simulations we varied the tuning of the inhibitory cells to the inputs, and whether or not an excitatory feedback connection (grey) from the output neuron back to the inhibitory neuron was present or not.

The network studied has a single output unit that receives inputs from both an excitatory and inhibitory population. Unlike the network used by Clopath et al. (2016) the network used included a feedback connection between the output neuron and the inhibitory neurons to make it more similar to the connectivity in the L5b-L4 loop (Figure 4.1). The output neuron was modelled as a threshold-linear neuron, and the activity of the network was represented using firing rates rather than spike times. The rate of the output neuron was the following:

$$R(t) = \left[\sum_{i=1}^{N_E} W_i^E E_i(t) - \sum_{i=1}^{N_I} W_i^I \left[I_i(t) + w_b R(t-1) \right] \right]_+ , \tag{4.1}$$

where $E_i(t)$ and $I_i(t)$ were the time dependent activity of the ith excitatory input

and the ith inhibitory input respectively. $i \in [1, N_E]$, and W_i^E and W_j^I were the weights of the excitatory and inhibitory synapses. w_b is the weight of the excitatory feedback weight onto the inhibitory input neurons. $[\cdot]_+$ sets all negative values to zero so that the output neuron's firing rate is nonnegative. All weights were constrained to be positive so that excitatory neurons always exerted an excitatory effect on the output neuron and the inhibitory neurons remained inhibitory throughout the simulation. A network diagram for the network we used can be seen in Figure 4.1, where the circles are excitatory rate neurons, and stars are inhibitory rate neurons. The light grey boxes represent groups of neurons that receive inputs from a particular whisker, and the neuron is coloured according to the input stream or whisker input that the unit responded to. The output neuron received inputs from all of the excitatory and inhibitory neurons in the network, and the connection weights determined the response strength of the output neuron to particular whiskers (*i.e.*, groups of input streams), or particular input streams. The diagram shows two groups with two input streams each, but in simulations six groups with ten input streams each were utilised.

4.2.2 Input streams

The original model is a rate model used noisy, sparse, uncorrelated inputs generated from an Ornstein-Uhlenbeck process with its negative values set to zero. An Ornstein-Uhlenbeck process is described by

$$\tau_m dx_t = (\mu - x_t)dt + \sigma dW_t, \tag{4.2}$$

where $\tau_m = 50$ms is the time constant, $\mu = 0$ is the drift term, $\sigma = 1$ is the standard deviation of dWt which are increments of the Wiener process with mean 0 and variance proportional to dt. After generating an Ornstein-Uhlenbeck process $u(t)$, we generated the input stream, $I(t)$, by subtracting $c = 0.1$ to increase the sparseness

of the activity, and setting all negative values to 0 as follows,

$$I(t) = [u(t) - c]_+ . \tag{4.3}$$

Activity generated this way is continuously active (Figure 4.2A). In order to better capture key aspects of barrel cortex such as the activity observed during the first few postnatal weeks, we made several modifications. Since we know that each whisker is individually innervated by $\sim$ 100 nerves (Rice et al., 1986; Ebara et al., 2002), and that during the first postnatal week the spontaneous activity recorded in mice is sparse and spatially forms a patchwork pattern (Yang et al., 2009; Khazipov et al., 2013; Mizuno et al., 2018), we changed the input to the network by grouping the input streams into $G = 6$ groups each with $N_G = 10$ input streams for a total of $N_{\text{tot}} = N_G \cdot G = 60$ input streams (Figure 4.2B). In the course of the simulation, each group was selected at random to be active for an event duration T_E. During this time any input stream, n, belonging to the active group is allowed to be active, and all other input streams were set to zero. For all the simulations, unless stated otherwise, groups were not co-active. The network received inputs from S input streams that are divided into G groups. Each group was active for one second at a time, and the order in which the groups were active was random.

4.2.3 Tuning of the input excitatory and inhibitory populations

We explored the way in which the receptive field of the inhibitory neurons affect the excitatory cell's receptive field, and whether it made a difference to add a feedback excitatory connection. To study this, we assumed that the tuning of excitatory and inhibitory input neurons were fixed throughout the simulation.

We assumed, for simplicity, that each excitatory input corresponded directly to one input stream. In other words, there were no excitatory inputs that responded to input streams in different groups. Although this is unlikely to be the case in the biological system, this simplifying assumption allowed us to study the effect of the

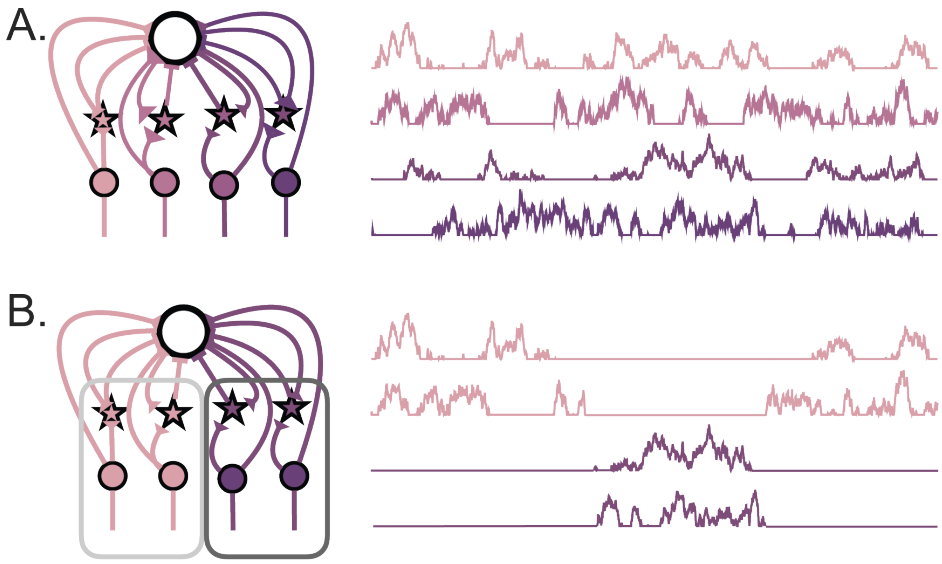


Figure 4.2: **Input activity.** The network is fed with activity generated from a rectified Ornstein-Uhlenbeck process. **A:** Continuous uncorrelated input streams of the sort used by (Clopath et al., 2016). **B:** Input streams were divided into groups of neurons corresponding to sensory inputs coming from the same whisker. Inputs were only allowed to be co-active with other input streams in the same group.

inhibitory connectivity on the selectivity, or lack thereof, acquired over time by the output unit.

The tuning of the inhibitory neurons was varied from responding equally to all input streams (unspecific), to having one inhibitory neuron per input stream (one-to-one). Tuning motifs with intermediate levels of specificity were also explored, including the case where the inhibitory neuron responded to all input streams within a whisker group, and the case where the inhibitory neuron had a Gaussian tuning to input streams with a particular mean and standard deviation. We describe each one in more detail below and summarise them in Table 4.1.

4.2.3.1 Unspecific tuning

In the case of unspecific tuning we assumed that the inhibitory neuron would be equally active regardless of the input stream. Since the overall mean activity in the network stayed constant throughout the simulation, this tuning profile was equivalent to the profile of a single inhibitory population that was tonically active in response to amount of background activity in the circuit. As a result, this inhibitory

Tuning type	# Units	Firing rate	Figure
Unspecific	1	0.1	Fig 4.3
Unspecific w/ feedback	1	$0.1 + R(t)$	Fig. 4.4
Group	G	$\sum_{i=1}^{N_G} E_i^{(g)}(t), g \in [1, G]$	Fig. 4.5
Group w/ feedback	G	$\sum_{i=1}^{N_G} E_i^{(g)}(t) + R(t), g \in [1, G]$	Fig. 4.6
Gaussian	2G	$\sum_{i=1}^{N_E} E_i(t) f(x),\ f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x-\mu^2}{2\sigma^2}}$	(*not shown*)
Gaussian w/ feedback	2G	$\sum_{i=1}^{N_E} E_i(t) f(x) + R(t)$	Fig. 4.7
One-to-one	N_{tot}	$E_i(t)$	Fig. 4.8
No Inhibition	0	–	Fig. 4.9

Table 4.1: Summary of inhibitory tuning studied.

motif was modelled as a single inhibitory neuron with a constant firing rate but whose connection to the output neuron was plastic.

4.2.3.2 Group-specific, stream-unspecific tuning

One of the features of our model was that the inputs are split into groups. We explored the possibility that instead of an inhibitory population that was tonically active regardless of the source of the input, that it would be active in response to all the input streams belonging to a single group. As a result, in this case we had G inhibitory neurons, each one receiving inputs from a single group.

4.2.3.3 Gaussian

Although it is possible in a cortical area like barrel cortex that neurons could be specific to inputs from a single whisker, it is more likely that each neuron would receive inputs from mostly one whisker but also some inputs from other whiskers. We implemented this by using a Gaussian tuning instead of a Group-specific tuning. The centres for the Gaussian tuning were distributed equally across all input streams, and had variance $\sigma^2 = 25$.

4.2.3.4 One-to-one

Finally, we considered a network where the inhibitory inputs responded specifically to a single input stream. As a result, there would be an inhibitory neuron that received exactly the same input as each excitatory neuron. In this case, it was possible for excitatory and inhibitory inputs onto the output excitatory neuron to balance each other out (Clopath et al., 2016).

4.2.4 Weight initialisation

The weights the strength of the connections between the input neurons (both excitatory and inhibitory) onto the output neuron. The inhibitory weights, W_i^I, were always initialised randomly, so each weight was drawn from a uniform distribution $w \sim U[0,1]$, and divided by the sum of all the weights so that their sum would equal to one. For most of the chapter, unless stated otherwise, the excitatory weights were initialised to be Gaussian, centred around input stream 30, with variance $\sigma^2 = 16$. Where explicitly stated in the text, a random initialisation was used for the excitatory neurons as well.

4.2.5 Plasticity rules

The network has two interacting plasticity rules. The excitatory plasticity rule had the form

$$\Delta w(t) \propto R(t)x(t) \quad (+ \text{ normalization}), \tag{4.4}$$

where R is the firing rate of the output neuron, and x is the firing rate of the excitatory input neuron. We used divisive normalization of the weights at each step.

The inhibitory plasticity rule had the form,

$$\Delta w(t) \propto (R(t) - \rho)x(t), \tag{4.5}$$

where ρ was the target firing rate and $x = I(t) + w_b R(t-1)$ was the inhibitory input

firing rate. If there was a feedback connection, w_b was set to 0.5, and in that case, the input firing rate would be dependent on the output firing rate at the previous step.

The learning rate used for the excitatory learning rule was 0.01, and for the inhibitory learning rule was 0.1

4.2.6 Modelling the formation of barrels

In this variation of the network, we used eight output neurons instead of just one. The excitatory weights were initialised to be Gaussian, as before, but each output neuron had a different centre (see Figure 4.11B).

4.3 Results

4.3.1 Unspecific inhibition

We first looked at a network in which there was a single inhibitory unit with a constant firing rate (Figure 4.3A). The weights of most excitatory inputs except for those corresponding to a single whisker weakened (4.3B). The weight for the single inhibitory input remained approximately stable throughout the simulation (Figure 4.3D). The receptive field of the output neuron was obtained by looking at how much it responded to each excitatory input. In Figure 4.3D, we plotted the receptive field at ten points in the simulation. The first row shows the gaussian initialisation of the excitatory weights at the beginning of the simulation, and we observed that the output neuron forms a stable representation of the fourth whisker group in the simulation, although it showed no consistent preference for any of the input streams belonging to that whisker. In Clopath et al. (2016), unspecific inhibition was observed to impose competition between the inputs and lead to the robust development of receptive fields to a single excitatory input neuron, whereas here, the output neuron refined its initial Gaussian tuning to a receptive field that was

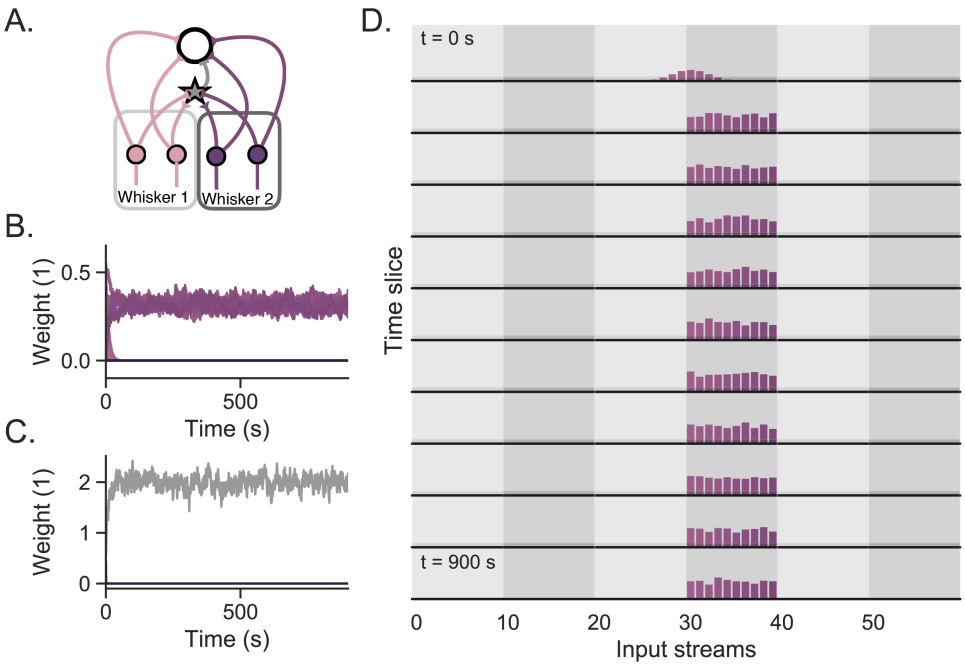


Figure 4.3: **Unspecific inhibition without feedback.** (**A:**) Schematic of the network connectivity for two (out of six) groups each with two (out of ten) input units. Each input is fed directly to a single excitatory input neuron which in turn sends inputs to a single inhibitory neuron (grey star) and an output excitatory neuron. The inhibitory neuron is connected to the output neuron and the connections from the excitatory and inhibitory neurons to the output neuron are plastic. (**B:**) Excitatory weights over time. Each line represents the weight of the excitatory connection strength from a single input neuron. (**C:**) Inhibitory weights over time. This plot has a single weight as we simulated a single interneuron. (**D:**) The evolution of the response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron responds equally to all input neurons from a single whisker within the first 50 seconds.

specific to a single whisker group but had no preference to any of the input streams belonging to that group.

One of the features of the L5b-L4 loop is that there is a feedback excitatory input onto the L5b inhibitory neuron. As a result, we added a feedback excitatory connection from the output neuron onto the inhibitory neuron (Figure 4.4A). The feedback connection was implemented by adding a weighted firing rate of the output unit to the inhibitory unit's firing rate. The excitatory and inhibitory weight changes throughout the population were qualitatively identical to the case without feedback (Figure 4.4B,C). Likewise, the evolution of the receptive field was very

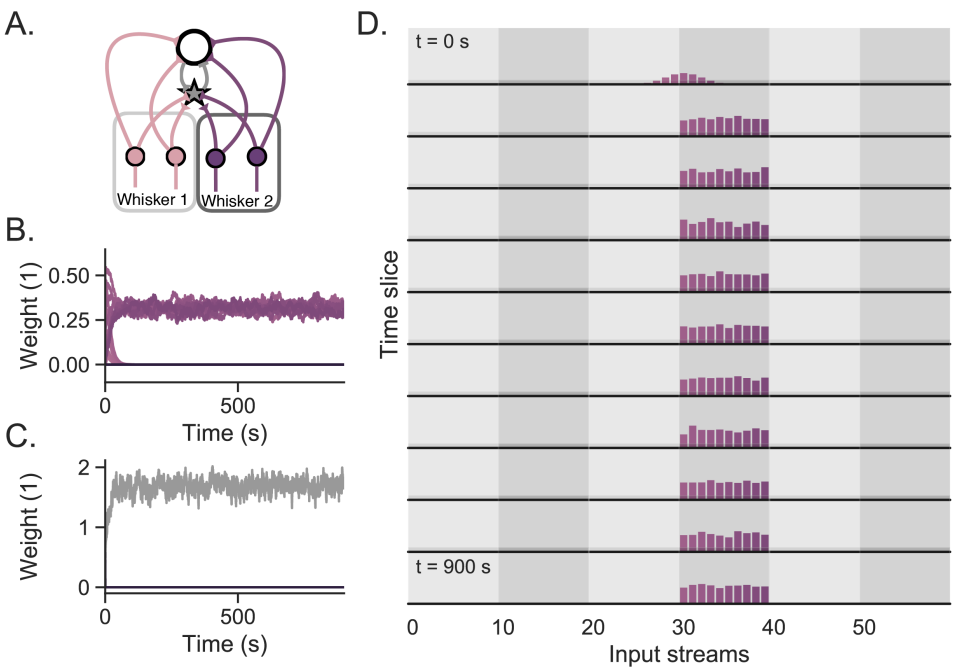


Figure 4.4: **Unspecific inhibition with feedback.** (**A:**) Schematic of the network connectivity with a single inhibitory neuron that responds equally to all input streams. It also receives a feedback connection from the output excitatory neuron. (**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The evolution of the response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron responds equally to all input neurons from a single whisker within the first 50 seconds.

similar to the receptive field of the output neuron in a network without excitatory feedback (Figure 4.4D).

4.3.2 Group inhibition

In the next set of simulations, we had a single inhibitory unit for each whisker input. We hypothesised that since the activity of the input streams belonging to a single whisker was identical to the activity of the input streams in Clopath et al. (2016), that by using an unspecific inhibitory input within each whisker might lead to the output neuron forming a receptive field to a single input stream. We first tried the network with no feedback excitation (Figure 4.5A). In this network, the output neuron no longer formed a stable receptive field. The excitatory weights from units belonging to the same group strengthened and weakened together, and there

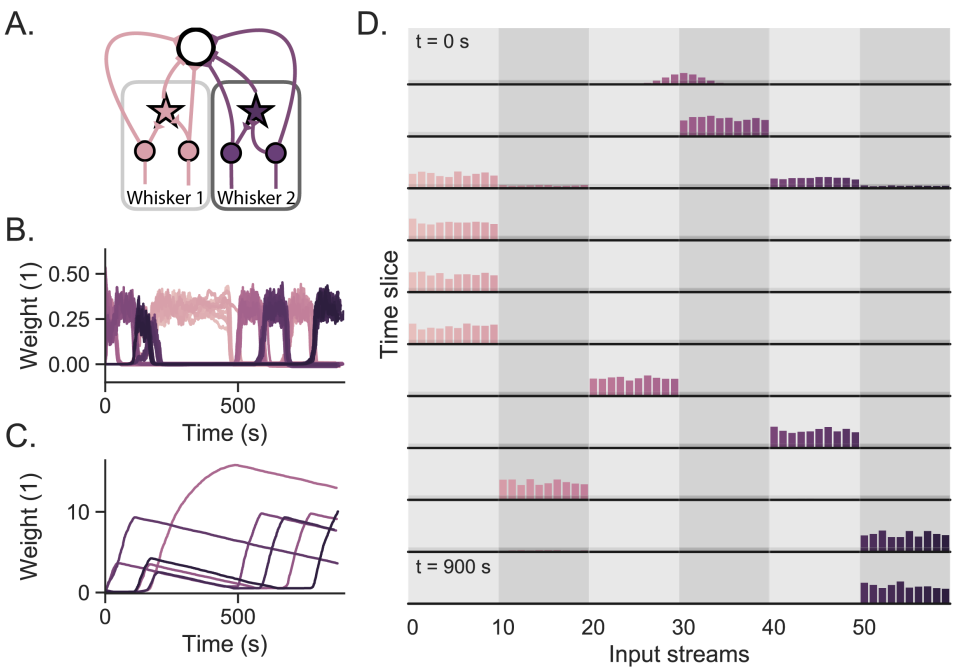


Figure 4.5: **Group Inhibition without feedback.** (**A:**) Every excitatory neuron in a given group is connected to a single inhibitory input unit. Shown are two groups (out of six simulated) each with two (out of ten) input units. (**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The evolution of the response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron responds to a single whisker at a time, but does not form a stable representation of any whisker.

were sharp transitions whereby different groups were strengthened (Figure 4.5B). The changes in the inhibitory weights were less sharp, and instead increased when the excitatory weights were strong and began a slow decrease after that group's excitatory connections had weakened (Figure 4.5C). The sharp transitions in the excitatory connection weights was reflected by the receptive field snapshots as at any given time slice the output neuron responded to a single group (Figure 4.5D). We noted that the next excitatory group to strengthen was always a group whose corresponding inhibitory neuron had weight zero. As a result of the weak inhibitory connection, the excitatory cells corresponding to that whisker would have been able to activate the output neuron more than the other inputs, and due to the normalisation of the excitatory weights led to a shift in the receptive field. The rates of change in the weights is slower for the inhibitory connections than it is for the excitatory

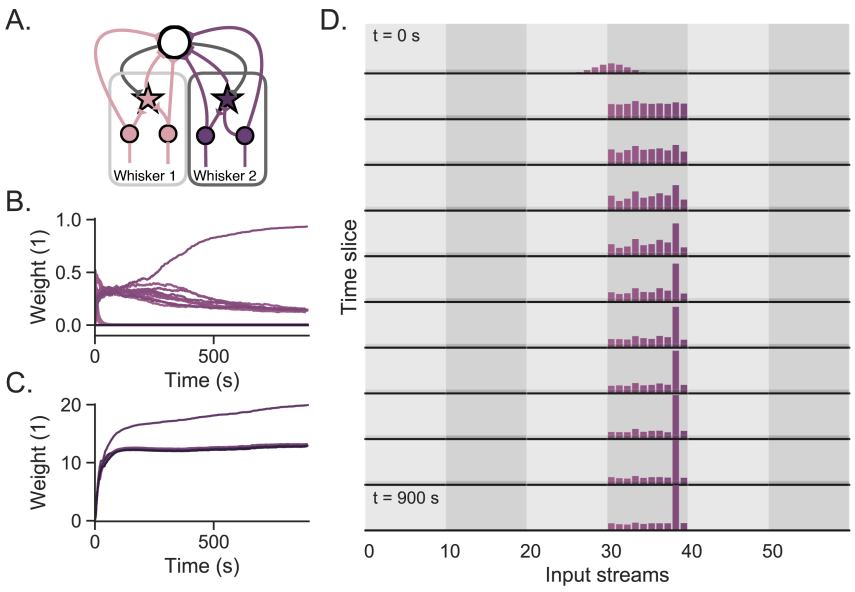


Figure 4.6: **Group Inhibition with feedback.** (**A:**) Every group is represented by one inhibitory input unit that receives inputs from all inputs streams in that group, and a feedback connection from the output neuron. Shown are two groups (out of six simulated) each with two (out of ten) input units. Each inhibitory neuron receives inputs from (**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The evolution of the response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron responds equally to all input neurons almost from the beginning of the simulation, and over the course of the simulation develops a stronger response for a single input stream in that whisker.

connections, which may seem contradictory given that the inhibitory learning rate was ten times higher than the excitatory learning rate. However, this may be due to the divisive normalisation that was in place only for the excitatory connections, but this needs to be examined more carefully in future work.

The addition of an excitatory feedback connection to the network stabilised the receptive field of the output neuron (Figure 4.6A). Within the first 100 seconds of the simulation, most excitatory connections had gone to zero, by 200 seconds, there was a single excitatory weight that strengthened more than the other connections (Figure 4.6B,D). Even though the output neuron responded primarily to a single input stream, it still responded more strongly to other input streams in the same whisker compared to input streams in other whiskers (Figure 4.6B,D). On the other hand, for the inhibitory units, the connection for the inhibitory neuron receiving

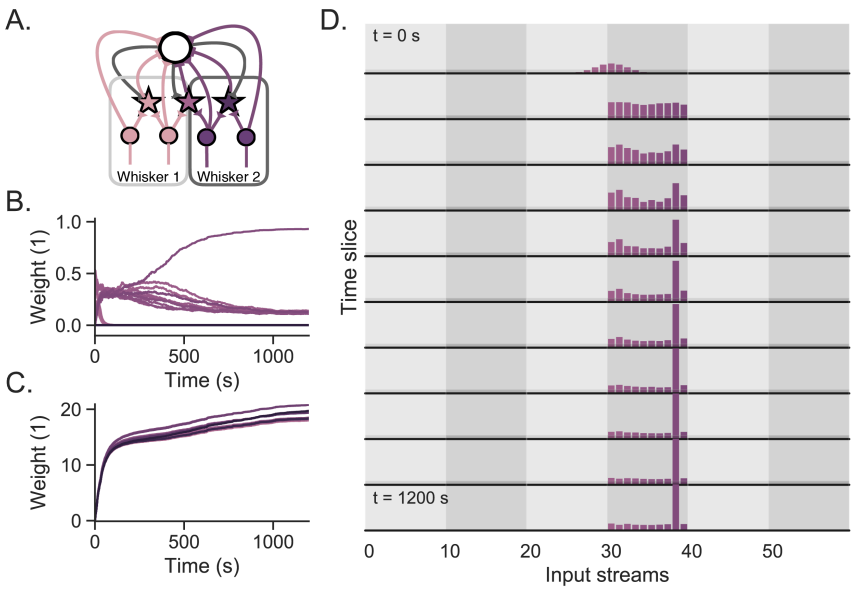


Figure 4.7: **Gaussian Inhibition with feedback.** **A:** Each inhibitory input unit in this network has a Gaussian tuning over input streams. Shown are two groups (out of six simulated) each with two (out of ten) input units. As each inhibitory unit has a Gaussian tuning, here we represent this by adding a third unit that sits between the two whisker groups. The feedback excitatory input is shown in grey. (**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron forms a representation of a single input stream from a single group.

inputs from the whisker that the output neuron responded to kept strengthening throughout the simulation, although the rate at which this happened slowed down considerably after 100 seconds (Figure 4.6C). All inhibitory connections remained positive throughout the simulation.

4.3.3 Gaussian inhibition

One criticism with group inhibition is that if the inhibitory units have already developed a selectivity for a single input group, there is no logical reason why the same would not be the case for the output neuron. As a result, we relaxed the structure of the inhibition to have a Gaussian tuning for input streams regardless of their whisker or group identity. We found that the results were very similar to the Group tuning, but only show the results for Gaussian tuning with feedback (Figure

4.7). We illustrate the Gaussian nature of inhibitory neurons by drawing an interneuron that is shared between two neighbouring whisker groups (Figure 4.7A). We found that the output neuron formed a receptive field for a single whisker and to a single input stream within that whisker group (Figure 4.7B), but that the inhibitory weights seem to continue to increase together throughout the simulation 4.7C). Throughout the simulation we find that the output neuron first responds equally to all input streams in a group and then proceeds to slowly respond preferentially to a single input stream in that group. This result suggests a function for the L5b-L4 loop in S1BF. Assuming that the L5b SST+ interneurons respond unequally to different whisker stimuli, even in the case when they respond to more than a single whisker stimulus, the reciprocal translaminar connection could help refine thalamic connections onto L4 SSNs in barrel cortex.

4.3.4 Edge cases

At the other extreme of unspecific inhibition is perfect one-to-one inhibition, where each inhibitory unit receives input from precisely one input stream (Figure 4.8A). We found that even when we initialised the excitatory and inhibitory weights to have a Gaussian tuning, this tuning was lost within the first 100 seconds of the simulation (Figure 4.8B,C). Once the tuning has been lost, there seems to be little variability in the output neuron's unspecific tuning throughout the simulation (Figure 4.8D) This is consistent with the results obtained by Clopath et al. (2016), where the case when inhibition perfectly balances out excitation, then no receptive field forms regardless of whether it had an initial tuning. We see that this is still true in our model, in spite of the activity being structured differently.

Finally, we wondered what would happen to the receptive field where the output neuron received no inhibition (Figure 4.9A). We found that if the initial Gaussian tuning was centred in between two groups, the output unit would respond only to these two groups, and would initially oscillate between responding more to one group than the other until eventually responding only to a single group (Figure

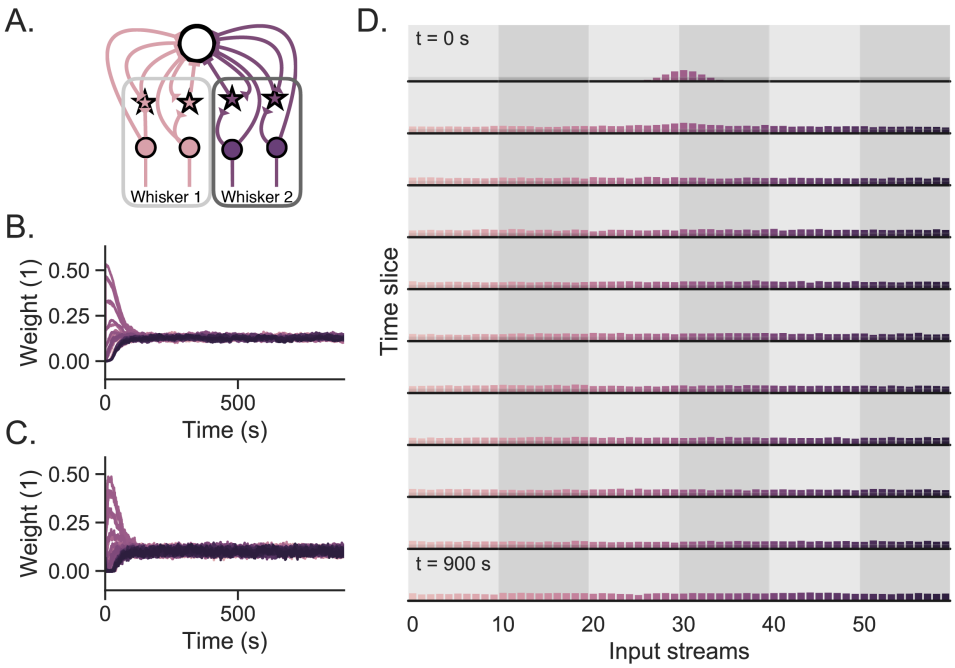


Figure 4.8: **One-to-One inhibition** In **A:** The network was simulated with each inhibitory neuron for each excitatory neuron. As each inhibitory unit corresponds to a single input stream, here each one is coloured with its respective colour. (**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The response of the output neuron to each input stream at ten timepoints of the simulation. At the end of the simulation, the output neuron responds equally to all input streams.

4.9B). Although this selectivity to a single whisker is what we ultimately want the output unit to develop, the initial instability in its selectivity between groups makes it suboptimal for barrel cortex where barrels have a well-defined spatial structure and neurons belonging to a particular barrel reliably respond to one whisker.

