Action selection in the striatum: Implications for Huntington's disease

Adam Tomkins

Department of Computer Science

University of Sheffield

This dissertation is submitted for the degree of

Doctor of Philosophy

October 2015

I would like to dedicate this thesis to the neonatal neurosurgeons in the paediatric neurology department at Alder Hay, who gave me the chance to contribute to science, and to whom I owe everything.

Acknowledgements

I am extremely grateful to Dr. Eleni Vasilaki for giving me the opportunity to carry out my research, and for her valuable guidance and eternal support throughout. Without her patience, understanding and dedication, this thesis could not have happened.

I am also grateful to Professor Kevin Gurney for encouraging my transition into computational neuroscience, and for continued support in the field. I would like to give special thanks to Dr. Mark Humphries who has encouraged much of my research, and provided plentiful technical assistance.

I would like to further extend my gratitude to Dr. Mark Hepple, Prof. Georg Struth and Dr. Fabio Ciravegna, in the department of computer science, who have all eased my journey through my studies. I would also like to thank the EPSRC for funding my PhD.

Finally, I would like to thank my family and friends for their support, particularly the friends I have made during my time here in Sheffield.

Abstract

Although the basal ganglia have been widely studied and implicated in signal processing and action selection, little information is known about the active role that the striatal microcircuit plays in action selection in the basal ganglia-cortical-thalamic loops. To address this knowledge gap we use a large scale three dimensional spiking model of the striatum, combined with a rate coded model of the basal ganglia-cortical-thalamic loop, to asses the computational role the striatum plays in action selection. We identify robust transient phenomena generated by the striatal microcircuit, which temporarily enhances the difference between two competing cortical inputs. We show that this transient is sufficient to modulate decision making in the basal ganglia-thalamo-cortical circuit.

We also find that the transient selection originates from a novel adaptation effect in single striatal projection neurons, which is amenable to experimental testing. Finally, we compared transient selection with models implementing classical steady-state selection. We challenged both forms of model to account for recent reports of paradoxically enhanced response selection in Huntington's disease patients. We found that steady-state selection was uniformly impaired under all simulated Huntington's conditions, but transient selection was enhanced given a sufficient Huntington's-like increase in NMDA receptor sensitivity. I propose a mechanistic underpinning to a novel neural compensatory mechanism, responsible for improved cognition in severe neuro-degeneration. Thus, our models provide an intriguing hypothesis for the mechanisms underlying the paradoxical cognitive improvements in manifest Huntington's patients, which is consistent with recent behavioural data.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

In this thesis we will discuss the problem of action selection. We cover the mechanisms in the mammalian brain which have evolved to solve the problem, and discuss what we can learn about such mechanisms by developing computational models. This research is build upon models of biologically-relevant spiking neural networks, used to explore how the basal ganglia (BG) achieve action selection in vertebrates, with the result of discovering novel signal selection mechanisms in the striatal microcircuit. This thesis further explores what happens to these novel selection processes as the neural mechanisms for action selection are compromised. I use a phenomenological model of the progressive neuro-degenerative condition, Huntington's disease (HD), to examine the interaction between neural dynamics and behavioural selection. I build on theoretical models of action selection in the basal ganglia, and challenge improved models to match experimental data. We find that the novel signal selection mechanisms that are revealed, are capable of giving a mechanistic explanation for recent, paradoxical behavioural experiments with Huntington's Disease patients.

I first presented my research at the International Conference on Artificial Neural Networks in 2012 with the paper 'How degrading networks can increase cognitive functions', published in the Lecture Notes of Computer Science (Tomkins et al., 2012) (Appendix A), as a preliminary study to the work continued in this thesis. I followed up the results of this study with a larger, more detailed role of signal selection in the basal ganglia, in the 2013 paper, 'Transient and steady-state selection in the striatal microcircuit', published in the journal of Frontiers in Computational Neuroscience (Tomkins et al., 2013)(Appendix B). Other abstracts concerning our research have appeared in the Journal of Neurology, Neurosurgery and Psychiatry (Beste et al., 2012) and Frontiers in Computational Neuroscience (Tomkins et al., 2011).

In this thesis I will be discussing and expanding on the ideas presented in these papers. This thesis will give a detailed account of the role which the basal ganglia play in signal selection. Through the lens of Huntington's disease, I discuss how computational models of signal selection in the basal ganglia can be linked to behavioural studies in selection experiments.

In the introduction, we will discuss the fundamental problem of action selection, and how computational neuroscience is well suited to provide insight into behavioural and neurological phenomena. I demonstrate how computational models can provide insights into biological neural processes, by examining signal selection performed in abstract circuits, that have been suggested to be theoretically supported by the basal ganglia. I examine and discuss the tools used, and the general progress made by this relatively young discipline. The chapter ends with a review of how we can apply the tools of computational neuroscience to re-examine paradoxical results of traditional behavioural neuro-scientific experiments.

In chapter 2 I dissect the field of computational neuroscience and establish the principal research areas relevant to this thesis. I detail the evolution of reduced neuron models, used to build complex and informative models of brain anatomy, on which we investigate neuroscientific principles. We establish the action selection field, and discuss previous attempts to study the problem of selection, and review competing hypotheses in the decision making domain.Chapter 4 details the computational abstractions detailed in this thesis, to clarify how they are used, and to what end. The chapter is designed to recap the tools and expand on their roles in the research, and to clarify the progression of results to be presented in later paragraphs.

Chapter 3 details how reduced neuron models, the basic building blocks of computational neuroscience, have been forged together to create large scale models of the striatal network. I cover modelling the specific neuron types, their connection dynamics, and the synaptic transmission seen in the most advanced striatum models. Additionally, I discuss how a complex genetic disease such as Huntington's can be modelled phenomenologically in the existing striatum models. Chapter 5 covers the high level concepts of action selection through signal selection in the basal ganglia, and details the various protocols designed to assess signal selection in the striatum. Here I first show the distinct selection mechanisms exhibited by the striatal microcircuit in response to competitive input signals. I show the existence of distinct selection mechanisms, which occur during different plausible connectivity schemes in models of the striatal microcircuit.

In chapter 6 I focus on one mode of signal selection in the striatum, termed *'Transient Selection'* (TS), which I first introduced in the 2012 Conference Paper (Tomkins et al.,

2012)(Appendix A). I consider which connectivity regimes transient selection emerges from, detail its root causes, and continue to investigate the effect transient selection has on action selection in decision tasks. I examine how *transient selection* in the striatum is amplified by the Basal ganglia-thalamo-cortical loops, causing a robust impact on action selection performance. I finish by detailing the experimental predictions made by these computational results. In chapter 7 I consider *'Steady-State Selection,'* (SS) a distinct mode of signal selection supported by our striatal model, first discussed in the 2013 Frontiers of Computational Neuroscience journal (Tomkins et al., 2013)(Appendix B). We cover the necessary striatal conditions needed to elicit steady-state selection, as well as the effect that the phenomena exert on signal selection in the striatum. I follow up with a discussion of the results, with experimental applications.

In chapter 8 I discuss how these two distinct selection mechanisms manifest in diseased states of the brain. I show that phenomenological models of cell atrophy and NMDA receptor over expression, caused during Huntington's Disease, manifest with interesting experimental predictions. We show how elicited transient behaviour could be used to justify the paradoxical increase in performance during recent selection experiments Beste et al. (2008). This novel insight provides a mechanism of network-level compensation, which could both support new reports of improved selection in auditory decision making tasks, while preserving a theoretical framework, capable of explaining the reduced capabilities usually associated with Huntington's disease. Finally, I conclude our discussion with the major results of the thesis, and consider how computational neuroscience can continue to provide theoretical insights into the practical problems faced by neuroscience today. I follow up with a discussion on further work in the field which deserves attention.

1.1 Action selection is a fundamental problem for all intelligent agents

Action selection is a fundamental problem faced by any intelligent agent, from rats, through to robots, and of course, humans. To survive in a dynamic environment, it is necessary to be able to reliably and robustly decide how to act, be that consciously or otherwise. In this research I have explored the problem of Action Selection from multiple viewpoints. I examine the computationally abstract abilities of Neural Networks to act as generalised decision circuitry. I next investigate the biological basis for signal selection using accurately connected

simulated tissue models of the striatal microcircuit, to investigate how the connectivity enforces competitive dynamics.

Action selection can be seen as the most critical and basic task required to be tackled by any autonomous, intelligent agent. At its most basic, it can be described as deciding what to do next, given a set of possible actions by incorporating partial knowledge of the environment. Action selection has been the focus of much research, through various different fields, with a long history of research in psychology and ethology. More recently, developments are emerging from the fields of artificial intelligence, neuro-economics and computational neuroscience. Each field brings its own set of tools and assumptions, forging complementary in roads into our understanding of decision making. In this thesis I will discuss and develop the role of computational neuroscience as a key research field for the action selection problem.

There are a variety of open research questions concerning action selection, which often differentiate the various research fields. Problematically action selection is a nebulous concept that can be tackled at multiple levels of description. Currently, even the definition of what constitutes as a discrete 'Action' is contentious. One perspective considers that intelligent agents consist of a hierarchy of decision complexity, each controlled competing subsystems. In this approach each system can have its own notion of distinct 'actions'. One outcome of this approach is the problem of action-value comparison. When deciding which action to take, a system must have some mechanism, or common currency by which to determine the most beneficial action. It is not yet clear how the actions described in traditional ethology, could be weighed against the actions described by economics, or artificial intelligence.

With such fundamental disagreements, it is no surprise that there has been tremendous backing for domain-specific research into the complexities of intelligent decision making. In contrast to ethology, economics, and artificial intelligence, neuroscience has given us a window into functioning autonomous agents. The mammalian brain is a general decision making system, expressing each conceivable action in a common, comparable language, that of neural spikes and spike timings. Decades of research have mapped specific structures, hierarchies and pathways geared towards fast and efficient action selection in the mammalian brain.

Decoding neural transmission enables the study of both biological and theoretical mechanisms of generalised action selection. The currency of spikes gives an elegant framework for the open problems of action comparison, and action representation. In this thesis I will show how we can gain insights into biological action selection mechanisms using computational models of spiking neural networks. I discuss how computational neuroscience can provide

insight into behavioural and neurological studies, by examining the affects of phenomenological models of Huntington's disease on action selection, and discuss what this can teach us about natural action selection frameworks.

These are the features that makes neuroscience such an advantageous lens through which to study the problem of action selection.

1.2 Action selection is a core problem in neuroscience

As neuroscience provides such an advantageous and budding vantage to study action selection, it is no surprise that there has been steady progress into unravelling complex structures and processes in the brain, which contribute to the phenomenon of action selection.

Within the bounds of neuroscience, we must first set out to define some neuro-physiological characteristics of action selection, to act as a guide to the discovery and mapping of potential structures in the brain, on to domain of action selections. It has been suggested (Redgrave et al., 1999) that in order to facilitate effective action selection, the mechanism should:

- Only allow the expression of a single action out of a set of mutually exclusive actions.
- Have an accompanying mechanism to determine the value or *salience* of each action.
- Express the action with the highest salience, given the environmental state and some possible internal states.
- Provide 'clean switching' to express an action in a timely and decisive manor.
- Allow each expressed action full control of the available motor resources, and prevent out-competed actions from interference.

With these properties in mind, neuroscientists have progressed to find candidate mechanisms for action selection.

1.3 The approach of computational neuroscience

While neuroscience has an ever expanding range of specialist technical hardware to utilise in the battle for understanding the brain, the work of the computational neuroscientist is completed with mere mathematics and brute computation. As a burgeoning discipline, we seek to model complex biological processes using contributions from the branches of dynamical systems and information theory, often combined with advances parallel computing.

Computational neuroscience consists of the endeavours to further understand the information content and processing carried out by the nervous system, across multiple levels of abstraction and over a continuum of structural scales, including the sub-neuronal circuitry, up to modular brain-spanning networks. Computational neuroscience has come into its own as a field with a huge potential in both hypothesis generation and testing for traditional neuroscientific advances, as well as pushing the bounds of current knowledge in its own right.

In this thesis I will discuss and utilise the following three main facets of computational neuroscience:

- The derivation mathematical representations of neural interaction, as first described in Hodgkin and Huxley (1952), and advanced by Izhikevich (2007a).
- The development and identification of network topologies which can exhibit the required properties of action selection.
- The testing of macroscopic ramifications of neural computations performed by anatomical models of spiking neural networks.

Using these three main approaches I will investigate the role of the striatum, and the basal ganglia in the domain of action selection.

The complex questions asked by computational neuroscience, are tackled using a mixture of three general model cases, the descriptive, the mechanistic and the interpretive model forms.

Descriptive models are succinct descriptions of large amounts of experimental data, accurately summarising how neurons and neural circuitry behave. Being based, often loosely, on anatomical data, these models aim to capture the behaviour of a system (neuron or circuit), but not necessarily explain the behaviour. In chapter 3 we see an example of descriptive modelling, where I will be describing the neuron junctions and developing a novel model of saturating synapses for point model neurons.

In contrast, mechanistic models focus on how aspects of the nervous system operate. They consider how computational roles are performed, based on anatomical and physiological observations, such as connectivity. In chapter 4 I use mechanistic models to decode the functions performed by the basal ganglia microcircuit, and the larger basal ganglia-thalamocortical loops.

Finally, interpretive models use computational constructs to explore the cognitive and behavioural significance of features in the nervous system, and discuss why nervous systems

operate in the way they do. In chapters 5 through 8 I introduce interpretive models of striatum function in relation to action selection and interpret the roles of striatal microcircuit phenomena, with respect to elicited behaviour. We further use the results of a phenomenological model of Huntington's disease related cell-atrophy, to understand the behavioural effects of cognitive decline in symptomatic HD patients.

In this thesis I will utilize each of the three model types to address different aspects of the same problem, creating a more complete picture of action selection mechanisms in the basal ganglia.

1.3.1 The core assumptions of computational neuroscience and the scope of thesis

As a field, Computational Neuroscience seeks to gain a mechanistic understanding of the cognitive processes which govern our everyday experiences. In extension to this, it could be said that the ambition of such an undertaking is to explain how cognition could occur from the behaviour of the neurons, networks and connections therein. The motivation behind computational neuroscience is to understand the nervous system in sufficient detail that, by studying individual neurons, artificially simulate accurate reproductions of complex neural processes, to gain a further understanding of neural form and function. By reducing the nervous system to its constituent parts, i.e. neurons and synapses, I inherently represent an ontological reductionist view, a reasonably consensual position in the scientific community.

On a more grounded and practical sense, as computational neuroscientists, we must make decisions as to the relative importance of various facts and observations about the nervous system we study. We have at our disposal an enormous array of detailed physiological and anatomical observations, down to precise morphologies of individual neurons, and the complex biochemistry of synaptic interaction. Conversely, we have limited, though growing computational resources at our disposal. Because of this trade off we must decide as to the detail we need concern and constrain ourselves with, to garner meaningful results with limited resources.

Neuroscience is often reduced to the study of neurons, synapses and action potentials. This neuro-centric view could at first glance seem unfounded, as it has been estimated that neurons make up only 25% of the cells in the basal ganglia, with the remaining 75% being made up of the glial cells (of which with 67% are oligodendrocytes, 22% are astrocytes, and 10% microglia (Karlsen and Pakkenberg, 2011)). However, although glial cells have recently been shown to play a role in some physiological processes like breathing, the traditional

studied neuron types have been shown to play a disproportionally larger role in information processing, transmission and control. While glial cells contribute importantly to the central nervous system, they are not considered in the majority of information-processing centred research, and are as such absent from this thesis too, due to the lack of a computational role.

So too is there contention as to the relevant aspects of individual neurons. We have at our disposal, very complex models of neuron morphology, synaptic interactions and three dimensional connectivity trees, and yet remains constrained by computational means, to consider only the most relevant aspects of neurobiology, as to gather tractable, relevant and timely results. In this thesis I will be using a combination of reduced (2 variable) model neurons capable of reproducing a wide array of neural firing patterns. We combine these with detailed models of a range of synaptic interactions, and accurate connectivity statistics to create our large models of the striatal microcircuit. This gives us the capabilities to investigate accurate and detailed neuron firing behaviour without the need to resort to computationally expensive multi-compartmental neuron models, such as the Hodgkin and Huxley (HH) models (Hodgkin and Huxley, 1952).

We are building on the implicit assumption that trains of action potentials are the units of information-processing that matter in the brain. Spike are elicited from the excitable cell bodies, and propagated through axons, throughout the neural networks of the CNS. There is a tremendous amount of evidence to suggest that the neural spike train is the most important information-processing aspect of the nervous systems, although as with all complex systems there are other less abundant mechanisms at play, such as analogue transmission of membrane potential which can also carry information across a specific set of synapses.

1.3.2 The progress of computational neuroscience in action selection

The basal ganglia have long been considered as an important region of the brain, and has been implicated in many important tasks, such as motor control and reinforcement learning. Crucially for this thesis, the basal ganglia have been shown to be a crucial mechanism in action selection (Prescott et al., 1999), and as such have received a lot of attention from both neuroscientists and computer scientists alike.

Early investigations sought to discover the complex connectivity seen between the various sub-nuclei contained within the basal ganglia. The basal ganglia contain a complex network of at least 6 sub nuclei, with internal interactions having both excitatory and inhibitory connections. Further, to this, as an important nucleus in the mid brain, the basal ganglia have their place within the larger brain networks, with links to various areas including the cortex,

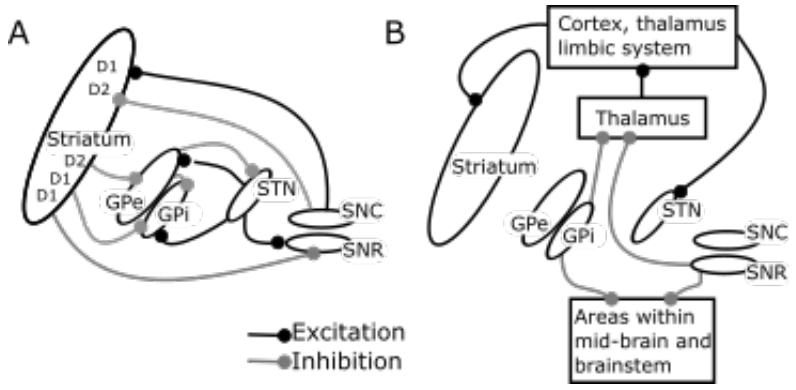

Fig. 1.1 **The anatomical connectivity of the basal ganglia. (A)**. The internal anatomical connectivity of the distinct nuclei which comprise the Basal Ganglia. D1 and D2 denote two separate classes of dopamine receptors. **(B)** The external anatomical connectivity of the Basal Ganglia nuclei, showing their relation to important external structures such as the thalamus, the cortex and areas of the brain stem. Excitatory connections are shown in black, and inhibitory shown in grey. Abbreviations: GPe-Globus Palidus external; GPi-Globus Paladus internal; STN-Subthalamus Nucleus; SNC-Substantia Nigra pars compacta; SNR-Substantia Nigra pars Reticular Adapted from Gurney et al. (2001)

the thalamus and the brain stem. Uncovering these internal and external loops has lead to the development of ever-improving anatomical maps of the basal ganglia, such as the map shown in figure 1.1, reproduced from Gurney et al. (2001).

With detailed mappings of anatomical interconnectivity to build on, computational neuroscientists have been able to abstract away biological detail, and investigate the range of functional roles which could be supported by the underlying structure. These high level computational analyses have provided many models of the functional roles of each sub-nucleus in the basal ganglia. Further, details about the connection types have been uncovered, with the introduction of diffuse and direct connections, allowing for more complex On-surround and off-centre networks for selective dis-inhibition. With such a complex internal-network, several approaches to the basal ganglia function have been formulated. A commonly accepted interpretation of the network is to divide the circuitry into two opposing pathways through the Basal Ganglia. Figure 1.2 shows the complementary 'Direct' and 'Indirect' pathways, which are hypothesised to play role in fast action switching through activation and dis-inhibition.

Using the proposed models of basal ganglia connectivity there have been copious simulations the network of nuclei and its varied connection types, in an effort to discover just how the behaviour of the system will behave to incoming stimuli. Models of this ilk have

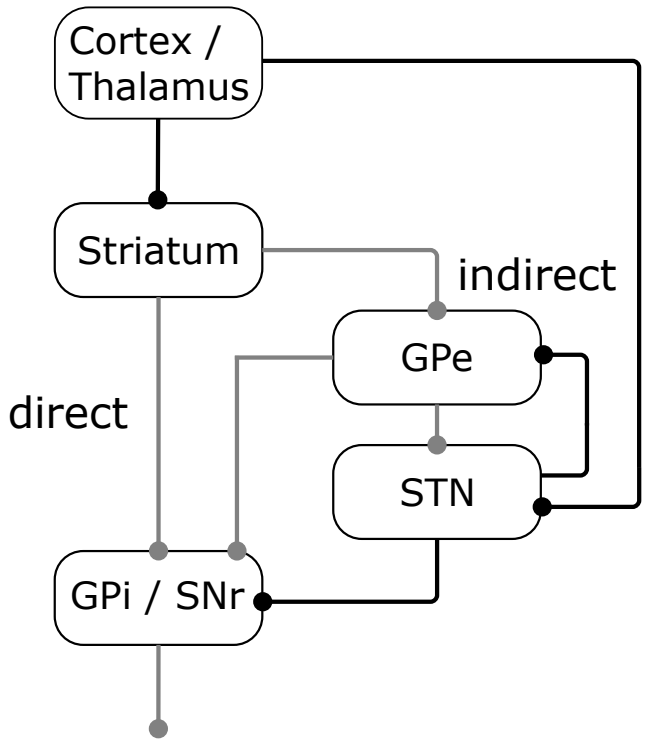

Fig. 1.2 **A proposed functional anatomy of the basal ganglia.** A simplified view of the functional anatomy of the Basal ganglia. Two distinct paths can be seen, as the direct and indirect pathway from the Striatum through to the substantia nigra pars reticulata (SNr) and globus pallidus external (GPe). Competing pathways are theorised to act as counter balancing to control selective disinhibition. Excitatory and inhibitory pathways are denoted by black and grey lines respectively. Adapted from Gurney et al. (2001).

generally simplified each nucleus to a discrete number of channels, each simulated using a simple leaky integrator differential equation to represent the mean-field activation level of each nucleus. These studies have been instrumental in exploring the role of complex connectivity in response to evolving neural inputs.

In contrast to developing high level conceptual models of the functional connectivity of the basal ganglia, neuroscientists have also delved deep into the micro-structure of elements of the basal ganglia and discovered some detailed and complex micro-circuitry.

For example, the excitatory spiny projection neurons (SPN), and inhibitory fast spiking inter-neurons (FSI), play two very different roles. SPNs make up $\geq 90\%$ of striatal neurons, and are medium sized at $\sim$ 15microns, with large dendritic trees with a diameter of about $\sim$ 500 microns (Humphries et al., 2009a). They receive cortical input, and project inhibitory connections outside the striatum. SPNs are quiet unless stimulated by cortical inputs. In contrast, FSIs are twice the size of SPNs, and are tonically active, firing spontaneously (Szydlowski et al., 2013). FSIs connect to SPNs within the striatum, and serve to modulate SPN excitability, by inhibitory connections close to the SPN soma (Humphries et al., 2009b). These two neuron types create a fine control mechanism for activity in the striatum, and have been shown to be differentially targeted in neuro-degenerative diseases (Reiner et al., 1988).

Although useful in elucidating the computational role of the entire basal ganglia, high-level models necessarily miss the intricate details of local signal processing done by such complex neural circuits.

Regardless of such an interesting microcircuit, and being the largest nucleus in the Basal Ganglia, the computational role striatum has generally been relegated to a mere input vector with little internal processing in and of itself. In this thesis I utilise a detailed spatial model of striatum tissue, constructed with various cell types, an array of connection morphologies, combined with a statistical approach to cell spatial distribution derived anatomical observation. This extra level of detail lets us investigate the complex processing done in the striatum, and discover interesting new forms of signal selection in the striatum, which I show, can improve signal selection in the Basal ganglia as a whole. In chapters 5-6 I see how these models can be embedded into a further model of cortical-thalamus loops to discover how processing in the basal ganglia is amplified throughout the wider-brain networks.

1.4 The results of our study.

Although there is a lot of speculation and investigation on the role of the basal ganglia in action selection, there is very little known about the role of the complex microcircuit of the

striatum. As the striatum is the largest nucleus in the basal ganglia, we would expect to find a computational purpose to its size and complexity. In this thesis I set out to discover more about how the striatum facilitates selection, with interesting and perhaps surprising results.

In this thesis I detail two active forms of signal selection which arise from a combination of the implicit structure of the striatum, and the neuron types of which it is composed. Along the way I develop a new model for synaptic saturation, improving the biological-realism of the previous spiny projection neuron models.

In chapter 6, I detail the phenomena I discovered, termed 'transient selection', in which competing signals in overlapping, distributed channels in the striatum, experience a temporary modification to the encoded salience of each competing signal. The signal with the largest salience undergoes a positive boost of activation, while simultaneously the least salient, and thus out competed signal, experiences a reduction in firing rate. We show that this transient selection emerges from both the circuit connectivity, and the intrinsic membrane properties of the spiny projection neuron. We go on to show that the transient selection is sufficient to alter decision making performance in the wider decision circuits of the basal ganglia-thalamo-cortical loops. This demonstrates a concrete computational application of complex striatal micro-circuitry, in the wider action selection mechanisms.

In chapter 7, I detail a second form of striatal selectivity termed 'Steady State selection', in which competing signals experience a continued alteration to activation levels. Most noticeably the out-competed signal is suppressed in proportion to the strength of the most salient signal, possibly allowing for the quick switching required by our definition of action selection, as stated earlier. We continue to show that the steady-state selection is a feature of striatum networks which have a high connection density.

These findings shed light onto the mysterious workings of the striatal microcircuit, and give a distinct computational role for differentially connected tissue densities.

1.4.1 Applying theory to practice: A Huntington's disease case study

Apart from being a locus of action selection in the vertebrate brain, the striatum also plays a central role in several neuro-degenerative diseases, such as Parkinson's and Huntington's disease.

Huntington's disease is a well studied hereditary neurological disorder which progressively leads to cognitive decline and behavioural abnormalities, such as involuntary writhing movements called chorea. Neurologically, Huntington's disease is characterised by a pro-

gressive loss of GABAergic spiny projection neurons in the neostriatum, with glial cells remaining intact (Reiner et al., 1988).

This co-location gives us the opportunity to further investigate the role of the striatum, by examining the behavioural and cognitive effects on patients with such diseases (Humphries et al., 2010). In this thesis I aim to unravel some functions kept in control by a healthy striatum, by comparison with the degraded striatum state seen in HD patients.

Recently, the behavioural-based paper by Beste et al. (2007), discovered that in contrast to a general decline in cognitive and behavioural performance in symptomatic Huntington's Disease patients, there are some specific tasks in which symptomatic patients show improvements. Behavioural studies have shown improved performance (increased reaction times and reduced error rates) in symptomatic HD patents over both healthy controls and pre-symptomatic HD patients. This paradoxical result of improved performance, coinciding with diminished neurological structure, posits an interesting question on the processes of selection that are carried out in the diminished striatum of HD patients (Craufurd and Snowden, 2002).

As the striatum is widely affected by cell atrophy during neuro-degenerative diseases such as Huntington's disease, we can use our detailed model of the striatum, and a phenomenological model of HD related cell atrophy to investigate the effect of degradation on selection in the striatum. We apply theoretical measurements of selection in the model basal ganglia, to a diminished striatum model, inspired by HD degradation, to compare the modifications to selectivity caused by cell atrophy and increased NMDA receptor sensitivity associated with the HD affected striatum Feigin et al. (1995).

In chapter 8 of this thesis, I compare the role of transient and steady-state selection. We challenged both forms of model to account for the paradoxical response selection results of (Beste et al., 2008). They found that manifest Huntington's disease patients were both faster and less error prone than controls in a simple two-choice reaction-time task. As Huntington's disease primarily results in striatal damage (Miller et al., 2012), this suggests the hypothesis that changes in the striatum directly affect response selection. We expand on the role of the striatum in signal selection, by describing a framework for signal selection that can account for both the typical decline in performance for most tasks under Huntington's conditions (Ho et al., 2003), and a mechanism for increased performance under the same conditions.

Our parametric study results suggest that an increase in NMDA receptor expression, combined with progressive cell atrophy, can improve the competitive signal dynamics. Thus, the cognitive performance of the manifest HD patients might be linked to enhanced *transient selectivity* in the striatum, while normal cognitive decline can be linked to the uniform decline in selectivity seen in steady-state selectivity.

Chapter 2

Literature Review

This thesis looks at the problem of action selection at multiple levels of abstraction, introducing novel approaches to modelling synaptic transmission and phenomenological models of disease. As such this section will draw on a wide range of topics from which the main thesis is built, starting from the basic elements of computational neuroscience, up to the neurological basis of Huntington's disease pathology.

I start by describing the perspective that computational neuroscience applies to the central nervous system. I detail the levels of abstraction that computational models can address, and identify that this thesis will span several layers of Marr's framework (Marr and Poggio, 1976). I continue to introduce the excitable cell, and explore the biophysical properties of the neuron. Building on an understanding of basic neuron function, I examine the evolution of neuron models, and show the basic derivation of a cell as an electrical circuit. Expanding the function of simple cell models, I detail the role of adaptation in neurons and describe how synaptic transmission is included in cell models. Moving up the spatial scales, I detail how populations of the derived model neurons are used in network models, and discuss their predictive power. Finally I build on the role of computational models by discussing the neurological effects of Huntington's diseases, and layout the aims of this thesis to bridge abstract neuron models and behavioural experiments.

2.1 The brain and its constituents

The neuroscientist Gerald Edelman described the human brain as the most physically complex object in the known universe, and as such, often the first problem with studying the brain, is deciding where to start. David Marr proposed a structural hierarchy of spatial resolution

when considering function and computation in the brain (Marr and Poggio, 1976). The hierarchy consists of the following loosely defined spatial scales:

- The CNS (whole brain)
- Brain Modules (Cerebellum, basal ganglia, hippocampus etc)
- Nuclei/regions within modules
- Small neuronal circuits
- Neurons
- Synapse and membranes
- Intracellular signalling

Marr's analysis suggested that, as we seek to describe and dissect the computational role of the brain, we must consider not only the computation we are studying, but also the algorithm and implementation used to undergo the computation. We must consider a generic neural mechanism to support the proposed algorithm, as well as the 'hardware' implementation of the suggested neural mechanism, such as small neuronal circuits. This framework supports both top down 'prescriptive' investigation from computational necessities and bottom up 'mechanism mining' approaches derived from biological observation. For function mapping in the CNS we should consider:

- Computation - i.e. where to look next
- Algorithm - i.e. find location of maximal salience
- Neural mechanism - i.e. competitive neural layer
- Implementation - i.e. circuits in striatum in basal ganglia

Computation can be studied at each level of the spatial hierarchies. Abstract computation occurs at the whole brain level, between interconnected modules over the scale of centimetres, and simultaneously within the complex cellular signalling pathways only micrometers apart. Due to these vastly different spatial scales, we require an array of tools and different computational models to study the nebulous phenomena of biological computation. Generally more complex mathematical neuron models are required at decreasing spatial scales, and larger, more general network-models as we increase spatial scales.

In this thesis we take a multi-level approach to studying the computational mechanisms in the brain, introducing novel models of synaptic input. I use reduced neuron models to study computation arising from small neuronal circuits in the striatum. I focus on the striatum, as the substrate that drives signal selection through the Basal ganglia, an important subset of nuclei which are crucial to action selection in vertebrates (Redgrave et al., 1999; Stephenson-jones et al., 2011).

2.2 Theory of spiking of neuron models

Neurons are a highly specialised class of cells, which have evolved specific morphologies and features to facilitate the computation and transmission of information. Figure 2.1 shows the general structure of a typical neuron.

The cell body, or soma, of a neuron has developed an intricate network of dendrites. Dendrites receive and respond to the chemical signalling at synapses with the axons of other neurons by converting chemical triggers at synapses on the dendrites, to electrical signals. A dendrite will act as conductor to propagate electrochemical stimulation from distant inputs toward the cell body, or soma, for further processing. Dendrites may also play a critical role in integrating synaptic inputs, and performing pre-somatic computation (London and Häusser, 2005), actively modifying somatic action potential generation.

The soma integrates synaptic input from the entire dendrite tree. With great enough stimuli, the soma will produce an 'action potential', an electrochemical spike in potential across the neuron membrane. These action potentials are actively maintained and conducted down the cell axon. As seen in figure 2.1, the axon is a long projection from the soma which allows for the propagation of action potentials over long distances, to connect with other neuron dendrites using specialised protrusions called synapses. A single axon can form thousands of synapses and innervate multiple parts of the CNS. Schwann cells help insulate the axon, aiding the long range information transmission capabilities of the neuron.

Synapses are formed between co-localised neurons, from the axons of one neuron, onto the dendrites of another. Synapses can be present along the axon structure, or at axon terminals. These structures allow for the transmission of electrical and electrochemical signals between disparate cells. Pre-synaptic action potentials cause a release of neurotransmitters into the extracellular space between cells called the synaptic cleft. Specific post-synaptic channels are triggered by the presence of neurotransmitters, causing a cascade which results in a raised membrane potential in the post-synaptic dendrites. By this method information transmission can be quickly achieved between neurons unidirectionally across the synapse.

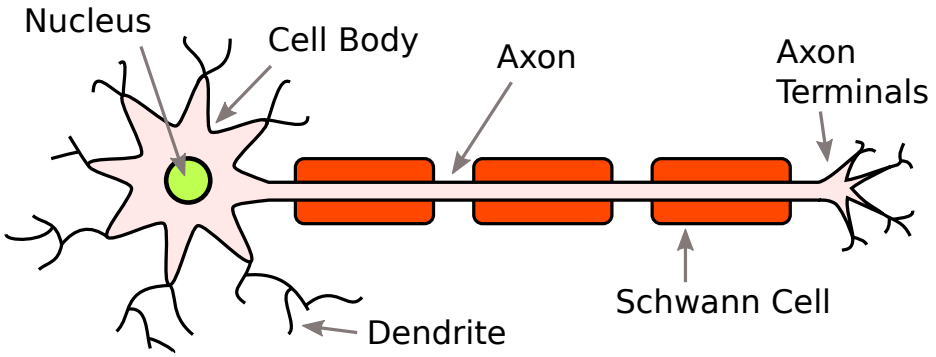

Fig. 2.1 **General structure of a neuron.** Information flows from left to right. Synaptic transmission is received at the dendrites. The soma processes synaptic input, releasing an action potential down the axon if input is sufficient. Schwann cells insulate the axon to improve transmission speeds.

A gap junction allows for the transmission of bidirectional information, creating a direct electrical membrane connection, allowing for passive two way information transmission.

The complex electrochemical signal processing achieved by neurons is a direct result of highly specialised cellular function, and can be measured and modelled in terms of basic electrical theory (Hodgkin and Huxley, 1952).

2.2.1 The electrical properties of neurons

Akin to other cell types, the cell body contains a mixture of positive and negatively charged ions, as well as molecules such as amino acids, proteins and higher cell structures. At rest the cell body maintains an excess concentration of negative charge, with the negatively charged ions clustering on the inside of the cell membrane. Due to electrostatic forces, an equal amount of positive charge from the extracellular membrane accumulates on the outside of the cell membrane. This differential is termed the membrane potential.

By convention, we define the potential of the extracellular fluid to be zero, and as such, in a neuron at rest, the build up of negatively charged ions on the inner membrane surface causes the cell to have a small negative charge, V_m. At rest the membrane potential is in a electrochemical equilibrium, such that the net flow of positive and negative ions through ion channels in the cell membrane is electrically equivalent. This equilibrium can be actively maintained by ion-pumps, and passively maintained by the opening and closing of ion channels.

A potential difference is maintained between the inner and outer membrane surface. The membrane is considered electrically active, with a resting potential, E_R. Ion channels in the lipid layer selectively allow a flow of charge across the membrane, acting as a resistive force R_m. The membrane capacitance C_m describes the build up of charge across the membrane. These elements can be combined to describe a basic electrical circuit supported by local the cell body membrane.

From these measurements we can derive a model cell formula. Using Ohm's law to describe the current I_r flowing over the resistive component of our circuit, we can describe the voltage difference over the ion channel as the difference between the resting potential E_R and the membrane potential V_m, as:

$$I_r = \frac{E_R - V_m}{R_m} \tag{2.1}$$

We can further define the capacitative current I_C by differentiating capacitance $Q = C_m V_m$, to derive:

$$I_C = C_m \frac{dV_m}{dt} \tag{2.2}$$

As current is conserved, we have:

$$C_m \frac{dV_m}{dt} = \frac{E_R - V_m}{R_m} \tag{2.3}$$

By setting $E_R = 0$, we can simplify our model to:

$$C_m \frac{dV_m}{dt} = \frac{-V_m}{R_m} \tag{2.4}$$

and derive the analytical solution as:

$$V_m(t) = V_m(t_0) e^{-(t-t_0)/C_m R_m} \tag{2.5}$$

given the initial $V_m(t_0)$. We can see that the rate of change of the model is determined by the value $C_m R_m$. When t = $C_m R_m$, $V_m(t) = \frac{V_m(t_0)}{e}$. $C_m R_m$ fixes the characteristic time constant τ_m of the membrane. We define $\tau_m = C_m R_m$. This is a useful equivocation, as physiologically it is easier to measure the time constant of a membrane, than the capacitance and resistance of its components. Using this, the analytical solution of the model can be written:

$$V_m(t) = V_m(t_0) e^{-(t-t_0)/\tau_m} \tag{2.6}$$

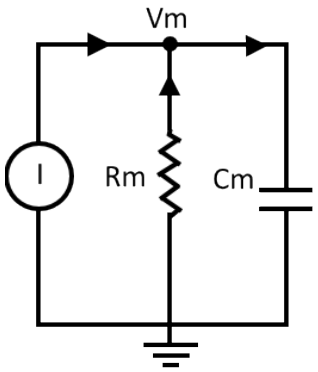

Fig. 2.2 **Membrane dynamics expressed as an electrical circuit.** A simple circuit can reproduce membrane potential dynamics. C_m is the membrane capacitance, R_m is the membrane resistance of ions flowing through ligand gated channels. V_m is the evoked membrane potential, and I_{inj} is the current injection. The extra-cellular potential is designated as $0mv$, represented by earth.

We can expand this formalism to become a leaky integrator, given an injection current, I_{inj}, so that we have:

$$C_m \frac{dV_m}{dt} = \frac{-V_m}{R_m} + I_{inj} \tag{2.7}$$

which can be expressed using the time constant τ as:

$$\tau_m \frac{dV_m}{dt} = -V_m + R_m I_{inj} \tag{2.8}$$

with the solution when I_{inj} is constant as:

$$V_m(t) = R_m I_{inj}(1 - e^{-t-t_0}/\tau_m) + V_m(t_0)e^{-(t-t_0)/\tau_m} \tag{2.9}$$

The leaky integrator demonstrates the basic principles of membrane potential dynamics in a in vitro cell membrane with current injection. In reality the complex behaviour of a cell depends on the biophysical characteristics and electro-chemical machinery of the cell. While this model is instructive, it lacks any mechanism for action potential generation. To derive a biologically accurate model of spike generation, we need to include a formalism for the complex interplay of various membrane ion channels.

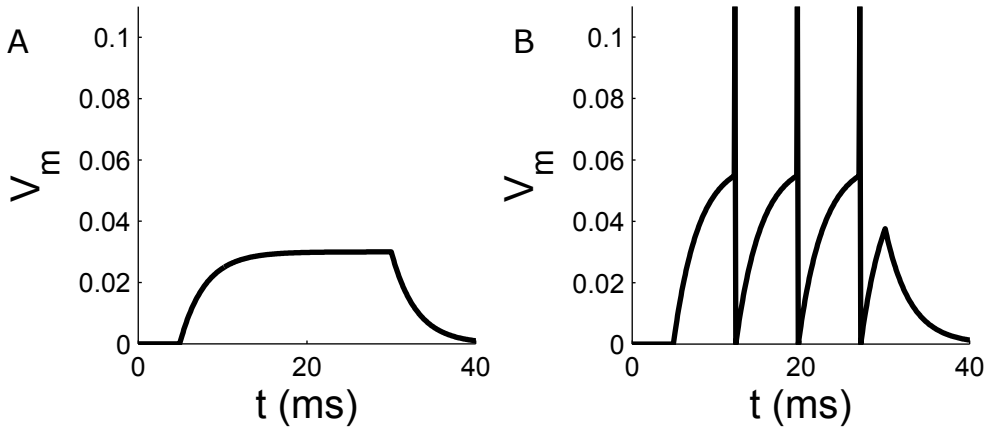

Fig. 2.3 **Membrane potential behaviour in leaky integrate and fire neuron models. (A)** Sub-threshold membrane potential at equilibrium in response to constant current input. **(B)** Spiking behaviour in integrate and fire neurons in response to super-threshold current injection. Membrane potential given in arbitrary units. firing threshold = 0.05. Spikes are artificially imposed for demonstration purposes. R = 1, τ_m = 10ms, and $V_r = 0$

2.2.2 Leaky integrate and fire neuron model

Leaky Integrators by themselves do not behave in the manor of typical neurons, as they merely act as a low-pass filter over the input received, and evolve accordingly. Due to this, they do not make a good approximation of a single neuron, but instead can be used to approximate the amalgamated behaviour of neuron channels, in which the activation of whole subsections of neurons carry the useful information, rather than the individual timed activations of specific neurons. These models provide a scalable and efficient tool, and thus used as models of channel activation in the striated channels of neurons in the basal ganglia.

To further approximate an individual neuron behaviour, we need to introduce and record the phenomena of neural spiking.

Biological neurons maintain an intra-cellular membrane potential by means of complex continuous interactions between the sodium and potassium ion channels in the cell membrane. When the input to the cell is large enough, a cascade of cellular behaviours is triggered, which results in a sudden surge in membrane potential, and equally sudden decay. This ‘spike’ in membrane potential can be propagated down the cell axon, and acts as an informational unit, eventually releasing chemical messages over intra-cellular junctions. This form of cellular communication is believed to be the backbone of the information processing occurring in the central nervous system.

To introduce spiking behaviour to the Leaky Integrator, we have to augment the equations of a leaky integrator with a non-linear threshold function. The spiking threshold function replaces the role of the complex action-potential generating cellular cascades, with a far more simple trigger function. This completes the basic Leaky Integrate and Fire neuron model (LIF) as:

$$\tau_m \frac{dV_m}{dt} = -V_m + R_m I_{inj} \tag{2.10}$$

with a reset condition for spiking generation as:

$$\text{if } V_m > V_\theta \text{ then } V_m \leftarrow V_r$$

such that V_m is the membrane potential, V_θ is the threshold level to elicit a spike, and V_r is the reset membrane potential.

The combination of a leaky integrator with the threshold, gives us the formulation of the Leaky Integrate and Fire Neuron, as discussed in Izhikevich and Edelman (2008). This simple formulation of the LIF suggests that the firing rate of the neuron is directly related to the magnitude of the input current. However a biological neuron has no such strict relation, and instead has implicit restraints placed on the achievable firing rate, due to the intrinsic cellular dynamics required to generate the spike. It takes time for ion channels to replenish the supply of Sodium and Potassium ions required for regular spiking behaviour.

In computational models we can emulate the biological restraints on firing rate, by introducing the concept of a 'refractory period' after a neuron spike. This can take the form of an absolute refractory period, whereby the membrane potential is forcibly held at the membrane rest potential for a specified time after firing, as shown in figure 2.4. Stopping the evolution of the membrane potential limits how close spikes can occur, regardless of the input current. Alternatively we can achieve a 'relative refractory period' such that the membrane potential is reset after firing to be lower than the resting potential. This still allows the neuron to fire, but forces each spike to be delayed after the previous due to the extra time and input needed to raise the membrane potential back above the firing threshold.

In practice both of these methods provide a natural maximum firing rate, which can be tuned to match biological measurements, and allow for a larger array of possible neural behaviours.

The LIF model neuron was the first step in developing computationally efficient models of artificial neuron behaviour. The mathematical simplicity of the model has led it to be very useful in studying the role of spiking neurons in plasticity and learning, as well as extracting

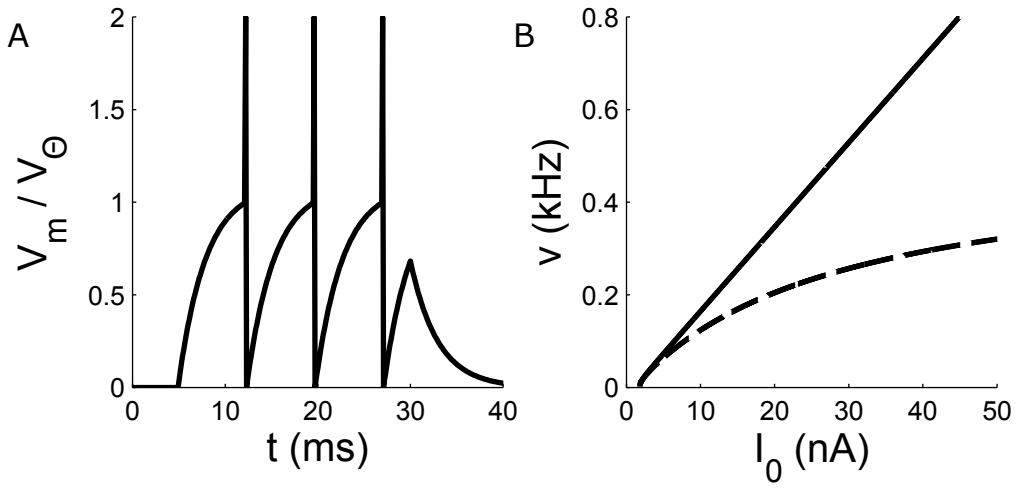

Fig. 2.4 **Refractory periods constrain in LIF firing frequency.** **(A)** Time course of the membrane potential of an integrate-and-fire neuron driven by constant input current $I_0 = 1.5$. The voltage V(t) is normalized by the value of the threshold $\vartheta = 1$. **(B)** Gain function. The firing rate ν of an integrate-and-fire neuron without (solid) and with absolute refractoriness of $\delta_{\text{abs}} = 4$ ms (dashed) as a function of a constant driving current I_0. Other parameters are R = 1, τ_m = 10ms, and $V_r = 0$.

emergent computation from large simulations of neural tissue. The LIF has been instrumental in the development of computational neuroscience.

However, the simplicity in the formulation of the LIF, does restrict the range of behaviours that the model can produce. While it was very important to capture the phenomena of spiking in a model neuron, the LIF is restricted to producing regular spiking patterns in response to a fixed input stimulus. It is important to know that there are many different forms of biological neurons, which can respond very differently to the same injection current. This variability is known as adaptation, and is a fundamental represents a limit of the LIF model.

2.2.3 Hodgkin, Huxley, and the circuit model of neurons

Computational neuroscience was born out of the hands on study of the physical and electrical properties of the central nervous system. The 'computation' in computational neuroscience does not necessitate that such research is done with the aid of computers, but rather that we study the computational power of the brain as a whole, and its minutest constituents.

Computational neuroscience became a field of its own, with the development of artificial models of neuron behaviour, derived from a combination of the study of the physiological measurements of the squid giant axon, and basic circuit theory (Hodgkin and Huxley, 1952).

Using electrical stimulation of a giant squid axon, Hodgkin and Huxley where able to observe an action potential travel down a squid axon. They where able to expand on a simple electrical circuit model, to derive a set of equations which when stimulated with a simulated injection current replicated the leaky integration seen in the squid axon. In ground breaking research they reproduced the phenomenon of a travelling action potential with a model derived from basic electronic principles, work for which they received a Nobel prize.

The Hodgkin-Huxley (HH) model expresses each component of the neuron as a electrical component. As above, the cell membrane is represented by V_m, the capacitance C_m. Voltage-gated ion channels depend on voltage and time and are represented as electrical conductances g_n (where n represents a specific ion channel i.e. Na or K. leak channels are represented as a linear conductance g_L. The electrochemical gradients driving the ions flow are represented by voltage sources E_n whose voltages depend on the ratio of the intra- and extracellular concentrations of the specific ionic species, n. Ion pumps are represented by current sources I_p.

As with the leaky integrator, the current flowing through the membrane is denoted:

$$I_C = C_m \frac{dV_m}{dt} \tag{2.11}$$

with the added current flowing through a given ion channel denoted:

$$I_n = g_n(V_m - E_r) \tag{2.12}$$

where V_n denotes the reversal potential for the given ion channel. For a typical cell with sodium and potassium channels, the current flowing through the cell can be denoted as:

$$I = C_m \frac{dV_m}{dt} + g_K(V_m - V_K) + g_{Na}(V_m - V_{Na}) + g_L(V_m - V_L) \tag{2.13}$$

where I is the total membrane current per unit area, C_m is the membrane capacitance per unit area and g_K and g_{Na} are the potassium and sodium conductances per unit area. V_K and V_{Na} are the potassium and sodium reversal potentials, and g_L and V_L are the leak conductance per unit area and leak reversal potential, respectively. V_m, g_{Na}, and g_K, are time dependant, and where the g_{Na}, and g_K also depend explicitly on voltage.

Hodgkin and Huxley further described the Ion channel models using a set of differential equations, to represent ion channel activation and inactivation as:

$$\frac{dm}{dt} = \alpha_m(1 - m) - \beta_m m \tag{2.14}$$

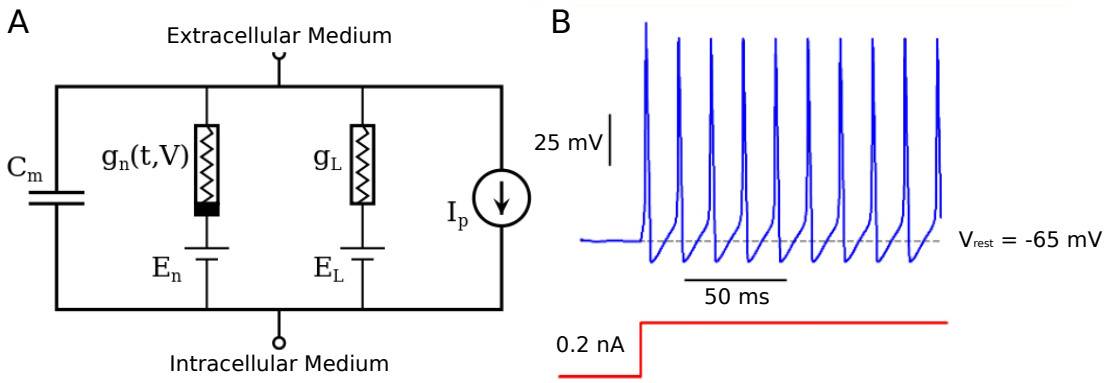

Fig. 2.5 **Intrinsic spiking behaviour in the Hodgkin and Huxley model neuron. (A)** The circuit diagram underpinning the Hodgkin and Huxley model neuron. Leak currents are represented as variable resistors, with individual ground states. Ionic equilibrium drives membrane potential evolution to produce spikes. **(B)** Intrinsic regular spiking behaviour of a Hodgkin and Huxley model neuron in response to a 0.2nA current injection.

$$\frac{dh}{dt} = \alpha_i(1-h) - \beta_i h \tag{2.15}$$

$$\frac{dn}{dt} = \alpha_i(1-n) - \beta_i n \tag{2.16}$$

where α_n and β_i are voltage dependant rate constants for the i-th ion channel. These combine with the specific ion channel state variables to describe the total current through he membrane as:

$$C_m\frac{dV}{dt} = \bar{g}^{Na}m^3h(E_{Na}-V) + \bar{g}^{K}n^4(E_K-V) + \bar{g}_L(E_L-V) + I_{inj} \tag{2.17}$$

where $\bar{g}_n$ is the maximal value of the conductance. n, m, and h are dimensionless quantities between 0 and 1 that are associated with potassium channel activation, sodium channel activation, and sodium channel inactivation, respectively. Fur further information on the voltage dependant activation, see Hodgkin and Huxley (1952).

