Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Abstract

Due to its strikingly regular structure, the cerebellum is widely thought to implement a universal neuronal computation. The leading candidate is the ‘adaptive filter’ which is analogous to an analysis-synthesis filter whose output weights are modified by a simple synaptic learning rule. In this formulation, the cerebellar granular layer forms part of the analysis pathway, and is commonly assumed to implement a spatio-temporal recoding where inputs are recombined into an expanded set of output signals. The nature of the recoding is unknown, although its dense connectivity suggests that circuit-level mechanisms play an important role, a view supported by simulations of recurrent neural networks. By developing computational simulations of neural network models of the cerebellar granular layer, I examine how the structure of neural networks enables them to effectively generate adaptive filter basis signals, and relate this to the known granular layer microcircuit. ‘Cerebellum-like’ structures in sharks and electric fish are thought to be the evolutionary precursor to the cerebellum, and have been characterised as adaptive filters which cancel the predictable component of a sensory signal. The sophistication of recoding implemented by cerebellum-like structures appears to increase through evolutionary time in a way that parallels increasingly recurrent connectivity. Networks constructed using a neural network training algorithm demonstrate the potential versatility of the granular layer circuit, whereas a more realistic ‘winner-take-all’ network reproduces some of its experimentally known properties.
Acknowledgements

I would like to express sincere gratitude to the following people, all of whom contributed to the completion of this thesis.

My two PhD supervisors, both of whom provided invaluable guidance and support throughout my doctorate. In particular I would like to thank Professor James Sneyd for his unfailingly candid and pragmatic advice, and Professor John Montgomery for many a long, occasionally meandering but always interesting scientific conversation. To think it started with an ‘old yarn about a shark’s brain’.

Professor David Bodznick for kindly inviting me to visit his laboratory at Woods Hole and for introducing me to experimental electrophysiology.

Owen and Kathy Young, my surrogate New Zealand family.

Thanks to everyone who reviewed draft manuscripts, especially Rebecca Turner, who has the patience of a saint and eyes of a hawk.

I also gratefully acknowledge the funding received from University of Auckland Doctoral Scholarship.

Finally, my biggest thanks go to my family; my parents Phillip and Hazel and my sister Katie.
Acknowledgements
Contents

Abstract iii

Acknowledgements v

1 Introduction 1

2 The Cerebellum and Cerebellum-Like Structures 5

2.1 Structure of the Cerebellum 6

2.2 Function of the Cerebellum 8

2.3 Experimental Paradigms 9

2.3.1 Classical Eyeblink Conditioning 9

2.3.2 Vestibulo-Ocular Reflex 10

2.4 Cerebellum-Like Structures 10

2.5 Adaptive Filter Model of the Cerebellum 12

2.6 The Cerebellar Granular Layer 14

2.6.1 Structure of the Granular Layer 14

2.6.2 Function of the Granular Layer 15
2.6.3 Granule Cells in Cerebellum-Like Structures .......................... 17

3 Computational Modelling of Biological Neural Networks 19

3.1 Firing Rate Models ......................................................... 20
3.1.1 Derivation ................................................................. 20
3.2 Recurrent Networks .......................................................... 23
3.3 Training Recurrent Networks .................................................. 26
3.3.1 Echo State Networks ...................................................... 26
3.3.2 FORCE Learning .......................................................... 28
3.4 Forward Models ............................................................... 30
3.4.1 Encoding Forward Models Using Recurrent Networks ................. 32
3.4.2 Discussion ................................................................. 35

4 Computational Architecture of the Granular Layer of Cerebellum-Like Structures 37

4.1 Introduction ................................................................. 40
4.2 Modelling ................................................................. 46
4.2.1 Granular Layer Model ..................................................... 46
4.2.2 Mossy Fibre Input .......................................................... 46
4.2.3 Performance Measure ...................................................... 48
4.3 Results ................................................................. 48
4.3.1 Single Cell ................................................................. 48
4.3.2 Single Layer Feedforward .................................................. 49
4.3.3 Random Recurrent Network ........................................ 50
4.3.4 Trained Recurrent Network ...................................... 51
4.3.5 Numbers of Neurons .............................................. 52
4.4 Discussion and Conclusions ........................................ 54
4.5 Appendix ............................................................. 56
  4.5.1 Neuron Model ...................................................... 56
  4.5.2 Mossy Fibre Signals ............................................. 57
  4.5.3 Performance Measure .......................................... 57
  4.5.4 Output Signal .................................................... 57
  4.5.5 Trained Recurrent Network .................................... 58

5 Sequential Pattern Formation in the Cerebellar Granular Layer 61
  5.1 Introduction ........................................................ 64
  5.2 Sequential Winner-Take-All Mechanism .......................... 66
    5.2.1 The Granular Layer Microcircuit ............................ 67
    5.2.2 Hysteresis-Loop Bursting ................................... 68
    5.2.3 Activity-Dependent Inhibition ............................... 69
    5.2.4 Parallel Fibre Signals ....................................... 69
  5.3 Firing Rate Model .................................................. 70
    5.3.1 Mossy Fibre ................................................... 73
    5.3.2 Granule Cells .................................................. 73
    5.3.3 Golgi Cell ..................................................... 74
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.4</td>
<td>Activity-Dependent Inhibition</td>
<td>75</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Time Constants</td>
<td>75</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Synaptic Weights</td>
<td>76</td>
</tr>
<tr>
<td>5.4</td>
<td>Results</td>
<td>76</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Adaptive Filter Basis</td>
<td>79</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Spiking Model</td>
<td>81</td>
</tr>
<tr>
<td>5.5</td>
<td>Discussion and Conclusions</td>
<td>83</td>
</tr>
<tr>
<td>5.6</td>
<td>Appendix</td>
<td>85</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Spiking Model</td>
<td>85</td>
</tr>
<tr>
<td>6</td>
<td>Discussion and Conclusions</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td><strong>Bibliography</strong></td>
<td>91</td>
</tr>
</tbody>
</table>
## List of Figures

2.1 The connectivity of the cerebellar cortex ...................................... 6
2.2 Adaptive filter model of the cerebellum ....................................... 13

3.1 Firing rate neuron model ............................................................ 21
3.2 Recurrent network model .............................................................. 23
3.3 Outputs from recurrent neural networks ........................................ 25
3.4 An echo state network learns the Mackey-Glass time series .............. 28
3.5 Internal models mimic the input-output characteristics of motor apparatus 30
3.6 Cerebrocerebellar communication loop ........................................ 31
3.7 Training a recurrent neural network using the FORCE learning algorithm . 33
3.8 A recurrent network learns a fixed-delay target signal ...................... 33
3.9 Training recurrent networks to encode forward models ...................... 34

4.1 Cerebellum-like structures implement an adaptive filter in order to suppress electrosensory reafference ......................................................... 42
4.2 Four granular layer models ............................................................. 43
4.3 Granular layer model, performance measure and mossy fibre signals ........ 47
4.4 Single cell granular layer network ........................................ 49
4.5 Single layer feedforward granular layer network ......................... 49
4.6 Single layer feedforward granular layer network with long-delay neurons . 51
4.7 Random recurrent network .................................................. 51
4.8 Trained recurrent granular layer network ................................ 52
4.9 Network performance against the number of neurons .................. 53

5.1 Schematic of the sequential winner-take-all mechanism ................ 66
5.2 Firing rate neuron model ..................................................... 70
5.3 Mathematical model of the sequential winner-take-all mechanism ........ 71
5.4 Mossy fibre input signal and granule and Golgi cell activation functions . 72
5.5 Firing rate model containing four granule cells .......................... 77
5.6 Firing rate model containing 10 granule cells ........................... 78
5.7 Firing rate model containing 20 granule cells with firing thresholds sampled at random ................................................................. 78
5.8 Firing rate model containing 20 granule cells with random synaptic weights 79
5.9 Three sequential winner-take-all networks combined in parallel .......... 80
5.10 Postsynaptic current consisting of the sum of two exponentials ........ 81
5.11 Spiking model of the sequential winner-take-all network ............... 82
List of Tables

4.1 Neurons in the cerebellum and cerebellum-like structures . . . . . . . . . . 41

5.1 Parameters for the firing rate model. . . . . . . . . . . . . . . . . . . . . . 75
Understanding the central nervous system is a hugely challenging task. Not only is it extraordinarily vast, the human brain containing around $10^{11}$ tiny nerve cells [1], but complexity exists at every scale — from genes encoding the proteins which form ion channels, up to the connectivity of whole brain regions. How is it possible to gain any understanding of such a complex system of interconnected parts?

Fortunately, the brain is not a randomly connected homogeneous mass, but a hierarchical collection of components, each displaying a degree of structural regularity. At a very broad level, the mammalian brain can be subdivided into three components: hindbrain, midbrain and forebrain. The hindbrain is largely associated with vegetative functions such as breathing, the midbrain with reflexive movements and the forebrain with higher cognitive functions such as memory and decision making. The cerebellum — the subject of this thesis — is considered part of the hindbrain, and has historically been thought of as a motor structure, although more recently it has been associated with a plethora of functions including language and even cognition [2].

The cerebellum is especially interesting due to the unique manner in which it is wired into the rest of the brain. Its entire output, and most of its input, is routed via higher brain
centres, meaning that it effectively sits on a side loop. This has led to the idea that the cerebellum ‘sculpts’ or ‘refines’ signals generated by motor centres, an idea which is supported by the poor motor co-ordination displayed by patients with cerebellar dysfunction [3].

Another interesting feature of the cerebellum is the strikingly uniform architecture of the cerebellar cortex, consisting of a small number of types of neurons connected in a geometrically repeating pattern. This uniform connectivity, along with its numerous connections to higher brain centres, has led to the idea of the cerebellar ‘chip’ [4]. Just as an identical microchip might be found within your mobile phone, microwave oven or car, the cerebellar chip is ‘plugged into’ other parts of the brain. The behavioural function of the cerebellar chip depends on its external connectivity, even though its internal computation is the same throughout. In this way, the cerebellum is characterised by its computational, rather than behavioural, function.

The leading candidate for the cerebellar computation is the ‘adaptive filter’, a device that implements an error-minimising signal transformation [5]. According to the adaptive filter hypothesis, the cerebellum acts as an analysis-synthesis filter which decorrelates its inputs from a teaching signal by adjusting a set of modifiable output weights. The overall function of the cerebellum is well characterised as an adaptive filter, although the mechanistic details of its operation, and the nature of the neural coding on which it relies, remain unclear.

The subject of this thesis is the computational function of the granular layer of the cerebellum. The granular layer lies on one of the cerebellum’s two input pathways, and is usually assumed to perform a ‘recoding’ of input signals, the first step in the adaptive filter’s two-step process. Despite (or perhaps as a result of) containing over 50% of the brain’s neurons, the (presumably crucially important) mechanism implemented by the granular layer is not known. In this thesis, I use a number of theoretical and computational tools to gain insight into this enigmatic part of the central nervous system. The ultimate goal of this research is to determine the mechanism by which the granular layer circuit implements a spatial and temporal recoding of its input.

Any investigation of the cerebellar computation requires a detailed understanding of existing biological knowledge of the cerebellum, as well as relevant computational modelling techniques. In Chapter 2, I review the scientific knowledge of the cerebellum of relevance to system-level theories of its function. In Chapter 3, I introduce important computational techniques used in later chapters.
In the experimental approach to neuroscience, we typically examine small components of the nervous system — perhaps individual neurons, synapses or even ion channels — and then speculate on their contribution to the overall function of the larger system to which it belongs. A contrasting approach, made possible by recent advances in machine learning and neural networks, is to engineer a simulation (in software) which mimics behaviours of the nervous system. Examining the algorithms implemented by such simulations is a source of hypotheses about how the brain works. This idea is explored in Chapter 3 where I use neural network training algorithms to build recurrent networks which implement forward models in a manner analogous to the cerebellum.