4.3.5 Random

So far, the excitatory weights onto the output neuron were initialised with a Gaussian tuning that was centred at the intersection of two whisker groups. Although this setup allowed us to probe the effect of changing the inhibitory tuning on the output receptive field, receptive field dynamics could change depending on the initial conditions. We predicted that whenever the output neuron's formation of a stable receptive field was independent of initial conditions, but that the initial tun-

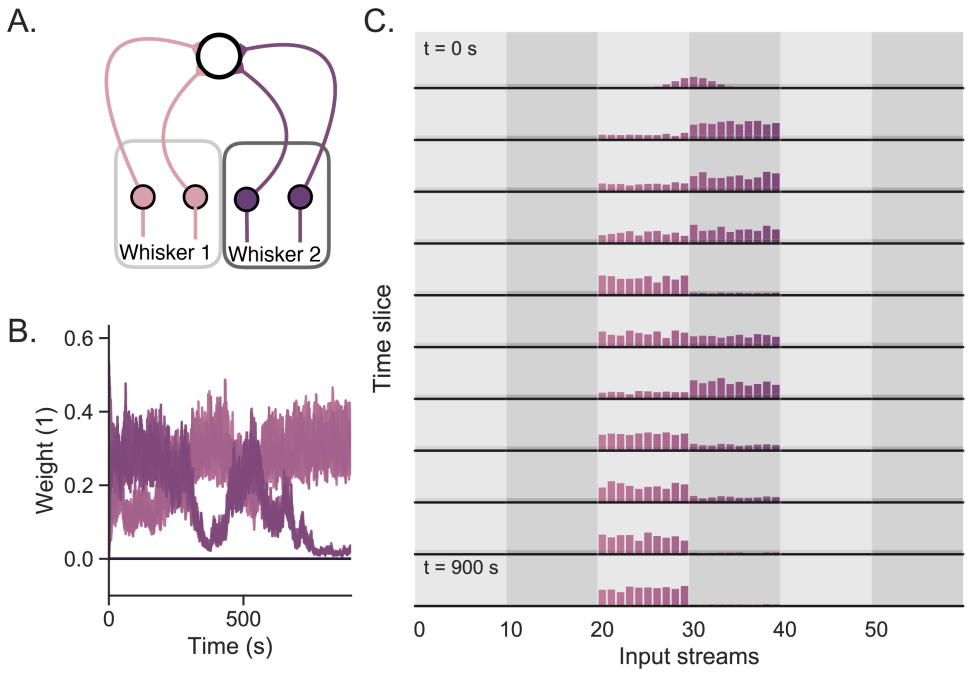


Figure 4.9: **No inhibition A:** This is a purely feedforward excitatory network with no inhibition.(**B:**) Excitatory weights over time. (**C:**) Inhibitory weights over time. (**D:**) The response of the output neuron to each input stream at ten timepoints of the simulation. The output neuron's tuning oscillates between the two groups it was initially tuned to.

ing would bias the output neuron to form a receptive field to a particular whisker.

To probe this question the network was initialised with random excitatory connection weights for the case when inhibitory neurons had a Gaussian tuning with and without excitatory feedback. At the end of each simulation whiskers were ranked according to how much the output neuron responded to each group. Furthermore, we kept track of the whisker that the output neuron responded to the most at each time step of the simulation. We plotted the percent of the output's response that belongs to each whisker where Group 1 is the whisker that showed the maximum response in a network with Gaussian inhibitory tuning and no feedback excitation (Figure 4.10A). We observed that most of the simulations responded primarily to the first whisker, but in some simulations, the response was split between the top two groups. When we tracked the identity of the whisker that the output neuron responded most strongly to throughout the simulation, we found

that the output neuron responded strongly to most groups throughout the course of the simulation (Figure 4.10B). This is consistent with our observation that without feedback excitation the receptive fields were unstable, but that the output neuron quickly switched between responding from one whisker to another one.

This analysis was repeated for the Gaussian inhibitory tuning with feedback excitation and found that at the end of the simulation, the output neuron responded only to the top group (Figure 4.10C), and that the output neuron responded to the top group for most of the simulation (Figure 4.10D). Therefore, the relationship between the inhibitory response tuning and the output neuron's receptive field stability did not depend on the initial weight configurations.

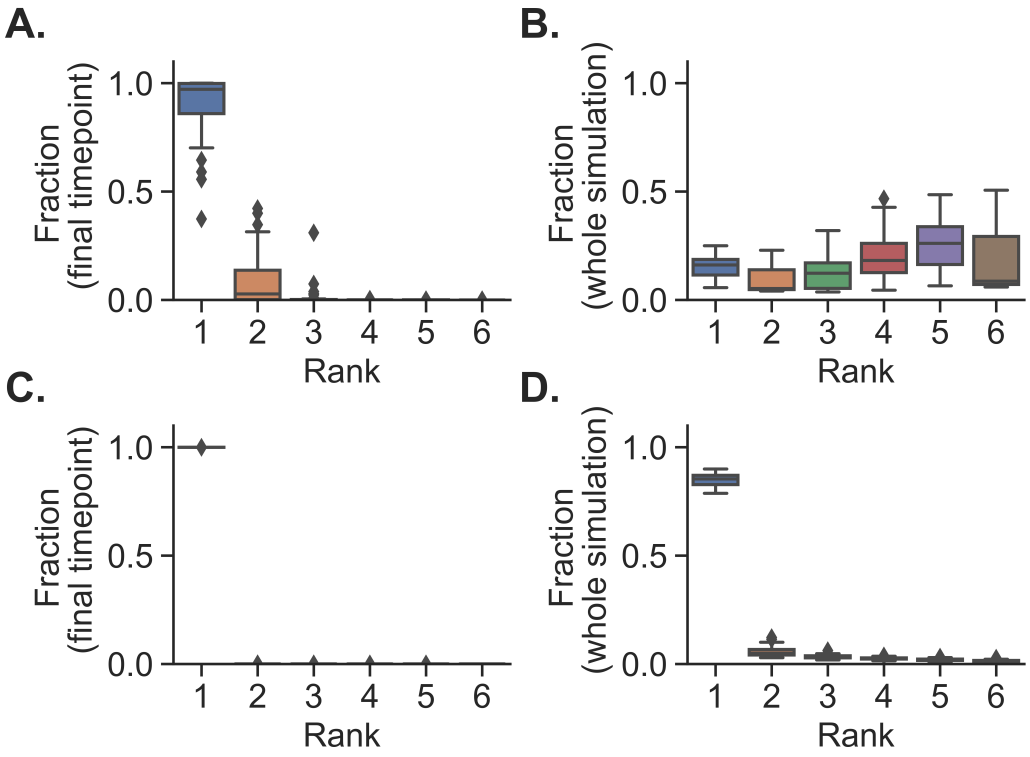


Figure 4.10: The effect of random tuning. This figure shows the result of 1000 simulations for a network with gaussian inhibitory tuning with and without feedback. **A,C:** The proportion of the output neuron's response to each whisker group at the end of a simulation. Here, group 1 on the x-axis marks the whisker group the output neuron responded to the most and the 6 is the whisker it responded to the least for a network with a feedback connection and one without respectively. **B,D:** The proportion of the output neuron's response to each whisker group integrated over the entire simulation.

4.3.6 Forming a barrel field

We showed that the receptive field of the output neuron depended on the response tuning of the inhibitory population but was independent from the initial weights of the excitatory inputs. Nonetheless, initialising the system with a Gaussian tuning reliably biased the output neuron to develop a preference for a particular group. These results do not demonstrate that a network with multiple output neurons would form discrete barrel-like modules. We modified our network to have multiple output neurons (Figure 4.11A). The only difference between output neurons was that their initial excitatory weight matrix had the Gaussian centred around a different input stream.

The simulation was run for 1000 seconds, and the initial and final excitatory and inhibitory receptive fields for each output neuron are plotted in different colours (Figure 4.11B). All of the neurons responded to a single whisker at the end of the simulation, and neurons with similar initial weight matrices were more likely to form a receptive field to the same whisker or a neighbouring whisker. However, output neurons that responded to the same whisker input formed similar receptive fields, instead of learning to respond to different input streams within that barrel. The fact that neurons responding to the same whisker had a very similar tuning to the input streams in that whisker was problematic, however, as we want different neurons in the same barrel to respond to different features of the sensory input. The problem seems to be that output neurons all receive the same input and the only difference between the output neurons is the initial tuning to the excitatory input streams. This suggests that there must be a mechanism to change the input slightly between neurons in the same barrel, which could happen if they received inputs from slightly non-overlapping inhibitory and excitatory populations or if there was lateral inhibition.

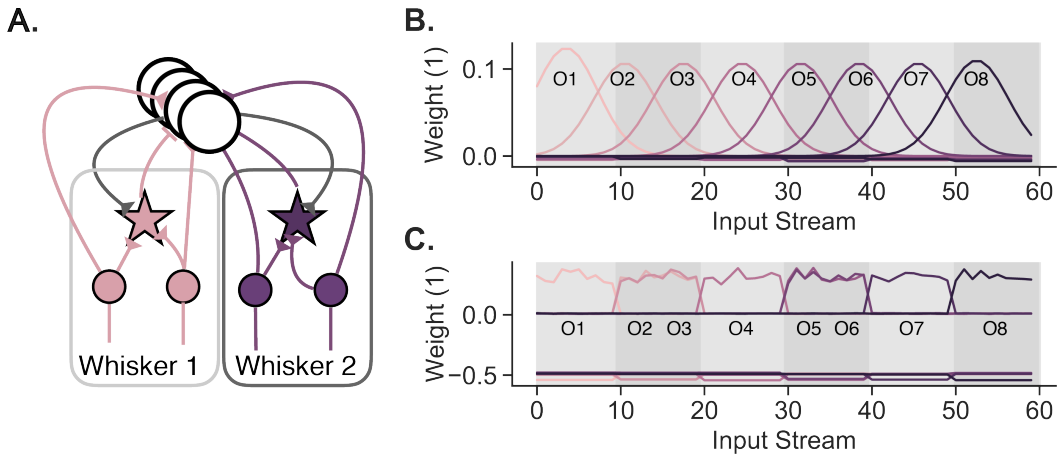


Figure 4.11: **Simulations of multiple output neurons still form group-specific responses.** (**A:**) Eight output neurons were added to the network to see whether different output neurons could form preferences to different whiskers. (**B:**) The excitatory (positive) and inhibitory (negative) weights onto each output neuron (labelled O1 to O8), each one represented in a different colour. Each output neuron had an initial Gaussian tuning centred around different input stream. (**C:**) shows the tuning of each output neuron at the end of the simulation. Output neurons that respond to the same group form similar responses to all ten input streams in that group.

4.4 Discussion

Our model suggested that even in the absence of all inhibition, if the activity coming from each whisker is temporally non-overlapping, the output neuron will learn to respond preferentially to a single whisker. However, the output neuron becomes selective to a single feature of a particular whisker when the inhibitory inputs are tuned to specific whiskers or have a Gaussian tuning, and there is feedback excitation onto the inhibitory neuron. Mitrukhina et al. (2015) showed that neurons in barrel cortex respond to multiple whiskers in P0 and P1, and these responses become confined to one or two whiskers by P7 in rats. The refinement in responses is accompanied by structural changes in the organisation of thalamic afferents in barrel cortex (Agmon et al., 1993; Rebsam et al., 2002). Furthermore, Figure 5 in Agmon et al. (1993) shows that while thalamic afferents seem to cluster into groups in L4, they are much more diffuse in L5b. This suggests that while L4 neurons would develop specific responses to a single whisker, L5b SST+ interneurons may

get inputs from more than one whisker providing support for the Gaussian tuning to whisker inputs scenario we modelled. As a result, we propose that the L5b-L4 loop, in conjunction with patchwork-like activity may contribute to the polarisation of dendrites and refinement of excitatory cell responses to whisker inputs from P4 to P7.

The interaction of excitatory and inhibitory plasticity leads to a wide range of behaviours that are not obvious from the individual rules alone. This chapter adds to the already diverse behaviours described by Clopath et al. (2016). A full analysis of the parameter dependencies of the system we proposed in this chapter are required in order to fully understand why feedback excitation stabilises excitatory receptive fields. Our results suggest that when inhibitory responses are less specific than excitatory responses, in the presence of excitatory feedback connections onto inhibitory populations, the excitatory receptive field is stabilised and further refined. As a result, in this chapter, we delineated a set of conditions under which the L5b-L4 loop could play a functional role in refining barrel cortex responses to sensory stimuli.

Part II

Synaptic plasticity in Down Syndrome

Chapter 5

When development goes wrong: A mouse model for Down Syndrome

5.1 What is Down Syndrome?

Down Syndrome in humans is usually caused by a partial or complete triplication of chromosome 21, which contains between 200 and 300 genes depending on the report (Asim et al., 2015; NIH, 2019). Accordingly, individuals with Down Syndrome tend to have a wide variety of health problems ranging from heart disease to intellectual disabilities and reduced life expectancy (Diamandopoulos & Green, 2018). It is unclear how the genotype of Down Syndrome maps onto the wide variety of phenotypes that are observed as this relationship depends not only on the identity of the genes that are being triplicated, but possibly also changes in gene expression, epigenetic factors, and gene regulation (Patterson, 2009; Lyle et al., 2008). Nonetheless, some genes can be linked to specific aspects of the pathophysiology of Down Syndrome. For example, the triplication of the gene APP (amyloid protein precursor) is thought to be linked to the high incidence of Alzheimer's Disease in individuals with Down Syndrome (Zis & Strydom, 2018).

In the 1950s, the life expectancy for Down Syndrome was less than 20 years of age, and this increased to 50 years by the 2000s (Bittles & Glasson, 2004), and

the research priorities with regards to Down Syndrome have shifted accordingly. A longer life expectancy has resulted in a growing demand for treatments that can improve quality of life, such as treatments to ameliorate the learning difficulties faced by individuals with Down Syndrome. There are some promising therapy strategies that have begun to emerge from animal research, and there is hope to use these with humans in the near future (Patterson, 2009). Nonetheless, the mechanisms underlying these learning difficulties are still not well understood, and these will be the focus of the rest of the thesis.

5.2 Mouse models of Down Syndrome

There are practical and ethical reasons that make it difficult to do research on people, so a lot of effort has been devoted to creating mouse models of Down Syndrome (DS). Obvious challenges with creating a mouse model of DS include the large number of triplicated genes that cause the disease, and the fact that some of the phenotypes associated to DS will be unique to humans (Liu et al., 2011). However, there are large regions that have been largely conserved across the mouse genome and the human Chromosome 21 (Hsa21) which has enabled the creation of several mouse models of DS (Reeves et al., 1986). Many studies have demonstrated phenotypic parallels between these mouse models and individuals with DS (Patterson, 2009). Through work on mouse models, we hope to uncover biological mechanisms that can lead to new pharmacological treatments (Liu et al., 2011).

Most of the different mouse strains used to study DS are trisomic in a subset of the mouse genes corresponding to HSA21, which are largely located in the mouse Chromosome 16 (MMU16), but also MMU17 and MMU10 (Reeves et al., 1986; Liu et al., 2011; Gupta et al., 2016). Of the vast majority of studies around these mice models have been centred around a particular strain called Ts65Dn, which is trisomic for around 90 coding genes that are HSA21 orthologues. However, they also have an extra segment of genes that are not DS-related (Herault et al., 2017; Liu

et al., 2011; Patterson, 2009; Reinholdt et al., 2011). There are other mouse models of Down Syndrome, but they have been less studied (Gupta et al., 2016). Future studies should compare the performance of these different models across a set of behavioural and electrophysiological studies but in this thesis we focused on the Ts65Dn mouse model of DS.

In addition to the genetic similarities between the Ts65Dn animals and Down Syndrome, in order to be a good model for DS, these animals have to exhibit some of the behavioural phenotypes of DS. Although there are some behaviours that cannot be studied in mice such as articulate language communication, there are others involving memory consolidation, motor skills, and learning for which there are standard behavioural tests in mice. In standard tasks like spontaneous alternation, the Morris Water maze, and the radial arm maze, Ts65Dn mice show increased locomotor activity and impaired memory (Faizi et al., 2011; The Jackson Laboratory, 2019). Furthermore, although they do not show any large impairments in social behaviours, they do show impairments in social learning (Powers et al., 2018).

Even though genetic modifications in mice that are similar to those in DS result in analogous behavioural phenotypes, it still does not explain how the genetic modifications lead to changes that produce the phenotype. A way to bridge this gap is to study an intermediate stage, and in the brain, this means looking at the network of synaptic connections and electrophysiological characteristics of Ts65Dn in comparison to their wildtype counterparts, in order to identify therapeutic targets that can eventually be translated to treatments in humans.

5.2.1 Altered plasticity and inhibition in Ts65Dn mice

Since most people with Down Syndrome suffer from learning impairments, changes in synaptic plasticity have been a central part of DS research. Understanding how learning happens differently in DS is complicated by the fact that the changes effected by DS are global, are present throughout development, and involve the interaction of many different factors. One of the key components required for learning

in the brain is the ability for synapses to change their strength, and one of the key brain regions for learning and memory is the hippocampus. Behavioural and electrophysiological studies have shown that Ts65Dn animals show changes in plasticity in hippocampus (e.g. Hyde et al., 2001; Liu et al., 2011).

It has been suggested that excessive inhibition is the cause of impaired LTP in Ts65Dn animals (Contestabile et al., 2017). Not only do these animals present fewer and larger excitatory synapses (Kurt et al., 2000; Belichenko et al., 2004), they also have an increased number of certain inhibitory neurons due to the overexpression of *Olig1* and *Olig2*, which are two genes involved in setting the number of inhibitory neurons during brain development (Chakrabarti et al., 2010; Rueda et al., 2012). Costa & Grybko (2005) showed that adding picrotoxin, which is a $GABA_A$ receptor antagonist, rescues LTP in Ts65Dn animals, and this is accompanied with an improvement in the novel object recognition task (Fernandez et al., 2007), and in long-term object memory (Colas et al., 2017). In addition to $GABA_A$-receptor antagonists, using $GABA_B$-receptor antagonists can also rescue long-term potentiation (Kleschevnikov et al., 2012).

The inhibition disregulation in DS is complicated by recent results that suggest that GABA can, through the activation of $GABA_A$, effect a depolarising change in the postsynaptic membrane potential caused by the upregulation of a sodium-chloride transporter NKCC1 (Deidda et al., 2015). When this transporter is present, it increases the internal chloride concentration, so when $GABA_A$ receptors open, instead of chloride entering cells, chloride leaves the cell causing a depolarising change in the cell membrane (Ben-Ari et al., 2007). The shift from depolarising to inhibitory GABA occurs due to an activity-dependent up-regulation of KCC2, which is a K^+/Cl^- cotransporter type 2 that extrudes chloride from a cell (Ouardouz & Sastry, 2005). Furthermore, adding bumetanide, a substance that blocks NKCC1 transporters and restores the inhibitory effect of GABA, also rescues LTP (Deidda et al., 2015). Although it is evident that inhibition is related to the plasticity impairment in Ts65Dn animals, it is unclear how adding a GABA-receptor antagonist that

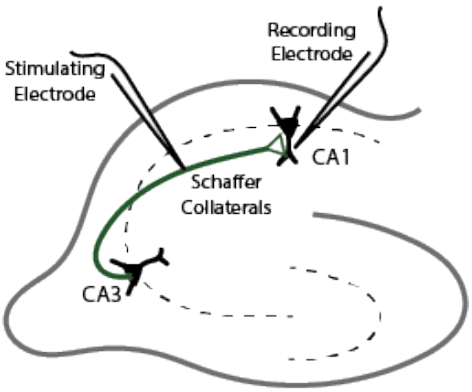


Figure 5.1: **Schematic of experimental recordings:** (**A**:) Diagram of hippocampus, with the placement of the extracellular stimulation electrodes close to the Schaffer collaterals and the recording electrode in CA1.

reduces inhibition, and adding bumetanide that restores inhibition can both rescue plasticity.

5.3 Experimental results

LTP and LTD induction was measured at the synapses from the Schaffer collaterals onto CA1 neurons. This experimental preparation is widely utilised because it is easily accessible for *in vitro* slice experiments. The stimulating electrode is placed directly above Schaffer collaterals and the responses in CA1 can be measured by placing a recording electrode in CA1 (Figure 5.1). For extracellular plasticity experiments, the stimulation strength is first adjusted so that the electric pulse elicits half of the maximal response recorded in CA1. To induce changes in the strength of the response, there are several protocols that have been identified in wildtype mice at these synapses and typically involve varying the frequency of extracellular stimulation.

The experiments used a theta-burst stimulation (TBS) protocol to induce LTP, which consisted of five bursts of four pulses at 100 Hz, and a period of 200 ms between each burst. LTD was induced by stimulating the slices 900 times at 1 Hz, or in other words 1 Hz stimulation for 15 minutes. The recording electrode measures changes in the field potentials that reflect the population activity, which

not only contains the response due to the Schaffer collateral activation, but also what is known as a population spike that reflects the summed action potentials evoked by the stimulation of the Schaffer collaterals. As such, the initial slope of the recorded field excitatory post-synaptic potential (fEPSP) slope is used as a proxy of the strength of the connections at the synapses from the Schaffer collaterals onto CA1 neurons (Deidda et al., 2015).

The first set of experiments used sequential inductions of LTP (seqLTP) to measure the saturation of LTP. The baseline response of the slice was recorded for ten minutes stimulating once every 30 seconds. Then they induced LTP using TBS, and waited thirty minutes between each subsequent induction. The 30-minute interval between inductions allowed the responses to reach their new stable level after a brief transient period. The average fEPSP as a percentage of baseline ($\pm$SEM) for Ts65Dn ($n = 7$) and wildtype ($n = 7$) slices showed that the the responses recorded in wildtype slices were consistently larger than those in Ts65Dn, but the difference across genotypes decreased for the second and third inductions. (Figure 5.1A). In the second and third induction, the slices show the same percentage increase with respect to the stable fEPSP level immediately preceding the LTP induction (Figure 5.1B). However, the response amplitudes recorded from wildtype slices are still larger than the the responses recorded from Ts65Dn slices after the third LTP induction (Figure 5.1A).

The second experiment used LTD to desaturate synapses before inducing LTP. As before, LTD was induced sequentially four times allowing for a thirty minute period between inductions to allow the fEPSP slope amplitudes to reach a new stable baseline. The only difference was that at the end of the four LTD inductions, LTP was then induced using TBS. The average fEPSP slope changes with respect to baseline for both genotypes overlap almost perfectly for LTD and the difference in LTP is reduced (Figure 5.3A). The amount of LTP achieved in the sequential inductions of LTP in the first experiment with respect of the response immediately before induction and the LTP attained after desaturation through LTD are plotted (Figure

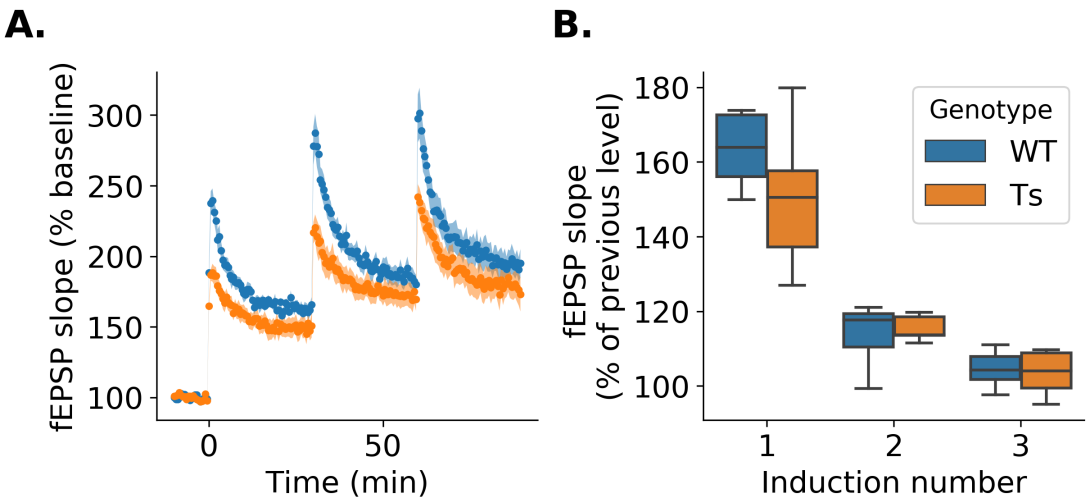


Figure 5.2: **Sequential LTP (seqLTP) experiments.** (**A:**) Average fEPSP slope as a percentage of the baseline for wildtype (in blue; $n = 7$), and for Ts65Dn (in orange; $n = 7$). (**B:**) fEPSP slope as a percentage of the previous level at each induction for wildtype (blue) and Ts65Dn (orange)

5.3B).

These results suggest that Ts65Dn slices are more saturated at rest compared to their wildtype counterparts, and that desaturation strategies, like LTD inductions used, can help recover the plasticity in these slices. The saturation of the Ts65Dn slices could be due to (i) a change in the machinery required to strengthen synapses that limits the total amount of LTP that can be induced, or in other words, the range of plasticity is smaller for Ts65Dn synapses, or (ii) synapses in Ts65Dn begin at a higher baseline strength than those in wildtype. Given the genetic complexity of DS, and the multiple aspects of brain function that are known to change in Ts65Dn animals, it is almost impossible to explain where the apparent saturation in Ts65Dn comes from, especially since it is unclear what saturation means in biological terms. In the next few chapters, we explore the use of mathematical models as a way to build a quantitative framework that can serve as a tool that enables experimentalists to think more concretely about the different possibilities and identify potential research directions.

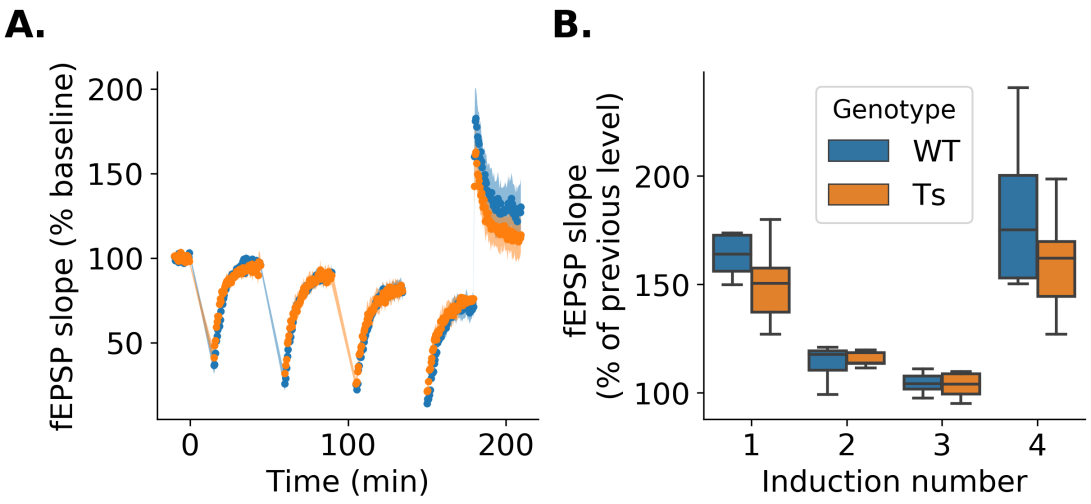


Figure 5.3: **Sequential LTD (seqLTD) experiments.** (**A:**) Average fEPSP slope as a percentage of the baseline for wildtype (in blue; $n = 7$), and for Ts65Dn (in orange; $n = 7$). (**B:**) fEPSP slope as a percentage of the previous level for wildtype (blue) and Ts65Dn (orange). The first three inductions are the results from the seqLTP experiments, and the fourth column is the fEPSP slope change after first performing four consecutive LTD inductions

5.4 Outline of Part II

We present three different models that approach the data from different perspectives and at differing levels of abstraction. The first model that we discuss is called statLSTP (statistical long-term plasticity), which is a model that proposes a statistical goal that synapses are trying to achieve during plasticity (Costa et al., 2017). We assumed that even with the myriad of changes at a cellular and network level, that the hippocampal networks in Ts65Dn mouse would try to optimise the same goal as those in their wildtype counterparts. Importantly, statLTSP had a straightforward and distinct interpretation for both the maximum level of LTP and the baseline levels of saturation, which are the two primary hypotheses to test. However, we found that statLTSP was not suitable for extracellular recordings, both because it was difficult to extract the necessary parameters from the data, and because of the mathematical properties of the model.

We then proceeded to use a phenomenological model in Chapter 7, whose purpose was to capture the qualitative properties of the data alone. Comparing the

parameter fits to the data, we discovered that the transient component of the recordings immediately following LTP induction, also called the post-tetanic potentiation (PTP), was more predictive of the genotype of the slice than the stable level of LTP attained. This finding has implications to experimental design and data analysis, but we wondered whether it could also reveal anything about the mechanisms of plasticity in Ts65Dn slices.

In Chapter 8, we build a third model based on the Moran process model for genetic variation in evolutionary biology. The model specifically linked the amplitudes of PTP and LTP. This model fit the data qualitatively, and was then used to predict mechanistic changes in plasticity induction in Ts65Dn. Furthermore, we used this model to formalise the idea of saturation mathematically. The Moran process model for synaptic plasticity has potential to help bridge between intracellular and extracellular plasticity experiments, as well as serve as a theoretical link between models of learning-and-memory and electrophysiological data.

Chapter 6

Applying statLTSP to extracellular LTP data

6.1 Introduction

Signals sent from one neuron to another one have been modelled as a single real-valued quantity in the models employed so far. However, the model in this chapter looks at the different components involved in neurotransmission. When a presynaptic neuron fires an action potential, the depolarisation at the synaptic terminals leads to the release of vesicles that contain neurotransmitters (*e.g.* glutamate or GABA) in sites of the neuron's cell membrane called release sites. The packets, or quanta, of neurotransmitters that are released into the synaptic cleft can then bind to receptors on the postsynaptic membrane that will result in changes in the postsynaptic membrane potential. The effect on the postsynaptic membrane potential due to a single neurotransmitter-containing vesicle will be referred to as the quantal amplitude (Del Castillo & Katz, 1954).

Synapses can change the strength of the signals that they transmit in response to the patterns of activity they receive. These changes are manifested through changes in the number of release sites for vesicles N, changes in the presynaptic probability of release, p, or changes in the quantal amplitude, q (Bekkers & Clements, 1999;

Del Castillo & Katz, 1954). There is variability, even within the same system and preparation, in whether plasticity is expressed presynaptically through changes in p or postsynaptically through changes in q (Larkman et al., 1992). statLTSP is a statistical framework proposed by Costa et al. (2017) that proposes that the expression loci of plasticity depends on the transmission statistics of the synapse at the time of the plasticity induction and the optimal transmission statistics. This theory was used successfully to explain several datasets recorded in different setups (*i.e.* from different laboratories) and in different brain areas.