The Hodgkin and Huxley framework details the biophysics of the membrane, guided by a qualitative examination of ionic currents. This formalisation of the kinetic theory of gating dynamics lead to the quantitative model of membrane dynamics, enabling an explicit model of activation and inactivation dynamics of the ion channels. As shown in figure 2.5, this breakthrough enabled mathematical models of biological processes to systematically reproduce the detailed dynamics of action potential generation.

The Hodgkin-Huxley conductance based model has, with modern computation, been instrumental in computational modelling, and has stood the test of time. The model has been used to study complex cellular-electrical interactions such as dendritic computation, and has been augmented with further concepts, such as intrinsic channel noise (Linaro et al., 2011), bringing even greater biological accuracy. However, further developments have also brought us much simpler models of neuron models (Gerstner and Kistler, 2002) which can be used to study very different computational properties of neurons and networks, without the computational expense of innate accurate spike generation through ion channel interaction.

2.3 Neuron models dictate model behaviour

Since the conception of computational neuroscience, we have seen a plethora of neuron models being developed (Hodgkin and Huxley, 1952; Izhikevich, 2004; Saarinen et al., 2006; Scharstein, 1979). Further to the ground breaking Hodgkin and Huxley conductance based model we have seen the adoption of voltage based models and stochastic models of neuron behaviour. Each model type often lends itself to a distinctive use-case, often trading biological-realism for computational efficiency to suit the task at hand. Modifications to the HH model have been used to study the intricate interplay of dendritic morphology, utilising formalisations such as cable theory to discover the subtle role of dendritic computation in single model neurons. On the opposite side of the scale, reduced model neurons such as the Leaky Integrate and fire model have lead the way in large scale models of neural substrates. Large organisations of simple neurons have shed light on the emergent computations performed in large volumes of densely connected tissue. In this thesis we largely focus on a mid-way approach to modelling, using a class of models called 'Two-variable' neuron models, which allow us to approximately replicate a large variety of spiking behaviours, while remaining computationally efficient enough to simulate large volumes of tissue, on non-prohibitive time scales. Here I will discuss the developments which lead to reduced neuron models and how they capture neuron behaviour.

2.3.1 Adaptation is an important feature of biological neurons

Biological neurons fulfil a huge variety of roles throughout the central nervous system, and as such are found with a complex set of morphologies and intrinsic behaviours. The simplicity of the LIF model, while capturing the fundamental building blocks of neuronal

communication, cannot be tuned to reproduce the plethora of neuronal spiking behaviour we can observe in vivo.

The complex computations that are performed in the CNS depend on specific neuron behaviour, which is influenced by the electrical properties and morphological features of individual neurons. This poses an issue with most neuron models, which are in effect a single point model neuron with no explicit morphology.

In the mammalian brain, we can find pyramidal neurons, spiny neurons, etc. which exhibit very different behaviours in response to stimuli, as shown in figure 2.6. Adaptation in how neurons respond to stimuli allow for tailored computation over input stimuli, which is expressed by phenomena such as chattering, bursting and short term plasticity. Adaptation to stimuli has been shown to be incredibly important with in the sensory systems such as vision, and as such is a key requirement to any biologically realistic model neuron.

2.3.2 Two variable neuron models

In a discussion of neuron models, Izhikevich (2004) proposes that 'having a network of computationally efficient and biologically plausible cortical spiking neurons interconnected according to the principles of known anatomy of the neocortex should be the goal of every scientist exploring information processing in the mammalian brain'. In this analysis, Izhikevich (2004) argues that the LIF neuron model should be avoided at all costs, given its limits reproducing variable neuron behaviour. Instead, for any meaningful analysis of biological relevance, a neuron model which can reproduce a variety of neural behaviour must be used.

To satisfy the twin desires of both biological realism through varied neural behaviours, and the utilitarian desire for computational efficiency, we must incorporate the power of adaptation without the shackles of complex morphology. We can build upon the foundations of the LIF model - a single-variable model - by incorporating an additional tracked variable, the 'adaptation variable'. It can be shown that the adoption of a second tracked variable, suitably interacting with the membrane potential, is sufficient to reproduce a large array of the neuronal variability with only nominal computational overheads.

One such model to incorporate an adaptation variable, is the Izhikevich model neuron (Izhikevich, 2003), which utilises a further quadratic equation to capture the intrinsic properties underpinning non-linear membrane potential dynamics. There is debate within the neuroscience community as to whether a quadratic or exponential (i.e. (Brette and Gerstner, 2005)) adaptation term replicates biological phenomena more closely, but both variations succeed in hugely increasing the expressible behaviours of model neurons. We implement

the quadratic form of (Izhikevich, 2003), for compatibility with previous work (Humphries et al., 2009b, 2010).

In the biophysical form of the Izhikevich model neuron, v_m is the membrane potential and the adaptation variable u is the contribution of the neuron class's dominant ion channel:

$$C\frac{dV_m}{dt} = k(V_m - V_r)(V_m - V_\theta) - u + I \tag{2.18}$$

$$\frac{du}{dt} = a[b(V_m - V_r) - u] \tag{2.19}$$

with reset condition

$$\text{if } V_m > V_\theta \text{ then } V_m \leftarrow c,\ u \leftarrow u + d$$

where in the equation for the membrane potential (3.19), C is capacitance, V_r and V_θ are the resting and threshold potentials, I is a current source, and c is the reset potential. Parameter a is a time constant governing the time scale of the recovery due to the dominant ion channel. Parameters k and b are derived from the I-V curve of the target neuron behaviour, where b describes how sensitive the recovery variable u is to fluctuations in the membrane potential V_m. Parameter d describes the after spike reset of recovery variable u, and can be tuned to modify the rate of spiking output.

The two-variable class of model neurons allows the expression of a wide variety of behaviours in response to fixed current injection, as shown in figure 2.6. To study the emergent computation, the single point neuron must be augmented with a model of the synapse, to allow for inter-neuron communication.

2.3.3 Models of synaptic transmission

Here we introduce development in of mathematical models of synaptic interaction. Firstly, we will only consider quickly acting synapses which directly gate ion channels to allow current flow. Figure 2.8A shows the morphology of a ligand gated receptor. Receptors have an associated neurotransmitter (such as glutamate or GABA) which supply the ligand, a chemical component which activates the channel. These channels are also called ionotropic receptors as they directly mediate ion flow, as opposed to metabotropic receptors which act using slower metabolic mechanisms. Experiments using current injection have been used to study the behaviour of ligand gated receptors.

Current injection has been instrumental in the investigation of neuron behaviour and collection of empirical data to create and tune computational models (Linden, 2001). Figure

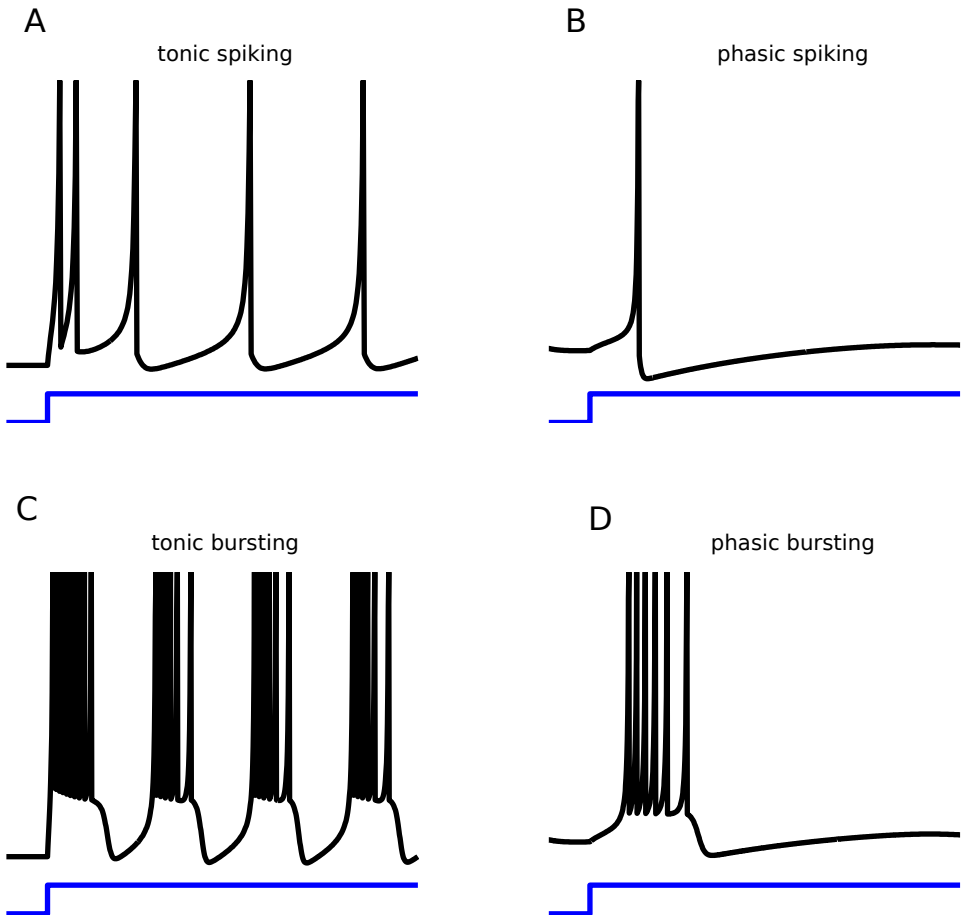

Fig. 2.6 **Adaptive spiking behaviour generated by the Izhikevich model neuron** The Izhikevich model neuron is capable of reproducing realistic spike train adaptation in response to constant current injection, through parameter fitting. **(A)** Tonic Spiking Behaviour with parameters $a = 0.02, b = 0.2, c = -65, d = 6$. **(B)** Phasic Spiking Behaviour with parameters $a = 0.02, b = 0.25, c = -65, d = 6$. **(C)** Tonic Bursting with parameters $a = 0.02, b = 0.2, c = -50, d = 2$. **(D)** Phasic Bursting with parameters $a = 0.02, b = 0.25, c = -55, d = 0.05$. Adapted from Izhikevich (2003)

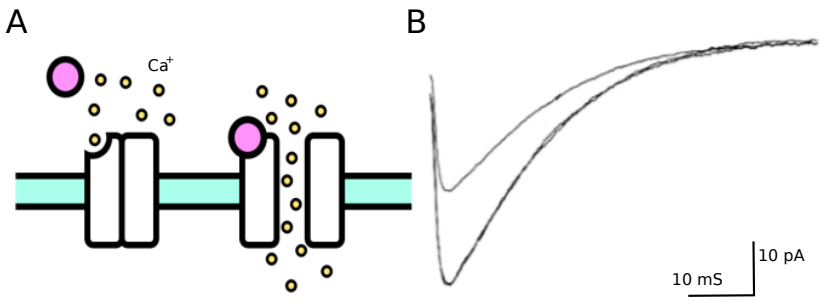

Fig. 2.7 **Synaptic dynamics in current clamp experiments**. **(A)** Ligand gated ion channel. Ions are permitted through the membrane when the correct number of ligands are bound to the receptor. **(B)** Post-synaptic AMPA EPSP from a current-clamp experiment (reproduced with permission from Linden (2001)). Trace shows the reduction in injected current required to maintain a constant current, as such this shows the inverse of the elicited post-synaptic potential.

2.7 B show data from a current injection experiment to measure evoked post-synaptic currents. A voltage clamp is used to dynamically force a constant membrane potential in the soma, using a closed feedback circuit. Figure 2.7B details the change in the current injected by the current clamp in response to an excitatory post synaptic potential (EPSP) reaching the soma. This measurement gives us an inverse picture of the shape and character of an incoming EPSP. We can use this data to derive mathematical models of synapse conductance, to improve our model neurons.

Ionotropic receptors result in EPSPs that can be described in a similar way to other membrane currents, using Ohm's Law in a conductance based formalism

$$I_{syn}(t) = g_{syn}(t)(E_{syn} - V_m(t)) \tag{2.20}$$

where g_{syn} is the synaptic conductance, and E_{syn} is a reversal potential and depends on the type of synapse in question.

This model can then be easily incorporated in the synaptic model using:

$$C_m \frac{dV_m}{dt} = I_{ion} + I_{syn} + I_{inj} \tag{2.21}$$

where I_{ion} represents both the active and passive 'leak' currents.

The most commonly ionic types are the glutamatergic NMDA and AMPA, and the GABAergic $GABA_A$ receptor families. NMDA is a slow-acting channel with a voltage-

dependant conductance, in contrast AMPA is a very fast acting channel, usually co-localized with NMDA receptors. Model parameters must be tuned accordingly.

In figure 2.8B we can see that the rise time τ_{rise} is much shorter than the decay time τ_{decay} in evoked post-synaptic currents. In a simple model of EPSPs, we assume that τ_{rise} is negligible, and set $\tau_{rise} = 0$. This gives rise to the single exponential synaptic model of synaptic conductance:

$$g_{syn}(t) = g_{max} e^{-t/\tau_{decay}} \tag{2.22}$$

where τ_{decay} parametrises the decay time of the EPSP. This model is often used in the LIF model, but neglects some of the intricacies of natural synaptic dynamics, such as the rise time of the EPSP.

The dual exponential model of conductance dynamics can explicitly tune the rise and decay times. This can be useful when detailed, small time scale simulations need to track recorded EPSPs closely. This can be formulated as:

$$g_{syn}(t) = g_{max} \left(e^{-t/\tau_{decay}} - e^{-t/\tau_{rise}} \right) \tag{2.23}$$

where τ_{rise} and τ_{decay} are used to explicitly tune the rise and decay times of the EPSPs, and t is time in *ms*. While very powerful, this model is considered too computationally intensive to utilise in many models, and often overlooked for the simpler alpha synapse model.

The alpha synapse allows for a non-negligible rise time, while remaining simple to tune, with a single time constant τ_{peak}, which can be tuned to dictate the rise and decay time.

$$g_{syn}(t) = g_{max} \frac{t}{\tau_{peak}} e^{-t/\tau_{peak}} \tag{2.24}$$

Models of synaptic saturation

A popular current method of implementing synaptic saturation in neuron models, is derived from a simple kinetic model of receptor binding and unbinding (Wang, 1999). An example of a kinetics based approach is detailed in Destexhe et al. (1994), which is summarised as follows. At the arrival of a pre-synaptic spike, neurotransmitter molecules T are released into the synaptic cleft, binding to the post-synaptic receptors R with the kinetics:

$$R + T \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} TR* \tag{2.25}$$

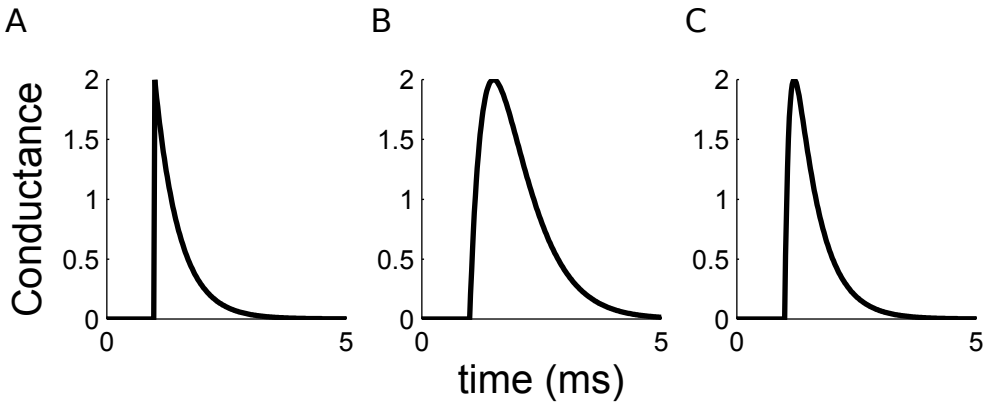

Fig. 2.8 **Synaptic models drive membrane response.** Synaptic models shape synaptic input, to produce more realistic membrane dynamics with synaptic integration **(A)** Single exponential model with $\tau = 0.005s$. **(B)** Dual exponential function with $\tau_{rise} = 0.001s$ and $\tau_{decay} = 0.005$. **(C)** Alpha synapse with $\tau = 0.005s$

where R and $TR*$ represent the bound and unbound receptors, and α and β represent the binding and unbinding rate constants. If r represents the fraction of unbound receptors, then these dynamics are:

$$\frac{dr}{dt} = \alpha[T](1-r) - \beta r \tag{2.26}$$

given $[T]$, the concentration of neurotransmitter. This model depends on tracking the level of saturation at each individual synapse on each of the neurons present in the full scale model. In chapter 3 we show show how lumping the traditional individual saturating synapse memory into a single lumped variable per channel per neuron, successfully approximates the neuronal effects of saturation, while providing the computational efficiency required to run saturating neurons in large scale models.

2.3.4 Neural coding schemes

Much work has gone into deciphering the language of the brain. While patterns of action potentials are important for communication between neurons, it is not as clear how the brain encodes information in a wider context. Three important hypotheses for information encoding are ‘Rate Coding’ and ‘Temporal Coding’ and ‘Population Coding’.

The rate coding model of information encoding states that, as the intensity of a stimulus increases, the frequency of elicited action potentials increases to reflect the intensity information. This hypothesis requires that most of the information about a stimulus can be encoded

by modulations in the firing rate of a neuron. In this case, neuron output can be treated as a stochastic output, with which the specific timing of action potential does not specifically encode information.

Rate coding was first proposed in Adrian (1926), to describe the output of sensory neurons connected to muscle fibres under increasing strain. In time, rate coding has been successfully demonstrated in many sensory systems, including vision and navigation. Since rate coding depends on the accumulation of spike information, one major disadvantage present in rate coding is reaction times due to the time taken to integrate new information, however this approach leaves the coding robust to noise in the precise spike timings.

Problematically, the firing rate $r(t)$ can never be exactly determined in a finite number of trials, and there is no unique way to approximate $r(t)$. Figure 2.9 shows a variety of ways to calculate the firing rate $r(t)$ over a single trial spike train. Where possible, more accurate approximations could be derived by averaging over multiple trials.

A simple way to extract an estimate of the firing rate is to divide the train into a series of discrete time-bins of length Δt, count the number of spikes in each bin, and divide by Δt. In reality, this calculates the spike-count firing rate over each time bin, such that the firing rate $r(t)$ within a given bin is estimated by this spike-count-rate. Figure 2.9B shows the binned firing rates for a 100ms bin size. This method will always produce discrete firing rates in multiples of Δt, and will always return a discontinuous firing rate. This method has the drawback that firing rates will arbitrarily depend on the bin placement, and that smaller bin sizes increase the temporal resolution but decrease the resolution for distinguishing different rates.

To avoid the arbitrary influences of bin size and placement, we can expand on the previous measurement by introducing sliding window function, moving the window of size Δt across the spike train, as shown in the jagged curve of figure 2.9 C. This function can be expressed as the sum of a window function over times t for $i = 1,2,3...,n$ when the n spikes in the particular sequence occurred,

$$r_{approx}(t) = \sum_{i=1}^{n} w(t - t_i) \tag{2.27}$$

where the window function is

$$w(\tau) = \begin{cases} 1/\Delta t & \text{if } -\Delta t \leq t \leq \Delta t/2 \\ 0 & \text{otherwise.} \end{cases} \tag{2.28}$$

While the windowing function provides a rate which provides better temporal resolution, it is worth remembering that rates obtained for times that separated by less than a bin width will be correlated as they will both contain some of the same spikes.

Equation 2.27 can also be expressed as the integral of the window function multiplied by the neural response ρ:

$$r_{approx}(t) = \int_{-\infty}^{\infty} d\tau w(\tau)\rho(t-\tau) \tag{2.29}$$

This equation represents a linear filter, and the filter kernel, the window function w, specifies how the spike train ρ, defined as

$$\rho(t) = \sum_{i=1}^{n} \delta(t-t_i) \tag{2.30}$$

contributes to the firing rate $r(t)$. The jagged curve seen in figure 2.9C is due to the discontinuous shape of the linear filter in equation 2.28. A linear filter can be used with almost any window function that goes to zero at $\tau = 0$, with an integral of 1.

We can obtain a smooth firing rate approximation by using a continuous windowing function, such as a Gaussian curve:

$$w(\tau) = \frac{1}{\sqrt{2\pi\sigma_w}} e^{-\frac{\tau^2}{2\sigma_w^2}} \tag{2.31}$$

Here σ_w controls the temporal resolution of the calculated firing rate, similar to Δt. This continuous window is used to create the continuous firing rate estimation seen in figure 2.9D.

In contrast to rate coding, temporal coding posits that information can be carried in the precise timing of action potentials, and in the time between spikes (inter-spike interval) (Gautrais and Thorpe, 1998; Rullen and Thorpe, 2001), which would not be ignored and intentionally smoothed out of the rate code as noise. As support for temporal coding, sensory systems such as sound localisation provide examples of millisecond processing, triggered by sparse neural spiking. In these examples, initial spiking to a sensory input is required to carry information about a stimulus within short enough time-frames that a rate coding model cannot easily account for.

Population encoding posits that rather than a single neuron encoding information, populations of neurons can encode precise information about a stimuli, and remain robust to individual neuronal noise (Hemmen and Schwartz, 2008; Page and Duffy, 2003). In a population, if each neuron has a distribution of responses over a set of inputs, it can be shown

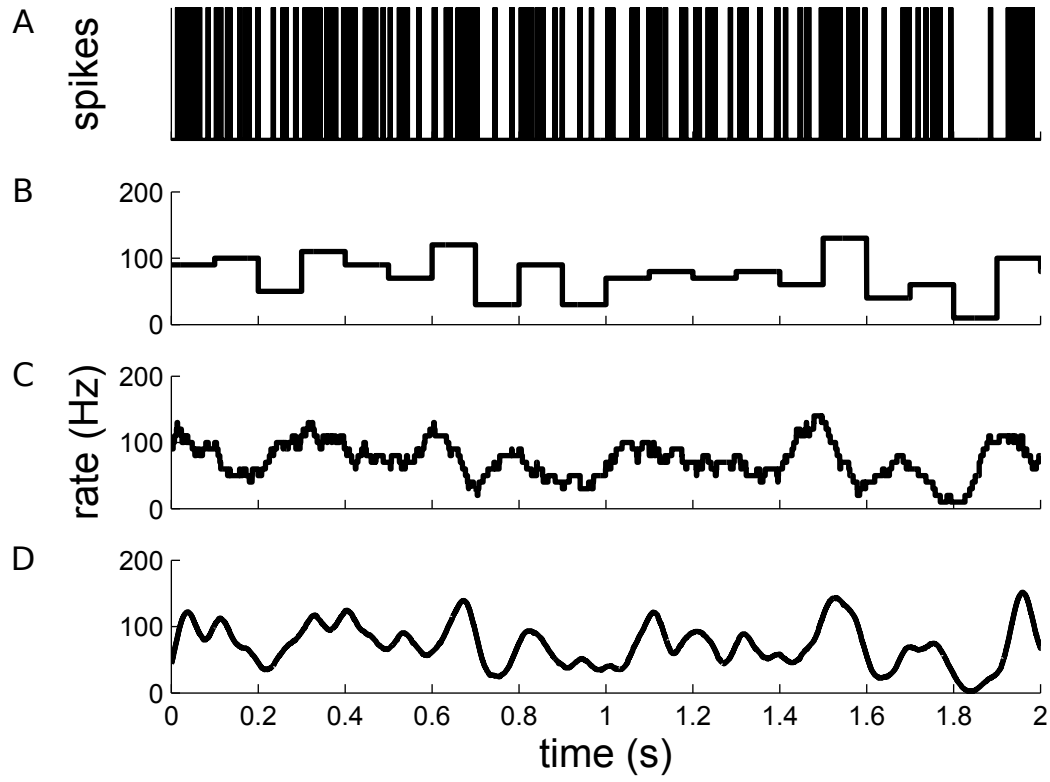

Fig. 2.9 **Firing rate metrics must approximate instantaneous firing frequencies** **(A)** Precise Spike Timings **(B)** Discrete-time firing rate obtained by counting spikes in non-overlapping time-bins of size $\Delta t = 100ms$. **(C)** Approximate firing rate determined by a sliding rectangular window of size $\Delta t = 100ms$. **(D)** Approximate firing rate using a sliding Gaussian window on $\sigma = 100ms$

that the combination of neuronal outputs can be combined to encode specific attributes of the input stimuli. Population encoding has been widely studied in the visual system in the medial temporal as a method of encoding direction information about a stimulus.

In practice, rate coding and temporal coding likely co-exist (Huxter et al., 2003), with theoretical studies suggesting that differential connectivity regimes imply differential coding schemes (Clopath and Büsing, 2009). Each encoding has inherent drawbacks and unique capabilities which are required for the global information processing in the CNS (Masuda and Aihara, 2003; Prescott and Sejnowski, 2008). In this paper we focus on the rate coding of distinct channels found in the basal ganglia, but show how precise spike timing in response to a changing stimulus can provide extra information, and aid in decision making and action selection.

2.4 Action selection is a well studied problem

The problem of action selection has been studied from a variety of disciplines including ethology, psychology, neurobiology, computational neuroscience, artificial intelligence, and robotics. Action selection has also been addressed as decision making, behavioural switching and motor program selection. Selection is intertwined with the study of attention, where large attentional shifts are linked with behavioural changes. Research into action selection can be grouped into the problem of selecting for optimality, or the problem of control architecture to general action selection. Behavioural studies have been performed to test the optimality of human decision making in noisy-tasks (Cassey et al., 2013), and have shown that human evidence accumulation is not always in line with optimal strategy. This thesis deals with the architecture of action selection, its computational design and its biological implementation, with optimal strategies usually the purview of cortical structures not address in this discussion. There are theoretical investigations which suggest that this division is artificial, and suggest that, as the evolution of control structures occurred simultaneously with the evolution of selection, that these two aims are the same. Bogacz and Gurney (2007) suggest that the basal ganglia can support a form of optimal behavioural switching, for evidence accumulation for more than 2 tasks. They suggest that a configuration of the basal ganglia sub-nuclei perform an implementation of the MSPRT statistical test. However, I would suggest that this study also conforms to the division of optimality and structure, due to the reliance of the MSPRT implementation on a very specific interpretation of connectivity in the basal ganglia, as indeed there are many, as shown in figure 2.10.

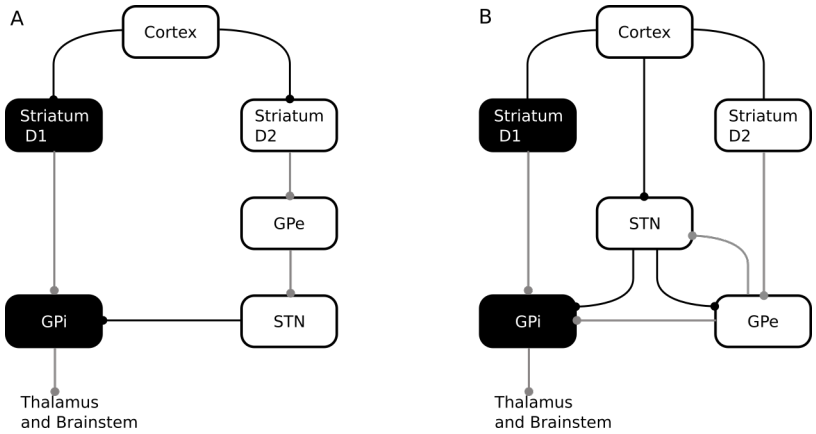

Fig. 2.10 **Competing connectivity regime in the basal ganglia A** The direct and indirect pathway model of BG anatomy proposed by (Albin et al., 1989) **B** The proposed new functional anatomy of the BG, from (Gurney et al., 2001) The selection pathway (shaded boxes) causes the disinhibition of the BG's projection targets. One of the actions of the control pathway is to scale the level of activity in the selection pathway. The STN plays a role in both pathways. Excitatory connections in black, inhibitory in grey. Reproduced from (Gurney et al., 2001)

The concept of action selection in artificial agents, such as robots lends it self to a clean and defined description of action and action selection. However in biological systems the notion of discrete action selection is controversial, as the assumptions of behavioural decomposition have no logical atomistic formulation in continuous action spaces. It is not obvious for instance, whether while taking a drink, the act of drinking should be considered the action, or the individual actions of moving ones arm, grasping a glass, etc. It is often unclear when studying biological systems what the 'units of selection' are composed of. One traditional school of thought evokes the concepts of behavioural hierarchies (Grafton and Hamilton, 2007), wherein upon the selection of a high level task, such as drinking, cascades of lower level tasks are selected and in turn executed by further sub-actions. This framework of behavioural decomposition depends on the existence of some set of sensorimotor primitives with access to the effector systems, such as muscles, ultimately constraining the atomistic action set.

Alternatively, recent research has considered a distributed approach to a action selection, removing the specific need to an explicit switching system. A dynamical systems approach could encode desired action-chains as networks of attractor states, removing the need to explicit encode actions, states and transitions (Eliasmith, 2005; Ijspeert et al., 2002; Montebelli et al., 2008). Artificial agents have been built using both behavioural decomposition and

attractor state dynamics, providing an interesting proof of concept for each alternative, but as yet provide nowhere near the complexity or robustness of biological systems. Biologically however the question of the style and modularity of control remains an open question. In this thesis we consider the role of both simple neural circuits and complex networks of ganglia with respect to solving the action selection problem.

An intelligent agent must solve three major issues to survive in a changing environment. First, Information must be gathered from the environment and encoded in such a way as that the information can be processed internally. Secondly, agents must combine and process this abstract information, and create strategies for action and survival. Lastly these strategies must be translated into actions to influence the environment, and continue the sensory-behavioural loop. We focus on the latter problem of how to select between competing actions. It is imperative for natural agents to limit mutually exclusive actions, and as such a mechanism is needed to reduce the available action set, to a reduced action plan.

Neurons are selective to input by their very nature, and so to qualify as an action selection mechanism, stringent requirements must be met to determine whether a subsystem is specifically used in action selection. Redgrave et al. (1999) have suggested that in order to be considered an agent of action selection, the subsystem must exhibit the following properties:

- (i) the system of interest should have inputs that carry information about both internal (to the body) and external (outside the body) cues relevant to decision-making
- (ii) there should be some mechanism that allows calculation of the salience that should be attached each available action,
- (iii) there should be mechanisms that allow for the resolution of conflicts between competing actions based on their relative salience,
- (iv) the outputs of the system should be configured so as to allow the expression of winning actions whilst disallowing losers.

2.4.1 Action selection reduces to signal selection

In an abstract agent, we have to consider how competing actions to a given stimulus are encoded, in order to determine how to select the most appropriate action. In Chapter 1 we discussed the various information encoding schemes used throughout the brain. Interestingly it seems that information encoding schemes vary depending on the task, while the touch-receptors in leeches encode pressure using a continuous representation using a rate coded

representation (Kandel et al., 2000), more specialised tasks use a more complex temporal coding, in which information is encoded in more that just the average firing rate (Gautrais and Thorpe, 1998).

When given the task of selecting between competing actions, it is a fair assumption that a common information-encoding scheme is shared between competing actions, such that individual action value can be directly compared, to select the most appropriate task (Seo et al., 2012). To this end we may use the rate-coding scheme, which specifies that increased firing rates encode for increased action utility or salience to the agent.

Simplifying the panoply of possible agent actions, to an efficient array of directly comparable continuous variables, grants a variety of advantages, both in biological terms, and computational. Primarily the simplification of actions to mere signals allows the direct comparison of disparate concepts, with regard to an ultimate utility to an agent, allowing a central mechanism for decision making to control limited physical resources. This simplification addresses point (iii) in the computational requirements of an action selection mechanism, by simplifying the concept of action comparison to the comparison of signal magnitudes.

In a biologically relevant agent, it is not simply enough to suggest that an action with maximum utility is selected (computationally trivial with a max() function), but an underlying neural mechanism must be provided for reducing the plethora of actions, to a reduced set of actions to be carried out.

As discussed in chapter 1, there is a variety of neuron types in the brain, with a range of different morphologies and computational functions. This toolbox of neuron functions allows for theoretical construction of complex and powerful networks, which are conducive to competitive dynamics such as mutual inhibition, and targeted dis-inhibition as shown in figure 2.11.

2.5 Models of selection in artificial agents

Artificial neural circuits can be constructed to satisfy the requirement of action selection mechanisms to both (i) provide a mechanism to resolve conflicts between salient actions and (ii) suppress all but the selected action. Such simplified circuits provide a useful insight into the use of excitatory and inhibitory dynamics to achieve signal selection in spiking neural networks and can be taken as useful templates which can be applied to all spiking neuron models.

Figure 2.11 demonstrates some simple mutually connected neural architectures which can support these features of action selection. In each architecture, the circles represent

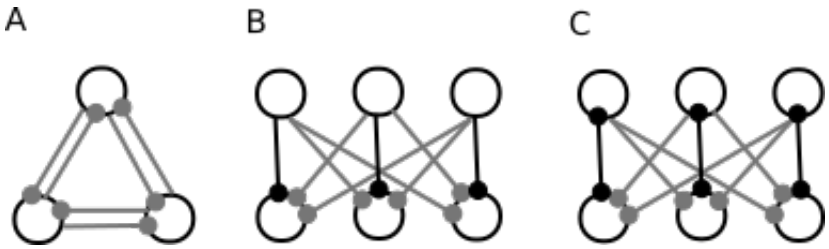

Fig. 2.11 **Distributed artificial neural networks can achieve signal selection.** Distributed inhibition enables a simple signal suppression mechanism, but scales poorly **(A)** Reciprocal recurrent inhibition architecture **(B)** Feed-forward architecture **(C)** positive feedback connectivity. Gray lines represent inhibition, black represent excitation. Double headed lines represent mutual connections.

neuron with the dashed and solid lines representing inhibitory and excitation connections respectively. Figure 2.11A details a fully connected, all inhibitory network, termed reciprocal recurrent inhibition (RRI). The RRI architecture is a typical architecture in action selection as it is capable of exhibiting full selection of a winner, absence of distortion and clean switching (Redgrave et al., 1999). RRI also exhibit hysteresis, allowing a selected action to remain selected in the presence of equally strong competition. The RRI function by the use of positive feedback. Disproportionally strong input to any of the cells, will inhibit the other cells, and thus receive less inhibition in return.

Figure 2.11 B shows a feed-forward network. In this case a distinction is made between the input neuron layer which amalgamates the input from external sources, and the output neuron layer which encode for the action to be taken. Here each input neuron inhibits competing output neurons while excitation is given to their cosponsoring output neuron. This architecture will provide action selection without the feedback loops seen in RRI networks. Figure 2.11 C combines the feedback approach with the abstraction of input and output layers, enabling the clean selection and hysteresis of the RRI network.

A clear disadvantage of the network architectures described by figure 2.11 is the cost of scaling the network to larger decision networks, due to the fully connected nature of the architecture. Figure 2.12 A describes a method of connection called a Winner-Take-All (WTA) network such that each new neuron only adds two connections, making this a uniquely efficient and scalable system, with heavy emphasis on the centralised control.

Winner-Take-All (WTA) is a computational principle, in which neurons compete with each other, for sole activation. WTA computation can be performed by a variety of neural networks, using recurrent networks, and inhibition. Computationally, for a network to exhibit WTA behaviour, after a finite time, only a single neuron should be activated, which should

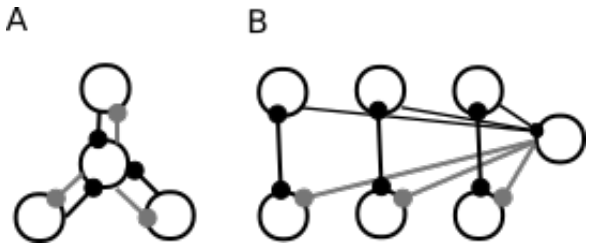

Fig. 2.12 **Centralised inhibition provides efficient signal selection.** Centralised inhibition networks provide an efficient scalable action selection mechanism, but require fine balancing of inhibition and excitation **(A)** Feed forward network with feedback and centralised inhibition **(B)** A selection mechanism with centralised inhibition, which is a driven by the incoming channel activations. Continuous inhibition requires strong enough excitation to overcome inhibition. Gray lines represent inhibition, black represent excitation. Double headed lines represent mutual connections.

correspond to the most salient input to the network. In this classical form of WTA, only the neuron with the highest activation stays active while all other neurons shut down, however other variations that allow more than one neuron to be active do exist, for example the soft winner take-all, by which a power function is applied to the neurons.

Figure 2.12 B describes a network architecture which introduces a global inhibition signal, drastically increasing the scalability of the network. These architectures, while being more efficient in terms of connections will require a more complex inhibition network, such as a tuned global inhibition signal to balance the excitation.

Local decision making networks such as these have been studied in the cortex (Wang, 2002), and can be shown to process competing information, such as encoding for stimuli direction. Recurrent inhibition can provide a candidate mechanism for the slow time integration of sensory stimuli seen in the cortex, and can be applied for the formation of categorical choices in a decision-making neocortical networks.

A rate-encoded scheme for signal selection lends itself naturally to Winner-Take-All network dynamics. WTA networks have been extensively investigated as an effective signal selection mechanism. Early models used analogue inputs and outputs on transistors (Oster and Liu, 2006) and progressed to models with spiking neuron inputs using non-leaky integrate and fire models. Oster and Liu (2006) provides a good start point for modelling action selection in neural networks, as it provides a review of the current simple network architectures, as shown below, leaving interesting mechanisms worth investigating. Later work in Oster et al. (2009) further discusses spiking inputs to a winner takes all network, using non-leaky Integrate and

fire neurons with spike trains generated from both constant inputs causing regular firing and also input neurons driven by poisson distributions for spike generation.

While these architectures are simple in their formulations, they provide insights into the restraints that a biological mechanism for action selection will face. Efficiency of scaling, and robustness to connection disruption are serious issues that any successful neuron mechanism for selection will have to overcome. The metabolic sizeable cost of highly interconnected networks could be a guiding factor in the development of more efficient, scalable centralised methods of action selection. Ringo (1991) noted that the degree of interconnectivity in the brain is restricted by certain geometric factors. In this study larger brains did not demonstrate as high a degree of interconnectivity as smaller brains. Also, Leise (1990) have noted that modularity plays a role as a shared design pattern in the CNS of vertebrate and invertebrates, with both functional and anatomical local compartments. These compartments of up to 1mm show higher connectivity within themselves than externally, due to the higher cost of high bandwidth traffic over long distances.

Together these observations on neuroanatomy suggest that evolutionary refinement of robust and cost effective selection mechanism could prefer the more efficient methods of decentralised action selection over large distances in the brain, with costly highly interconnected microcircuits reserved for localised competition dynamics.

In this thesis we examine this possibility by considering the role of the highly interconnected inhibitory microcircuits of the striatum, within the longer range channel wise inhibition, excitation and dis-inhibition networks of the basal ganglia, uniting these two opposing architectures at, different scales of abstraction.

The giant Mauthner Cells (M-cells), present in the brain stem of most fish provide an instructive example of complementary action selection mechanisms at different levels of processing. M-Cells are known to play a role the ‘C-start’ escape manoeuvre used by many fish as the primary mechanism to avoid predation and other threats. Eaton et al. (1995) have argued that the M-cell does not play an active role in behavioural comparison, but instead acts to quickly inhibit all behaviours but C-Start, prohibiting involved decision making, and enabling fast action selection for rapid escape. Evidence shows that removal of the M-cell does not stop the C-start behaviour, only slows down selection. Due to the similarity of brain stem organisation across vertebrates, similar mechanisms may play a role in vertebrates (Lingenhöhl and Friauf, 1994). In chapter 5 we consider how a continuum of urgency and then need for quick suppression could be locally enforced by the striatal microcircuit.

It has been suggested that behavioural switching methods could be ‘hard wired’ in fixed-priority hierarchies, with conflict resolution determined by the priority of distinct levels

of the neuraxis (spinal cord, hindbrain, midbrain, etc.). Prescott et al. (1999) reviewed the evidence that higher levels can suppress and modulate the behaviour of lower levels in vertebrate defence systems. However, fixed-priority hierarchies have failed to explain the complex interplay of behavioural systems, such as breading, feeding and fleeing, found in adult vertebrae. In vertebrae, such behaviours are influenced by the presence of social dominance hierarchies, and behaviour is adaptable beyond the capability of fixed action switching.

It is interesting to note, that despite the computational cost of reciprocal recurrent inhibition networks mentioned above, several examples of that architecture have been documented in many areas of the vertebrate brain (Windhorst, 1996). RRI may occur in local action selection mechanisms across the vertebrae brain (Gallistel, 1981), with top-down modulation from cortical systems (Deco and Rolls, 2005; Desimone and Duncan, 1995). These mechanisms are likely to only play influential roles over small local subsystems of selection due to their high metabolic and computational cost.

2.5.1 The role of the basal ganglia in decision making

Finding the neural substrate for the process of selection is key to furthering our understanding of decision-making (Ding and Gold, 2013), action selection (Grillner et al., 2005; Mink, 1996), planning (Houk and Wise, 1995), action sequencing (Jin and Costa, 2010), and even working memory (Gruber et al., 2006). One unifying proposal, formulated first by Redgrave et al. (1999), suggests that the basal ganglia forms just such a generic selection mechanism (Prescott et al., 1999; Redgrave et al., 1999). This proposal neatly explains why the basal ganglia have been hypothesised to contribute to each of the above functions, but specifying the exact computational process of selection by the basal ganglia is challenging (Berns and Sejnowski, 1998; Gurney et al., 2001; Leblois et al., 2006).

The basal ganglia are a well studied collection of interconnected sub-cortical nuclei, composed of a variety of neuron types, capable of supporting a variety of neural architectures (Wilson, 2010). The basal ganglia consist of the striatum, subthalamic nucleus (STN), globus pallidus (GPe) in the forebrain and the Substantia nigra pars (SNr) reticulata and the Substantia nigra pars compacta in the midbrain. There is an emerging neuroscientific consensus that the basal ganglia are required for both enabling desired actions, and inhibiting undesired actions. It appears that the basal ganglia are well placed as a general action selection mechanism, fulfilling the computational requirements set out in Redgrave et al. (1999).

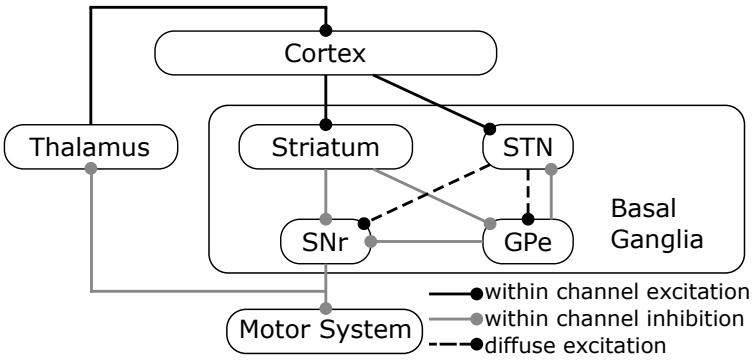

Fig. 2.13 **A computational view of basal ganglia selection circuitry.** The SNr provides selective disinhibition to the motor systems. Loops within the decision circuitry provide feed back, and modulate inhibition. Gray lines are inhibitory, Black, excitatory. Solid lines represent within-channel action scope, and dashed lines represent diffuse or tonic scope. Within the box represents the core basal ganglia nuclei, along with external thalamo-cortical loops. Abbreviations: GPe-Global Palidus external; STN-Subthalamus Nucleus; SNr-Substantia Nigra pars Reticular. Inspired by Gurney et al. (2001)

As the main input nucleus to the basal ganglia, the striatum receives cortical projections from functional subsystems of the forebrain which may represent 'requests for access' to motor control resources. Simultaneously the striatum receives afferent connections from a range of subsystems associated with motivation and sensory processing. This rich combination of external stimuli, internal drive and behavioural effectors grants the basal ganglia computational processing power over the exact relevant information needed to competently support action selection. Amalgamating these inputs,the striated channel architecture of the striatum could support an efficient, central representation of signal saliency, as required for action representation and comparison.

Importantly, the main output nuclei of the basal ganglia are tonically active, releasing a continuous flow of inhibition to areas of the brain directly or indirectly involved with generating movement. The tonic flooding of inhibition to movement associated areas acts to continuously suppress undesired movement. This mechanism enables the basal ganglia to successfully gate voluntary action control, using selective disinhibition, generated by the basal ganglia network.

The basal ganglia selection circuitry, when combined with the greater decision loops to the thalamus and cortex can be said to implement a variation of the feed forward selection described in figure 2.12. In contrast to the feed forward selection seen in figure 2.12D, the output of the basal ganglia uses selective dis-inhibition to release the tonic suppression of

involuntary action in the motor system. In this case the role of tonic inhibition is implemented by the subthalamic nucleus, by providing tonic excitation to the inhibitory SNr neuron populations, with the GPe mediating the balance of selective dis-inhibition (Gurney et al., 2001). Further, to the feed forward mechanisms employed in the BG networks, compartmental dis-inhibition can also be found at sub-mm scales within BG nuclei, such as the striatum.

Figure 2.13 highlights the role of the cortex as providing executive function to the decision making process. The development of the new cortical circuits in modern mammals (Butler and Hodos, 2005) cemented the supplementary role of the cortex in action selection. Evidence with decerebrated rats (Gallistel, 1981) suggest that the cortex enhances the motor systems lower down the neuraxis, enabling a layered control architecture. This provides multiple levels of sensory motor interaction, without being a necessary part of action selection (Redgrave et al., 1999). Experiments with decorticate rats and cats suggest that while the animals become more stimulus-bound and lack skilled motor control, they are still able to eat, drink and groom under appropriate motivational stimulus, displaying many behaviours of normal social adults. This suggests that the cortex is not the locus of action selection, but acts with the subcortical systems in the action selection loop. Here the cortex provides context in the form of experience and reward prediction to otherwise stimulus-coupled behaviour (Krawczyk, 2002; Ridderinkhof et al., 2004).

2.6 Using models of the basal ganglia to understand action selection

Studies investigating the role of the basal ganglia in action selection have exponentially increased in the last decade (Frank, 2011). The structural complexity of the basal ganglia has become a perfect candidate for the computational study of the action selection, which is fundamentally a computational problem to solve. The computation performed by the basal ganglia for action selection has been studied broadly (Graybiel, 1995; Mink, 1996), with initial computational models coming from Barto (1994); Doya (1999), expanding into the increasingly complex network models of (Frank, 2011; Gurney et al., 2001; Humphries et al., 2008; Prescott et al., 2006).

Early models primarily focused on elucidating the internal roles of the basal ganglia nuclei, establishing the formalised network of inhibitory and excitatory pathways. Building on the early successes of mapping form onto function, basal ganglia studies diverged into investigating the BG structure at different spatial scales. Studies examine everything from

the biophysical properties of constituent neurons (Chuhma et al., 2011; Tozzi et al., 2011; Xia et al., 2011), up to the abstract algorithmic properties of centralised switching methods (Gurney et al., 2004; Prescott et al., 1999). Further biologically inspired research has focused on unravelling the role that the basal ganglia has in the computational domains of dopamine in reinforcement learning (Bogacz and Larsen, 2011; Dayan and Balleine, 2002), and the intricate effects of neuro-degenerative diseases such as Parkinson's and Huntington's disease (Moustafa and Gluck, 2011; Redgrave et al., 2010).

In an important early paper Redgrave et al. (1999) pointed out that since the basal ganglia have been conserved across vertebrates, that the network could act be the vertebrate solution to the problem of action selection, this spurred the development of biologically inspired network-level architectures (Gurney et al., 2001) which aimed to map the biological structures onto functional algorithms. In these network level formulations of the basal ganglia, the BG is designed to act as a multiple-input 'channel-wise' general selector. Each channel is fed an abstract signal from the cortex, which is directly mapped as the salience of the represented action, by means of leaky integrator approximations of cell populations. The primary function of the Basal ganglia is to inhibit these channels down-stream using tonically active GABAergic neurons. The striatum then relays the signal salience by enabling selective dis-inhibition through the SNr. This system allows voluntary actions to become active while maintaining tonic inhibition on all other channels. Early models divided the functional anatomy of the BG into the direct and indirect pathway, allowing the direct pathway to dis-inhibit a channel, while the indirect pathway suppresses unwanted variation. Later the role of the indirect pathway was recast as a control process to allow for 'capacity scaling', enforcing the strict selection of a single action regardless of channel volume (Gurney et al., 2001).

Despite development of knowledge in the anatomical and computational function in the basal ganglia, there was no formal, theoretical framework to algorithmically describe the basal ganglia circuits. Bogacz and Gurney (2007) demonstrated that the computations performed by the cortex and the basal ganglia where exactly those required to implement the computation defined by an asymptotically optimal statistical test for decision making: the multihypothesis sequential probability ratio test (MSPRT). The MSPRT framework created a new way to understand the computation in the basal ganglia during decision making in highly practised tasks, with the framework predicting neural behaviour which matched experimental data.

Developing a maturing understanding of the established network-level functionality, computational models became free to confidently explore the network connectivity features

with respect to the various functional roles the basal ganglia have been implicated in, such as learning and memory formation. Early models of the BG assume that the channel values striatum represents the signal salience, directly as passed on from the cortex, other interpretations of salience have been modelled. Other approaches investigated the role of the basal ganglia in reinforcement learning (Graybiel, 1995), as the striatum contains a high proportion of Dopaminergic D1 and D2 expressing neurons. Dopamine has been heavily implicated in memory formation and synaptic plasticity and is necessary for the acquisition of stimulus-response associations (Choi et al., 2005; Smith-Roe and Kelley, 2000). Theoretical studies of reinforcement learning mapped dopamine as a potential neurotransmitter to encode for prediction-error signals (Montague et al., 1996) and have investigated the role in dopaminergic busting behaviour as the signals that drive learning in the BG (Schultz, 1998).

Studies have looked closer at the role of dopamine modulation on the SPN populations in the striatum (Centonze et al., 2003) which are associated with theoretical models of reinforcement models. Relationships have been discovered between the D1 and D2 pathways, which suggest modulation effects on risk behaviour and decision making. The discovery of active roles for the neuronal populations in larger decision network, allows us to break away from simplified network models, and incorporate more predictive power into our models, with studies further looking into the biochemical effects of neurotransmitters (Blomeley et al., 2009).

In contrast to the diversification of function in the basal ganglia, larger scale, more physiologically realistic versions of the basal ganglia where developed to understand the role of the complex structure within the nuclei of the basal ganglia. As physiological data about the structure and composition of the distinct nuclei improved, research started re-examining the computational roles performed within the BG nuclei, and within the cells themselves, often from the perspective of action selection. More in depth modelling of the SPN populations have discovered fundamental bistable membrane properties (Wilson and Kawaguchi, 1996). Often overlooked in simplified models, these discoveries provide novel insights into dynamic assemblies (Humphries et al., 2009b), and cell interaction (Koos et al., 2004), which could dictate network dynamics in larger decision networks.

As the intricacies of cell interaction and neurotransmitter modulates are being unravelled, the startling discrepancy of the role of the striatum became apparent. The striatum is the largest of the BG nuclei, comprising a multitude of cell types. The striatum contains self organising local inhibitory microcircuits which receiving massive cortical input from the cortex, pertaining to motor control, motivation, and memory. In spite of the physiological

evidence of the complexity of the striatum, it has generally been relegated to a mere relay of cortical signals through the BG architecture.