The granular layer of the cerebellum consists of an extremely dense recurrently connected network, consisting of excitatory and inhibitory neurons connected with feedback loops. Computational studies of recurrent neural networks have shown them to be capable of sophisticated and complicated dynamical behaviour, and recent advances have developed learning mechanisms which allow such networks to implement a variety of input-output transformations. In Chapter 3, I investigate the properties of such networks in the context of the types of computations we expect the granular layer to implement.

Comparative neuroscience is of particular relevance when attempting to determine the high-level computational principles underlying nervous system structures. Such fundamental principles are, presumably, shared by species across the phylogenetic spectrum. Study of the cerebellar computation benefits from a number of ‘cerebellum-like’ structures which share much of the cerebellum’s architecture and are found in most vertebrates. In particular, the study of sensory processing nuclei in sharks and electric fish has been valuable in revealing the workings of the adaptive filter.

Similarities and differences in structure and function point to fundamental principles underlying the granular layer mechanism. In Chapter 4 I develop a computational framework, based on a recurrent neural network model, which allows the cerebellum to be compared to the cerebellum-like structures of sharks and electric fish. This chapter is a verbatim reproduction of a paper published in the the journal Cerebellum [6].

The architecture of the granular layer is certainly recurrent, but unlike the networks developed in the previous chapters, not randomly so. For example, the ratio of inhibitory to excitatory neurons and their relative connectivity is quite distinctive. In Chapter 5 I develop a detailed model of the granular layer based on a ‘winner-take-all’ network, which is not only functional as the input layer to an adaptive filter, but relies on the granular layer’s unique structure. This chapter is a copy of a journal paper that, at the time of writing, is under review.
The Cerebellum and Cerebellum-Like Structures

The cerebellum has long received considerable interest from physiologists and neuroscientists, resulting in a wealth of data regarding its histology, connectivity and electrophysiology, and a great deal of speculation about its behavioural and computational function. Any study of cerebellar function must be positioned within this vast (and occasionally conflicting) sea of information.

In this chapter I review the cerebellum’s anatomy and microcircuitry, on which most theories of cerebellar function rest, and describe the understanding that has been gained from controlled experiments, in particular studies of the eyeblink reflex. I introduce the leading system-level theory of cerebellar function, the ‘adaptive filter’, and relate this to experiments on ‘cerebellum-like’ structures in fish. Finally, I describe the granular layer of the cerebellum in detail, and review theories of its function.
2.1 Structure of the Cerebellum

Of foremost prominence is the highly regular microcircuit of the cerebellar cortex, which consists of excitatory and inhibitory neurons arranged in a geometrically uniform pattern. An acquaintance with the unique contribution of each cell type, and the significance of the connections between them, is crucial to developing an understanding of the cerebellar computation.

Much of our present understanding of the cerebellum’s unique connectivity originates from a systematic investigation in the 1960s by John Eccles, Masao Itô and János Szentágothai which they laid out in their book *The Cerebellum as a Neuronal Machine* [7]. The structure of the cerebellar microcircuit, reproduced from work by Rodolfo Llinás [8], is depicted in Figure 2.1.

The cerebellum comprises an outer layer, the cerebellar cortex, wrapped around a core comprising a small number of cerebellar deep nuclei. Outputs from the cerebellum originate from the deep nuclei, which project to the brainstem via bundles of fibres called
cerebellar peduncles. The ultimate targets of cerebellar output include the motor cortex and brainstem motor centres. Most inputs to the cerebellum are also routed via the cerebellar peduncles. The cerebellum receives a huge variety of signals, including vestibular, somatosensory, auditory and visual information. Further inputs are received via the inferior olive, a nucleus located in the medulla.

The cerebellar cortex consists of a large intricately folded sheet of tissue with a highly uniform neural connectivity. It is divided into three distinct layers: the granular layer, Purkinje cell layer and molecular layer.

The granular layer (described in more detail in Section 2.6) is the innermost layer and is notable for containing a huge number of tiny, densely packed granule cells, accounting for over half the neurons in the entire brain. Granule cells’ thin, unmyelinated axons ascend from the granular layer into the molecular layer, where they bifurcate to form long, narrow filaments termed parallel fibres. The granular layer contains smaller numbers of other cell types including inhibitory Golgi cells and excitatory unipolar brush cells.

The other major cell type in the cerebellar cortex is the Purkinje cell, whose cell bodies form the Purkinje cell layer. Purkinje cell dendrites branch profusely in the molecular layer, forming dense two dimensional arborisations, arranged in parallel planes. Parallel fibres pass through these planes at right angles, making synaptic connections at Purkinje cell dendritic spines. The molecular layer also contains small inhibitory interneurons termed basket cells and stellate cells. The only output from the cerebellar cortex is conveyed by Purkinje cell axons, which project to the deep cerebellar nuclei and vestibular nuclei, forming inhibitory connections with their targets.

External input to the cerebellar cortex is conveyed via two distinct pathways: mossy fibres and climbing fibres. Mossy fibres convey signals originating from many parts of the brain including the cerebral cortex, vestibular nuclei, reticular formation and deep cerebellar nuclei [9]. They branch extensively when they enter the granular layer, where they form synapses with several hundreds of granule cells and Golgi cells.

Climbing fibres originate in the inferior olive, which despite lying in the medulla, sends its output entirely to the cerebellum. Each climbing fibre sends collaterals to the deep cerebellar nuclei, before entering the cerebellar cortex where it splits into roughly 10 branches. Each branch makes connections with a single Purkinje cell, wrapping around it and making several hundred synapses along its length.
2.2 Function of the Cerebellum

As with other brain structures, early insight into the function of the cerebellum came largely from patients with selective damage to the brain. Cerebellar dysfunction is associated with deficient motor activity; patients typically display erratic, poorly timed and uncoordinated movements. Dysmetria, defined as a consistent over- or undershooting of limb or eye movements, is a characteristic symptom of cerebellar dysfunction [10].

Patients with cerebellar dysfunction are usually capable of making a full range of movements, albeit with reduced precision. In the extreme case of cerebellar agenesis, where a cerebellum is completely absent, patients usually display impaired motor function but may otherwise be capable of leading a relatively normal life [11]. These observations led to the idea that the cerebellum does not initiate movements, but contributes to their smooth execution, perhaps generating control signals for movements which are too quick to be influenced by sensory feedback.

There is good evidence that the cerebellum has a key role in motor learning, especially tasks related to precise timing and the refinement of accurate movements. Various experiments involving lesions of a number cerebellar pathways have implicated the cerebellar cortex, as well as the nuclei, in the storage of learned motor skills. Many researchers have promoted the idea that the cerebellum acts to learn and store ‘internal models’ of motor apparatus [12, 13]. More recently, a whole host of non-motor functions have been ascribed to the cerebellum. Functional magnetic resonance imaging has revealed cerebellar activity during behaviours related to language, attention and cognition [2]. However, the extent of the cerebellum’s contribution to such high-level behaviours remains a matter of some controversy.

Its structural uniformity has prompted the idea that the function of the cerebellum is to perform a single, characteristic computation [4]. Much like a microchip, the ‘cerebellar chip’ describes a reusable architecture which may be ‘plugged in’ to other brain centres. From this viewpoint, identifying the behavioural function of the cerebellum becomes redundant — its function could be motor, sensory or cognitive, depending on how it is connected to the rest of the brain. The leading candidate for the cerebellar computation is the adaptive filter, which is described in detail in Section 2.5.
2.3 Experimental Paradigms

Much of our understanding of the mechanisms underlying cerebellar learning has been informed by a relatively small number of controlled experimental paradigms. A definitive cerebellar contribution has been found in classical eyeblink conditioning, the vestibuloocular reflex, the opto-kinetic reflex, saccadic eye movement and voluntary limb movement.

2.3.1 Classical Eyeblink Conditioning

Classical conditioning, famously discovered by Ivan Pavlov through his experiments on the salivary reflex of dogs, is a learning process by which an initially neutral stimulus acquires the capacity to elicit a reflexive response. The cerebellar contribution to one form of classical conditioning has been definitively established through a series of experiments on the eyeblink reflex [14].

In a typical eyeblink conditioning experiment, an unconditioned stimulus (US) consisting of a puff of air to the eyeball reflexively elicits the unconditioned response (UR), an eyeblink. The conditioned stimulus (CS), which precedes the US by a fixed period of time, is usually an auditory tone. After repeatedly pairing the CS and US, a learnt association is formed such that an eyeblink occurs in anticipation of the US. A distinction is made between ‘delay’ conditioning, in which the CS and US overlap, and ‘trace’ conditioning in which the CS terminates before the onset of the US.

Our understanding of the contribution of the cerebellum to eyeblink conditioning has come from a number of electrophysiological experiments. By making careful lesions to interrupt specific pathways, experimenters have been able to develop a detailed picture of signalling pathways and sites of learning [15–17].

The conditioned eyeblink reflex is triggered by activity within one of the cerebellar nuclei. The cerebellar nucleus is stimulated by the conditioned stimulus directly, but simultaneously inhibited by input from Purkinje cells. The termination of Purkinje cell activity releases the nucleus from inhibition, resulting in a precisely timed eyeblink. Purkinje cells receive stimulation from parallel fibres, which convey a wide range of signals, and from climbing fibres, which fire strongly in time with the air puff.

Learning has been demonstrated in both the cerebellar cortex [15] and the cerebellar nuclei
In the cortex, activity-dependent long term depression at parallel fibre-Purkinje cell synapses results in Purkinje cell activity coincident with climbing fibre stimulation. Learning in the cerebellar nucleus has been demonstrated in experiments in which the cortex is removed after learning. The conditioned response remains, timed to be coincident with the conditioned stimulus rather than the unconditioned stimulus.

### 2.3.2 Vestibulo-Ocular Reflex

The cerebellum has been shown to be involved in the adaptation of various ocular reflexes, including the vestibulo-ocular reflex (VOR), opto-kinetic reflex (OKR) and saccadic eye movement. These reflexes are invoked in order to maintain stable vision in spite of movements of the head and the environment. While the cerebellum is not responsible for initiating these movements, it has an important role in adapting them to maintain their precision. The region of the cerebellum involved with these reflexes is the vestibulocerebellum, its phylogenetically oldest part.

The VOR is a reflex which maintains stable eye gaze during head movements. Movement of the head is sensed by semicircular canals in the vestibular system which trigger a reflex arc, via the vestibular nuclei, to motoneurons supplying the eye muscles. Rotation of the head in one direction elicits a rotation of the eyes in the other, maintaining direction of eye gaze [19].

The cerebellum contributes to adaptation of the VOR through inhibition of the vestibular nuclei by Purkinje cells in the flocculus, a part of the vestibulocerebellum. Mossy fibre input to the flocculus is known to include vestibular afferents; climbing fibres convey *inter alia* retinal slip signals (indicating motion of the visual field across the retina). Experiments which disrupt the vestibular-visual relationship have shown that a remarkable degree of adaptation is possible. In an extreme case, subjects wearing vision-reversing goggles experience a complete reversal of the reflex [20].