One of the features of statLTSP is that it explicitly defines and fits a target or upper bound for plasticity, and could therefore be used to diagnose the difference in plasticity expression between Ts65Dn and wildtype slices. The two hypotheses that we wanted to test with statLTSP were that (i) the target mean response of the synapses in Ts65Dn animals are lower than that of wildtype animals and that (ii) synapses in Ts65Dn slices start at a higher baseline strength relative to the maximal strength of the connections. These hypotheses can be directly compared within the statLTSP framework, as this framework explicitly defines the target (*i.e.* an upper and lower bound) of LTP and LTD and the the baseline strength could be calculated directly from the parameter extraction from the data.

6.2 Methods

6.2.1 statLTSP

statLTSP belongs to a class of top-down models plasticity that begin by proposing a function for plasticity. Other models of this type, for example, propose that the function of plasticity is to maximise information transmission (Toyoizumi et al., 2005), or enable the brain to represent information according to some principle like sparse coding (Olshausen & Field, 1996) and predictive coding (Friston & Kiebel, 2009). statLTSP was chosen because it allowed us to explicitly test two hypotheses about the origin of the differences in plasticity between wildtype and Ts65Dn

animals.

To understand how statLTSP works, we first need to introduce the notion of synaptic transmission as a probability distribution. Even in the absence of plasticity, the response of a neuron to a presynaptic spike varies. The probability of the signal having a particular magnitude has been approximated using a binomial distribution (Del Castillo & Katz, 1954). The shape of the distribution depends on the number of release sites (N), the probability of release (p) and the quantal amplitude (q). The mean and variance of the binomial distribution of synaptic transmission are given by the following expressions:

$$\mu = Npq, \tag{6.1}$$

$$\sigma^2 = Nq^2p(1-p). \tag{6.2}$$

The mean response amplitude will increase linearly with respect to any change in p or q (Equation 6.1, and Figure 6.1A). On the other hand, the change in the variance is different for changes in p than it is for changes in q (Figure 6.1B). When the value of q increases or decreases, the variance will vary linearly. However, when $p = 0.5$ the variance will be maximal, and the variance will decrease to zero as $p \to 0$ or $p \to 1$ (Equation 6.2). Costa et al. (2017) used the fact that changes in p and q have distinct effects on the shape of the binomial distribution to propose a theory that could predict when a synapse would undergo changes presynaptically, and when plasticity would be expressed postsynaptically (Figure 6.1C).

statLTSP proposes that the goal of synaptic plasticity is for synapses to become as reliably strong (or weak) as possible. Maximal reliability is is achieved when the variance is minimised, which happens at $p = 0$ or $p = 1$, and statLTSP predicts that synapses will change their weight optimally to approach this reliable transmission state. Formally, synapses whose responses are binomially distributed will change p and q in order to minimise the difference between their response distributions P^{PSP} and a Dirac delta function, $P^{\text{bound}} = \delta(X - \phi)$, where ϕ is the target quantal

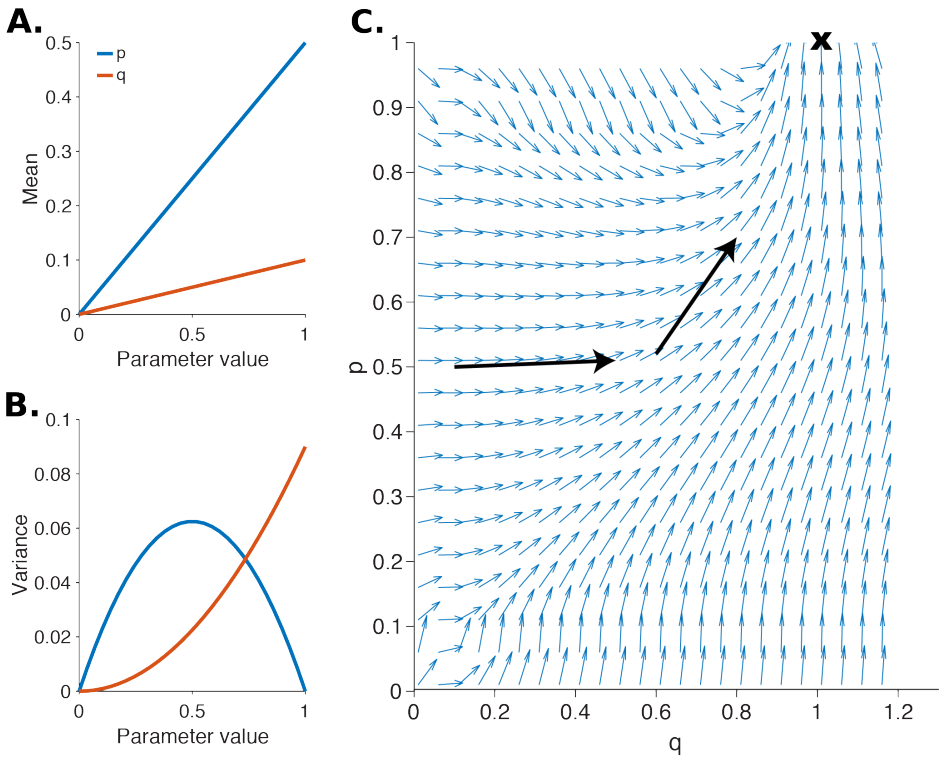


Figure 6.1: **Schematic of the model.** (**A**:) Changes in the average response with respect to p when $N = 1$ and $q = 0.5$ (blue) and with respect to q when $N = 1$ and $p = 0.1$ (red). (**B**:) Change in the variance of the response with respect to p and q. (**C**:) Predicted changes in p and q and two example trajectories with different starting points shown in black plotted on top of the full landscape for a given target at $(1, 1)$ (black x). Blue arrowheads show the predicted directions, but not magnitude of change, according to statLTSP.

amplitude. The similarity of the distributions is measured using the KL-divergence, which is a metric that measures the information loss between a probability distribution and a reference distribution from an information theoretic perspective (Bishop, 2006, Ch. 1).

The KL-divergence between the distribution of postsynaptic responses P^{PSP} and the Dirac delta function P^{bound} is

$$KL(P^{bound}||P^{PSP}) = \int P^{bound} \ln \frac{P^{bound}}{P^{PSP}} dX, \tag{6.3}$$

where X is the postsynaptic potential, and $P^{PSP} \sim \mathcal{N}(Npq, q^2Np(1-p))$ is the Gaussian approximation of the binomial distribution. This is not the only metric to compare distributions, and it has been shown that the landscape of trajectories can vary widely depending on the chosen metric (Surace et al., 2018). However, we continued using the KL-divergence given the success of Costa et al. (2017) at fitting a wide variety of datasets.

In order to find the optimal trajectory from its current state to the target distribution, the model proposes that synapses adjust p and q using gradient descent, which is an optimisation technique that can be used to find the minimum of a function. In gradient descent, the parameter values are iteratively adjusted to approach the minimum. The gradient is the amount that the value of the function, in our case the KL-divergence, will change if the parameter values are changed a small amount. If the gradient is zero, then it means that a minima or maxima has been reached.

The gradients for this system are the partial derivatives of Equation 6.3 with respect to p and q. In other words, the change in p after plasticity is $\partial KL/\partial p \cdot \delta\pi$ and the change in q is $\partial KL/\partial q \cdot \delta\pi$, where $\delta\pi$ is the magnitude of change induced by the plasticity induction protocol. Depending on the initial condition, the plasticity change will be primarily postsynaptic (horizontal change) or presynaptic (vertical change) (Figure 6.1C). Once the synapse reaches the target response distribution, then additional plasticity inductions will result in no changes as a minimum occurs

when $\partial KL/\partial p = \partial KL/\partial q = 0$.

6.2.2 Fitting statTLSP to data

The data was fit using Matlab code provided by Costa et al. (2017). The fitting algorithm requires knowing the values of p, q, and N before and after plasticity induction assuming N does not change. It then uses these values to fit the target, ϕ, that would best fit the observed changes with a genetic algorithm fitting procedure from Matlab's Global Optimization Toolbox. In order to use any fitting algorithm, one needs to input a cost function that is minimised. In this case, the error was calculated as follows. Given an initial target ϕ_0, and the pre-induction p_{pre} and q_{pre}, the trajectory of the data through q-p space using a Forward Euler Method. The simulation is stopped when the simulated weight reaches the measured synaptic weight or when $p_{\text{sim}}q_{\text{sim}} \geq p_{\text{post}}q_{\text{post}}$. The error is defined as the angle between the line connecting the initial starting point (q_{pre}, p_{pre}) and the observed (q_{post}, p_{post}), and the line connecting (q_{pre}, p_{pre}) and the simulated (q_{sim}, p_{sim}).

6.3 Results

6.3.1 Obtaining model parameters from the extracellular data

To fit statLTSP to the data, the model parameters (p, q, and N) need to be calculated. In the papers modelled in Costa et al. (2017), the parameter values were extracted from the distribution of recorded post-synaptic responses. After calculating the mean and variance of the recordings, one can use Equations 6.1 and 6.2 to solve for the values of the parameters. However, there are three unknowns and only two equations. As a result, it is necessary to fit the value of one of the parameters separately. Larkman et al. (1992) fit N separately, because it is assumed that N stays constant for the duration of a plasticity experiment, and that experiments in the same brain region using similar methods, record responses from a similar number

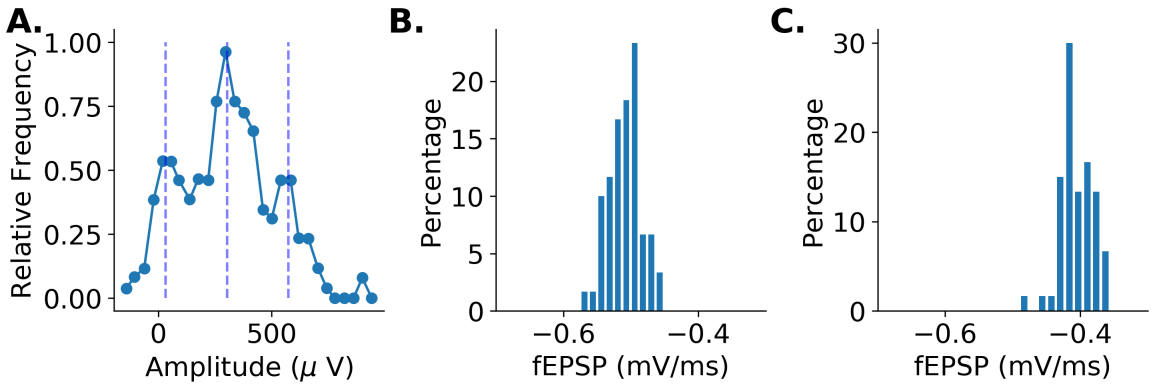


Figure 6.2: **Fitting quantal amplitude using histograms of responses.** (**A:**) Histogram of 250 responses from an intracellular recording. There are three equidistant peaks at a distance of approximately $280\mu V$ which represents the quantal size of the synapse marked by the three vertical dotted lines. Data obtained from Larkman et al. (1992) by hand and replotted here. **B,C:** Two representative histograms of 60 extracellular responses measured as the slope of the initial response. There are no easily distinguishable peaks in the responses.

of synaptic sites.

Unlike the data used by Costa et al. (2017), the data presented here was collected using extracellular recordings we had to fit N, since the value of N in extracellular recordings probably exceed the number of release sites in intracellular recordings. Furthermore, the responses to stimulation measured were not deflections in membrane potential, but rather the slope of the initial population response often referred to as field EPSPs (fEPSP). Even though the units are different, we reasoned that the slope which is in units of mV/s would be proportional to the changes in membrane potential evoked by stimulating the Schaffer collaterals.

In order to estimate N, we began by attempting to replicate the method proposed by Larkman et al. (1992), where they first fit the quantal amplitude for one representative slice. N was estimated from a single slice by recording post-synaptic responses for an extended amount of time in order to have at least 200 responses. These were then used to construct a histogram of the frequency of different response strengths, and this revealed a multimodal histogram with equidistant peaks (Figure 6.2A). Such a distribution is expected since the postsynaptic response is thought to be a sum of many quanta of neurotransmitters, and as such the response ampli-

tudes should be a multiple of the quantal amplitude. Therefore, we should be able to obtain q by measuring the distance between two consecutive peaks.

When we attempted to do this with the available data, it was not possible to discern multiple peaks from the histograms of responses (Figure 6.2B,C). This may have been due in part to the fact that we only had at most 60 responses for any given slice while Larkman et al. (1992) had at least 200 responses. However, it is also possible that since we used extracellular responses which reflect an aggregated response over many synaptic sites, each with possibly a different quantal amplitude, it may be impossible to obtain multimodal histograms such as the one described by Larkman et al. (1992) (Figure 6.2A). Even if the extracellular recordings were completely noiseless, since the responses would be pooled over many synaptic sites, by the Central Limit Theorem, we would still perhaps obtain a Gaussian distribution rather than a multimodal distribution.

We then decided to look at the estimates for p and q for different Ns (Figure 6.3). We found that for small N the fit for the probability of release is extremely high which is highly inconsistent with the initial values fit by Costa et al. (2017). In the experimental data used by Costa et al. (2017), the maximum p was 0.89, so we chose an N for which this was also true for our data. For the rest of this section, the values used were $N = 400$ for Ts65Dn data and $N = 500$ for WT data, since we know that there are approximately 24% more glutamatergic synapses in wildtype animals than in Ts65Dn animals (Dierssen et al., 2003).

6.3.2 Fitting statLTSP to the data

Once p and q were calculated, the optimal target that best described the trajectory in q-p space of the data was fit using code shared with us by the authors. The first step to fit the target was to normalise the data by the maximum value of q. However, since we wanted to compare the target for WT and Ts65Dn, the maximum value of q across all data was used to normalise the data even though the data corresponding to each genotype were then fit separately. Contrary to our hypothesis that the target

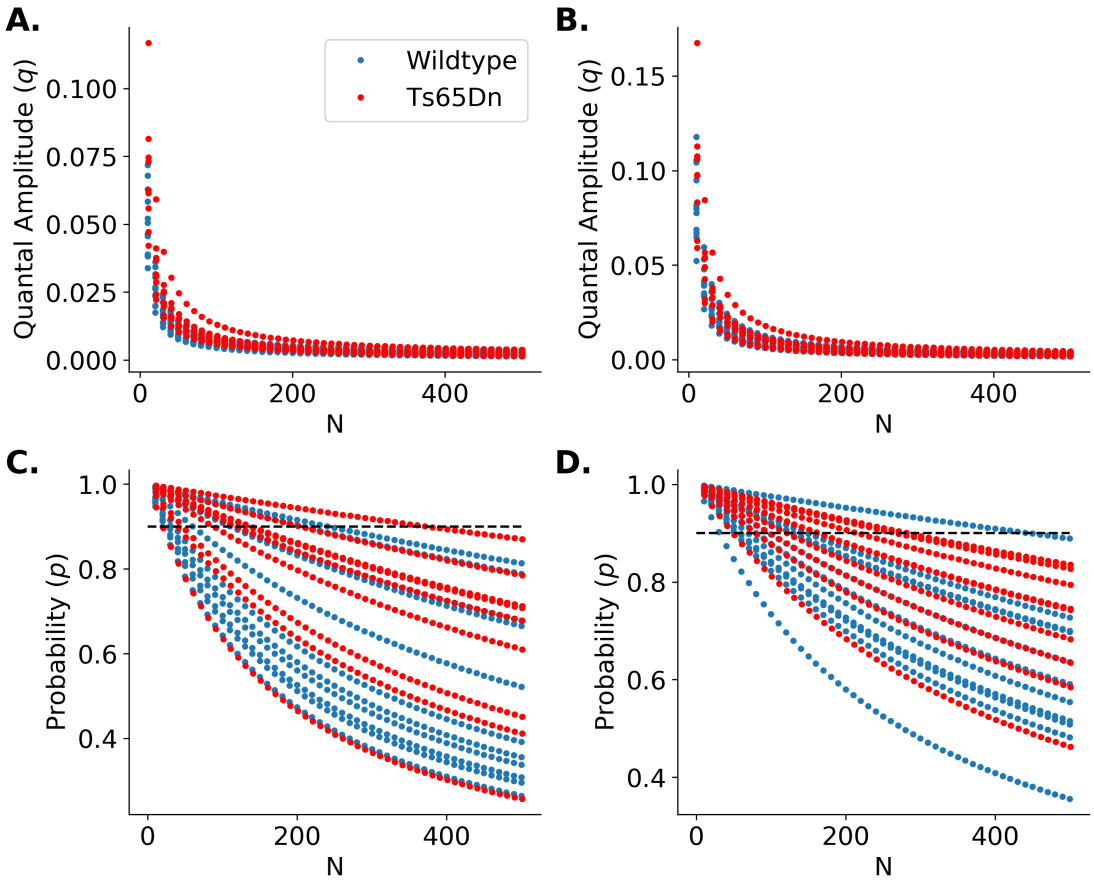


Figure 6.3: **Data estimates for p and q for different values of N.** (**A,B**:) Estimate of q for different values of N at baseline and after plasticity induction respectively. Fits for individual slices are shown in red for wildtype and in blue for Ts65Dn. (**A,B**:) Estimate for the release probability (p) is shown for different values of N at baseline and after plasticity induction. A horizontal dotted black line is plotted at the maximum value of p observed in the data fit in (Costa et al., 2017). Values of N for which all the fits were under this line were chosen for fitting.

for wildtype synapses would be larger than that of the Ts65Dn synapses, the best-fit target for the Ts65Dn synapses was 1.25 whereas the best target for widltype synapses was 1.02 (Figure 6.4A,C). Several confounding factors limit the strength of this result including the fact that we used an arbitrary number of synaptic sites for fitting.

The errors in the fit ranged from 0 to 90 degrees, and most of the Ts65Dn data fits had smaller errors than the fits for wildtype slices (Figure 6.4B,D). In general, the errors in our fits were larger than those reported in Costa et al. (2017), but this was expected due to the noisier nature of extracellular recordings, and our uncer-

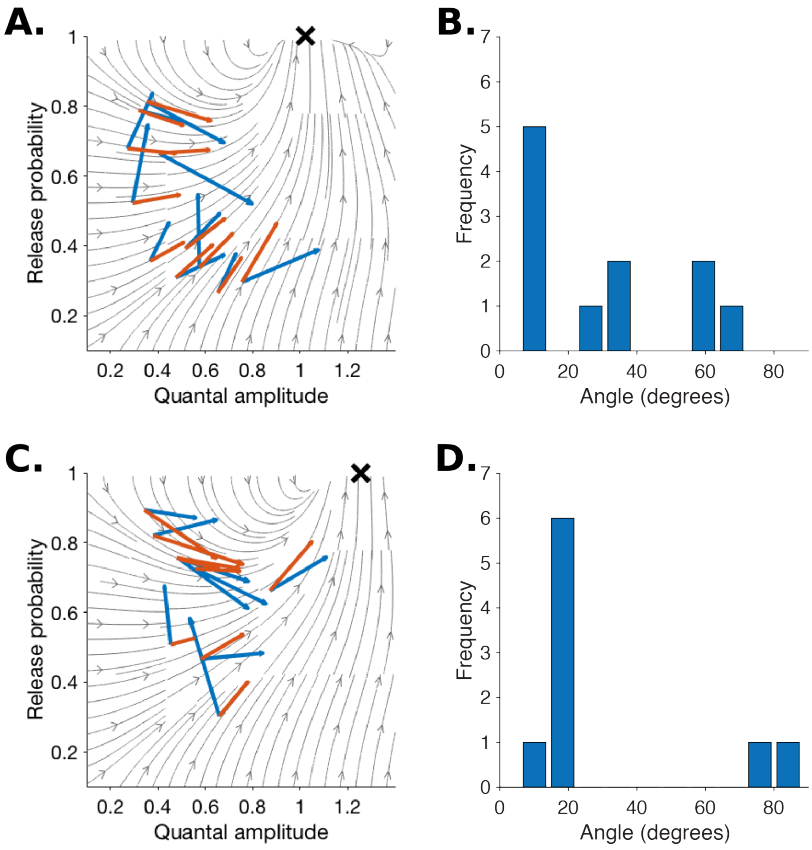


Figure 6.4: **statLTSP** The data was fit to (Costa et al., 2017). (**A,C:**) The observed (blue) and predicted (red) trajectories in q-p space for the wildtype and the Ts65Dn slices respectively. The endpoint of the red arrow is where the weight pq is the same as the observed mean $p_{post}q_{post}$ in the data. The grey arrows in the background show the predicted directions of change and the black x shows the location of the target. (**B,D:**) Error distributions for the wildtype and Ts65Dn data.

tainty in our fits of N. The distribution of errors both in the paper and in our fits are multimodal with the larger peak around smaller errors between 0 and 40 degrees. Although the errors suggested that perhaps the fits of the model were adequate, this result was difficult to assess because of the uncertainty in the calculation of the model parameters and the fact that we had few data points for each genotype. However, since LTP was induced three times for each slice, we decided to test whether statLTSP could explain the data not only in a single LTP induction, but also for all three sequential inductions of LTP.

6.3.3 Fitting statLTSP to multiple induction data

Multiple plasticity induction data is challenging, if not impossible, to obtain with intracellular recordings as performing multiple LTP inductions would require maintaining a stable intracellular recording for more time than is usually possible. As such, the sequential inductions of plasticity presented a unique opportunity to see what happens to the release probabilities and the quantal amplitude after a second and a third LTP induction. Using the same method described above, we fit p and q for all the consecutive inductions of LTP, but we treated each induction as an independent data point. We found that the trajectories of the data in q-p space were very different to those predicted by statLTSP, as can be seen from two example slices (Figure 6.5A). The directions predicted for the second and third LTP inductions sometimes even moved in the direction opposite of the predicted direction. We found, however, that the distribution of angular errors failed to reflect the failure of statLTSP in capturing the changes in the data, as more than half of the errors were still below 40 degrees (Figure 6.5B). We reasoned that one of the reasons that the error distribution did not capture how badly statLTSP seemed perform at capturing the changes in the sequential inductions of the data was that the error for each individual induction was not so bad, but when comparing the predicted trajectories to the actual movement of p and q after each plasticity induction it seemed that the data changed in ways that could not be predicted using statLTSP.

We noticed that the length of the predicted arrow grew smaller at each induction, whereas the length of the trajectory in the data, shown in blue, did not decrease as dramatically. The endpoint of the red arrow is determined by the observed mean in the data after the induction, $p_{\text{post}}q_{\text{post}}$, which means that the average response shown by a synapse at the (p, q) coordinates of a blue arrow would be equivalent to a synapse at the (p, q) coordinates of the tip of the red arrow. We measured the relative difference, D, between the length of the actual and predicted changes in q-p

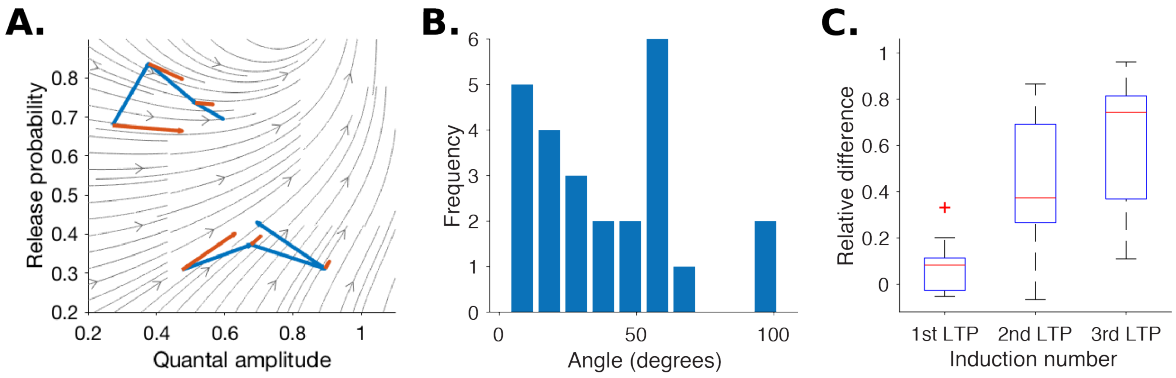


Figure 6.5: **Problem fitting using statLTSP.** (**A:**) Two example slices with three sequential LTP inductions. The blue arrows show the trajectories fit to the data, the red arrows are the predicted changes by the model, and the light grey lines show the gradients of the model at different points. (**B:**) Histogram of the error measured as the angle between the predicted and the real changes in p and q for each LTP induction. (**C:**) The relative difference in the magnitude of the predicted ($|\vec{v}^{\text{pred}}|$) and real ($|\vec{v}^{\text{data}}|$) changes in $q - p$ space calculated as $(|\vec{v}^{\text{data}}| - |\vec{v}^{\text{pred}}|)/(|\vec{v}^{\text{data}}| + |\vec{v}^{\text{pred}}|)$.

space (Figure 6.5C). The relative difference was defined as,

$$D = \frac{|\vec{v}^{\text{data}}| - |\vec{v}^{\text{pred}}|}{|\vec{v}^{\text{data}}| + |\vec{v}^{\text{pred}}|}, \tag{6.4}$$

where $|\vec{v}^{\text{pred}}|$ was the magnitude of the predicted change and $|\vec{v}^{\text{data}}|$ was the magnitude of the change observed in the data. The relative difference can take values $D \in [-1, 1]$, and the closer it is to 0, the more similar the two magnitudes were. As a result, we can see from Figure 6.5C that the predicted and real magnitudes of change were similar in the first LTP induction, but the difference increased for the second and third inductions. Furthermore, since they increased in the positive direction, it showed that the changes in the data were larger than the predicted magnitude of change. This result suggested that contrary to the predictions of statLTSP, synapses do not continue to travel in the 'optimal' trajectory towards the target after one LTP induction, and that statLTSP only captures the magnitude of the changes in p and q for the first LTP induction.

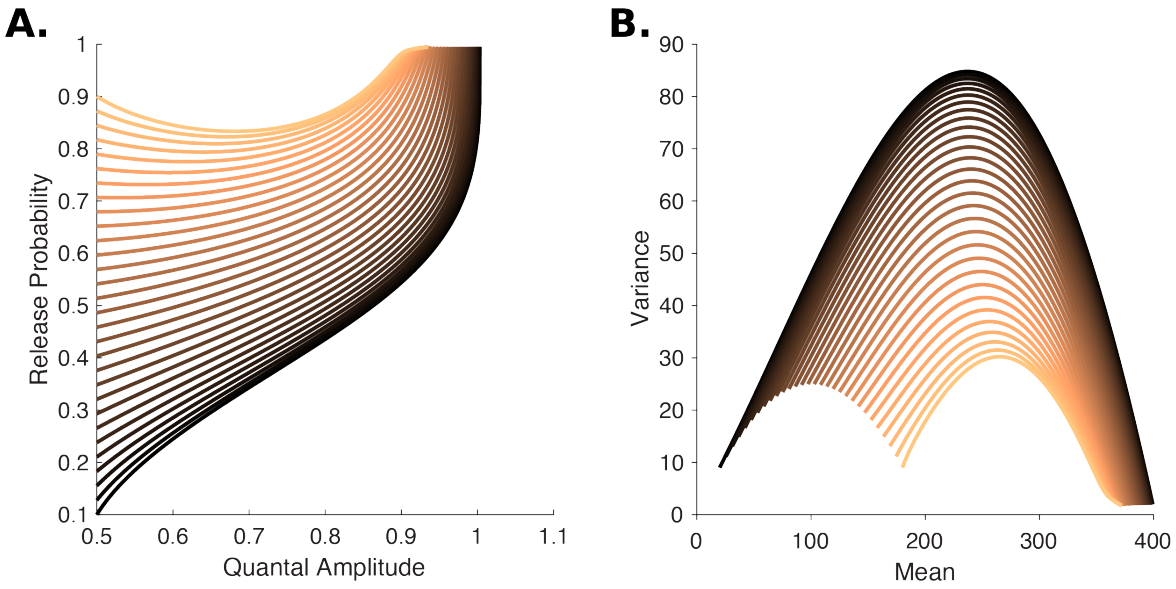


Figure 6.6: **Translating trajectories in q-p space into changes in variance and mean.** (**A:**) 30 trajectories through q-p space towards the target at $q = 1$. (**B:**) The same trajectories are replotted in mean-variance space. All simulations were stopped when $p \geq 1$.

6.3.4 Using mean and variance to test the predictions of statLTSP

One of the difficulties when using statLTSP for extracellular data was the uncertainty and imprecision with which the model parameters were obtained from the data. Given this uncertainty, it was difficult to conclude how much of the error in the fitting process to attribute to errors in the parameter values. However, all of the results from Costa et al. (2017) can be expressed equally in terms of the changes in the mean and variance of synaptic responses after LTP induction similar to the variance-mean analysis used for short-term plasticity (*e.g.* Ikeda et al., 2008). In particular, the premise of statLTSP is that synapses will tend to a state of minimal variance. However, to reach this state, weaker synapses are predicted to first increase their variance before decreasing to a state of maximum reliability or minimal variance. In Figure 6.6A we show a series of trajectories in the q-p space we have been working with and the same trajectories plotted in μ-σ^2 space in Figure 6.6B.

To test whether the sequential LTP inductions were in line with the predictions of statTLSP, the ratio of the mean of the post-LTP induction responses to the mean

of the pre-LTP induction responses are plotted for the three hippocampal datasets used by Costa et al. (2017) and the sequential induction results for wildtype and Ts65Dn animals (Figure 6.7A). The distribution of the ratio of the variances was also calculated for the same datasets (Figure 6.7B). In general, the ratios of the mean and variance for the intracellular datasets used by Costa et al. (2017) were larger than those in the extracellular recordings, which may be due to differences in the recording setups. For the sequential induction data, the increase in the mean is greatly diminished in the second and third inductions for both wildtype and Ts65Dn, which would be indicative that the synapses reached their target. This intuition was confirmed for wildtype synapses as the variance ratio was larger than 1 for the first two inductions, but then was closer to 1 at the third induction, suggesting that the increase in the variance was slowing down. However, the pattern was different for Ts65Dn synapses as the variance increased after the first induction, stayed the same after the second induction, and then increased a lot in half of the data points after the third induction. This suggested that the statLTSP framework did not adequately predict the changes in variance of Ts65Dn synapses even though it may adequately describe the changes in the wildtype synapses.