By combing accurately tuned reduced neuron models, with detailed three-dimensional connectivity statistics, Humphries et al. (2009b) investigated the computational role of the striatal microcircuit itself. They discovered that the striatal microcircuit was capable of dynamically generating cell assemblies using dopaminergic modulation, supporting the channel models described in early models.

Further investigations into the structure of the striatum Humphries et al. (2010) has revealed some important insights into the role of accurate connectivity statistics in emergent circuit dynamics. Humphries et al. develop their simulation using physiological measurements of the neuron types present in the striatum. Using biological spiny projection neuron measurements, they probabilistically build dendritic trees to determine three dimensional connectivity. Using this 3D structure, they were able to discover how connectivity statistics influence global striatal dynamics. The model predicted two novel network properties. Firstly that the dominant GABAergic input to an SPN arises from neurons with somas at the edge of its dendritic field, creating effective microcircuit ranges. Secondly that Fast Spiking Interneurons are intra-connected on two different spatial scales: locally by gap junctions and distally by synapses.

Building on this novel and accurate model of the striatum, this thesis looks at how the detailed striatal microcircuit can play an active role in action selection, elevating its status from a mere relay, to an active and additive component in the human decision-making process, based on microcircuit dynamics.

2.7 The effect of Huntington's disease on action selection

Huntington's disease is a progressive autosomal-dominant neuro-degenerative disease which primarily affects motor co-ordination, but also results in progressive mental decline and a severe behavioural modification. HD onset usually occurs in middle age, as the patients exhibit chorea and dystonia, poor-coordination, cognitive decline, and behavioural difficulties.

Sufferers of HD can be characterised as being symptomatic or pre-symptomatic carriers of HD. Pre-symptomatic carriers (pHD) do not yet show the physical manifestations of HD, such as chorea, but carry the gene responsible for HD. Patients with pHD provide an important role in diagnosis, mapping and treatment of HD early on in the disease progression.

The progression of HD varies within patients, but typical onset periods are observed across patients. Kirkwood et al. (2001) categorised 6 periods, in which first, involuntary

movements are grouped alone as the earliest reported symptom. This is followed by mental and emotional symptoms, including sadness and depression, followed third by clumsiness, sexual problems, lack of motivation, and suspiciousness/paranoia. The disease then progresses through a variety of physical symptoms including unsteadiness, trouble holding onto things, trouble walking, changes in sleeping patterns, delusions and hallucinations accompanied by intellectual decline, and memory loss. Late-stage HD patients tend to experience speech difficulty and weight loss, followed by a loss of bowel and bladder control (Kirkwood et al., 2001).

Craufurd and Snowden (2002) investigated the Neuro-physiological and neuropsychiatric aspects of Huntington's disease, and found that the vast majority of patients exhibited neuropsychiatric symptoms, the most prevalent being dysphoria, agitation, irritability, apathy, and anxiety (Kirkwood et al., 2001). Craufurd and Snowden (2002) proposed that the neuropsychiatric symptoms are relatively independent of cognitive and motor aspects of the disease and explored the link between neuropsychiatric symptoms and frontal-striatal circuitry.

Animal models have been used to uncover the biological symptoms of Huntington's disease, with some studies focusing on the neurological altercations imposed by the huntingtin mutation. Animal HD models have revealed alterations in the function of neuronal networks, particularly those linking the cerebral cortex and striatum (Eidelberg and Surmeier, 2011).

Paulsen (2011) suggests that cognitive decline in pHD patients can become evident as much as 15 years prior to the time at which motor diagnosis is given. This decline corresponds to biological markers of HD, such as progressive brain atrophy. Longitudinal studies have found that tasks measuring mainly attention, object-space perception and executive functions can adequately assess the progression of HD disease, while other functions do not significantly deteriorate. Interestingly, declines in functions associated with the basal ganglia, i.e. attention, working memory, and learning of random associations are the earliest cognitive manifestations in pHD patients.

The explicit links between the progression Huntington's Disease, cell death in the striatum, and progressive loss off precise voluntary control make Huntington's disease a prime target for studying the relationship between action selection and the basal ganglia.

2.7.1 Neurological effects of Huntington's disease

A major underlying cause of the mental and behavioural decline seen with HD, has been attributed to the progressive cell loss in the striatum (Feigin et al., 1995). It has been

shown that dysfunctions in striatal circuits and the cortico-striatal pathway occur during the progression of HD, before there is significant cell loss. Interestingly, Cepeda et al. (2007) suggest that morphological changes in the striatum are linked to alterations in the intrinsic functional properties of the spiny projection neurons, such as increased NMDA receptor sensitivity. However, increased receptors sensitivity requires further abnormalities in the cortico-striatal inputs to express the full HD phenotype of behavioural changes. Cepeda et al. (2007) emphasise the role of neuronal circuits in the expression of the HD phenotype, and point out that cortico-striatal abnormalities are a key problem to understand, forwarding our knowledge and treatments of HD.

Mouse models of HD show that the early increase in NMDA receptor expression contributes to phenotype onset in Huntington's disease (Milnerwood et al., 2010). Specifically increased extra-synaptic NMDA receptor expression in mouse models of HD and can be observed before and after phenotype onset, and correlate with the CAG mutant severity.

The excitotoxicity hypothesis links the progressive cell atrophy seen in the striatum, with the increase in NMDA receptor sensitivity. Excitotoxicity is the pathological process by which nerve cells are damaged or killed by excessive stimulation by neurotransmitters such as glutamate. This occurs when glutamate receptors, such as the NMDA and AMPA receptors, are over activated by glutamatergic release. NMDA is an excitotoxin which binds to these receptors, which can cause excitotoxicity by allowing high levels of calcium ions (Ca2+) to enter the cell. Ca2+ influx into cells activates a number of enzymes, which go on to damage cell structures such as components of the cytoskeleton, membrane, and DNA.

The excitotoxicity hypothesis suggests that increased firing rates over long periods, in NMDA expressing cells such as the SPN, cause cells to atrophy prematurely. Evidence supports a role for neuronal damage arising as a result of excessive activation of glutamate receptors, up-regulated by excitatory amino acids in the pathogenesis of HD (Fan and Raymond, 2007). It has been shown that neurons expressing high levels of NMDA receptors are lost early from the HD afflicted striatum. Further to NMDA, studies have hypothesised a link between dopamine in the striatum, and progressive cell death (Jakel and Maragos, 2000).

The disproportionate cell atrophy seen in the striatum during HD progression, combined with the progressive behavioural control issues associated with the HD phenotype, indicate a clear effect on action selection, due to disruption in the BG decision circuits.

2.7.2 Huntington's disease and action selection

Studies of mouse models of HD have shown changes in the spiking behaviour of SPN neurons (Miller et al., 2010, 2012). Results indicate an important role in both bursting and correlated burst firing for information processing in SPNs, and that dysregulation of the SPN neural coding scheme, moreover, is a key component of HD pathophysiology regardless of the severity of HD symptoms.

There is considerable evidence that changes in striatal SPNs differentially effect the behaviour of D1 and D2 expressing neurons (Andre et al., 2011). As D1 and D2 expressing neurons may have distinct roles in the BG (Kravitz et al., 2012), changes in the fine balance of activation and suppression attributed to SPN populated control pathways could be associated with the HD behavioural phenotype. Andre et al. (2011) determined that in mice models of HD, differential D1 and D2 expression in different pathways manifested in behavioural change. They concluded that increased stereotypic behaviours in early stage HD could be explained by increased glutamate activity and DA tone in direct pathway neurons, whereas hypokinesia at later stages could result from reduced input onto these neurons.

Interestingly, experiments with pre-clinical HD patients have shown that increased activation in thalamocortical pathways during motor learning tasks can compensate for cell loss in the caudate in pre-symptomatic HD patients (Feigin et al., 2006). Although it is remarked that these mechanisms would not be sufficient to compensate for the HD phenotype, they can play a role in slowing the apparent severity during the progressive onset. These interesting results demonstrate the importance of network dynamics on maintaining normal function during abnormal development. Further, studies show that preclinical compensation goes beyond a simple shift of activity from pre-motor to parietal regions, but involving multiple compensatory mechanisms in executive and cognitive motor areas (Kloppel et al., 2009).

Regardless of temporary compensation methods, cognitive functions are generally thought to deteriorate globally in late stages of various neuro-degenerative disorders. However, building on the compensatory mechanisms shown in preclinical HD trials, Beste et al. (2008) discovered a surprising example of behavioural improvement for symptomatic HD patients over pre-symptomatic patients. Using an auditory signal detection task, Beste et al. were able to demonstrate that pre-symptomatic HD patients and healthy controls performed worse compared with symptomatic HD patients. Behavioural, event-related potentials and were recorded during task performance, and show that both behavioural performance and neurophysiological correlates of auditory sensory memory and attentional reorientation can occur in late stages of HD.

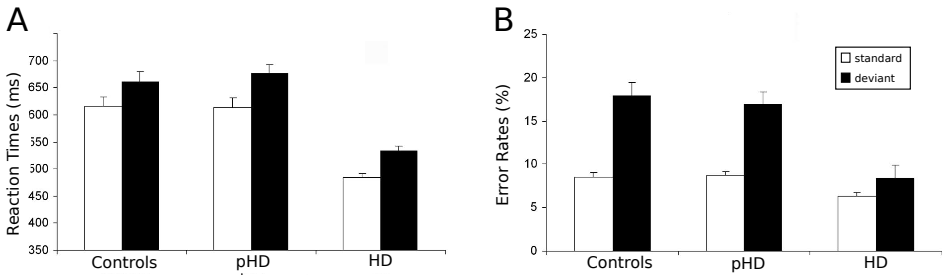

Fig. 2.14 **Reaction times in Huntington's Disease patients.** Behavioural Data for pHD,HD and healthy cases in an auditory selection test. **(A)** Mean reaction time(error bars indicate SEM) for the standard and deviant stimuli. **(B)** Mean error rate(error bars indicate SEM) for the standard and deviant stimuli. Reproduced with permission from Beste et al. (2008)

Interestingly, these novel results shine some doubt onto the optimality of decision making in the basal ganglia. While Bogacz and Gurney (2007) suggest that the basal ganglia perform asymptotically optimal decision making, these results suggest that changes in the striatum can improve selection in the basal ganglia in specific tasks. I will discuss this possibility further in the discussion chapter, and detail how this result need not contradict decision optimality of the basal ganglia.

Figure 2.14 summarises the magnitude of the results in Beste et al. (2008). Figure 2.14A shows the reaction times of healthy, pre-symptomatic and symptomatic HD patients, showing a statistically significant improvement in the reaction time of symptomatic HD patients. Similarly, figure 2.14B shows a reduction in error rates for HD patients. It its worth noting that the difference between the healthy and pre-symptomatic HD cases is not significant, which rules out the selection-effect being due to the mutated huntingtin gene. Rather, the effect only becomes apparent during late stage HD phenotype expression, which is characterised by extensive cell atrophy in the striatum.

Beste et al. (2008) hypothesize that increased NMDA activity recorded in HD patients (Fan and Raymond, 2007; Milnerwood et al., 2010) might facilitate signal selection at the striatal level by altering winner-take-all population dynamics performed by the striatal microcircuit, and basal ganglia as a whole. These results provide a powerful and instructive counter-example to the view that late stage neuro-degeneration is necessarily related to a global decline in cognitive abilities in HD.

In this thesis we aim to investigate the active role of the striatum in action selection, and challenge complex biophysical models of the striatum to provide a mechanistic explanation of the improved selectivity recorded by Beste et al. Novel results in chapters 5-6 show how

selectively enhanced cognitive functioning could emerge simultaneously with otherwise impaired cognitive functioning under phenomenological models of Huntington's disease.

Chapter 3

Construction of Striatum Models

3.1 Constructing the basal ganglia cortico-thalamus model

To study the contribution of striatal dynamics to the selection mechanism of the whole basal ganglia, we used a population-level implementation of a basal-ganglia thalamo-cortical loop model (Humphries and Gurney, 2002). Figure 3.1 schematically illustrates the loop model, and the connectivity of the response-representing populations.

Each population activity evolves with a single common model in equation 3.1, which is expanded in complexity by the specification of each populations specific input specified in equations 3.4 - 3.18. Each specific population is defined by the evolving population activity a, the input equation I, and the output firing equation y. The average activity a of all neurons comprising a channel's population changes according to

$$\tau \frac{da}{dt} = -a(t) + I(t) \tag{3.1}$$

where τ is a time constant and I is summed, weighted input. We used $\tau = 10$ ms throughout. The normalised firing rate y of the unit is given by a piecewise linear output function

$$y(t) = F(a(t), \theta) = \begin{cases} 0 & \text{if } a(t) \leq \theta \\ a(t) - \theta & \text{if } \theta < a(t) < 1 - \theta \\ 1 & \text{if } a(t) \geq 1 - \theta \end{cases} \tag{3.2}$$

with threshold θ.

The following describes net input I_i and output y_i for the i^{th} channel of each structure, with n channels in total. The full model was thus given as:

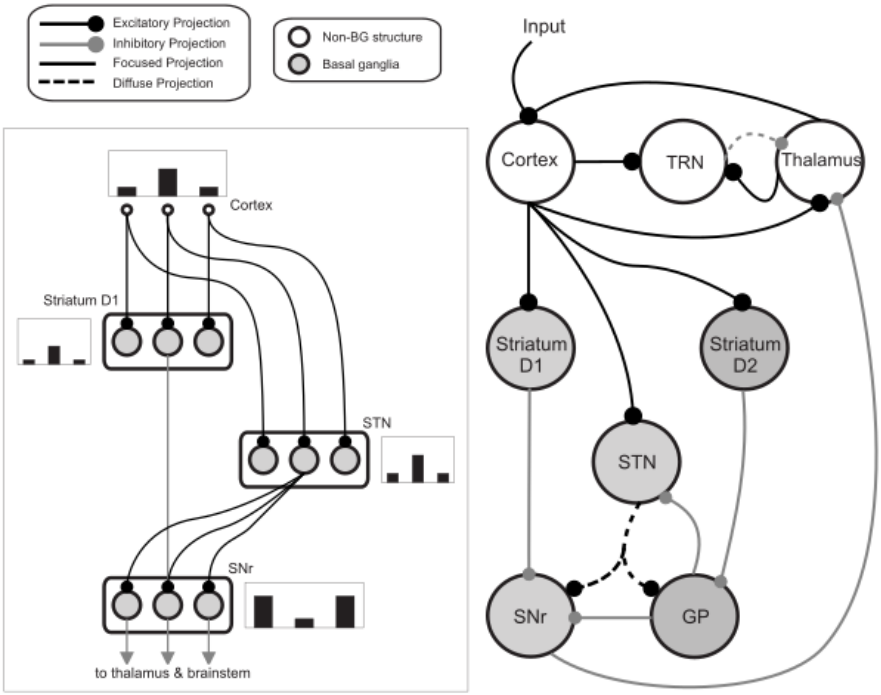

Fig. 3.1 **Basal ganglia thalamo-cortical loop model**. The main circuit (right) embeds the basal ganglia into a thalamo-cortical feedback loop. Each nucleus contains multiple response-representing populations. Within the basal ganglia, the circuit can decomposed into an off-centre, on-surround network (left): three populations are shown, with example activity levels in the bar charts to illustrate the relative contributions of the nuclei. Note that, for clarity, full connectivity is only shown for the second population. Briefly, the selection mechanism works as follows. Constant inhibitory output from substantia nigra pars reticulata (SNr) provides an 'off' signal to its widespread targets in the thalamus and brainstem. Cortical inputs, representing competing salience, are organised in separate populations, which project to corresponding populations in striatum and subthalamic nucleus (STN). The balance of focussed inhibition from striatum and diffuse excitation from STN results in the most salient input suppressing the inhibitory output from the corresponding SNr population, signalling 'on' to that SNr population's targets. Tonic dopamine levels in the striatum set the ease with which the channels are selected, and subsequently switched between following further salient inputs. For quantitative demonstrations of this model see (Gurney et al., 2001; Humphries and Gurney, 2002). GP: globus pallidus; SNr: substantia nigra pars reticulata; STN: subthalamic nucleus; TRN: thalamic reticular nucleus.

$$\text{Cortex}: \quad I_i^{ctx} = y_i^{thal} + c_i, \tag{3.3}$$

$$y_i^{ctx} = F(a_i^{ctx}, 0), \tag{3.4}$$

$$\text{Thalamus}: \quad I_i^{thal} = y_i^{ctx} - y_i^{SNr} - 0.1 y_i^{TRN} - 0.7 \sum_{j \neq i}^{n} y_j^{TRN}, \tag{3.5}$$

$$y_i^{ctx} = F(a_i^{thal}, 0), \tag{3.6}$$

$$\text{TRN}: \quad I_i^{TRN} = y_i^{thal} + y_i^{ctx}, \tag{3.7}$$

$$y_i^{TRN} = F(a_i^{TRN}, 0), \tag{3.8}$$

$$\text{Striatum D1}: \quad I_i^{d1} = y_i^{ctx}(1 + \lambda_1), \tag{3.9}$$

$$y_i^{d1} = F(a_i^{d1}, 0.2), \tag{3.10}$$

$$\text{Striatum D2}: \quad I_i^{d2} = y_i^{ctx}(1 - \lambda_2), \tag{3.11}$$

$$y_i^{d2} = F(a_i^{d2}, 0.2), \tag{3.12}$$

$$\text{Subthalamic nucleus}: \quad I_i^{stn} = y_i^{ctx} - y_i^{gp}, \tag{3.13}$$

$$y_i^{stn} = F(a_i^{stn}, -0.25), \tag{3.14}$$

$$\text{Globus pallidus}: \quad I_i^{gp} = 0.9 \sum_{j}^{n} y_j^{stn} - y_i^{d2} \tag{3.15}$$

$$y_i^{gp} = F(a_i^{gp}, -0.2), \tag{3.16}$$

$$\text{SNr}: \quad I_i^{snr} = 0.9 \sum_{j}^{n} y_j^{stn} - y_i^{d1} - 0.3 y_i^{gp}, \tag{3.17}$$

$$y_i^{snr} = F(a_i^{snr}, -0.2), \tag{3.18}$$

Net input was computed from the outputs of the other structures: the thalamus, (thal), the the Thalamic reticular nucleus, (TRN), the Globus Pallidus (gp),Striatal D1 neurons (d1),Striatal D2 neurons (d2), subthalamic nucleus (stn), Substantia nigra reticular (SNr), except driving input c_i to channel i of cortex (ctx). The striatum was divided into two populations, one of projection neurons with the D1-type dopamine receptor, and one of projection neurons with the D2-type dopamine receptor. Many converging lines of evidence from electro-physiological and anatomical studies support this functional split into D1- and D2-dominant projection neurons and, further, that the D1-dominant neurons project to SNr, and the D2- dominant neurons project to GP (Gerfen et al., 1990; Matamales et al., 2009; Surmeier et al., 2007).

In line with the projection neuron model described above, the model included opposite effects of activating D1 and D2 receptors on striatal projection neuron activity: D1 activation facilitated cortical efficacy at the input, while D2 activation attenuated this efficacy (Humphries et al., 2009b; Moyer et al., 2007). The mechanism for this mirrored that of the spiking projection neuron model in using simple linear factors. Thus, if the relative activation of D1 and D2 receptors by tonic dopamine are $\lambda_1, \lambda_2 \in [0,1]$, then the increase in efficacy due to D1 receptor activation was given in equation 3.10 by $(1+\lambda_1)$; the decrease in efficacy due to D2 receptor activation was given in equation 3.12 by $(1-\lambda_2)$. Throughout we set $\lambda_1 = \lambda_2 = 0.2$, simulating tonic levels of dopamine.

The negative thresholds ensured that STN, GP, and SNr have spontaneous tonic output (Humphries et al., 2006). We simplified the model here compared to Humphries and Gurney (2002) by delivering input only to cortex, to represent the salience-driven response selection, rather than to cortex, striatum and STN; both models gave qualitatively the same results. We used exponential Euler to numerically solve this system, with a time-step of 1 ms.

We used $n = 8$ channels in total, with two of those channels (4 and 5) receiving non-zero inputs, mimicking the input protocol used for the striatal network model, which is designed to abstractly simulate the two choice reaction-time task performed in Beste et al. (2008). Baseline inputs $c_4 = c_5 = 0.3$ were delivered at simulation onset. A step in input c_5 occurred between 100 and 200 time-steps: a small step of $c_5 = 0.5$ or a large step of $c_5 = 0.7$. The ability for the model to select was assessed during this step period. As in prior models (Berns and Sejnowski, 1998; Gurney et al., 2001; Humphries and Gurney, 2002; Humphries et al., 2006), selection was assessed by observing the change in activity on each SNr channel, as this output provides the tonic inhibition of thalamic and brainstem structures and is thought to gate the execution of actions (Redgrave et al., 1999). Here, successful selection of a channel was defined as the SNr output falling to zero.

3.2 Constructing the striatal microcircuit

A particular unknown is the computational role of the basal ganglia's input nucleus, the striatum. The striatal GABAergic projection neurons comprise the vast majority of cells and are connected by local collaterals of their axons (Wilson and Groves, 1980). The lack of layers or of clear axial preferences in the direction of dendrites or axons suggests that striatal tissue is homogeneous in all three dimensions (Humphries et al., 2010). Such GABAergic connectivity naturally lends itself to the idea that the striatum forms a vast recurrent network that, locally, implements a winner-takes-all computation (Alexander and Wickens, 1993;

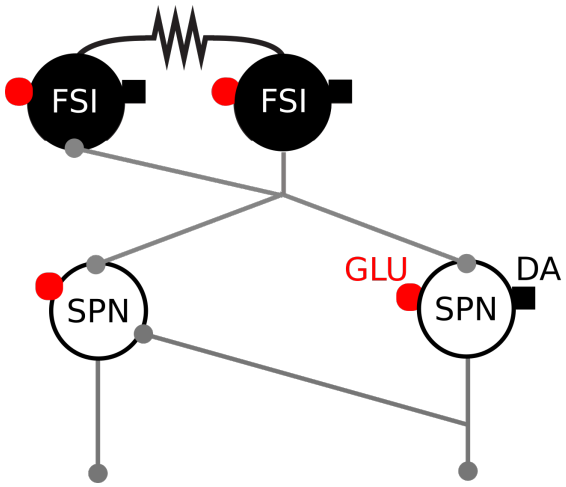

Fig. 3.2 **GABAergic striatal microcircuit**. Input to the striatum comes from glutamatergic (GLU: red circle) fibres originating in the cortex, thalamus, hippocampal formation and amygdala, and dopaminergic (DA: black square) fibres from brainstem dopaminergic neurons. The projection neurons (SPNs) are interconnected via local collaterals of their axons projecting to other nuclei of the basal ganglia. The fast-spiking inter-neurons (FSIs) can form dendro-dendritic gap junctions (black, excitatory) between them and are also connected by standard axo-dendritic synapses. All these intra-striatal axo-dendritic connections (grey) are GABAergic and hence inhibitory Each neuron type is represented with a different fill colour, FSI-Black and SPN-White.

Fukai and Tanaka, 1997; Wickens, 1997). The weak strength of synapses between the projection neurons (Czubayko and Plenz, 2002; Jaeger et al., 1994; Tunstall et al., 2002) is difficult to reconcile with this proposal (Plenz, 2003), as they suggest projection neuron output can only modulate ongoing activity and not outright inhibit their targets.

I study here an updated version of our prior full-scale model of striatum (Humphries et al., 2009b, 2010). Compared to those models, the model here brings together the three-dimensional anatomy model from Humphries et al. (2010) with an updated version of the dopamine-modulated projection neuron model from Humphries et al. (2009b), and a novel saturating synapse model introduced in Tomkins et al. (2012).

3.2.1 Spiking neuron models in the striatum

The basic model neuron used in the large scale striatal model is derived from the model neuron proposed in (Izhikevich, 2003), which was extended to encompass the effects of dopamine modulation on intrinsic ion channels and synaptic input in (Humphries et al., 2009b).

In the biophysical form of the Izhikevich model neuron, v is the membrane potential and the "recovery variable" u is the contribution of the neuron class's dominant ion channel:

$$C\frac{dv}{dt} = k(v - v_r)(v - v_t) - u + I \tag{3.19}$$

$$\frac{du}{dt} = a[b(v - v_r) - u] \tag{3.20}$$

with reset condition

$$\text{if } v > v_{peak} \text{ then } v \leftarrow c,\ u \leftarrow u + d$$

where in the equation for the membrane potential (3.19), C is capacitance, v_r and v_t are the resting and threshold potentials, I is a current source, and c is the reset potential. Parameter a is a time constant governing the time scale of the recovery due to the dominant ion channel. Parameters k and b are derived from the I-V curve of the target neuron behaviour, where b describes how sensitive the recovery variable u is to fluctuations in the membrane potential v. Parameter d describes the after spike reset of recovery variable u, and can be tuned to modify the rate of spiking output.

Modelling the striatum projection neurons

The projection neuron model parameter values and their source are given in Table 3.1. Parameters C, d, v_t, and the AMPA synaptic conductance g_{ampa} (see below) were found by searching for the best-fit to the f-I curve and spiking input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model.
Humphries et al. (2009b) show how this model can capture key dynamical phenomena of the projection neuron: the slow-rise to first spike following current injection; paired-pulse facilitation lasting hundreds of milliseconds; and bimodal membrane behaviour emulating up- and down-state activity under anaesthesia and in stimulated slice preparations.
Updating on the work done in Humphries et al. (2009b), I retuned the model neurons, to compensate for the addition of the new model of synaptic saturation, to again approximate the behaviour of the 189-compartment Hodgkin-Huxley based Medium Spiny Neuron model of Moyer et al. (2007). Adding synaptic saturation to the underlying neuron model required retuning the parameters g_{ampa} for the base model, without dopamine, $\beta 1$ for D1 expression and $\beta 2$ for D2 expression, in order to realign the saturating neuron behaviour with that of the multi-compartmental SPN model, with respect to both unmodulated and dopamine modulated synaptic input.

Using the tuned SPN model from (Humphries et al., 2009a), I adopted the F-I tuning variables without change, as with respect to constant current injection, the model response has not been changed. Building from this base, I must further tune the model parameters regarding the effects of input-output frequency curves, as supplied in (Moyer et al., 2007).

This breaks down into three separate stages, tuning first g_{ampa}, the maximal conductance of the AMPA channel, from which the conductances of the NMDA and GABA channels are derived, leading onto tuning $\beta 1$ and $\beta 2$ fitting the D1 and D2 modulated synapses respectively to the modulated F-F curves. The dopamine level $\phi 1$, $\phi 2$ was set to 0.8, assuming near maximal end equal receptor activations. We do not use maximal receptor activation as it is not clear how easily high concentrations of dopamine or dopamine receptor antagonists diffuse into the synaptic cleft after being introduced into the artificial cerebro-spinal fluid of the preparation (Moyer et al., 2007).

The tuning process I employ requires minimising the difference between the firing frequencies of the multi-compartmental Moyer model neuron and the Lumped-Saturating model neuron, by measuring the firing frequencies over a series of overlapping time windows, against the firing frequencies provided by Moyer et al. (2007) for a given set of input frequencies. Multiple time windows allowed us to smooth out the any first-firing rate anomalies, focusing on the stable firing rates.

The proposed normalised error of a given time window E_w is given by:

$$E_w = \sum_f |\frac{N_w^f - M_f}{M_f}| \tag{3.21}$$

over the range of input frequencies f as used in the Moyer experiment, where w is the index of the window, defining the starting time of the individual window at w seconds. N_w denotes the output firing frequency of the model neuron given an input of fHz, the time window w, M_f denotes the output firing frequency of the multi compartmental tuned Moyer model for the given input of fHz. Tuning parameters to minimize E_w can provide a good fit of the model neuron to the multi compartmental model firing frequencies, but can lead to an over fitting to the specific behaviour in a given time window, causing the neuron output to drift as a result of the varying time taken for a neuron behaviour to stabilise. I.e. tuning to reproduce exact first firing times, will introduce a difference in later firing rates under a specific input.

The tuning found that reducing the error E_w over a single time window can cause the behaviour of the neuron to drift over time away from the desired firing rates of the Moyer data, therefore we suggest that the total error of individual overlapping windows over a

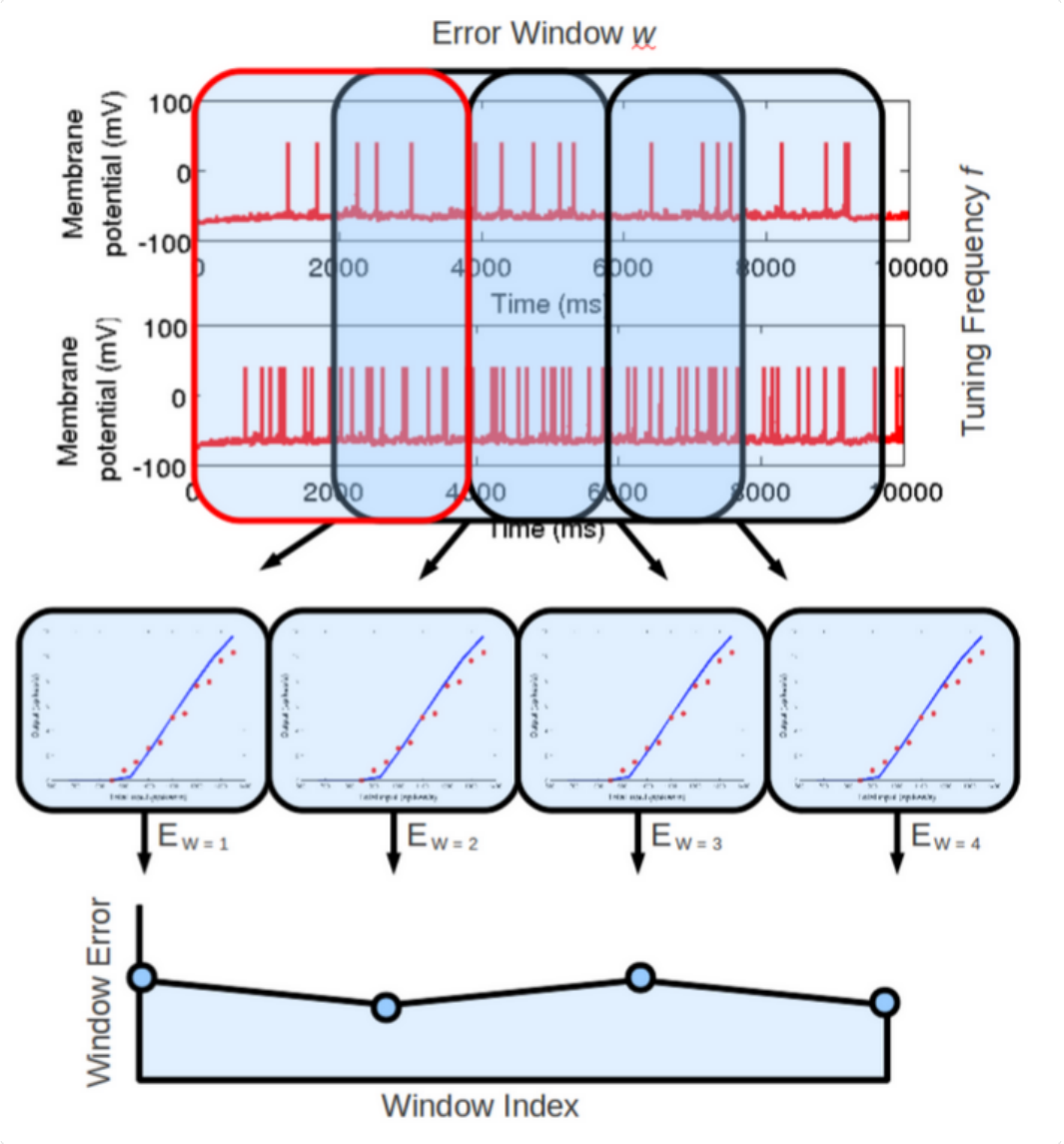

Fig. 3.3 **Tuning output rates of the reduced SPN model**. Calculating the error resulting from a set of tuning parameters. The red window represents the first of the sliding windows beginning at t = 0. 4s. Windows are used, each lasting 4 seconds, starting 1 second after the previous window. For each window, the error between this and the model SPN of Moyer et al. (2007) is calculated as shown. The total error of the parameter set over the supplied frequencies is then calculated as the average of the window errors.

longer run time is minimised to find parameters that have a low and stable error throughout prolonged neuron firing.
In order to reduce this over fitting, the above error function E_w is applied to an array of n of overlapping time windows over an extended time, starting at time t = 0, with the aim of minimising the total error over the sum of the individual window errors, creating a tuning that more accurately reproduce the Moyer fittings at any post-stable time in the neuron firing. Normalising this error with respect to the number of time windows produces the final error measure:

$$E_{total} = \sum_{n=1}^{N} \frac{E_w^n}{N} \tag{3.22}$$

The error calculation is illustrated in figure 3.3, showing how for a single set of parameters, the associated error can be calculated.
Due to the variable nature of the stochastic spike generator used, E_{total} was calculated repeatedly with different random spike generator seeds, with the parameter value producing the lowest average error selected as the most robust value.
Using the described tuning methods, the parameters for g_{ampa}, $\beta 1$ and $\beta 2$ where tuned to closely reproduce the F-F curves produced by the Moyer et al. multi-compartmental model, for the intrinsic Izhikevich neuron model with synaptic input, followed by building on this, with the dopamine modulated D1 and D2 receptive neuron models.

Modelling fast-spiking interneurons

For the FSI model, Equation 3.20 for the u term is given by (Izhikevich, 2007b)

$$\dot{u}_{fs} = \begin{cases} -au_{fs} & \text{if } v_{fs} < v_b, \\ a[b(v_{fs} - v_b)^3 - u_{fs}] & \text{if } v_{fs} \geq v_b, \end{cases} \tag{3.23}$$

which enables the FSI model to exhibit Type 2 dynamics, such as a non-linear step at the start of the current-frequency curve between 0 and 15-20 spikes/s, as seen in figure 3.4 B. Further discussion on the FSI model used in the striatal microcircuit can be found in (Humphries et al., 2009b); the FSI model parameters are reproduced in Table 3.2.

Modelling dopaminergic modulation of intrinsic ion channels

Tonic levels of dopamine in the striatum modulate the excitability of the projection neurons and fast-spiking inter-neurons (Mallet et al., 2006; Nicola et al., 2000). Our network model

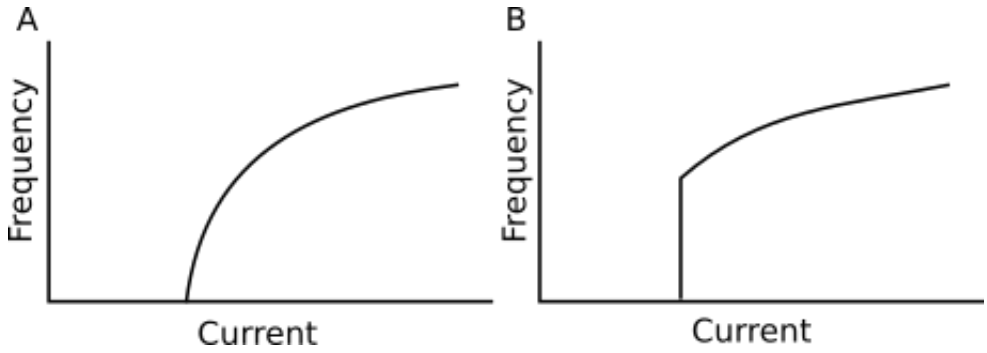

Fig. 3.4 **Type 1 and type 2 neuron dynamics F-I curve. (A)** Type one dynamics, with a gradual increase in firing rate in response to input frequency. The SPN exhibits type one dynamics **(B)** Type two dynamics, with a non-linear jump in firing rates. The FSI model exhibits type two dynamics

incorporates modulation by tonic dopamine through the relative activation levels of D1 and D2 receptors. These levels are modelled using the method proposed in (Humphries et al., 2009b), where complex membrane dynamics are subsumed by linear transforms with only two parameters ϕ_1, $\phi_2 \in [0,1]$, describing the proportion of D1 and D2 receptor activation, respectively. Throughout we used $\phi_1 = \phi_2 = 0.3$.

For activation of D1 receptors on projection neurons we used the linear mappings:

$$v_r \leftarrow v_r(1 + K\phi_1) \tag{3.24}$$

and

$$d \leftarrow d(1 - L\phi_1), \tag{3.25}$$

which respectively model the D1-receptor mediated enhancement of the inward-rectifying potassium current (KIR) (3.24) and enhancement of the L-type Ca^2+ current (3.25).

For activation of D2 receptors on projection neurons we used the linear mapping:

$$k \leftarrow k(1 - \alpha\phi_2) \tag{3.26}$$

which models the small inhibitory effect on the slow A-type potassium current, increasing the neuron's rheobase current (Moyer et al., 2007).

With these mappings, the model neuron is able to accurately capture the effect of D1 or D2 receptor activation on both the f-I curves and spiking input-output functions of the Moyer et al. (2007) compartmental model of the projection neuron.

Table 3.1 Intrinsic parameters for the projection model

Parameter	Value	Source
a	0.01	(Mahon, 2000);(Izhikevich, 2007b)
b	-20	(Izhikevich, 2007b)
c	-55 mV	"
k	1	"
v_r	-80 mV	"
v_{peak}	40 mV	"
C	15 pF	(Humphries et al., 2009b)
v_t	-30 mV	"
d	91	"
K	0.0289	"
L	0.331	"
α	0.032	"

Dopamine modulated fast spiking inter-neurons in the striatal network only express the D1-family of receptors (Centonze et al., 2003). Activation of this receptor depolarises the neuron's resting potential (see Humphries et al., 2009b for further details). Thus, we used the linear mapping of the resting potential:

$$v_r \leftarrow v_r(1-\eta\phi_1) \tag{3.27}$$

3.2.2 Modelling synaptic transmission

Synaptic input comprises the source of current I in Equation 3.19:

$$I = I_{ampa} + I_{gaba} + B(v)I_{nmda}. \tag{3.28}$$

where I_{ampa}, I_{gaba}, I_{nmda} are current input from AMPA, GABA and NMDA receptors respectively, and $B(v)$ is a term that models the voltage-dependent magnesium plug in the NMDA receptors. Compared to the projection neuron, FSIs receive no NMDA receptor input from cortex, have a moderately larger AMPA conductance (Table 2), but do receive input via local gap junctions (see below).

Each synaptic input type z (where z is one of ampa, nmda, gaba) is modelled by

$$I_z = \bar{g}_z h_z(E_z - v), \tag{3.29}$$

Table 3.2 Intrinsic parameters for the fast spiking inter-neuron model. Dimensions are given where applicable. See Humphries et al. (2009b) for details.

Parameter	Value	Source
a	0.2	Izhikevich (2007a)
b	0.025	"
d	0	"
k	1	"
v_{peak}	25 mV	"
v_{b}	-55 mV	"
C	80 pF	Tateno et al. (2004)
c	-60 mV	"
v_{r}	-70 mV	"
v_{t}	-50 mV	"
η	0.1	fitted to Bracci et al. (2002)
ε	0.625	fitted to Gorelova et al. (2002)

where $\bar{g}_z$ is the maximum conductance and E_z is the reversal potential. We use the standard single-exponential model of post-synaptic currents

$$\dot{h}_z = \frac{-h_z}{\tau_z}, \quad \textit{and} \qquad h_z(t) \leftarrow h_z(t) + S_z(t), \tag{3.30}$$

where τ_z is the appropriate synaptic time constant, and $S_z(t)$ is the number of pre-synaptic spikes arriving at all the neuron's receptors of type z at time t.

Given that one interest here is in the possible roles of striatal NMDA sensitivity in Huntington's disease, we paid careful attention to two complexities of the NMDA receptor: its non-linear voltage-gating, and its saturation. The term $B(v)$ in Equation 3.28, which models the voltage-dependent magnesium plug in the NMDA receptors, is given by (Jahr et al., 1990)

$$B(v) = \frac{1}{1 + \frac{[Mg^{2+}]_0}{3.57} exp\,(-0.062v)}, \tag{3.31}$$

where $[Mg^{2+}]_0$ is the equilibrium concentration of magnesium ions.

Lumped synaptic saturation

Previous large scale spiking neuron models of the Striatum (Humphries et al., 2009b, 2010) utilised a simple unbounded synaptic transition model, without considering the effects of synaptic saturation upon the cell input. Synaptic saturation provides additional constraints on the firing-behaviour and prevents neurons from unrealistic over excitation.

Table 3.3 Synaptic and gap junction parameters for the striatal network. See Humphries et al. (2009b) for details.

Parameter	Value	Source and notes
E_{ampa},E_{nmda}	0 mV	Moyer et al. (2007)
E_{gaba}	-60 mV	Moyer et al. (2007)
τ_{ampa}	6 ms	Moyer et al. (2007)
τ_{nmda}	160 ms	Moyer et al. (2007)
τ_{gaba}	4 ms	Moyer et al. (2007)
τ FSI gap	5	fitted to (Galarreta and Hestrin, 1999)
$[Mg^{2+}]_0$	1 mM	Jahr et al. (1990)
g_{ampa} Ctx-SPN	0.4 nS	tuning (see main text)
g_{ampa} Ctx-FSI	1 nS	Fits linear rise in EPSC data from (Gittis et al., 2010)
g_{nmda} Ctx-SPN	0.2 nS	Fixed by maintaining the 2:1 AMPA:NMDA ratio from (Moyer et al., 2007)
g_{gaba} SPN-SPN	0.75 nS	Koos et al. (2004)
g_{gaba} FSI-SPN	3.75 nS	mean 5-fold increase compared to SPN-SPN (Koos et al., 2004); 3$\times$ increase of PSP (Planert et al., 2010)
g_{gaba} FSI-FSI	1.1 nS	Gittis et al. (2010)
g FSI gap	5 nS	fitted to (Galarreta and Hestrin, 1999)
β_1	0.5	tuning (see main text)
β_2	0.3	tuning (see main text)

The fast receptors $GABA_A$ and *AMPA* have a short binding time of typically a few milliseconds, requiring relatively strong inputs to reach levels which would be noticeably affected by synaptic saturation. Alternatively the binding time of the NMDA receptor can last up to 200 milliseconds (Moyer et al., 2007), resulting in saturation effects being observed at low rates of pre-synaptic spiking. Modelling the effects of saturation provides a natural balance to the problem of unnatural over excitation.

Here I detail the lumped synapse saturating model introduced in Tomkins et al. (2012). The lumped synapse model, tracks a single conductance per synapse type, per neuron, instead of tracking one conductance per synapse, per neuron. It is a computationally simple, step and decay exponential synapse model modulated by a saturating coefficient (equation 6.7) with the benefit of considerable efficiency gains in running-time and memory usage for large scale simulations, over more mathematically and computationally complex models. The lumped saturating synapse model, as described by equations 3.28, and 3.34, represents a linear model that can be summed over the total number of synapses, giving the ability to keep track of the neuron conductance of GABA, NMDA and AMPA using only a single value for each,

hence referring to this as the 'lumped saturating synaptic model.' Further, to this, the model is analytically solvable.

As above, in the lumped synapse model, each synaptic input type z, where z represents one of either AMPA, GABA or NMDA, is expressed by:

$$I_z = \bar{g}_z h_z (E_z - v), \tag{3.32}$$

where $\bar{g}_z$ represents the maximal conductance for the input type z, and E_z is the reversal potential associated with that type. As discussed we use the standard single exponential model of post synaptic currents, and expand the number of incoming spikes to include a saturation term:

$$\dot{h}_z = \frac{-h_z}{\tau_z}, \quad \textit{and} \qquad h_z(t) \leftarrow h_z(t) + \left[1 - \frac{h_z(t)}{N_z}\right] S_z(t), \tag{3.33}$$

Where τ_z represents the appropriate time constant and $S_z(t)$ is the total number of pre-synaptic inputs of type z arriving at any of the neuron synapses of type z at time t. During simulations the value of $S_z(t)$ is computed using our spike event generator, depending on the number of neuron connections N_z and their mean firing rate r spikes/s. The synaptic saturation term applied here resembles the dynamics of the soft-bounded synapse detailed in Vasilaki et al. (2009a), scaling the change of the increase with respect to the maximum value of bound receptors.

Here we interpret h_z as the number of active synapses over the whole neuron, at each time step, the number of incoming spikes $S_z(t)$ activates that number of connections, which become progressively unbound at the rate τ_z. With this understanding we incorporate saturation into the lumped synapse model by bounding the size of the step increase in h_z by the proportion of available connections, governed by N_z, which is equivalent to number of individual connections to the model neuron.

It is worth noting that the proposed lumped saturating synaptic model reduces to the basic lumped synaptic model (Humphries et al., 2009b), if the saturation term is removed, as is effectively accomplished in the case that $N_z \gg h_z(t)$.

As glutamate can remain locked into the NMDA receptor for 100 milliseconds or more (Lester et al., 1990), so the pool of available receptors becomes rapidly saturated at high afferent firing rates. To capture this we introduce a mean-field model of synaptic saturation where we interpret the term h_z in Equation 3.30 as the number of active receptor groups over the whole neuron. Each step in h_{nmda}, following a number of spikes $S_{nmda}(t)$, activates that number of receptor groups, which decays with a time constant τ_{nmda}. To introduce saturation,

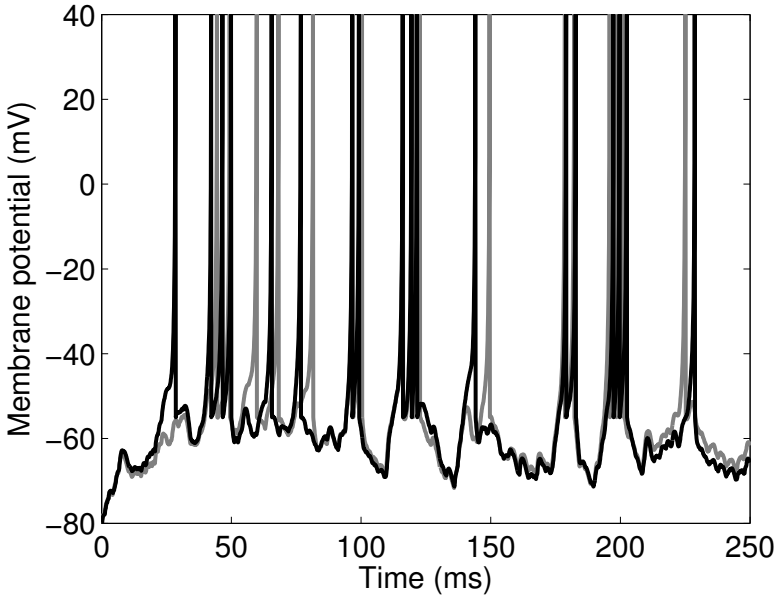

Fig. 3.5 **Equivalence of lumped- and individually-saturating synapse models.** Each model neuron is an Izhikevich neuron tuned to replicate the behaviour of a multi-compartmental spiny projection neuron. Each neuron is supplied with the same input spike train. (Black) Model neuron with our lumped saturating synapse model. (Gray) Model neuron with the kinetic model of synaptic saturation.

we bound the size of the step by the proportion of available groups. Together, these concepts give us the NMDA model:

$$\dot{h}_{nmda} = \frac{-h_{nmda}}{\tau_{nmda}}, \quad \textit{and} \qquad h_{nmda}(t) \leftarrow h_{nmda}(t) + \left[1 - \frac{h_{nmda}(t)}{N_{nmda}}\right] S_{nmda}(t). \tag{3.34}$$

As well as introducing this saturation of the NMDA synapses, we also removed the $1/\tau_s$ scaling of post-synaptic current amplitude used in (Humphries et al., 2009b). This allowed the model synaptic conductances to be the same order of magnitude as their experimental counterparts. Consequently, we re-tuned g_{ampa} by fitting the input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model, following the protocol in (Humphries et al., 2009b). We obtained equally good fits to those found previously with a value of $g_{ampa} = 0.4$. Figure 3.5 shows the approximately equivalent spiking behaviour of the lumped and individually saturating synapse when attached to an Izhikevich models, tuned to replicated the spiking behaviour of the 189-compartmental model of a SPN (Moyer et al., 2007).

Modelling dopaminergic modulation of synaptic transmission

Following the projection neuron models in Humphries et al. (2009b), we add D1 receptor modulation of NMDA receptor evoked EPSPs by

$$I_{nmda}^{D1} = I_{nmda}(1+\beta_1\phi_1), \tag{3.35}$$

and we add D2 receptor modulation of AMPA receptor evoked EPSPs by

$$I_{ampa}^{D2} = I_{ampa}(1-\beta_2\phi_2), \tag{3.36}$$

where β_1 and β_2 are scaling coefficients determining the relationship between dopamine receptor occupancy and the effect magnitude (Table 3.3). Due to the addition of saturating NMDA synapses, we also re-tuned these parameter values by fitting the input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model under D1 and D2 receptor modulation of synaptic inputs, following the protocol in Humphries et al. (2009b).

Finally, following the model in (Humphries et al., 2009b), we add D2 receptor modulation of GABAergic input to FSIs by

$$I_{gaba}^{fsi} = I_{gaba}(1-\varepsilon_2\phi_2). \tag{3.37}$$

Modelling gap junction connections

A gap junction between FSIs i and j is modelled as a compartment with voltage v_{ij}^*, which has dynamics

$$\tau\dot{v}_{ij}^* = (v_i - v_{ij}^*) + (v_j - v_{ij}^*), \tag{3.38}$$

where τ is a time constant for voltage decay, and v_i and v_j are the membrane potentials of the FSI pair. The current introduced by that cable to the FSI pair is then

$$I_{gap}^*(i) = g(v_{ij}^* - v_i) \qquad I_{gap}^*(j) = g(v_{ij}^* - v_j), \tag{3.39}$$

where g is the effective conductance of the gap junction. The total gap junction input I_{gap} to a FSI is then the sum over all contributions I_{gap}^*.

3.2.3 Creating the striatum microcircuit model

Our model captures the connections within the GABAergic microcircuit in striatum, illustrated in Figure 3.2. We simulated a large-scale model representing a three-dimensional cuboid of the striatum in the adult rat at one-to-one scale, containing every projection neuron and fast-spiking inter-neuron present in the biological tissue. We used a density of 89000 projection neurons per mm^3 (Oorschot, 1996) and a FSI density of 1% (see Humphries et al., 2010 for discussion). The paper assumed projection neurons were evenly split between D1 and D2 receptor dominant types, and without any spatial bias. Hence, we randomly assigned half of the projection neurons to be D1-type and half to be D2-type.

In the Results we predominantly report the results of simulations using a 300 μm on the cube side, giving 2292 projection neurons and 23 FSIs. Other sizes are noted explicitly where used.

Connectivity in the *physical* model

To connect the neurons we used two different models *The physical model* allows us to study striatal processing, which depends on the specific distance dependant connectivity seen in the microcircuit. As a spatial model, this connectivity mode suffers from edge effects, resulting in distant neurons receiving fewer connections than central neurons. We can study distance dependant effects at the centre, but results will be effected at the edges.

In the *physical* model we used distance-dependent functions for probability of connection between each element of the microcircuit. These functions were derived from overlap of dendritic and axonal arbors, and are given in (Humphries et al., 2010) for each connection type in the microcircuit.