### 2.4 Cerebellum-Like Structures

The cerebellum is an evolutionary ancient structure, and its distinctive morphology is highly conserved across vertebrates. In addition to the cerebellum itself, the nervous systems of most vertebrates contain architecturally similar ‘cerebellum-like’ structures. There is good evidence that the original cerebellum-like structures were the evolutionary
precursor to the cerebellum. For this reason, the study of cerebellum-like structures has been a useful means to gain insight into the function of the cerebellum [21]. Note that, according to the established usage of the term, the cerebellum itself is not considered a cerebellum-like structure.

Known cerebellum-like structures include the dorsal cochlear nucleus (DCN) in mammals, the electrosensory lateral line lobe (ELL) in weakly electric fish and the dorsal octavolateralis nucleus (DON) in sharks. The ELL and DON will feature in Chapter 4, where we develop a computational framework allowing the structure and function of cerebellum-like structures to be compared.

Most cerebellum-like structures, including the ELL and DON, are sensory nuclei, lying directly on the afferent sensory pathway. This simple connectivity has made them highly conducive to experimental study, and unlike the cerebellum, allowed their physiological function to be characterised. Experiments have demonstrated that the function of these structures is to act as noise filters, removing the predictable component of the incoming sensory signal [22].

For example, sharks possess a highly sensitive electrosense, which they use primarily for locating prey. All creatures emit an electric field, including the shark itself, and so the ascending electrosensory signal consists of behaviourally relevant (‘exafferent’) signal superimposed on a much stronger self-generated (‘reafferent’) signal. The reafferent signal is primarily generated by the animal’s own breathing motion, which results in a slowly varying predictable waveform. Experiments have demonstrated the function of a cerebellum-like sensory processing nucleus, the DON, to remove this reafferent component of the signal [23].

The cerebellum-like sensory nuclei in fish have been shown to act as adaptive filters, which assemble a prediction of the reafferent stimulus from an array of signals conveyed by parallel fibres. Section 2.5 describes the adaptive filter mechanism in more detail.

The distinguishing feature of the cerebellum and cerebellum-like structures is a distinctive molecular layer consisting of closely packed parallel fibres and the highly branching dendrites of ‘principal’ neurons. Being the output neurons of cerebellum-like structures, principal neurons are equivalent to cerebellar Purkinje cells, although unlike Purkinje cells they are not inhibitory. Like the cerebellum, principal neurons receive two distinct sets of inputs, one of which is transmitted by parallel fibre signals in the molecular layer. The other input consists of sensory stimuli received from the periphery and transmitted via deep layers of the nucleus. A major difference between cerebellum-like structures and the
cerebellum is a lack of climbing fibres.

As in the cerebellum, parallel fibres are the axons of numerous tiny granule cells, although the granule cells are generally located externally rather than in a layer beneath the molecular layer. Like the cerebellum, many cerebellum-like structures contain inhibitory Golgi cells as well other cells such as unipolar brush cells [24].

2.5 Adaptive Filter Model of the Cerebellum

The distinctive structure of the cerebellar cortex has long fascinated neuroscientists, its highly regular microcircuit motivating the idea of a universal cerebellar computation. Soon after Eccles et al. [7] recorded the neuroanatomy of the cerebellum, David Marr and James Albus independently proposed an algorithm for the cerebellar microcircuit [25, 26]. Refinements of this theory, stimulated in particular by the recognition that the cerebellum processes time-varying signals related to the control of movement, culminated in the adaptive filter model [27].

In engineering terms, a filter is a device which receives a time-varying input signal and processes it to generate an output signal. For example, a controller which transforms a signal representing a movement intention into an instruction for a piece of motor apparatus. An adaptive filter is able to modify the transformation by adjusting some of its parameters. The adaptive filter model of the cerebellum is an implementation of such a mechanism based on the cerebellar microcircuit. (Although strictly speaking, the adaptive filter model applies only to the cortex of the cerebellum.)

An analysis-synthesis filter consists of a set of fixed filters which analyse an input signal into a large number of components (see Figure 2.2). For example, these could be band-pass filters which decompose the signal into frequency components. The components are subsequently synthesised into an output signal, typically by a weighted linear sum. In the adaptive filter model of the cerebellum, the granular layer performs analysis of mossy fibre signals, and the Purkinje cells synthesise an output signal [5].

Mossy fibre signals are distributed amongst the granule cells, and subsequently transmitted along parallel fibres. Each Purkinje cell receives synaptic input from many thousands of parallel fibres (PF), integrating their signals to generate an output signal. The strength of each PF-Purkinje cell synapse determines the contribution each PF signal makes to the output. Thus, the output of the cerebellar cortex is the weighted sum of parallel fibre sig-
Each Purkinje cell also receives a signal from a climbing fibre, which is assumed to represent an error in the output signal. A learning rule, termed the ‘covariance learning rule’, modifies the PF-Purkinje synaptic weights in order to reduce the size of the error signal [28]. The rule operates as follows: if a parallel fibre is positively correlated with the climbing fibre, its synaptic weight is decreased; if negatively correlated, its weight is increased. Provided that the relationship between system output and error signal is direct (in particular, without significant time delays), this results in a reduction in the error signal [4].

The adaptive filter mechanism has been shown to be consistent with a number of known behaviours including saccadic eye movements [29], the vestibulo-ocular reflex [30] and classical eyeblink conditioning [31]. For example, in delay eyeblink experiments, mossy fibre signals representing the conditioned stimulus (usually a continuous tone) predict an error signal representing the unconditioned stimulus (a puff of air to the eyeball). A conditioned response (an eyeblink) is learnt over a number of trials. Electrophysiological experiments have shown that plasticity at PF-Purkinje cell synapses results in Purkinje cell output which generates the correct timing of this response.

The cerebellum-like structures in fish, whose function is to remove reafference from an
electrotransient signal (Section 2.4), have also been clearly characterised as adaptive filters. In place of climbing fibres, the sensory processing nuclei receive sensory stimuli directly from the periphery. Mossy fibre signals consist of a host of proprioceptive, corollary discharge and electrotransient signals which are likely to be correlated with the sensory reafference. By decreasing the gain of synapses of parallel fibres which are co-active with the sensory signal, the predictable component of the signal is removed.

2.6 The Cerebellar Granular Layer

The granular layer is perhaps the most tantalising part of the cerebellum. Containing roughly 50 billion granule cells, difficulty in making recordings from such tiny cells has made unravelling its function challenging.

2.6.1 Structure of the Granular Layer

The granular layer lies on one of the cerebellar cortex’s two input pathways, and contains granule cells, Golgi cells and other cells such as unipolar brush cells and Lugaro cells. Incoming mossy fibres branch extensively upon entering the cortex, forming synaptic projections termed rosettes. Connections between mossy fibres, granule cells and Golgi cells are made in small bundles of nerve fibre terminals termed glomeruli. A glomerulus is composed of a glial sheath encasing a mossy fibre rosette, granule cell dendrites and Golgi cell dendrites. Each mossy fibre contacts many thousands of granule cells and up to 10-12 Golgi cells [32, 33].

Each granule cell projects a thin, unmyelinated axon upwards to the molecular layer, where it splits into two branches which extend horizontally to form parallel fibres. Parallel fibres run for around 3 mm in each direction, passing at right-angles through the planar dendritic trees of Purkinje neurons, making synaptic connections at dendritic spines.

Golgi cells are the main inhibitory interneurons of the granular layer. They are much bigger than granule cells, and much fewer in number. Golgi cell axons target granule cell dendrites within the glomeruli, and their dendrites extend to the molecular layer, where they receive input from parallel fibres, and to the glomeruli where they receive input from mossy fibres. Golgi cells therefore make both feedforward and feedback connections with granule cells.
Other granular layer cells include Lugaro cells, which are inhibitory interneurons, and unipolar brush cells which are excitatory.

### 2.6.2 Function of the Granular Layer

In the influential Marr-Albus framework of cerebellar function, the granular layer is assumed to perform an ‘expansion recoding’ which increases the diversity of the patterns of activity encoded by the parallel fibres [25, 26]. The adaptive filter model is agnostic as to the mechanism by which diversity of parallel fibre signals is achieved, but the large mossy fibre-granule cell divergence and the complex interconnectivity of the granular layer has led many researchers to speculate on potential mechanisms.

The recoding scheme envisaged by Marr and Albus would be classified as a ‘spatial’ recoding, meaning that multiple mossy fibre signals are recombined non-linearly. This type of recoding is often referred to as ‘multi-modal’, especially when the signals are sensory related. On the other hand, in ‘temporal’ recoding a mossy fibre stimulus results in a number of signals with varied temporal profiles. Because of the requirement of most well-studied cerebellar functions (such as the eyeblink reflex) to generate the precisely timed outputs associated with motor activity, granular layer models are typically designed to generate a temporal recoding.

**Spectral Timing**

Grossberg proposed a mechanism based on granule cells with varied response characteristics [34]. His model was developed in the context of the eyeblink reflex, where the mossy fibres deliver a signal representing the onset of the conditioned stimulus, and the output signal triggers a correctly-timed conditioned response. Granule cells are simultaneously stimulated by the incoming mossy fibre signal, but due to differences in integration rates, parallel fibre signals consist of a spectrum of activation peaks.

The timescales of the types of motor activities usually associated with the cerebellum are of the order of hundreds of milliseconds. For example, in eyeblink conditioning experiments, the typical delay between conditioned and unconditioned stimuli may be up to around 500 ms, far above physiologically realistic values for the response times of neurons. Spectral timing is therefore unlikely to be the mechanism responsible for generating parallel fibre signals in these cases.
On the other hand, at short timescales a mechanism based on neurons with slow response characteristics is quite plausible. Indeed, such a mechanism has been proposed in the electrosensory lateral line lobe, a cerebellum-like structure in weakly electric fish, where the delays involved in reaference suppression are of the order of tens of milliseconds [35].

**Tapped Delay Line**

In electrical engineering, a delay line is a device which introduces a delay between its input and output. A tapped delay line joins a number of delay lines in a chain, resulting in a series of outputs with increasing delays. Moore et al. proposed a tapped delay line mechanism for the generation of parallel fibre signals in the cerebellum, with a number of neurons, each producing a small delay, connected in series [36].

The cerebellar cortex does not have the necessary connectivity to implement a tapped delay line. Moore et al. suggested that it may be implemented in the pontine nuclei, thus introducing delays into the mossy fibre signals, and relegating the granule cells to mere relay elements. This possibility certainly cannot be ruled out on experimental grounds, although it does not provide any explanation for the granular layer’s curious architecture or the huge numbers of granule cells.

**Recurrent Inhibitory Network**

It is well known that recurrent inhibitory networks, consisting of interconnected excitatory and inhibitory neurons, are theoretically capable of generating a wide variety of complex behaviours, including oscillations and storage of memory (see also Section 3.3.1). Buonomano and Mauk [37] suggested that the granule-Golgi-granule feedback loop constitutes such a network in the cerebellar granular layer, and that parallel fibres encode a population vector of dynamic time-varying waveforms.

Recurrent inhibitory networks, connected at random, have been shown to be effective at generating parallel fibre signals suitable for the eyeblink reflex [31, 38]. However, it remains to be confirmed that the connectivity of such networks is consistent with the known connectivity of the granular layer, and with experimental recordings of granule and Golgi cell activity. In Chapter 4 we develop these ideas further in the context of cerebellum-like structures in sharks and electric fish.
2.6 The Cerebellar Granular Layer

Other Theories

Other theories of granular layer function have been proposed. For example, D’Angelo and De Zeeuw introduced the ‘time window matching’ hypothesis which suggests a key role for the feedforward mossy fibre-Golgi-granule cell circuit [39]. Such theories, while consistent with the adaptive filter model, do not propose a central role for the granular layer in generating the temporally diverse signals required by the adaptive filter.