If the synapses in wildtype animals behaved in a way that was consistent with the hypothesis set forth by the model in μ-σ^2 space, then what could explain the failure of statLTSP in capturing the trajectories of the slices in q-p space? It is possible that our inability to fit N properly is at the root of our difficulties, but this explanation may be too simplistic because of the way in which the model fails to predict the sequential induction data. Using the wrong value of N should not affect the general shape of the trajectories, but the predicted trajectories were completely different from the changes predicted by the model. This suggested that perhaps the model itself is only suitable to model data that have been recorded intracellularly (either paired patching or a single patch with extracellular stimulation), but not for extracellular recordings. In order to address this question, we spend the rest of the chapter discussing what happens to the behaviour of the system for larger values

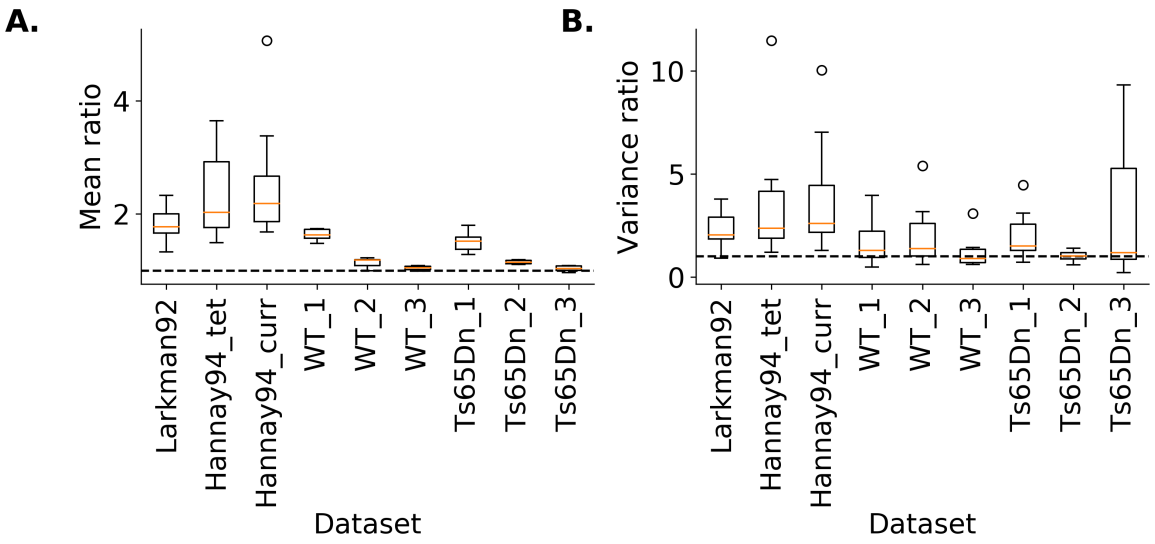


Figure 6.7: **Comparing the induced changes in variance and mean across datasets.** (**A:**) The ratio of the mean in data after induction to the variance before induction is plotted for three hippocampal datasets used by Costa et al. (2017), as well as the three sequential inductions on wildtype and Ts65Dn slices recorded in the Cancedda lab. (**B:**) Similar to (**A**), but for the variance. In both plots, there is a horizontal line at the ratio of 1. Any data points below this line indicate that the mean or variance decreased after induction and vice versa.

of N.

6.3.5 Can statLTSP be used on extracellular recordings?

In order to address the question of whether statLTPS is a valid model for extracellular recordings, we needed to understand the effect of varying N on the predictions and behaviour of the model. Costa et al. (2017) does not discuss N in depth, but we noticed that N is present in both the expressions of the mean and variance of the binomial model of neurotransmitter release (Equations 6.1 and 6.2). Sometimes it is the case that two parameters will vary proportional to each other, and it is possible to reduce the number of parameters in the equation by introducing a rescaling factor. However, it is not possible to rescale p and q to absorb N in Equations 6.1 and 6.2. This means that N adds an extra degree of freedom independent to p and q. In this section we review the derivation of the gradients from Equation 6.3. Then, we carry out a fixed-point analysis, which was not shown in (Costa et al., 2017) and is a

standard way to study systems of ordinary differential equations. Finally, we study how the system changes for different values of N.

Costa et al. (2017) directly substitutes the Dirac delta function $P^{\text{bound}} = \delta(X - \phi)$ into the expression for the KL-divergence (Equation 6.3), and obtains the following expressions for the partial derivatives of the KL-divergence with respect to p and q:

$$\frac{\partial KL}{\partial p} = \frac{1-2p}{2p(1-p)} + \frac{Npq - \phi)(p(Nq - 2\phi) + \phi)}{2N(1-p)^2p^2q^2}, \tag{6.5}$$

$$\frac{\partial KL}{\partial q} = \frac{1}{q} + \frac{\phi(\phi - Npq)}{(p-1)Npq^3}. \tag{6.6}$$

However, it is not possible to derive the stationary point $(q, p) = (\phi, 1)$ from these expressions (see Appendix B). However, it was possible that the reason that we failed to confirm that the stationary points for the system occur at $(\phi, 1)$ could be due to the fact that the KL-divergence of any distribution to a Dirac delta function is ill-defined, since the Dirac delta function makes sense only as a limit. As a result, the stationary point can be obtained by calculating the KL-divergence between the distribution of the postsynaptic responses, P^{PSP}, and a Gaussian distribution with mean, ϕ, and variance, σ_b^2. We do all the same calculations and only at the end do we take the variance, $\sigma_b^2 \to 0$.

If we let $P^{\text{PSP}} \sim \mathcal{N}(\mu, v)$, where $\mu = Npq$ and $\sigma^2 = Nq^2p(1-p)$, and $P^{\text{bound}} \sim \mathcal{N}(\phi, \sigma_b^2)$, then,

$$KL(P^{bound}||P^{PSP}) = \int P^{bound} \ln \frac{P^{bound}}{P^{PSP}} dX \tag{6.7}$$

$$= \int P^{\text{bound}} \ln \frac{v}{\sigma_b} dX - \int P^{\text{bound}} \frac{(X-\phi)^2}{2\sigma_b^2} dX + \int P^{\text{bound}} \frac{(X-\mu)^2}{2v^2} dX \tag{6.8}$$

$$= \ln \frac{\sigma}{\sigma_b} - \frac{1}{2} + \frac{1}{2\sigma^2} \int P^{\text{bound}} (x-\phi+\phi-\mu)^2 dX \tag{6.9}$$

$$= \ln \frac{\sigma}{\sigma_b} - \frac{1}{2} + \frac{1}{2\sigma^2} \int P^{\text{bound}} [(x-\phi)^2 + 2(x-\phi)(\phi-\mu) + (\phi-\mu)^2] dX \tag{6.10}$$

$$= \ln \frac{\sigma}{\sigma_b} - \frac{1}{2} + \frac{1}{2v^2} [\sigma_b^2 + (\phi-\mu)^2] \tag{6.11}$$

$$= -\ln \sigma_b + \ln \sigma + \frac{\sigma_b^2}{2\sigma^2} + \frac{(\phi-\mu)^2}{2\sigma^2} - \frac{1}{2} \tag{6.12}$$

Then, the partial derivative with respect to μ is,

$$\frac{\partial KL}{\partial \mu} = -\frac{2(\phi-\mu)}{2\sigma^2}, \tag{6.13}$$

$$= 0 \quad \text{as } \mu \to \phi \tag{6.14}$$

which goes to zero as $\mu \to \phi$.

The partial derivative with respect to σ^2 is,

$$\frac{\partial KL}{\partial \sigma^2} = \frac{1}{2} \frac{\partial}{\partial \sigma^2} \left[\ln \sigma^2 + \frac{\sigma_b^2}{\sigma^2} + \frac{(\phi-\mu)^2}{\sigma^2} \right] \tag{6.15}$$

$$= \frac{1}{2} \left[\frac{1}{\sigma^2} - \frac{\sigma_b^2}{\sigma^4} - \frac{(\phi-\mu)^2}{\sigma^4} \right] \tag{6.16}$$

$$= \frac{1}{2} \left[\frac{\sigma^2 - \sigma_b^2 - (\phi-\mu)^2}{\sigma^4} \right] \tag{6.17}$$

$$= 0 \quad \text{as } \sigma^2 \to \sigma_b^2, \text{ and } \mu = \phi \tag{6.18}$$

In statLTSP we work in $q-p$ space, but the result is the same since we can simply do a coordinate change from the partial derivatives with respect to σ^2 and

μ to the derivatives with respect to p and q. The Jacobian of the transformation, $\mu = Npq$ and $\sigma^2 = Nq^2p(1-p)$ is,

$$J = \begin{bmatrix} \frac{\partial \mu}{\partial q} & \frac{\partial \mu}{\partial p} \\ \frac{\partial \sigma^2}{\partial q} & \frac{\partial \sigma^2}{\partial p} \end{bmatrix} = N \begin{bmatrix} p & q \\ 2qp(1-p) & q^2(1-2p) \end{bmatrix}, \tag{6.19}$$

which is bounded and defined for all q and p, so therefore, the gradient $\partial KL/\partial p$ and $\partial KL/\partial q$ also vanish at the stationary point. We note, however, that when one does the change of coordinates, $\mu = \phi$ means that $Npq = \phi$. Since $p \to 1$, then, $q = \phi/N$ rather than $q = \phi$ as was reported by Costa et al. (2017).

An adapted quiver plot with the system's nullclines was plotted to confirm that $(\phi, 1)$ behaves as a stable fixed point (Figure 6.8). Quiver plots usually plot the magnitude and direction of the gradient at a given point in space, and can be used to see the way that variables will evolve over time given a particular starting point. Since the absolute value of the gradient grows towards infinity close to $(\phi, 1)$, we decided to plot only the direction and normalise all vectors to have the same length. Since the gradients in p are increasing disproportionately faster than the gradients in q close to the attractor, the arrows in the vicinity of this point are almost vertical, which suggests that synapses would reach the maximum $p = 1$ not only at $q = \phi$ but also in the neighbourhood of this target.

The solid lines in Figure 6.8 were the nullclines, or the curves along which $\partial KL/\partial p = 0$ (blue), and $\partial KL/\partial q = 0$ (red). The system has a stationary point wherever the two nullclines meet. The nullclines approach each other near the point $(\phi, 1)$, reinforcing our intuition that this point indeed behaves as a stationary point of the system. One of the reasons that the atypical behaviour of this system close to $(\phi, 1)$ did not affect the results in Costa et al. (2017) was that the data they fit occupied a different part of the $q - p$ plane that was less affected by the unbounded growth of the gradients close to $p = 1$.

As noted above, one inaccuracy is that $(\phi/N, 1)$ is the correct stationary point for the system, so the results in Costa et al. (2017) are valid only for $N = 1$. While this

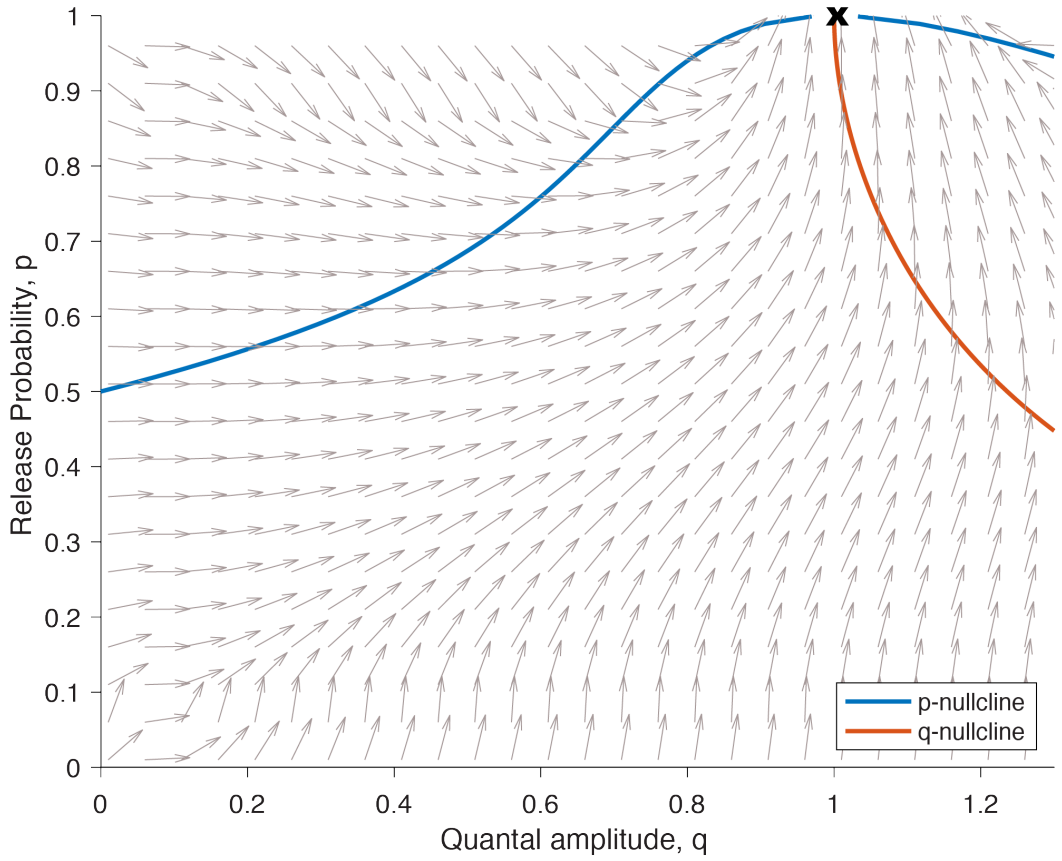


Figure 6.8: **Phase space for $N = 1$ and $\phi = 1$.** This figure shows a quiver plot where each arrow shows the direction of flow at each point of the figure. Since the gradients grow towards infinity close to $p = 1$, we normalised all the arrows to have the same length. The nullclines of a dynamical system are the lines at which $dKL/dq = 0 = dKL/dp$ and are shown in blue and red respectively.

may be an acceptable simplification for intracellular recordings where we expect the number of synaptic sites to be reduced, and in Costa et al. (2017) estimates ranged from $N \in [3, 12]$, this would be very different for extracellular data where a large population of synapses is contributing to the responses. In the following section we studied the effect of $N > 1$ on the phase space.

6.3.6 How does the system behave for $N > 1$?

We began by showing that the moment that $N > 1$, the attractor at $(\phi, 1)$ disappears. For $N > 1$,

$$\begin{aligned}
\left.\frac{\partial KL}{\partial q}\right|_{(1,\phi)} &\to \frac{1}{\phi} + \lim_{p\to 1^-}\left(\frac{\phi(\phi - Np\phi)}{(p-1)Np\phi^3}\right), \\
&\to \frac{1}{\phi} + \lim_{p\to 1^-}\left(-\frac{1-Np}{(1-p)N\phi}\right), \\
&\to \frac{1}{\phi} + \infty \\
&\to \infty.
\end{aligned}$$

Since the absorbing state for the system as $p \to 1$ was no longer $(1, \phi)$ for $N > 1$, we first attempted to find whether there was still an attractor in the system in q, by solving Equation B.8 with respect to q.

$$\frac{\partial KL}{\partial q} = \frac{1}{q} + \frac{\phi(\phi - Npq)}{(p-1)Npq^3} = 0 \tag{6.20}$$

$$(p-1)Npq^2 - \phi(\phi - Npq) = 0, \text{ for } p \neq 0, 1, \quad q \neq 0 \tag{6.21}$$

$$[Np(p-1)]q^2 + (Np\phi)q - \phi^2 = 0, \tag{6.22}$$

which is a quadratic function in terms of q, that becomes linear when $p = 1$ and for which $q = \frac{\phi}{N}$. Although this approximation may be true in the close vicinity of $p = 1$, one of the conditions for this to be true was $p \neq 1$. The formal way to find the roots of this equation at the limit $p \to 1$ is the quadratic formula, and found that the zeros of Equation 6.22 as $p \to 1$ are actually $q = \pm\infty$. Due to this inconsistency, numerical computations were used to confirm our intuition that for values close to $p = 1$, $dKL/dq \sim 0$ close to $q* = \phi/N$, and found that the trajectories to tend to approach $(q*, 1)$ and can be confirmed visually by looking at the nullclines (see Figure 6.9).

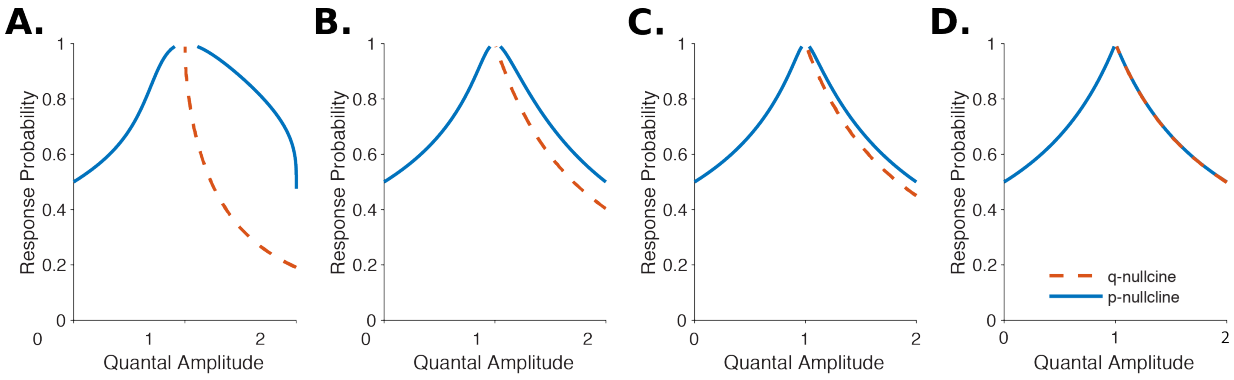


Figure 6.9: **Nullclines for different values of** N. The p-nullcline and q-nullclines for (**A**:) $N = \phi = 1$, (**B**:) $N = \phi = 5$, (**C**:) $N = \phi = 10$, and (**D**:) $N = \phi = 500$.

6.3.7 The effect of $N > 1$ on the predicted directions

The target gets renormalised by N, such that the trajectories in q-p space tend towards $(\phi/N, 1)$, so it is straightforward to predict the change in the target for different N. This alone, however, is not enough to understand the impact of increasing N on the model predictions, because it is the gradient at each point in space that determines the changes induced by plasticity. Next we evaluated the effect of increasing N on the gradients away from the target.

First, we focused on looking at the effect of N on the nullclines. Solving for q, we obtained four equations. The p-nullclines for the system were:

$$q = \frac{Np^2\phi \pm \sqrt{N(p-1)p(-1-(-4+N)p+(-4+N)p^2)\phi^2)}}{Np(1+(-3+N)p+2p^2)}, \tag{6.23}$$

for $(p-1)\left(Np^2t \pm \sqrt{N(p-1)pt^2((N-4)p^2-(N-4)p-1}\right) \neq 0$ and $Np((N-3)p+2p^2+1) \neq 0$.

The q-nullclines for the system were:

$$q = \frac{\pm\phi - \frac{\phi\sqrt{4+(N-4)p}}{\sqrt{N}\sqrt{p}}}{2(p-1)}, \tag{6.24}$$

for $N(p-1)p \neq 0$ and $\phi\sqrt{N}\sqrt{p} \pm \phi\sqrt{(N-4)p+4} \neq 0$.

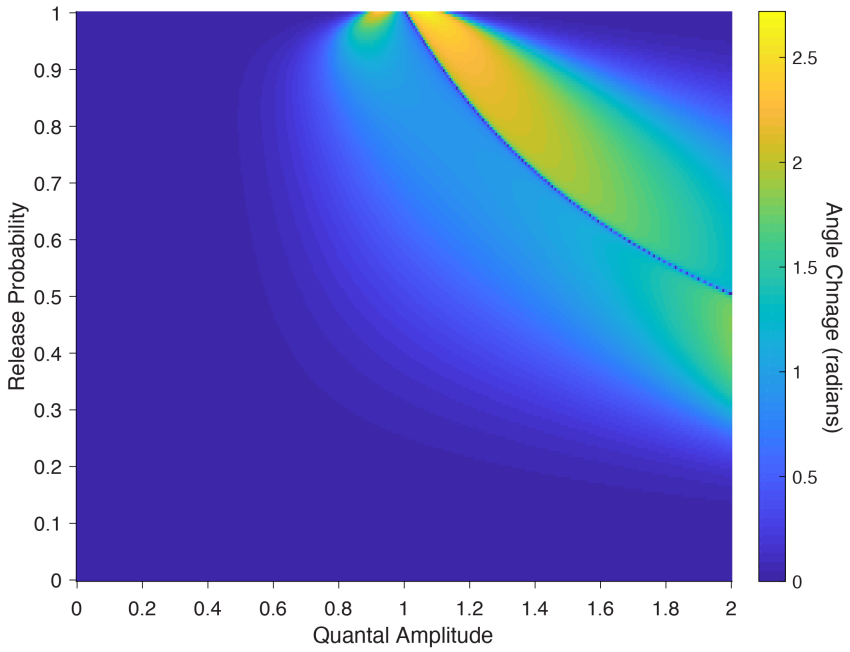


Figure 6.10: N **changes the gradients unequally in** $q - p$ **space.** The angular difference between the predicted directions of change in p and q at each point for $N = \phi = 1$, and $N = \phi = 100$.

We plotted the q- and p-nullclines for different values of N, but for simplicity we did so for the same *de facto* target $q* = \phi/N$ (Figure 6.9). The phase plane landscape was similar to the left of the fixed point, but the left brances of the q-nullclines and p-nullclines move closer together as $N = \phi$ increase. We expected the general behaviour of the system to change drastically close to where these two nullclines meet. To test this hypothesis, we calculated the gradient for $N = 1$ and $N = 100$, and plotted the angle between the two predicted directions. These angles are plotted in Figure 6.10 as a heatmap, where blue areas changed little in their predictions, and yellow areas changed the most.

6.4 Discussion

In this chapter we revisited the statLTSP model proposed by Costa et al. (2017), and analysed it in a context that was outside of the scope of the original study. First, we found that a feature data that is not captured by statLTSP is the predicted length

of the trajectory in space for the second and third induction. The proposition that synapses tend towards a particular target optimally using gradient descent is therefore not the complete story, because the data travels further in space than would be predicted following the first plasticity induction (Figure 6.5C). Since synapses are not updating p and q optimally according to statLTSP, this raises the question of what distance in q-p space represents in terms of information or computations in neuronal circuits.

Secondly, we explored the impact of N on the predictions by the model, and found that the current metric is not appropriate for extracellular recordings as the behaviour of the system changes drastically for larger N due to the fact that for larger N, the p-nullcline and q-nullcline grow closer giving rise to a *de facto* line of fixed points. This is particularly troublesome for the model, as it suggests that for larger N we are no longer guaranteed an approach to a state of maximal response and minimal variance, which contradicts the theoretical premise of statLTSP.

In order to overcome these issues, it may be necessary to choose a different metric, whereby the gradients are bounded and defined for all the points of interest. For example, it may even be possible to do this with the KL-divergence if a different target distribution were chosen like for example a Gaussian distribution with a small variance centred around the target. Alternatively, Surace et al. (2018) proposes different distance metrics that could be employed. While we identified several limitations of statLTSP including types of datasets that this model is unsuitable for, we also found some evidence that the computational principle of statLTSP may hold. In particular, when we looked at the changes in mean and variance for the sequential induction of LTP in recordings of wildtype slices, the variance seemed to increase in the first two inductions and stay the same or decrease in the third induction while the mean increased the most after the first induction, and while it continued to grow in the second and third induction, the magnitude of change decreased (Figure 6.7). This is precisely the trend that one would expect from the trajectories in mean and variance predicted by statLTSP (Figure 6.6).

Even though the existence of a line of fixed points for large N is certainly problematic, using a different metric may not be sufficient to make statLTSP a useful model to understand extracellular data. As the number of synaptic sites increases, the variance of the responses will be lower according to the Law of Large Numbers and the Central Limit Theorem. As the overall response variance decreases the larger the number of synaptic sites that are stimulated, this raises a second, theoretical challenge to applying statLTSP to extracellular responses. If the overall expected variability is lower for large N, we would expect it to be harder to explain any change in the mean or variance of the responses through changes in p and q. As a result, we posit that statLTSP may be suited only to fit data collected *in vitro* where the N would be lower, and the effects of individual synapses may be higher.

Our unsuccessful attempt to fit the extracellular LTP recordings as well as the mathematical analysis of the model for $N > 1$ suggest that while statLTSP may be useful to describe plasticity changes recorded using intracellular recordings, this method cannot be used on extracellular plasticity data. The part of the extracellular recording that is used as a proxy of the population response to stimulation is the slope of the fEPSP, which means that the range of possible values of the recorded responses is limited and subject to measurement errors. As the number of recorded extracellular responses increases, changes in the mean and variance in response to changes in p and q would also be much smaller as the term N would dominate the expression. As such, even though the nullclines intersect to form a line of fixed points for large N contradicts the premise of statLTSP whereby responses go to a state of maximum reliability, *de facto* the change in the variance for different points on the line of fixed points would probably be minimal due to the large N. As a result, we suggest that statLTSP should be used only with intracellular recordings where the number of release sites would be small, and where changes in p and q would have a measurable impact on the recorded responses.

Chapter 7

Phenomenological model

In this project, we were interested in using mathematical modelling as a tool to understand the differences in plasticity expression in Ts65Dn animals, a well-studied mouse model of Down Syndrome, compared to their wildtype littermates. Our first attempt to fit the data we received from the Cancedda lab was with statLSTP, which is a statistical framework that has successfully predicted presynaptic and postsynaptic changes in data before (Costa et al., 2017). However, all the datasets that were previously fit in Costa et al. (2017) were recorded using intracellular recordings. This led to several problems when we were adapting statLSTP to extracellular recording data. Firstly, in an intracellular recording one records responses from a much smaller number of synaptic sites. The numbers of synapses being recorded from was not a factor considered in the original paper, but their analysis was based on the assumption that this number was close to 1. When the analysis of the model was carried out to extend the model to multiple release sites, we found a number of instabilities in the system that suggested statLSTP is not suitable to describe extracellular datasets. In this chapter, we build a simple phenomenological model or a model that was intended to solely capture the qualitative features of the data (Figure 7.1). In addition to differences in the stable LTP levels induced in Ts65Dn and wildtype slices, there was a larger difference in the transient post-tetanic potentiation (PTP).

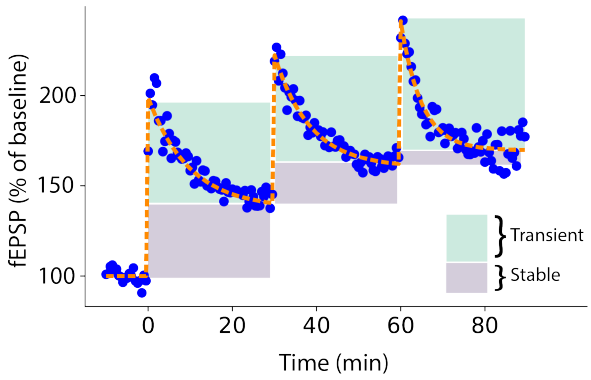


Figure 7.1: **Phenomenological model.** Data of an example experiment (blue circles). The phenomenological model fit (orange dotted line) that reproduces the main features of the changes in fEPSP slope responses in the slices. Notably, the height of the post-tetanic petentiation (transient peaks) at each induction step and that of the stable component of LTP are highlighted by the green and purple shaded boxes respectively.

7.1 Methods

7.1.1 Fitting the phenomenological model

First, we explored the data through a phenomenological model, that is, a model whose purpose was to reproduce the qualitative trends of the data without necessarily paying heed to its biological applicability. This type of model is thus not aimed to inform us about the data in terms of biological quantities like synapse numbers or the properties of inhibition. Instead, the model was aimed to identify aspects of the data that were the most different between wildtype and Ts65Dn.

The stereotypical curve of field extracellular post-synaptic potential (fEPSP) slopes observed during a LTP protocol consisted of an initial baseline during which fEPSP slopes are fairly constant. Following the induction protocol we saw a large deflection and subsequent decay down to a new, usually larger, fEPSP slope. As such, the data were fit with a function of the following form:

$$f(t) = \begin{cases} g, & \text{if } t < t' \\ g + A + B \cdot \exp(-|t - t'|/\tau), & \text{if } t \geq t', \end{cases} \tag{7.1}$$

where g is the baseline response, A is the stable increase in the response, B is the amplitude of the transient response, τ is the time constant of the decay of the transient, and t' is the time of induction. We could use a similar fit, but with an extra term of the form $A_n + B_n \cdot \exp(-|t - t'_n|/\tau_n)$ for every induction. Since the data we analysed was collected using three sequential plasticity inductions, the full form of the model was the following:

$$f(x) = \begin{cases} \text{baseline}, & \text{if} \quad t < t' \\ \text{baseline} + A_1 + B_1 \exp(-|t - t'|/\tau_1) & \text{if} \quad t' \leq t < t'' \\ \text{baseline} + A_1 + A_2 + B_2 \exp(-|t - t''|/\tau_2) & \text{if} \quad t'' \leq t < t''' \\ \text{baseline} + A_1 + A_2 + A_3 + B_3 \exp(-|t - t'''|/\tau_3) & \text{if} \quad t''' \leq t. \end{cases} \tag{7.2}$$

The model was fit to the data using the lmfit package in Python, and an example fit is shown in Figure 7.1.

7.1.2 Decision trees

In Section 7.2.3, we wanted to find the parameter that could best classify an experiment into whether it was recorded from a wildtype or a Ts65Dn animal. Decision trees are a machine-learning approach for classification where the class of a particular datapoint is decided by a series of if-then statements (Bishop, 2006, Ch. 14.4). There are many other approaches that can be used for classification like support vector machines (SVMs) (Cortes & Vapnik, 1995), logistic regression (Cramer, 2003), the perceptron (Rosenblatt, 1958; Freund & Schapire, 1999), and artifical neural networks (Basheer & Hajmeer, 2000). However, the primary purpose of a classifier is to take a new datapoint and assign it to the correct class. Here, however, we were interested in the subset of parameters that could best separate our data by their genotype. As such, we chose to use a decision tree as it was the most transparent classifier and provided a hierarchy of parameters that were useful to classify the data by the genotype of the slice.

We used a decision tree to classify slices into wildtype and Ts65Dn using the parameters in the model (Equation 7.2). The advantage of using a decision tree is that they are highly interpretable and it is clear what features of the data are most important for classification. Each if-then statement is called a fork and they split the data into two branches based on a split point or a value. The accuracy of the decision tree can be increased by allowing the decision tree to be deeper or, in other words, having additional forks after the first decision. Decision trees are trained to find the optimal split, which is defined as a split point that best divides the data into two homogeneous categories. We allowed for a maximum depth of 2 forks because of our small dataset. The models were evaluated by using a leave-one-out cross-validation where we trained the model on all the data except for one data point, and saw whether it could correctly predict the data point that was left out (Bishop, 2006, Ch. 1.3).