First, we grow sample dendritic trees, branching dendrites according to a stochastic branching algorithm, which is bounded by the known properties of the dendritic trees, to ensure realistic morphology and branching density for a single dendritic tree. Second, each dendritic tree is expanded into a 3-dimensional space, by replacing dendritic sections with cylindrical volumes. Each cylinder connection reduced the diameter of the dendrite, taping with respect to the distance from the soma. By summing over the dendritic volume, we can obtain the total volume of dendrite, as a function of distance from the soma. Using many repetitions of the dendritic growth algorithm we determine a continuous function, to map the proportion of occupied dendritic space, with respect to the same distance (figure 3.6 D). Using the same techniques, we also develop volumetric density function for axons. We may

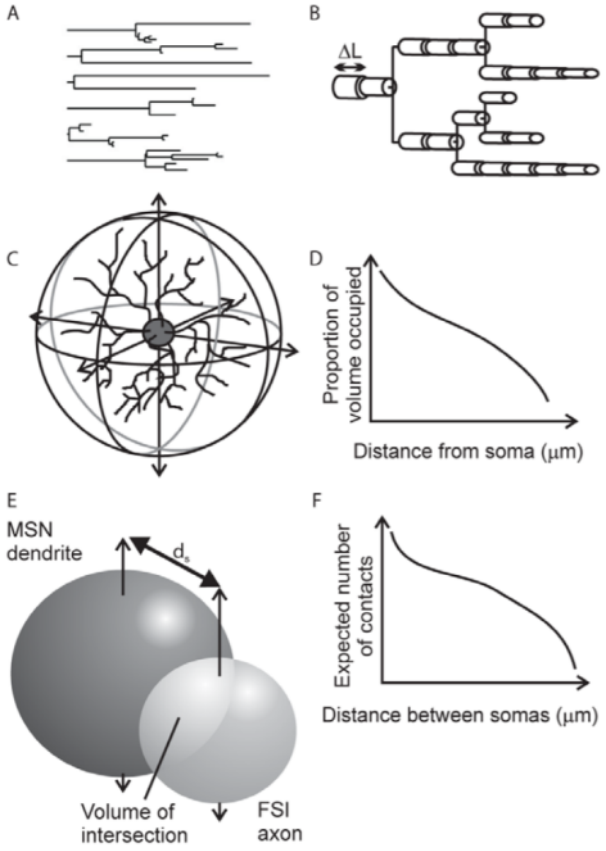

Fig. 3.6 **Striatum microcircuit model creation steps.** Reproduced from Humphries et al. (2010). Panels A–F show in order the steps involved in moving from a dendrite model to a probability function of contacts between two neuron types. **A** They create complete dendrograms using a stochastic algorithm, bounded by known properties of the dendrites. (A) shows all six dendritic trees of the complete dendrogram for a single SPN. **B** Each segment is modelled as a cylinder, with successive cylinder diameters tapering with distance from the soma. Summing over all branches gives the total volume of dendrite (or axon) at each distance from the soma. **C** They then compute the proportion of spherical volume occupied by dendrite (or axon) at each distance from the soma. **D** Expected values for occupied volume are computed over many repetitions of the growth algorithm resulting in a continuous function of volume occupancy. **E** We find the intersecting volume between the dendrite and axon spherical fields for each distance d_s between somas. **F** For each $1\mu m^3$ voxel, given its distance from the respective somas, we compute the probability of intersection between neurites (dendrite-axon or dendrite-dendrite) from the volume occupancy functions (in D). We then sum over all probabilities to get the expected number of intersections between neuron pairs as a function of distance between their somas. The resulting functions construct these networks. Reproduced from Humphries et al. (2009a).

then use the distance between pairs of neurons in the model, to calculate the probability of spatial contact, as a function of dendritic and axonal distance, as shown in figure 3.6 E.

Finally, For each $1\mu m^3$ voxel of space, we compute the probability of intersection between neurites (dendrite-axon or dendrite-dendrite) from the volume occupancy functions (in 3.6 D), given its distance from the respective somas. For the entire network, we must calculate the probability of connection between each plausible set of neurons, depending on their somatic-distance. We sum over all probabilities of collision to get an expected number of intersections between each of the possible neuron pairs, as a function of distance between their somas. Using the expected number of connections from each neuron pair, we can build a statistically bounded model of connectivity on the three dimensional striatum space.

Due to the size constraints in simulation, and the distance dependant connectivity function, only neurons in the centre of the simulation are guaranteed to have an anatomically accurate number of connections, as can be seen in figure 3.7 A. Spatial connectivity has been preserved over total connectivity to examine the effects of local circuit dynamics.

Connectivity in the *random* model

Alternatively, we study the *random model*, in which explicit distance dependence is removed, and each neuron is connected across the tissue, such that each neuron receives on average the same number of connections, as can be seen in figure 3.7 B. This removes edge effects, bust also any distance dependant computation. We use this model to study population phenomena which require high connection density, such as population inhibition, removing edge-effect sparsity.

In the *random* model we ignored distance, and simply made connections to each neuron at random until the correct number of incoming connections of each type was made. The target number of connections were derived from the mean values obtained from the central neurons of the three-dimensional connectivity model in (Humphries et al., 2010), and taken from column 1 of Table 5 in that paper: SPNs $\rightarrow$ 1 SPN: 728; FSIs $\rightarrow$ 1 SPN: 30.6; FSIs $\rightarrow$ 1 FSI: 12.8; FSI gap junctions per FSI: 0.65. In contrast to the *physical* model, spatial connectivity is sacrificed in order to maintain accurate connection counts for all neurons in the network.

We use a combination of both models to examine the striatal dynamics governed by both local and global connectivity, and show how increasing the size of the physical model approaches the dynamics of the random model, while keeping local circuity phenomena.

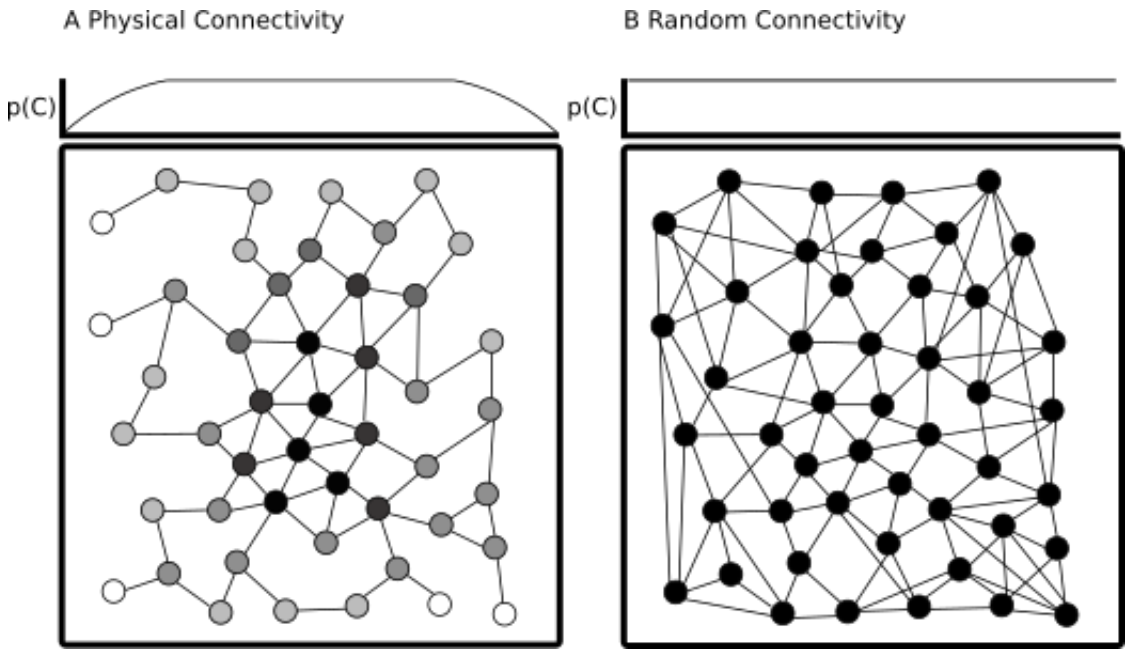

Fig. 3.7 **Probability of connection in the physical and random models.** A cartoon representation of connection in the striatal model. Each circle represents a neuron, with the colour representing the degree of connectivity that neuron recieves, with white low, and black high. **A** The physical model of connectivity, in which the neurons closer to the edge of the network are forced by the edge effects of the network, to have reduced connectivity. This causes a non-constant probability of connection (p(C)) across the network. **B** The random model ensures each neuron is statistical as highly-connected as any other, resulting in a constant P(C). Network distances are ignored in this regime, to remove edge effects.

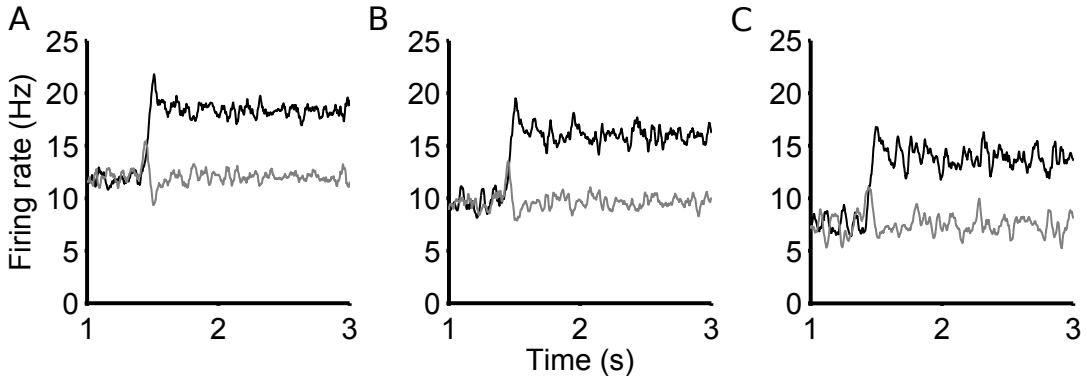

Fig. 3.8 **Increasing cell atrophy.** Cell atrophy is modelled by removing a percentage of neurons from the healthy network, by lesioning all connections and fixing atrophied membrane potentials. **(A)** A 'healthy' striatum model, with 0% cell atrophy and g_{nmda} : $g_{ampa} = 0.5$ ratio. **(B)** A 'degraded' striatum model with a 40% cell atrophy. **(C)** A 'heavily degraded' striatum model with a 70% cell atrophy. All traces are in response to a 'ramping decision protocol' with a base input of $5Hz$ and a salient signal of $6Hz$.

3.3 Phenomenological model of Huntington's disease in the striatum

Progressive cell atrophy

Huntington's disease disproportionally effects the spiny projection neurons on the striatum. The HD phenotype is associated with extensive cell atrophy in the striatum, primarily effecting the loss of the SPN population (Feigin et al., 1995; Portera-Cailliau et al., 1995). Cell atrophy occurs progressively throughout the life of a HD patient, but the exact trajectory of this cell atrophy is hard to study. We model the extensive cell death in the striatum using our simulated striatum circuitry. As with a healthy striatum, the network is constructed to reproduce the anatomical connectivity. We then randomly remove a percentage of the neurons, such that the volume of SPNs in the given simulated area is reduced by x percent. We cut all connections to the atrophied neurons, and hold their membrane potential fixed. By modelling the striatum under increasing levels of SPN atrophy, we can investigate the network level effects seen under HD-like progression. Initially we see an expected reduction in network excitability(Figure 3.8 A-C) under atrophy, as well as a reduction in transient phenomena. Cell atrophy has a wide impact on behavioural function in many neuro-degenerative and is key to understanding HD symptoms. A phenomenological model of cell atrophy allows us to

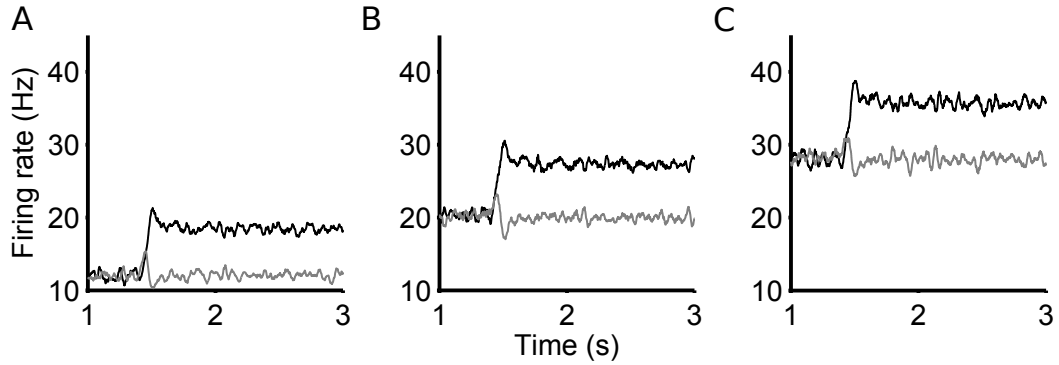

Fig. 3.9 **Increasing NMDA receptor sensitivity in a phenomenological model of HD degradation.** NMDA receptor sensitivity is modelled by up regulating the NMDA conductance, increasing cell excitability. Synaptic NMDA conductance is expressed in ratio to the AMPA conductance. **(A)** A 'healthy' striatum model, with 0% cell atrophy and $g_{nmda} : g_{ampa} = 0.5$ ratio. **(B)** An 'affected' striatum model with a 50% increase in NMDA receptor sensitivity, $g_{nmda} : g_{ampa} = 0.75$. **(C)** A 'heavily affected' striatum model with a 100% increase in NMDA receptor sensitivity, $g_{nmda} : g_{ampa} = 1$. All traces are in response to a 'ramping decision protocol' with a base input of $5Hz$ and a salient signal of $6Hz$.

untangle the changing complex network interactions and behavioural-modifications due to shifting cell-type ratios, not expressible in rate coded models.

Increased NMDA receptor sensitivity

Additionally, studies on the behaviour of spiny projection neurons in animal models of HD, have shown an increase in the NMDA receptors expression HD. This causes an increase in sensitivity to glutamate, which in turn raises the excitability of the cell (Milnerwood et al., 2010). This up-regulation of synaptic sensitivity has been proposed as the mechanism of improved selection seen in Beste et al. (2008).

We adjust the conductance of the fast-acting NMDA receptor on, $CTX \Rightarrow SPN$ connection, with respect to the fixed synaptic AMPA conductance, to model the HD associated modifications, increasing the excitability of model cells.

In the healthy model of the striatum, the sensitivity of the NMDA receptor is encoded as a ratio between the conductance of the AMPA and NMDA conductances, with $g_{nmda} : g_{ampa} = 0.5$ (Table 3.3). To phenomenologically represent the increasing NMDA sensitivity, we progressively increase the $g_{nmda} : g_{ampa} \Rightarrow 1$ ratio such that the NMDA conductance is raised, without effecting the synaptic AMPA conductance.

The combination of cell atrophy and NMDA sensitivity opens up a proposed 'landscape of degradation', of 'Atrophy-Excitation' possibilities. As it is not known exactly how a patient's striatum evolves in time with regard to increasing cell atrophy and NMDA sensitivity, we simulate each progression independently, and explore the possible trajectories a degrading striatum could progress through.

This two-factor atrophy-excitation phenomenological model of Huntington's Disease allows us to test the speculations of Beste et al. (2008). In their work they suggested that NMDA up-regulation would improve the 'sharpness' of selection behaviour, while demonstrating that NMDA up regulation seen in pHD patients, does not show improved selection. We can test if NMDA up-regulation is sufficient to improve selection in some cases, and investigate the mechanisms which may under pin increased selection only seen in late-stage HD patients.

Chapter 4

Experimental Protocol and Results Roadmap

This thesis is built upon the closely interlocked application of a variety of network models, connectivity regimes, and cortical stimulus design. To clarify the progression of the following results chapters, this chapter will provide a short summary of the procedures followed, and computational tools used in each chapter.

4.1 Network models used in the experimental protocols

In this work, we use two complementary network models. Primarily we use a spiking model of the striatal microcircuit, to study intrinsic striatal selectivity mechanisms. Secondarily, we also use a rate coded model of the basal ganglia-thalamo-cortical decision circuits, to embed the insights gained from the microcircuit model, into the larger decision networks in the brain. These networks are conceptually divided along the lines of complexity and physical size. The microcircuit model is a complex reproduction of the 3D striatal microcircuit (Humphries et al., 2010), simulating between 300 - 1000 μm^3 of virtual tissue. The basal ganglia-thalamo-cortical model is a firing rate population model of the basal ganglia circuitry which stretches across the brain, and contains a simplified model of the striatum (Humphries and Gurney, 2002). The basal ganglia-thalamo-cortical has been shown to exhibit the hallmarks of selection required for reliable, dynamic action selection in vertebrates.

Striatal microcircuit model

The striatal microcircuit model is a network of spiking neurons developed to study accurate spiking dynamics of the striatal microcircuit (Humphries et al., 2010). This model includes a variety of cell types (SPNs,FSIs), with explicit synaptic models, including gap junctions. The model is largely comprises two-variable spiking neurons (Izhikevich, 2003) which have been tuned to accurately reproduce the firing behaviour of multi-compartmental SPN models (Moyer et al., 2007), which receive spiking input from simulated cortical channels. This model is unique in its attempts to recreate the 3D GABAergic microcircuit, using accurate statistical models of connectivity.

The microcircuit model has been used to investigate selectivity in the striatal microcircuit (Tomkins et al., 2012, 2013), and is used extensively throughout this thesis. We utilized this model in chapters 5,6,7,8, as the main network to study competition in the striatum. We primarily use a striatal model representing a 300 μm^3 cube of simulated tissue. In chapter 7, we examine a larger network of 1 mm^3, to examine steady-state selection in large, physically connected models.

The striatal microcircuit can support two different connectivity motifs, both driven by experimental observations, which exhibit different spiking dynamics. We investigate both motifs, to better understand supported selection mechanisms in the striatum, and the conditions under which it can emerge. Here, I reiterate the technical differences between network configurations covered in the previous chapter, in order to illustrate they funder mental difference, and how each configuration is used in garnering different results. Ultimately we see that each configuration has its unique use case, adding to the general understanding of striatal function.

4.1.1 Network connectivity in the striatal microcircuit

The striatal microcircuit can be configured to use either a distance dependant *physical model* of connectivity, or a *random model* of connectivity. Each model reproduces different biological parameters, and allow for different compromises, with each displaying distinct spiking phenomena related to selectivity.

Physical connectivity

The physical model of connectivity enforces a strict distance-dependant connectivity motif, such that connection probability between projection neurons is determined by the distance between each soma. These probabilities are calculated by expanding 3 dimensional models

of dendritic trees, to calculate statistical collision rates between neural fields (Humphries et al., 2010).

Tomkins et al. (2012) exclusively uses the physical model of connectivity, to discover the transient phenomena which motivated further research into selectivity. Tomkins et al. (2013) further investigates the behaviour of the physically connected model to unravel the role transient selectivity measure. In this thesis, the physical model is primarily examined in chapters 5, 6, and 8, but also shows up in chapter 7 to examine the existence of steady-state selectivity in the physical model, under dense connectivity.

In 300 μm^3 models of the striatum, the physical model exhibits transient selection mechanisms. However, it is shown in chapter 7 that in 1 mm^3 models of the striatum, physical connectivity can support stead-state selectivity, due to an increased SPN-SPN connection density, as a result of increased volume. As we simulate a finite patch of the striatum, enforcing distance dependent connections results in connectivity edge-effects, such that neurons on the edge of the simulated striatum have reduced synaptic input from neighbouring projection neurons. Edge effects can be reduced by increasing the network size, at the cost of computational efficiency. For convenience, we can use a non-distance dependant model to remove edge effects, and keep small network sizes.

Random connectivity

The random connectivity model is used to remove the edge-effects seen in the distance dependant connectivity model. Random connectivity ensures that all neurons statistically receive the same number of synaptic inputs, determined by the average number of synapse that a neuron in the centre of the distance dependant model would receive. While this connectivity motif cannot support distance dependant computations, it does support computation that is dependant on uniform dense connectivity, such as continuous channel suppression.

The random model is studied in Tomkins et al. (2013), and is employed in this in chapters 7,8, to investigate steady-state selectivity in the striatum. Steady-state selectivity is analogous to winner-take-all selection mechanisms, and has been theoretically suggested to be supported by the striatal microcircuit, given dense inhibitory connectivity.

The random model is a useful computational simplification. We show that larger models of the physical model can support steady-state selection as seen in the random model, but at much greater computational cost, due to increased network size. The random model allows us to study steady-state selection at smaller network sizes, with theoretical support that the behaviour of the physical model approaches the random model at larger network sizes.

4.1.2 Cortical input stimulus to the striatal microcircuit

The striatum receives massive cortical input, and acts as a locus to combine and process motivational and situational information. To investigate striatal dynamics, we model spiking cortical input into competing, overlapping striatal populations.

The exact form of the cortical signals, is determined by our cortical input model. We developed 2 cortical input paradigms to investigate striatal processing. We focus on 2 populations of neurons in the striatum, and thus model two cortical input channels. We abstractly model a two-choice perceptual task, in which each channel represents a distinct, mutually exclusive perceptual choice.

Cortical step inputs to the striatal microcircuit

We first investigated selection in the striatal microcircuit in Tomkins et al. (2012), using cortical step inputs. The stepped cortical inputs required instant salience changes in the cortex, which could represent sudden stimulus onset. We found that this input exhibits transient selection and increased signal differentiability. These early results motivated the further investigations into selectivity in the striatum in Tomkins et al. (2013), with a focus on showing how these transients can improve action selection.

Step cortical inputs are only used in chapter 5, to motivate the investigation into transient phenomena in response to more biologically relevant cortical inputs.

Ramping cortical inputs to the striatal microcircuit

We expand on the step cortical inputs in Tomkins et al. (2013), by developing a ramped cortical input protocol. This protocol mimics neural correlates of evidence accumulation seen in the cortex, in response to simple decision tasks.

Initially each population receives an equal salient signal, analogous to cortical circuit priming. At a stimulus onset, both population salience rise. At some point in the evidence accumulation, the stimulus can be seen to favour one decision, and at this point, the least salient signal is reduced to baseline firing again. In the behavioural experiments in Beste et al. (2008), patients where required to judge the length of a tone to be either short (200ms) or long (400ms). This corresponds to our protocol, with the both salience rising at the onset of the tone, followed by a drop in the incorrect salience after 200ms, as the outcome become clear.

The ramped cortical input protocol is introduced for out main results, in chapter 6. This input stimulus design is used as the backbone to selectivity research, and further used in

chapter 7 to discuss stead-state selection, and chapter 8 to discuss transient and steady-state selection in HD conditions.

4.1.3 Basal ganglia-thalamo-cortical model

In contrast to the spiking striatal microcircuit model, we also employ a rate coded model of the basal ganglia-thalamo-cortical loops. Developed to investigate the action selection mechanisms of the basal ganglia, we use this model to support our hypothesis that signal selection in the striatum modifies action selection in the basal ganglia Humphries and Gurney (2002).

The striatal microcircuit model provides a perfect platform to discover innate striatal phenomena, yet we must show that selectivity in the striatum is effective in the greater decision circuits.

We use the rate coded model in Tomkins et al. (2013), to justify the computational power of transient selection, which is covered in chapter 6. The rate coded model does not intrinsically capture the transient section of the striatal microcircuit as it contains none of the crucial structure of the full striatum model. As such we must artificially induce transients in the striatum of the rate coded model, by injecting small positive and negative inputs, coinciding change in cortical signal salience

The rate coded model is supplied with a cortical step input, which approximates the cortical input provided to the full microcircuit model, to encode a two choice decision task in the basal ganglia.

4.2 Results Roadmap

Here I motivate the research progression, and emphasise the computational tools, as discussed above, which we deploy at each step in the process.

In chapter 5 we review the results of our first paper, Tomkins et al. (2012), which motivated further research into selection mechanisms in the striatal microcircuit. We detail how the *physical* model of the *striatal microcircuit* responds to *step cortical input*. We find that under these constrains, the striatal model demonstrates a transient phenomenon, in response to changing salience of cortical signals.

Building on these results, we show construct our hypothesis of selection in the striatum. We propose that the striatal microcircuit plays an active role in signal selection in the basal ganglia decision circuits. Other studies have suggested that the striatum may be capable

of winner-take-all style computation, which does not match our observations of transient-selection seen in sparsely connected networks.

We develop two metrics for measuring selection in the striatum, in order to determine if:

- transient selectivity can improve signal selection in the basal ganglia
- the striatal microcircuit can support traditional winner-take-all dynamics

In chapter 6 we expand on the role of transient selectivity found in Tomkins et al. (2012), in order to validate the hypothesis that transient selectivity could improve signal selection in the basal ganglia. We detail the response of a *physical* model of the *striatal microcircuit* to a more biologically realistic model of evidence accumulation using *ramped cortical input* signals.We demonstrate how transient selectivity remains in response to slow ramping inputs, and confirm that the transient phenomenon is a general response to competitive inputs, and is robust over a wide range in two-channel decision task salience.

We demonstrate how the transient phenomena is specific to the SPN neuron model, and is not reproduced in pyramidal neurons. This result suggests that transient selectivity is a striatum-specific statistical mechanism. We further show that the negative transient is emergent of network level competitive inhibition, using SPN- and FSI-afferent lesions. We use the *rate coded* model of the basal ganglia-thalamo-cortical decision circuit to show that transient selection can improve signal selection in the SNr. We show that transient selection can both reduce reaction times, and increases clean selection, validating our first hypothesis.

In chapter 7, we extend our understanding of selection in the striatum, and demonstrate that a *randomly connected* model of the *striatal microcircuit* will exhibit winner-take-all like behaviour in response to *ramping cortical input*. We discover that steady-state selectivity is effective the striatal network, given dense SPN-SPN connectivity, with SPN inhibition facilitating signal suppression. We show that dense connectivity will result in steady-state connectivity, and as such, we challenge the *physical model* of the striatum to exhibit steady-state selection. We find that in sufficiently large distance dependant network models, steady-state suppression is achieved.

Having demonstrated that the striatal microcircuit can support two complementary forms of selection, we sought to find where these mechanisms differ. We have shown that transient selectivity arises from internal membrane dynamics, whilst steady-selectivity depends on SPN-SPN connection density. Based on these distinct mechanistic causes, we hypothesised that each mechanism would evolve differentially in response to changes to the striatal network. In chapter 8, we expose our physical and random models of the striatum, to a phenomenological model of Huntington's disease.

We find that steady-state selectivity declines in all HD-related states, with increasing losses of selectivity correlating with increased disease severity. This is in line with the majority of behavioural studies in Huntington's disease. Interestingly, we find that transient selectivity is able to support areas of improved selectivity under Huntington's disease conditions. We suggest that transient selectivity could provide a neural mechanism for a paradoxical late-stage Huntington's disease compensatory mechanism discovered by Beste et al. (2008). Our model predicts that non-symptomatic HD patients would not see increased selection performance, and suggest that behavioural improvements will only manifest in late stage Huntington's disease, which is in line with the findings of Beste et al. (2008). We discuss the implications of these discoveries further, in chapter 9.

Chapter 5

The Striatal Microcircuit Actively Enhances Signal Selection

In previous studies of action selection mechanisms in the Basal Ganglia, the focus is often on the computational role of the complex network of nuclei within the basal ganglia. Previous efforts have focused on deriving the computational framework represented by the complex interaction of excitatory and inhibitory inter-nucleus pathways throughout the basal ganglia.

In our study of action selection in the basal ganglia, we focus on the Striatum, the largest individual nucleus of the basal ganglia, which integrates information from across the cortex. The basal ganglia have efferent connections to the substantia nigra pars reticular (SNr) and the globus pallidus (GP), serving as the main input nuclei to the basal ganglia formation as a whole. Having such a central role in the action selection circuits, as well as being the single largest nuclei in the basal ganglia, suggest that the role of the striatum may extend further than a mere informational relay towards the lower basal-ganglia. Unfortunately, the exact role of the striatum, and the complex striatal microcircuit within, is poorly understood with regards to the decision making process.

Recent efforts to demonstrate the computational anatomy of the basal ganglia, focus on a strict channel wise interpretation of information flow through the structure. The simplification of channel-wise information processing allows the task of action selection to be accurately reduced to the problem of channel-wise signal selection. While representing a useful simplification to understand the computational role of the BG connectivity, representing each channel as a single leaky integrator, necessarily misses information processing done at the individual neuron level.

The striatal microcircuit contains a plethora of cell types and modes of connection, which are abstracted away by the channel-wise mean-field modelling approach. In reality each

informational processing channel depends on the interaction of both excitatory and inhibitory cell populations, with distinct firing regimes and spatially distributed connectivity motives.

We use the detailed model of the striatal microcircuit described in chapter 3, to investigate the channel-wise interactions in the striatum. By expanding the definition of a channel in the striatum to incorporate the network of Spiny Projection Neurons and Fast spiking inter-neurons, we can study the aggregate channel behaviour as a direct consequence of spatially distributed, laterally interacting, collections of spiking neurons. Via simulations, we find that there are indeed signs of competitive signal processing in the complex striatal microcircuit that cannot arise organically in a mean field approximation of channel dynamics. This discovery shows an important role for the striatum itself, but may have far reaching effects in the entire basal ganglia.

5.1 Striatal microcircuits actively modify cortical signal representation.

In Tomkins et al. (2012), we investigate the role of the striatal microcircuit in signal selection. We show that the striatum does more than merely reflect signal salience through the BG circuit, but also actively modifies cortical representation. In our preliminary study, we use a step input signal to represent cortical input of a two-choice perceptual decision making task. In response to changing signal salience, a transient, active signal modification is elicited in the channels of the striatal microcircuit, as shown in figure 5.1. Here I discuss these early results which motivated the further investigation covered in Tomkins et al. (2013).

The striatum model was divided up into three populations (Figure 5.1A), two physically close SPN populations representing the two competing responses, which we refer to throughout as *channels*, and the remaining background neurons, which are given a constant input. Neurons were randomly divided into the two channels, with 40% of the neurons in channel 1 and 2 respectively, and the remaining 20% of cells were labelled “background" neurons.

Figure 5.1 B shows the smoothed spiking output of two competing striatal channels. After the divergence of the salient inputs. A transient over representation of the input can be seen around $t = 1500s$ on channel a, along with a transient suppression of the out-competed signal, b. The robust transient phenomena elicited by the microcircuit suggests that the striatum actively contributes to the decision making circuits of the basal ganglia, in response to competing inputs representing diverging salient signals.

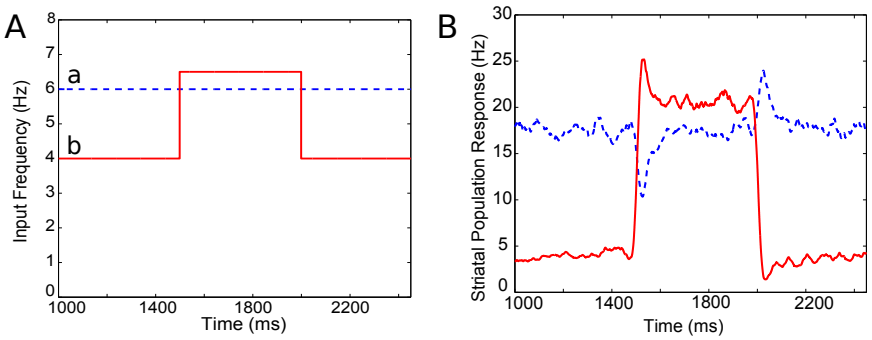

Fig. 5.1 **Selectivity in the striatal microcircuit (A)** Selection Protocol: The cortical input to each striatal population, driving signal A (*dashed*) and signal B (*solid*).**(B)** The healthy network response: Cell atrophy = 0%, NMDA:AMPA ratio = 0.5. Traces represent the mean firing rates of the striatal responses to the spiking cortical. The signal responses are divided up into three defined areas: Primed signal period $P = 1000$ to 1500 ms, before signal competition is introduced. Transient signal period $T = 1500$ to 1700 ms, where signal transients are induced. Stable Signal period $S = 1700$ to 1900 ms, corresponding loosely to a steady-state (no transients).

We propose that the step input presented in 5.1A could be used to model a sudden and immediate stimulus onset which requires a quick and clean action selection. We suggest that the elicited transient phenomena to step inputs is a mechanism to enforce quick action selection to sudden stimuli. These phenomena create a network-level urgency-related selection mechanism complementary to neuron-level fast selection mechanisms such as M-cells in fish.

Building on the success of establishing an active signal-processing role in the striatum (Tomkins et al., 2012), we further sought to investigate the role of the striatum in a wider context. Here we discuss expanding the cortical-signal representation to represent less artificial inputs, as well as investigate further mechanisms for signal processing in the basal ganglia-thalamo-cortical loops.

5.2 Striatal signal processing occurs in response to models of evidence accumulation

In Tomkins et al. (2013), cortical input to the model was designed to emulate the response selection component in a general two-choice task, where a (possibly noisy) stimulus taking one of two values is observed over time and a choice made between the two corresponding responses. In such a task, we propose that the two responses are made salient by the

onset of each trial; and after a perceptual decision is made about the stimulus value, then the corresponding response increases in salience. This generic set-up was inspired by the experimental procedures of Beste et al. (2008), in which participants were asked to distinguish between short (200 ms) and long (400 ms) auditory tones, using a distraction paradigm. Inputs followed a ramping trajectory to simulate evidence accumulation and increasing decision confidence (Asaad et al., 2000). Simple two-choice perceptual decision tasks have proven to be very instructive in unravelling the complexities of cortical decision making. We previously showed that transient selection can be seen in response to stepped cortical inputs which mimic a two choice decision task (Tomkins et al., 2012). Evidence accumulation in decision making has shown to have direct cortical correlates, and has motivated many behavioural decision making studies. In further research, we have investigated how evidence accumulation time effects this novel form of selection in response to stepped cortical inputs.

The ramping input protocol is illustrated in Figure 5.2A. Each response population received a priming input at a background rate for 1500 ms, causing them to reach a steady state of firing activity. At 1500 ms, channel 1, (black) received a ramping input for a time of 50ms, raising the salience towards a new steady state, the most salient cortical input to the striatum. During the 50 ms ramping time, channel 2 also received a ramping input, matching that of channel 1 for 25 ms. Following this, the signal to channel 2 decreased back to the background rate, describing the evidence accumulation trajectory of an out-competed action.

Rates were specified for each cortical spike train input to each projection neuron and FSI model. Both neuron models received the equivalent of 250 input spike trains (see Humphries et al., 2009b for details).

We measured how the striatal microcircuit performed channel wise signal selection on the cortical inputs, using this simple protocol, inspired by the auditory decision task performed in Beste et al. (2008). However due to the abstract nature of the input protocol we use, applied to a generic simulation of the striatal microcircuit, the selection measured in these results could be applied to any channel-wise decision task throughout the striatum, and is not limited to auditory processing.

5.3 Metrics for selection

We define “selectivity" in the striatum as the ability to robustly distinguish competing signals. Using two modes of biologically-derived connectivity (*physical* and *random*), we can show that the striatum demonstrates two complementary modes of selectivity, which we measure

with specific metrics. These selection metrics are applied to the output of each channel, which is characterised by a zero-phase filtered mean firing rate.

5.3.1 Transient selectivity

In the *physical* striatum model, we see a temporary boosting of the most-salient signal, accompanied with a temporary suppression of the least-salient competitive signal (Figure 5.2C). The physical model of the striatum enforces distance dependant connectivity statistics, ensuring that neurons in the centre of the simulation have biologically-accurate connectivity counts.

These transient phenomena present a boost of the difference in salience between the two competing signals. We identify two key regimes: (i) $\Delta^{S(1,2)}$, the maximum difference between the two signals during the transient peaks; (ii) S^1, S^2, the mean stable activity level of each channel after the transient period dissipates. Thus, the total transient selectivity, between 0 and 1, is defined as:

$$TS = 1 - \frac{S^1 - S^2}{\Delta^{S(1,2)}}, \qquad 0 \leq TS \leq 1 \tag{5.1}$$

$\Delta^{S(1,2)}$ is calculated as the *maximum* difference between the firing rates of Channel 1 and channel 2 over the transient window ($t = 1500 : 2000$ ms). This enables the measure to allow for cases in which the largest perturbations from the mean are not temporally coincident, either due to reliable intrinsic dynamic properties of the network, or statistical fluctuations therein. Transient selection approaches 0 in the case that $\Delta^{S(1,2)} \Rightarrow (S^1 - S^2)$, such that the maximum transient phenomena become indistinguishable from normal channel firing. Conversely, transient selection approaches 1 in the case that $\Delta^{S(1,2)} \gg (S^1 - S^2)$, where transients are increasing distinguished from equilibrium firing rates.

5.3.2 Steady-state selectivity

In the *random* striatum connectivity, the striatum network can exhibit signal suppression on its least-salient channel due to sustained inhibition by the most salient channel. In the random model of the striatum, distance dependence is removed from connectivity statistics, meaning that all neurons in the simulation receive the biologically accurate number of connections, but possible spatial processing is removed.

Steady-state selectivity is measured on the least-salient channel, as the percentage reduction in the mean channel firing rate after the rise in salience of the most-salient signal.

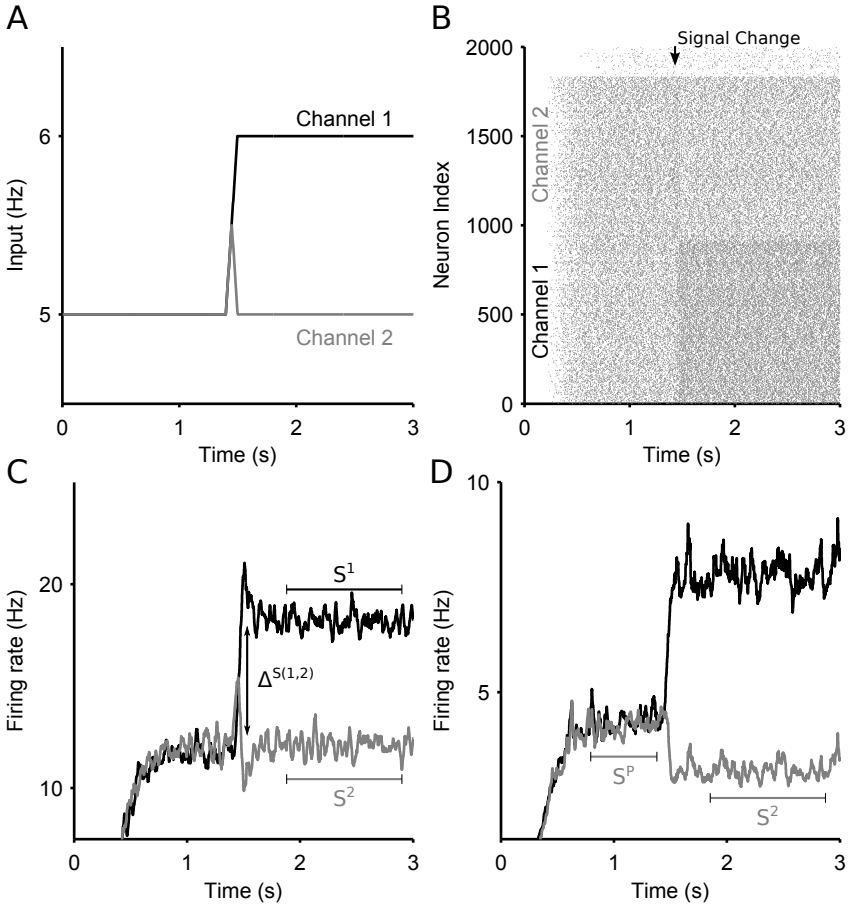

Fig. 5.2 **Measures of selectivity in striatal output**. **(A)**. Ramping cortical input into the striatum model. Two channels are driven by input spike trains, demonstrating signal selection between most-salient (channel 1) and least-salient (channel 2) striatal signals. **(B)**. Raster plot of the striatum microcircuit output for a single selection experiment. Increased firing can be seen in channel 1 at the onset of the ramped input in panel A. **(C)** A sample striatal output of the *physical* network, showing a zero-phase filter of the mean spiking output from the two competing channels in response to the ramped input in panel A. Annotations demonstrate the measures used in the transient selectivity measure. S^1,S^2: stable firing rate; $\Delta^{S(1,2)}$: maximum of the difference between the two channels firing rates over the transient period. **(D)** A sample striatal output of the *random* network, in response to the same input. Annotations demonstrate the measures used in the steady-state selectivity measure. S^P: pre-step stable firing rate.

An example of steady-state selectivity in the *random* network can be seen in Figure 5.2D. We define (S^P) as mean the stable firing rate of the primed channel 2 before the increase in competition, and from this we calculate the steady-state selectivity (SS) as:

$$SS = 1 - \frac{S^2}{S^P} \tag{5.2}$$

Steady-state selectivity measures the signal suppression, with respect to the primed firing rate. Steady-state selectivity tends to 0, as $S^2 \Rightarrow S^P$, and tends to 1 with $S^2 \ll S^P$

Using these metrics we can determine how these selection phenomena evolve with respect to changes in the striatal microcircuit model, such as size and degradation. We have demonstrated that the striatum exhibits selection, and can support two separate selectivity mechanisms. To validate that the striatum plays an active role is action selection in the BG loops, we must show that these phenomena are robust, and modify selection behaviour downstream, in the SNr. In chapter 6, we examine the underlying mechanisms of transient selectivity, and investigate the impact transient phenomenon have on action selection in the basal ganglia.

Chapter 6

Transient Selection can Improve Action Selection

We have shown in chapter 5 that the striatal microcircuit model supports two distinct selectivity mechanisms. We have proposed two selectivity measures, to quantify the selection phenomena (Tomkins et al., 2012). In order to further substantiate the hypothesis, that the striatum plays an active role in action selection in the basal ganglia, through innate selective signal dynamics, we must show that:

- Selectivity is robust to cortical input regimes
- Selectivity is amplified in the basal ganglia-thalamo-cortical loops
- Selectivity ultimately improves signal suppression in the SNr

In this chapter we investigate transient selectivity in a physically connected model of the striatal microcircuit, in response to ramping cortical input, to determine the underlying neural mechanism of transient selectivity. We further examine the impact of introducing signal transients into a rate coded model of basal ganglia-thalamo-cortical decisions circuits.

We sought insight into the potential for signal competition within the striatum by examining the dynamics of our three-dimensional network model. We first explored the effect on striatal output of competing inputs to two projection neuron populations. These inputs were intended to emulate the changes in cortical signals representing two alternative responses in a generic two-choice decision-making task.

6.1 Transient selection is effective in sparsely connected networks

We examine the selectivity mechanisms which emerge in a 300 μm^3 cube of physically-connected striatal tissue. We find that distance dependant connectivity in the striatal model implicitly supports the combination, short term signal suppression, and over expression we have termed transient selectivity. In response to ramping divergent cortical input (inset), Figure 6.1A shows the mean firing rate of each corresponding striatal population. After the divergence in inputs at $t = 1.5s$, a transient increase of the firing rate is elicited in channel 1 (black) and a transient suppression of the firing rate is elicited in channel 2. This transient suppression occurs despite no change in the input to channel 2. Moreover, channel 2 activity rapidly returns ($\sim$ 100 ms) to its pre-step firing rate. The result of transient selection in response to ramped cortical inputs are consistent with the response to step inputs investigated in Tomkins et al. (2012), and provide a generalisation of the applicability of the transient response to a wider range of decision task formulations.

Transient selection is robust

Having shown that transient selection is robust to spiking cortical input design, we must show that the transient response is present in a wide range of possible cortical inputs. Previously, we have shown the transient phenomena in response to a specific cortical input design with a background salience of 5Hz, and an increased salience of 6Hz, a ramped increase of 1Hz. In Figure 6.1B we examine an input space with baseline salience from 3.5Hz up to 7Hz, and step inputs from 0.5 up to 4Hz

We found that the elicited transient selection was robust over the parametric input space of input signal strength vs size of signal increase. We comprehensively simulated a large range of baseline input rates and step sizes, for the two channel inputs after the signal divergence. Figure 6.1B shows that transient selection could be robustly elicited for any step size over 0.5 Hz when the baseline input rate exceeded $\sim$ 4 Hz. In this terminology, figure 6.1A is run using a base rate of 5Hz and a step size, or signal difference, of 1Hz, which is within the threshold for stable transient phenomena.

We conclude, that transient selection will be employed in signal selection tasks with any suitably salient competitive signals.

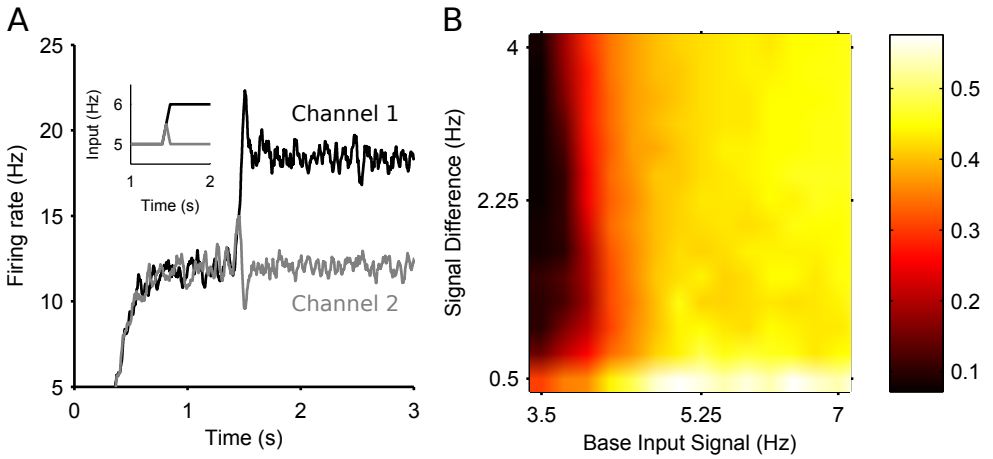

Fig. 6.1 **Transient selection of competing input signals by the striatum. (A)**. Mean firing rate of the two output channel populations in the experiment in response to the ramped input protocol (inset); individual spike trains have been convolved with a zero-phase digital filter to create smooth firing rates without lag. **(B)**. Mean transient selection landscape colour coded such that brighter colours represent higher selectivity. Landscape shows the mean transient selectivity averaged over 30 trials as a function of base input signal and step in signal difference during competition.

6.1.1 Transient selection is due to both circuit and intrinsic membrane properties

We further investigated the mechanisms underlying the positive and negative transient changes in population activity, as seen in Figure 6.1A. We found that the positive transient was produced by single neuron dynamics, whereas the negative transient was due to network connectivity. This can be seen in Figure 6.2A-B. Lesioning the SPN-SPN connections failed to remove the positive transient effects, suggesting the transient is a product of internal membrane dynamics. We demonstrate the mechanism in Figure 6.2C, and investigate the neural underpinning in figure 6.3. Conversely, we see that removing SPN-SPN connection removes the negative transient but did not prevent the positive transient, which suggests the negative transient is a product of local inhibitory SPN-SPN connectivity.

Positive transient is due to SPN dynamics

To confirm the positive transient was a single neuron phenomenon, we simulated an individual projection neuron model receiving many trials of the same stepped input protocol, and averaged its responses. The resulting peri-stimulus time histogram (Figure 6.2 C) shows that the neuron had a clear transient increase in em firing probability immediately after the step of

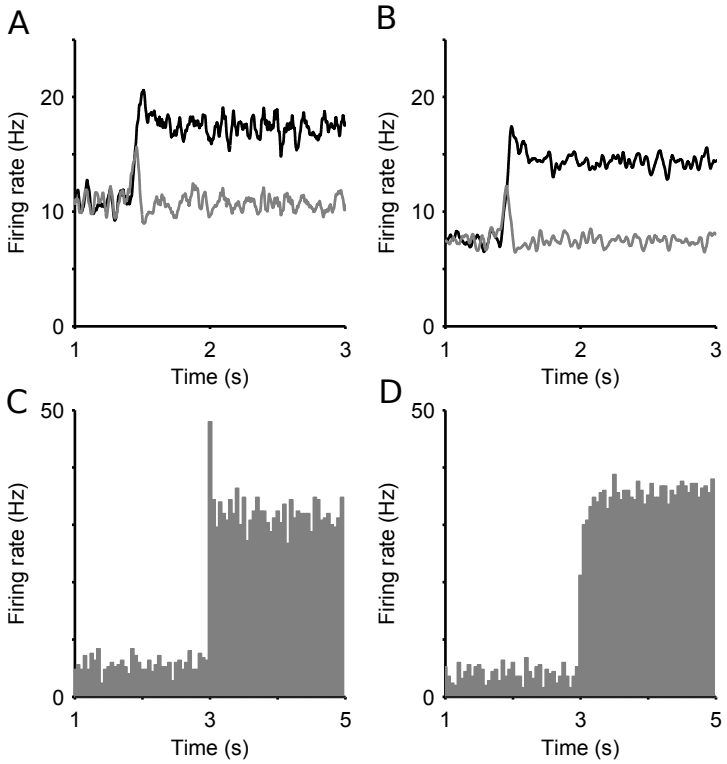

Fig. 6.2 **Positive transients are specific to the SPN model.** **(A)**. Striatal output with lesioned projection neuron connections. **(B)**. Striatal output with all intra-striatal connections lesioned. **(C)** Peri-stimulus time histogram of a single projection neuron output, averaged over 50 steps of spiking input from $r = 4$ Hz to $r = 7.2$ Hz (onset $t = 3$s), exhibiting transient behaviour. **(D)**. Peri-stimulus time histogram of a single regular-spiking cortical neuron model, averaged over 50 steps of spiking input from $r = 0.75$ Hz to $r = 3$ Hz, with no transient behaviour. Model parameters given in (Izhikevich, 2007a).

input. Running the same test on a model of a cortical regular-spiking pyramidal neuron, with input scaled to produce approximately the same steady-state rates, showed no such transient increase in firing probability after a step in input (Figure 6.2D). Thus, the transient increase in population activity observed in a single trial of the network is a statistical phenomenon of synchronous spiking of many projection neurons, and seemingly dependent upon properties particular to the striatal projection neuron.

We sought to elucidate these properties by injecting sequential current steps directly into the projection neuron model and observing the behaviour of the membrane voltage v and slow current u. Figure 6.3 A shows that a step in current applied to an already depolarised membrane triggers a rapid double spike, followed by slower regular spiking. Figure 6.3 B plots the corresponding trajectory of the slow current u: the initial depolarising injection makes the slow current u increasingly negative, thus slowly charging the membrane potential v (Figure 6.3 A; see Equation 3.19). The subsequent step of injected current increases the membrane potential rapidly, and the contribution of the large, negative u ensures a rapid pair of spikes time-locked to the current step. However, once spiking has been initiated, the equilibrium value of u is less negative than immediately before the current step. Consequently, the smaller contribution of the slow current u ensures a comparatively slow spike rate in the steady-state.

To show that the slow current u is critical, we examined the dependence of this spiking "adaptation" on the parameters of the slow current. We repeated the sequential-step current injection protocol for a range of step-sizes, and measured the adapting response as $f_{ratio} = F_{first}/F_{last}$, the ratio of the first and last inter-spike intervals after the current step. A value of $f_{ratio} > 1$ thus indicates an adaptation. We found that the adaptation response appeared with a second current step above ~ 50 pA (blue curves in Figure 6.3 C,D).

Figure 6.3C shows that the adaptation response disappeared if we reduced the effective time constant of the slow current (increased a), allowing the slow current to recover faster after spiking. Figure 6.3 D shows that the adaptation response also disappeared if we reduced the gain b of the slow current The transient phenomena thus depends critically on the slow current u.