2.6.3 Granule Cells in Cerebellum-Like Structures

Most cerebellum-like structures do not incorporate a granular layer. Instead, a distinct granule cell mass with similarities to the granular layer of the cerebellum is usually present in a nearby but separate brain region [21]. Like the cerebellum, mossy fibres contact granule cells and Golgi cells, and granule cell axons extend into the molecular layer to form parallel fibres.

In sharks and electric fish, Golgi cells are distant from the molecular layer and so Golgi cells do not receive synaptic connections from parallel fibres. For this reason, unlike the cerebellum there is no evidence of a granule-Golgi-granule cell inhibitory feedback loop. Many other minor differences exist between cerebellum-like structures and the cerebellum, the significance of which is not yet known.
3

Computational Modelling of Biological Neural Networks

The nervous system lends itself to study by means of mathematical and computational simulation. Computational modelling may be performed at various levels of detail, from individual ion channels and synapses up to large networks of neurons and entire brain regions. A major challenge is to choose a suitable level of detail for the problem at hand.

This topic of this thesis is the mechanisms implemented by neural circuits in order to carry out their computational functions. As such, I focus on neural connectivity, signal flow and synaptic plasticity, rather than the detailed behaviour of neurons at a cellular or molecular level. Of course, neglecting such details assumes that they do not have an important influence at a system level. These assumptions will be examined in more detail in Section 3.1 and in Chapters 4 and 5, where models are developed in detail.

Another approach dispenses with the aim of biological realism, and seeks to understand general computational principles by examining the behaviour of artificial neural networks. In ‘reservoir computing’ a recurrent neural network is generated at random, resulting in a system which displays such properties as dynamical memory. This idea is explained in
more detail in Section 3.3.1, and applied to models of the cerebellum in Chapter 4.

3.1 Firing Rate Models

When modelling a circuit consisting of a large networks of neurons, it is usual to assume that the dynamics of individual neurons and synapses is not important to the behaviour of the system. Moreover, rather than tracking the times of individual action potentials, information is assumed to be encoded by an average spike rate.

There are many advantages to using firing rate models rather than more detailed models. Firstly, firing rate models are computationally efficient, especially when compared to conductance based models such as the Hodgkin-Huxley equations. This is especially important when modelling large networks consisting of hundreds or thousands of neurons. Secondly, by neglecting cellular level details, we isolate properties of the system which are functions of its connectivity alone. Thirdly, \textit{in vivo} electrophysiological recordings of neurons have, generally speaking, shown a strong correlation between behaviour and spike rates [40]. Fourthly, artificial neural networks can be easily formulated in a manner equivalent to rate coded biological neurons, allowing results from reservoir computing to be incorporated. This will be important in Chapter 4 where we develop echo state network models of the cerebellar granular layer.

3.1.1 Derivation

A typical formulation of the firing rate model is illustrated in Figure 3.1. A neuron is characterised by two variables: $r_i(t)$, representing spike rate, and $z_i(t)$ representing an internal level of activation. The interpretation of $z_i(t)$ depends on the exact method of derivation, and may be associated with either a membrane potential or a synaptic current. However, in the contexts that firing rate models are usually applied, the physical interpretation of the variables is not considered to be important.

The following derivation is adapted from Ermentrout and Terman [41]. We firstly assume that a presynaptic spike arriving at $t = 0$ results in a postsynaptic potential $\Phi_i(t)$ which decays at a rate $\tau_i$. The waveform of the potential is assumed to depend only on the presynaptic neuron $i$. If the rise time of the potential is much smaller than the decay
3.1 Firing Rate Models

Figure 3.1. Firing rate neuron model. Neuron $i$ receives synaptic connections from neurons delivering synaptic drives $z_j(t)$. These are weighted by synaptic strengths $w_j$ and combined to form synaptic current $I_i(t) = \sum_j w_j z_j(t)$. Current is converted to a spike rate $r_i(t) = F_i(I_i(t))$, where $F_i$ is a nonlinear activation function. Finally, synaptic drive $z_i(t)$ is calculated according to the equation $\tau_i \frac{dz_i(t)}{dt} = -z_i(t) + r_i(t - \Delta_i)$ where $\tau_i$ is the time constant associated largely with synaptic currents and $\Delta_i$ is the delay caused by axonal action potential transmission.

If neuron $i$ is spiking at a rate given by $r_i(t)$ then the synaptic drive is given by:

$$z_i(t) = \int_0^t \Phi_i(t-s)r_i(s-\Delta_i)\,ds = \int_0^t \frac{1}{\tau_i} e^{-(t-s)/\tau_i}r_i(s-\Delta_i)\,ds,$$

where $\Delta_i$ represents the axonal delay in the spike arriving at the synaptic terminal. This is equivalent to:

$$\tau_i \frac{dz_i(t)}{dt} = -z_i(t) + r_i(t - \Delta_i). \quad (3.1.1)$$

Synaptic current received from presynaptic neuron $j$ is determined by $z_j(t)$, scaled by a weight variable $w_{ij}$ representing the synaptic strength of the connection between neurons $j$ and $i$. Synaptic currents are assumed to sum linearly, resulting in a total current $I_i(t)$:

$$I_i(t) = \sum_j w_{ij} z_j(t), \quad (3.1.2)$$

time, we can approximate it as a single exponential:

$$\Phi_i(t) = \frac{1}{\tau_i} e^{-t/\tau_i}, \quad t > 0.$$
Synaptic current is translated into a mean firing rate by a nonlinear activation function $F_i(t)$:

$$ r_i(t) = F_i \left( I_i(t) \right). $$

(3.1.3)

The activation function can take many forms, but for biologically realistic networks it usually has a finite range preventing the spike rate exceeding a maximum value or becoming negative. Common activation functions include a binary step function, which implements a threshold dividing the neuron between active and inactive states, and smooth sigmoidal functions which are approximately linear for small values but saturate at large values.

Combining Equations 3.1.1, 3.1.2 and 3.1.3 results in a system of $n$ equations which describes a network of $n$ neurons:

$$ \tau_i \frac{dz_i}{dt} = -z_i(t) + F_i \left( \sum_j w_{ij}z_j(t - \Delta_i) \right). $$

(3.1.4)

In this formulation, the variable $z_i(t)$ is often termed synaptic drive. Inputs to the network can be modelled by additional neurons with prescribed waveforms $z_j(t)$.

This derivation is only one of many ways to derive a firing rate model. Other derivations — such as those based on populations of neurons, or on averaged solutions to conductance-based models — result in the same or very similar equations. For this reason, the physical interpretation of the modelled variables is often ignored.

Despite the simplicity of Equation 3.1.4, highly complicated dynamics are possible even with a small number of neurons, meaning that analytic methods are not usually feasible. Instead, an approximation is calculated using a numerical method such as the Euler method.

How is Equation 3.1.4 used in practice? Broadly speaking, there are two approaches. In the ‘bottom up’ approach, values for $w_{ij}$ are chosen in order to build a specific biological microcircuit. The output of the network can then be validated using electrophysiological data. This is an approach used in Chapter 5 in order to build a ‘winner-take-all’ model of the cerebellar granular layer. In a ‘top down’ approach, a large network is generated with connection weights $w_{ij}$ sampled at random. We use this approach in Chapter 4.
3.2 Recurrent Networks

Figure 3.2. Recurrent network model. The activity of each neuron $z_i(t)$ follows a firing rate equation and is stimulated by external input $u(t)$ and input from other neurons, as determined by input matrix $[w_{in}^i]$ and connection matrix $[w_{ij}]$.

3.2 Recurrent Networks

Traditional artificial neural networks are feedforward networks, where signalling progresses in only one direction, from input to output. In contrast, a recurrent neural network includes feedback loops in its connectivity, a feature ubiquitous amongst neural circuits found in biology, including the cerebellar cortex.

Equation 3.1.4 can be used to simulate a wide variety of neural networks, including recurrent networks. The connectivity is determined by the connection weight matrix $W = [w_{ij}]$ (Figure 3.2). For example, consider the following firing rate model:\(^1\)

\[
\tau \frac{dz_i}{dt} = -z_i(t) + \tanh \left( \sum_j w_{ij} z_j(t) + w_{in}^i u(t) \right), \tag{3.2.1}
\]

where we have assumed that there is no synaptic delay and a uniform value for the time constant $\tau = 100 \text{ ms}$. $u(t)$ is an input signal:

\[
u(t) = \exp \left( -\frac{1}{2} \left( \frac{t - t_0}{\alpha} \right)^2 \right), \tag{3.2.2}\]

\(^1\)This equation differs slightly from Equation 3.1.4 since the firing rate function tanh can take on negative as well as positive values. However, the equations can be considered equivalent by making the transformation $z_i \rightarrow z_i + 1$ and introducing a constant bias input $\sum_j w_{ij}$ to each neuron. Thus in Equation 3.2.1 $z_i$ represents a deviation around a mean firing rate. The bias current can be equivalently (and perhaps more realistically) considered a neuron-dependent firing threshold.
defining a Gaussian waveform of width $\alpha = 10\,\text{ms}$ centred at $t_0 = 100\,\text{ms}$.

For example, two neurons connected in a mutual excitatory-inhibitory pair are defined by:

$$W = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$  \hfill (3.2.3)

Such a network behaves as a coupled oscillator (Figure 3.3 (a)). Similarly, three neurons consisting of one inhibitory and two excitatory neurons behave as an oscillator, albeit of much lower frequency (Figure 3.3 (b)):

$$W = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}. \hfill (3.2.4)$$

For larger networks, it becomes unwieldy to construct connection matrices representing specific network topologies. Instead, random recurrent networks may be generated from matrices whose elements are sampled at random. Figure 3.3 (c) shows a five-neuron network corresponding to a random matrix.

More complicated behaviour, including multiple stable steady-states and chaotic dynamics, can be achieved by introducing neurons with varying firing thresholds. This could be achieved by introducing a new parameter to Equation 3.1.4, but it is usually simpler to instead introduce a ‘bias’ neuron which has a constant output of one. Varying the weights of the connection from the bias neuron is equivalent to varying firing thresholds.

Figure 3.3 (d) shows the output of such a network consisting of 20 neurons with random connection matrix $W$ and input matrix $W_{in}$. The output of the network is chaotic.

As these examples illustrate, recurrent neural networks are able to store memory of their input, and exhibit complicated dynamic behaviour. In fact, it has been shown that any dynamical system can be approximated arbitrarily closely by a recurrent neural network [42].
Figure 3.3. Outputs from recurrent neural networks. Each network is stimulated by an identical stimulus consisting of a brief pulse. (a) Two neurons connected as an excitatory-inhibitory pair generate oscillations. (b) Three neurons, two excitatory and one inhibitory, generate slower oscillations. (c) A five-neuron network with random connection weights which generates stable oscillations. (d) A larger network with random network weights which generates chaotic output. This network contains 20 neurons, and two inputs. The second input is a ‘bias’ input taking a constant value of one, connected via random positive and negative weights, which simulates random firing thresholds.