7.2 Results

7.2.1 Parameter fit distributions are similar across genotypes

A sample fit of the phenomenological model, and a sample fit is shown in Figure 7.1. We then studied the distributions of the fitted parameters (Figure 7.2), and found a large overlap in the distributions of the stable levels of LTP across strains, but a larger difference in the amplitude of the post-tetanic potentiation (PTP) given by the parameters B1, B2, and B3. The post-tetanic potentiation (PTP) is a transient increase in the response amplitude immediately following the plasticity induction protocol, and it is thought to be a primarily presynaptic phenomenon related to an increase in calcium in the presynaptic terminal following tetanisation (Sastry et al., 1986; Delaney et al., 1989; Zucker et al., 1991). The differences in the distributions of the PTPs in the recordings of wildtype and Ts65Dn slices suggested that perhaps one of the contributing factors to learning impairment in Down Syndrome could be presynaptic.

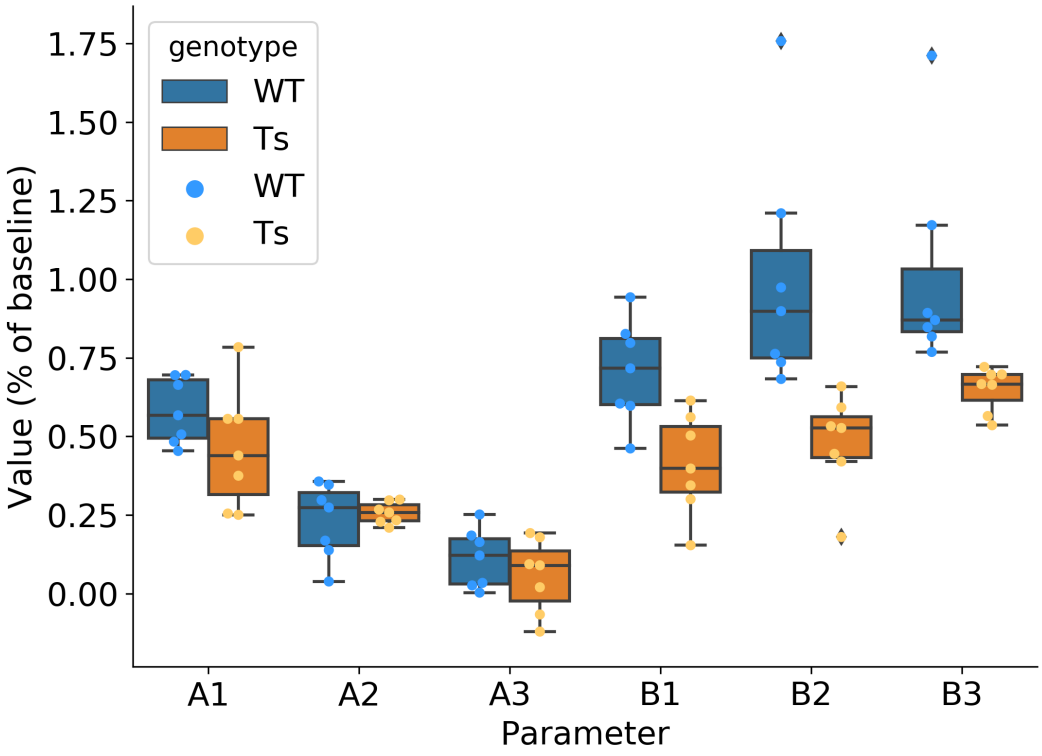


Figure 7.2: **Parameter distribution of phenomenological model.** The distribution of parameter fits is displayed as a boxplot in blue for wildtype slices and in orange for Ts65Dn slices. The first three parameters (A1, A2, A3) are the levels of stable plasticity at each induction, and the last three parameters (B1, B2, B3) are the levels of the transient peak of plasticity at each induction. The horizontal line inside each coloured box is the value of the median for the data, the bottom and top of each box are the 25th and 75th percentiles, and the whiskers show the full range of the data. If there are any points that are more than $1.5(Q3 - Q1)$ below the 25th percentile (Q1) or above the 75th percentile (Q3). The individual values of the fits for each slice are also shown with horizontal jitter so points do not overlap.

7.2.2 The data confirm the patterns of the phenomenological model parameters

We then returned to the data to ensure that the differences in PTP amplitude were not due to the way we had defined and fit the phenomenological model. We calculated the baseline as the mean of the fEPSP slopes 10 seconds before each plasticity induction, and the PTP as the maximum response within 2 seconds after plasticity induction (Figure 7.3). We observed that the stable LTP for WT and Ts65Dn (Figure

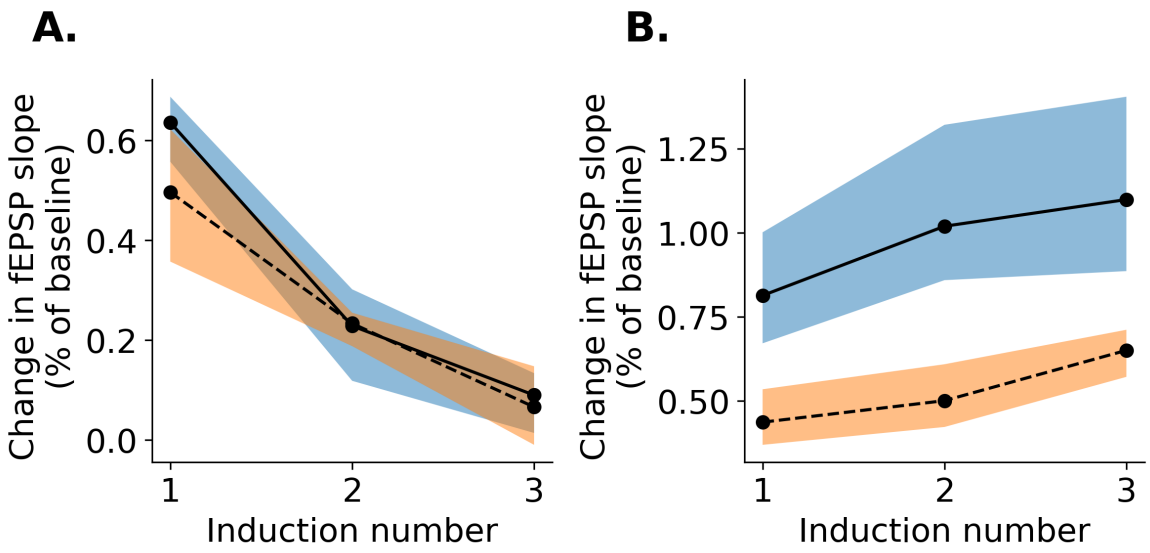


Figure 7.3: **Confirmation of the model using data.** (**A:**) Mean level of stable plasticity induced in WT (solid line; blue envelope) and Ts65Dn (dotted line; orange envelope) with the shaded bars representing the 95% confidence interval of the mean calculated using bootstrapping. (**B:**) Mean level of transient plasticity induced at each step.

7.3A), were less distinct for the two genotypes than the PTP (Figure 7.3B). Furthermore, the PTP amplitude increased for both wildtype and Ts65Dn over the three inductions, whereas the amount of induced LTP decreased after the first induction. The PTP amplitude seemed to be inversely correlated with the amount of LTP induced, so we wondered whether the differences in the PTPs observed between WT and Ts65Dn were unrelated to the learning impairment in DS. In the next section we show that regardless of its implications to learning, the PTP could accurately predict whether a data point was recorded from a WT or a Ts65Dn slice.

7.2.3 The transient response to a plasticity protocol is most predictive of genotype

Our finding that the PTP amplitude could distinguish between genotypes was counterintuitive, because the post-tetanic potentiation has been thought of as an experimental artefact and thus, not involved in learning and memory. In this section, we compared the ability of the LTP and of the PTP amplitudes to distinguish be-

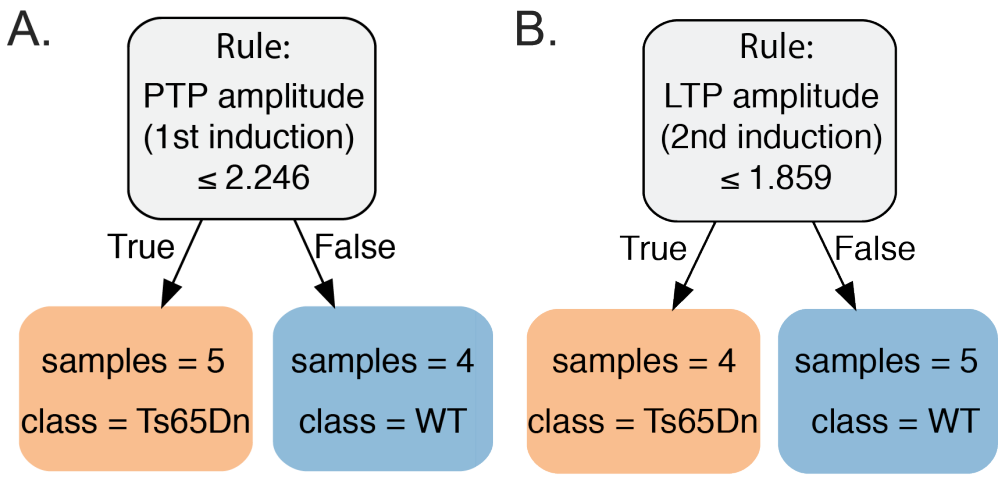


Figure 7.4: **Decision trees for classification.** The two diagrams show the structure of trees that were equally good at classifying the training set into the strain of each slice, but performed differently on the test set. The first box of the tree is the rule used to classify slices into their respective strains. (**A:**) Decision tree that was only allowed to use the transient levels of plasticity, and this decision tree used the PTP following the first LTP induction. (**B:**) Decision tree that was only allowed to use the stable levels of plasticity, and used the stable level after the second LTP induction.

tween WT and Ts65Dn data. Many methods have been developed for classification, but here we chose to train a decision tree as they can be used to easily interpret a hierarchy in the parameters of the model. However, due to the small size of the dataset, we found that different training instances would lead to slightly different parameters to be chosen as the first node of the decision tree (see Methods).

Since we were interested in understanding how informative the transient was over the stable component of plasticity about the genotype, we trained two decision trees. First, we classified based on the three transient PTP amplitudes (Figure 7.4A). This decision tree could predict the training set with an accuracy of 87%. The second tree was allowed to classify based on the three stable LTP amplitudes(Figure 7.4B), and performed with a lower accuracy of 58%. Finally, we found that a decision tree on all six factors did not outperform the first decision tree, classifying with an 87% accuracy. Interestingly, when we allowed the decision tree to have more than one decision node, the accuracy fell in all cases to 82%, 47%, and 80%, indicating that after one level, the decision tree starts to overfit the training set, *i.e.* the decision tree predicted well on the training set, but was worse at generalising to unseen examples

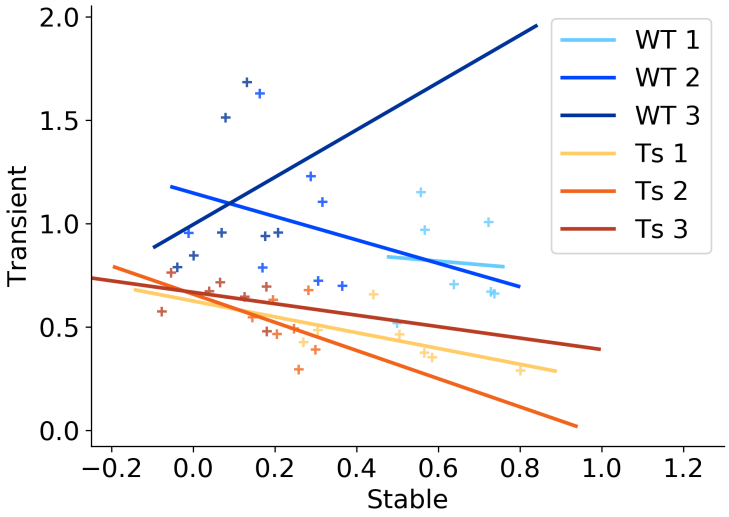


Figure 7.5: **Correlations between the stable and transient level of plasticity.** Each point in this plot represents a single induction experiment, with the tones of blue representing wildtype slices at each LTP induction and tones of orange representing Ts65Dn slices at each LTP induction. The lines are the lines of best fit for every induction type. The majority of inductions showed a slight negative correlation between the stable and transient components.

in the test set.

7.2.4 The transient response to a plasticity protocol is not predictive of stable level of plasticity

We noticed that while the amplitude of LTP decreased after the first induction, the PTP amplitude increased (Figure 7.3). To see whether there is a relationship between the level of the stable and transient components of plasticity, we looked at the correlations between the two. We plotted the LTP amplitudes against the PTP amplitudes for all plasticity induction experiments (Figure 7.5). We found a weak negative correlation for most of the inductions across the two genotypes. As such, even though the transient component seemed to be best at identifying the two strains, how the PTP can help us understand the changes in memory and learning in Down Syndrome, which are presumably mediated by the stable plasticity changes, remained unclear.

7.3 Discussion

Here, we used a phenomenological model to identify aspects of the data that could help us understand the differences in plasticity expressed by Ts65Dn animals and their wildtype counterparts. We analysed the explanatory power of each parameter and found that the post-tetanic potentiation was the most informative factor about the strain of the slice. When we trained a decision tree to classify the data into the strain of the slice it was recorded from, the decision tree achieved 87% accuracy based only on the PTP height, in contrast to a 58% accuracy when the decision tree was forced to classify based on the stable level of LTP. This means that the amplitude of the PTP is a much more reliable quantity to measure if we had to correctly classify a datapoint into its genotype. Unfortunately, the level of PTP was only weakly correlated to the level of LTP, which is the aspect of the data that we would like to understand. In the next chapter, we develop a model that links the transient and stable components of plasticity to better understand whether the PTP amplitude can be used to understand the biological mechanisms underlying learning impairments in DS.

The phenomenological model revealed that the post-tetanic potentiation amplitude was the parameter that best differentiated the data according to genotype. PTP amplitude is usually not discussed as being relevant for learning and memory due to its transient nature, so these results raise the question as to whether PTP amplitude is only good from a classification perspective, or whether it could be related to learning and thus relevant to the study of learning impairments in Down Syndrome. In the next chapter, we develop a model with a mathematical link between the amplitude height and the change in the weight due to plasticity induction to begin to address this question.

Chapter 8

Using a Moran Process for LTP and LTD

In Chapter 7, we used a phenomenological model and found that the post-tetanic potentiation (PTP), or the transient increase in response amplitude following plasticity induction, was a better predictor of the mouse genotype than the final LTP amplitude. However, even though the transient may be useful to classify data correctly into their respective genotypes, it may not explain the underlying learning impairment in Down Syndrome. Still, our results are important because the PTP is often ignored in the learning-and-memory literature. Intuitively, it makes sense to disregard the PTP, as learning and memory should be mediated by long-lasting changes in the brain, and the PTP was thought of as an artefact of the plasticity induction protocol. We also found a weak negative correlation between the PTP amplitude and LTP, but we would have expected that if the PTP were involved in determining the amount of LTP induced, that LTP and PTP would be positively correlated. In this chapter, we test the possibility that the difference in the PTP amplitudes could be related to the learning phenotype of Down Syndrome using a stochastic model that mechanistically linked the transient and stable components of plasticity.

We began by looking for a model that satisfied the following criteria. The model

had to (i) have well-defined stable and transient components, (ii) have a link between the transient and the stable response, (iii) be dynamic and capable of capturing the temporal characteristics of the data, and (iv) be analytically tractable. These criteria made finding such a model challenging because most models of plasticity do not explicitly model the PTP at all, and the only model that could capture transients is a relatively complicated system of three differential equations and 60 model parameters (Ziegler et al., 2015). Since our dataset was small (seven recordings for each genotype), and we did not have data on the parameter values for Ts65Dn animals, we decided to use a different model. We adapted a model used in the field of evolutionary biology called the Moran process (Moran, 1958). The Moran process is a simple model originally employed to study genetic variability in fixed populations that has few equations, is easy to simulate, and has mathematically proven properties that enabled us to develop computationally efficient ways to fit the model to the data.

8.1 Methods

In this section we first introduce the Moran process and its mathematical properties. Then we discuss how we adapted the Moran process to the context of synaptic plasticity.

8.1.1 The Moran process

The Moran process is a stochastic model that has a finite population of size N with two types of individuals (r and R). At each time step, a random individual is chosen to reproduce, and another random individual is chosen to die. Consequently, the size of the population stays constant, but the way that individuals are picked affects how the composition of the population changes over time. This model has been used to model phenomena in evolutionary biology such as neutral drift where neutral mutations, or mutations that don't confer any particular advantage or dis-

advantage, can spread through a population and replace the original allele (Moran, 1958; Allen & Tarnita, 2012). The Moran process can also be modified to capture selection, where an allele does have an advantage over another one, and mutations where alleles can mutate when they reproduce (Blythe & McKane, 2007). In the next section we review some of the known mathematical properties of the Moran process.

8.1.2 Neutral Moran Process

We began by exploring the mathematical properties of a neutral Moran process, in which each unit in the population is equally likely to replicate or die. We can rewrite this process as a Markovian process with a transition matrix of the size of the population. Each state in the transition matrix represents a different combination of r and R such that $r + R = N$. One extreme of the transition matrix will have $R = N$, its neighbouring state will have $R = N - 1$ and $r = 1$, and so on until the state $r = N$. The transition from one state to the next can be formulated as follows:

$$p_{i,i-1} = \frac{N-i}{N} \cdot \frac{i}{N} \tag{8.1}$$

$$p_{i,i} = 1 - p_{i,i-1} - p_{i,i+1} \tag{8.2}$$

$$p_{i,i+1} = \frac{i}{N} \cdot \frac{N-i}{N} \tag{8.3}$$

The probability of going from a state where $R = i$ to a state where $R = i - 1$ is the probability of choosing an individual of type R to die, and an individual of type r to reproduce (Equation 8.1). Because both of these conditions need to be fulfilled, then the probability of this transition is given by the product of these probabilities. In the case of a neutral Moran process, all members of the population are equally likely to be selected, which means that $p_{i,i-1} = p_{i,i+1}$, *i.e.* the probability of transitioning from a state $R = i$ to $R = i - 1$ is the same as that of transitioning from $R = i$ to $R = i + 1$.

8.1.3 Neutral Moran Process with selection

In the previous section we assumed that both individuals were equally likely to be selected, but this is not always the case. To capture differences in the fitness of the individual, or the probability that it will be chosen to reproduce, we introduce a Moran process with selection. If we assume that the R individuals have fitness f_i and the r individuals have fitness g_i, then, the transition probability from one state to the next is the following:

$$p_{i,i-1} = \frac{g_i(N-i)}{f_i \cdot i + g_i(N-i)} \cdot \frac{i}{N} \tag{8.4}$$

$$p_{i,i} = 1 - p_{i,i-1} - p_{i,i+1} \tag{8.5}$$

$$p_{i,i+1} = \frac{f_i \cdot i}{f_i \cdot i + g_i(N-i)} \cdot \frac{N-i}{N} \tag{8.6}$$

Note that if $f_i = g_i$, this reduces to the transition probabilities of the neutral case (Eqns. 8.1 - 8.3). We can derive the fixation probabilities in this case as well. Let $\alpha_i = p_{i,i+1}$, $\beta_i = p_{i,i-1}$, $\gamma_i = \beta_i/\alpha_i$, where x_i is the fixation probability of a synapse with $R = i$. We introduce a new variable $y_i = x_i - x_{i-1}$. We can recursively define the fixation probability as follows:

$$x_i = \beta_i x_{i-1} + (1 - \beta_i - \alpha_i)x_i + \alpha_i x_{i+1}, \tag{8.7}$$

We can rearrange this state to obtain:

$$\beta_i(x_i - x_{i-1}) = \alpha_i(x_{i+1} - x_i) \tag{8.8}$$

$$\beta_i y_i = \alpha_i y_{i+1}. \tag{8.9}$$

Using this, we can write an expression for x_m,

$$x_m = \sum_{i=1}^{m} y_i \tag{8.10}$$

as well as an expression for y_k,

$$y_k = x_1 \cdot \prod_{l=1}^{k-1} \gamma_l \tag{8.11}$$

Combining Equations 8.10 and 8.11 we obtain that,

$$x_i = \sum_{m=1}^{i} y_m = x_1 + x_1 \sum_{j-1}^{i-1} \prod_{k=1}^{j} \gamma_k \tag{8.12}$$

Since $x_N = 1$, then we can substitute that into Equation and find that

$$x_1 = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k}, \tag{8.13}$$

and we can substitute this back into Equation to obtain the following general expression for x_i:

$$x_i = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} + \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} \cdot \sum_{j-1}^{i-1} \prod_{k=1}^{j} \gamma_k \tag{8.14}$$

$$= \frac{1 + \sum_{j-1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} \tag{8.15}$$

Having an analytic expression for the fixation probability was crucial to fitting our model to data (Section 8.1.7).

8.1.4 Moran process with mutations

The benefit of the neutral Moran process is that it reaches a stable state in finite time. To adapt the Moran model to describe plasticity in synaptic populations, we needed to include a way to perturb the system away from its steady state. In the literature, one simple modification to the Moran process that enables a population to escape fixation is adding 'mutations'. In the mutations case, the individual chosen to reproduce has a probability m that its offspring will be of the other type. This can

be expressed as a transition matrix,

$$\begin{bmatrix} m_{1\to1} & 1-m_{1\to1} \\ 1-m_{2\to2} & m_{2\to2} \end{bmatrix}, \tag{8.16}$$

where $m_{1\to1} \in [0,1]$ is the probability that if an individual was chosen to reproduce and belonged to R its offspring will be R as well, and $m_{2\to2}$ is the probability that an r individual will have an offspring of type r.

Depending on the values chosen for $m_{1\to1}$, and $m_{2\to2}$, the probability of finding a particular number of individuals of one kind changes (Allen & Tarnita, 2012) (Figure 8.1). We simulated 1000 populations each of $N = 500$ for 2000 steps after initialising each population with 50% of each type (*i.e.* 50% R and 50% r). Without mutations, the distribution is bimodal, and it has two absorbing states at $R = N$ and $r = N$ (Figure 8.1A). For values of $m_{1\to1}$, and $m_{2\to2}$ less than 1, the distributions we obtained were U-shaped (Figure 8.1B), and as mutation rates increased the distribution became unimodal (Figure 8.1C). One can derive the stationary distributions of the Moran process using a Fokker-Planck continuous approximation (see Janvin, 2017).

Even though in the context of genetics, it makes sense to use the term 'mutation', this is less appropriate in the context of synaptic plasticity. Instead of individual types, we will now talk about readily-releasable vesicles (R) and reserve vesicles (r), and the 'mutations' should be thought of as 'rates of change' of a vesicle changing from being readily-releasable to becoming reserve and vice versa, and will be written in quotations for the remainder of the chapter.

8.1.5 The Moran process as a model of plasticity

To use the Moran process as a model of synaptic plasticity, we combined the Moran process with selection and the Moran process with mutations. We assumed that each population represents a single synapse. We thought of the individual components of a presynaptic synapse as vesicles that are readily releasable (R) or in a

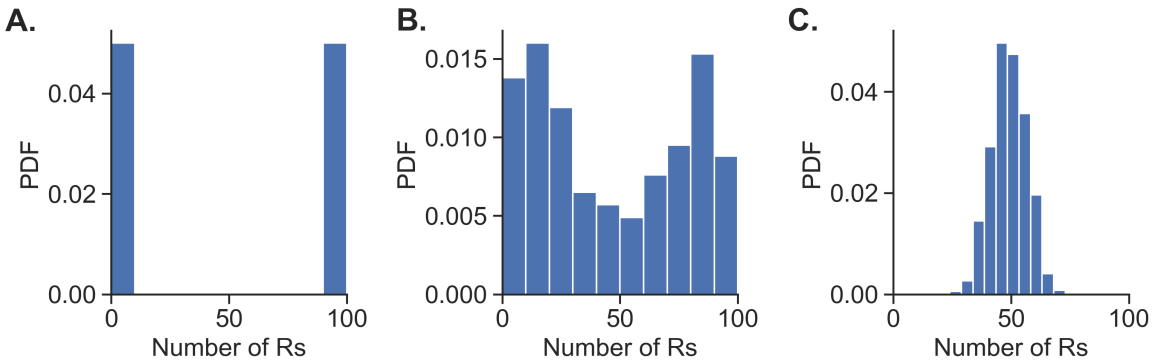


Figure 8.1: **Population distributions for different mutation rates** (**A:**) The probability distribution of a population with 100 individuals to have N individuals of R when there are no mutations introduced. The two extremes are stable attractors. (**B:**) A mutation probability of 0.01 is introduced. (**C:**) Mutation probability increased to 0.2.

reserve (r) (Figure 8.2A). Note that these two populations could equally represent components of the postsynaptic synapse as active receptors (R) or inactive receptors (r) that are in the synaptic terminal but have been internalised, without any loss of generality. We used this analogy as it made biological sense that vesicles or receptors would be in constant flux, and that their relative proportions would determine the response strength of a synapse. If we impose an additional assumption that the total number of vesicles or receptors in a single terminal will be fixed, then all the requirements for a Moran process are fulfilled.

We assumed that when the system was at rest, or at times when the synapses were not undergoing plasticity, they underwent a Moran process with selection. We know that a Moran process with selection fixates in finite time, or in other words, reach the state with all Rs or all rs in finite time. As a result, synapses in our model had two stable states, which were analogous to the high and low states of synapses (O'Connor et al., 2005). Each state had a transition probability to a neighbouring state, but once one of the two extreme states was entered, the probability that they will remain there is equal to 1 (Figure 8.2B). This suggested that synapses at rest remain at either a high or a low state, and that for synapses to undergo plasticity, there needs to be some perturbation away from the absorbing states. When we allowed readily-releasable vesicles, R, to become reserve vesicles, r, and vice versa, we ef-

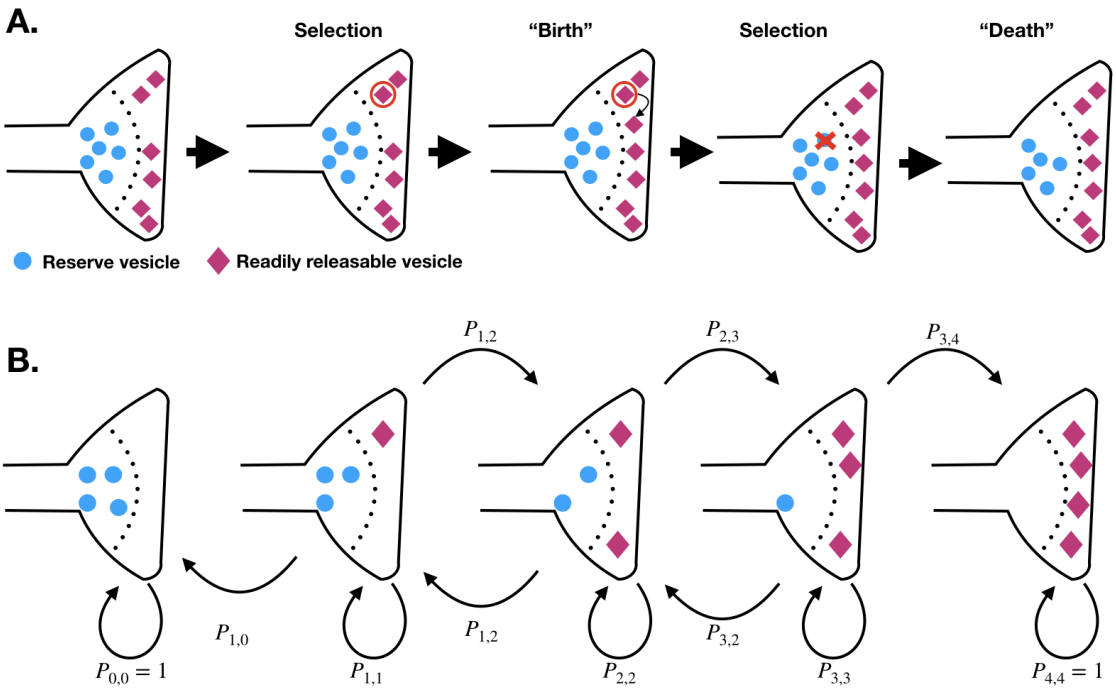


Figure 8.2: **Moran model schematic.** (**A:**) A single time step of the Moran process is illustrated. Each synapse is comprised of R readily-releasable vesicles (magenta diamonds), and r reserve vesicles (blue circles). First a vesicle is chosen at random and duplicated. Then a random vesicle is chosen at random and eliminated. At the end of the step the total number of vesicles will remain the same. (**B:**) An alternative description of the model is as a transition between states with varying numbers of each type of vesicles. Each transition has a probability $p_{i,j}$ of going from $R = i$ to $R = j$. The system can only transition between adjacent states, and once $R = 0$ or $R = N$, where N is the number of vesicles, then this state is stable.

fectively get rid of these two absorbing states. In other words, even when a vesicle was chosen for 'reproduction', the vesicle type that was added had a nonzero probability of becoming of the opposite type (Figure 8.3A). In practice, the changes in the state of a vesicle allowed synapses to escape the two absorbing states (Figure 8.3B). We modelled the end of a plasticity induction by reducing the probability of switching vesicle types back to 0, and returned to the dynamics of a Moran process with selection.

The probability of fixation was set by the proportion of R and r in the population (Equation 8.15), the stationary distribution of the Moran process with 'mutations' will be directly related to the new steady state of the synapse. In the neutral Moran

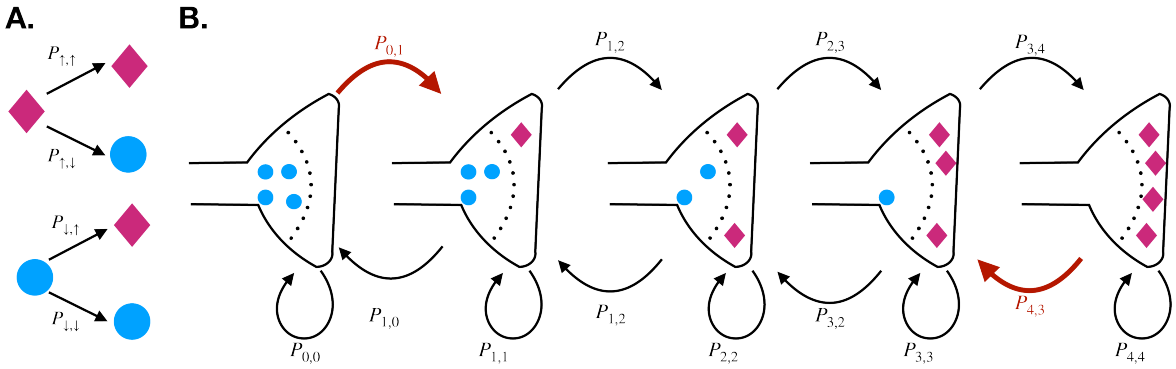


Figure 8.3: **Adding mutations to the Moran process gets rid of the two steady states.** (**A:**) A mutation in the Moran process means that when a vesicle is chosen to be duplicated, the new inserted vesicle can be of the same type with a certain probability or of the opposite type. (**B:**) Adding a mutation changes the transition probabilities and it makes P_{01} and $P_{4,3}$ nonzero as shown in the red arrows.

process, the probability of fixation was equal to the proportion of Rs wan the population. However, this did not capture the data because there would be no PTP. As a result, we included a selection curve where the probability of a vesicle type to be chosen to 'reproduce' depends on how many of that type of vesicle are in the synapse. The selection curve we chose was a sigmoidal function,

$$f_i = \frac{0.3}{1 + \exp(-k(i - m))} + 0.85, \tag{8.17}$$

where k was the slope, m was the midpoint, and it was a sigmoidal function that goes from 0.85 to 1.15.