Negative transient is due to network level effects

Lesioning SPN-SPN connections abolishes the negative transient (Figure 6.3A), this suggests the negative transient emerges from SPN-SPN local inhibition. The neurons contributing to the positive transient, contribute to a network effect, in order to synchronously suppress activity in competing channels.

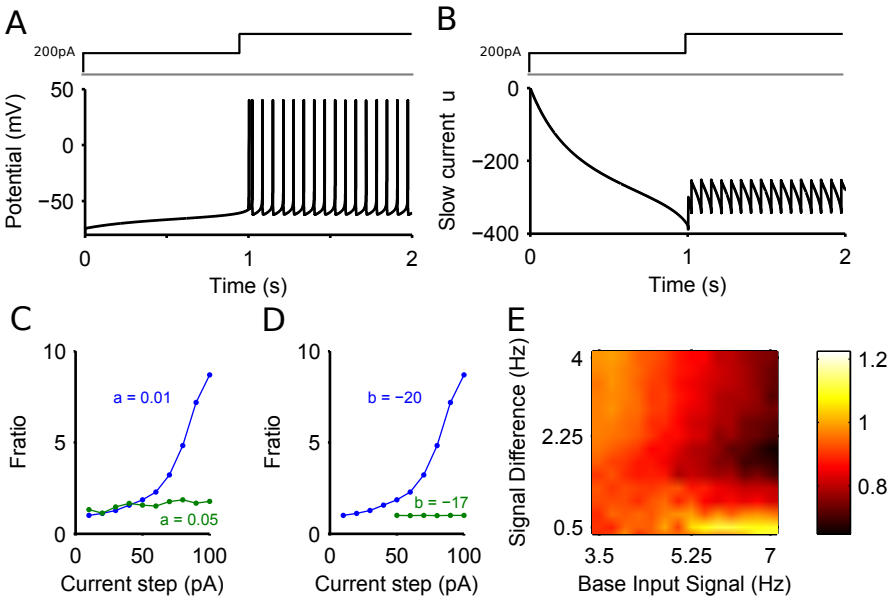

Fig. 6.3 **Sources of the positive and negative transients.** **(A)** The membrane potential (v) of the projection neuron model in response to a depolarising current injection (200 pA) followed by a further step in current at 1 second. **(B)** The corresponding changes in the slow current (u). **(C)** f_{ratio} in the projection neuron model as a function of current step size and slow current decay constant $1/a$ ms. **(D)** f_{ratio} in the projection neuron model as a function of current step size and slow current gain b. **(E)**. The effect of projection neuron connection lesions on the negative transient. Landscape of negative transients measured as ratio of the maximum negative transient peak over the steady state, plotted as a function of base input rate vs signal difference. A ratio of 1 supports no suppression, due to the pre and post signal activations being equal.

To test this observation, we simulated the model with lesioned SPN-SPN collaterals, over a range of baseline cortical signal firing rates, and a range of salient step sizes (protocol in Figure 5.2A). For each combination of base salience, and step size, we calculated the size of the negative transient that was elicited. Figure 6.3E shows that the negative transient was indeed abolished for a wide-range of values for the input firing rates. The negative transient is calculated as the ratio of pre- and post-signal divergence activity, i.e. S^P/S^2 (defined in figure 5.2). We can see that as $S^2 == S^P$, the ratio approaches 1, indicating no negative transient. As S^2 is reduced, the ratio increases >1. However, a sufficiently large baseline firing rate and step in firing rate could still result in a negative transient (right side of Figure 6.3E). Thus, it seems that sufficient cortical drive of the FSI population (which inhibits the projection neurons) also contributes to the negative transient in projection neuron population activity.

6.2 Transient selectivity is amplified by decision circuits.

We mimicked the ability of the striatum microcircuit to produce transient phenomena using an input injection into the striatum of the rate coded model, as seen in figure 6.4. As a rate coded model, the basal ganglia-thalamo-cortical model is unable to intrinsically reproduce the selection effects seen in the cortical microcircuit. We introduce a phenomenological model of transient elicitation, in response to cortical inputs using current injection. At $t = 100$ we injected external inputs into each striatal channel in the model, forcing a transient increase or decrease as appropriate in the corresponding channels. The most salient channel received a short burst of positive input, the least low salient channel received a short negative input.

Transient sizes were extracted from the striatal microcircuit traces, and reproduced in the rate coded model, to phenomenologically reproduce the transients elicited in the spiking neuron model. Individual transients were calculated as the percentage change in the firing rate of the circuit during the transient period compared to the stable firing rate achieved post-transient. This allowed us to gauge the role of the complex striatal dynamics, on the selection properties of the entire basal ganglia-cortex loop.

Mimicking striatal dynamics in a rate coded basal ganglia-thalamo-cortical loop

Though the previous result demonstrates the existence and origin of transient selection within the striatum, it is not sufficient to show a causative effect of transient selection on decision-making and in the larger basal ganglia-thalamocortical loops. To address this issue, we asked whether transient signals in the striatum would enhance the signal selection basal

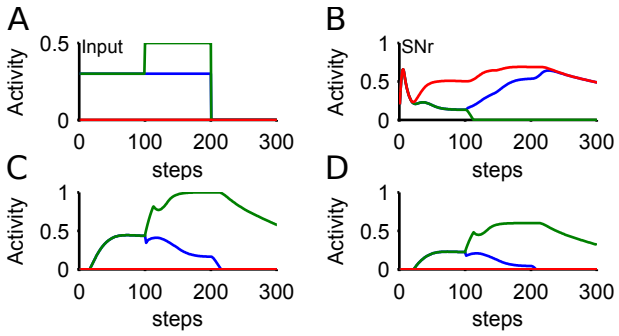

Fig. 6.4 **Transient injections in basal ganglia-thalamo-cortical loop**. Panels **(A)-(D)** show an example simulation of the loop model that included emulation of the transient selection signals originating in the striatum (transient size: 50%; thalamo-cortical loop gain g=2). **(A)** Cortical input to the rate-coded model, mimicking the selection protocol used in the striatal microcircuit selection experiments. **(B)** Corresponding SNr output response for three populations: no input (red); baseline only (blue); and baseline-plus-step (green). The input steps thus caused clear selection by forcing the SNr output to zero. **(C)** Evoked response in the rate coded striatal D1 neurons, showing the effect of the injected transient at $t = 100$. **(D)** Evoked response in the rate coded striatal D2 neurons.

ganglia circuit. Here we consider selection to mean that the output of a substantia nigra pars reticulata (SNr) population falling from its tonic rate to zero, this dis-inhibiting the desired action. In particular, we hypothesised that the transient signals in striatum would be amplified in the complete basal ganglia-thalamo-cortical loop, and thus directly influence the output of the basal ganglia.

To test this, we reproduced our two-choice decision task, using our rate-coded model of population activity in the basal ganglia-thalamo-cortical loop (Humphries and Gurney, 2002). The model received inputs to two populations of cortico-striatal neurons (Figure 6.4A), mimicking the protocol used in our full-scale striatum model. An example of the subsequent SNr outputs are illustrated in Figure 6.4B. At the time of the step in input to one population, we emulated the subsequent transient signals observed in our full-scale model, in the rate coded mode, by brief salient injections of activity, in order to elicit a transient in the otherwise smooth striatal population dynamics, resulting in mimicked transient behaviours. These injections correspondingly produced small, brief positive and negative transients in the output of those striatal populations, for both D1 and D2-type projection neurons, which can be seem in Figures 6.4C-D). Note that the subthalamic nucleus populations also received the cortical input signals, but not the transient signals.

6.2.1 Transient selection is sufficient to reduce Time-To-Selection

We found that a small positive transient elicited in the striatal population, like that shown in figure 6.4C was sufficient to change the speed and persistence of selection (Figure 6.5A-D). Increasing the size of the elicited transient forced the circuit to achieve quicker selection. Correspondingly, Figures 6.5A,C show that increasing the size of transients injected into the model striatum decreased the time to selection. These changes were found irrespective of the size of input step, or of the closed-loop gain g of the positive thalamocortical feedback loop (Chambers et al., 2011) (When g=1, this loop is a perfect integrator, while with g=2, there is an amplifying feedback loop, and with g=0.5, the loop is a leaky integrator) Figures 6.5B,D examine the proportion of time that an action remains selected after the jump in cortical salience. Figure 6.5B demonstrates that, with low cortical input, a thalamocortical feedback loop always improves signal selection, with improving selection corresponding to increasing transient sizes. In a perfect integrator configuration, the presence of a large transient is able to enable action selection where it would have otherwise not have occurred. Figure 6.5D demonstrates that under high signal salience, transients still marginally improve signal selection. Thus, transient signals in the striatum are sufficient to significantly modulate selection by the basal ganglia.

6.3 Discussion of results

We found a novel form of transient selection supported by the striatal network. This emerged from our three-dimensional network of sparse, weak feedback connectivity between the striatal projection neurons and dense, strong feed forward inputs from the fast-spiking interneurons. We observed that rapidly increasing the ongoing input to one of two competing populations of projection neurons caused a transient peak of activity in that population and a synchronous transient dip in activity of the other. The dip lasted around 100 milliseconds before the activity returned to its pre-step level, thus showing no steady-state competitive effect between the two populations.

Using a population-level model of the complete basal ganglia-thalamo-cortical loop, we showed that transient selection in the striatum was sufficient to enhance selection by the entire circuit (as determined by suppression of SNr output). The presence of transient selection both increased the speed at which the whole circuit resolved a competition between salient inputs, and increased the circuit's ability to persist with the selected input. Both effects were observed for either perfect-integrator or amplifying feedback in the thalamo-cortical loop.

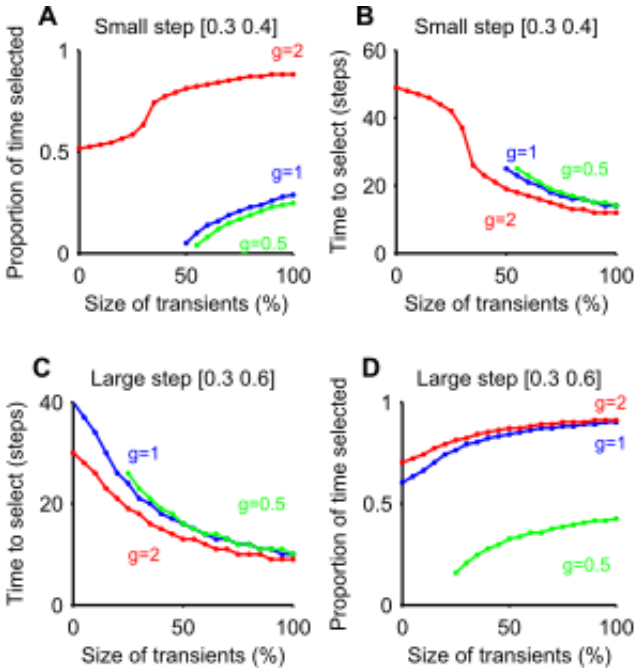

Fig. 6.5 **Transient selection in striatum is amplified by basal ganglia-thalamo-cortical loop (A)** Proportion of time an action was selected, as a function of transient size. Transient size is expressed as a proportion of the steady-state firing rate achieved without the transient. Step values indicate the cortical input before and after the step in input. Parameter g: closed-loop gain of the thalamocortical loop $g = 1$ is a perfect integrator, $g > 1$ is an amplification loop and $g < 1$ is a leaky integrator. **(B)** Proportion of time an action was selected, given a small input step. **(C)** Time delay before selection achieved, as a function of transient size, for large input step. Delay is given between the step in cortical input and the corresponding SNr population reaching zero output. **(D)** Time delay before selection achieved, as a function of transient size, for small input step. Results are generated deterministically, by the equations 3.4-3.5

The origin of the transient selection had two components. The positive transient in the population activity was driven by single neuron adaptation. We found that a further step in input to an already depolarised projection neuron caused a spike followed by rapid decrease in spiking probability. This implies that the positive transient observed in the population activity was a statistical effect: that, across a whole population of projection neurons, a sub-set of neurons were sufficiently depolarised at the time of stepped input to show this adaptation effect in synchrony, and thus cause a transient peak in population activity.

The negative transient in the population activity was a subsequent network effect of the positive transient: the synchronised spiking of the neurons participating in the positive transient was sufficient to drive a dip in activity in their target neurons in the other population.

Optimality of decision making in the basal ganglia

Bogacz and Gurney (2007) suggest that the basal ganglia may compute for asymptotically optimal decision making. However, it would seem that the results of Beste et al. (2007) as well as the results of our study, suggest that the decision making can be improved upon, through certain pathological modifications. Primarily, it must be mentioned, that the statistical test suggested by Bogacz and Gurney (2007), is the MSPRT decision task, which is designed to compute over more than two perceptual choices, and as such is not directly applicable here, yet the example is instructive. As a sequential test, we see that the data is continuously gathered, until a decision point is met. The transient phenomena we see, provides a mechanism for increased information transmission in reduced time, and as such provides for a quicker decision making task. However, we can argue that in order to benefit from this mechanism, we sacrifice the optimality of the basal ganglia as a whole, as decision making is impaired almost across the board, under Huntington's like degradation.

While the transient mechanism can improve selection, the phenomena would not necessarily improve decision in all tasks, and may be constrained by the requirements of the auditory selection task. We see that for longer evidence-accumulation based tasks, the steady-state selectivity mechanism provides a more robust, information rich source, for longer periods of time. We also see a longer selection time, as well as a shorter time-to-selection. This result may prove a hindrance under certain circumstances, such as quick-changing sequential action selection tasks. As such, I suggest that these results do not detract from the overall optimality of the basal ganglia, but instead show that sacrificing general optimality, may produce locally improved circumstance for constrained, task specific domains, however, this would reduce the overall efficacy of the general selection mechanism of the basal ganglia.

Experimental predictions of transient selection

Direct experimental observation of transient selection is challenging. The positive transient in population activity could only be observed on a single trial given sufficient simultaneous sampling of neurons within that population, a situation unlikely to occur with current recording technology. However, we showed that the basic mechanism underlying the positive transient in the population activity could be observed through sequential steps of current injection into a single neuron model. Thus, our model makes a tractable experimental prediction: that there exists a regime of long, sequential steps of current into the projection neuron soma that will elicit a rapid burst of two or more spikes followed by slower regular firing. If such a regime exists, it would provide evidence in favour of the existence of transient selection mechanisms in the striatal network.

We have demonstrated the efficacy of transient selection in the model striatum, and validated our hypothesis concerning selection in the striatum impacting action selection in the basal ganglia with this specific form of selection. We have seen in chapter 5, that the striatum is also able to support a second form of signal selection. In chapter 7 we set out to understand steady-state selectivity, and determine how the striatum can support winner-take-all selectivity as well as transient selectivity.

Chapter 7

Steady-state selectivity affects signal selection in dense networks

Having established the existence and mechanics of the transient selection phenomena, we sought to understand the conditions under which our striatal model could also support a steady-state competitive effect, akin to classical winner-takes-all dynamics (Fukai and Tanaka, 1997; Hartline and Ratliff, 1958; Mao and Massaquoi, 2007). Steady-state competition could plausibly arise in the striatum as each projection neuron receives a sufficient volume of weak synapses from other projection neurons to continuously modulate ongoing activity (Chuhma et al., 2011; Guzman et al., 2003; Humphries et al., 2010).

We investigate the relationship between *steady-state suppression* and network connection density, and hypothesize that competitive signal suppression emerges from increasing connectivity in the striatal network. We determine whether the phenomena can be observed in both a densely connected *random* network connectivity, and a larger *physical* network connectivity (due to increasing connection density with size).

We have investigated the root causes of *steady-state selection*, to determine whether it is an emergent property of SPN-SPN connectivity in the striatum,by lesioning SPN-SPN collaterals and FSI-SPN collaterals.

7.1 Local SPN inhibition facilitates steady-state selection

Prior debates about selection in the striatum have focussed on winner-takes-all modes of computation (Plenz, 2003; Wickens, 1997). In order to compare striatal selection with this more common form of selection computations, we sought to understand whether our striatal model could show winner-takes-all types of dynamics.

Neurally-inspired models of winner-take-all dynamics are often based on fully-connected or dense randomly-connected networks (Alexander and Wickens, 1993; Fukai and Tanaka, 1997; Hartline and Ratliff, 1958; Mao and Massaquoi, 2007; Yim et al., 2011). We thus simulated our striatal model with random connectivity, in which each neuron type received, on average, the same number of connections, and the connections were made by choosing source neurons at random from across the three-dimensional model. The target number of connections was based on the expected number of connections of a projection neuron and FSI in the centre of a 1 mm^3 network, according to the computational anatomical estimates of Humphries et al. (2010). In this way, the randomly-connected model was more densely connected relative to the distance-dependent model. Thus, while closer to the topology usually studied for steady-state selection, the randomly-connected model still retained connection statistics consistent with the estimates obtained in Humphries et al. (2010).

We tested the randomly-connected model with the same stepped input protocol as the physically-connected model, emulating a cortical two-choice decision task (Figure 7.1A inset).

Figure 7.1A shows an example of the mean population firing rates in the randomly-connected striatum model, with evident steady-state selection. The population receiving the stepped cortical input increases its firing rate, as per the increased cortical salience. Simultaneously the least salient population correspondingly has its activity suppressed, to below pre-ramping levels. Firing rate suppression in the least salient signal represents an active striatal modification of cortical signal representation.

We found that the magnitude of steady-state selection was dependent on the size of the baseline firing rate of the cortical inputs, and the salience increase in the strongest channel. Figure 7.1B shows that the most drastic steady-state selection occurred for between signals with low baseline cortical input salience (3.5-4 Hz), and large input steps (4Hz increase). Represented in yellow, this steady state selection of 0.8 and higher represents a strong winner-takes-all like response of nearly complete suppression ($\sim 80\%$) of the activity of the least salient signal, with no decrease in cortical representation.

Steady state selection can be shown to depend on local inhibitory SPN-SPN connections. Figure 7.1C demonstrates the effect of lesioning the connections between projection neurons, where steady-state selection is destroyed. This has important implications for rate-coded, feed forward models of the basal ganglia. Models which do not encode the inhibitory local-dynamics of the striatum, will not exhibit this WTA-like behaviour. We propose that rate

coded channel models must at least employ lateral inhibition, to capture some steady-state suppression dynamics.

WTA networks require that only the most salient signal is able to fire, having suppressed all other competitors. Steady-state selectivity does not inherently model a classical WTA network, but instead implements a form of soft-bounded WTA computation, such that the winning channel suppresses the less salient channels, as a function of the most salient input. Larger salient inputs approach pure WTA behaviour. This further models WTA behaviour, when placed in the larger decision circuits of the Basal Ganglia-Cortico-Thalamus circuits, in which, recurrent feedback enforces, a strict single channel activation.

FSI inhibition can be shown to improve steady-state selection. Figure 7.1D demonstrates the effect of lesioning the FSI input to the projection neurons. Lesioned FSI-SPN connections reduces, but does not eliminate the steady-state selection. Interesting, this reinstates a transient period. This suggests that mutual intra-channel inhibition by the projection neurons populations is responsible for the suppression effect seen in both the *random* and the larger *physical* networks.

7.2 Distance-dependent connectivity can support steady-state selection

To assess if such steady-state selection required homogeneous, random connectivity of the kind described above, we checked whether such selection could be found in the physical model of connectivity. Again using the same stepped input protocol, we simulated physical networks up to 1 mm^3, in order to increase the density of connectivity within the centre of the network, which scales with the number of neurons in the model (Figure 7.2B).

Figure 7.2A shows that steady-state selection could be observed for distance-dependent connectivity, given a sufficiently large model (here 1 mm^3). We found that the magnitude of steady-state selection increased monotonically with increasing network size (Figure 7.2D), approaching the steady-state selectivity seen in the *random model*. Figure 7.2B,C shows that in the *physical* model as the number of neurons increases as a function of network size so does the average number of connections each projection neuron receives. By contrast, the *random model* always has the same density of connections. The *physical* model's correspondence between the number of connections to a projection neuron and the effectiveness of steady-state selection suggests that such selection is dependent on the density of connections between projection neurons.

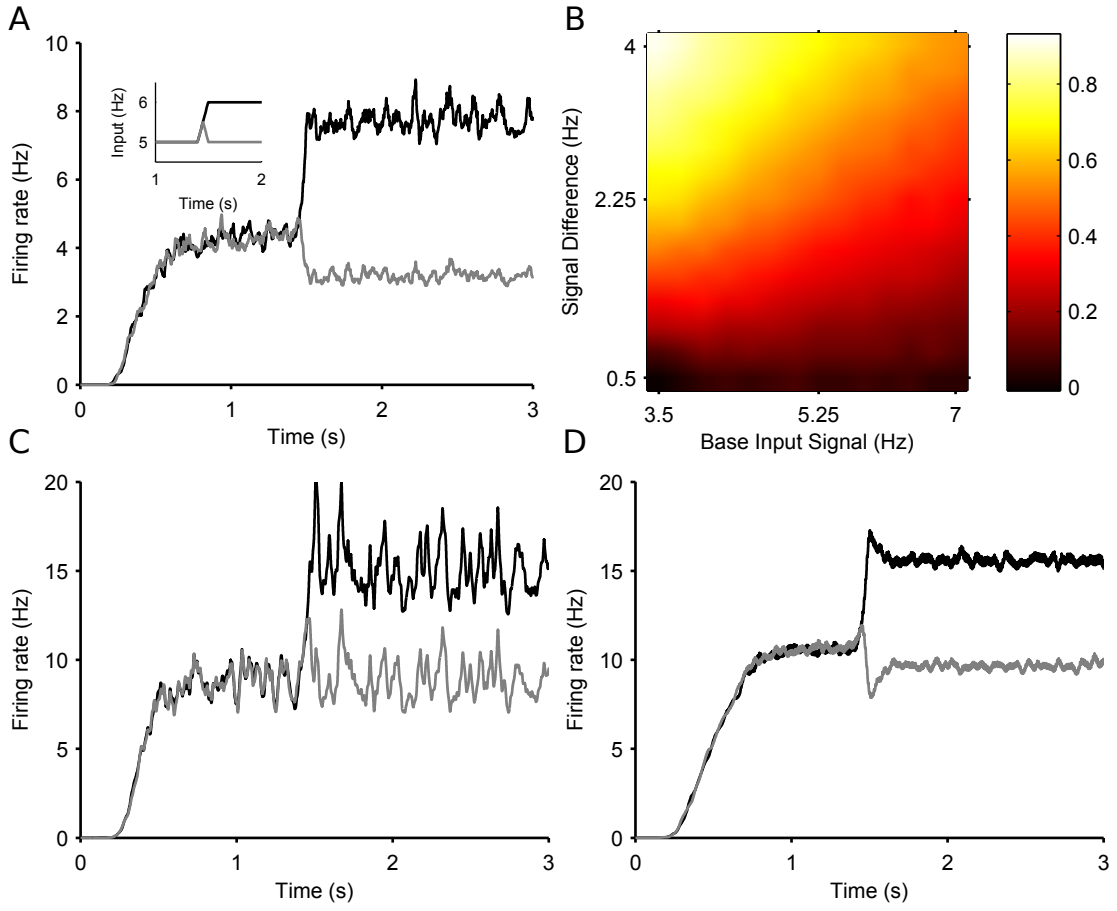

Fig. 7.1 **Steady-state selection in the randomly-connected striatum model. (A)** Smoothed mean firing rates of two projection neuron populations, in response to the ramped input protocol (inset). **(B)** The magnitude of steady-state selection as a function of baseline input and step size. The magnitude gives the fall in firing rate of the losing population as a proportion of its pre-step firing rate. Each magnitude is an average over 15 simulations. **(C)** Smoothed mean firing rates of two projection neuron populations, with SPN-SPN connections lesioned, in response to the same input as above. Steady-state selectivity is removed. **(D)** Smoothed mean firing rates of two projection neuron populations, with FSI-SPN connections lesioned, in response to the same input as before. Steady-state selection remains.

The model further suggests that it is only the increased density of connections that is key, and not an increase in recurrent connections between projection neurons. Figure 7.2E shows the absolute number of recurrent connections in the *physical* and *random* network configurations. Note that the number of bi-directional connections in the *random* network drops of as a function of network size due to the fact that each neuron receives a fixed number of connections regardless of the network size. By contrast we see a small rise in the number of bi-directional connections in the *physical* model. However, Figure 7.2F shows that in both random and physical networks the proportion of connections that are bi-directional falls with increasing network size. Thus, the increased effectiveness of steady-state selection is likely due to increased absolute connection density and not increased recurrent connections.

7.3 Discussion of results

We have documented a novel form of selectivity that is supported by the striatal microcircuit. *Steady-state selectivity* is distinct from *transient selectivity*, in that an 'out competed' signal is actively and continuously suppressed, in proportion to the salience of the 'selected signal'. We propose that *steady-state selection* provides a neural analogue to the various winner-take-all signal selection architectures, with signal suppression ensuring that the least salient cortical signal can not be selected by the disinhibition mechanisms of the basal ganglia decision network. This mechanism ensures a clean selection of the most salient signal.

Increasing the number of projection neuron synapses gave rise to steady-state competition, such that a stable increase in activity in one population caused a stable decrease in activity of the other population. These results are consistent with Yim et al. (2011) who reported a weakly-competitive effect between two populations of neurons in a randomly-connected inhibitory network of spiking neurons. They further showed that weak correlation between inputs to the network could enhance this effect. We advanced this result by showing that such steady-state competition could arise in both distance-dependent and randomly-connected networks, given either that we increased the physical size of our three-dimensional striatal network, and thus increase the density of connections, or randomly-connected the network based on the average connections of the most densely connected projection neuron.

Our models thus predict that this form of selective competition is dependent on the density of connections between projection neurons. Whether the striatum is ever as sparsely connected as in our distance-dependent model, or ever as densely connected as in the homogeneous random model is an open question. It is possible that both forms of selection exist depending on local inhomogeneities in striatal tissue. We know that many aspects of the

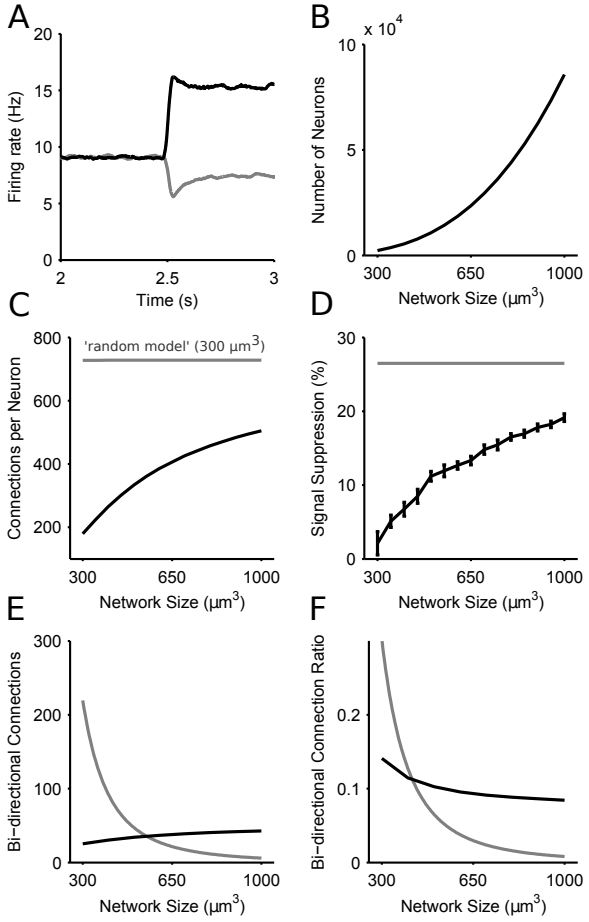

Fig. 7.2 **Steady-state selection in the physical model of the striatal microcircuit.** **(A)** Mean firing rate of two projection neuron populations in a 1 mm^3 model, with 89,749 total simulated neurons. **(B)** Number of simulated neurons as a function of network size. **(C)** Average number of connections per neuron as a function of network size. The *physical* network (black) approaches the density of connections seen in the random network (grey) with increased network size. **(D)**Magnitude of steady-state selection as a function of network size. All simulations used the inputs [5,6] Hz. Magnitude is the percentage suppression in the average firing rate of the losing channel after the competitive signal onset ($t = 2.5$ s). Shown in grey is the steady-state selectivity seen in the random model for a network of size 300 μm^3. Bars set at $\pm$ 2 s.d, computed over 15 repeats. **(E)** Number of bi-directional connections as a function of network size. The total number of pairs of reciprocal connections in the *physical* model are shown in black, and the *random* model in grey. Bi-directional pairs decrease in the physical model with increasing network sizes, due to the fixed number of connections each neuron receives. **(F)** The ratio $R_{bi}$of bi-directional connections to the total number of connections a neuron makes for the *physical* model (black) and the *random* model (grey).

striatum show gradients of density across the network, including the dorsal-ventral gradient of inter-neuron populations (Kubota and Kawaguchi, 1993) and the rostro-caudal gradient of FSI gap junctions (Fukuda, 2009). Correspondingly, it is plausible that there exists a gradient of projection neuron connection density.

We also note that the recent report by Oorschot et al. (2013) of projection neuron collaterals making synapses on to the somas of other projection neurons can only enhance both forms of competition. Such GABAergic somatic synapses are likely to shunt all dendritic input to the soma, thus providing powerful feedback inhibition. For transient selection, this could result in a larger negative transient; for steady-state selection, this could result in more depressed activity in the losing population. Open questions here include the relative density of such somatic synapses originating from projection neurons, and whether they have specific functional targets such as specifically occurring between projection neurons in competing populations.

Both forms of striatal selection mechanisms ultimately influence selection mediated by the whole basal ganglia network and expressed via their output nuclei (including SNr). This expression is via disinhibition (Berns and Sejnowski, 1998; Chevalier and Deniau, 1990; Gurney et al., 2001; Humphries et al., 2006; Redgrave et al., 1999), such that: increased activity of a striatal population inhibits the tonic inhibitory output of a SNr population, thus representing the selection of their represented signal (Figure 3.5). We showed that transient selection in the striatal populations is sufficient to enhance selection by disinhibition from SNr (Figure 6.4). This occurs because the most salient input causes a transient increase of activity in the corresponding striatal population and consequently transiently decreases the output of the corresponding SNr population. This fall is sufficient to allow activity to grow in the target thalamocortical loop, which in turn projects to the original striatal population, further increasing its activity – thus the positive feedback loop amplifies the transient changes in striatum. The effect of steady-state selection in the striatum on the whole basal ganglia is more straightforward. The long-lasting drop in output of all losing striatal populations comparatively reduces their inhibition of the corresponding SNr populations. Consequently, the fall in output of the SNr population representing the winning signal is enhanced compared to its competitors.

We have shown that large physical networks of the striatum are able to express both *transient* and *steady-state* selectivity, suggesting that these mechanisms can work together under the right network conditions. We have discussed how the striatum has several documented connectivity gradients, and as such, that it is plausible that these two selection mechanisms exist in the human striatum, and may work together to actively improve signal selection.

With a detailed understanding of the mechanistic underpinning of two forms of active selection in the striatum, we sought to understand how these mechanisms can be differentiated. Using our knowledge of the distinct underlying mechanisms, we hypothesise that each mechanism would evolve differently under observed neurological changes in the striatal network. In chapter 8, we consider the neuro-degenerative Huntington's disease, which has been shown to heavily effect the striatal microcircuit,coincident with severe cognitive and behavioural effects. We test our models, under simulated HD conditions and predict that the observed selectivity mechanisms in the striatum will evolve differentially, under HD-like degradations. This hypothesis would have important physiological implications, and could be predictive of some recent paradoxical behavioural results.

Chapter 8

Striatal selectivity may explain Huntington's disease results

We have demonstrated that the striatum plays an active role in action selection in the basal ganglia. Using the two forms of selectivity that we have uncovered, we have given a distinct role to the complex striatal microcircuit, showing that the circuit can support two forms of complementary selectivity which impact action selection by enhancing selective disinhibition through the SNr. We have detailed the underlying mechanisms which can give rise to such phenomena, and have hypothesised that due to their distinct mechanisms, that each form selectivity may respond differently to changes in the striatal microcircuit.

A hypothesised differential selectivity response to striatal modification, raises some interesting questions, and enables some powerful explanatory tools. Striatal degradation is one of the main symptoms of Huntington's disease, and associated with this, is extensive behavioural change. However, one recent experiment has shown that some cognitive abilities can improve during late stage Huntington's disease, coincident with massive striatal degradation. The underlying mechanisms of these results are as yet, unknown. We investigate the hypothesis that the novel forms of selection seen in our striatal microcircuit models, could evolve differentially under HD-like striatal degradation. We posit the hypothesis that our distinct selection mechanisms, may underpin the distinct behavioural changes seen in HD patients, enabling both general behavioural decline, while supporting specific cognitive improvements.

In what follows we discuss the simulation results of our model and interpret them as potential mechanisms explaining the findings of Beste et al. (2008). To compare *transient* and *steady-state selection* we challenged both forms of model to account for the paradoxical response selection results of Beste et al. (2008). In their work, manifest Huntington's disease

patients were both faster and less error prone than controls on a simple two-choice reaction-time task. As Huntington's disease primarily results in striatal damage, this suggests the hypothesis that changes in the striatum directly affect response selection.

We expand on the role of the striatum in signal selection, by describing a framework for signal selection that may account for both the typical decline in performance for most tasks under Huntington's conditions (Ho et al., 2003), and a mechanism for increased performance under the similar degraded conditions for different tasks. We thus explored how Huntington's disease-like changes to our striatum models could affect both transient and steady-state selection, and sought whether the effect on either form of selection could explain the results of Beste et al. (2008), while also accounting for the usual cognitive impairment in Huntington's disease (Ho et al., 2003; Lawrence et al., 1998).

8.1 Comparing selection mechanisms: paradoxical selection enhancement in Huntington's disease

We have established that two contrasting forms of selection can be supported by the striatal circuit, depending on the type and density of connectivity. We then sought insight into how the two forms of selection could be distinguished. In particular, we hypothesised that they would make different predictions about how changes to the striatum would alter response selection. In order to test this hypothesis, we sought an experimental data-set that could provide a basis for testing our predictions.

Beste et al. (2008) have recently shown a rare example of paradoxical cognitive enhancement in a neurological disorder. They reported that manifest Huntington's disease patients reacted quicker, and were less error prone on a simple two-choice auditory task, than healthy controls or pre-manifest Huntington's disease patients. As Huntington's disease is primarily characterised by widespread loss of striatal projection neurons (FSI populations have been shown to be more resistant to HD modifications (Ghiglieri et al., 2012)), and exhibit increased sensitivity of NMDA receptors on striatal projection neurons (Fan and Raymond, 2007), these results suggest the hypothesis that one or both of these changes to the striatum lead to enhanced selection. As such we look into excitotoxicity as a possible candidate for the paradoxical improvements investigated.

We thus simulated both *transient* and *steady-state selection* under HD-like changes to the striatal model, and searched for evidence of enhanced selection. We emulated increased NMDA receptor sensitivity by increasing the conductance of the NMDA synapse (we report

this as the ratio of the NMDA:AMPA conductances), and separately emulated the cell loss by randomly removing a specified percentage of projection neurons. We did this to explore a wide range of plausible simulated Huntington's disease conditions. Across both changes, we mapped the change in transient and steady-state selection in response to the same input protocol (baseline 5 Hz, step 1 Hz).

8.2 Transient selection enhancement in Huntington's disease

We assessed the impact of Huntington's-like changes on *transient selection* using the same *physical* model network as that used for Figure 6.1. Figure 8.1 shows that transient selection could be diminished by the loss of projection neurons alone, yet could be enhanced by the simultaneous increase in NMDA conductance. Thus the model predicts a region of Huntington's-like conditions where the deleterious effect of cell loss can be overcome, by a compensatory increase in selectivity, when cell atrophy is combined with increased sensitivity of NMDA receptors.

Figure 8.1A shows an example improvement in transient selectivity in a striatum with high cell atrophy, of 81% and a high NMDA receptor excitability, at an NMDA:AMPA conductance ratio of 0.95, almost twice a healthy striaum. At $t = 1.5s$ we can see defined transient phenomena, distinct from the noisy steady-state firing seen after $t = 2s$. Conversely, figure 8.1B shows the suppression of the transient selectivity under a high cell atrophy of 81%, and only a small increase in excitability to a NMDA:AMPA ratio of 0.55. Under these conditions, both the positive and the negative transient seen in figure 8.1A, have been reduced. These examples show that the Transient Selectivity range of $\sim$0.10 over the excitotoxicity landscape in Figure 8.1C, corresponds to dramatic changes in the striatal output, which in turn, is amplified in the basal ganglia thalamocortical loop. Figure 8.1C examines the transient selectivity over the excitotoxicity landscape. Here yellow patches represent areas of high transient selectivity, which have shown an improvement over the healthy striatum. This landscape shows that there exists, a combination of NMDA excitability, and cell atrophy which combine to improve the transient selectivity.

Further, as we have shown in Figure 6.5, small modifications in the transient size in the striatum will modulate the signal selection speed in the wider basal ganglia networks. We can conclude that there are possible combinations in the progression of cell atrophy and

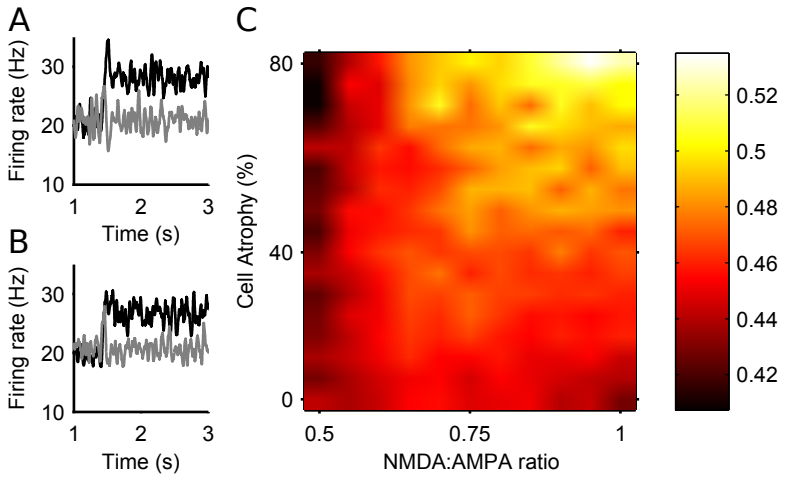

Fig. 8.1 **Transient selection can be enhanced in simulated Huntington's disease (A).** An example of enhanced transient selection in a Huntington's-like condition (81% cell atrophy, 0.95 NMDA:AMPA ratio) **(B)** An example of the loss of transient selection in a Huntington's-like condition (81% cell atrophy, 0.55 NMDA:AMPA ratio). **(C).** Selection landscape for NMDA:AMPA conductance ratio against cell atrophy. Colour coded such that brighter colours represent better transient selectivity in the striatal model. Magnitudes are means over 30 simulations. The control, healthy-state model is in the bottom left-hand corner (NMDA:AMPA = 0.5; 0% atrophy).

NMDA sensitivity which, given the same cortical inputs, will achieve signal selection faster, and more accurately than the healthy striatum model.

8.3 Steady-state selection consistently degrades in simulated Huntington's disease

To assess the impact of Huntington's-like changes on steady-state selection, we used the randomly-connected model to measure if the suppression of the losing population is sufficient to be detectably modulated by the Huntington's-like changes.

Figure 8.2 A shows the presence of steady-state selectivity in a degraded striatum with cell atrophy of 65%. Cell atrophy reduces the steady-state selectivity vs the healthy state striatum, as well as a predictable reduction in general firing rate. Figure 8.2 B shows an example of a degraded striatum with 75% cell atrophy and a 100% increase of NMDA receptor sensitivity. Here steady state selectivity can be seen to have reduced, regardless of the NMDA receptor sensitivity compensating for the reduced firing rates due to cell atrophy.

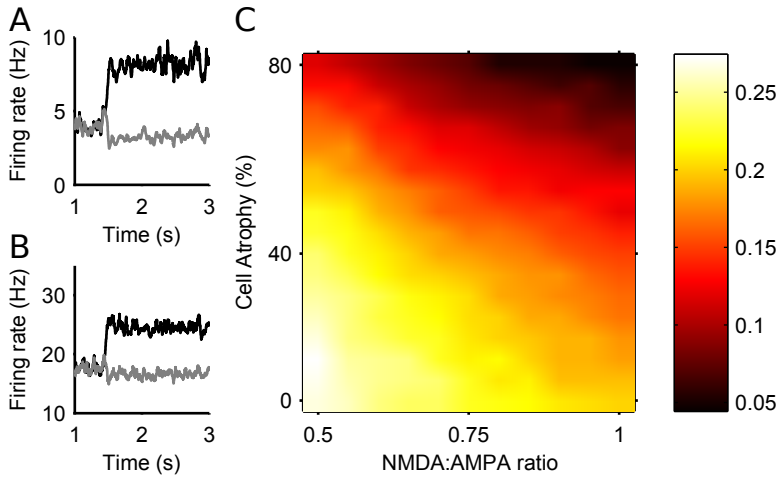

Fig. 8.2 **Steady-state selection declines under simulated Huntington's disease. (A)** An example of reduced signal suppression in the striatum with high cell atrophy (65% cell loss, NMDA:AMPA ratio 0.5). **(B)** An example of removed signal suppression in the striatum with high degradation (75% cell loss, NMDA:AMPA ratio 1). (C) Magnitude of signal suppression over all simulated Huntington's conditions. Magnitudes are means over 15 simulations. The control, healthy-state model is in the bottom left-hand corner (NMDA:AMPA = 0.5; 0% atrophy).

Interestingly, unlike transient selectivity, we show that there is no compensation mechanism which can recuperate steady-state selectivity under our simulated HD-like modifications. Figure 8.2 C shows that the landscape of degradation for steady-state selection. We see that the healthy striatum represents the pinnacle formation for steady-state selection, and any HD-like modification smoothly reduces the signal suppression phenomena. Steady-state selection is uniformly diminished by all Huntington's-like changes, whether in isolation or any combination of cell atrophy and NMDA receptor sensitivity increase. m As steady-state selection has been shown to depend on the density of SPN-SPN connections in the striatum, and is independent of neuron dynamics, it is no surprise that cell atrophy causes reduced steady-state signal selection.

8.4 Transient selection alone could explain enhanced selection in Huntington's disease

We sought to determine whether transient and steady-state selection could be differentiated by their predictions for how changes to the striatal circuit would affect selection. To this end, we asked if Huntington's-like changes of increased NMDA receptor sensitivity and loss

of projection neurons could account for Beste et al. (2008)'s report of enhanced selection by Huntington's disease patients. In terms of our models, we asked if either *transient* or *steady-state selection* would improve due to these Huntington's-like changes to the striatum.

As one might expect a priori, simply removing projection neurons and thus reducing connectivity between them impaired both forms of selection. Increasing NMDA receptor sensitivity also impaired *steady-state selection*, and thus this form of selection predicted that all Huntington's-like changes impair selection, a result which is inconsistent with the report by Beste et al. (2008). Surprisingly, however, we found that for transient selection, increased NMDA receptor sensitivity could more than compensate for cell loss and actually enhance selection. We also found that transient selectivity was only clearly improved with both high cell degradation and increased excitability, and thus not in pre-symptomatic-like conditions. Thus, alteration of transient selection and not steady-state selection in striatum is consistent with enhanced performance of symptomatic Huntington's disease patients compared to controls and pre-symptomatic patients.

Beste et al. (2008) noted that this enhanced response selection was paradoxical, as Huntington's disease patients are consistently worse than age-matched controls across a range of cognitive decision-making tasks (Bamford et al., 1995; Ho et al., 2003; Knopman and Nissen, 1991; Lawrence et al., 1998). Our models offer two potential explanations for why Huntington's disease related changes in striatum are usually associated with cognitive impairment but could also lead to paradoxical cognitive enhancement. First, suppose that all regions of striatum engaged by cognitive tasks implement transient selection. Our model shows that there are limited combinations of NMDA receptor sensitivity increase and cell atrophy where transient selection is enhanced compared to the healthy case; for most combinations transient selection is deteriorated compared to the healthy-state. Thus, one hypothesis is that there is a continuum of NMDA receptor sensitivity increase and cell atrophy across the striatum, and the Beste et al. (2008) task engaged a region of striatum with enhanced transient selection, whereas most tasks engage regions of the striatum with deteriorated transient selection. Second, suppose instead that different regions of striatum use transient or steady-state selection dependent on the local density of projection neuron connections. Our models shows that steady-state selection is always deteriorated by any Huntington's-like change to the striatum. Consequently, this suggests the hypothesis that the Beste et al. (2008) task engaged a region of the striatum using (enhanced) transient selection, whereas most cognitive tasks engage a region of striatum using steady-state selection, and thus are always deteriorated in Huntington's disease patients compared to the healthy-state.

Chapter 9

Discussion

In this thesis, and throughout my publications I have advanced the use of abstract computational models to investigate complex biological systems, examine neurological function, and interpret behavioural data. However, researchers in biological sciences are often sceptical of using *in silico* computational modelling to interpret biological processes. A common objection asserts that all computational models are flawed, and unable to wholly represent a biological process, especially something as multifaceted as Huntington's disease. I will endeavour to show the hollowness of this assertion.

Apart from in silico models, Huntington's disease is often studied using a variety of animal models, which have provided a valuable insight into the pathology and neurology of HD. However, as with in silico models, each non-human model has been demonstrated to be fundamentally flawed. We argue, even though no model is perfect, in silico or in vivo, that sufficiently detailed models can prove to be instructive and play a role in both hypothesis testing, and hypothesis generation (Schneider, 2013). As George E. P. Box once said, 'Essentially, all models are wrong, but some are useful'.

As a biological analogue to our cell atrophy model, excitotoxic models of HD in mice focus on striatal cell loss (Coyle and Schwarcz, 1976; Wang and Qin, 2006). The most prevalent excitotoxic HD mouse model utilises the NMDA receptor agonist, quinolinic acid (Roberts et al., 1993), which presents with selective neuro-degeneration of GABAergic neurons, while largely sparing inter-neurons. Excitotoxic models have contributed to the understandings of a variety of mechanisms related to cell atrophy and age-dependant onset (Roberts et al., 1993; Sun et al., 2003), yet they are fundamentally limited by the ongoing difficulty to study degenerative progression of the disease, as well as failing to replicate the widespread neuropathology observed in the human condition. Our model has been influenced by discoveries from excitotoxic studies in HD, and has in turn exposed theoretically plausible disease-state progressions in the striatum under HD-like conditions.

Similarly, generic models of HD have enabled studies of disease progression using the htt gene (Cepeda et al., 2010; Heng et al., 2008). Genetic reproductions of the htt gene have demonstrated that motor symptoms and cell death might be preceded by neuronal dysfunctions. Problematically however, symptoms develop gradually over many months and do not develop a sufficient expression of disease to examine late stage HD morbidity. Alternatively knock-in models are faithful in terms of genetic causation, but show mild symptoms, which can be confounded by ageing. We can see that murine models are capable of reproducing some symptoms of HD, but are unable to unravel the complex information-processing deficits which coincide with such symptoms. With a disassociated model of time and degradation, we directly simulated degradation in striatum models, to observe neuronal dysfunction with regards to specific desegregate states. This abstraction enables us to test the information-processing capability of tailored degrading neural-networks, which cannot be easily done with mouse models.

In general mouse models have failed to confirm any correlation between neuronal intra-nuclear inclusions, cell death, and the pattern of selective degeneration (Fusco et al., 1999), with the result that the exact role of these phenomena in HD pathogenesis is still unclear. These results have spurred research into non-human primate models of HD, which express more human-like HD symptoms.

Ghiglieri et al. (2012) conclude that selective neuro-degeneration might result not only from the specific distribution of toxic htt products but also from intrinsic neuronal properties, such as those studied in this thesis. We have affirmed that computational models can play an important role in elucidating the causes and contributions of the Huntington's disease phenotype, and as such, can be seen as another useful model to understand the brain and its behaviour, both in hypothesis generation and testing.

From this perspective, using in silico models as part of the arsenal of tools to study complex biological processes, we have contributed, in this thesis, to the understanding of neuron modelling and striatal function. We have provided a mechanistic hypothesis for active striatal compensation mechanisms, with which we interpret recent counter-intuitive behavioural results from Huntington's disease patients.

9.1 Contributions and results

A novel model of synaptic saturation

We have introduced a novel model of synaptic saturation, we termed the *lumped-saturating synapse model*. We have shown that this model accurately reproduces the spiking behaviour of more computationally intensive saturating models, such as kinetics based approaches, which individually track per-synapse saturation (Destexhe et al., 1994; Wang, 1999). We reduce per-synapse models, using a mean-field approximation of per-neuron saturation (Tomkins et al., 2012).

Our saturating synapse model is desirable when modelling receptor types with long binding times, such as the NMDA receptor (Moyer et al., 2007). Saturation incorporates a soft bound on receptor activation resembling soft bounded synapses detailed in Vasilaki et al. (2009a). This model curtails over excitatory firing, to within biological constraints. For our striatum model, lumped-synaptic saturation reduced overall network excitation and brought our synaptic conductance parameters in line with the same order of magnitude as experimentally recorded conductances.

Active selection facilitated by striatum microcircuit

Neural computation is a metabolically expensive task (Laughlin et al., 1998), and the striatum is the largest nuclei in the basal ganglia. Thus, one would expect that the complex striatal microcircuits play an active role in the basal ganglia, yet many computational studies of decision making use models which cannot express active selection.

Our results, using detailed spiking models of the striatal microcircuits, substantiated the hypothesis that the striatum can have an active role in cortical signal processing. We determine that striatal processing is further amplified in the basal ganglia. Our work illustrates that signal selectivity is capable of improving action selection performance in the entire basal ganglia-thalamocortical decision networks. We contrast our work with other computational models of the basal ganglia in action selection and neuro-degeneration disease (i.e, (Grosse-Wentrup and Contreras-Vidal, 2007)), in which the complex recurrent striatal microcircuit is reduced to a two layer feed forward network. This abstraction necessarily removes the emergent selection mechanisms of the striatal microcircuit, and as such must result to implementing selection as a simple arg-max function. One can see that this is a poor approximation of striatal Winner-Take-All style behaviour, devoid of biological realism.

We determine the root causes of emergent selectivity in the striatum, and stress the importance of accurate spiking neuron models. For *transient selectivity*, we demonstrate

that the transient phenomena are due a combined production of the neuron model, combined with network dynamics. The positive transient is generated from adaptation in first-spike generation in already-depolarised neurons. Such adaptation is specific to the SPN model and cannot be reproduced in pyramidal neurons, as the SPN specific parameters are critical. In contrast, the negative transient is formed from network dynamics, suppressing channel firing in response to the positive transient, and thus requires FSI connectivity. Models using rate coded representations, or LIF neurons (i.e. (Grosse-Wentrup and Contreras-Vidal, 2007)), could not see this form of active selection, due to a lack of realistic spiking behaviour, and realistic network interactions.

We have supported the claim that detailed anatomical models are crucial to discover emergent computation in complex tissue structures, providing a strong basis for computational models as tools for both hypothesis generation, and testing.

Selection in the striatum is mediated by network connectivity

The spiking striatal microcircuit supports two distinct signal-selection mechanisms, *transient* and *steady-state selection*. We suggest that active selection in the striatum is achieved by the modulation of the signal salience of cortical input signals. Selection reshapes cortical action-representations to become more favourable for quick and reliable selection.

Theoretical models of action selection in artificial neural networks have suggested that a winner-take-all topology presents an efficient and scalable action selection mechanism (Maass, 2000), with many studies suggesting that the basal ganglia implement a WTA algorithm (Gurney et al., 2001; Joel et al., 2002; Redgrave et al., 1999). WTA selection is often used as a convenient mechanism in 'biologically-inspired' studies, without theoretical confirmation of its existence in the striatum (Lee et al., 1999).