3.3 Training Recurrent Networks

In the field of machine learning, a typical task is to train a neural network to learn a specified input-output relationship. The ‘backpropagation’ algorithm was one of the earliest methods to be discovered [43]. By repeated application of the chain rule, the relationship between network weights and output error is calculated, allowing network weights to be optimised by gradient descent. The original algorithm only applied to feedforward networks, but variations such as ‘backpropagation through time’ and ‘real-time recurrent learning’ allowed the algorithm to be extended to recurrent networks [44]. Such algorithms were found to be limited to small networks or very simple learning tasks, due to the tendency to get trapped in local optima.

3.3.1 Echo State Networks

Recurrent neural networks are particularly hard to train, due to the complicated relationship between network weights and network output. A novel approach to recurrent neural network training was independently developed by Wolfgang Maass [45] and Herbert Jaeger [46]. While differing in terminology — liquid state machine vs echo state network — their approaches were effectively the same [47], involving restricting training to a limited set of output neurons. The remainder of the neurons have fixed connection weights sampled at random. Echo state networks have been successfully trained to learn a wide variety of nonlinear systems.

A recurrent network is randomly generated, and input signals connected at random. This part is termed the ‘reservoir’ and its connection weights are not altered during training. An output neuron is connected to each of the reservoir neurons, and each of the connection weights calculated by linear regression.

An echo state network consists of $N$ neurons and $M$ input signals represented by $N$- and $M$-dimensional vectors $\mathbf{x}(t)$ and $\mathbf{u}(t)$, where $t = 1 \ldots T$ is the timestep. Network connectivity is defined by $N \times N$ weight matrix $\mathbf{W}$ and $N \times M$ input weight matrix $\mathbf{W}_{\text{in}}$. The state vector $\mathbf{x}$ evolves according to the following update equation:

$$\mathbf{x}(t+1) = -\mathbf{x}(t) + f(\mathbf{W}_{\text{in}}\mathbf{u}(t) + \mathbf{W}\mathbf{x}(t)),$$  \hspace{1cm} (3.3.1)

where $f$ is an activation function, typically a sigmoid such as tanh.
The output signal is given by

\[ y(t) = W_{\text{out}} x(t), \]  

(3.3.2)

where \( W_{\text{out}} \) is an \( 1 \times N \) matrix of output weights. The aim of the method is to determine \( W_{\text{out}} \) such that \( y(t) \) best approximates a prescribed target signal \( y_{\text{target}}(t) \).

\( W_{\text{out}} \) is calculated using regularised linear regression as follows. First, an \( N \times T \) matrix \( X \) is constructed by stacking together the \( T \) column vectors \( x(t) \). Similarly, \( Y_{\text{target}} \) is a \( 1 \times T \) matrix formed from \( y_{\text{target}}(t) \). Next, calculate \( W_{\text{out}} \):

\[
W_{\text{out}} = Y_{\text{target}} X^T (XX^T + \beta I)^{-1}.
\]  

(3.3.3)

The regularisation factor \( \beta \) is important in order to avoid overfitting, and without which the output weights can become unrealistically large.

The weight matrix \( W \) is generated by sampling values at random. The exact statistical distribution is not important, although usually \( W \) is sparse. An important property of the network is the so-called ‘echo state’ property, which demands that in the absence of external input, network activity should eventually decay to zero. This can be achieved by ensuring that the spectral radius of \( W \) (its maximal absolute eigenvalue) is sufficiently small.

Improvements in performance can be achieved using output feedback combined with teacher forcing. Connections are made from the output \( y(t) \) to the reservoir according to a random feedback matrix \( W_{\text{fb}} \), and during training the output neuron is driven with the target signal \( y_{\text{target}}(t) \).

**Example: Mackey-Glass System**

In order to demonstrate the capabilities of an echo state network, Jaeger and Haas trained a network to generate the output of the Mackey-Glass equation [48]. The Mackey-Glass system is a non-linear dynamical system which can display a variety of chaotic and periodic behaviour and has become a model system for testing machine learning algorithms.

Figure 3.4 shows the results of training a network consisting of 1000 neurons against the Mackey-Glass time series. The network was trained against 2000 ms of data using output forcing, the output weights calculated using linear regression, and then the simulation continued. The figure shows a 400 ms sample of output, along with the output from three
Figure 3.4. The Mackey-Glass time series (green) and output from the echo state network after training (blue) (top). The activity of three neurons in the reservoir (bottom).

of the reservoir neurons. The time-series was discretised at a 1 ms timestep.

Examples such as this demonstrate how a random recurrent network is capable of encoding a large variety of dynamical systems. Training is limited to output weights, which is more biologically realistic than algorithms which require adjustments to connections deep within the network.

3.3.2 FORCE Learning

FORCE learning is a neural network training technique based on the ideas of echo state networks proposed by Sussillo and Abbott [49]. In contrast to the technique described in the previous section, which was based on a discrete time model, it incorporates a continuous time model equivalent to a simple firing rate neuron model (Section 3.1). It also incorporates an online learning rule, which is not only more biologically realistic, but enables a much greater degree of stability.

Like the echo state network from the previous section, FORCE learning requires a reservoir of randomly connected neurons. The reservoir consists of a sparsely connected set of inhibitory and excitatory neurons, tuned such that the network generates spontaneous chaotic activity. In general, chaotic activity is considered an undesirable property in neural networks, whereas it is a key requirement for FORCE learning.
3.3 Training Recurrent Networks

An output neuron receives connections from the reservoir, and sends its output back to the reservoir via a set of random feedback connections. The network output is a linear weighted sum of the activities of the reservoir neurons. The FORCE algorithm adjusts these weights in order to approximate a target function $f(t)$:

$$z(t) = w^T r(t),$$

where $z(t)$ is the network output, $w$ is the output weight vector and $r(t)$ is the vector of reservoir neuron activity.

The FORCE learning algorithm updates the output weights online, meaning that they are updated continually during learning. Unlike traditional online learning algorithms, in which output error slowly converges to zero during learning, FORCE learning adjusts output weights rapidly, ensuring that the error is always close to zero, and the output feedback is kept close to the target function. Thus early in the learning process, the weight vector $w$ fluctuates rapidly before converging to a stable value.

The weight update algorithm is based on recursive least-squares. $P(t)$, a running estimate of the inverse of the correlation matrix of the network rates, is calculated as follows:

$$P(t) = P(t - \Delta t) - \frac{P(t - \Delta t) r(t) r^T(t) P(t - \Delta t)}{1 + r^T(t) P(t - \Delta t) r(t)},$$

$$P(0) = (1/\alpha) I,$$

where $\Delta t$ is the interval of time between updates of the readout weights and $\alpha$ is a regularisation parameter. Error is defined as

$$e(t) = w^T (t - \Delta t) r(t) - f(t),$$

and weights are updated using the following rule:

$$w(t) = w(t - \Delta t) - e(t) P(t) r(t).$$

Sussillo and Abbott demonstrated the FORCE learning algorithm using a number of tasks, including periodic, discontinuous, chaotic and noisy targets, as illustrated in [49]. In the next section, I demonstrate how this learning algorithm may be used in the context of modelling the cerebellum.
3.4 Forward Models

In the theory of motor control, an ‘internal model’ is a virtual representation which mimics some aspect of the operation of motor apparatus. For example, an ‘inverse model’ takes a motor signal representing an intended limb movement, and translates it into a signal representing a motor command for the appropriate muscles (Figure 3.5 (a)). Masao Itô first proposed that the cerebellum encodes internal models, which through a learning process enable the body to adapt to changes in the dynamic characteristics of limbs, muscles and joints [50].

The cerebellum is known to make an important contribution to the co-ordination of motor
activity, but does not itself initiate movements. Voluntary limb movements are initiated by brain centres in the cerebral cortex, which issue instructions to the motor cortex. The motor cortex acts as a controller, generating motor commands which ultimately drive motor apparatus. The cerebellum sits on a side loop from the motor cortex, forming the cerebrocerebellar communication loop (Figure 3.6) [3].

What is the contribution of the cerebellum to the generation of movement signals? Many theories have been postulated, such as that the cerebellum behaves as a Smith predictor which compensates for the delays in feedback signals, or a Kalman filter for state estimation [3]. Porrill et al. [4] review in detail the evidence for various types of proposed models, and conclude that while many are unlikely to be biologically realisable, the cerebellar microcircuit is highly suited to encode forward models which predict the sensory consequences of motor activity (Figure 3.5 (b)).

The forward model theory of cerebellar function is supported by evidence from cerebellum-like structures in sharks and weakly electric fish, which have been clearly demonstrated to encode forward models [21]. These structures cancel reafference by generating a signal which predicts and cancels the sensory consequences of self-generated activity. In sharks, the DON cancels electrosensory stimuli generated by breathing motion, whereas in electric fish, the ELL cancels the stimulus caused by an electric organ discharge. In these cases, the forward model encodes a simple relationship between command signal and sensory signal, representing the delays related to signal transduction and the dynamics of muscle movement or electric organ discharge. In contrast, sensory predictions made by the cerebellum would be expected to be more complicated, perhaps involving multiple
input signals relating proprioceptive and sensory information, which may be combined in various combinations.

### 3.4.1 Encoding Forward Models Using Recurrent Networks

In this section, I investigate the capabilities of recurrent networks in encoding forward models. I do this by using a neural network learning algorithm to construct networks that reproduce a prescribed input-output transformation. I follow a method developed by Sussillo and Abbott [49] which they term FORCE learning. This method is based on ideas developed in the context of echo state networks, in which a random recurrently connected network serves as a dynamical reservoir. A suitable output signal is extracted by means of a synaptic learning mechanism on the synaptic weights of a readout neuron.

Figure 3.7 illustrates the neural network learning mechanism. All neurons are governed by a simple firing rate equation (see Section 3.1), except for the output neuron whose activity is a linear sum weighted by output weight vector \( w_{\text{out}} \). Network connection weights \( w \) and input connection weights \( w_{\text{in}} \) are sampled randomly such that the network spontaneously generates chaotic activity. Feedback from the output neuron is fed back to the network via feedback connections \( w_{\text{fb}} \).

In order to train the network, an input signal \( u(t) \) and associated target signal \( f(t) \) are selected. Training the network is achieved by applying the input signal \( u(t) \) to the network, while rapidly making adjustments to the output weights \( w_{\text{out}} \) in a manner which keeps the output signal \( z(t) \) close to the target signal \( f(t) \).

Figure 3.8 shows a network learning to produce an output spike a short fixed time interval after receiving an input spike. Initially, output weights and output signal are zero. After learning is complete, the network spontaneously generates an output which closely matches the target signal.

Figure 3.9 Illustrates the output of a recurrent network containing 200 neurons after being trained to encode a variety of forward models. In each case, the target signal (pink line) is related to the input signal (blue line) by a simple rule.

The cerebellum is often associated with functions which require the generation of precisely timed signals. For example, in eyeblink conditioning experiments, the cerebellum generates a signal which determines the onset of an eyeblink motor command, a fixed time interval after a conditioned stimulus (Section 2.3.1). Likewise, experiments on rea-
Figure 3.7. Training a recurrent neural network using the FORCE learning algorithm. A recurrent network, determined by network connection weights $w$ and input connection weights $w_{\text{in}}$, serves as a dynamical reservoir. Output weights $w_{\text{out}}$ and feedback weights $w_{\text{fb}}$ determine connectivity to an output neuron. All weights have fixed random values, apart from $w_{\text{out}}$ which is continuously updated according to the learning rule. The network is stimulated by input $u(t)$, which causes network activity leading to output signal $z(t)$. The learning rule updates $w_{\text{out}}$ so as to reduce the error in the approximation of $f(t)$ by $z(t)$.