A single synapse is not enough to capture the qualitative features of the experimental data, because it can only oscillate between two states, and given that it is probabilistic, we cannot ensure that an LTP protocol will indeed set the synapse to its high state but rather only guarantee that the probability of this happening is larger. This is consistent with the experimental data as extracellular recordings measure population-level responses to electrical stimulation, which corresponds to the activation of many individual synaptic sites, each undergoing similar changes in their internal composition of vesicles and receptors. Furthermore, the experimentally measured response would be equivalent to the sum of all the synapses

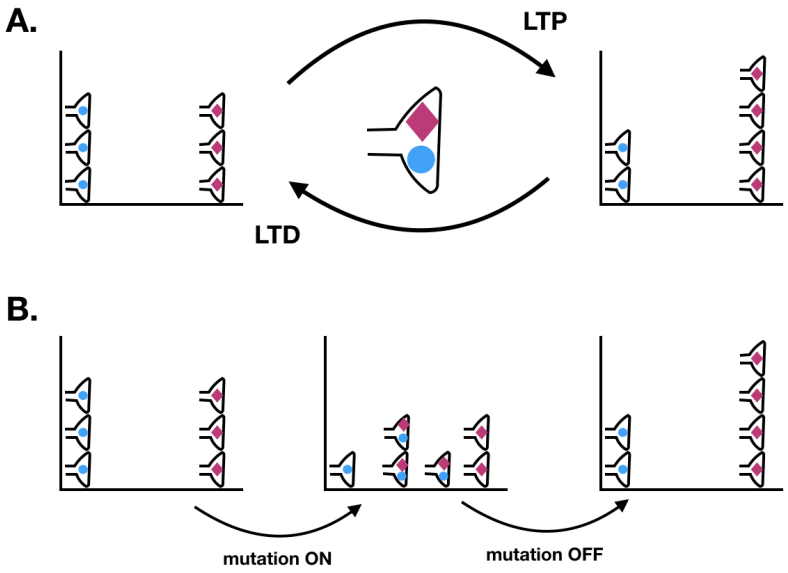


Figure 8.4: **LTP and LTD as changes in the distributions of synapses.** Individual synapses in this model go to one of two stable states. As such, individual synapses are not able to capture the gradual changes in the responses observed in the extracellular recordings. (**A:**) shows a schematic of a distribution of synapses where some proportion of the total are in their high state and the others are in their low state. In the model, LTP is equivalent to increasing the proportion of strong synapses in the steady state. (**B:**) Schematic summarising the three steps of plasticity induction in the model. During a plasticity induction, the 'mutation' is turned on, which allows synapses to move away from the two extremes, and at the end of the induction, they once again go back to one of the two stable states. However, since the plasticity induction pushed them away from the steady state, synapses have the ability to change their steady state from high to low or low to high.

weighted by their respective synaptic strength given by the proportion of active receptors or vesicles in each synapse. In experiments, performing a plasticity induction protocol changes the response strength to an electrical stimulation. In our model that would be equivalent to changing the proportion of synapses in a high or low state. As a result, increasing the proportion of synapses in a high state would be equivalent to LTP, and the reduction of synapses in a high state would be equivalent to LTD (Figure 8.4A). Furthermore, the transition between states is determined by the rates of change where synapses are driven towards an intermediate state in which they have a mix of readily-releasable and reserve vesicles (Figure 8.4B).

In summary, we adapted a Moran process to describe synaptic plasticity of large

populations of synapses. It assumes that at rest, synapses follow a Moran process with selection such that synapses remain in either a high or a low state. During a plasticity protocol, vesicles are allowed to change from R to r, which creates a new equilibrium distribution for the state of each synapse. At the end of this period, synapses will once again follow a Moran process with selection. Key to this model, however is the fact that the stationary distribution determined by the mutations will determine the average proportion of synapses that will end in a high or low state. Therefore, this model has a stable state set by the neutral Moran process, a transient state set by the 'mutation' rates, and a probabilistic link between these two.

8.1.6 Simulating the Moran process

The most intuitive approach to simulate a Moran process is to randomly select a vesicle for elimination and one for reproduction (see Figure 8.2A). However, this approach is computationally expensive, as it requires using nested conditionals. A more efficient method to simulate the Moran process emerges from the fact that the system is Markovian, or in other words, that the state at time $t+1$ depends only on the state at t, and the reproduction and elimination can be summarised as a transition matrix. Given that there are $R = k$ readily releasable vesicles at time t, the only possible states at time $t+1$ are $k-1$, k, or $k+1$, the transition matrix is tridiagonal.

We calculated the transition matrix, Tr, which was an $(N+1) \times (N+1)$ matrix where N was the total number of receptors in a synapse. The elements $Tr_{i,j}$ of this matrix were the probability that if at time t, there were i readily-releasable vesicles in the synapse, at time $t+1$ there would be j readily-releasable vesicles, where $j \in \{i-1, i, i+1\}$. There was a different transition matrix depending on whether we were simulating the period of time between plasticity inductions (see Figure 8.2B) or the plasticity induction protocol (see Figure 8.3B). Furthermore, we used a different plasticity transition matrix for LTP and LTD. As a result, we used one

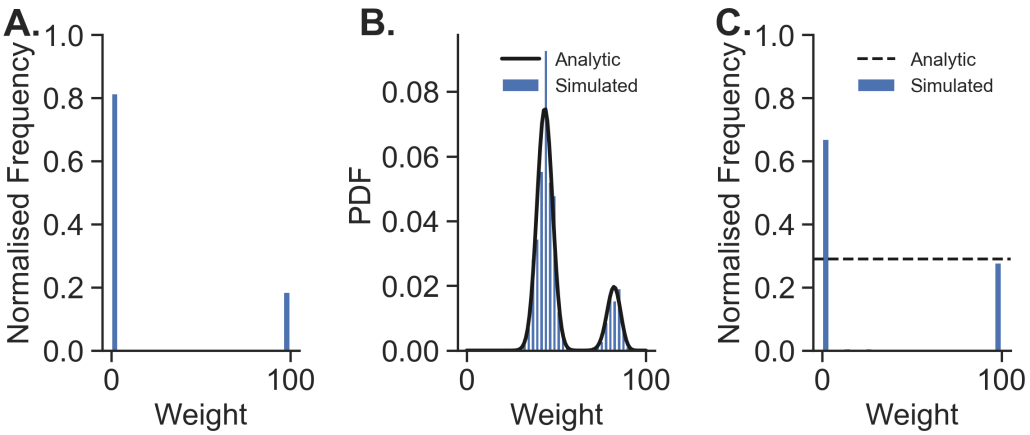


Figure 8.5: **Fitting the Moran process analytically.** (**A:** Distribution of synaptic weights at baseline where only 20% of synapses are at a high state. (**B:**) Distribution of synaptic weights immediately following the simulated plasticity induction in blue. The expected distribution of synaptic weights calculated analytically is plotted in black. (**C:**) The new equilibrium distribution of weights after plasticity induction in the simulation shown in blue. The dotted horizontal black line shows the proportion of strong synapses at weight 100 calculated analytically.

transition matrix for the Moran process with selection, Tr_{sel}, one transition matrix for LTP, Tr_{LTP}, and one transition matrix for LTD, Tr_{LTD}.

To simulate an LTP protocol, we started by setting an initial proportion of synapses in a high state, w_0. Each synapse had a total of N receptors. The plasticity protocol consisted of 150 time steps in which the weight of each synapse was simulated using the transition matrix for LTP, followed by a period of 8000 time steps between inductions during which the neutral transition matrix was used.

Some parameters, *e.g.* the duration of the plasticity protocols, and the time intervals between inductions were fixed in accordance with the data. Other parameters were fit. We fit the slope, k, and the midpoint, m of the selection curve as these would determine the shape of the PTP. Additionally, we fit the parameters for Tr_{LTD}, and Tr_{LTP}, which we refer as the Low-to-Low and the High-to-High rates in the rest of the chapter, because these determined the stationary distribution of the plasticity protocol. Finally we fit the baseline proportion of high and low synapses.

8.1.7 Fitting the parameters of the Moran process to data

In Figure 8.4B we show the transition from a stable state into a distribution of synapses with a mix of R and r vesicles after the plasticity protocols were turned on (*i.e.* Tr_{LTP} was used instead of Tr_{neutral} in the simulation), followed by a second transition to a new steady state. In this section we derive an analytic expression for these two transformations. These analytic expressions were needed for fitting, because numeric simulations of a single sequential LTP induction took about five minutes to run in Python.

For the first transformation, we used the property that given a transition matrix Tr_{LTP}, where each element gives the probability of being in a particular state in one time step, the elements of Tr^m_{LTP} give the probability of being in state $R = i$ in m time steps. As a result, assuming we started from the steady state (a synapse is either at a low or a high state), the distribution of synapse weights at the end of the plasticity induction was given by:

$$\text{pdf} = (1 - w_0)Tr^m_{\text{LTP}}[1][:] + w_0 Tr^m_{\text{LTP}}[N+1][:], \tag{8.18}$$

where w_0 was the initial proportion of strong synapses. Given an initial distribution of synaptic weights (Figure 8.5A), this probability distribution matched the simulated distribution of synaptic weights (Figure 8.5B).

After the plasticity induction, the state of the synapse will be driven by a Moran process with selection. We derived the expression for the fixation probabilities (Equation 8.15) that depended only on the initial number of Rs present in the synapse. Since we knew the distribution of synaptic weights at the end of the plasticity distribution (Equation 8.18), we calculated the response amplitude after all synapses reached a high or low state, and the predicted proportion of strong or high synapses matched the simulation results (Figure 8.5C).

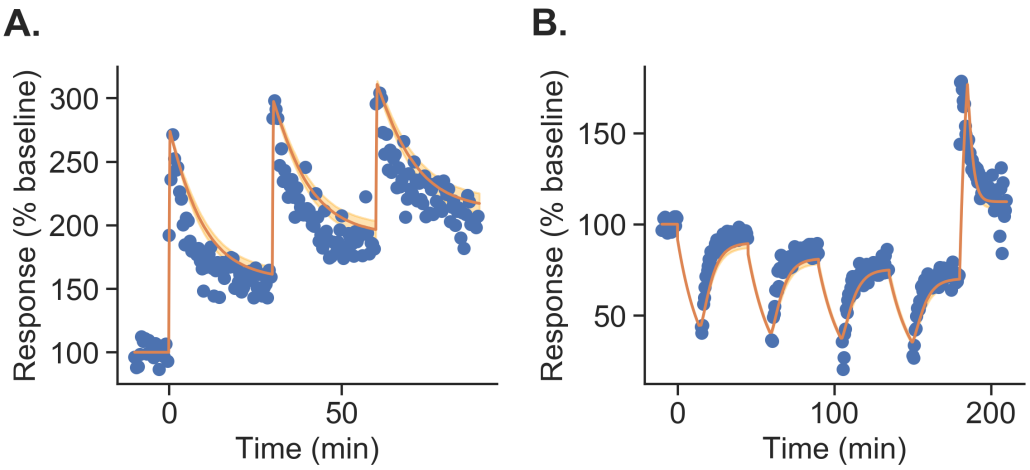


Figure 8.6: **Example fits.** (**A:**) Example model fit to a sequential LTP experiment. The recorded response at each timepoint is shown in blue, and the simulation results are shown in orange. Since the model is stochastic, the simulations were run 100 times, and the shaded region is the mean $\pm$ the standard deviation. (**B:**) Same as in (**A**) for data collected from an experiment where LTD was induced four times followed by a single LTP induction.

8.2 Results

8.2.1 The Moran Process fits experimental data qualitatively

In order to verify that the Moran process could be used as a general framework to study extracellular plasticity data, we first fit the model to both LTP and LTD experiments. The Moran Process model was fit to data from individual slices using the procedure described in Section 8.1.7, for both the sequential LTP experiments (seqLTP), which consisted of three inductions of LTP with 30 seconds of waiting time between inductions, and the sequential LTD experiments (seqLTD), which consisted of four inductions of LTD followed by a single LTP induction using TBS. An example fit for each protocol type is shown in Figure 8.6. In both cases, there was a close correspondence between the model's fit to the transient and the stable level of LTP and LTD, with a slightly better fit for the LTD protocol. In spite of only using the transient height and the stable plasticity changes to fit the data (six datapoints in total), the model also captured the exponential decay between the peak or trough to the steady state for LTP and LTD, respectively.

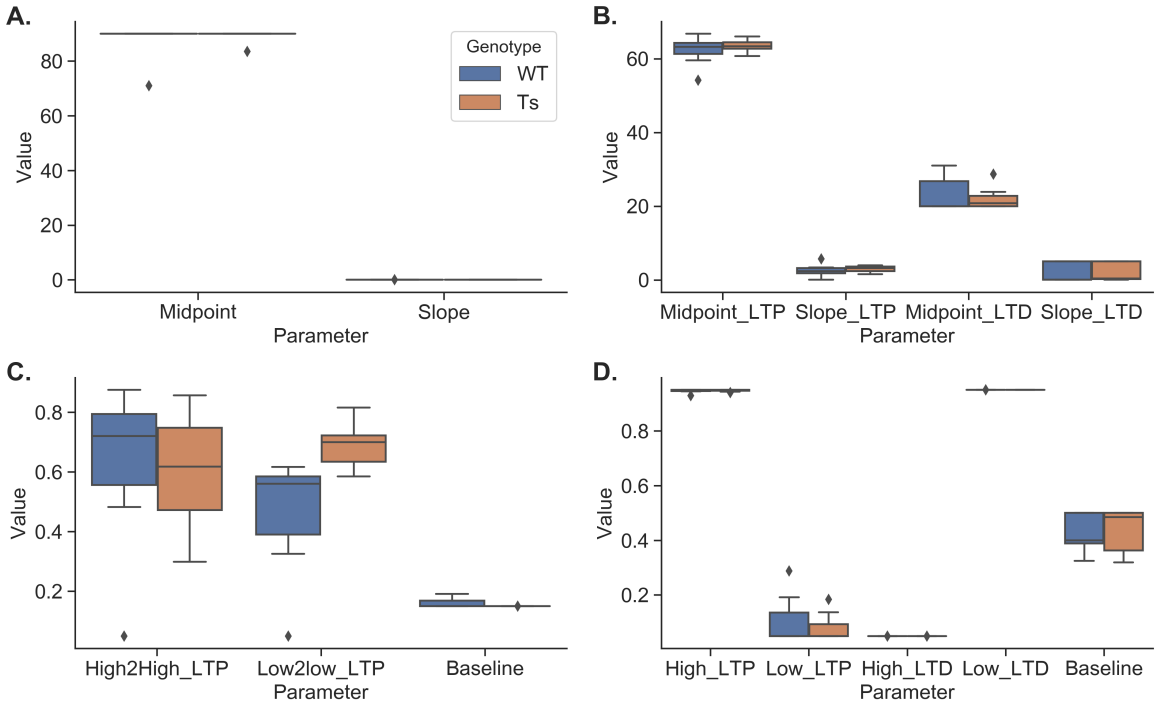


Figure 8.7: **Parameter distributions A, B:** Parameter distributions for the midpoint and slope for the selection curves in the sequential LTP experiment fits and the sequential LTD fits for wildtype (blue) and Ts65Dn (orange) slices. **C,D:** Parameter distributions for the baseline and mutation rates in the sequential LTP experiment fits and the sequential LTD fits for wildtype (blue) and Ts65Dn (orange) slices

We fit each slice experiment separately, and compared the distributions of parameter values that were used to fit wildtype and Ts65Dn slices for both seqLTP and seqLTD experiments. Since the seqLTD experiments also included a separate LTP induction, to fit the seqLTD we needed four additional parameters as compared to the LTP experiments. We found that for seqLTP experiments, there was virtually no difference in the values that were fit for wildtype and Ts65Dn animals (Figure 8.7A). For the seqLTD experiments, we found a wider spread for the parameter values that were fit, but little difference across genotypes (Figure 8.7B).

Two other parameters were informative about the transient height of the LTP and LTD experiments. These were the High-to-High rate and the Low-to-Low rate that were used during the plasticity inductions. There was a marked difference between the genotypes for the Low-to-Low parameter compared to the other two parameters (Figure 8.7C). The Low-to-Low parameter determines the probability

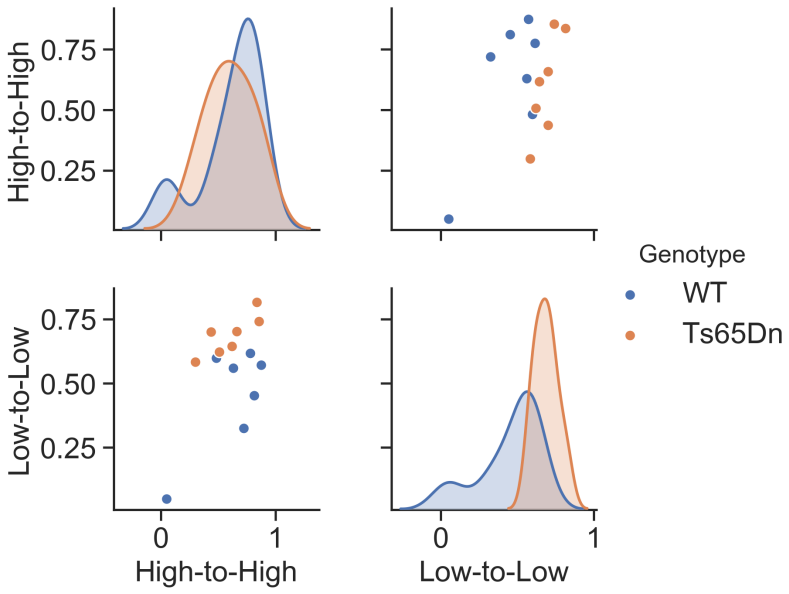


Figure 8.8: **Combinations of parameter fits can separate slices by genotype** Subplots on the diagonal show the distribution of the parameter values for wildtype (blue) and Ts65Dn slices (orange). Each row and column represents a parameter in the model, so the subplots that are not on the diagonals show the correlations between two parameters. Each point corresponds to a slice experiment coloured by strain.

that a reserve vesicle is introduced, so the higher the Low-to-Low parameter is, the more resistance a synapse at a low state will have to going into a high state. These results suggested that the decrease in LTP induced in Ts65Dn slices could be due to synapses being less likely to strengthen. Unlike the differences in the parameters for the seqLTP experiments, the plasticity parameters for seqLTD were more similar for the two genotypes (Figure 8.7D).

Even though there was a difference in the distributions of the Low-to-Low parameter in the model fits to the seqLTP experiments, there was still an overlap in the parameter fits. In other words, even though on average they were different, the Low-to-Low parameter itself is not enough to perfectly distinguish slice genotypes. As a result, we decided to look at whether parameter combinations could separate the data points by genotype. We used scatter plots to look at the relationships between two variables (Figure 8.8). Since the fits for the median and the slope of the selection curve had little variability, we focused on the High-to-High rate,

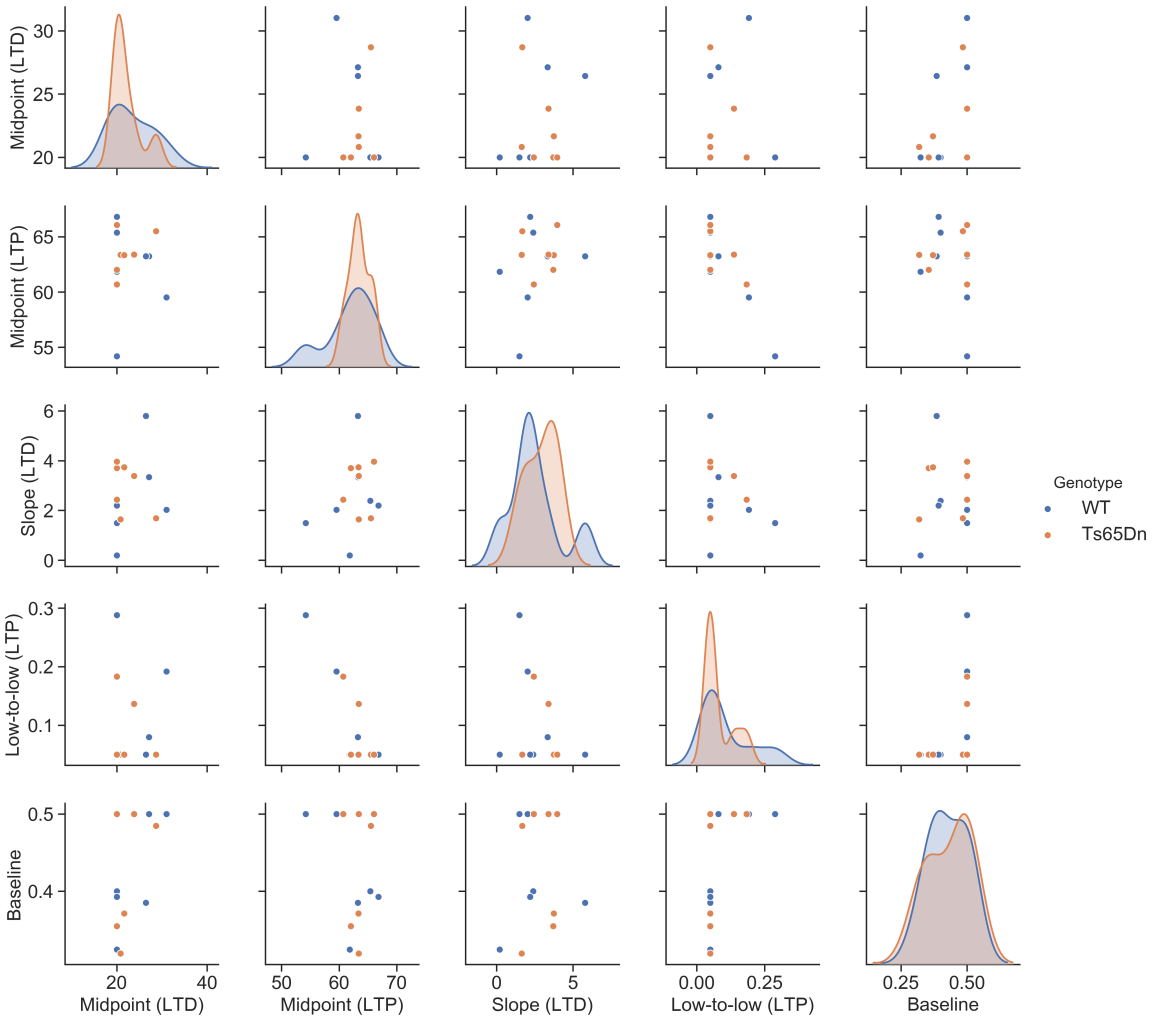


Figure 8.9: **Parameters fit to seqLTD experiments are not different across genotypes.** Subplots on the diagonal shows the distribution of the parameter values for wildtype (blue) and Ts65Dn slices (orange). In off-diagonal plots, each point corresponds to a slice experiment coloured by strain, and shows relationships between the fits of two parameters.

Low-to-Low rate. Even though for both individually (plots on the diagonal), the distributions of wildtype and Ts65Dn are overlapping, the scatter plot with the values of the High-to-High and Low-to-Low parameters for each slice lead to higher discriminability between the two genotypes.

We also looked at the pair-wise correlations between parameters for the parameters fit to the seqLTD experiments. The parameters we looked at were the midpoint of the selection curve, the slope of the LTD selection curve, the Low-to-Low and the baseline parameters (Figure 8.9). None of the pair-wise relationships between pa-

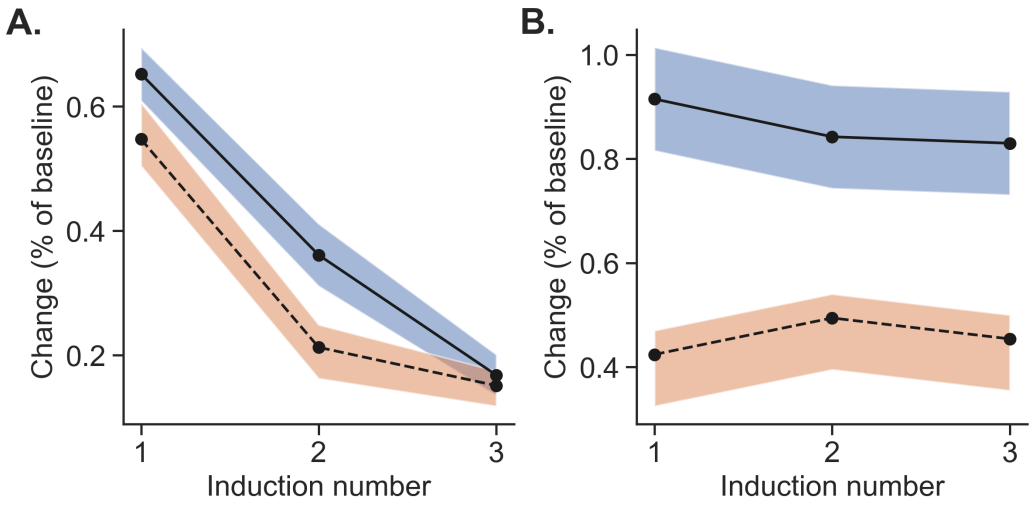


Figure 8.10: **Comparison to the data.** (**A:**) Mean changes in the stable levels of LTP from simulations using the parameters fit to the data for wildtype (solid line; blue), and Ts65Dn (dotted line; orange). The shaded regions are the standard deviation. (**B:**) Mean changes in transient levels of LTP in the simulations.

rameters could separate wildtype from Ts65Dn. This confirmed that in the seqLTD experiments, synapses in wildtype and Ts65Dn slices responded more similarly to the plasticity induction. In other words, these results suggested that performing LTD first, enabled synapses in Ts65Dn slices to become less resistant to strengthening as compared to synapses in WT slices. As a result, the distributions of parameters reflected patterns observed in the data. For seqLTP experiments, where we observed differences in the amplitude of LTP for WT and Ts65Dn animals, the parameter fits were also different. However, for seqLTD experiments where all slices behaved in similar ways regardless of genotype, the best fit parameters were also similar. Our results thus far suggested that synapses in Ts65Dn animals have a higher resistance to strengthening, and as a result fewer synapses responded to the LTP induction protocol compared to WT synapses. However, LTD makes allows Ts65Dn synapses and WT synapses respond to the LTP protocol.

8.2.2 Evaluating the performance of the model

As a comparison to the data, we plotted the height of the transient and stable components as a percentage of the baseline for each induction (Figure 8.10). The simulated LTP for both genotypes qualitatively resembled the data, as the stable height at the first induction was larger for wildtype than Ts65Dn slices, but by the third induction, the amount of LTP induced was similar across genotypes (Figure 8.10A). On the other hand, we found that the trend in the PTP amplitude shown by the model was qualitatively different from the trend observed in the data. The PTP amplitude had increased for both genotypes after the first LTP induction in the data (Figure 7.3). The simulated PTP amplitudes in the simulations for the Ts65Dn increased between the first and second induction, but decreased for all inductions in the simulations of WT data (Figure 8.10B). However, like in the data, the wildtype PTP amplitudes were larger than the Ts65Dn transient heights at every induction, and the difference in the PTP amplitudes between genotypes was also larger than the differences in the stable LTP amplitudes. Therefore, a model where the PTP amplitude was linked to the LTP amplitude, captured the main features of the data including the trends that we had previously concluded suggested that the PTP and the LTP amplitudes were unrelated.

8.2.3 Saturation determines stable LTP levels

Since the Moran Process model captured the data qualitatively, we then wanted to apply the model to understand the concept of saturation. In the data we observed that the amount of LTP that can be induced in response to the same induction protocol was much smaller after the first induction. This decrease in LTP induction makes intuitive sense as synapses cannot continue to strengthen forever, and if synapses reach their maximum weight, then they can no longer change in response to plasticity induction protocols. We asked whether the smaller LTP amplitude in Ts65Dn animals is because their synapses are closer to saturation. This is a question that is impossible to answer using extracellular recordings, because all the responses are

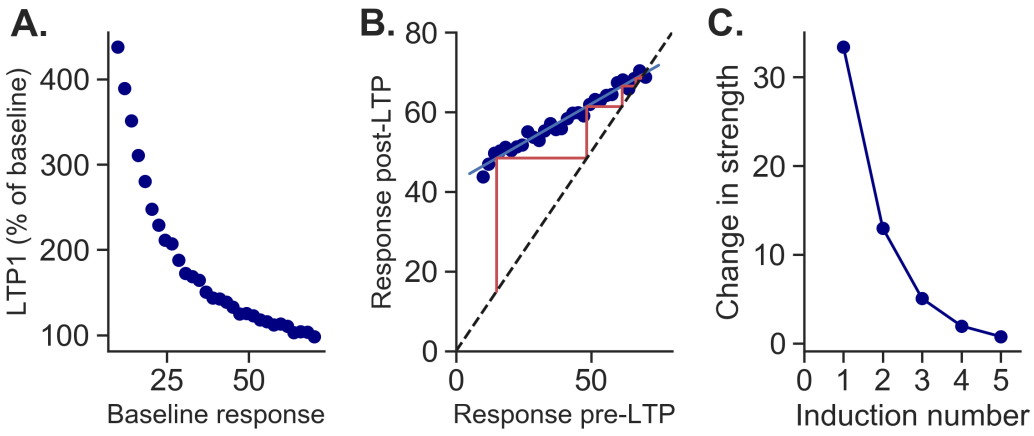


Figure 8.11: **Saturation levels predict amount of stable LTP.** (**A:**) The amount of LTP induced as a percentage of the baseline plotted against the baseline response strength. (**B:**) The response strength after LTP induction given a starting response strength. The black dotted line is the identity line. The red lines show the sequential predicted changes in the LTP given sequential inductions of LTP. (**C:**) The change in response strength at each LTP induction.

measured as a percentage of the baseline response, and are only relative measures of synaptic strength.