Studies suggested that each SPN receives sufficient recurrent, weak synapses to continuously modulate its ongoing activity (Chuhma et al., 2011; Guzman et al., 2003; Humphries et al., 2010). We have confirmed that the local inhibitory microcircuits in the striatum can support a weak version of WTA-like signal selection, consistent with the weakly-competitive effect between inhibitory populations reported by Yim et al. (2011).

Selection in the striatal microcircuit is mediated by network connectivity density. We demonstrate that distance dependant connectivity (as introduced by Humphries et al. (2010)), with a variable connection-density will elicit a transient selection response to rapid-change cortical salience. However, we further demonstrate that selection exists on a gradient, and thus the distance dependant striatum model can elicit *steady-state selectivity*, with increasing connection density, due to corresponding to increasing network size.

Alternatively random models with a fixed statistical connection-density, also exhibit steady-state selection behaviour, acting as an upper bound for *steady-state selection* in the distance dependant model.

The striatum is known to exhibit connectivity gradients such the dorsal-ventral inter-neuron population gradient (Kubota and Kawaguchi, 1993), and the rostro-caudal gradient of FSI gap junctions (Fukuda, 2009). Thus, it is plausible that a SPN connectivity gradients exists capable of supporting both selection regimes.

Steady-state selection decays in simulated HD conditions

Each selection mechanism emerges from distinct neuron-level and network-level effects. We hypothesized that changes in the striatum would result in differential changes in the expression of each selectivity measure, based on their distinct underlying mechanisms.

Huntington's disease provided to be an instructive case study for neurological changes in the striatum (Fan and Raymond, 2007; Ghiglieri et al., 2012). The Huntington's disease phenotype generally results in diminished action selection capabilities (Freeman et al., 1996; Knopman and Nissen, 1991; Peltsch et al., 2008; Sprengelmeyer et al., 1995; Stout et al., 2001), co-incident with progressive striatal degeneration (Fan and Raymond, 2007; Garside et al., 1996; Leise, 1990; Milnerwood et al., 2010; Portera-Cailliau et al., 1995). We developed a phenomenological model of excitotoxic-induced degradation in the striatum, to test the effects of HD-associated striatal changes upon the intrinsic striatal selection mechanisms.

Excitotoxicity in the striatum is associated with increased NMDA receptor sensitivity (Fan and Raymond, 2007; Milnerwood et al., 2010), and progressive cell atrophy (Leise, 1990; Portera-Cailliau et al., 1995). It is difficult to quantify how NMDA receptor sensitivity co-evolves with cell-atrophy over time, thus we model these symptoms independently. We create an exhaustive 'landscape of degradation' through which a striatum may progress, as a result of disease degeneration. Huntington's disease progression varies between cases (Feigin et al., 1995; Kirkwood et al., 2001; Lemiere et al., 2004), and HD-like degradation can progress differentially throughout the striatum (Kowall et al., 1987; Reiner et al., 1988). Our landscape allows us to consider the gamut of possible paths that degradation could proceed through.

Interestingly, we find that, no matter the degradation-trajectory, *steady-state selectivity* always declines, with the healthy network representing peak selective expression. This is in line with the majority of experimental findings (Bamford et al., 1995; Feigin et al., 1995;

Lemiere et al., 2004; Paulsen, 2011; Reilmann et al., 2011). While behaviour can degrade at different rates, generally all progressions incur a performance decline.

Our models predict that reduced selectivity in the striatum, will lead to a decline in the functionality of the basal ganglia decision circuit. Reduced steady-state selection reduces the absolute, and relative signal salience of action-associated cortical channels. A reduction in the absolute salience of cortical signals will lead to slower action selection (Liu and Pleskac, 2011), due to the increased time needed to integrate cortical signals to a level sufficient enough to suppress SNr output. Our behavioural predictions of reduced reaction times match general experimental results in visual (Peltsch et al., 2008), and attentional (Sprengelmeyer et al., 1995) tasks. Peltsch et al. (2008), have shown that HD patients exhibit both longer and more variable saccadic reaction times. They continue to report that increasing reaction times and error rates were highly correlated with disease severity, reflecting the prediction of our degradation landscape.

Further, to this, our model suggests that a reduction in the relative salience, due to a weakening of the suppression effect, will reduce the likelihood of clean-action selection. Behaviourally this could result in hesitation, involuntary behavioural effects (Reilmann et al., 2011; Swerdlow et al., 1995) and reduced behavioural-timing (Freeman et al., 1996). Confounding both of these, we hypothesize that learning in the striatum will be impaired by reduced selectivity, due to decaying firing rates.

Standard models of plasticity, such as STDP strengthen connections between incidentally firing neurons. Under these learning strategies, steady-state selectivity would serve to reduce the firing rates, of the least salient actions. This implies that the potential for strengthening the least salient connections is reduced. This would further suggest that steady-state selection gates action selection learning to concern primarily the selected action. In this case, positive, dopamine modulated reinforcement learning will act to either strengthen or weaken only the action which lead to action, and as such, to dopamine release. Experimentally our predictions of reduced learning may be reflected by the results of Stout et al. (2001).They found that people with HD may have had difficulties learning or remembering gambling strategies, and comment that their findings are consistent with current models of frontal-subcortical circuits. I suggest that this would be a fruitful avenue for future research in the further work section.

We suggest that characteristic severity-linked mental and behavioural decline generally seen in Huntington's disease patients, is influenced by degrading information processing capabilities in the HD-afflicted striatum. Our *steady-state selectivity* measure provides one metric to quantify this progressive loss of signal selection, as decline in steady-state selectivity is consistent with behavioural decline in patients.

Transient selection could explain paradoxical behavioural improvement under HD conditions

Alternatively, we demonstrated that less densely connected regions of the striatum, can support a weaker form of selection, termed *transient selection*. Transient selection enacts a temporary modification the cortical inputs. We observed that rapidly increasing the ongoing input to one of two competing populations of projection neurons caused a transient peak of activity in that population, and a synchronous transient dip in activity of the other.

We illustrate that transient section is elicited due to two components. The positive transient is a result of adaptation in the single neuron model, with step input causing fast first-spiking behaviour, followed by reduced firing probability. This result suggests a statistical dependence, such that across the striatal channels, there must exist a subset of neurons sufficiently depolarised to create this synchronous fast-firing response. We determined that the presence of transient selection in the basal ganglia-thalamo-cortical loops improved the speed at which the whole circuit resolved a competition between salient inputs. Transient selectivity simultaneously improved the circuit's ability to persist with the selected input, reducing the noise in signal selection. We demonstrate these effects in two parametrisations of the thalamo-cortical loop, encoding perfect-integrator or feedback-amplifying thalamo-cortical configurations.

Recent behavioural experiments have indicated that despite overwhelming cognitive and physical decline seen in HD patients, in at least one case, HD patients may out perform their healthy, and pre-symptomatic counterparts. Beste et al. (2008) investigated sensory memory and attentional processes in an auditory distraction paradigm ,with symptomatic and pre-symptomatic Huntington's patients. They found that The symptomatic HD group exhibited better behavioural performance, with lower error rates and shorter reaction times, compared with both healthy controls and pre-symptomatic Huntington's disease patients.

These results were is stark contrast to the established view of a global decline of all affected cognitive abilities, under Huntington's disease conditions (Craufurd and Snowden, 2002; Feigin et al., 1995; Paulsen, 2011), which we have mapped onto models of universal decline of steady-state selectivity under HD conditions.

To investigate a mechanistic explanation of the phenomena of increased behavioural performance in under HD conditions, we subjected our distance dependant models of the striatum to the phenomenological models of Huntington's disease discussed earlier. In contrast to densely connected networks, we found that sparsely connected distance dependant networks generated a HD-degradation landscape which supported areas of both increased and decreases selectivity.

Beste et al. (2008) suggest that changes in the proportions of SPN and FSI populations may sharpen the inhibitory winner-take-all network of the striatum. NMDA receptors play a key role in signal selection in the striatum (Calabresi et al., 2000), and are the most affected by HD alterations, with FSIs being more resistant (Cepeda et al., 2007). Our phenomenological model of Huntington's disease expresses this altered relationship, enabling us to test these mechanistic hypotheses.

Our model is crucially validated by two observations raised by Beste et al. (2008), namely:

- Pre-symptomatic Huntington's disease patients demonstrate no improvement in selection experiments, regardless of their likely increased sensitivity to glutamate, over that of healthy patients,
- improved selection must be first appear in late-stage Huntington's disease, and cannot be a product of previous compensatory mechanisms.

Firstly, The landscape of degradation is validated by the observation that pHD patients do not see the same behavioural improvement as HD patients, regardless of HD-related up-regulated NMDA receptor expression. Our model indicates that increased NMDA sensitivity alone, is not sufficient to improve transient selectivity, in agreement with the results of Beste et al. (2008).

Secondly, Our results support the conclusion that paradoxical improvements must first emerge in late stage Huntington's disease. The degradation landscape provides support for improved transient selection, and thus improved behavioural performance, only with increased NMDA receptor sensitivity is expressed in conjunction with widespread cell atrophy, characteristic of late stage Huntington's disease.

Our models provide a mechanistic insight into a previously unknown compensatory mechanism which manifests only during late-stage Huntington's disease. This mechanism is distinct from other recorded compensatory mechanisms (Beste et al., 2007; Feigin et al., 2006; Paulsen et al., 2004) which are limited to pre symptomatic Huntington's disease patients. It still remains to be seen how the majority of the literature reports widespread cognitive decline, characteristic of steady state selectivity, while sensory memory and attention processes in an auditory distraction paradigm reflected the improved selectivity landscapes seen with transient selectivity. Beste et al. (2008) suggest that the altered intra-striatal neuronal dynamics could produce this patchwork of parallel declines and improvements of cognition functions. We propose two possible solutions to this dichotomy.

One hypothesis proposes that there is a continuum of NMDA receptor sensitivity increase and cell atrophy across the striatum, and the Beste et al. (2008) task engaged a region of

striatum with enhanced transient selection, whereas most tasks engage regions of the striatum with deteriorated transient selection.

Secondly, we could conceive instead, that different regions of striatum use transient, steady-state selection or a combination of the two, dependent on the local density of projection neuron connections.

Problematically, direct experimental observation of transient selection would be challenging. However, as we have confirmed the mechanism underpinning transient selectivity, it could be possible to observe the positive transient through sequential steps of current injection into a single neuron. We thus make the explicit experiential hypothesis, that there exists a regime of long, sequential steps of current into the projection neuron soma that will elicit a rapid burst of two or more spikes followed by slower regular firing. The existence of this regime would support the existence of transient selection within the striatal microcircuit.

9.2 Future Work

In this thesis we used two distinct models of the striatum and the basal ganglia, a rate coded model (Gurney et al., 2015; Humphries and Gurney, 2002), and a spiking neuron model (Humphries et al., 2010). The rate coded model of the basal ganglia was necessary in order to model signal selection throughout the full cortico-striatal decision circuits. However, to examine the contribution of cell dynamics and emergent network behaviour from striatal microcircuits, a more detailed spiking model of the striatum was required. The phenomena of *transient* and *steady-state selectivity* could not have been seen in a purely rate coded model of the striatum. Winner-take-all behaviour such as that seen in *steady-state selectivity* has been theorised to be supportable by the striatum, but a detailed model was required to see the behaviour as an intrinsic property of the network.

Spiking models of Cortico-Striatal pathways

Anatomically connected spiking networks of the striatal microcircuit have shed light on computation emerging from network connectivity, and neural dynamics (Tomkins et al., 2012, 2013). We demonstrated that selectivity in the striatum is propagated through rate coded models of the basal ganglia, improving action selection timing and reliability. Similarly, complex cortical microcircuits carry out extensive computation on multi-modal data in the form of input-sensitive spike trains (Maass and Markram, 2004). We suggest that increasing the biological realism of cortical processing in our model, with spiking neurons Izhikevich

(2004), would further enlighten how striatal dynamics are propagated and processed in cortical microcircuits, in conjunction to striatal processing.

Current spiking models of the cortico-striatal pathways fail to capture the complexity of decision circuit processing covered in Tomkins et al. (2013). Studies often are incapable of investigating these complexities, due to insufficiently complex neuron models (Stewart et al., 2012), or simplified connectivity regimes (Chersi et al., 2013), and thus, although informative in other ways, cannot adequately study the role that active selection in the striatum plays in cortico-striatal dynamics.

We suggest that, in light of the success of anatomical-connectivity models in the striatum, similar connectivity approaches should be applied to further brain regions with complex microcircuit dynamics, to best investigate the interplay of microcircuit-computation and brain-wide selection circuitry. Accurate spiking models can further be used to investigate specific effects on spiking neurons, such as neuromodulation, and spike-dependant plasticity, which can play important roles in dynamic processing, action selection and learning.

Dopamine modulation could dynamically modulate cortical representations

Dopamine is a neurotransmitter often associated with reward-based learning which plays an important role in the basal ganglia. Integration of dopamine and glutamate neuro-transmission within the striatum microcircuit is hypothesized to enable learning and action control by shaping synaptic plasticity and modifying cellular excitability (Shiflett and Balleine, 2011). Dopamine is necessary for motivation and goal directed behaviour (Palmiter, 2008), and is critical to novel action acquisition (Knowlton et al., 1996) and habit formation (Wang et al., 2011). Mechanistically. Previous computation studies of the striatal microcircuit has shown that dopamine modulates dynamic cell assembly in the striatum, indicating a possible role is signal selection and channel gating (Humphries et al., 2009b).

A recent study has shown that tonic levels of dopamine in the striatum can encode a ramping signal that provides a continuous estimate of animal proximity to distal rewards (Howe et al., 2013). Further, to this they found that the reward signal can also be modulated by the relative size of the delivered rewards consistent with previous studies (Hollerman and Schultz, 1998; Tobler et al., 2005). In continuous dopamine measurements during navigation, Howe et al. (2013) have discovered quick remapping of the dopamine gradient to reflect changes is reward position.

We propose that the continuous value-mapping implemented by tonic dopamine in the striatum could provide dynamic contextual integration of reward value and proximity, over the cortical action-representation.

As our model captures how dopamine modulates the excitability of projection neurons and fast-spiking inter-neurons (Mallet et al., 2006; Nicola et al., 2000), we hypothesise that continuous dopamine modulation will actively modify the selection dynamics we have observed in our model. Dopaminergic firing rate modification could allow signal selection to seamlessly integrate continuously varying reward-proximity information, with top-down cortical action salience, to enhance dynamic behavioural switching in a context-responsive manor.

Striatal selectivity could enhance plasticity in cortico-striatal projections

In complement to tonic dopamine in the ventral striatum, dopamine is a critical neuromodulator of short- and long-term plasticity in the cortico-striatal circuits. There is strong evidence that activity in cortico-striatal circuits changes during the learning of novel actions (Knowlton et al., 1996), with dopamine being differentially involved in early and late stages of action learning (Costa, 2007). Striatal plasticity could be critical in motor learning and motor control (Charpier and Deniau, 1997). Excitation of striatal neurons produces a disinhibition in pre-motor networks, thus long term potentiation in excitatory striatal inputs support movement initiation.

In studies of the rat visual cortex, Sjöström et al. (2001) has shown that plasticity changes depend on the exact timings of spike-pairs between pre- and post-synaptic neurons, which can be expressed in models of Spike Timing Dependant Plasticity (STDP) proposed by Fiete et al. (2010). Models of STDP consist of long term depression (LTD) which encodes for uncorrelated spiking, where a post-synapitc cell fires before the pre-synaptic, and long term potentiation (LTP), encoding correlation, where a post-synaptic cell fires in response to pre-synaptic firing. Time between the pairs of spikes modulates the overall weight change. In vitro experiments have shown that cortico-striatal plasticity implements a form of dopamine modulated spike-timing dependent plasticity (STDP) Schulz et al. (2010).

Our spiking model of the striatum, when expanded to include explicit models of plasticity in the cortico-striatal pathway could provide insight into the multiple reinforcement learning algorithms that have been mapped onto the cortico-striatal circuits Ito and Doya (2011). We propose that *Transient* and *steady-state selectivity* will modulate cortico-striatal plasticity, through advantageous suppression and over-expression or cortical inputs.

We hypothesize that *steady-state selectivity* will support novel action-acquisition through suppression of competition, and encourage rapid dopamine-modulated learning, through increased firing rates in salient channels. Steady-state selectivity should act to reduce noisy competition in novel learning situations and facilitate clean-selection. Salient channels are

over-expressed, reducing inter spike intervals, and thus up-regulating STDP weight changes. We hypothesize that transient increases in dopamine, will improve signal-differentiation in steady-state selection, enforcing a stricter winner-take-all policy through mutual inhibition.

Accurate STDP models may predict catastrophic excitotoxicity

STDP models encode for pair-firing protocols well, but fail to capture plasticity dynamics higher-frequency pair-protocols in which it is no longer obvious which 'pairs of spikes should be considered significant (Bi and Wang, 2002; Froemke et al., 2006). Pfister and Gerstner (2006) propose a model of STDP which considers interactions between triplets of spikes to determine weight change dynamics. These experiments show that the 'triplet rule' is capable of reproducing early experimental STDP results (Bi and Poo, 1998), and contemporary high-frequency protocol results (Wang et al., 2005). At low firing rates, the triplet rule expresses Spike Timing Dependant Plasticity consistent with Sjöström et al. (2001), however, at high firing rates, conventional hebbian plasticity prevails (Vasilaki and Giugliano, 2012, 2014), ensuring LTP under high frequency stimulation.

Phenotypic Huntington's Disease presents with increased NMDA receptor expression, and as a consequence, generally increased projection neuron excitability. We hypothesize that the excitotoxicity hypothesis could be intrinsically linked with hebbian dominated high-frequency long-term potentiation seen in the 'triplet-rule' formulation of STDP'. High-frequency firing stimulated by increased NMDA receptor sensitivity would encourage hebbian LTP. We propose that as NMDA receptor sensitivity increases, due to Huntington's Disease pathology, the increased projection neuron excitability drives plasticity up into the high-frequency LTP regions of the triplet-rule, creating a positive feedback loop, which drives abnormal firing behaviour.

Increased, abnormal firing leads to a build-up of calcium in the neuron which will eventually interfere with normal cell pathways, and mitochondrial function (Fan and Raymond, 2007). This will in turn trigger cell apoptosis, or 'programmed cell death'. This potential link between accurate models of spike timing dependant plasticity and programmed cell death would have important experimental and theoretical ramifications. This possibility could open new avenues of research into slowing the progression of Huntington's Disease symptoms, by chemically reducing the potential for runaway-plasticity and abnormal firing feedback in the striatum, by for example, using tonic dopamine antagonists, or inhibition through deep-brain stimulation.

Final Remarks

In this thesis, we have built complex three-dimensional models of neuronal tissue, from synapses, up to substrates. We have shown how the abstract behaviour of sub-neuronal processes can shape, complex network level behaviour, and ultimately enable enhanced cognitive and behavioural performance. Our results demonstrated the explanatory power of using simple computational elements to discover and complex compensatory mechanisms which could play a role in neuro-degenerative diseases, and may offer valuable medical insights, for treatment and diagnosis. In time, with the development of computational neuroscience as a field, I hope that theoretical explorations of emergent behaviour, such as this, may one day help make the leap, from developing computational neuroscience to understanding consciousness. I, for one, welcome our new robot overlords.

Appendix A

How Degrading Networks Can Increase Cognitive Functions.

Lecture Notes in Computer Science
Volume 7552, 2012, pp 185-192
2012

Authors

Adam Tomkins[1]
Mark Humphries[1,2,3]
Christian Beste[4]
Eleni Vasilaki [1]
Kevin Gurney[2]

Institutions

1 Department of Computer Science, The University of Sheffield, United Kingdom
2 Department of Psychology, The University of Sheffield, United Kingdom
3 Department of Cognitive Studies, Ecole Normale Superieure, France
4 Department of Biopsychology, Ruhr-University Bochum, Germany

Abstract

Huntington's is a genetic, progressive neuro-degenerative disease, causing massive network degradation effecting the medium spiny neurons of the striatum in the basal ganglia (a set of sub-cortical nuclei, believed to be critical for action selection). Despite substantial striatal cell atrophy, some cognitive functions have been shown to improve in manifest Huntington's disease patients over healthy and pre-symptomatic Huntington's disease patients. Using a detailed model of the striatal microcircuit, we show that combining current ideas about the underlying causes of the disease could lead to the counter-intuitive result of improved competitive network dynamics for signal selection.

A.1 Introduction

The striatum is the primary input structure to the basal ganglia (BG) system, located in the base of the mammalian forebrain. The BG is strongly connected to the cerebral cortex and the thalamus, and implicated in many crucial functions such as motor control and reinforcement learning.

Huntington's disease (HD) is a neuro-degenerative genetic disorder which manifests with progressive cell death in the striatum. Despite the severe cell atrophy associated with HD, patients demonstrate increased proficiency in a limited set of cognitive tasks. More specifically, Mismatch Negativity (MMN) studies Beste et al. (2008) have shown that patients with symptomatic HD outperform both pre-symptomatic HD patients and healthy control groups, with manifest Huntington's disease patients exhibiting lower error rates for standard and deviant stimuli.

We propose that this increased performance could be explained as a product of the striatal network structure combined with the excitotoxicity hypothesis, which assumes that in HD neuron firing rates progressively increase, eventually leading to cell death. Using a detailed model of the striatal microcircuit taken from Humphries et al. (2010), modified to account for synaptic saturation, we demonstrate in simulations a potential explanation to the paradox of degrading striatal networks' boosted performance in cognitive tasks. We model two

cortical inputs (representing competing actions) to two corresponding sub-groups of the striatal microcircuit, and observe the subgroup mean firing rates (the 'output' of the striatal microcircuit). The microcircuit temporally suppresses the activity of the striatal neurons that receive input from the less salient (weaker) cortical signal, while enhancing the activity of the striatal neurons that receive input from the more salient (stronger) cortical signal. To quantify the level of this temporary enhancement of the firing rate difference between the two subgroups, we introduce a measurement that we term *Transient selectivity*.

Our parametric study results suggest that a pathogenic increase in NMDA receptor expression, while encouraging cell atrophy, can improve the competitive signal dynamics, due to increased neuron activity. Thus, the cognitive performance of the manifest HD patients might be linked to an enhance of the Transient selectivity in the striatum.

A.2 Striatal Model

We simulate a large scale striatum model, which comprises of medium spiny neurons Izhikevich (2003), and models of fast spiking inter-neurons detailed in Humphries et al. (2009b). The model represents a 0.5 mm^3 slice of the striatal microcircuit of an adult rat. The connectivity patters of the network have been derived from neurological data where possible, and represents a biologically relevant three-dimensional connectivity regime reflecting that of the striatal microcircuit. The striatal model implements an on-centre, off-surround feedforward network, and as such there is diffuse connectivity from the cortex to the FSIs and a focussed ('channel-wise') connectivity to SPNs. The basic model neuron used in the network is based on the model neuron proposed in Izhikevich (2003), further extended to encompass the effects of dopamine modulation in Humphries et al. (2009a). For a detailed description of the network, the connectivity and the corresponding parameters we refer the reader to Humphries et al. (2010). To this model, we have added synaptic saturation, explained below.

Striatal neurons receive cortical inputs via multiple synaptic connections. For simulation efficiency, we model these connections as a single bundle for each synaptic input type (AMPA, GABA or NMDA). To increase the biological realism of the striatal model Humphries et al. (2010), we assume starndard synaptic dynamics according to:

$$I_z = \bar{g}_z h_z (E_z - v), \tag{A.1}$$

where z represents one of either AMPA, GABA or NMDA, $\bar{g}_z$ represents the maximal conductance for the input type z, E_z is the reversal potential associated with that type. The

conductance dynamics h_z are modelled by a standard exponentially decaying term, but with the addition of a saturation effect, encapsulated in the term in square brackets below:

$$\dot{h}_z = \frac{-h_z}{\tau_z}, \text{ and } h_z(t) \leftarrow h_z(t) + \left[1 - \frac{h_z(t)}{N_z}\right] F_z(t). \tag{A.2}$$

Parameter τ_z is a time constant, $F_z(t)$ is the total number of pre-synaptic firing events of of type z arriving at any of the neuron synapses of type z at time t, and N_z the maximum number of synapses (or receptor groups). In our simulations, all parameters are as in Humphries et al. (2009a) with the addition of $N_z = 250$.

The synaptic saturation term in square brackets (Eq. A.1, A.2) closely approximates individually saturating synapse neuron behaviour based on kinetic models of receptor binding and unbinding such as those implemented in Destexhe et al. (1994) (simulation data not shown). The approximation is true for neuron firing rates less than 100 Hz. The mathematical form resembles the dynamics of a soft-bounded synapse, see for instance Vasilaki et al. (2009b) and references therein.

The striatal model is written in C with interface to Matlab code. We run our simulations with a discretization time step of 0.1 ms.

A.3 Results

In what follows, we establish a link of the striatal model with HD, introduce signal selectivity and quality measurements, and explore the evolution of the disease in a network representing a virtual.

A.3.1 Linking the Model with Huntington's Disease

theory of excitotoxicity suggests that cell atrophy in the striatum is caused by increased sensitivity of NMDA receptors in the medium spiny neurons. This leads to excessive firing of the neurons, eventually resulting in early cell death (DiFiglia, 1990). Adopting this hypothesis, we model excitotoxicity by up-regulating the NMDA conductivity versus that of AMPA, increasing the excitability of the model cell. In addition to this mechanism, we simulate cell atrophy by randomly removing increasing percentages of medium spiny neurons from the network, severing all relevant connections.

We assume that each striatal sub-population's SPNs A (dashed) and B (solid) are driven by two competing cortical signals a and b correspondingly (see Figure A.1A). These cortical signals conceptually represent two alternative choices. The remaining SPNs receive constant

background noise at 3.5 Hz. Each FSI receives input from the cortex. Initially, population A receives an input $a = 5.0Hz$ while population B receives an input of $b = 3.5Hz$, implying that choice a is selected. Then, 1500 ms afterwards (to allow the network to reach a state of equilibrium), the signal to population B increases to $b = 5.5Hz$ suggesting competition between the two choices with choice b being the more salient, and therefore the 'winning' choice. Finally, choice b reduces to background rate, suggesting that once more choice a is the winner. Figure A.1B shows the mean firing rate response of the striatal sub-populations for a 'healthy' network (0% cell death, normal NMDA:AMPA conductivity ratio of 0.5, according to Humphries et al. (2009a)). Striatal response signals R_A and R_B are calculated over the total spikes per millisecond of each population, using a running mean with a square window of size 100 ms which processes data in both the forward and reverse directions (Matlab command filtfilt). Subjecting the striatal microcircuit to a spectrum of plausible receptor excitability levels and comprehensive degrees of network degradation, we produced a hypothetical landscape through which the striatal microcircuit might progress under progressive HD regimes.

A.3.2 Quantifying Selectivity

Selectivity is defined as the ability to distinguish competing signals in the striatal microcircuit. Here we discuss measurements that quantify such signals for different combinations of cell atrophy and excitability.

Transient Selectivity

We identify two regimes of interest following a switch of the winning signal (for instance at 1500 ms, see Figure A.1A), based on simulation data of the healthy network. The first regime, where the transient phenomena occur, we term *Transient Period* (T = 1500-1700 ms). The second one, when signals seem to have reached a loosely defined 'steady-state', we term the *Stable Period* (S = 1700-1900 ms). We then define Transient selectivity TS as follows:

$$TS = 1 - \frac{\bar{S}^B - \bar{S}^A}{\Delta^{BA}}, \tag{A.3}$$

where Δ^{BA} represents the maximum difference that signals R^B and R^A have in the Transient period (see also Figure A.1B) and $\bar{S}^A$, $\bar{S}^B$ being the temporal means, within the *Stable Period* S, of the signals R^A and R^B correspondingly. For simplicity, we present the definition as specific to the first signal inversion (1500 ms) but it can be generalised to capture the second signal inversion (with appropriate re-labelling).

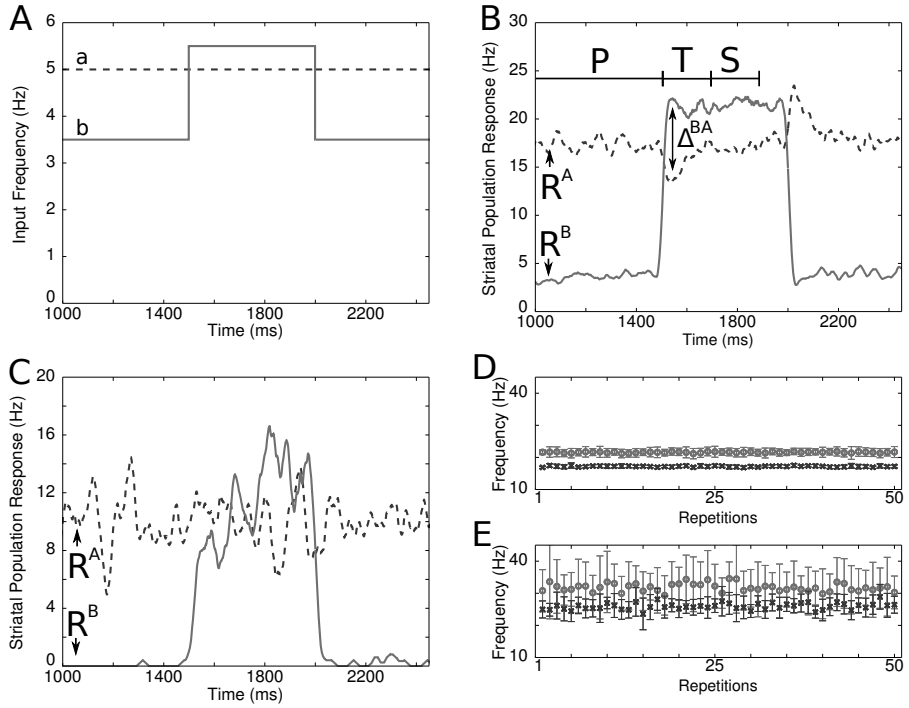

Fig. A.1 **(A)** Selection Protocol: The cortical input to each striatal population, driving signal A (*dashed*) and signal B (*solid*).**(B)** The healthy network response: Cell atrophy = 0%, NMDA:AMPA ratio = 0.5. The mean firing rates of the striatal sub-populations are depicted as R^A (*dashed*) and R^B(*solid*). The signal responses are divided up into three defined areas: Primed signal period $P = 1000$ to 1500 ms, before signal competition is introduced. Transient signal period $T = 1500$ to 1700 ms, where signal transients are induced. Stable Signal period $S = 1700$ to 1900 ms, corresponding loosely to a steady-state (no transients). **(C)** An extreme atrophy case with Cell atrophy = 90%, NMDA:AMPA ratio = 0.5. Large noise-driven fluctuations in the signal and slow system response are exhibited under high cell atrophy situations, without high receptor excitability. **(D,E)** Margin of Separation: The mean S^A (x marker) and S^B (o marker), over 50 repetitions. Error bars show standard deviation. (D) shows the Margin of Separation fo the healthy case, and (E) shows an extreme case (Cell Atrophy = 95%, NMDA:AMPA ratio = 1). The Margin of Separation between S^A and S^B is reduced, despite high NMDA receptor expression.

Signal Degradation

Figure A.1C shows the striatal population response of a 'virtual patient' with NMDA:AMPA conductivity ratio of 0.5 (normal condition) but high cell atrophy of 90%, representing cell atrophy inflicted without excitotoxicity. We can see a substantial difference in these signals versus the signals of a healthy network. As the cell atrophy increases, the reliability of the signal representations becomes diminished, generally exhibiting large noise driven fluctuations. We therefore measure the quality of the striatal signals using a measure of the coefficient of variation by defining the signal Degradation D as:

$$D = \frac{\sigma^P}{\mu^P}, \tag{A.4}$$

where μ and σ correspond to the temporal mean and variance of the signal over the P = 1000-1500 ms, that we term *Primed period*, before the cortical signals change. This is to capture the noise lever in the signal without taking competition-related effects into account.

Separation Margin

With increasing cell atrophy, we observe that the distance between the firing rates of the two populations in the Stable Period S decreases. We measure the *Margin of Separation M* of the population responses by calculating the temporal means and variances of the signals R^A and R^B in the period S, for 50 repetitions. Figure A.1D,E illustrate this process. In Figure A.1D we show the results for 50 repetitions of the simulation for a healthy network and in Figure A.1E for a network corresponding to a last stage HD patient (Cell Atrophy = 95%, NMDA:AMPA ratio = 1). The variance in each of these repetitions comes from many factors such as randomness in firing times, removing neurons from the simulation (cell death) etc. We then average the mean and standard deviation metrics over the 50 repetitions of the simulation, resulting in an average mean and standard deviation of the signals in the Stable Period, that we term μ^A, σ^A, μ^B, σ^B. We calculate the Margin of Separation between these average signals as $M = (\mu^B - 2\,\sigma^B) - (\mu^A + 2\,\sigma^A)$.

A.3.3 The Landscape of Selectivity

We mimic the excitotoxicity hypothesis by up-regulating the NMDA conductivity versus that of AMPA from 0.5 (health subject) to 1, consequently increasing excitation in the model cell. In addition to this mechanism, we simulate cell death by randomly removing from the simulation increasing percentages of medium spiny neurons, from 5% up to 95%. As we

do not know the exact progression of the intertwined mechanisms, we implement the two independently, exploring the 2D parameter space, covering different possible progressions of the disease. Each point in this 2D space is calculated by averaging over 50 repetitions of the simulation (with identical parameters).

Figure A.2A shows the evolution of the signal Degradation D of the striatal population responses for different combinations of cell atrophy and receptor excitability. The plot shows that, with increasing cell death while keeping the neuron excitability low, Degradation of the signal increases, representing cell atrophy caused distinct from, or weakly linked to excitotoxicity. However for high excitability, the higher firing rates (which are captured by the denominator of the Degradation) lead to signals of better signal quality, as increasing cell excitability compensates for rising cell death. Figure A.2B shows Transient selectivity, again averaged over 50 repetitions. We observe that, despite the cell death, as the NMDA:AMPA conductivity increases, Transient selectivity improves, an improvement which we attribute to higher firing rates of the network. We have looked into combinations of high excitability and high cell death and found that in individual cases, the suppression mechanism is impaired in the sense that the Transient selectivity measurement reflects only the enhancement of the 'winning' signal, without suppression of the 'losing' signal. In Figure A.2C we show the Margin of Separation. The plot shows that despite the transient enhancement of the winning signal, with increased cell atrophy, the firing rates of the two population are not always distinguishable in the Stable Period. Finally in Figure A.2D we combine the results of all previous figures, filtering the Transient selectivity landscape (Figure A.2B) by imposing a cut-off in all parameter combinations where the degradation $D > 0.3$ or the margin of separation $M < 0$. White coloured areas indicate impaired signal selectivity. This combined landscape suggests that there is a regime where, despite the increasing cell death, signal selectivity is effectively improved versus a healthy network. Results presented here refer exclusively to the first signal inversion (at 1500 ms), but we have checked that results stay qualitatively the same for the second signal inversion (2000 ms).

A.4 Conclusions

It was generally assumed that, with widespread degradation of the striatal micro-circuit, a universal decline in cognitive abilities would ensue. Recent data Beste et al. (2008) alter this view, indicating that symptomatic HD patient outperform both pre-symptomatic HD patients and healthy control groups in Mismatch Negativity studies. To investigate this unexpected result, we establish a link between the striatal microcircuit model in Humphries

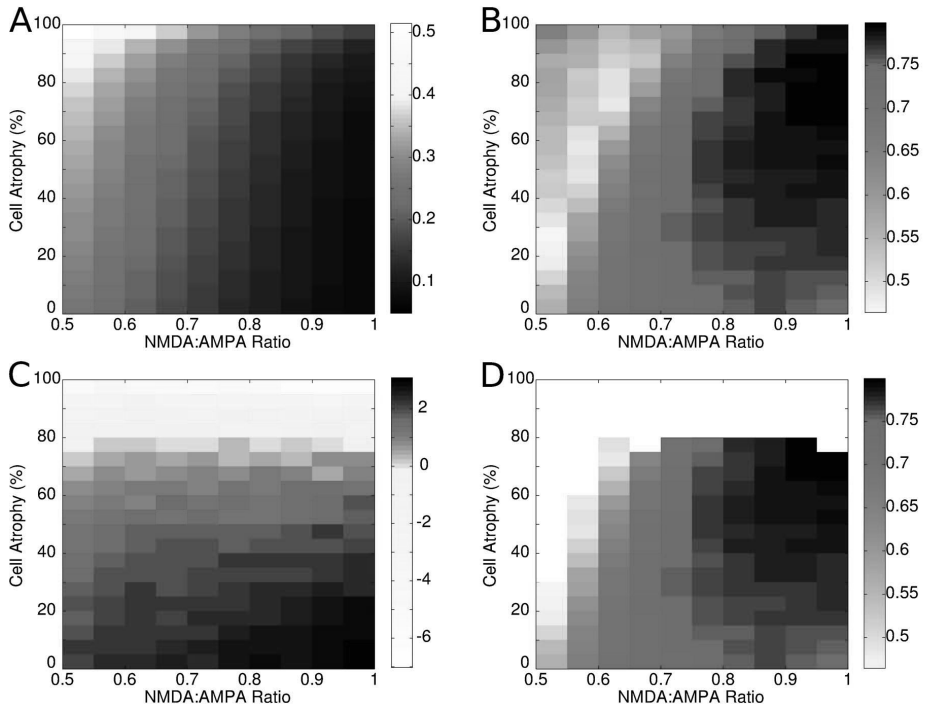

Fig. A.2 The 2D 'Selectivity Landscape' for combinations of increasing cell atrophy (0% - 95%) and excitability(NMDA:AMPA conductivity ratio 0.5 - 1) for metrics defined in section A.3.2. Light colour indicates values of poor performance while dark colour indicates good performance (Minimal signal Degradation, improved Transient selectivity and a high Margin of Separation). (A) Degradation D for signal R^A averaged over 50 trials. The Degradation measure is applied only to the Population response of A during the Primed period. (B) Transient selectivity averaged over 50 trials (C) Separation Margin. (D) Transient Selectivity in B truncated such that areas for which Degradation $D > 0.3$ or Separation Margin $M < 0$ have been removed (white coloured areas).

et al. (2010) and Huntington's disease via a set of simulations. Our parametric study suggest that progressive cell death, combined with the excitotoxicity hypothesis might paradoxically improve aspects of signal selection and representation in the striatum. We expect that the observed transient phenomena will be amplified in the greater cortical-striatal loops, encouraging faster and more reliable behavioural switching in specific auditory selectivity tasks.

Appendix B

Transient and Steady-state Selection in the Striatal Microcircuit

Frontiers in Computational Neuroscience
2013

Authors

Adam Tomkins[1]
Mark Humphries[1,2,3]
Christian Beste[4]
Eleni Vasilaki [1]
Kevin Gurney[2]

Institutions

1 Department of Computer Science, The University of Sheffield, United Kingdom
2 Department of Psychology, The University of Sheffield, United Kingdom
3 Department of Cognitive Studies, Ecole Normale Superieure, France
4 Department of Biopsychology, Ruhr-University Bochum, Germany

Abstract

Although the basal ganglia have been widely studied and implicated in signal processing and action selection, little information is known about the active role the striatal microcircuit plays in action selection in the basal ganglia-thalamo-cortical loops. To address this knowledge gap we use a large scale three dimensional spiking model of the striatum, combined with a rate coded model of the basal ganglia-thalamo-cortical loop, to asses the computational role the striatum plays in action selection. We identify a robust transient phenomena generated by the striatal microcircuit, which temporarily enhances the difference between two competing cortical inputs. We show that this transient is sufficient to modulate decision making in the basal ganglia-thalamo-cortical circuit. We also find that the transient selection originates from a novel adaptation effect in single striatal projection neurons, which is amenable to experimental testing. Finally, we compared transient selection with models implementing classical steady-state selection. We challenged both forms of model to account for recent reports of paradoxically enhanced response selection in Huntington's disease patients. We found that steady-state selection was uniformly impaired under all simulated Huntington's conditions, but transient selection was enhanced given a sufficient Huntington's-like increase in NMDA receptor sensitivity. Thus our models provide an intriguing hypothesis for the mechanisms underlying the paradoxical cognitive improvements in manifest Huntington's patients.

B.1 Introduction

Finding the neural substrate for the process of “selection" is key to furthering our understanding of decision-making (Ding and Gold, 2013), action selection (Grillner et al., 2005; Mink, 1996), planning (Houk and Wise, 1995), action sequencing (Jin and Costa, 2010), and even working memory (Gruber et al., 2006). A unifying proposal is that the basal ganglia forms just such a generic selection mechanism (Prescott et al., 1999; Redgrave et al., 1999); this proposal neatly explains why the basal ganglia have been hypothesised to contribute to each of these functions. But specifying the computational process of selection by the basal ganglia is challenging (Berns and Sejnowski, 1998; Gurney et al., 2001; Leblois et al., 2006).

A particular unknown is the computational role of the basal ganglia’s input nucleus, the striatum. The striatum’s GABAergic projection neurons comprise the vast majority of cells and are connected by local collaterals of their axons (Wilson and Groves, 1980). The lack of layers or of clear axial preferences in the direction of dendrites or axons suggests that striatal tissue is homogeneous in all three dimensions (Humphries et al., 2010). Such GABAergic connectivity naturally lends itself to the idea that the striatum forms a vast recurrent network that, locally, implements a winner-takes-all computation (Alexander and Wickens, 1993; Fukai and Tanaka, 1997; Wickens, 1997). The weak strength of synapses between the projection neurons (Czubayko and Plenz, 2002; Jaeger et al., 1994; Tunstall et al., 2002) is difficult to reconcile with this proposal (Plenz, 2003), as they suggest projection neuron output can only modulate ongoing activity and not outright inhibit their targets.

Here we report an alternative form of transient selection that can occur in weak, sparse networks of striatal projection neurons. Using our three-dimensional network model of distance-dependent connections in the striatal microcircuit (Humphries et al., 2009b, 2010), we explored the effect on striatal output of competing inputs to two projection neuron populations. We found that rapidly stepped input to one population caused a transient competitive effect on the two populations’ outputs, which disappeared after around hundred milliseconds. In response to the same inputs, we also found that sufficiently dense striatal connectivity could result in steady-state competition, where the post-step equilibrium activity of each population reflects the inhibition of one by the other.

To compare transient and steady-state selection we challenged both forms of model to account for the paradoxical response selection results of Beste et al. (2008). They found that manifest Huntington’s disease patients were both faster and less error prone than controls on a simple two-choice reaction-time task. As Huntington’s disease primarily results in striatal damage, this suggests the hypothesis that changes in the striatum directly affect response

selection. We expand on the role of the striatum in signal selection, by describing a framework for signal selection that may account for both the typical decline in performance for most tasks under Huntington's conditions Ho et al. (2003), as well as a mechanism for increased performance under the same conditions. We thus explored how Huntington's disease-like changes to our striatum models could affect both transient and steady-state selection, and sought whether the effect on either form of selection could explain the results of Beste et al. (2008), while also accounting for the usual cognitive impairment in Huntington's disease (Ho et al., 2003; Lawrence et al., 1998).

B.2 Material & Methods

We study here an updated version of our prior full-scale model of striatum (Humphries et al., 2009b, 2010). Compared to those models, the model here brings together the three-dimensional anatomy model from (Humphries et al., 2010) with an updated version of the dopamine-modulated projection neuron model from (Humphries et al., 2009b).

B.2.1 Spiking neuron models

The basic model neuron used in the large scale striatal model is derived from the model neuron proposed in (Izhikevich, 2003), which was extended to encompass the effects of dopamine modulation on intrinsic ion channels and synaptic input in (Humphries et al., 2009b).

In the biophysical form of the Izhikevich model neuron, v is the membrane potential and the "recovery variable" u is the contribution of the neuron class's dominant ion channel:

$$C\dot{v} = k(v - v_r)(v - v_t) - u + I \tag{B.1}$$

$$\dot{u} = a[b(v - v_r) - u] \tag{B.2}$$

with reset condition

$$\text{if } v > v_{peak} \text{ then } v \leftarrow c,\ u \leftarrow u + d$$

where in the equation for the membrane potential (B.1), C is capacitance, v_r and v_t are the resting and threshold potentials, I is a current source, and c is the reset potential. Parameter a is a time constant governing the time scale of the recovery due to the dominant ion channel. Parameters k and b are derived from the I-V curve of the target neuron behaviour, where b

describes how sensitive the recovery variable u is to fluctuations in the membrane potential v. Parameter d describes the after spike reset of recovery variable u, and can be tuned to modify the rate of spiking output.

Projection neuron model

The projection neuron models' parameter values and their source are given in Table B.1. Parameters C, d, v_t, and the AMPA synaptic conductance g_{ampa} (see below) were found by searching for the best-fit to the f-I curve and spiking input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model (Humphries et al., 2009b).

In Humphries et al. (2009b) we showed how this model can capture key dynamical phenomena of the projection neuron: the slow-rise to first spike following current injection; paired-pulse facilitation lasting hundreds of milliseconds; and bimodal membrane behaviour emulating up- and down-state activity under anaesthesia and in stimulated slice preparations.

Fast-spiking interneuron model

For the FSI model, Equation B.2 for the u term is given by (Izhikevich, 2007b)

$$\dot{u}_{fs} = \begin{cases} -au_{fs} & \text{if } v_{fs} < v_b, \\ a[b(v_{fs} - v_b)^3 - u_{fs}] & \text{if } v_{fs} \geq v_b, \end{cases} \tag{B.3}$$

which enables the FSI model to exhibit Type 2 dynamics, such as a non-linear step at the start of the current-frequency curve between 0 and 15-20 spikes/s. Further discussion on the FSI model used in the striatal microcircuit can be found in (Humphries et al., 2009b); the FSI model parameters are reproduced in Table B.2.

Dopaminergic modulation of intrinsic ion channels

Tonic levels of dopamine in the striatum modulate the excitability of the projection neurons and fast-spiking interneurons (Mallet et al., 2006; Nicola et al., 2000). Our network model incorporates modulation by tonic dopamine through the relative activation levels of D1 and D2 receptors. These levels are modelled using the method proposed in (Humphries et al., 2009b), in which complex membrane dynamics are subsumed by linear transforms with only two parameters ϕ_1, $\phi_2 \in [0, 1]$, describing the proportion of D1 and D2 receptor activation, respectively. Throughout we used $\phi_1 = \phi_2 = 0.3$.

For activation of D1 receptors on projection neurons we used the linear mappings:

$$v_r \leftarrow v_r(1 + K\phi_1) \tag{B.4}$$

Table B.1 Intrinsic parameters for the projection model

Parameter	Value	Source
a	0.01	(Mahon, 2000);(Izhikevich, 2007b)
b	-20	(Izhikevich, 2007b)
c	-55 mV	"
k	1	"
v_r	-80 mV	"
v_{peak}	40 mV	"
C	15 pF	(Humphries et al., 2009b)
v_t	-30 mV	"
d	91	"
K	0.0289	"
L	0.331	"
α	0.032	"

and

$$d \leftarrow d(1 - L\phi_1), \tag{B.5}$$

which respectively model the D1-receptor mediated enhancement of the inward-rectifying potassium current(KIR) (B.4) and enhancement of the L-type Ca^2+ current (B.5).

For activation of D2 receptors on projection neurons we used the linear mapping:

$$k \leftarrow k(1 - \alpha\phi_2) \tag{B.6}$$

which models the small inhibitory effect on the slow A-type potassium current, increasing the neuron's rheobase current (Moyer et al., 2007).

With these mappings, the model neuron is able to accurately capture the effect of D1 or D2 receptor activation on both the f-I curves and spiking input-output functions of the Moyer et al. (2007) compartmental model of the projection neuron.

Dopamine modulated fast spiking inter-neurons in the striatal network only express the D1-family of receptors (Centonze et al., 2003). Activation of this receptor depolarises the neuron's resting potential (see Humphries et al., 2009b for further details). Thus we used the linear mapping of the resting potential:

$$v_r \leftarrow v_r(1 - \eta\phi_1) \tag{B.7}$$

Table B.2 Intrinsic parameters for the fast spiking interneuron model. Dimensions are given where applicable. See Humphries et al. (2009b) for details.

Parameter	Value	Source
a	0.2	Izhikevich (2007a)
b	0.025	"
d	0	"
k	1	"
v_{peak}	25 mV	"
v_{b}	-55 mV	"
C	80 pF	Tateno et al. (2004)
c	-60 mV	"
v_{r}	-70 mV	"
v_{t}	-50 mV	"
η	0.1	fitted to Bracci et al. (2002)
ε	0.625	fitted to Gorelova et al. (2002)

B.2.2 Synaptic models

Synaptic input comprises the source of current I in Equation B.1:

$$I = I_{ampa} + I_{gaba} + B(v)I_{nmda}. \tag{B.8}$$

where I_{ampa}, I_{gaba} , I_{nmda} are current input from AMPA, GABA and NMDA receptors respectively, and $B(v)$ is a term that models the voltage-dependent magnesium plug in the NMDA receptors. Compared to the projection neuron, FSIs receive no NMDA receptor input from cortex, have a moderately larger AMPA conductance (Table 2), but do receive input via local gap junctions (see below).

Each synaptic input type z (where z is one of ampa, nmda, gaba) is modelled by

$$I_z = \bar{g}_z h_z (E_z - v), \tag{B.9}$$

where $\bar{g}_z$ is the maximum conductance and E_z is the reversal potential. We use the standard single-exponential model of post-synaptic currents

$$\dot{h}_z = \frac{-h_z}{\tau_z}, \quad \text{and} \qquad h_z(t) \leftarrow h_z(t) + S_z(t), \tag{B.10}$$

where τ_z is the appropriate synaptic time constant, and $S_z(t)$ is the number of pre-synaptic spikes arriving at all the neuron's receptors of type z at time t.

Given that one interest here is in the possible roles of striatal NMDA sensitivity in Huntington's disease, we paid careful attention to two complexities of the NMDA receptor: its non-linear voltage-gating, and its saturation. The term $B(v)$ in Equation (B.8), which models the voltage-dependent magnesium plug in the NMDA receptors, is given by (Jahr et al., 1990)

$$B(v) = \frac{1}{1 + \frac{[\mathrm{Mg}^{2+}]_0}{3.57}\exp(-0.062v)}, \tag{B.11}$$

where $[\mathrm{Mg}^{2+}]_0$ is the equilibrium concentration of magnesium ions.

As glutamate can remain locked into the NMDA receptor for 100 milliseconds or more (Lester et al., 1990), so the pool of available receptors becomes rapidly saturated at high afferent firing rates. To capture this we introduce a mean-field model of synaptic saturation where we interpret the term h_z in Equation B.10 as the number of active receptor groups over the whole neuron. Each step in h_{nmda}, following a number of spikes $S_{nmda}(t)$, activates that number of receptor groups, which decays with a time constant τ_{nmda}. To introduce saturation, we bound the size of the step by the proportion of available groups. Together, these concepts give us the model:

$$\dot{h}_{nmda} = \frac{-h_{nmda}}{\tau_{nmda}}, \quad \text{and} \qquad h_{nmda}(t) \leftarrow h_{nmda}(t) + \left[1 - \frac{h_{nmda}(t)}{N_{nmda}}\right] S_{nmda}(t). \tag{B.12}$$

As well as introducing this saturation of the NMDA synapses, we also removed the $1/\tau_s$ scaling of post-synaptic current amplitude used in (Humphries et al., 2009b). This allowed the model synaptic conductances to be the same order of magnitude as their experimental counterparts. Consequently, we re-tuned g_{ampa} by fitting the input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model, following the protocol in (Humphries et al., 2009b). We obtained equally good fits to those found previously with a value of $g_{ampa} = 0.4$ (results not shown).