Figure 3.8. A recurrent network learns a fixed-delay target signal. The input consists of a random sequence of narrow spikes. The target is a sequence of broader spikes, delayed by a short fixed interval. Initially, learning is off, and the output weights $w_{\text{out}}$ and output signal are zero. When learning is turned on, the norm of the output weight vector increases, and the output closely follows the target signal. After learning is turned off, the output continues to match the target.
Figure 3.9. Training recurrent networks to encode forward models. A recurrent network of 200 neurons is trained using the FORCE learning algorithm to generate a number of forward models. One or more input signals (consisting of a random train of spikes) are shown in blue, and the target signal is shown in pink. The output of the network after training is shown in green. (a) The target is a broad spike a fixed delay after each input spike. (b) An output spike is generated only when two input spikes are received simultaneously. (c) An output spike is generated only when exactly one input spike is received. (d) High frequency noise is added to the target signal. (e) Adding noise of a similar frequency to the target signal disrupts learning. (f) Five target signals with a range of delays with respect to the input.
ference suppression in sharks [51] and electric fish [52] require the generation of a signal whose timing matches an artificial stimulus timed with respect to the breathing cycle or the electric organ discharge command respectively. These transformations are illustrated in Figure 3.9 (a), where the target signal is a broad spike a fixed time interval after each input spike.

In addition to temporal transformations, the cerebellum is thought to implement spatial transformations, where multiple inputs are combined in various combinations. (In fact, this was the focus of the original Marr-Albus theory of cerebellar function). For example, the integration of multiple proprioceptive, sensory and motor inputs would certainly be required for the generation of signals related to the movement of multiple interacting joints. The multi-modal integration of mossy fibre inputs has recently been discovered in the cerebellar granular layer [53].

Simple spatial transformations include logical operations such as co-incidence detection (the logical AND of two inputs), which is illustrated in Figure 3.9 (b). A slightly more challenging operation is the detection of precisely one of two inputs (the logical XOR of two inputs), Figure 3.9 (c).

In general, we would expect signals in biological circuits to include a large component of noise, either intrinsically generated or as part of a sensory signal. Fig 3.9 (d) and (e) show learning in the presence of high frequency and low frequency noise respectively. High frequency noise, which might be caused by the stochastic nature of spike trains, does not disrupt the learning. Where the noise is of a frequency similar to the target signal itself, the learning is less successful.

In adaptive filter models of the cerebellum, it is usually assumed that the recurrent network constituting the granular layer generates a varied but fixed set of basis signals. For example, signals which have a peak of activity at a range of delays with respect to the input signal [54]. Figure 3.9 (f) shows the output of a network which has been trained to produce five Gaussian waveforms.

### 3.4.2 Discussion

In the previous section, I outlined a number of examples of forward models generated by recurrent neural networks. While many of these might seem simple, they represent typical transformations required in making predictions of the sensory consequences of motor activity, and hence the types of transformations associated with the cerebellum.
It is important to note that the learning algorithm used is not biologically realistic. As noted by Sussillo and Abbott, the learning rule is not synapse-specific and requires changes to synaptic strength at a physiologically unrealistic speed [49]. As such, I do not claim that this type of algorithm underlies cerebellar learning; rather that recurrent neural networks (constructed by some unknown developmental or learning program) are capable of the types of tasks usually associated with the cerebellum.

It is natural to ask what sort of connectivity results from the learning algorithm, and by what means it generates the desired output. However, as pointed out in [49], only the output weights are modified during the learning process, the rest remaining at random values. For example, in a 100 neuron network, 90% of connections are fixed. It is therefore unlikely that many conclusions could be drawn by examining the post-learning network weights.

Nevertheless, this learning algorithm is a useful technique for building models of biological neural networks. In the next chapter, which compares the performance of various granular layer models against a number of learning tasks, I construct a model of the cerebellar granular layer using this technique. By combining a network constructed using the FORCE learning algorithm with an output neuron which learns according to the adaptive filter, we arrive at a realistic model of the cerebellum.
As explained in Chapter 2, the cerebellum shares many neuroanatomical properties with cerebellum-like structures in sharks and electric fish. On the other hand, there are differences in connectivity, as well as probable differences in function, between these structures. In this chapter I develop a series of computational models which allow these differences to be examined.

This chapter was published as a research paper in *The Cerebellum* [6]. Minor clarifications requested by the thesis examiners have been annotated as footnotes to the text.
Abstract

In the adaptive filter model of the cerebellum, the granular layer performs a recoding which expands incoming mossy fibre signals into a temporally diverse set of basis signals. The underlying neural mechanism is not well understood, although various mechanisms have been proposed, including delay lines, spectral timing and echo state networks. Here, we develop a computational simulation based on a network of leaky integrator neurons, and an adaptive filter performance measure, which allows candidate mechanisms to be compared. We demonstrate that increasing the circuit complexity improves adaptive filter performance, and relate this to evolutionary innovations in the cerebellum and cerebellum-like structures in sharks and electric fish. We show how recurrence enables an increase in basis signal duration, which suggest a possible explanation for the explosion in granule cell numbers in the mammalian cerebellum.
4.1 Introduction

The adaptive filter model of the cerebellum has become widely accepted, neatly explaining its computational function as well as accounting for its distinctive neural circuitry [5, 27]. At the core of the adaptive filter model is an ‘analysis-synthesis’ filter, whereby incoming mossy fibre signals are recoded into an expanded set of basis signals which are then recombined into an output signal conveyed by Purkinje cell axons. The adaptive filter mechanism is mediated by an anti-Hebbian learning rule which implements LTD and LTP at synapses from parallel fibres onto Purkinje cells, depending on the coincidence of parallel and climbing fibre input. Given a suitable error signal, delivered by climbing fibres and encoded as complex spikes, this mechanism results in a covariance learning rule [28] which minimises the error in the output signal.

In addition to the cerebellum itself, there exist a range of ‘cerebellum-like’ structures which share the cerebellum’s core architecture: a distinct molecular layer containing uniformly arranged parallel fibres arising from a densely-packed granular layer. Examples include the dorsal cochlear nucleus in mammals as well as a number of sensory processing nuclei in various fish species [21, 55]. Amongst these, the electrosensory lateral line lobe (ELL) in electric fish [52, 56] and dorsal octavolateralis nucleus (DON) in sharks [57] have been clearly characterised as adaptive filters.

The function of such cerebellum-like sensory processing nuclei is to cancel the ‘reafferent’ component of an incoming sensory signal, which is the predictable component of the signal caused by the animal’s own voluntary movements. Over time, the sum of parallel fibre input closely matches and cancels the predictable component of the signal, leaving a residual component consisting of behaviourally relevant exafferent stimuli, which is passed on to higher brain centres. That this capability emerges naturally from a local synaptic learning rule highlights the elegant simplicity of the adaptive filter model.

It is commonly proposed that the function of the cerebellum is to encode ‘internal models’ of motor and sensory apparatus [12]. In particular, a ‘forward model’ encodes the predicted sensory consequences of motor action. This is predictive in the sense that it is based on the corollary copy of the motor command system in advance of the feedback that actually occurs as a result of the movement itself. However, given the complexity of motor control, perhaps the simplest illustration of how the cerebellar circuit implements an adaptive filter in order to build a forward model is the suppression of electrosensory reaference by the cerebellum-like sensory nuclei of fish.
Table 4.1. Many neurons are common to the cerebellum and cerebellum-like structures such as the electrosensory lateral line lobe (ELL) in electric fish and dorsal octavolateralis nucleus (DON) in sharks.

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mossy fibre</td>
<td>Nerve fibre conveying external input to the granular layer. In the ELL and DON, mossy fibre signals consist largely of proprioceptive signals, corollary discharges and higher modalities of sensory information.</td>
</tr>
<tr>
<td>granule cell</td>
<td>The most numerous cell in the granular layer, granule cells are excitatory neurons whose axons form parallel fibres which terminate on principal neurons.</td>
</tr>
<tr>
<td>Golgi cell</td>
<td>Inhibitory interneuron of the granular layer which, in the cerebellum, makes feedforward and feedback connections with granule cells. In the ELL and DON, only feedforward connections have been reported.</td>
</tr>
<tr>
<td>unipolar brush cell</td>
<td>Excitatory neuron of the granular layer which produces a distinctly prolonged response to stimulation. Reported in the cerebellum and ELL but not in the DON.</td>
</tr>
<tr>
<td>principal cell</td>
<td>General term for the cell which constitutes the output of a cerebellum-like structure. Equivalent to the Purkinje cell in the cerebellum.</td>
</tr>
</tbody>
</table>

As well as a passive electrosense, used for navigation and sensing prey, weakly electric fish emit electric pulses for communication and active electrolocation. These electric organ discharges results in an unwanted ‘ringing pattern of activation’ of the passive electroreceptors [35, 58]. This reafferent stimulus is suppressed by the electrosensory lateral line lobe (ELL), a cerebellum-like sensory processing nucleus which implements an adaptive filter as described below and illustrated in Figure 4.1.

The output of the ELL is transmitted by the axons of ‘principal cells’ which are the functional equivalent of cerebellar Purkinje cells. Inputs to a principal cell are of two types. Firstly, an excitatory afferent input conveys electrosensory signals from the periphery. Secondly, many thousands of parallel fibres convey, inter alia, recoded versions of the EOD command [56]. This recoding is thought to be performed by the eminentia granularis posteria (equivalent to the cerebellum’s granular layer) which receives EOD command copies as mossy fibre input [59].

Each time the fish issues an EOD command, the principal neuron is stimulated simultaneously by both sets of inputs. Importantly, because the EOD command is a brief stereotyped pulse, the temporal profile of both sets of inputs is near identical each time. Anti-Hebbian plasticity at synapses from parallel fibres onto principal cells causes parallel fibres which are active simultaneously with the afferent stimulus to be modulated in a manner which depresses the activity of the principal neuron [60]. Over time, the synaptic
Figure 4.1. Cerebellum-like structures implement an adaptive filter in order to suppress electrosensory reafference. Mossy fibre signals (red) are recoded by the granular layer and then conveyed by parallel fibres to a principal neuron. Plasticity at synapses from parallel fibres onto principal neurons results in cancellation of the incoming electrosensory signal. In this example, the electrosensory signal is entirely reafference, so the output is zero. Mossy fibres convey a wide variety of proprioceptive signals, efference copy and higher levels of sensory information, but we show here only those signals correlated with the reafferent signal. In weakly electric fish (A), relevant mossy fibre signals consist of a relatively sparse electric organ discharge (EOD) command copy whereas in sharks (B) mossy fibres convey a temporally rich set of breathing-related motor commands.
4.1 Introduction

![Diagram](image)

**Figure 4.2.** Four granular layer models. (a) A single cell network is included as a reference case. (b) A feedforward network consists of a single layer of cells with random time constants and input weights. (c) A random recurrent network consists of cells with identical time constants but random positive and negative connection weights. (d) A trained recurrent network consists of a network whose connection weights have been selected using the FORCE learning algorithm. In (b), (c) and (d) there are exactly 60 cells per mossy fibre.

weights of the synapses converge to values such that the sum total of the parallel fibre contribution forms an negative image of the afferent input. This results in cancellation of the sensory signal, and zero output from the principal neuron.

Of course, in the animal’s environment, the sensory stimulus does not consist merely of interference caused by electric organ discharge, but includes behaviourally relevant stimuli, such as fields generated by prey or conspecifics. However, because such stimuli are not correlated with the EOD command, they do not contribute to net changes in synaptic strength. Once the reafferent signal has been cancelled, such uncorrelated signals constitute the output of the sensory nucleus.