The model was simulated for the set of parameters of the best fit to wildtype data varying only the baseline response. We then plotted the stable component of LTP as a percentage of the baseline against the true baseline response amplitude (Figure 8.11A). We observed that weaker baseline responses led to larger LTP amplitudes. When we plotted the response strength after induction against the baseline response before induction, we found that all the points fell roughly on a line with a slope that is lower than 1 (Figure 8.11B). We noticed that at the level of LTP inductions, the Moran process model of synaptic plasticity behaved like a map in a discrete-time dynamical system, where each time-step corresponded to a plasticity induction. The map will have the form,

$$a(t+1) = m \cdot a(t) + b, \tag{8.19}$$

where $a(t)$ is the response amplitude at induction number t where $t \in \mathcal{Z}^+, 0 < m < 1$ is the slope of the line, and $b > 0$. The point where $a* = a(t+1) = a(t)$ is a stable

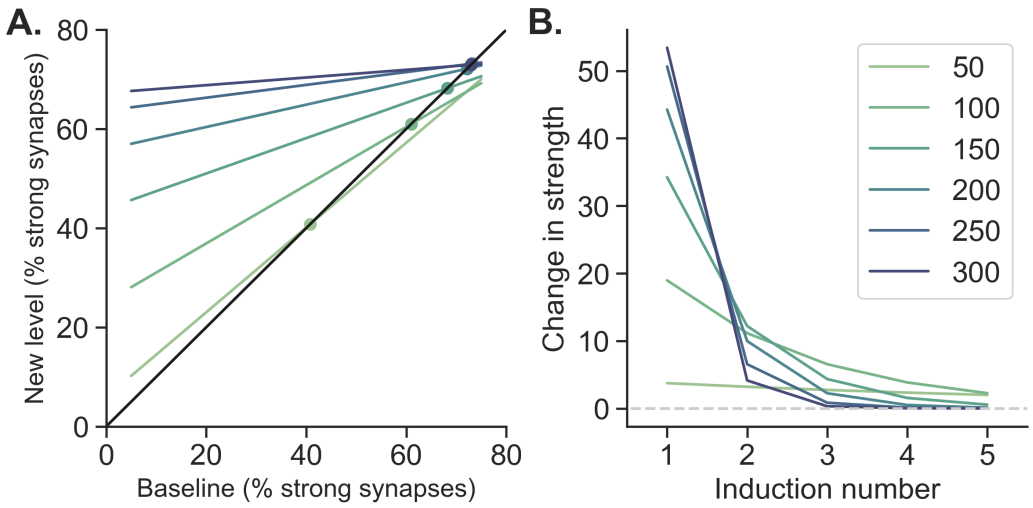


Figure 8.12: **Different plasticity induction lengths influences saturation rates.** (**A:**) The strength of the response after induction is plotted agains the initial response strength for different plasticity induction durations (in simulation steps). The black line is the identity line, and the intersection between each line and the identity line is a stable fixed point. (**B:**) The response strength change after induction for sequential inductions, where each colour corresponds to a different plasticity induction duration.

fixed point of the system, because if we start below $a*$, the next plasticity induction will bring it back up towards the fixed point and vice versa. We found that the size of the increment at each step decreases nonlinearly (Figure 8.11C). As a result, the model supported the hypothesis that if a synapse is close to the stable fixed point or saturation, the same LTP induction protocol will produce a smaller change in the response strength.

The speed at which we expect to reach the fixed point was determined by the slope of the line m. For small m values, the response post-LTP after the first induction will already be close to the fixed point. On the other hand, for values of m close to 1, the change at each induction will be small. We predicted that we could vary the slope of the map by using different durations of the plasticity protocol. The length of the plasticity protocol determined how close the distribution of synaptic weights was to the stationary distribution of the Moran process. We found that the slope of the map was steeper for short durations than for longer plasticity protocols, which corresponds to smaller changes at each induction step (Figure 8.12A).

The fixed point was lower for shorter plasticity inductions, which suggested that even though the stationary distribution was determined by the mutation matrix, the duration of the plasticity protocol determined whether or not synapses will be able to reach this distribution. The duration of the plasticity protocol also determined how quickly plasticity saturated (Figure 8.11B). Although longer durations fit the experimental results, these results also suggested that the regime of learning *in vivo* might behave differently as the plasticity events may be shorter in duration and have more incremental and modest changes in synaptic strengths.

8.3 Discussion

In this chapter we built a model that could capture the the temporal changes to the responses to extracellular stimulation in LTP and LTD experiments by adapting the Moran process. Our model originates in evolutionary biology, and although it is relatively simple, and was fit using only six time points, the model captured the time-course of both LTP and LTD experiments and the distinct trends in the stable and transient components of LTP. The model supported the idea that the smaller the LTP amplitude was, the closer the synapses in a population were to saturation. Furthermore, it predictsed that the PTP is not just an artefact as has been thought before, but that by understanding the mechanisms that produce a particular PTP amplitude, we may gain insights into the learning deficits of Down Syndrome.

There were aspects of the data that the model did not fully capture. In particular, the data showed an increase in the transient response following the second and third LTP induction, whereas the simulations showed a decrease in the response. This is possibly due to the fact that all of the synapses in the simulated network were independent, whereas we know that this is not the case for the synapses recorded in the experiments. In the brain, neurons are interconnected and synaptic changes may be coupled, as is clear from the fact that network dynamics, measured by the shape and duration of the population spike in extracellular recordings, change in

response to plasticity inductions (Bliss & Lømo, 1973).

Although the ability of synapses that behave according to the Moran process to learn and store memories was not the emphasis of this work, the synapses in our model are similar to several models of complex synapses and plasticity that have been previously published (Fusi et al., 2005; Lahiri & Ganguli, 2013; Nguyen-Vu et al., 2017). In these models, a synapse has multiple states and transition probabilities between them much like in the diagram shown in Figure 8.2B. In these models, the analysis focus was on the impact of complex synapses on learning and memory. In Fusi et al. (2005), memory capacity is quantified in a network that has complex synapses, and Lahiri & Ganguli (2013) explores different topologies of the synaptic states and their effects on memory capacity and signal to noise ratios. In Nguyen-Vu et al. (2017), complex synapses are used to explain a particular set of learning impairments in a VOR task. Where these papers focused on the functional advantages of having multi-state complex synapses, one thing that sets our model apart from theirs is that we can capture the time-course of the weight changes in response to plasticity induction protocols, that the cascade models and similar models ignore. The similarity in the synaptic structures of cascade-like models and ours suggests that there is an opportunity to develop a general model that can be used to bridge between electrophysiological experiments, and theoretical work on memory, and should be explored in future studies.

A second opportunity that our model provides is to provide a bridge between intracellular and extracellular plasticity experiments. In Chapter 6, we found that a statistical model that was successful to predict changes in intracellular plasticity experiments could not explain extracellular data adequately. What the contributions of individual cells are to the responses measured in population responses are currently unknown. The Moran process model presents us with an opportunity to explore this question as we can keep track of the changes in the contribution of individual neurons to the population response. We found that there were a diversity of responses from synapses that remain in their initial state throughout the simu-

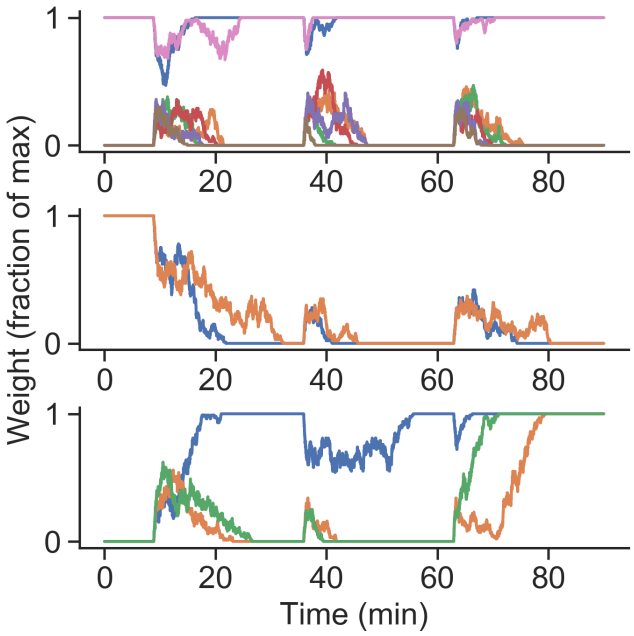


Figure 8.13: **Changes in individual synapse strengths over time.** Individual synapse strengths, or the proportion of RRVs in the synapse are plotted over time. Some synapses move away from the steady state at the time of induction, but return to their orignal state (*top*), other synapses transition from an up-state to a down-state (*middle*), and some synapses switch from a down-state to an up-state in one of the inductions (*bottom*).

lation, and other synapses that switch from being in a down-state to an up-state or vice versa (Figure 8.13). We predict that unlike intracellular recordings where an LTP protocol is thought to induce an increase in the responses of the postsynaptic neuron, in extracellular recordings, given the lower specificity of the input pairing, some synapses will strengthen LTP induction, but there may be synapses whose weights do not change or whose weights decrease.

Understanding how individual cells contribute to LTP, and especially the PTP, can lead to better models of learning and memory. For example, in our model, the 'population' response saturated or stopped changing, but the same population response could be measured even if the identity of the synapses that were in an up-state changed. In fact, if one takes a model like the Hopfield model for memory, at an experimental level, the population response of a Hopfield network would be constant as half of the neurons respond to every stimuli. Observing no change

in an LTP experiment does not necessarily imply that there was no change in the representation of a stimuli that is saved in the connections of the network.

At a practical level, our model suggests that the transient is indicative of the amount of LTP that will be induced, and that it might be a much better predictor of slice strain. As the differences in the transient height were much larger than in the stable height, we suggest that a more robust measure of the efficiency of particular treatments should take into account the changes in the transient height as well. For example, Costa & Grybko (2005) reported a difference in LTP between Ts and wildtype with TBS protocols used here but not with high-frequency stimulation, which is a different protocol used for LTP induction. In their data, the transient peak size is very different between strains in the TBS protocol, but not in the HFS protocol. This suggests that there is a protocol dependence to plasticity induction as well, and in the future it would be interesting to look at how the parameters in the Moran model correspond to changes in individual protocols as it might give us insights into the different pathways that each pattern of stimulation activates, and whether it can induce plasticity more effectively or less effectively.

Beyond helping us understand the differences between wildtype and Ts65Dn animals, the Moran model of synaptic plasticity is well-positioned to bridge between experimental data and other theoretical models of plasticity, metaplasticity and learning. On the one hand, the Moran model can reproduce the temporal dynamics of plasticity experiments, and on the other hand, it shares many similarities with existing cascade-like models of learning and memory. However, a plausible biological implementation of the Moran model is unclear, and more experiments are required in order to understand how different parameters are related to the particular features of the induction protocols.

Chapter 9

Conclusion

"It is good to have an end to journey towards; but it is the journey that matters, in the end. "

-Ursula LeGuin

There is a growing notion that the purpose of computer simulations of the brain is to enable experiments to be carried out virtually, and that "a meticulous virtual copy of the human brain would enable basic research on brain cells and circuits or computer-based trials" (Requarth, 2003). However, obtaining a meticulous virtual copy of the developing mouse brain would require an extraordinary number of experiments, and would defeat the purpose of using 'virtual' experiments to reduce the number of animal experiments. Throughout this thesis we have provided examples that a meticulous virtual copy of the brain is not required to make models useful, and that useful models are often simple. We found that building models that could address specific questions was sufficient to design future experiments and provide new interpretations of the data. Altogether, we propose that modelling and experiments can be used together and that specific questions and hypotheses can serve as bridges between both approaches.

9.1 Summary and Discussion

In this thesis, we explored different modelling strategies that can be used to help deepen our understanding of experiments and generate experimentally testable hypotheses. Synaptic plasticity and GABAergic systems were central to the two experimental systems studied. In Part I, we studied a transient inhibitory circuit that exists for a week in the developing barrel cortex before disappearing (Marques-Smith et al., 2016), and in Part II, we studied the changes in synaptic plasticity in Down Syndrome where it has been suggested that excess or altered GABAergic transmission could be contributing to the impairment in plasticity (Deidda et al., 2015; Contestabile et al., 2017).

In Chapter 3 we used spiking neuronal networks to test the hypothesis that activity dependent mechanisms could drive the connectivity changes observed in the first postnatal week. The biggest challenge with this approach was the number of parameters that are needed to simulate a spiking network. Even though it was possible to choose parameter values based on previous modelling studies and experimental data, it is clear that the brain in the first week after birth is very different to the adult brain. Using parameter values that may or may not be similar to those in the developing brain limits the model's applicability to provide new insights into the data as it introduces a significant amount of uncertainty into the modelling results.

However, by focusing on the effects of parameters that are important to development such as the level of activity in the neural network, we were able to identify a link between the amount of input activity and the connection strength of the inhibitory neuron. This link allowed us to resolve apparent conflicts in the experimental data and provided a very firm set of experiments that are now being carried out in the Butt lab. In particular, we predicted that the ION transection increases rather than decreases activity in barrel cortex during the first postnatal week, and that the amount of activity in the Celsr3 cortex isolé animal would be comparable to that in widltype animals.

In Chapter 4, we shifted our attention away from the mechanisms that drive the changes in connectivity to the potential role that they may play during development. Since the loop is present in barrel cortex but not in primary visual cortex we reasoned that this inhibitory circuit could be involved with processes specific to barrel connectivity. Even though by P4, the structural plasticity critical period of thalamic afferents onto L4 has ended (Fox, 1992), from P5 to P7, there continue to be changes in L4 like the polarisation of L4 neuron dendrites towards the centre of the barrels (Espinosa et al., 2009). We tested the hypothesis that the L5b-L4 loop could be involved in the refinement of the thalamocortical connectivity into L4. We predicted that as long as the inhibitory neurons in L5b are tuned to primarily one whisker input, then they can help generate and enforce a stable receptive field for SSNs in L4. While these results did not generate experimentally testable predictions, they did place constraints on the type of inhibition that the L5b SST+ interneurons should provide. These predictions seem consistent with published data on thalamocortical afferents in L5b during this time period (Agmon et al., 1993). Finally, this model provides a theoretical example of a developmental cortical circuit scaffold, because even though we show that the output cell formed a receptive field even in the absence of inhibition, the reciprocal connections between the inhibitory neurons and the output neuron enforced the emergence of particular connectivity patterns.

In Part II of the thesis we studied synaptic plasticity in Down Syndrome. Down Syndrome in humans is caused by a triplication of chromosome 21 and as a result it is a complex disorder where the exact relationship between the genetic origin and the phenotypic manifestation of the disease are not completely understood. There is evidence that GABAergic transmission is altered in a mouse model of DS called Ts65Dn (Contestabile et al., 2017). Modelling was used to formalise the hypothesis that Ts65Dn synapses are more saturated at baseline, or in other words, the baseline connection strength of synapses is closer to the maximum connection strength than wildtype synapses. Even though we know that GABA probably plays a central

role in the synaptic plasticity changes observed experimentally, we chose abstract models that did not explicitly model GABAergic transmission. Part of the rationale to do this was the observation from Part I that using models with many parameters whose values probably differ from the values obtained from wildtype animals greatly limited the interpretability of the simulation results.

In Chapter 6 we fit the data using statLTSP, which is a model that proposes that long-term potentiation (LTP) induces a change in the synaptic weight that corresponds to the optimal path for the synapse to reach a particular response distribution. We chose this model because it fits both the baseline and the optimal target so it would automatically provide a measure of synapse strength saturation. However, just like in Chapter 3, we once again faced difficulties obtaining the model parameter values from data, even though this time we only had three parameters to fit. We observed that statLTSP did not predict the changes in release probability, p, and quantal amplitude, q, beyond the first LTP induction. However, we found that the angular error measure used by Costa et al. (2017) failed to fully capture the discrepancy between the predictions from the model and the data for the second and third induction. We then measured the length of the predicted trajectory and compared it to the actual length and found that the model's predicted magnitude of change in q-p space worsened after the first LTP induction. These results contradicted the prediction of statLTSP that the synapse optimally approaches the target weight and distribution, and instead, suggested that the probability of release, p, and the quantal amplitude, q, will continue to change after the first LTP induction, even though the mean of the responses saturated.

The second part of the chapter focused on extending the analysis of statLTSP to large values of release sites, N, as we noticed that this analysis was missing from the original work. We were surprised to find that the system did not have a stationary point at $(\phi, 1)$, but rather a point that acted *de facto* as an attractor of the system. The fact that the target is not a stable fixed point, and that gradients are unbounded complicated the analysis of the system as it is not particularly well-

behaved. Furthermore, when we increased N, the location of the attractor and the shape of the landscape changed. The most striking feature of these changes was that the p-nullcline and the q-nullcline approach each other as N gets larger and eventually gave rise to a line of attractors. This suggested that as the number of release sites increase, the system no longer predicts that the response distribution will approach a distribution centred at the target with minimal variance. However, in practice, the predicted direction of change in p and q in the part of the phase space where most of the data in Costa et al. (2017) was found would be unaffected by the changes in the nullclines due to the increase in N.

In Chapter 7, we moved away from statLTSP and used a phenomenological model that captured the qualitative features of the data. Unlike in Chapter 6 where we fit three biologically relevant parameters from the data, in this chapter we chose a function that could reproduce the shape of the graph of fEPSP slope changes during an experiment. The function we chose was a step function plus an exponentially decaying function. We hoped to find differences in the distributions of the parameters across slices with different genotypes that could then provide clues of the biophysical quantities that could be relevant to study next. We were surprised to find that the post-tetanic potentiation (PTP) height was the most consistently different parameter between wildtype and Ts65Dn slices. We proposed that since PTP is thought to be mediated presynaptically, this could suggest that the reason that plasticity was impaired in Ts65Dn could be located presynaptically. Nonetheless, we wondered whether the differences in PTP height was linked to the deficits in LTP induction, as there were no suggestions of this in the literature, and the correlations between PTP and LTP were weak at best.

In Chapter 8 we adapted a model that had been used in evolutionary biology that linked the transient and stable components of the data. We used this model as a proof of principle that it is possible for there to be a link between these two aspects of the data, and that they could fit the qualitative aspects of the data with relatively few parameters and using few data points. Using this model we tested

the notion that the size of the LTP induction could be tied to the saturation of a slice. Furthermore, given that the model was stochastic, individual synapses could switch from being in a high state to a low state or vice versa following a plasticity induction even if the mean response stayed the same. We wonder if the synapse-level stochasticity but slice-level stability could be related to the fact that p and q continued to change for the second and third LTP inductions even though statLTSP would have predicted otherwise. Finally we found that in the Moran model of synaptic plasticity, the sequential inductions of LTP could be interpreted as a map in discrete dynamical systems with a single stable fixed point. As a result, in addition to fitting the data qualitatively, this model was able to provide an explanation as to why the amount of LTP that can be induced decreases dramatically after the first induction.

9.2 Future Directions

There are several ways to expand each one of the models presented in this thesis. In Chapter 3, we proposed experimentally testable hypotheses that are being carried out in the Butt lab. Furthermore, we included very few neuronal types in the spiking neural network, and future work could incorporate additional types of neurons that are thought to play an important role during development like subplate and PV+ interneurons. For example, a simple extension in the model could look at mechanisms through which SST+ and PV+ interneurons compete to inhibit L4 SSNs, and how this changes the current predictions of the model. Chapter 4 was a largely exploratory chapter. One of the advantages of rate networks and the simple neuron model used, is that it is much easier to do analytics on the model's behaviour. Future work on this model could focus on analytically deriving the conditions under which the L5b-L4 loop could help refine barrel maps in L4 during development.

In Part II we used three models to help us understand the differences in learn-

ing between wildtype and Ts65Dn animals. When we carried out an analysis of statLTSP in Chapter 6, we found that the original paper had not discussed the effect of varying the number of release sites on the predicted changes in presynaptic and postsynaptic components leading to a particular postsynaptic response. One of the surprising findings of our analysis was that as the number of release sites increases, a line of fixed points emerges so the system no longer predicts that there is a single target distribution of responses. Although in practice it does not affect any of the results presented in Costa et al. (2017), this suggests that perhaps a distinct, more well-behaved metric should be chosen for statLTSP.

In Chapter 8, we adapted the Moran process from evolutionary biology to describe extracellular synaptic plasticity experiments. We hope that one of the applications of the Moran model of synaptic plasticity will be to help us understand the differences between the plasticity experiments recorded intracellularly and those recorded extracellularly. Even though the results from both kinds of experiments are often discussed as reflecting the same phenomenon, our experience extending statLTSP to data recorded from extracellular setups challenges this perception. There are key differences in the methodology to obtain plasticity data with extracellular and intracellular recordings that make it likely for these experiments to provide insights into different aspects of synaptic plasticity. A striking example is the difference in the the strength of electrical stimulation used to collect the baseline response in both setups. In intracellular plasticity recordings, baseline responses are obtained using minimal stimulation, where about half of the electrical stimulations fail to evoke a postsynaptic response (Isaac et al., 1996). On the other hand, in extracellular recordings, baseline responses are collected using electrical pulses that evoke responses where the amplitude of the response is half of the maximum response (Deidda et al., 2015).

Finally, future experiments should be carried out to clarify the relationship between the PTP and LTP amplitudes. One way to do this is by studying the relationship between the induction protocol and the PTP. Costa & Grybko (2005) found

that there were differences between wildtype and Ts65Dn animals only with one of the two protocols used, and looking at their figures, it seems that the same applies to the differences in the PTP levels induced by each protocol. Recent work used nonlinear dynamical systems to systematically determine the relationship between the stimulation protocol and synaptic plasticity changes (Gastaldi et al., 2018). In a similar spirit, we could design different stimulation protocols that could help us determine the relationship between the transient and stable components of plasticity and their dependence to, for example, the frequency and duration of stimulation.

9.3 On making useful models and on making models useful

My thesis has used five distinct models to study and understand two datasets. The primary goal to build these models was to make them insightful by providing experimentally testable predictions. However, I quickly learned that making models useful was difficult, starting from the inherent methodological differences between theoretical and experimental approaches. In experiments one begins with a complex system and the challenge is to design a controlled experiment where the effect of varying specific factors can be understood. In modelling one can choose the amount of complexity one starts with, and the insights arise from the behaviour of the model, the aspects of the data it can capture, and most importantly where it fails. This may sound counterintuitive, but understanding what makes a model fail makes the next, more complex, iteration of the model more interpretable, as we then know exactly what we have gained by the added complexity.

The main challenge I faced in trying to build useful models came from this methodological difference, because my models were often missing a large portion of the complexity that experimentalists know exists in the brain. ”Can you just add (fill in the blank) to your model?” became my most dreaded question. While it was usually possible to add an extra component to the model, it was not always clear to

me whether adding this extra component would simply make my model describe the data better, or whether it would provide any real insights into the biological system. As Phillips (2015) puts it:

> Often in cell biology, when theoretical modeling takes place, it is as a figure 7 reflection on experiments that have already been done, with data fitting providing a metric of success. Figure 1 theory, by way of contrast, is about living dangerously by turning our thinking into formal mathematical predictions and confronting that math with experiments that have not yet been done.

As much as possible, I wanted to learn how to do figure 1 theory, because I wanted to create models useful scientifically, rather than making models useful for publishing. Nonetheless, as Chapter 7 and Chapter 8, building purely descriptive models is sometimes the first step to making a more insightful data. This is the case especially when we want to build a formal mathematical framework that can inform us about particular features of the data.

In the end, what worked best was to focus on the questions. Theoretical and experimental approaches are opposite in some ways, but questions provided a common middle ground. Understanding the questions that my experimental collaborators were most interested in finding answers to, but were limited by resources or technology, was key to finding the niche where my models could be the most useful. Communication was challenging and there were many meetings that seemed to go around in circles, but keeping the common goals and questions at the centre of the conversation was crucial to reaching a mutual understanding and defining future steps. It is true that useful models can fill in gaps in knowledge that experiments cannot reach. However, to make models useful, it may be the case that we should aim to create modest models, or in other words, models that do not explain everything but rather provide a scaffold that future experiments and models can continue to build upon.

Appendix A

Table of Abbreviations

Abbreviation	Meaning
SP	Subplate
SSN	Spiny Stellate Neuron
SST+	Somatostatin-positive
PV+	Parvalbumin-positive
Th	Thalamus
L4, L5b...	Layer 4, Layer 5b ...
P1,P2...	Postnatal day 1, Postnatal day 2,...
ION	infraorbital nerve
LIF	leaky integrate-and-fire
STDP	spike-timing dependent plasticity
iSTDP	inhibitory spike-timing dependent plasticity
eSTDP	excitatory spike-timing dependent plasticity
PCA	Principal Components Analysis
PC1, PC2 ...	first principal component, 2nd principal component...
DS	Down Syndrome
WT	wildtype
Ts65Dn	mouse model for Down Syndrome
CA1	hippocampus area
fEPSP	field excitatory post-synaptic potential
PTP	post-tetanic potentiation
LTP	long-term potentiation
LTD	long-term depression
Tr	transient
St	stable
PSP	post-synaptic potential
statLTSP	statistical long-term synaptic potentiation
seqLTP	sequential LTP inductions
seqLTD	sequential LTD inductions with one LTP induction

Appendix B

Stationary points of the original statLTSP formulation

In this Appendix, we outline the calculations shown in the Methods of Costa et al. (2017), and show that it is not possible to confirm the stationary point using the partial derivative of the KL-divergence between a Gaussian and a Dirac-delta function.

Equation 6.3 can be expanded as follows:

$$KL(P^{bound}||P^{PSP}) = \int P^{bound} \ln \frac{P^{bound}}{P^{PSP}} dX \tag{B.1}$$

$$= \int \partial_\phi \ln \frac{\partial_\phi}{\mathcal{N}(\mu, \sigma^2)} dX, \tag{B.2}$$

$$= \int \partial_\phi \ln \partial_\phi dX - \int \partial_\phi \ln \mathcal{N} \tag{B.3}$$

$$= \int \partial_\phi \ln \partial_\phi dX - \left[\ln \frac{1}{\sigma\sqrt{2\pi}} + \ln \exp\left(-\frac{(\phi-\mu)^2}{2\sigma^2} \right) \right] \tag{B.4}$$

$$= \int \partial_\phi \ln \partial_\phi dX - \ln \frac{1}{\sqrt{2\pi}} + \ln \sigma + \frac{(\phi-\mu)^2}{2\sigma^2} \tag{B.5}$$

$$= \int \partial_\phi \ln \partial_\phi dX - \ln \frac{1}{\sqrt{2\pi}} + \ln \sqrt{Nq^2p(1-p))} + \frac{(\phi-Npq)^2}{2Nq^2p(1-p)}. \tag{B.6}$$

The first two terms do not play a role as they do not contain N, p, or q. In order to find the gradients we ignored the first two terms and took the derivative of the last two terms with respect to p and to q. Doing so, we obtained the following

expressions for the partial derivatives of KL:

$$\frac{\partial KL}{\partial p} = \frac{1-2p}{2p(1-p)} + \frac{Npq-\phi)(p(Nq-2\phi)+\phi)}{2N(1-p)^2p^2q^2}, \tag{B.7}$$

$$\frac{\partial KL}{\partial q} = \frac{1}{q} + \frac{\phi(\phi-Npq)}{(p-1)Npq^3}. \tag{B.8}$$

The model's predictions on how p and q change after a plasticity induction is determined by these equations. Given an initial point in $q-p$ space, p changes according to Equation B.7, and q changes according to Equation B.8.

Costa et al. (2017) states that there is a stationary point at $q=\phi$ and $p=1$, so the first step we took was to confirm that it was indeed a stationary point but were surprised to find that $(\phi,1)$ does not meet the requirements to be a stationary point. By definition, a stationary point is a point where the gradient is zero, or in other words, where $\partial \mathrm{KL}/\partial p = 0 = \partial \mathrm{KL}/\partial q$. Equations B.7 and B.8 are both undefined at $p=1$ so we calculated the values of the gradients at the limit as $p \to 1$.

We began by showing that $\partial KL/\partial q \to 0$ as $(q,p) \to (\phi,1)$

$$\begin{aligned}\left.\frac{\partial KL}{\partial q}\right|_{(1,\phi)} &\to \frac{1}{\phi} + \lim_{p\to 1^-}\left(\frac{\phi(\phi-Np\phi)}{(p-1)Np\phi^3}\right)\\ &\to \frac{1}{\phi} + \lim_{p\to 1^-}\left(-\frac{1-Np}{(1-p)Np\phi}\right)\end{aligned}$$

If $N=1$, then the limit is $-1/\phi$, so $\partial KL/\partial q = 1/\phi - 1/\phi = 0$.

Next we checked whether $\partial KL/\partial p \to 0$ as $(q,p) \to (\phi,1)$.

$$\left.\frac{\partial KL}{\partial p}\right|_{(\phi,1)} \to \lim_{p\to 1^-}\left(\frac{1-2p}{2p(1-p)}\right) + \lim_{p\to 1^-}\left(\frac{(Np\phi-\phi)(p(N\phi-2\phi)+\phi)}{2N(1-p)^2p^2\phi^2}\right)$$

if we let $N = 1$,

$$\to -\infty + \lim_{p\to 1^-}\left(\frac{\phi^2(p-1)(-p+1)}{2(1-p)^2p^2\phi^2}\right)$$

$$\to -\infty + \lim_{p\to 1^-}\left(-\frac{(1-p)^2}{(1-p)^2}\cdot\frac{1}{2p^2}\right)$$

$$\to -\infty - \frac{1}{2}$$

$$\to -\infty$$

This contradicted the definition of a stationary point, so the statement that $(\phi, 1)$ is a stationary point of the system is false. However, we noticed that $(\phi, 1)$ behaves as an attractor. Since $\left.\frac{\partial KL}{\partial p}\right|_{(\phi,1)} \to -\infty$, the predicted direction of change in p is a positive update in gradient descent. In effect, this means that all trajectories close to $p = 1$ will approach $p = 1$, but unlike a typical fixed point, or line of fixed points, the velocity of approach will increase infinitely close to this line. Furthermore, $\left.\frac{\partial KL}{\partial q}\right|_{(\phi,1)} = 0$ when $N = 1$. Since $-\frac{\partial KL}{\partial q} \to -\infty$ for $p \to 1$, and $-\frac{\partial KL}{\partial q} \to \infty$ for $q > \phi$, the point $(\phi, 1)$ is stable in q as well. Therefore, in practical terms, all trajectories will approach $(1, \phi)$ as if it were a stationary point. In Chapter 6, we start by assuming that the distribution of the bound has finite width and taking the limits as variance goes to zero only after calculating the partial derivatives.

Bibliography

Abbott, B. et al. (2016). Observation of Gravitational Waves from a Binary Black Hole Merger. Physical Review Letters, 116, 061102.

Abraham, W. C. & Bear, M. F. (1996). Metaplasticity: the plasticity of synaptic plasticity. Trends in Neurosciences, 19, 126–130.

Agmon, A., Yang, L., O'Dowd, D., & Jones, E. (1993). Organized growth of thalamocortical axons from the deep tier of terminations into layer IV of developing mouse barrel cortex. J. Neurosci.; Journal of Neuroscience, 13, 5365–5382.