Dopaminergic modulation of synaptic input

Following the projection neuron models in (Humphries et al., 2009b), we add D1 receptor modulation of NMDA receptor evoked EPSPs by

$$I^{\mathrm{D1}}_{\mathrm{nmda}} = I_{\mathrm{nmda}}(1 + \beta_1 \phi_1), \tag{B.13}$$

Table B.3 Synaptic and gap junction parameters for the striatal network.

Parameter	Value	Source and notes
E_{ampa},E_{nmda}	0 mV	Moyer et al. (2007)
E_{gaba}	-60 mV	Moyer et al. (2007)
τ_{ampa}	6 ms	Moyer et al. (2007)
τ_{nmda}	160 ms	Moyer et al. (2007)
τ_{gaba}	4 ms	Moyer et al. (2007)
τ FSI gap	5	fitted to (Galarreta and Hestrin, 1999)
$[Mg^{2+}]_0$	1 mM	Jahr et al. (1990)
g_{ampa} Ctx-SPN	0.4 nS	tuning (see main text)
g_{ampa} Ctx-FSI	1 nS	Fits linear rise in EPSC data from (Gittis et al., 2010)
g_{nmda} Ctx-SPN	0.2 nS	Fixed by maintaining the 2:1 AMPA:NMDA ratio from (Moyer et al., 2007)
g_{gaba} SPN-SPN	0.75 nS	Koos et al. (2004)
g_{gaba} FSI-SPN	3.75 nS	mean 5-fold increase compared to SPN-SPN (Koos et al., 2004); 3$\times$ increase of PSP (Planert et al., 2010)
g_{gaba} FSI-FSI	1.1 nS	Gittis et al. (2010)
g FSI gap	5 nS	fitted to (Galarreta and Hestrin, 1999)
β_1	0.5	tuning (see main text)
β_2	0.3	tuning (see main text)

and we add D2 receptor modulation of AMPA receptor evoked EPSPs by

$$I_{\text{ampa}}^{\text{D2}} = I_{\text{ampa}}(1-\beta_2\phi_2), \tag{B.14}$$

where β_1 and β_2 are scaling coefficients determining the relationship between dopamine receptor occupancy and the effect magnitude (Table B.3). Due to the addition of saturating NMDA synapses, we also re-tuned these parameter values by fitting the input-output functions of the Moyer et al. (2007) 189-compartment projection neuron model under D1 and D2 receptor modulation of synaptic inputs, following the protocol in (Humphries et al., 2009b).

Finally, following the model in (Humphries et al., 2009b), we add D2 receptor modulation of GABAergic input to FSIs by

$$I_{\text{gaba}}^{\text{fsi}} = I_{\text{gaba}}(1-\varepsilon_2\phi_2). \tag{B.15}$$

Gap junctions

A gap junction between FSIs i and j is modelled as a compartment with voltage v_{ij}^*, which has dynamics

$$\tau\dot{v}_{ij}^* = (v_i - v_{ij}^*) + (v_j - v_{ij}^*), \tag{B.16}$$

where τ is a time constant for voltage decay, and v_i and v_j are the membrane potentials of the FSI pair. The current introduced by that cable to the FSI pair is then

$$I^*_{\text{gap}}(i) = g(v^*_{ij} - v_i) \qquad I^*_{\text{gap}}(j) = g(v^*_{ij} - v_j), \tag{B.17}$$

where g is the effective conductance of the gap junction. The total gap junction input I_{gap} to a FSI is then the sum over all contributions I^*_{gap}.

B.2.3 Striatum network model

Our model captures the connections within the GABAergic microcircuit in striatum, illustrated in Figure B.2. We simulated a large-scale model representing a three-dimensional cuboid of the striatum in the adult rat at one-to-one scale, containing every projection neuron and fast-spiking interneuron present in the biological tissue. We used a density of 89000 projection neurons per mm^3 (Oorschot, 1996) and a FSI density of 1% (see Humphries et al., 2010 for discussion). We assumed projection neurons were evenly split between D1 and D2 receptor dominant types, and without any spatial bias. Hence we randomly assigned half of the projection neurons to be D1-type and half to be D2-type.

In the Results we predominantly report the results of simulations using a 300 μm on the side cube, giving 2292 projection neurons and 23 FSIs. Other sizes are noted explicitly where used.

To connect the neurons we used two different models. In the *physical* model we used distance-dependent functions for probability of connection between each element of the microcircuit. These functions were derived from overlap of dendritic and axonal arbors, and are given in (Humphries et al., 2010) for each connection type in the microcircuit.

In the *random* model we ignored distance, and simply made connections to each neuron at random until the correct number of incoming connections of each type was made. The target number of connections were derived from the mean values obtained from the central neurons of the three-dimensional connectivity model in (Humphries et al., 2010), and taken from column 1 of Table 5 in that paper: SPNs $\rightarrow$ 1 SPN: 728; FSIs $\rightarrow$ 1 SPN: 30.6; FSIs $\rightarrow$ 1 FSI: 12.8; FSI gap junctions per FSI: 0.65.

B.2.4 Selection competitions

Cortical input to the model was designed to emulate the response selection component in a general two-choice task, where a (possibly noisy) stimulus taking one of two values is observed over time and a choice made between the two corresponding responses. In such

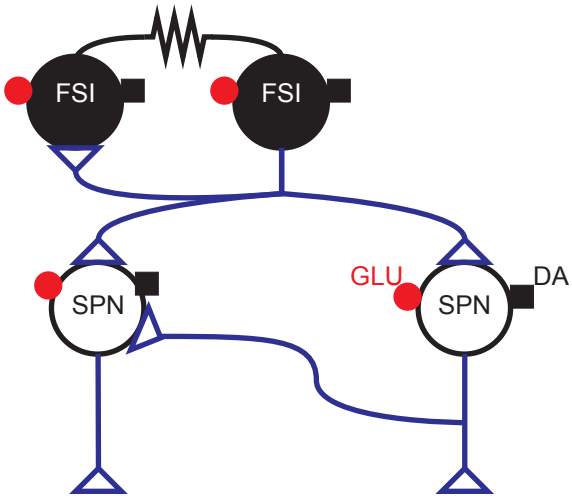

Fig. B.1 **GABAergic striatal microcircuit**. Input to the striatum comes from glutamatergic (GLU: •) fibres originating in the cortex, thalamus, hippocampal formation and amygdala, and dopaminergic (DA: ■) fibres from brainstem dopaminergic neurons. The projection neurons (SPNs) are interconnected via local collaterals of their axons projecting to other nuclei of the basal ganglia. The fast-spiking interneurons (FSIs) can form dendro-dendritic gap junctions between them and are also connected by standard axo-dendritic synapses. All these intra-striatal axo-dendritic connections (△) are GABAergic and hence inhibitory.

a task, we propose that the two responses are made salient by the onset of each trial; and after a perceptual decision is made about the stimulus value, then the corresponding response increases in salience. This generic setup was inspired by the experimental procedures of Beste et al. (2008), in which participants were asked to distinguish between short (200 ms) and long (400 ms) auditory tones, using a distraction paradigm. Inputs followed a ramping trajectory to simulate evidence accumulation and increasing decision confidence (Asaad et al., 2000). We previously showed that transient selection can be seen in response to stepped cortical inputs (Tomkins et al., 2012).

The striatum model was divided up into three populations, two physically close SPN populations representing the two competing responses, which we refer to throughout as *channels*, and the remaining background neurons given a constant input. Neurons were randomly divided into the two channels, with 40% of the neurons in channel 1 and 2 respectively, and the remaining 20% of cells were labelled “background" neurons.

The input protocol is illustrated in Figure B.2A. Each response population received a priming input at a background rate for 1500 ms, causing them to reach a steady state of firing activity. At 1500 ms, channel 1, (black) received a ramping input for a time of 50ms, raising the salience towards a new steady state, when it became the most salient cortical input to the

striatum. During the 50 ms ramping time, channel 2 also received a ramping input, matching that of channel 1 for 25 ms. Following this, the signal to channel 2 decreased back to the background rate, describing the evidence accumulation trajectory of an out-competed action.

Rates were specified for each cortical spike train input to each projection neuron and FSI model. Both neuron models received the equivalent of 250 input spike trains (see Humphries et al., 2009b for details).

We measured how the striatal microcircuit performed channel wise signal selection on the cortical inputs, using this simple protocol, inspired by the auditory decision task performed in Beste et al. (2008). However due to the abstract nature of the input protocol we use, applied to a generic simulation of the striatal microcircuit, the selection measured in these results could be applied to any channel-wise decision task throughout the striatum, and is not limited to auditory processing.

B.2.5 Metrics for selection

We define “selectivity" in the striatum as the ability to robustly distinguish competing signals. The striatum demonstrates two complementary modes of selectivity, which we measure with different metrics. These selection metrics are applied to the output of each channel, which is characterised by a zero-phase filtered mean firing rate.

Transient Selectivity

Given a competitive split in cortical input, we see a temporary boosting of the most-salient signal, accompanied with a temporary suppression of the least-salient competitive signal (Figure B.2C). This transient phenomena presents a boost of the difference in salience between the two competing signals. We identify two key regimes: (i) $\Delta^{S(1,2)}$, the maximum difference between the two signals during the transient peaks; (ii) S^1, S^2, the mean stable activity level of each channel after the transient period dissipates. The total transient selectivity, between 0 and 1, is defined as

$$TS = 1 - \frac{S^1 - S^2}{\Delta^{S(1,2)}}, \qquad 0 \leq TS \leq 1 \tag{B.18}$$

$\Delta^{S(1,2)}$ is calculated as the *maximum* difference between the firing rates of Channel 1 and Channel 2 over the transient window ($t = 1500 : 2000$ ms). This enables the measure to allow for cases in which the largest perturbations from the mean are not temporally coincident, either due to reliable intrinsic dynamic properties of the network, or statistical fluctuations therein.

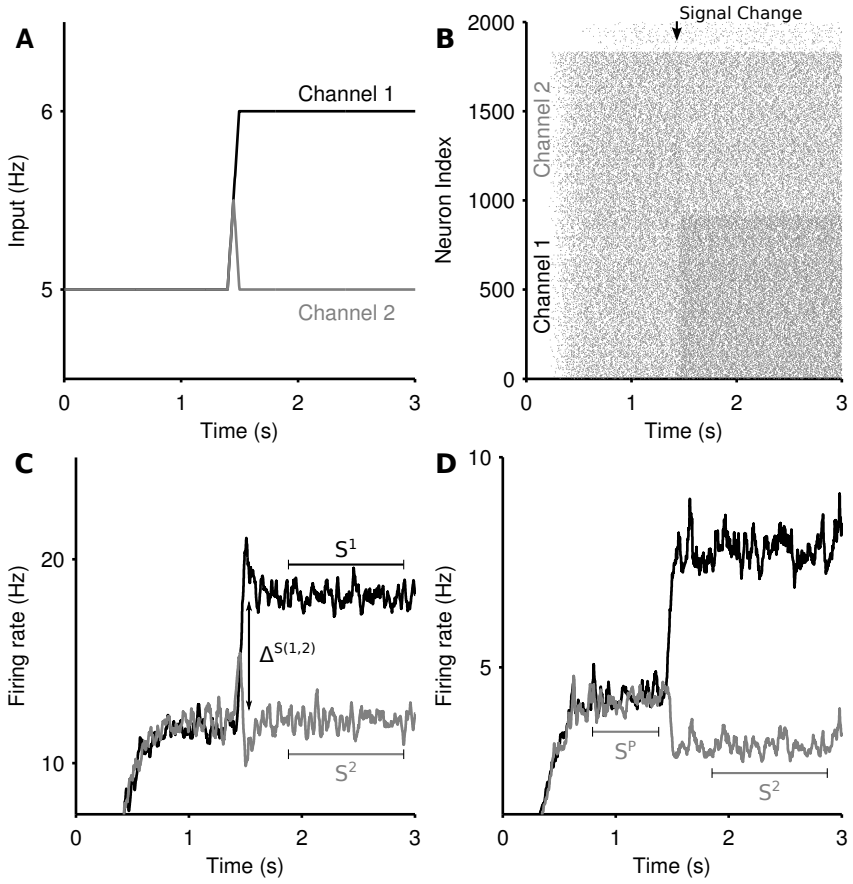

Fig. B.2 **Measures of selectivity in striatal output**. **(A)**. Ramping cortical input into the striatum model. Two channels are driven by input spike trains, demonstrating signal selection between most-salient (channel 1) and least-salient (channel 2) striatal signals. **(B)**. Raster plot of the striatum microcircuit output for a single selection experiment. Increased firing can be seen in channel 1 at the onset of the ramped input in panel A. **(C)** A sample striatal output of the *physical* network, showing a zero-phase filter of the mean spiking output from the two competing channels in response to the ramped input in panel A. Annotations demonstrate the measures used in the transient selectivity measure. S^1,S^2: stable firing rate; $\Delta^{S(1,2)}$: maximum of the difference between the two channels firing rates over the transient period. **(D)** A sample striatal output of the *random* network, in response to the same input. Annotations demonstrate the measures used in the steady-state selectivity measure. S^P: pre-step stable firing rate.

Steady-state selectivity

The striatum network can exhibit signal suppression on its least-salient channel due to sustained inhibition by the most salient channel. Steady-state selectivity is measured on the least-salient channel, as the percentage reduction in the mean channel firing rate after the rise in salience of the most-salient signal. An example of steady-state selectivity in the *random* network can be seen in Figure B.2D. We define (S^P) as the stable firing rate of the primed channel 2 before the increase in competition, and from this we calculate the steady-state selectivity (SS) as:

$$SS = 100 \times \left(1 - \frac{S^2}{S^P}\right). \tag{B.19}$$

B.2.6 Basal ganglia-thalamocortical loop model of transient selection

To study the contribution of the transient striatal dynamics to the selection mechanism of the whole basal ganglia, we used the population-level implementation of our basal-ganglia thalamo-cortical loop model (Humphries and Gurney, 2002). Figure B.3 schematically illustrates the loop model, and the connectivity of the response-representing populations.

The average activity a of all neurons comprising a channel's population changes according to

$$\tau \dot{a} = -a(t) + I(t) \tag{B.20}$$

where τ is a time constant and I is summed, weighted input. We used $\tau = 10$ ms throughout. The normalised firing rate y of the unit is given by a piecewise linear output function

$$y(t) = F(a(t), \theta) = \begin{cases} 0 & a(t) \leq \theta \\ a(t) - \theta & \theta < a(t) < 1 - \theta \\ 1 & a(t) \geq 1 - \theta \end{cases} \tag{B.21}$$

with threshold θ.

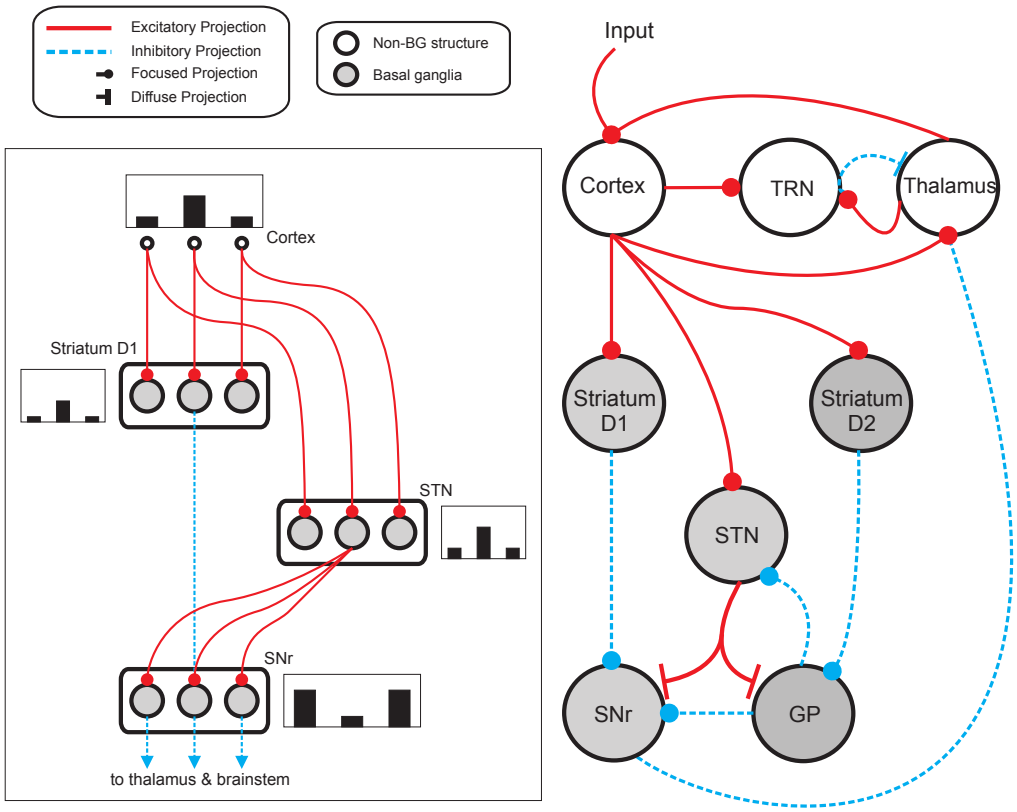

Fig. B.3 **Basal ganglia thalamo-cortical loop model**. The main circuit (right) embeds the basal ganglia into a thalamo-cortical feedback loop. Each nucleus contains multiple response-representing populations. Within the basal ganglia, the circuit can decomposed into an off-centre, on-surround network (left): three populations are shown, with example activity levels in the bar charts to illustrate the relative contributions of the nuclei. Note that, for clarity, full connectivity is only shown for the second population. Briefly, the selection mechanism works as follows. Constant inhibitory output from substantia nigra pars reticulata (SNr) provides an 'off' signal to its widespread targets in the thalamus and brainstem. Cortical inputs representing competing saliences are organised in separate populations, which project to corresponding populations in striatum and subthalamic nucleus (STN). The balance of focussed inhibition from striatum and diffuse excitation from STN results in the most salient input suppressing the inhibitory output from the corresponding SNr population, signalling 'on' to that SNr population's targets. Tonic dopamine levels in the striatum set the ease with which the channels are selected, and subsequently switched between following further salient inputs. For quantitative demonstrations of this model see (Gurney et al., 2001; Humphries and Gurney, 2002). GP: globus pallidus; SNr: substantia nigra pars reticulata; STN: subthalamic nucleus; TRN: thalamic reticular nucleus.

The following describes net input I_i and output y_i for the i^{th} channel of each structure, with n channels in total. The full model was thus given by (Humphries and Gurney, 2002):

$$
\begin{aligned}
\text{Cortex:} \quad I_i^{ctx} &= y_i^{thal} + c_i, \\
y_i^{ctx} &= F(a_i^{ctx}, 0), \\
\text{Thalamus:} \quad I_i^{thal} &= y_i^{ctx} - y_i^{SNr} - 0.1 y_i^{TRN} - 0.7 \sum_{j \neq i}^{n} y_j^{TRN}, \\
y_i^{ctx} &= F(a_i^{thal}, 0), \\
\text{TRN:} \quad I_i^{TRN} &= y_i^{thal} + y_i^{ctx}, \\
y_i^{TRN} &= F(a_i^{TRN}, 0), \\
\text{Striatum D1:} \quad I_i^{d1} &= y_i^{ctx}(1 + \lambda_1), \\
y_i^{d1} &= F(a_i^{d1}, 0.2), \\
\text{Striatum D2:} \quad I_i^{d2} &= y_i^{ctx}(1 - \lambda_2), \\
y_i^{d2} &= F(a_i^{d2}, 0.2), \\
\text{Subthalamic nucleus:} \quad I_i^{stn} &= y_i^{ctx} - y_i^{gp}, \\
y_i^{stn} &= F(a_i^{stn}, -0.25), \\
\text{Globus pallidus:} \quad I_i^{gp} &= 0.9 \sum_{j}^{n} y_j^{stn} - y_i^{d2} \\
y_i^{gp} &= F(a_i^{gp}, -0.2), \\
\text{SNr:} \quad I_i^{snr} &= 0.9 \sum_{j}^{n} y_j^{stn} - y_i^{d1} - 0.3 y_i^{gp}, \\
y_i^{snr} &= F(a_i^{snr}, -0.2),
\end{aligned}
$$

Net input was computed from the outputs of the other structures, except driving input c_i to channel i of cortex. The striatum was divided into two populations, one of projection neurons with the D1-type dopamine receptor, and one of projection neurons with the D2-type dopamine receptor. Many converging lines of evidence from electrophysiological and anatomical studies support this functional split into D1- and D2-dominant projection neurons and, further, that the D1-dominant neurons project to SNr, and the D2- dominant neurons project to GP (Gerfen et al., 1990; Matamales et al., 2009; Surmeier et al., 2007).

In line with the projection neuron model described above, the model included opposite effects of activating D1 and D2 receptors on striatal projection neuron activity: D1 activation facilitated cortical efficacy at the input, while D2 activation attenuated this efficacy

(Humphries et al., 2009b; Moyer et al., 2007). The mechanism for this mirrored that of the spiking projection neuron model in using simple linear factors. Thus, if the relative activation of D1 and D2 receptors by tonic dopamine are $\lambda_1, \lambda_2 \in [0,1]$, then the increase in efficacy due to D1 receptor activation was given by $(1+\lambda_1)$; the decrease in efficacy due to D2 receptor activation was given by $(1-\lambda_2)$. Throughout we set $\lambda_1 = \lambda_2 = 0.2$, simulating tonic levels of dopamine.

The negative thresholds ensured that STN, GP, and SNr have spontaneous tonic output (Humphries et al., 2006). We simplified the model here compared to Humphries and Gurney (2002) by delivering input only to cortex, to represent the salience-driven response selection, rather than to cortex, striatum and STN; both models gave qualitatively the same results. We used exponential Euler to numerically solve this system, with a time-step of 1 ms.

We used $n = 8$ channels in total, with two of those channels (4 and 5) receiving non-zero inputs, mimicking the input protocol used for the striatal network model, which is designed to abstractly simulate the two choice reaction-time task performed in Beste et al. (2008). Baseline inputs $c_4 = c_5 = 0.3$ were delivered at simulation onset. A step in input c_5 occurred between 100 and 200 time-steps: a small step of $c_5 = 0.5$ or a large step of $c_5 = 0.7$. The ability for the model to select was assessed during this step period. As in prior models (Berns and Sejnowski, 1998; Gurney et al., 2001; Humphries and Gurney, 2002; Humphries et al., 2006), selection was assessed by observing the change in activity on each SNr channel, as this output provides the tonic inhibition of thalamic and brainstem structures and is thought to gate the execution of actions (Redgrave et al., 1999). Here, successful selection of a channel was defined as the SNr output falling to zero.

Modelling transient selection in the rate-coded model

We mimicked the ability of the striatum microcircuit to produce transient phenomena using an input injection into the striatum of the rate coded model. At $t = 100$ we injected external inputs into each striatal channel in the model, forcing a transient increase or decrease as appropriate in the corresponding channels. Transient sizes were extracted from the striatal microcircuit traces, and reproduced in the rate coded model. Individual transients were calculated as the percentage change in the firing rate of the circuit during the transient period compared to the stable firing rate achieved post-transient. This allowed us to gauge the role of the complex striatal dynamics, generated by our microcircuit model and responsible for the transient selection mechanism, on the selection properties of the entire basal ganglia-cortex loop.

B.3 Results

In what follows we discuss the simulation results of our model and interpret them as potential mechanisms explaining the findings of Beste et al. (2008). We discuss two types of potential selection mechanisms that we have termed *transient* and *steady-state*.

B.3.1 Transient selection by the striatum

Transient selection emerges from the striatal microcircuit

We sought insight into the potential for competition within the striatum by examining the dynamics of our three-dimensional network model. We first explored the effect on striatal output of competing inputs to two projection neuron populations. These inputs were intended to emulate the changes in cortical signals representing two alternative responses in a generic two-choice decision-making task.

Figure B.4A shows the mean firing rate of each channel from the same example simulation. After the divergence in inputs at $t = 1.5$ s, a transient increase of the firing rate is elicited in channel 1, the most salient population, and a transient suppression of the firing rate is elicited in channel 2. This transient suppression occurs despite no change in the input to channel 2. Moreover, this population rapidly returns (~ 100 ms) to its pre-step firing rate. Consequently, we termed this phenomenon *transient* selection.

We found that the elicited transient selection was robust over a wide range of choices for the baseline input rate and the signal difference between the two channel inputs after the signal divergence. Figure B.4B shows that transient selection could be robustly elicited for any step size over 0.5 Hz when the baseline input rate exceeded ~ 4 Hz.

Transient selection is due to both circuit and intrinsic membrane properties

We further investigated the mechanisms underlying the positive and negative transient changes in population activity. We found that the positive transient was produced by single neuron dynamics, whereas the negative transient was due to network connectivity. This can be seen in Figure B.5A-B, where lesioning either the projection neuron connections or all the network connections abolished the negative transient but did not prevent the positive transient.

To confirm the positive transient was a single neuron phenomenon, we simulated an individual projection neuron model receiving many trials of the same stepped input protocol, and averaged its responses. The resulting peri-stimulus time histogram (Figure B.5C) shows that the neuron had a clear transient increase in firing probability immediately after the step of input. Running the same test on a model of a cortical regular-spiking pyramidal neuron, with

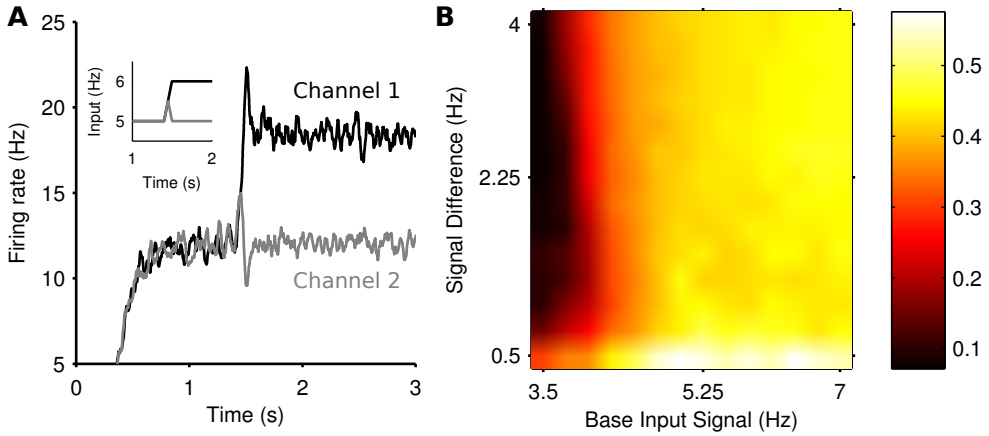

Fig. B.4 **Transient selection of competing input signals by the striatum. (A)**. Mean firing rate of the two output channel populations in the experiment in response to the ramped input protocol (inset); individual spike trains have been convolved with a zero-phase digital filter to create smooth firing rates without lag. **(B)**. Mean transient selection landscape colour coded such that brighter colours represent higher selectivity. Landscape shows the mean transient selectivity averaged over 30 trials as a function of base input signal and step in signal difference during competition.

input scaled to produce approximately the same steady-state rates, showed no such transient increase in firing probability after a step in input (Figure B.5D). Thus the transient increase in population activity observed in a single trial of the network is a statistical phenomenon of synchronous spiking of many projection neurons, and seemingly dependent upon properties particular to the striatal projection neuron.

We sought to elucidate these properties by injecting sequential current steps directly into the projection neuron model and observing the behaviour of the membrane voltage v and slow current u. Figure B.5E shows that a step in current applied to an already depolarised membrane triggers a rapid double spike, followed by slower regular spiking. Figure B.5F plots the corresponding trajectory of the slow current u: the initial depolarising injection makes the slow current u increasingly negative, thus slowly charging the membrane potential v (Figure B.5E; see Equation B.1). The subsequent step of injected current increases the membrane potential rapidly, and the contribution of the large, negative u ensures a rapid pair of spikes time-locked to the current step. However, once spiking has been initiated, the equilibrium value of u is less negative than immediately before the current step. Consequently, the smaller contribution of the slow current u ensures a comparatively slow spike rate in the steady-state.

To show that the slow current u is critical, we examined the dependence of this spiking "adaptation" on the parameters of the slow current. We repeated the sequential-step current injection protocol for a range of step-sizes, and measured the adapting response as $f_{ratio} = F_{first}/F_{last}$, the ratio of the first and last inter-spike intervals after the current step. A value of $f_{ratio} > 1$ thus indicates an adaptation. We found that the adaptation response appeared with a second current step above ~ 50 pA (blue curves in Figure B.5G,H). Figure B.5G shows that the adaptation response disappeared if we reduced the effective time constant of the slow current (increased a), allowing the slow current to recover faster after spiking. Figure B.5H shows that the adaptation response also disappeared if we reduced the gain b of the slow current The transient phenomena thus depends critically on the slow current u.

As lesioning only the connections between the projection neuron could abolish the negative transient (Figure B.5A), this suggested it arose from a network effect where the neurons contributing to the positive transient inhibited their targets. To test this observation, we simulated the model with lesioned projection-neuron collaterals for a range of baseline input firing rates and step sizes (protocol in Figure B.2A) and computed the size of the negative transient that resulted. Figure B.5I shows that the negative transient was indeed abolished for a wide-range of values for the input firing rates. However, a sufficiently large baseline firing rate and step in firing rate could still result in a negative transient (upper-right corner of Figure B.5I). Thus, it seems that sufficient cortical drive of the FSI population (which inhibits the projection neurons) also contributes to the negative transient in projection neuron population activity.

Transient selection is sufficient to alter decision making performance

Though the previous result demonstrates the existence and origin of transient selection within the striatum, it is not sufficient to show a causative effect of transient selection on decision-making. To address this issue, we asked whether such transient signals in the striatum could enhance the selection of input signals by the basal ganglia circuit. Here we consider selection to mean that the output of a substantia nigra pars reticulata (SNr) population falls from its tonic rate to zero. In particular, we hypothesised that the transient signals in striatum would be amplified in the complete basal-ganglia-thalamo-cortical loop, and thus directly influence the output of the basal ganglia.

To test this, we used our rate-coded model of population activity in the basal ganglia-thalamocortical loop (Humphries and Gurney, 2002). The model received inputs to two populations of cortico-striatal neurons (Figure B.6A), mimicking the protocol used in our full-scale striatum model. An example of the subsequent SNr outputs are illustrated in Figure

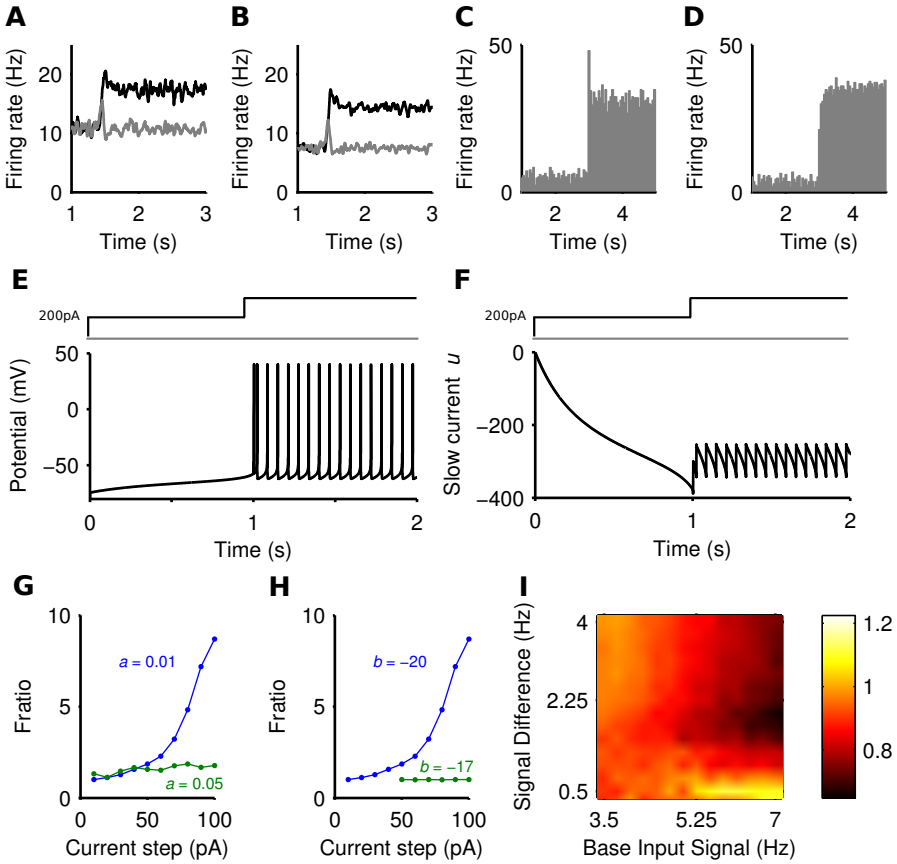

Fig. B.5 **Sources of the positive and negative transients.** **(A)**. Striatal output with lesioned projection neuron connections. **(B)**. Striatal output with all intra-striatal connections lesioned. **(C)** Peri-stimulus time histogram of a single projection neuron output, averaged over 50 steps of spiking input from $r = 4$ Hz to $r = 7.2$ Hz (onset $t = 3$s), exhibiting transient behaviour. **(D)**. Peri-stimulus time histogram of a single regular-spiking cortical neuron model, averaged over 50 steps of spiking input from $r = 0.75$ Hz to $r = 3$ Hz, with no transient behaviour. Model parameters given in (Izhikevich, 2007a). **(E)** The membrane potential (v) of the projection neuron model in response to a depolarising current injection (200 pA) followed by a further step in current at 1 second. **(F)** The corresponding changes in the slow current (u). **(G)** f_{ratio} in the projection neuron model as a function of current step size and slow current decay constant $1/a$ ms. **(H)** f_{ratio} in the projection neuron model as a function of current step size and slow current gain b. **(I)**. The effect of projection neuron connection lesions on the negative transient. Landscape of negative transients measured as ratio of the maximum negative transient peak over the steady state, plotted as a function of base input rate vs signal difference.

B.6B. At the time of the step in input to one population, we emulated the subsequent transient signals observed in our full-scale model by brief injections of further increased input to that striatal population and decreased input to the other. These correspondingly produced small, brief positive and negative transients in the output of those striatal populations, for both D1 and D2-type projection neurons (Figure B.6C-D). Note that the subthalamic nucleus populations also received the cortical input signals, but not the transient signals.

We found that a small positive transient elicited in the striatal population was sufficient to change the speed and persistence of selection (Figure B.6E-H). Figures B.6E,F show that signal selection was maintained for longer with increasing transient sizes. Correspondingly, Figures B.6G,H show that increasing the size of transients injected into the model striatum decreased the time to selection. These changes were found irrespective of the size of input step, or of the closed-loop gain g of the positive thalamocortical feedback loop (Chambers et al., 2011) (When g=1, this loop is a perfect integrator, while with g=2, there is an amplifying feedback loop.) Thus, transient signals in the striatum are sufficient to modulate selection by the basal ganglia.

B.3.2 Steady-state selection by the striatum

Prior debates about selection in the striatum have focussed on winner-takes-all modes of computation (Plenz, 2003; Wickens, 1997). In order to compare transient selection with this more common form of selection computations, we sought to understand whether our striatal model could show winner-takes-all types of dynamics; here we refer to these as “steady-state" selection, in contrast to “transient" selection, as the competition between inputs causes persistent changes to output firing rates.

Steady-state selection in a randomly-connected model

Neurally-inspired models of winner-take-all dynamics are often based on fully-connected or dense randomly-connected networks (Alexander and Wickens, 1993; Fukai and Tanaka, 1997; Hartline and Ratliff, 1958; Mao and Massaquoi, 2007; Yim et al., 2011). We thus simulated our striatal model with random connectivity, in which each neuron type received, on average, the same number of connections, and the connections were made by choosing source neurons at random from across the three-dimensional cuboid. The target number of connections was based on the expected number of connections of a projection neuron and FSI in the centre of a 1 mm^3 network, according to the computational anatomical estimates of Humphries et al., 2010 (see Methods). In this way, the randomly-connected model was more densely connected relative to the distance-dependent model. Thus, while closer to

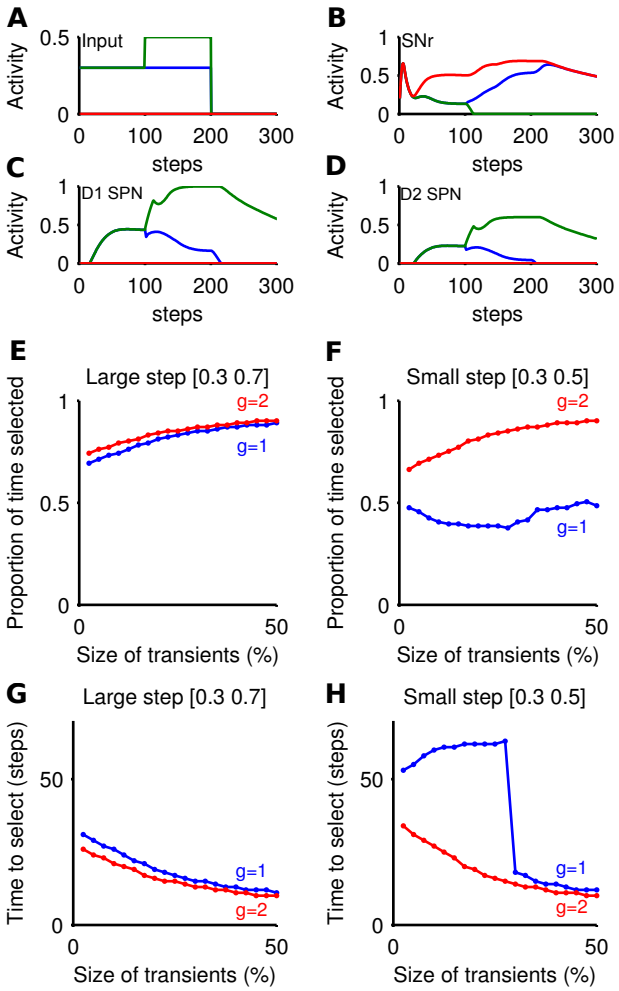

Fig. B.6 **Transient selection in striatum is amplified by basal ganglia-thalamo-cortical loop**. Panels **(A)-(D)** show an example simulation of the loop model that included emulation of the transient selection signals originating in the striatum (transient size: 50%; thalamo-cortical loop gain g=2). **(A)** Cortical input to the rate-coded model, mimicking the selection protocol used in the striatal microcircuit selection experiments. **(B)** Corresponding SNr output response for three populations: no input (red); baseline only (blue); and baseline-plus-step (green). The input step thus caused clear selection by forcing the SNr output to zero. **(C)** Evoked response in the rate coded striatal D1 neurons, showing the effect of the injected transient at $t = 100$. **(D)** Evoked response in the rate coded striatal D2 neurons. **(E)** Proportion of time an action was selected, as a function of transient size. Transient size is expressed as a proportion of the steady-state firing rate achieved without the transient. Step values indicate the cortical input before and after the step in input. Parameter g: closed-loop gain of the thalamocortical loop. **(F)** Proportion of time an action was selected, given a small input step. **(G)** Time delay before selection achieved, as a function of transient size, for large input step. Delay is given between the step in cortical input and the corresponding SNr population reaching zero output. **(H)** Time delay before selection achieved, as a function of transient size, for small input step.

the topology usually studied for steady-state selection, the randomly-connected model still retained connection statistics consistent with the estimates obtained in Humphries et al., 2010.

We tested the randomly-connected model with the same stepped input protocol as the physically-connected model (Figure B.2A). Figure B.7A shows an example of the mean population firing rates in the randomly-connected striatum model, with evident steady-state selection: the population receiving the stepped cortical input increases its firing rate, and the other population correspondingly decreases its firing rate despite receiving the same input throughout. We found that the magnitude of steady-state selection was dependent on the size of the baseline firing rate and input step. FigureB.7B shows that the most effective steady-state selection occurred for low baseline rates and large input steps, approaching a winner-takes-all like response of nearly complete suppression ($\sim 80\%$) of the losing population's activity.

Figure B.7B shows that lesioning the connections between projection neurons prevents steady-state selection. Figure B.7D shows that lesioning the FSI input to the projection neurons reduces but does not eliminate the steady-state selection, while also reinstating a transient period. This suggests that mutual intra-channel inhibition by the projection neurons populations is responsible for the suppression effect seen in both the *random* and the larger *physical* networks.

Distance-dependent connectivity can support steady-state selection

To assess if such steady-state selection required homogeneous, random connectivity of the kind described above, we checked whether such selection could be found in the physical model of connectivity. Again using the same stepped input protocol, we simulated physical networks up to 1 mm^3, in order to increase the density of connectivity within the centre of the network, which scales with the number of neurons in the model (Figure B.9B).

Figure B.9A shows that steady-state selection could be observed for distance-dependent connectivity, given a sufficiently large model (here 1 mm^3). We found that the magnitude of steady-state selection increased monotonically with increasing network size (Figure B.9D), approaching the steady-state selectivity seen in the *random model*. Figure B.9B,C shows that in the *physical* model as the number of neurons increases as a function of network size so does the average number of connections each projection neuron receives. By contrast, the *random model* always has the same density of connections. The *physical* model's correspondence between the number of connections to a projection neuron and the effectiveness of steady-

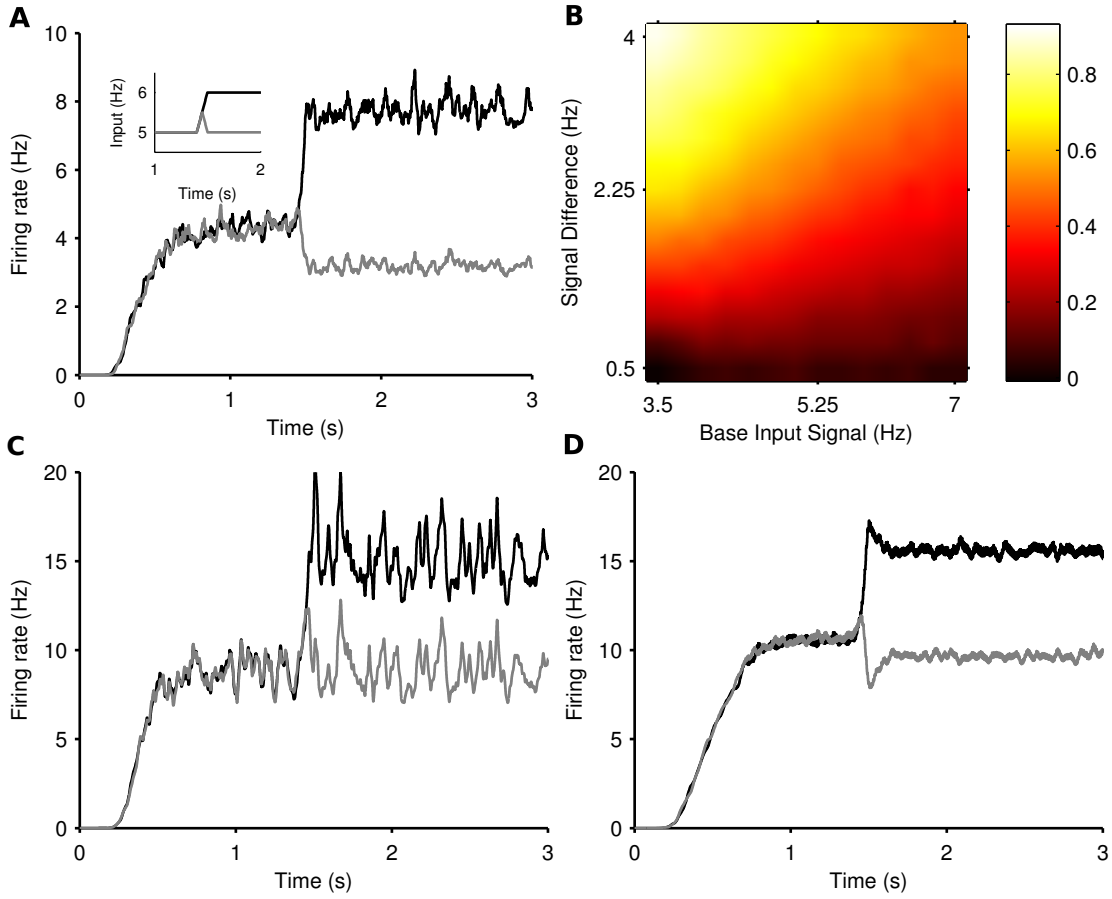

Fig. B.7 **Steady-state selection in the randomly-connected striatum model.** **(A)** Smoothed mean firing rates of two projection neuron populations, in response to the ramped input protocol (inset). **(B)** The magnitude of steady-state selection as a function of baseline input and step size. The magnitude gives the fall in firing rate of the losing population as a proportion of its pre-step firing rate. Each magnitude is an average over 15 simulations. **(C)** Smoothed mean firing rates of two projection neuron populations, with SPN-SPN connections lesioned, in response to the same input as above. Steady-state selectivity is removed. **(D)** Smoothed mean firing rates of two projection neuron populations, with FSI-SPN connections lesioned, in response to the same input as before. Steady-state selection remains.

state selection suggests that such selection is dependent on the density of connections between projection neurons.

The model further suggests that it is only the increased density of connections that is key, and not an increase in recurrent connections between projection neurons. Figure B.9E shows the absolute number of recurrent connections in the *physical* and *random* network configurations. Note that the number of bi-directional connections in the *random* network drops of as a function of network size due to the fact that each neuron receives a fixed number of connections regardless of the network size. By contrast we see a small rise in the number of bi-directional connections in the *physical* model. However, Figure B.9F shows that in both random and physical networks the proportion of connections that are bi-directional falls with increasing network size. Thus, the increased effectiveness of steady-state selection is likely due to increased absolute connection density and not increased recurrent connections.

B.3.3 Comparing selection mechanisms: paradoxical selection enhancement in Huntington's disease

We had established that two contrasting forms of selection can be supported by the striatal circuit, depending on the type and density of connectivity. We then sought insight into how the two forms of selection could be distinguished. In particular, we hypothesised that they would make different predictions about how changes to the striatum would alter response selection. In order to test this hypothesis, we sought an experimental data-set that could provide a basis for testing our predictions.

Beste et al. (2008) have recently shown a rare example of paradoxical cognitive enhancement in a neurological disorder. They reported that manifest Huntington's disease patients had faster and less error prone response selection on a simple two-choice auditory task than controls or pre-manifest Huntington's disease patients. As Huntington's disease is primarily characterised by widespread loss of striatal projection neurons(FSI populations have been shown to be more resistant to HD-modifications (Ghiglieri et al., 2012)), and increased sensitivity of NMDA receptors on striatal projection neurons (Fan and Raymond, 2007), these results suggest the hypothesis that one or both of these changes to the striatum lead to enhanced selection, and as such we look into excitotoxicity as a possible candidate for the paradoxical improvements investigated.

We thus simulated both transient and steady-state selection under Huntington's-like changes to the striatal model, and searched for evidence of enhanced selection. We emulated increased NMDA receptor sensitivity by increasing the conductance of the NMDA synapse (we report this as the ratio of the NMDA:AMPA conductances), and separately emulated the

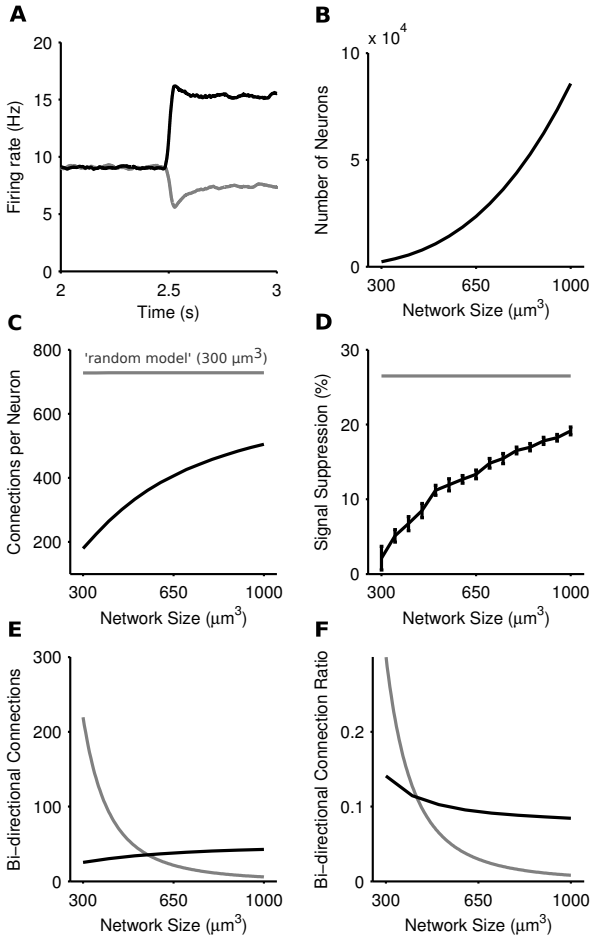

Figure 8. Steady-state selection in the physical model of the striatal microcircuit. **(A)** Mean firing rate of two projection neuron populations in a 1 mm^3 model, with 89,749 total simulated neurons. **(B)** Number of simulated neurons as a function of network size. **(C)** Average number of connections per neuron as a function of network size. The *physical* network (black) approaches the density of connections seen in the random network (grey) with increased network size. **(D)**Magnitude of steady-state selection as a function of network size. All simulations used the inputs [5,6] Hz. Magnitude is the percentage suppression in the average firing rate of the losing channel after the competitive signal onset ($t = 2.5$ s). Shown in grey is the steady-state selectivity seen in the random model for a network of size 300 μm^3. Bars set at $\pm$ 2 s.d, computed over 15 repeats. **(E)** Number of bi-directional connections as a function of network size. The total number of pairs of reciprocal connections in the *physical* model are shown in black, and the *random* model in grey. Bi-directional pairs decrease in the physical model with increasing network sizes, due to the fixed number of connections each neuron receives. **(F)** The ratio $R_{bi}$of bi-directional connections to the total number of connections a neuron makes for the *physical* model (black) and the *random* model (grey).