This example illustrates the importance of the temporal recoding implemented by the granular layer. Since the negative image is the weighted sum of parallel fibre signals, only negative images that lie within the span of such signals may be formed. Thus, in this case, the function of this recoding is clear: from a brief EOD command pulse (about 10 ms in duration), generate a temporally diverse set of parallel fibre signals spanning the time course of the interference (about 200 ms). Later in this paper we investigate potential mechanisms underlying this recoding, and the implications of the recoding on the capabilities of the adaptive filter.

A second example of a cerebellum-like structure whose function is to cancel electrosensory reafference is the DON, a sensory nucleus found in the shark. The functional architecture
of the DON is identical to that of the ELL, including a dense granular layer and learning implemented by plasticity of synapses of parallel fibres onto principal neurons [51]. As with the ELL, the function of the nucleus is to predict and cancel a self-generated electrosensory signal, but unlike weakly electric fish, sharks do not possess an electric organ. The source of the reafference is electrical fields generated by the animal’s movements, chiefly those associated with breathing motion [57, 61].

The reafferent stimulus consists of a large, slowly varying waveform which repeats each breathing cycle (about 2 seconds in duration). Mossy fibre signals convey, _inter alia_, efference copy of breathing-related motor commands. Since there exist breathing muscles activated throughout the breathing cycle, mossy fibre signals are distributed throughout the duration of the reafferent stimulus.

Figure 4.1 contrasts the two cases. While in both animals mossy fibre consists of a wide variety of sensory, proprioceptive and efference copy signals, only signals correlated with the reafferent signal are relevant to the adaptive filter. In electric fish, this is the EOD command; in sharks a diverse set of breathing commands. Later, we will argue that the distinction between ‘temporally sparse’ and ‘temporally rich’ mossy fibre input has important implications for the recoding implemented by the granular layer.

While the neural circuitry in the cerebellum proper differs slightly to the cerebellum-like structures described here, the role of the cerebellum in generating accurately timed output signals is clear [62]. For example, eyeblink conditioning experiments demonstrate the cerebellar circuit’s capability to generate eyeblink command signals which anticipate an aversive stimulus. It is widely assumed that this timing information is generated within the granular layer [63, 64].

The examples described here illustrate the role of the granular layer in increasing the temporal diversity of parallel fibre signals. But by what mechanism does the granular layer achieve this? In part due to the the small size of granule cells, and the consequent difficulty of taking recordings from them, experimental evidence is in short supply. Nevertheless, inspired by the granular layer’s strikingly dense interconnectivity, a number of mechanisms have been hypothesised.

Moore et al. [36] proposed that chains of neurons generate outputs with various delays, in the manner of a tapped delay line. More recently, mechanisms based on the more sophisticated idea of an ‘echo state network’ have gained currency.

Echo state networks developed from the observation that recurrent inhibitory networks,
incorporating neurons which send feedback signals to each other, are capable of generating complicated output signals [65]. When excited by an external stimulus, oscillations are produced which outlast the duration of the original stimulus [66]. The observation that the cerebellar granular layer circuit incorporates an inhibitory recurrent loop formed by connections between granule and Golgi cells has led to the conjecture that this is its underlying mechanism [31, 38].

Others propose that rather than emerging from network effects, the recoding is the result of intrinsic neuron properties. For example, ‘spectral timing’ is based on neurons which respond with various time constants [67]. The discovery of unipolar brush cells (UBCs), which respond to stimulation with a markedly prolonged output, may lend credence to this viewpoint.

The structure of the granular layer is not uniform across all cerebellum-like structures, nor even between different cerebellar microzones. In the cerebellum, UBCs are common in the vermis and especially in the vestibulocerebellum [24], and amongst cerebellum-like structures have been detected only in the dorsal cochlear nucleus (DCN) and ELL [21]. Recurrent connectivity between Golgi and granule cells is found throughout the granular layer of the cerebellum but is unlikely to not exist in either the DON or ELL. These differences lead to the conjecture that the granular layer recoding mechanism may be dependent on function, and later in this paper we speculate that they may relate to a hierarchy of evolutionary innovations in the granular layer.

The purpose of this study is to relate the structure of the granular layer, and hence the basis signal recoding, to the I/O transformation achieved by the filter. To this end, we develop:

i. A generic computational framework, based a network of leaky-integrator neurons. Parameters determine intrinsic neuron properties and network connectivity.

ii. A performance measure which quantifies the capability of the network. In effect, we calculate the accuracy to which the network output is capable of approximating a range of target signals.

By carefully prescribing model parameters, we construct four granular layer models (single cell, feedforward, random recurrent and trained recurrent; illustrated in Figure 4.2). Each model is tested against ‘sparse’ and ‘rich’ mossy fibre scenarios. We find that richness of mossy fibre input, as well as the complexity of the network structure, has a large influence on the capabilities of the adaptive filter.
Finally, we examine the effect of increasing the number of neurons in the network. We find that only with recurrent connectivity does increasing the number of neurons result in an unbounded increase in performance, and speculate that this relates to an explosion in granule cell numbers in the mammalian cerebellum.

4.2 Modelling

4.2.1 Granular Layer Model

We simulate the granular layer by a network of leaky integrator neurons (Figure 4.3). Each neuron in our model represents a granular layer cell (which might be a granule cell, Golgi cell or another cell such as a UBC) whose activity is described by a variable representing an instantaneous firing rate. Neuronal activity is driven by incoming connections from mossy fibres and internal connections from other granular layer cells. See appendix 4.5.1 for the equations describing the neuron model.

The model is specified by parameters which describe neural time constants $\tau_i$, incoming connections $w_{ij}^{\text{in}}$ and internal connections $w_{ij}$, where $w_{ij}^{\text{in}}$ represents the strength of the connection between mossy fibre $i$ and neuron $j$, and $w_{ij}$ represents the strength of the connection between neurons $i$ and $j$. Positive and negative values represent excitatory and inhibitory connections respectively. This framework allows us to specify a range of types of granular layer in a uniform manner.\(^1\)

4.2.2 Mossy Fibre Input

As discussed, we assume that the function of the cerebellar circuit is to implement an adaptive filter which generates a forward model of a target I/O transformation. In other words, to closely approximate a target signal given a specified mossy fibre input. We consider two scenarios for the set of mossy fibre inputs: a ‘sparse’ scenario, consisting of a single mossy fibre conveying a single discrete pulse, and a ‘rich’ scenario consisting of a number of mossy fibres, each of which conveys a discrete pulse at various delays (see Appendix 4.5.2).

\(^{1}\)Note that the suffix notation is the reverse of that in Chapter 3. If there are $N$ network neurons and $M$ mossy fibres, then we have network weights $w_{ij}$ for $i, j = 1 \ldots N$. On the other hand we have input weights $w_{ij}^{\text{in}}$ for $i = 1 \ldots M$ and $j = 1 \ldots N$. 
4.2 Modelling

**A** Granular layer model

- mossy fiber signals
- basis signals
- synaptic weights
- time constants

**B** Performance measure

- basis signals
- weighted sum
- target signal approximation
- error = ê

**C** Mossy fibre signals

- sparse
- rich

**Figure 4.3.** (A) The granular layer model. The granular layer (enclosed by box) receives mossy fibre signals $m_i(t)$ and outputs parallel fibre signals $r_i(t)$. The granular layer is specified by input connection weights $w_{ij}^{in}$, internal connection weights $w_{ij}$ and time constants $\tau_i$. (B) The adaptive filter learning rule is assumed to result in output weights $w_{ij}^{out}$ such that total output $\sum_i w_{ij}^{out} r_i(t)$ equals the least-mean-squares approximation to the target signal $f(t)$. (C) Two classes of mossy fibre input. A ‘sparse’ scenario consisting of a single mossy fibre signal, and a ‘rich’ scenario consisting of multiple mossy fibre signals distributed across the entire time frame.
Figure 4.3 (C) illustrates the waveforms of the two scenarios, which roughly correspond to the distinct sensory stimuli received by electric fish and sharks.

4.2.3 Performance Measure

In order to assess the efficacy of the adaptive filter, we measure how closely the output of the granular layer is capable of approximating a range of prescribed target signals. Later we discuss why this is an appropriate measure of performance, even though in many models of cerebellar function an explicit target signal does not exist.

We generate target signals by sampling at random from a distribution of smooth signals (see Appendix 4.5.3). By calculating the average error in approximating a large number of target signals, we calculate a value which represents the performance of the set of basis signals. In the adaptive filter model, the output of the Purkinje-like cell is the weighted sum of parallel fibre signals, and it can be shown that learning results in weights which minimise the least-mean-square difference between the output and target signal \( f(t) \) [68]. The least-mean-squares error \( e \) in this approximation then represents how effective the granular layer is at generating basis signals for approximating the target signal.

By averaging over a large number of such target signals, we calculate a value \( \hat{e} \) which describes the filter’s performance (see Appendix 4.5.4). A small value for \( \hat{e} \) indicates that the filter is capable of closely approximating a wide variety of target signals.

4.3 Results

4.3.1 Single Cell

For reference, we first consider a very simple granular layer network consisting of a single cell with a small time constant \( (\tau = 10 \text{ ms}) \). In this case, the granular layer does not perform any recoding, and parallel fibre signals are an almost exact copy of mossy fibre signals.

Figure 4.4 shows the results of this network when applied to the two mossy fibre scenarios. In neither case does the resulting basis set enable a good approximation to the target signal.
4.3 Results

Figure 4.4. A single cell granular layer network. In this and the following figures, the top line (red) shows the mossy fibre input signal(s) and the second line(s) (black) shows the output from the granular layer. The bottom three rows show three randomly generated target output signals (blue) as well as the best possible least-mean-squares approximation to each (black). Left: sparse mossy fibre input. Right: rich mossy fibre input. In neither case is a good approximation achieved.

Figure 4.5. A single layer feedforward granular layer network. With sparse mossy fibre input, the approximation is good only for a brief duration, whereas in the rich case the approximation is good for the full duration.

4.3.2 Single Layer Feedforward

We built a granular layer network which implements ‘spectral timing’, where the diversity of signals is generated by the variety of neuron time constants. The network consists of a single-layer feedforward network, with a variety of input weights and time constants. In our simulation, each mossy fibre diverges to 60 granular layer cells, with random time constants between $\tau_i = 1\,\text{ms}$ and $\tau_i = 150\,\text{ms}$, and input weights between $w_{in}^{ij} = 1$ and $w_{in}^{ij} = 20$.

As can be seen in Figure 4.5, the granular layer recoding increases the diversity of parallel fibre signals. In the rich scenario, the basis set is diverse enough to approximate the target signal over its entire time course. On the other hand, in the sparse mossy fibre scenario, the basis set is sufficient to approximate the target function only within a short time window with respect to onset of the mossy fibre signal.
The duration of this time window is, presumably, determined by the values of the intrinsic parameters $\tau_i$ and $w_{ij}^{in}$ that characterise the granular layer neurons. The electrophysiological properties of granular layer neurons are not well understood, but in general the maximum response of a neuron to a stimulating current is of the order of tens of milliseconds. In this simulation, we restricted the envelope of parameter values $\tau_i$ and $w_{ij}^{in}$ to enforce this criterion.

On the other hand, it is notable that a specific type of granular layer cell — the unipolar brush cell (UBC) — is characterised by generating a prolonged burst of firing in response to stimulation [35]. Motivated by this observation, we performed a further simulation which includes a much broader range of values input weights between $w_{ij}^{in} = 1$ and $w_{ij}^{in} = 3000$.