Akerman, C. J. & Cline, H. T. (2007). Refining the roles of GABAergic signaling during neural circuit formation. Trends in Neurosciences, 30, 382 – 389.

Allen, B. & Tarnita, C. E. (2012). Measures of success in a class of evolutionary models with fixed population size and structure. Journal of Mathematical Biology, 68, 109–143.

Anastasiades, P. G., Marques-Smith, A., Lyngholm, D., Lickiss, T., Raffiq, S., Kätzel, D., Miesenböck, G., & Butt, S. J. (2016). GABAergic interneurons form transient layer-specific circuits in early postnatal neocortex. Nature Communications, 7, 10584.

Asim, A., Kumar, A., Muthuswamy, S., Jain, S., & Agarwal, S. (2015). Down syndrome: an insight of the disease. Journal of Biomedical Science, 22.

Ask, M. & Reza, M. (2016). Computational models in neuroscience: How real are they? A critical review of status and suggestions. Austin Neurology & Neurosciences, 1, 1008.

Babadi, B. & Abbott, L. F. (2013). Pairwise analysis can account for network struc-

tures arising from spike-timing dependent plasticity. PLoS Computational Biology, 9, e1002906.

Basheer, I. A. & Hajmeer, M. N. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of microbiological methods, 43 1, 3–31.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. Neuron, 76, 695–711.

Bekkers, J. M. & Clements, J. D. (1999). Quantal amplitude and quantal variance of strontium-induced asynchronous EPSCs in rat dentate granule neurons. The Journal of Physiology, 516, 227–248.

Belichenko, P. V., Masliah, E., Kleschevnikov, A. M., Villar, A. J., Epstein, C. J., Salehi, A., & Mobley, W. C. (2004). Synaptic structural abnormalities in the Ts65Dn mouse model of down syndrome. The Journal of Comparative Neurology, 480, 281–298.

Ben-Ari, Y. (2001). Developing networks play a similar melody. Trends in Neurosciences, 24, 353–360.

Ben-Ari, Y., Gaiarsa, J.-L., Tyzio, R., & Khazipov, R. (2007). GABA: a pioneer transmitter that excites immature neurons and generates primitive oscillations. Physiological Reviews, 87, 1215–1284.

Bi, G. Q. & Poo, M. M. (1998). Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. Journal of Neuroscience, 18, 10464–10472.

Bishop, C. M. (2006). Pattern recognition and machine learning. (Springer).

Bittles, A. & Glasson, E. (2004). Clinical, social, and ethical implications of changing life expectancy in Down syndrome. Developmental Medicine and Child Neurology, 46, 282–286.

Bliss, T. V. P. & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. The Journal of Physiology, 232, 331–356.

Blythe, R. A. & McKane, A. J. (2007). Stochastic models of evolution in genetics, ecology and linguistics. Journal of Statistical Mechanics: Theory and Experiment, 2007, P07018.

Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J. M., Diesmann, M., Morrison, A., Goodman, P. H., Harris, F. C., & et al. (2007). Simulation of networks of spiking neurons: A review of tools and strategies. Journal of Computational Neuroscience, 23, 349–398.

Brito, C. S. N. & Gerstner, W. (2016). Nonlinear Hebbian Learning as a Unifying Principle in Receptive Field Formation. PLOS Computational Biology, 12, e1005070.

Brodland, G. W. (2015). How computational models can help unlock biological systems. Seminars in Cell & Developmental Biology, 47, 62–73.

Butt, S. J., Stacey, J. A., Teramoto, Y., & Vagnoni, C. (2017). A role for GABAergic interneuron diversity in circuit development and plasticity of the neonatal cerebral cortex. Current Opinion in Neurobiology, 43, 149–155.

Chakrabarti, L., Best, T. K., Cramer, N. P., Carney, R. S., Isaac, J. T., Galdzicki, Z., & Haydar, T. F. (2010). Olig1 and Olig2 triplication causes developmental brain defects in Down syndrome. Nature Neuroscience, 13, 927–934.

Che, A., Babij, R., Iannone, A. F., Fetcho, R. N., Ferrer, M., Liston, C., Fishell, G., & García, N. V. D. M. (2018). Layer I Interneurons Sharpen Sensory Maps during Neonatal Development. Neuron, 99, 98 – 116.

Clopath (2010). Voltage and spike timing interact in STDP - a unified model. Frontiers in Synaptic Neuroscience.

Clopath, C., Vogels, T. P., Froemke, R. C., & Sprekeler, H. (2016). Receptive field formation by interacting excitatory and inhibitory synaptic plasticity. BioRxiv, p. 066589.

Cohen-Cory, S., Kidane, A. H., Shirkey, N. J., & Marshak, S. (2010). Brain-derived neurotrophic factor and the development of structural neuronal connectivity. Developmental Neurobiology, 70, 271–288.

Colas, D., Chuluun, B., Garner, C. C., & Heller, H. C. (2017). Short-term treatment with flumazenil restores long-term object memory in a mouse model of Down syndrome. Neurobiology of learning and memory, 140, 11–16.

Constantinople, C. & Bruno, R. (2013). Deep Cortical Layers Are Activated Directly by Thalamus. Science, 340, 1591–1594.

Contestabile, A., Magara, S., & Cancedda, L. (2017). The GABAergic Hypothesis for Cognitive Disabilities in Down Syndrome. Frontiers in Cellular Neuroscience, 11, 54.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297.

Costa, A. C. & Grybko, M. J. (2005). Deficits in hippocampal CA1 LTP induced by TBS but not HFS in the Ts65Dn mouse: a model of Down syndrome. Neuroscience Letters, 382, 317–322.

Costa, R. P., Padamsey, Z., D'Amour, J. A., Emptage, N. J., Froemke, R. C., & Vogels, T. P. (2017). Synaptic Transmission Optimization Predicts Expression Loci of Long-Term Plasticity. Neuron, 96, 177–189.

Cramer, J. (2003). The Origins of Logistic Regression. SSRN Electronic Journal.

Daw, M. I. et al. (2007). Coordinated developmental recruitment of latent fast spiking interneurons in layer IV barrel cortex. Nature Neuroscience, 10, 453–461.

Debanne, D., Gähwiler, B., & Thompson, S. (1998). Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. Journal of Physiology, 15, 237–247.

Deidda, G., Parrini, M., Naskar, S., Bozarth, I. F., Contestabile, A., & Cancedda, L. (2015). Reversing excitatory $GABA_AR$ signaling restores synaptic plasticity and memory in a mouse model of Down syndrome. Nature Medicine, 21, 318–326.

Del Castillo, J. & Katz, B. (1954). Quantal components of the end-plate potential. The Journal of Physiology, 124, 560–573.

Delaney, K., Zucker, R., & Tank, D. (1989). Calcium in motor nerve terminals associated with posttetanic potentiation. Journal of Neuroscience, 9, 3558–3567.

Diamandopoulos, K. & Green, J. (2018). Down syndrome: An integrative review. Journal of Neonatal Nursing, 24, 235–241.

Dierssen, M., Benavides-Piccione, R., Martinez-Cue, C., Estivill, X., Florez, J., Elston, G., & DeFelipe, J. (2003). Alterations of Neocortical Pyramidal Cell Phenotype in the Ts65Dn Mouse Model of Down Syndrome: Effects of Environmental Enrichment. Cerebral Cortex, 13, 758–764.

Ebara, S., Kumamoto, K., Matsuura, T., Mazurkiewicz, J. E., & Rice, F. L. (2002). Similarities and differences in the innervation of mystacial vibrissal follicle-sinus complexes in the rat and cat: A confocal microscopic study. The Journal of Comparative Neurology, 449, 103–119.

Erzurumlu, R. S. & Gaspar, P. (2012). Development and critical period plasticity of the barrel cortex. European Journal of Neuroscience, 35, 1540–1553.

Espinosa, J. S., Wheeler, D. G., Tsien, R. W., & Luo, L. (2009). Uncoupling dendrite growth and patterning: single-cell knockout analysis of NMDA receptor 2B. Neuron, 62, 205–217.

Faizi, M., Bader, P. L., Tun, C., Encarnacion, A., Kleschevnikov, A., Belichenko, P., Saw, N., Priestley, M., Tsien, R. W., Mobley, W. C., et al. (2011). Comprehensive behavioral phenotyping of Ts65Dn mouse model of Down Syndrome: Activation of β1-adrenergic receptor by xamoterol as a potential cognitive enhancer. Neurobiology of Disease, 43, 397–413.

Fernandez, F., Morishita, W., Zuniga, E., Nguyen, J., Blank, M., Malenka, R. C., & Garner, C. C. (2007). Pharmacotherapy for cognitive impairment in a mouse model of Down syndrome. Nature neuroscience, 10, 411.

Filar, J. A. (2002). Mathematical models. Ph.D. thesis, UNESCO/EOLSS.

Fox, K. (1992). A critical period for experience-dependent synaptic plasticity in rat barrel cortex. J. Neurosci.; Journal of Neuroscience, 12, 1826–1838.

Frangeul, L., Kehayas, V., Sanchez-Mut, J. V., Fièvre, S., Krishna-K, K., Pouchelon, G., Telley, L., Bellone, C., Holtmaat, A., Gräff, J., & et al. (2017). Input-dependent regulation of excitability controls dendritic maturation in somatosensory thalam-

ocortical neurons. Nature Communications, 8, 2015.

Freund, Y. & Schapire, R. E. (1999). Large Margin Classification Using the Perceptron Algorithm. Machine Learning, 37, 277–296.

Friston, K. & Kiebel, S. (2009). Cortical circuits for perceptual inference. Neural Networks, 22, 1093–1104.

Fusi, S., Drew, P. J., & Abbott, L. (2005). Cascade Models of Synaptically Stored Memories. Neuron, 45, 599–611.

Gastaldi, C., Muscinelli, S. P., & Gerstner, W. (2018). Optimal stimulation protocol in a bistable synaptic consolidation model. arXiv preprint, arXiv:1805.10116.

Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2009). Neuronal dynamics: From single neurons to networks and models of cognition. (Cambridge University Press).

"Why we need computational models in biology". Ghani, N. (2016). <https://blogs.plos.org/thestudentblog/2016/07/29/why-we-need-computational-models-in-biology/>. Last Accessed: 6 May, 2019.

Gordon, J. & Stryker, M. (1996). Experience-Dependent Plasticity of Binocular Responses in the Primary Visual Cortex of the Mouse. Journal of Neuroscience, 16, 3274–3286.

Graupner, M. & Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. Proceedings of the National Academy of Sciences, 109, 3991–3996.

Gupta, M., Dhanasekaran, A. R., & Gardiner, K. J. (2016). Mouse models of Down syndrome: Gene content and consequences. Mammalian Genome, 27, 538–555.

Hebb, D. O. (1949). The Organization of Behavior. (Wiley & Sons).

Hennequin, G., Agnes, E. J., & Vogels, T. P. (2017). Inhibitory Plasticity: Balance, Control, and Codependence. Annual Review of Neuroscience, 40, 557–579.

Hensch, T., Fagiolini, M., Mataga, N., Stryker, M., Baekkeskov, S., & Kash, S. (1998). Local GABA Circuit Control of Experience-Dependent Plasticity in Developing

Visual Cortex. Science, 282, 1504–1508.

Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. Nature Reviews Neuroscience, 6, 877–888.

Herault, Y., Delabar, J. M., Fisher, E. M., Tybulewicz, V. L., Yu, E., & Brault, V. (2017). Rodent models in Down Syndrome research: Impact and future opportunities. Disease Models & Mechanisms, 10, 1165–1186.

Hille, B. (2001). Ion channels of excitable membranes, vol. 507. (Sinauer Sunderland, MA).

Hodgkin, A. L. & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology, 117, 500–544.

Hyde, L. A., Frisone, D. F., & Crnic, L. S. (2001). Ts65Dn mice, a model for Down Syndrome, have deficits in context discrimination learning suggesting impaired hippocampal function. Behavioural Brain Research, 118, 53–60.

Ikeda, K., Yanagawa, Y., & Bekkers, J. (2008). Distinctive Quantal Properties of Neurotransmission at Excitatory and Inhibitory Autapses Revealed Using Variance–Mean Analysis. Journal of Neuroscience, 28, 13563–13573.

Isaac, J. T., Hjelmstad, G. O., Nicoll, R. A., & Malenka, R. C. (1996). Long-term potentiation at single fiber inputs to hippocampal CA1 pyramidal cells. Proceedings of the National Academy of Sciences, 93, 8710–8715.

Itami, C. & Kimura, F. (2012). Developmental switch in spike timing-dependent plasticity at layers 4–2/3 in the rodent barrel cortex. Journal of Neuroscience, 32, 15000–15011.

Janvin, M. (2017). Survival of the flattest or arrival of the frequent? Mmathphys dissertation, University of Oxford.

Jiang, B., Trevino, M., & Kirkwood, A. (2007). Sequential development of long-term potentiation and depression in different layers of the mouse visual cortex. Journal of Neuroscience, 27, 9648–9652.

Kaila, K., Price, T. J., Payne, J. A., Puskarjov, M., & Voipio, J. (2014). Cation-chloride

cotransporters in neuronal development, plasticity and disease. Nature Reviews Neuroscience, 15, 637–654.

Kandell, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of Neural Science. (McGraw-Hill), 4th edn.

Kanold, P. (2009). Subplate neurons: crucial regulators of cortical development and plasticity. Frontiers in Neuroanatomy, 3.

Kanold, P. O. & Shatz, C. J. (2006). Subplate neurons regulate maturation of cortical inhibition and outcome of ocular dominance plasticity. Neuron, 51, 627–638.

Khazipov, R., Colonnese, M., & Minlebaev, M. (2013). Chapter 8 - neonatal cortical rhythms. In Neural Circuit Development and Function in the Brain, J. L. Rubenstein & P. Rakic, eds. (Oxford: Academic Press), pp. 131–153.

Khazipov, R. & Luhmann, H. J. (2006). Early patterns of electrical activity in the developing cerebral cortex of humans and rodents. Trends in Neurosciences, 29, 414–418.

Kimura, F. & Itami, C. (2019). A hypothetical model concerning how spike-timing-dependent plasticity contributes to neural circuit formation and initiation of the critical period in barrel cortex. Journal of Neuroscience.

Kirmse, K., Kummer, M., Kovalchuk, Y., Witte, O. W., Garaschuk, O., & Holthoff, K. (2015). GABA depolarizes immature neurons and inhibits network activity in the neonatal neocortex in vivo. Nature communications, 6, 7750.

Kleschevnikov, A. M., Belichenko, P. V., Faizi, M., Jacobs, L. F., Htun, K., Shamloo, M., & Mobley, W. C. (2012). Deficits in cognition and synaptic plasticity in a mouse model of Down syndrome ameliorated by $GABA_B$ receptor antagonists. Journal of Neuroscience, 32, 9217–9227.

Kondo, S., Yoshida, T., & Ohki, K. (2016). Mixed functional microarchitectures for orientation selectivity in the mouse primary visual cortex. Nature Communications, 7, 13210.

Kurt, M., Davies, D., Kidd, M., Dierssen, M., & Flórez, J. (2000). Synaptic deficit in the temporal cortex of partial trisomy 16 (Ts65Dn) mice. Brain Research, 858,

191–197.

Lahiri, S. & Ganguli, S. (2013). A memory frontier for complex synapses. In Advances in Neural Information Processing Systems 26, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger, eds. (Curran Associates, Inc.), pp. 1034–1042.

Larkman, A., Hannay, T., Stratford, K., & Jack, J. (1992). Presynaptic release probability influences the locus of long-term potentiation. Nature, 360, 70–73.

Lisman, J. (1989). A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. Proceedings of the National Academy of Sciences, 86, 9574–9578.

Liu, C., Belichenko, P. V., Zhang, L., Fu, D., Kleschevnikov, A. M., Baldini, A., Antonarakis, S. E., Mobley, W. C., & Yu, Y. E. (2011). Mouse Models for Down Syndrome-Associated Developmental Cognitive Disabilities. Developmental Neuroscience, 33, 404–413.

Luhmann, H. J., Kirischuk, S., & Kilb, W. (2018). The Superior Function of the Subplate in Early Neocortical Development. Frontiers in Neuroanatomy, 12, 97.

Lyle, R., Béna, F., Gagos, S., Gehrig, C., Lopez, G., Schinzel, A., Lespinasse, J., Bottani, A., Dahoun, S., Taine, L., & et al. (2008). Genotype–phenotype correlations in Down Syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. European Journal of Human Genetics, 17, 454–466.

Markram, H. et al. (2004). Interneurons of the neocortical inhibitory system. Nature Reviews Neuroscience, 5, 793–807.

Marques-Smith, A., Lyngholm, D., Kaufmann, A.-K., Stacey, J. A., Hoerder-Suabedissen, A., Becker, E. B., Wilson, M. C., Molnár, Z., & Butt, S. J. (2016). A transient translaminar GABAergic interneuron circuit connects thalamocortical recipient layers in neonatal somatosensory cortex. Neuron, 89, 536–549.

Marr, D. & Poggio, T. (1976). From understanding computation to understanding neural circuitry. A.I. memo, 357, 1–22.

Marx, M., Qi, G., Hanganu-Opatz, I. L., Kilb, W., Luhmann, H. J., & Feldmeyer, D. (2015). Neocortical Layer 6B as a Remnant of the Subplate - A Morphological Comparison. Cerebral Cortex, 27, 1011–1026.

McClamrock, R. (1991). Marr's three levels: A re-evaluation. Minds and Machines, 1, 185–196.

Miles, R., Tóth, K., Gulyás, A. I., Hájos, N., & Freund, T. F. (1996). Differences between somatic and dendritic inhibition in the hippocampus. Neuron, 16, 815–823.

Minlebaev, M. et al. (2011). Early gamma oscillations synchronize developing thalamus and cortex. Science, 334, 226–229.

Mitrukhina, O., Suchkov, D., Khazipov, R., & Minlebaev, M. (2015). Imprecise Whisker Map in the Neonatal Rat Barrel Cortex. Cerebral Cortex, 25, 3458–3467.

Mizuno, H., Ikezoe, K., Nakazawa, S., Sato, T., Kitamura, K., & Iwasato, T. (2018). Patchwork-Type Spontaneous Activity in Neonatal Barrel Cortex Layer 4 Transmitted via Thalamocortical Projections. Cell Reports, 22, 123–135.

Moran, P. A. P. (1958). Random processes in genetics. Mathematical Proceedings of the Cambridge Philosophical Society, 54, 60–71.

Morrison, A., Aertsen, A., & Diesmann, M. (2007). Spike-Timing-Dependent Plasticity in Balanced Random Networks. Neural Computation, 19, 1437–1467.

Morrison, J. E. (2007). The Astrolabe. (S. D. Myers Publishing Services).

Nakazawa, S., Mizuno, H., & Iwasato, T. (2018). Differential dynamics of cortical neuron dendritic trees revealed by long-term in vivo imaging in neonates. Nature communications, 9, 3106.

Neniskyte, U. & Gross, C. T. (2017). Errant gardeners: glial-cell-dependent synaptic pruning and neurodevelopmental disorders. Nature Reviews Neuroscience, 18, 658–670.

Nguyen-Vu, T. B., Zhao, G. Q., Lahiri, S., Kimpo, R. R., Lee, H., Ganguli, S., Shatz, C. J., & Raymond, J. L. (2017). A saturation hypothesis to explain both enhanced and impaired learning with enhanced plasticity. eLife, 6, e20147.

"Chromosome 21 - Genetics Home Reference". NIH (2019). <https://ghr.nlm.nih.gov/chromosome/21#resources>. Last Accessed: 28 April, 2019.

O'Connor, D., Wittenberg, G., & Wang, S.-H. (2005). Graded bidirectional synaptic plasticity is composed of switch-like unitary events. PNAS; Proceedings of the National Academy of Sciences, 102, 9679–9684.

Oh, W. C., Lutzu, S., Castillo, P. E., & Kwon, H.-B. (2016). De novo synaptogenesis induced by GABA in the developing mouse cortex. Science, 353, 1037–1040.

Olshausen, B. A. & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381, 607–609.

Ouardouz, M. & Sastry, B. (2005). Activity-mediated shift in reversal potential of GABA-ergic synaptic currents in immature neurons. Developmental Brain Research, 160, 78 – 84.

Owens, D. F., Boyce, L. H., Davis, M. B. E., & Kriegstein, A. R. (1996). Excitatory GABA Responses in Embryonic and Neonatal Cortical Slices Demonstrated by Gramicidin Perforated-Patch Recordings and Calcium Imaging. Journal of Neuroscience, 16, 6414–6423.

Patterson, D. (2009). Molecular genetic analysis of Down Syndrome. Human Genetics, 126, 195–214.

Pfister, J.-P. (2006). Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity. Journal of Neuroscience, 26, 9673–9682.

Phillips, R. (2015). Theory in Biology: Figure 1 or Figure 7? Trends in Cell Biology, 25, 723 – 729. Special Issue: Quantitative Cell Biology.

Pinon, M. C., Jethwa, A., Jacobs, E., Campagnoni, A., & Molnár, Z. (2009). Dynamic integration of subplate neurons into the cortical barrel field circuitry during postnatal development in the Golli-tau-eGFP (GTE) mouse. The Journal of Physiology, 587, 1903–1915.

Polleux, F. (2005). Genetic Mechanisms Specifying Cortical Connectivity: Let's MakeSome Projections Together. Neuron, 46, 395–400.

Powers, B. E., Santiago, N. A., & Strupp, B. J. (2018). Rapid forgetting of social

learning in the Ts65Dn mouse model of Down Syndrome: New evidence for hippocampal dysfunction. Behavioral Neuroscience, 132, 51–56.

Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., LaMantia, A., McNamara, J., & White, L. (2014). Neuroscience. (Sinauer, Sunderland, Mass), 4th edn.

Rebsam, A., Seif, I., & Gaspar, P. (2002). Refinement of thalamocortical arbors and emergence of barrel domains in the primary somatosensory cortex: a study of normal and monoamine oxidase a knock-out mice. Journal of Neuroscience, 22, 8541–8552.

Reeves, R. H., Gearhart, J. D., & Littlefield, J. (1986). Genetic basis for a mouse model of Down Syndrome. Brain Research Bulletin, 16, 803–814.

Reinholdt, L. G., Ding, Y., Gilbert, G. T., Czechanski, A., Solzak, J. P., Roper, R. J., Johnson, M. T., Donahue, L. R., Lutz, C., & Davisson, M. T. (2011). Molecular characterization of the translocation breakpoints in the Down Syndrome mouse model Ts65Dn. Mammalian Genome, 22, 685–691.

"Bringing a Virtual Brain to Life". Requarth, T. (2003). <https://www.nytimes.com/2013/03/19/science/bringing-a-virtual-brain-to-life.html>.

Riccomagno, M. M. & Kolodkin, A. L. (2015). Sculpting Neural Circuits by Axon and Dendrite Pruning. Annual Review of Cell and Developmental Biology, 31, 779–805.

Rice, F. L., Gomez, C., Barstow, C., Burnet, A., & Sands, P. (1985). AComparative analysis of the development of the primary somatosensory cortex: Interspecies similarities during barrel and laminar development. The Journal of Comparative Neurology, 236, 477–495.

Rice, F. L., Mance, A., & Munger, B. L. (1986). A comparative light microscopic analysis of the sensory innervation of the mystacial pad. I. Innervation of vibrissal follicle-sinus complexes. The Journal of Comparative Neurology, 252, 154–174.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65, 386–408.

Rubenstein, J. & Merzenich, M. M. (2003). Model of autism: increased ratio of exci-

tation/inhibition in key neural systems. Genes, Brain and Behavior, 2, 255–267.

Rubin, J., Lee, D., & Sompolinski, H. (2001). Equilibrium Properties of Temporally Asymmetric Hebbian Plasticity. Physical Review Letters, 86, 364–367.

Rueda, N., Florez, J., & Martínez-Cué, C. (2012). Mouse models of Down Syndrome as a tool to unravel the causes of mental disabilities. Neural Plasticity, 2012, 584071.

Sahara, S., Yanagawa, Y., O'Leary, D., & Stevens, C. (2012). The Fraction of Cortical GABAergic Neurons Is Constant from Near the Start of Cortical Neurogenesis to Adulthood. Journal of Neuroscience, 32, 4755–4761.

Sastry, B., Goh, J., & Auyeung, A. (1986). Associative induction of posttetanic and LTP in CA1 neurons of rat hippocampus. Science, 232, 988–990.

Shouval, H., Bear, M., & Cooper, L. (2002). A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. PNAS; Proceedings of the National Academy of Sciences, 99, 10831–10836.

Singer, Y., Teramoto, Y., Willmore, B. D., Schnupp, J. W., King, A. J., & Harper, N. S. (2018). Sensory cortex is optimized for prediction of future input. eLife, 7.

Sprekeler, H., Michaelis, C., & Wiskott, L. (2007). Slowness: An Objective for Spike-Timing–Dependent Plasticity? PLoS Computational Biology, 3, e112.

Stacey, J. A. (2018). The role of activity in remodelling area-specific neocortical circuits during development. Ph.D. thesis, University of Oxford.

Stevens, C. F., Tonegawa, S., & Wang, Y. (1994). The role of calcium–calmodulin kinase II in three forms of synaptic plasticity. Current Biology, 4, 687 – 693.

Stevens, K. A. (2012). The Vision of David Marr. Perception, 41, 1061–1072.

Sulis Sato, S., Artoni, P., Landi, S., Cozzolino, O., Parra, R., Pracucci, E., Trovato, F., Szczurkowska, J., Luin, S., Arosio, D., & et al. (2017). Simultaneous two-photon imaging of intracellular chloride concentration and pH in mouse pyramidal neurons in vivo. Proceedings of the National Academy of Sciences, 114, E8770–E8779.

Surace, S. C., Pfister, J.-P., Gerstner, W., & Brea, J. (2018). On the choice of metric in gradient-based theories of brain function. arXiv preprint, arXiv:1805.11851.

"Mouse Strain Datasheet - 001924". The Jackson Laboratory (2019). <https://www.jax.org/strain/001924>. Last Accessed: 28 April, 2019.

Thomas, M. S., Davis, R., Karmiloff-Smith, A., Knowland, V. C., & Charman, T. (2015). The over-pruning hypothesis of autism. Developmental Science, 19, 284–305.

Thompson, J. E. S. (1974). Maya Astronomy. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 276, 83–98.

Toyoizumi, T., Pfister, J.-P., Aihara, K., & Gerstner, W. (2005). Generalized Bienenstock-Cooper-Munro rule for spiking neurons that maximizes information transmission. Proceedings of the National Academy of Sciences, 102, 5239–5244.

Toyoizumi, T., Pfister, J.-P., Aihara, K., & Gerstner, W. (2007). Optimality Model of Unsupervised Spike-Timing-Dependent Plasticity: Synaptic Memory and Weight Distribution. Neural Computation, 19, 639–671.

Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron, 91, 260–292.

Tuncdemir, S. N., Wamsley, B., Stam, F. J., Osakada, F., Goulding, M., Callaway, E. M., Rudy, B., & Fishell, G. (2016). Early somatostatin interneuron connectivity mediates the maturation of deep layer cortical circuits. Neuron, 89, 521–535.

Valeeva, G., Tressard, T., Mukhtarov, M., Baude, A., & Khazipov, R. (2016). An Optogenetic Approach for Investigation of Excitatory and Inhibitory Network GABA Actions in Mice Expressing Channelrhodopsin-2 in GABAergic Neurons. The Journal of Neuroscience, 36, 5961–5973.

Van der Loos, H. & Woolsey, T. A. (1973). Somatosensory Cortex: Structural Alterations following Early Injury to Sense Organs. Science, 179, 395–398.

van Rheede, J. J., Richards, B. A., & Akerman, C. J. (2015). Sensory-Evoked Spiking Behavior Emerges via an Experience-Dependent Plasticity Mechanism. Neuron, 87, 1050–1062.

van Rossum, M., Bi, G., & Turrigiano, G. (2000). Stable Hebbian Learning from Spike Timing-Dependent Plasticity. J. Neurosci.; Journal of Neuroscience, 20,

8812–8821.

Vanderhaeghen, P. & Cheng, H.-J. (2010). Guidance Molecules in Axon Pruning and Cell Death. Cold Spring Harbor Perspectives in Biology, 2, a001859.

Vasilaki, E. & Giugliano, M. (2014). Emergence of connectivity motifs in networks of model neurons with short-and long-term plastic synapses. PloS one, 9, DOI: 10.1371/journal.pone.0084626.

Vickers, E. D., Clark, C., Osypenko, D., Fratzl, A., Kochubey, O., Bettler, B., & Schneggenburger, R. (2018). Parvalbumin-Interneuron Output Synapses Show Spike-Timing-Dependent Plasticity that Contributes to Auditory Map Remodeling. Neuron, 99, 720–735.

Vogels, T. P., Froemke, R. C., Doyon, N., Gilson, M., Haas, J. S., Maffei, A., Miller, P., Wierenga, C., Woodin, M., Zenke, F., & Sprekeler, H. (2013). Inhibitory synaptic plasticity: spike timing-dependence and putative network function. Frontiers in Neural Circuits, 7, 119.

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., & Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. Science, 334, 1569–1573.

Whittington, J. C. R. & Bogacz, R. (2017). An Approximation of the Error Backpropagation Algorithm in a Predictive Coding Network with Local Hebbian Synaptic Plasticity. Neural Computation, 29, 1229–1262.

Xie, X. & Seung, H. S. (2004). Learning in neural networks by reinforcement of irregular spiking. Physical Review E, 69.

Yang, J.-W. et al. (2009). Three patterns of oscillatory activity differentially synchronize developing neocortical networks in vivo. The Journal of Neuroscience, 29, 9011–9025.

Zenke, F., Agnes, E. J., & Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nature Communications, 6, 6922.

Zhou, L., Bar, I., Achouri, Y., Campbell, K., De Backer, O., Hebert, J., Jones, K.,

Kessaris, N., de Rouvroit, C., O'Leary, D., Richardson, W., Goffinet, A., & Tissir, F. (2008). Early Forebrain Wiring: Genetic Dissection Using Conditional Celsr3 Mutant Mice. Science, 320, 946–949.

Ziegler, L., Zenke, F., Kastner, D., & Gerstner, W. (2015). Synaptic Consolidation: From Synapses to Behavioral Modeling. Journal of Neuroscience, 35, 1319–1334.

Zis, P. & Strydom, A. (2018). Clinical aspects and biomarkers of Alzheimer's Disease in Down Syndrome. Free Radical Biology and Medicine, 114, 3–9.

Zucker, R. S., Dewey, K. R., Mulkey, R., & Tank, D. W. (1991). Presynaptic calcium in transmitter release and posttetanic potentiation. Annals of the New York Academy of Sciences, 635, 191–207.

Zucker, R. S. & Regehr, W. G. (2002). Short-Term Synaptic Plasticity. Annual Review of Physiology, 64, 355–405.