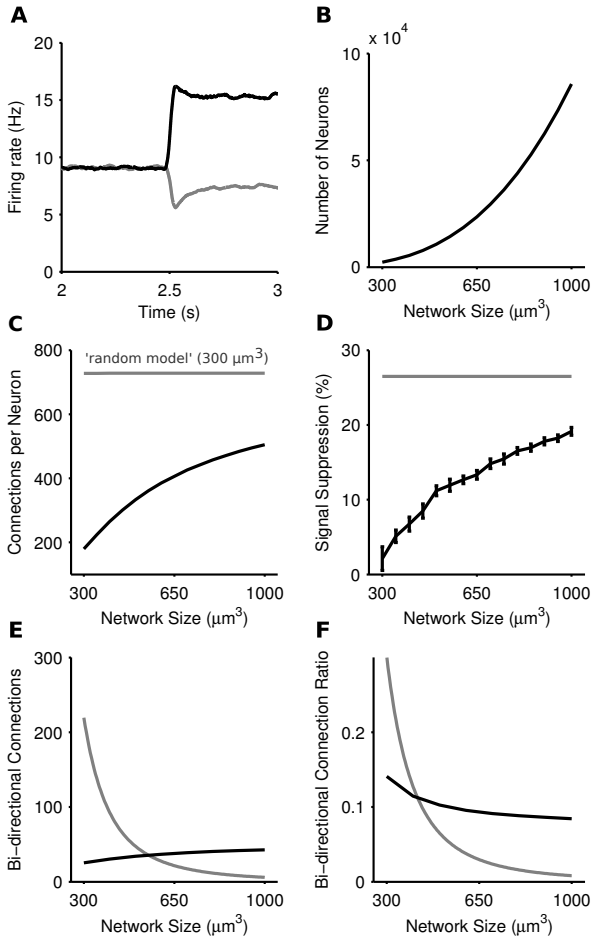

Fig. B.9 **Steady-state selection in the physical model of the striatal microcircuit.** **(A)** Mean firing rate of two projection neuron populations in a 1 mm^3 model, with 89,749 total simulated neurons. **(B)** Number of simulated neurons as a function of network size. **(C)** Average number of connections per neuron as a function of network size. The *physical* network (black) approaches the density of connections seen in the random network (grey) with increased network size. **(D)**Magnitude of steady-state selection as a function of network size. All simulations used the inputs [5,6] Hz. Magnitude is the percentage suppression in the average firing rate of the losing channel after the competitive signal onset ($t = 2.5$ s). Shown in grey is the steady-state selectivity seen in the random model for a network of size 300 μm^3. Bars set at $\pm$ 2 s.d, computed over 15 repeats. **(E)** Number of bi-directional connections as a function of network size. The total number of pairs of reciprocal connections in the *physical* model are shown in black, and the *random* model in grey. Bi-directional pairs decrease in the physical model with increasing network sizes, due to the fixed number of connections each neuron receives. **(F)** The ratio $R_{bi}$of bi-directional connections to the total number of connections a neuron makes for the *physical* model (black) and the *random* model (grey).

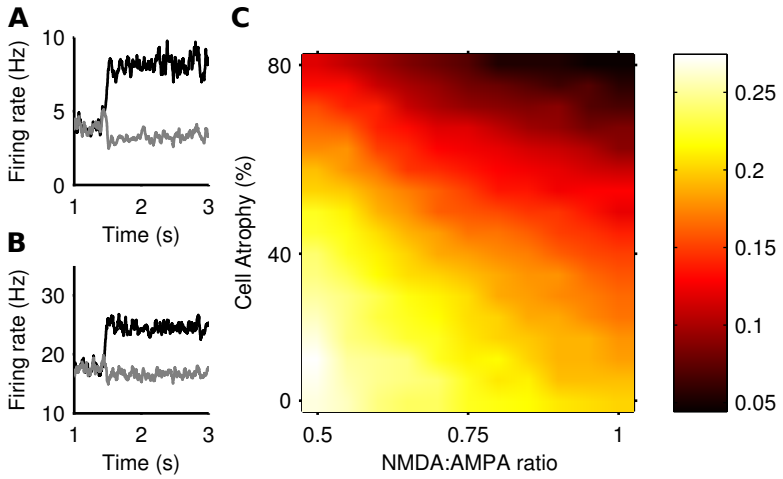

Fig. B.10 **Steady-state selection under simulated Huntington's disease. (A)** An example of reduced signal suppression in the striatum with high cell atrophy (65% cell loss, NMDA:AMPA ratio 0.5). **(B)** An example of removed signal suppression in the striatum with high degradation (75% cell loss, NMDA:AMPA ratio 1). (C) Magnitude of signal suppression over all simulated Huntington's conditions. Magnitudes are means over 15 simulations. The control, healthy-state model is in the bottom left-hand corner (NMDA:AMPA = 0.5; 0% atrophy).

cell loss by randomly removing a specified percentage of projection neurons. We did this to explore a wide range of plausible simulated Huntington's disease conditions. Across both changes, we mapped the change in transient and steady-state selection in response to the same input protocol (baseline 5 Hz, step 1 Hz).

Steady-state selection consistently degrades in simulated Huntington's disease

To assess the impact of Huntington's-like changes on steady-state selection, we used the randomly-connected model to ensure that the suppression of the losing population was sufficient to be detectably modulated by the Huntington's-like changes. Figure B.10 shows that steady-state selection was uniformly diminished by all Huntington's-like changes, whether in isolation or combination.

Transient selection enhancement in simulated Huntington's disease

We assessed the impact of Huntington's-like changes on transient selection using the same physical model network as that used for Figure B.4. Figure B.11 shows that transient selection could be diminished by the loss of projection neurons alone, yet could be enhanced by the simultaneous increase in NMDA conductance. Thus the model predicts a region

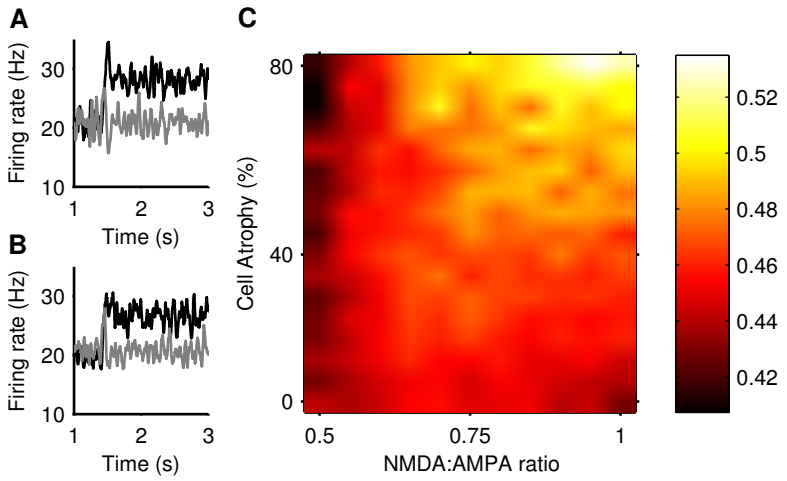

Fig. B.11 **Transient selection can be enhanced in simulated Huntington's disease (A)**. An example of enhanced transient selection in a Huntington's-like condition (81% cell atrophy, 0.95 NMDA:AMPA ratio) **(B)** An example of the loss of transient selection in a Huntington's-like condition (81% cell atrophy, 0.55 NMDA:AMPA ratio). **(C)**. Selection landscape for NMDA:AMPA conductance ratio against cell atrophy. Colour coded such that brighter colours represent better transient selectivity in the striatal model. Magnitudes are means over 30 simulations. The control, healthy-state model is in the bottom left-hand corner (NMDA:AMPA = 0.5; 0% atrophy).

of Huntington's-like conditions where the deleterious effect of cell loss can be more than compensated by the increased sensitivity of NMDA receptors.

Figure B.11A shows an example improvement in transient selectivity under high cell atrophy and a high excitability, whereas Figure B.11B shows the removal of the transient selectivity under high cell atrophy but only a small increase in excitability. These examples show that the Transient Selectivity range of $\sim$0.10 over the excitotoxicity landscape in Figure B.11C, corresponds to dramatic changes in the striatal output. Further, Figure B.6 shows that even small modifications in the transient size in the striatum will modulate the signal selection speed in the wider basal ganglia networks.

B.4 Discussion

We found a novel form of transient selection supported by the striatal network. This emerged from our three-dimensional network of sparse, weak feedback connectivity between the striatal projection neurons and dense, strong feedforward inputs from the fast-spiking interneurons. We observed that rapidly increasing the ongoing input to one of two competing populations of projection neurons caused a transient peak of activity in that population and a

synchronous transient dip in activity of the other. The dip lasted around 100 milliseconds before the activity returned to its pre-step level, thus showing no steady-state competitive effect between the two populations.

Using a population-level model of the complete basal ganglia-thalamo-cortical loop, we showed that transient selection in the striatum was sufficient to enhance selection by the entire circuit (as determined by suppression of SNr output). The presence of transient selection both increased the speed at which the whole circuit resolved a competition between salient inputs, and increased the circuit's ability to persist with the selected input. Both effects were observed for either perfect-integrator or amplifying feedback in the thalamo-cortical loop.

The origin of the transient selection had two components. The positive transient in the population activity was driven by single neuron adaptation. We found that a further step in input to an already depolarised projection neuron caused a spike followed by rapid decrease in spiking probability. This implies that the positive transient observed in the population activity was a statistical effect: that, across a whole population of projection neurons, a sub-set of neurons were sufficiently depolarised at the time of stepped input to show this adaptation effect in synchrony, and thus cause a transient peak in population activity.

The negative transient in the population activity was a subsequent network effect of the positive transient: the synchronised spiking of the neurons participating in the positive transient was sufficient to drive a dip in activity in their target neurons in the other population.

B.4.1 Two forms of selection competition

Having established the existence and mechanics of the transient selection phenomenon, we sought to understand the conditions under which our striatal model could also support a steady-state competition effect, akin to classical winner-takes-all (Fukai and Tanaka, 1997; Hartline and Ratliff, 1958; Mao and Massaquoi, 2007). Such steady-state competition could plausibly arise in striatum as each projection neuron receives sufficient weak synapses from other projection neurons to continuously modulate its ongoing activity (Chuhma et al., 2011; Guzman et al., 2003; Humphries et al., 2010).

We found that increasing the number of projection neuron synapses gave rise to steady-state competition where the stable increase in activity in one population caused a stable decrease in activity of the other population. These results are consistent with Yim et al. (2011) who reported a weakly-competitive effect between two populations of neurons in a randomly-connected inhibitory network of spiking neurons, and showed that weak correlation between inputs to the network could enhance this effect. We advanced this result by showing that such steady-state competition could arise in both distance-dependent and randomly-connected

networks, given either that we increased the physical size of our three-dimensional striatal network, and thus increase the density of connections, or randomly-connected the network based on the average connections of the most densely connected projection neuron.

Our models thus predict that the form of selection competition is dependent on the density of connections between projection neurons. Whether the striatum is ever as sparsely connected as in our distance-dependent model, or ever as densely connected as in the homogenous random model is an open question. It is possible that both forms of selection exist depending on local inhomogeneities in striatal tissue. We know that many aspects of the striatum shows gradients of density across the network, including the dorsal-ventral gradient of interneuron populations (Kubota and Kawaguchi, 1993) and the rostro-caudal gradient of FSI gap junctions (Fukuda, 2009). Correspondingly, it is plausible that there exists a gradient of projection neuron connection density.

We also note that the recent report by Oorschot et al. (2013) of projection neuron collaterals making synapses on to the somas of other projection neurons can only enhance both forms of competition. Such GABAergic somatic synapses are likely to shunt all dendritic input to the soma, thus providing powerful feedback inhibition. For transient selection, this could result in a larger negative transient; for steady-state selection, this could result in more depressed activity in the losing population. Open questions here include the relative density of such somatic synapses originating from projection neurons, and whether they have specific functional targets such as specifically occurring between projection neurons in competing populations.

Both forms of striatal selection mechanisms ultimately influence selection mediated by the whole basal ganglia network and expressed via their output nuclei (including SNr). As discussed in the Methods, this expression is via disinhibition (Berns and Sejnowski, 1998; Chevalier and Deniau, 1990; Gurney et al., 2001; Humphries et al., 2006; Redgrave et al., 1999): increased activity of a striatal population inhibits the tonic inhibitory output of a SNr population, thus representing the selection of their represented signal (Figure B.3). We showed that transient selection in the striatal populations is sufficient to enhance selection by disinhibition from SNr (Figure B.6). This occurs because the most salient input causes a transient increase of activity in the corresponding striatal population and consequently transiently decreases the output of the corresponding SNr population. This fall is sufficient to allow activity to grow in the target thalamo-cortical loop, which in turn projects to the original striatal population, further increasing its activity – thus the positive feedback loop amplifies the transient changes in striatum. The effect of steady-state selection in the striatum on the whole basal ganglia is more straightforward. The long-lasting drop in output of all

losing striatal populations comparatively reduces their inhibition of the corresponding SNr populations. Consequently, the fall in output of the SNr population representing the winning signal is enhanced compared to its competitors.

B.4.2 Experimental predictions of transient selection

Direct experimental observation of transient selection is challenging. The positive transient in population activity could only be observed on a single trial given sufficient simultaneous sampling of neurons within that population, a situation unlikely to occur with current recording technology. However, we showed that the basic mechanism underlying the positive transient in the population activity could be observed through sequential steps of current injection into a single neuron model. Thus our model makes a tractable experimental prediction: that there exists a regime of long, sequential steps of current into the projection neuron soma that will elicit a rapid burst of two or more spikes followed by slower regular firing. If such a regime exists, it would provide evidence in favour of the existence of transient selection mechanisms in the striatal network.

B.4.3 Transient selection alone could explain enhanced selection in Huntington's disease

We sought to determine whether transient and steady-state selection could be differentiated by their predictions for how changes to the striatal circuit would affect selection. To this end, we asked if Huntington's-like changes of increased NMDA receptor sensitivity and loss of projection neurons could account for Beste et al. (2008)'s report of enhanced selection by Huntington's disease patients. In terms of our models, we asked if either transient or steady-state selection would improve due to these Huntington's-like changes to the striatum.

As one might expect a priori, simply removing projection neurons and thus reducing connectivity between them impaired both types of selection. Increasing NMDA receptor sensitivity also impaired steady-state selection, and thus this form of selection predicted that all Huntington's-like changes impair selection, a result which is inconsistent with the report by Beste et al. (2008). Surprisingly, however, we found that for transient selection, increased NMDA receptor sensitivity could more than compensate for cell loss and actually enhance selection. We also found that transient selectivity was only clearly improved with both high cell degradation and increased excitability, and thus not in pre-symptomatic-like conditions. Thus, alteration of transient selection and not steady-state selection in striatum is consistent

with enhanced performance of symptomatic Huntington's disease patients compared to controls and pre-symptomatic patients.

Beste et al. (2008) noted that this enhanced response selection was paradoxical, as Huntington's disease patients are consistently worse than age-matched controls across a range of cognitive decision-making tasks (Bamford et al., 1995; Ho et al., 2003; Knopman and Nissen, 1991; Lawrence et al., 1998). Our models offer two potential explanations for why Huntington's disease related changes in striatum are usually associated with cognitive impairment but could also lead to paradoxical cognitive enhancement. First, suppose that all regions of striatum engaged by cognitive tasks implement transient selection. Our model shows that there are limited combinations of NMDA receptor sensitivity increase and cell atrophy where transient selection is enhanced compared to the healthy case; for most combinations transient selection is deteriorated compared to the healthy-state. Thus, one hypothesis is that there is a continuum of NMDA receptor sensitivity increase and cell atrophy across the striatum, and the Beste et al. (2008) task engaged a region of striatum with enhanced transient selection, whereas most tasks engage regions of the striatum with deteriorated transient selection. Second, suppose instead that different regions of striatum use transient or steady-state selection dependent on the local density of projection neuron connections. Our models shows that steady-state selection is always deteriorated by any Huntington's-like change to the striatum. Consequently, this suggests the hypothesis that the Beste et al. (2008) task engaged a region of the striatum using (enhanced) transient selection, whereas most cognitive tasks engage a region of striatum using steady-state selection, and thus are always deteriorated in Huntington's disease patients compared to the healthy-state.

Abbreviation	Definition	Page
AMPA	α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid	30
BG	Basal ganglia	1
CAG	Codon encoding for Glutamine	50
CNS	Central Nervous System	8
CTX	Cortex	57
*D*1	Dopamine receptor D1	9
*D*2	Dopamine receptor D2	9
DA	Dopamine	51
EPSP	Excitatory postsynaptic potential	30
FSI	Fast Spiking Inter-neuron	11
GABA	γ-Aminobutyric acid	13
GPe	globus pallidus external	9
GPi	globus pallidus internal	9
HD	Huntington's disease	1
HH	Hodgkin and Huxley neuron model	8
LIF	Leaky integrate and fire	22
MSPRT	Multi-Hypothesis Sequential Probability Ratio Test	46
NMDA	N-methyl-D-aspartate	30
pHD	pre-symptomatic Huntington's disease	48
RRI	Recurrent reciprocal inhibition	40
SEM	Standard error of the mean	52
SNr	substantia nigra reticular	9
SNc	substantia nigra compacta	9
SPN	Spiny projection neuron	11
SS	Steady-state selectivity	3
STN	Subthalamic nucleus	9
THAL	Thalamus	57
TRN	Thalamic reticular nucleus	57
TS	Transient Selectivity	2
WTA	Winner Take All	40

References

Adrian, E. D. (1926). The impulses produced by sensory nerve-endings. *Journal of physiology*, 62(1):33–51.

Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in neurosciences*, 12(10):366–375.

Alexander, M. E. and Wickens, J. R. (1993). Analysis of striatal dynamics: the existence of two modes of behaviour. *Journal of Theoretical Biology*, 163:413–438.

Andre, V. M., Cepeda, C., Fisher, Y. E., Huynh, M., Bardakjian, N., Singh, S., Yang, X. W., and Levine, M. S. (2011). Differential Electrophysiological Changes in Striatal Output Neurons in Huntington's Disease. *Journal of Neuroscience*, 31(4):1170–1182.

Asaad, W. F., Rainer, G., and Miller, E. K. (2000). Task-Specific Neural Activity in the Primate Prefrontal Cortex. *Journal of Neurophysiology*, 84(1):451–459.

Bamford, K. A., Caine, E. D., Kido, D. K., Cox, C., and Shoulson, I. (1995). A prospective evaluation of cognitive decline in early Huntington's disease: Functional and radiographic correlates. *Neurology*, 45(10):1867–1873.

Barto, A. G. (1994). Adaptive Critics and the Basal Ganglia. In Houk, James C., Davis, Joel L., Beise, D. G., editor, *Models of Information Processing in the Basal Ganglia*, volume 1994, chapter 11, page 394. MIT Press.

Berns, G. S. and Sejnowski, T. J. (1998). A computational model of how the basal ganglia produce sequences. *Journal of Cognitive Neuroscience*, 10:108–121.

Beste, C., Saft, C., Güntürkün, O., and Falkenstein, M. (2008). Increased cognitive functioning in symptomatic Huntington's disease as revealed by behavioral and event-related potential indices of auditory sensory memory and attention. *The Journal of Neuroscience*, 28(45):11695–702.

Beste, C., Saft, C., Yordanova, J., Andrich, J., Gold, R., Falkenstein, M., and Kolev, V. (2007). Functional compensation or pathology in cortico-subcortical interactions in preclinical Huntington's disease? *Neuropsychologia*, 45(13):2922–2930.

Beste, C., Tomkins, A., Vasilaki, E., Saft, C., Gurney, K., and Humphries, M. (2012). Increased cognitive functioning in manifest HD—empirical evidence and computational modelling. *Journal of Neurology, Neurosurgery & Psychiatry*.

Bi, G. Q. and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, 18(24):10464–72.

Bi, G. Q. and Wang, H. X. (2002). Temporal asymmetry in spike timing-dependent synaptic plasticity. *Physiology & behavior*, 77(4-5):551–5.

Blomeley, C. P., Kehoe, L., and Bracci, E. (2009). Substance P mediates excitatory interactions between striatal projection neurons. *The Journal of Neuroscience*, 29(15):4953–63.

Bogacz, R. and Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural computation*, 19(2):442–477.

Bogacz, R. and Larsen, T. (2011). Integration of reinforcement learning and optimal decision-making theories of the Basal Ganglia. *Neural computation*, 23(4):817–51.

Bracci, E., Centonze, D., Bernardi, G., and Calabresi, P. (2002). Dopamine excites fast-spiking interneurons in the striatum. *Journal of Neurophysiology*, 87:2190–2194.

Brette, R. and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of neurophysiology*, 94(5):3637–42.

Butler, A. B. and Hodos, W. (2005). *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation: Second Edition*. John Wiley and Sons.

Calabresi, P., Centonze, D., Gubellini, P., Marfia, G. a., Pisani, a., Sancesario, G., and Bernardi, G. (2000). Synaptic transmission in the striatum: From plasticity to neurodegeneration. *Progress in Neurobiology*, 61(3):231–265.

Cassey, T. C., Evens, D. R., Bogacz, R., Marshall, J. A. R., and Ludwig, C. J. H. (2013). Adaptive sampling of information in perceptual decision-making. *PLoS ONE*, 8(11).

Centonze, D., Grande, C., Usiello, A., Gubellini, P., Erbs, E., Martin, A. B., Pisani, A., Tognazzi, N., Bernardi, G., Moratalla, R., Borrelli, E., and Calabresi, P. (2003). Receptor subtypes involved in the presynaptic and postsynaptic actions of dopamine on striatal interneurons. *The Journal of Neuroscience*, 23(15):6245–54.

Cepeda, C., Cummings, D. M., André, V. M., Holley, S. M., and Levine, M. S. (2010). Genetic mouse models of Huntington's disease: focus on electrophysiological mechanisms. *ASN neuro*, 2(2):e00033.

Cepeda, C., Wu, N., André, V. M., Cummings, D. M., and Levine, M. S. (2007). The corticostriatal pathway in Huntington's disease. *Progress in Neurobiology*, 81(5-6):253–271.

Chambers, J. M., Gurney, K., Humphries, M., and Prescott, A. (2011). Mechanisms of choice in the primate brain: a quick look at positive feedback. In Bryson, J. J., Prescott, T. J., and Seth, A. K., editors, *Modelling Natural Action Selection*, pages 390–420. CUP, Cambridge, UK.

Charpier, S. and Deniau, J. M. (1997). In vivo activity-dependent plasticity at cortico-striatal connections: evidence for physiological long-term potentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 94(13):7036–7040.

Chersi, F., Mirolli, M., Pezzulo, G., and Baldassarre, G. (2013). A spiking neuron model of the cortico-basal ganglia circuits for goal-directed and habitual action learning. *Neural Networks*, 41:212–224.

Chevalier, G. and Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal function. *Trends in Neurosciences*, 13:277–280.

Choi, W. Y., Balsam, P. D., and Horvitz, J. C. (2005). Extended habit training reduces dopamine mediation of appetitive response expression. *The Journal of Neuroscience*, 25(29):6729–6733.

Chuhma, N., Tanaka, K. F., Hen, R., and Rayport, S. (2011). Functional Connectome of the Striatal Medium Spiny Neuron. *The Journal of Neuroscience*, 31(4):1183–1192.

Clopath, C. and Büsing, L. (2009). Connectivity reflects Coding : A Model of Voltage-based Spike-Timing-Dependent-Plasticity with Homeostasis. *Nature Neuroscience*.

Costa, R. M. (2007). Plastic corticostriatal circuits for action learning: What's dopamine got to do with it? In *Annals of the New York Academy of Sciences*, volume 1104, pages 172–191.

Coyle, J. T. and Schwarcz, R. (1976). Lesion of striatal neurones with kainic acid provides a model for Huntington's chorea. *Nature*, 263(5574):244–246.

Craufurd, D. and Snowden, J. (2002). Neuropsychological and neuropsychiatric aspects of Huntington's Disease. *Huntington's disease*.

Czubayko, U. and Plenz, D. (2002). Fast synaptic transmission between striatal spiny projection neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 99:15764–15769.

Dayan, P. and Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, 36(2):285–298.

Deco, G. and Rolls, E. T. (2005). Attention, short-term memory, and action selection: a unifying theory. *Progress in Neurobiology*, 76(4):236–56.

Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of Neuroscience*, 18:193–222.

Destexhe, A., Mainen, Z. F., and Sejnowski, T. J. (1994). An Efficient Method for Computing Synaptic Conductances Based on a Kinetic Model of Receptor Binding. *Neural Computation*, 6(1):14–18.

DiFiglia, M. (1990). Excitotoxic injury of the neostriatum: a model for Huntington's disease. *Trends in neurosciences*, 13(7):286–289.

Ding, L. and Gold, J. I. (2013). The basal ganglia's contributions to perceptual decision making. *Neuron*, 79:640–649.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, 12(7-8):961–974.

Eaton, R. C., Hofve, J. C., and Fetcho, J. R. (1995). Beating the competition: the reliability hypothesis for Mauthner axon size. *Brain, behavior and evolution*, 45(4):183–194.

Eidelberg, D. and Surmeier, D. J. (2011). Brain networks in Huntington disease. *The Journal of clinical investigation*, 121(2):484–492.

Eliasmith, C. (2005). A unified approach to building and controlling spiking attractor networks. *Neural computation*, 17(6):1276–1314.

Fan, M. M. Y. and Raymond, L. A. (2007). N-methyl-D-aspartate (NMDA) receptor function and excitotoxicity in Huntington's disease. *Progress in Neurobiology*, 81(5-6):272–93.

Feigin, A., Ghilardi, M. F., Huang, C., Ma, Y., Carbon, M., Guttman, M., Paulsen, J. S., Ghez, C. P., and Eidelberg, D. (2006). Preclinical Huntington's disease: Compensatory brain responses during learning. *Annals of Neurology*, 59(1):53–59.

Feigin, A., Kieburtz, K., Bordwell, K., Como, P., Steinberg, K., Sotack, J., Zimmerman, C., Hickey, C., Orme, C., and Shoulson, I. (1995). Functional Decline in Huntingtons-Disease. *Movement Disorders*, 10(2):211–214.

Fiete, I. R., Senn, W., Wang, C. Z. H., and Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*, 65(4):563–76.

Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current opinion in neurobiology*, 21(3):381–386.

Freeman, J. S., Cody, F. W. J., O'Boyle, D. J., Craufurd, D., Neary, D., and Snowden, J. S. (1996). Abnormalities of Motor Timing in Huntington's Disease. *Parkinsonism and Related Disorders*, 2(2):81–93.

Froemke, R. C., Tsay, I. A., Raad, M., Long, J. D., and Dan, Y. (2006). Contribution of individual spikes in burst-induced long-term synaptic modification. *Journal of neurophysiology*, 95(3):1620–1629.

Fukai, T. and Tanaka, S. (1997). A simple neural network exhibiting selective activation of neuronal ensembles: from winner-take-all to winners-share-all. *Neural computation*, 9:77–97.

Fukuda, T. (2009). Network architecture of gap junction-coupled neuronal linkage in the striatum. *Journal of Neurophysiology*, 29:1235–1243.

Fusco, F. R., Chen, Q., Lamoreaux, W. J., Figueredo-Cardenas, G., Jiao, Y., Coffman, J. A., Surmeier, D. J., Honig, M. G., Carlock, L. R., and Reiner, A. (1999). Cellular localization of huntingtin in striatal and cortical neurons in rats: lack of correlation with neuronal vulnerability in Huntington's disease. *The Journal of Neuroscience*, 19(4):1189–1202.

Galarreta, M. and Hestrin, S. (1999). A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature*, 402:72–75.

Gallistel, C. R. (1981). Précis of Gallistel's The organization of action: A new synthesis. *Behavioral and Brain Sciences*, 4(04):609.

Garside, S., Furtado, J. C., and Mazurek, M. F. (1996). Dopamine-glutamate interactions in the striatum: behaviourally relevant modification of excitotoxicity by dopamine receptor-mediated mechanisms. *Neuroscience*, 75(4):1065–74.

Gautrais, J. and Thorpe, S. (1998). Rate coding versus temporal order coding: A theoretical approach. *Biosystems*.

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., and Sibley, D. R. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science*, 250(4986):1429–1432.

Gerstner, W. and Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press.

Ghiglieri, V., Bagetta, V., Calabresi, P., and Picconi, B. (2012). Functional interactions within striatal microcircuit in animal models of Huntington's disease. *Neuroscience*, 211(0):165–184.

Gittis, A. H., Nelson, A. B., Thwin, M. T., Palop, J. J., and Kreitzer, A. C. (2010). Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. *Journal of Neurophysiology*, 30:2223–2234.

Gorelova, N., Seamans, J. K., and Yang, C. R. (2002). Mechanisms of dopamine activation of fast-spiking interneurons that exert inhibition in rat prefrontal cortex. *J Neurophysiol*, 88:3150–3166.

Grafton, S. T. and Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590–616.

Graybiel, A. M. (1995). Building action repertoires: Memory and learning functions of the basal ganglia. *Current opinion in neurobiology*, 5(6):733–741.

Grillner, S., Hellgren, J., Menard, A., Saitoh, K., and Wikstrom, M. A. (2005). Mechanisms for selection of basic motor programs – roles for the striatum and pallidum. *Trends in Neurosciences*, 28:364–370.

Grosse-Wentrup, M. and Contreras-Vidal, J. L. (2007). The role of the striatum in adaptation learning: A computational model. *Biological Cybernetics*, 96(4):377–388.

Gruber, A. J., Dayan, P., Gutkin, B. S., and Solla, S. A. (2006). Dopamine modulation in the basal ganglia locks the gate to working memory. *Journal of Computational Neuroscience*, 20:153–166.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001). A computational model of action selection in the basal ganglia {I}: {A} new functional anatomy. *Biological Cybernetics*, 85:401–410.

Gurney, K., Prescott, T. J., Wickens, J. R., and Redgrave, P. (2004). Computational models of the basal ganglia: from robots to membranes. *Trends in Neurosciences*, 27(8):453–9.

Gurney, K. N., Humphries, M. D., and Redgrave, P. (2015). A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. *PLoS Biology*, 13(1):e1002034.

Guzman, J. N., Hernandez, A., Galarraga, E., Tapia, D., Laville, A., Vergara, R., Aceves, J., and Bargas, J. (2003). Dopaminergic modulation of axon collaterals interconnecting spiny neurons of the rat striatum. *Journal of Neurophysiology*, 23:8931–8940.

Hartline, H. K. and Ratliff, F. (1958). Spatial summation of inhibitory influences in the eye of Limulus, and the mutual interaction of receptor units. *J Gen Physiol*, 41:1049–1066.

Hemmen, J. L. and Schwartz, A. B. (2008). Population vector code: A geometric universal as actuator. *Biological Cybernetics*, 98(6):509–518.

Heng, M. Y., Detloff, P. J., and Albin, R. L. (2008). Rodent genetic models of Huntington disease. *Neurobiology of disease*, 32(1):1–9.

Ho, A. K., Sahakian, B. J., Brown, R. G., Barker, R. A., Hodges, J. R., Ané, M. N., Snowden, J., Thompson, J., Esmonde, T., Gentry, R., Moore, J. W., and Bodner, T. (2003). Profile of cognitive progression in early Huntington's disease. *Neurology*, 61(12):1702–1706.

Hodgkin, A. L. and 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.

Hollerman, J. R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature neuroscience*, 1(4):304–9.

Houk, J. C. and Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cerebral Cortex*, 5:95–110.

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., and Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, 500(7464):575–9.

Humphries, M. D. and Gurney, K. (2002). The role of intra-thalamic and thalamocortical circuits in action selection. *Network*, 13:131–156.

Humphries, M. D., Gurney, K., and Redgrave, P. (2008). Cortico-striatal plasticity for action-outcome learning using spike timing dependent eligibility. *Nature Neuroscience*, 120(2002):2008.

Humphries, M. D., Lepora, N., Wood, R., and Gurney, K. (2009a). Capturing dopaminergic modulation and bimodal membrane behaviour of striatal medium spiny neurons in accurate, reduced models. *Frontiers in computational neuroscience*, 3(November):26.

Humphries, M. D., Stewart, R. D., and Gurney, K. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *The Journal of Neuroscience*, 26(50):12921–42.

Humphries, M. D., Wood, R., and Gurney, K. (2009b). Dopamine-modulated dynamic cell assemblies generated by the GABAergic striatal microcircuit. *Neural Networks*, 22:1174–1188.

Humphries, M. D., Wood, R., and Gurney, K. (2010). Reconstructing the Three-Dimensional GABAergic Microcircuit of the Striatum. *PLoS Computational Biology*, 6(11):e1001011.

Huxter, J., Burgess, N., and O'Keefe, J. (2003). Independent rate and temporal coding in hippocampal pyramidal cells. *Nature*, 425(6960):828–832.

Ijspeert, A. J., Nakanishi, J., and Schaal, S. (2002). Learning Attractor Landscapes for Learning Motor Primitives. *Advances in Neural Information Processing Systems*, pages 1547–1554.

Ito, M. and Doya, K. (2011). Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. *Current opinion in neurobiology*, 21(3):368–373.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 14(6):1569–72.

Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15(5):1063–1070.

Izhikevich, E. M. (2007a). *Dynamical Systems in Neuroscience*. MIT Press, Cambridge, MA.

Izhikevich, E. M. (2007b). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17(10):2443–52.

Izhikevich, E. M. and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9):3593–8.

Jaeger, D., Hitoshi, K., and Wilson, C. J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *Journal of Neurophysiology*, 72:2555–2558.

Jahr, E., Health, O., and Diego, S. (1990). Voltage Dependence of NMDA-Activated Predicted by Single-Channel Kinetics Macroscopic Conductances. *Nature*, (September).

Jakel, R. J. and Maragos, W. F. (2000). Neuronal cell death in Huntington's disease: a potential role for dopamine. *Trends in Neurosciences*, 23(6):239–245.

Jin, X. and Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, 466:457–462.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks*, 15(4-6):535–547.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science*, volume 4. McGraw-Hill Medical.

Karlsen, A. S. and Pakkenberg, B. (2011). Total numbers of neurons and glial cells in cortex and basal ganglia of aged brains with down syndrome-a stereological study. *Cerebral Cortex*, 21(11):2519–2524.

Kirkwood, S. C., Su, J. L., Conneally, P., and Foroud, T. (2001). Progression of symptoms in the early and middle stages of Huntington disease. *Archives of neurology*, 58(2):273–278.

Kloppel, S., Draganski, B., Siebner, H. R., Tabrizi, S. J., Weiller, C., and Frackowiak, R. S. J. (2009). Functional compensation of motor function in pre-symptomatic Huntingtons disease. *Brain*, 132(6):1624–1632.

Knopman, D. and Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia*, 29(3):245–254.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, 273(5280):1399–1402.

Koos, T., Tepper, J. M., and Wilson, C. J. (2004). Comparison of IPSCs evoked by spiny and fast-spiking neurons in the neostriatum. *Journal of Neurophysiology*, 24:7916–7922.

Kowall, N. W., Ferrante, R. J., and Martin, J. B. (1987). Patterns of cell loss in huntington's disease. *Trends in Neurosciences*, 10:24–29.

Kravitz, A. V., Tye, L. D., and Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature neuroscience*, 15(6):816–818.

Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience & Biobehavioral Reviews*, 26(6):631–664.

Kubota, Y. and Kawaguchi, Y. (1993). Spatial distributions of chemically identified intrinsic neurons in relation to patch and matrix compartments of rat neostriatum. *Journal of Computational Neuroscience*, 332:499–513.

Laughlin, S. B., Steveninck, R. R., and Anderson, J. C. (1998). The metabolic cost of neural information. *Nature neuroscience*, 1(1):36–41.

Lawrence, A. D., Sahakian, B. J., and Robbins, T. W. (1998). Cognitive functions and corticostriatal circuits: insights from Huntington's disease. *Trends Cognitive Sciences*, 2:379–388.

Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *Journal of Neurophysiology*, 26:3567–3583.

Lee, D. K., Itti, L., Koch, C., and Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature neuroscience*, 2(4):375–381.

Leise, E. M. (1990). Modular construction of nervous systems: a basic principle of design for invertebrates and vertebrates. *Brain Research Reviews*, pages 1–23.

Lemiere, J., Decruyenaere, M., Evers-Kiebooms, G., Vandenbussche, E., and Dom, R. (2004). Cognitive changes in patients with Huntington's disease (HD) and asymptomatic carriers of the HD mutation–a longitudinal follow-up study. *Journal of neurology*, 251(8):935–42.

Lester, R. A., Clements, J. D., Westbrook, G. L., and Jahr, C. E. (1990). Channel kinetics determine the time course of {NMDA} receptor-mediated synaptic currents. *Nature*, 346:565–567.

Linaro, D., Storace, M., and Giugliano, M. (2011). Accurate and fast simulation of channel noise in conductance-based model neurons by diffusion approximation. *PLoS Computational Biology*, 7(3).

Linden, D. J. (2001). The expression of cerebellar LTD in culture is not associated with changes in AMPA-receptor kinetics, agonist affinity, or unitary conductance. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24):14066–14071.

Lingenhöhl, K. and Friauf, E. (1994). Giant neurons in the rat reticular formation: a sensorimotor interface in the elementary acoustic startle circuit? *The Journal of Neuroscience*, 14(3 Pt 1):1176–1194.

Liu, T. and Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of neurophysiology*, 106(5):2383–2398.

London, M. and Häusser, M. (2005). Dendritic computation. *Annual review of Neuroscience*, 28:503–532.

Maass, W. (2000). On the computational power of winner-take-all. *Neural computation*, 12(11):2519–35.

Maass, W. and Markram, H. (2004). On the computational power of circuits of spiking neurons. *Journal of Computer and System Sciences*, 69(4):593–616.

Mahon, S. (2000). Role of a Striatal Slowly Inactivating Potassium Current in Short-Term Facilitation of Corticostriatal Inputs: A Computer Simulation Study. *Learning & Memory*, 7(5):357–362.

Mallet, N., Ballion, B., Le Moine, C., and Gonon, F. (2006). Cortical inputs and GABA interneurons imbalance projection neurons in the striatum of parkinsonian rats. *Journal of Neurophysiology*, 26:3875–3884.

Mao, Z. and Massaquoi, S. G. (2007). Dynamics of winner-take-all competition in recurrent neural networks with lateral inhibition. *IEEE Trans Neural Netw*, 18:55–69.

Marr, D. and Poggio, T. (1976). From understanding computation to understanding neural circuitry. *AI Memo*, 357:1–22.

Masuda, N. and Aihara, K. (2003). Duality of rate coding and temporal coding in multilayered feedforward networks. *Neural computation*, 15(1):103–125.

Matamales, M., Bertran-Gonzalez, J., Salomon, L., Degos, B., Deniau, J., Valjent, E., Hervé, D., and Girault, J. (2009). Striatal medium-sized spiny neurons: identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. *PloS one*, 4(3):e4770.

Miller, B. R., Walker, A. G., Fowler, S. C., Hörsten, S. V., Riess, O., Johnson, M. A., and Rebec, G. V. (2010). Neurobiology of Disease Dysregulation of coordinated neuronal firing patterns in striatum of freely behaving transgenic rats that model Huntington ' s disease. *Neurobiology of Disease*, 37(1):106–113.

Miller, B. R., Walker, A. G., Shah, A. S., Barton, S. J., Rebec, G. V., Cummings, D. M., Alaghband, Y., Hickey, M. A., Joshi, P. R., Candice, S., Zhu, C., Ando, T. K., André, V. M., Cepeda, C., B, J., Levine, M. S., Höhn, S., Dallérac, G., Faure, A., Urbach, Y. K., Nguyen, H. P., Hörsten, S. V., Blanc, P. L., Desvignes, N., Massioui, N. E., Brown, L., and Doyère, V. (2012). Dysregulated Information Processing by Medium Spiny Neurons in Striatum of Freely Behaving Mouse Models of Huntington ' s Disease Dysregulated Information Processing by Medium Spiny Neurons in Striatum of Freely Behaving Mouse Models of Huntington ' s Dis. *Journal of Neurophysiology*, (July 2008):2205–2216.

Milnerwood, A. J., Gladding, C. M., Pouladi, M. A., Kaufman, A. M., Hines, R. M., Boyd, J. D., Ko, R. W. Y., Vasuta, O. C., Graham, R. K., Hayden, M. R., Murphy, T. H., and Raymond, L. A. (2010). Early Increase in Extrasynaptic NMDA Receptor Signaling and Expression Contributes to Phenotype Onset in Huntington's Disease Mice. *Neuron*, 65(2):178–190.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in neurobiology*, 50(4):381–425.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, 16(5):1936–1947.

Montebelli, A., Herrera, C., and Ziemke, T. (2008). On Cognition as Dynamical Coupling: An Analysis of Behavioral Attractor Dynamics. *Adaptive Behavior*, 16(2-3):182–195.

Moustafa, A. a. and Gluck, M. a. (2011). Computational cognitive models of prefrontal-striatal-hippocampal interactions in Parkinson's disease and schizophrenia. *Neural networks : the official journal of the International Neural Network Society*.

Moyer, J. T., Wolf, J., and Finkel, L. H. (2007). Effects of dopaminergic modulation on the integrative properties of the ventral striatal medium spiny neuron. *Journal of neurophysiology*, 98(6):3731–48.

Nicola, S. M., Surmeier, J., and Malenka, R. C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annual Review of Neuroscience*, 23:185–215.

Oorschot, D. E. (1996). Total number of neurons in the neostriatal, pallidal, subthalamic, and substantia nigral nuclei of the rat basal ganglia: a stereological study using the cavalieri and optical disector methods. *Journal of Computational Neuroscience*, 366:580–599.

Oorschot, D. E., Lin, N., Cooper, B. H., Reynolds, J. N. J., Sun, H., and Wickens, J. R. (2013). Synaptic connectivity between rat striatal spiny projection neurons in vivo: Unexpected multiple somatic innervation in the context of overall sparse proximal connectivity. *Basal Ganglia*, 3:93–108.

Oster, M., Douglas, R., and Liu, S. (2009). Computation with spikes in a winner-take-all network. *Neural computation*, 21(9):2437–65.

Oster, M. and Liu, S. C. (2006). Spiking inputs to a winner-take-all network. *Advances in Neural Information Processing Systems*.

Page, W. K. and Duffy, C. J. (2003). Heading representation in MST: sensory interactions and population encoding. *Journal of neurophysiology*, 89(4):1994–2013.

Palmiter, R. D. (2008). Dopamine signaling in the dorsal striatum is essential for motivated behaviors: Lessons from dopamine-deficient mice. *Ann N Y Acad Sc*, 1129:35–46.

Paulsen, J. S. (2011). Cognitive impairment in Huntington disease: diagnosis and treatment. *Current neurology and neuroscience reports*, 11(5):474–83.

Paulsen, J. S., Zimbelman, J. L., Hinton, S. C., Langbehn, D. R., Leveroni, C. L., Benjamin, M. L., Reynolds, N. C., and Rao, S. M. (2004). fMRI biomarker of early neuronal dysfunction in presymptomatic Huntington's disease. *American Journal of Neuroradiology*, 25(10):1715–1721.

Peltsch, A., Hoffman, A., Armstrong, I., Pari, G., and Munoz, D. P. (2008). Saccadic impairments in Huntington's disease. *Experimental Brain Research*, 186(3):457–469.

Pfister, J. and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *The Journal of Neuroscience*, 26(38):9673–82.

Planert, H., Szydlowski, S. N., Hjorth, J. J., Grillner, S., and Silberberg, G. (2010). Dynamics of synaptic transmission between fast-spiking Interneurons and striatal projection neurons of the direct and indirect pathways. *Journal of Neurophysiology*, 30:3499–3507.

Plenz, D. (2003). When inhibition goes incognito: feedback interaction between spiny projection neurons in striatal function. *Trends in Neurosciences*, 26(8):436–43.

Portera-Cailliau, C., Hedreen, J. C., Price, D. L., and Koliatsos, V. E. (1995). Evidence for apoptotic cell death in Huntington disease and excitotoxic animal models. *The Journal of Neuroscience*, 15(5):3775–3787.

Prescott, S. A. and Sejnowski, T. J. (2008). Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *The Journal of Neuroscience*, 28(50):13649–13661.

Prescott, T. J., González, F. M., Gurney, K., Humphries, M. D., and Redgrave, P. (2006). A robot model of the basal ganglia: behavior and intrinsic processing. *Neural networks*, 19(1):31–61.

Prescott, T. J., Redgrave, P., and Gurney, K. (1999). Layered Control Architectures in Robots and Vertebrates. *Adaptive Behavior*, 7(1):99–127.

Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, 89:1009–1023.

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., Lehericy, S., Bergman, H., Agid, Y., Delong, M. R., and Obeso, J. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nature reviews. Neuroscience*, 11(NOvemBer):760–772.

Reilmann, R., Bohlen, S., Kirsten, F., Ringelstein, E. B., and Lange, H. W. (2011). Assessment of involuntary choreatic movements in Huntington's disease-Toward objective and quantitative measures. *Movement Disorders*, 26(12):2267–2273.

Reiner, A., Albin, R. L., Anderson, K. D., D'Amato, C. J., Penney, J. B., and Young, A. B. (1988). Differential loss of striatal projection neurons in Huntington disease. *Proceedings of the National Academy of Sciences of the United States of America*, 85(15):5733–5737.

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and cognition*, 56(2):129–40.

Ringo, J. L. (1991). Neuronal interconnection as a function of brain size. *Brain, behavior and evolution*, 38(1):1–6.

Roberts, R. C., Ahn, A., Swartz, K. J., Beal, M. F., and DiFiglia, M. (1993). Intrastriatal injections of quinolinic acid or kainic acid: differential patterns of cell survival and the effects of data analysis on outcome. *Experimental neurology*, 124(2):274–82.

Rullen, R. and Thorpe, S. J. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural computation*, 13(6):1255–1283.

Saarinen, A., Linne, M., and Yli-Harja, O. (2006). Modeling single neuron behavior using stochastic differential equations. *Neurocomputing*, 69(10-12):1091–1096.

Scharstein, H. (1979). Input-output relationship of the Leaky-integrator neuron model. *Journal of mathematical biology*, 8(4):403–420.

Schneider, M. V. (2013). In Silico Systems Biology. *In silico systems biology, Methods in Molecular Biology*, 1021:1–11.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, 80(1):1–27.

Schulz, J. M., Redgrave, P., and Reynolds, J. N. J. (2010). Cortico-striatal spike-timing dependent plasticity after activation of subcortical pathways. *Frontiers in Synaptic Neuroscience*, (JUL).

Seo, M., Lee, E., and Averbeck, B. B. (2012). Action Selection and Action Value in Frontal-Striatal Circuits. *Neuron*, 74(5):947–960.

Shiflett, M. W. and Balleine, B. W. (2011). Molecular substrates of action control in cortico-striatal circuits. *Progress in neurobiology*, 95(1):1–13.

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–64.

Smith-Roe, S. L. and Kelley, A. E. (2000). Coincident activation of NMDA and dopamine D1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. *The Journal of Neuroscience*, 20(20):7737–7742.

Sprengelmeyer, R., Lange, H., and Homberg, V. (1995). The pattern of attentional deficits in Huntington's disease. *Brain*, 118(1):145–152.

Stephenson-jones, M., Samuelsson, E., Ericsson, J., Robertson, B., and Grillner, S. (2011). Evolutionary Conservation of the Basal Ganglia as a Common Vertebrate Mechanism for Action Selection. *Current Biology*, 21(13):1081–1091.

Stewart, T. C., Bekolay, T., and Eliasmith, C. (2012). Learning to select actions with spiking neurons in the basal ganglia. *Frontiers in Neuroscience*, (JAN).

Stout, J. C., Rodawalt, W. C., and Siemers, E. R. (2001). Risky decision making in Huntington's disease. *Journal of the International Neuropsychological Society : JINS*, 7(1):92–101.

Sun, Z., Chen, Q., and Reiner, A. (2003). Enkephalinergic striatal projection neurons become less affected by quinolinic acid than substance P-containing striatal projection neurons as rats age. *Experimental neurology*, 184(2):1034–42.

Surmeier, D. J., Ding, J., Day, M., Wang, Z., and Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in Neurosciences*, 30(5):228–35.

Swerdlow, N. R., Paulsen, J., Braff, D. L., Butters, N., Geyer, M. A., and Swenson, M. R. (1995). Impaired prepulse inhibition of acoustic and tactile startle response in patients with Huntington's disease. *Journal of Neurology*, 58(2):192–200.

Szydlowski, S. N., Pollak Dorocic, I., Planert, H., Carlén, M., Meletis, K., and Silberberg, G. (2013). Target selectivity of feedforward inhibition by striatal fast-spiking interneurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(4):1678–83.

Tateno, T., Harsch, A., and Robinson, H. P. (2004). Threshold firing frequency-current relationships of neurons in rat somatosensory cortex: type 1 and type 2 dynamics. *J Neurophysiol*, 92:2283–2294.

Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715):1642–1645.

Tomkins, A., Humphries, M., Beste, C., Vasilaki, E., and Gurney, K. (2012). How Degrading Networks Can Increase Cognitive Functions. *Artificial Neural Networks and Machine Learning – ICANN 2012*, 7552:185–192.

Tomkins, A., Humphries, M. D., Vasilaki, E., and Gurney, K. N. (2011). How Degrading Networks Can Increase Select Cognitive Functions. *Frontiers in Computational Neuroscience*.

Tomkins, A., Vasilaki, E., Beste, C., Gurney, K., and Humphries, M. D. (2013). Transient and steady-state selection in the striatal microcircuit. *Frontiers in computational Neuroscience*, 7:192.

Tozzi, A., de Iure, A., Di Filippo, M., Tantucci, M., Costa, C., Borsini, F., Ghiglieri, V., Giampa, C., Fusco, F. R., Picconi, B., and Calabresi, P. (2011). The Distinct Role of Medium Spiny Neurons and Cholinergic Interneurons in the D2/A2A Receptor Interaction in the Striatum: Implications for Parkinson's Disease. *Journal of Neuroscience*, 31(5):1850–1862.

Tunstall, M. J., Oorschot, D. E., Kean, A., and Wickens, J. R. (2002). Inhibitory interactions between spiny projection neurons in the rat striatum. *Journal of Neurophysiology*, 88:1263–1269.

Vasilaki, E., Frémaux, N., Urbanczik, R., Senn, W., and Gerstner, W. (2009a). Spike-based reinforcement learning in continuous state and action space: when policy gradient methods fail. *PLoS computational biology*, 5(12):e1000586.

Vasilaki, E., Fusi, S., Wang, X.-J., and Senn, W. (2009b). Learning flexible sensori-motor mappings in a complex network. *Biological cybernetics*, 100(2):147–58.

Vasilaki, E. and Giugliano, M. (2012). Emergence of Connectivity Patterns in Networks of Model Neurons from Long-Term and Short-Term Plasticities. *Plos One*, pages 1–8.

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

Wang, H., Gerkin, R. C., Nauen, D. W., and Bi, G. Q. (2005). Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nature neuroscience*, 8(2):187–93.

Wang, L. H. and Qin, Z. H. (2006). Animal models of Huntington's disease: Implications in uncovering pathogenic mechanisms and developing therapies. *Acta pharmacologica sinica*, 27(10):1287–1302.

Wang, L. P., Li, F., Wang, D., Xie, K., Wang, D., Shen, X., and Tsien, J. Z. (2011). NMDA receptors in dopaminergic neurons are crucial for habit learning. *Neuron*, 72(6):1055–1066.

Wang, X. J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *The Journal of Neuroscience*, 19(21):9587–603.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968.

Wickens, J. (1997). Basal ganglia: structure and computations. *Network: Computation in Neural Systems*, 8:R77—-R109.

Wilson, C. J. and Groves, P. M. (1980). Fine structure and synaptic connections of the common spiny neuron of the rat neostriatum: a study employing intracellular inject of horseradish peroxidase. *Journal of Computational Neuroscience*, 194:599–615.

Wilson, C. J. and Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *The Journal of Neuroscience*, 16(7):2397–410.

Wilson, R. I. (2010). It takes all kinds to make a brain. *Nature Neuroscience*, 13(10):1158–1160.

Windhorst, U. (1996). On the role of recurrent inhibitory feedback in motor control. *Progress in neurobiology*, 49(6):517–587.

Xia, Y., Driscoll, J. R., Wilbrecht, L., Margolis, E. B., Fields, H. L., and Hjelmstad, G. O. (2011). Nucleus Accumbens Medium Spiny Neurons Target Non-Dopaminergic Neurons in the Ventral Tegmental Area. *Journal of Neuroscience*, 31(21):7811–7816.

Yim, M. Y., Aertsen, A., and Kumar, A. (2011). Significance of input correlations in striatal function. *PLoS Computational Biology*, 7:e1002254.