As can be seen in Figure 4.6, introducing such long-delay neurons into the network enables a further improvement by increasing the duration over which the target signal may be approximated. However, it is notable that the capability of the network remains limited by the intrinsic neuron properties.

### 4.3.3 Random Recurrent Network

In order to build a granular layer network which is capable of generating a temporally diverse output beyond the time window limitation described above, we introduce recurrent connections into the network. Recurrent neural networks — such as echo state networks [46] and liquid state machines [45] — are known to be capable of generating temporally varied yet predictable output. Furthermore, the recurrent nature of the granular layer has stimulated the construction of recurrent models of the cerebellum, such as the model proposed by [31]. Such networks generally consist of a randomly connected network of excitatory and inhibitory neurons, the distribution of network parameters being selected in order to generate a dynamically complex yet non-chaotic output.

We constructed a recurrent network consisting of 60 cells. Time constants of the cells were set to a uniform value of $\tau_i = 50$ ms. Each cell was connected to, on average, 10% of cells and connection weights were sampled from a zero-mean normal distribution, resulting in approximately equal numbers of excitatory and inhibitory connections. The weight matrix was then scaled to ensure that in the absence of stimulation the network activity decayed to a steady state.

As illustrated in Figure 4.7, the random recurrent network generates a set of basis signals which are temporally diverse across the entire time course of the target signal. This results
4.3 Results

Figure 4.6. A single layer feedforward granular layer network with long-delay neurons. Introducing cells with slow dynamics increases the duration over which a good approximation can be achieved, but this duration is constrained by the intrinsic neuron properties.

Figure 4.7. A random recurrent network. Inhibitory recurrent connections result in a set of basis signals of sufficient diversity to approximate the target signals reasonably well over the entire duration.

in a reasonably good approximation to the target signal.

4.3.4 Trained Recurrent Network

Is it possible — through a judicious choice of connection weights $w_{ij}$ — to construct a network which performs better than one constructed at random? A more sophisticated approach, developed by [49], allows a network to be constructed which generates an arbitrarily prescribed output. ‘FORCE learning’ is a neural network training algorithm which takes a random recurrent neural network, then trains the synaptic weights of a set of output neurons until their output matches a set of target waveforms.

We followed their approach in order to build a granular layer network which, in response to mossy fibre input, generates a set of Gaussian waveforms with varying delays with respect to the input signal (see Appendix 4.5.5). This type of basis filter has been used successfully in models of the eyeblink reflex [54]. A network of 60 neurons is sufficient to generate a set of twenty Gaussian waveforms spanning the full time course of the target
Figure 4.8. A trained recurrent granular layer network. Using the FORCE learning algorithm, a network was constructed which generates a set of Gaussian waveforms. This results in a very good approximation.

We tested this network against a sparse mossy fibre input. As illustrated in Figure 4.8, the resulting adaptive filter model is capable of approximating the target signal across its entire time course.

4.3.5 Numbers of Neurons

A striking feature of the cerebellar granular layer is the large number of neurons. Motivated by this observation, we tested how the number of neurons affects the performance of the adaptive filter.

We measured performance by calculating the maximum duration over which the output of the network is capable of accurately approximating target signals (see Figure 4.9, top). For each of three network types (feedforward, random recurrent and trained recurrent), and for numbers of neurons ranging from 1 to 200, the simulation was executed, and this maximum duration calculated.

The results are shown in Figure 4.9. Whereas the performance of the feedforward network quickly saturates at a value constrained by intrinsic neuron properties, the performance of the recurrent networks increases without bound. The performance of the trained recurrent network in particular is superior to the random recurrent network.
Figure 4.9. Network performance against the number of neurons. Network performance is defined as the maximum duration over which the least-mean-squares approximation remains close to the target (top). The performance of the feedforward network (red) quickly saturates, whereas the performance of the recurrent networks improves as more neurons are added. The trained recurrent networks (blue) shows consistently better performance than the random recurrent network (white).
4.4 Discussion and Conclusions

The granular layer models presented in this paper represent a hierarchy of increasing complexity, each level introducing a new innovation in network structure. The first simulation demonstrates the requirement for the granular layer: without any recoding, the capability for signal approximation is poor, even with a relatively rich mossy fibre input. The second simulation shows that while a simple feedforward granular layer is capable of approximating a target signal, the duration is constrained by the intrinsic neuron firing properties. A recurrent network enables good approximation of a large time window, and if we carefully construct the network connectivity, the duration is constrained only by the number of neurons in the network.

This hierarchy reflects a similar hierarchy of innovations in the evolution of the cerebellum and cerebellum-like structures. Of particular relevance are two innovations in the structure of the granular layer. The first is the unipolar brush cell, which exists in the mammalian cerebellum as well as the ELL and is an excitatory granular layer cell which exhibits a peculiarly long-latency burst response. [35] suggest that the function of UBCs is to produce parallel fibre signals active at durations beyond the time window of the efference copy of the temporally sparse EOD command, a result illustrated clearly by our modelling. We note that UBCs have not been reported in the DON [21].

The second innovation is the existence of recurrent inhibitory loops from parallel fibres to granule cells via Golgi cells, a feature of the cerebellum [69] but not reported in cerebellum-like structures such as the ELL or DON [21]. It has long been understood that recurrent neural networks are capable of storing a memory of network input, a necessary condition for the generation of temporally complex output. Our modelling demonstrates how recurrent connections enable the production of a set of basis signals whose temporal profile exceeds the restriction imposed by the intrinsic properties of the individual neurons.

Further, it is interesting to note the explosion in the number of granule cells as the cerebellum evolved in mammals. Our modelling suggests that, with the innovation of recurrence within the granular layer, the number of cells becomes the constraining factor in the generation of a temporally diverse set of basis signals.

We measured the performance of each of the granular layer models by testing how closely the generated set of basis signals could approximate an arbitrary target output signal. Below we discuss the implications of this, examine the assumptions implied by our modelling approach, and consider other possible measures of adaptive filter performance.
The adaptive filter model requires the existence of a teaching or error signal, which in the case of the cerebellum is conveyed by climbing fibres. However, in general an explicit signal representing a target output does not exist. For example, sensory processing nuclei receive a ‘teaching’ signal in the form of an afferent sensory signal, which consists of a superposition of predictable reafference and an unpredictable, behaviourally relevant stimulus. The target output of the adaptive filter is the negative image of only the reafferent component. In the cerebellum, the role of the teaching signal likely varies between cerebellar microzones, but is unlikely to encode an explicit target signal.

Nevertheless, regardless of external connectivity, the adaptive filter model implies that the learning rule will ultimately direct the adaptive filter to produce a stable and predictable output signal. The target signal referred to in our modelling represents this \textit{a priori} unknown signal, and it is a key assumption that the function of the granular layer is to produce a basis set sufficient to approximate it with good accuracy.

The measure of performance $\hat{e}$ represents how well the basis set can approximate a variety of smooth target signals. However, it should be noted that a large value for $\hat{e}$ does not necessarily predict that the network will perform poorly in a biologically realistic situtation — for example as illustrated in figure 4.6, a single layer feedforward network may result in a large value for $\hat{e}$, but is clearly capable of approximating target signals which are suitably short in duration, as may well be the case in a realistic EOD command situation.

By neglecting to include an explicit learning mechanism, we ignored various important adaptive filter performance criteria such as learning speed and robustness. A future study could extend our approach to including an explicit co-variance type learning rule, allowing these properties to be examined. A future study could, for example, determine whether the output of the granular layer is robust with the addition of noise.

It is known that recurrent inhibitory neural networks, such as echo state networks, are capable of producing diverse time-varying output [65], and number of studies have examined them in the context of the cerebellar granular layer [31, 38]. In particular, recent work has demonstrated how recurrent networks incorporating sufficiently strong recurrent inhibition are capable of generating a broad class of biologically useful filter functions [70]. In this paper, in addition to constructing the network at random, we took the somewhat novel approach of building a network such that the connectivity results in a prescribed set of output signals. This was made possible using the FORCE learning technique developed by [49].

It is satisfying that, even with a small number of neurons, it is possible to construct a
network whose performance is greater than a network constructed at random. Furthermore, by increasing the number of neurons, the capability of the network can be improved without limit. With a random network, performance is both much more variable and in general much poorer (see Figure 4.9). Future work will examine how imposing constraints on the connectivity of the network, informed by the known connectivity of the cerebellar granular layer, affects the performance of the network.

Our modelling has focussed on the temporal recoding required for well-studied cerebellar functions such as the eyeblink reflex [71] and vestibulo-ocular reflex [72], as well as reafference suppression in cerebellum-like sensory nuclei. We did not consider so-called ‘spatial recoding,’ where the output depends on specific combinations of signals. Evidence for such multimodal recoding in the cerebellum is not clear [5] although it has been detected in the ELL [73], and recent work has shown that coactivation of different sensory modalities results in increased granule cell firing rates and pathway-specific synaptic responses, termed a ‘biophysical signature’ of the input pathway [53].

4.5 Appendix

4.5.1 Neuron Model

We model each neuron as a leaky integrator whose activity is governed by a simple firing rate model [41]:

\[
I_j(t) = \sum_i w_{ij}r_i(t) + \sum_i w_{in}m_i(t),
\]

\[
\tau_j \frac{dx_j}{dt} = -x_j + I_j(t),
\]

\[
r_j(t) = \tanh(x_j),
\]

where \(x_j(t)\) and \(r_j(t)\) are the activity and firing rate of neuron \(j\) and \(m_i(t)\) is the activity of mossy fibre \(i\). Synaptic weights \(w_{ij}\) represent the strength of the connection between neuron \(i\) and \(j\) and \(w_{in}\) represents the strength of the connection between mossy fibre \(i\) and neuron \(j\). Each neuron has a time constant \(\tau_j\). The output of the granular layer is the parallel fibre activity \(r_j(t)\).\(^2\)

\(^2\)This model is different from the model presented in Chapter 3. Firstly, due to slightly different assumptions about the timescales of the pre- and postsynaptic currents, the order of the summation and activation function (tanh) is reversed (see [41]). Secondly, as discussed in Section 3.2, the activation
4.5 Appendix

4.5.2 Mossy Fibre Signals

Mossy fibres convey incoming signals $m_i(t)$ described by a Gaussian waveform:

$$m_i(t) = \exp \left( - \left( \frac{t - t_i}{\alpha} \right)^2 \right), \quad 0 < t < 1000 \text{ ms},$$

where $t_i$ represents the time of onset of the signal. Sparse mossy fibre input consists of a single waveform $m_1(t)$ with $t_1 = 100$. Rich mossy fibre input consists of five waveforms $m_0(t) \ldots m_4(t)$ where $t_i = 100 + 180i$.

4.5.3 Performance Measure

Target signals $f(t), 0 < t < 1000 \text{ ms}$, are sampled at random from a Gaussian process [74]. A Gaussian process is a distribution of functions characterised by a mean function $m(t)$ and covariance $k(t,t')$. We choose $m(t) = 0$ and the following covariance function:

$$k(t,t') = \exp \left( - \frac{1}{2} \left( \frac{t - t'}{\ell} \right)^2 \right),$$

which defines a distribution of smooth functions with characteristic timescale $\ell$. By discretising over a time interval $\Delta t = 1 \text{ ms}$, we define a normal distribution with mean 0 and covariance matrix $\Sigma$:

$$\Sigma_{ij} = \exp \left( - \frac{1}{2} \left( \frac{\Delta t(i - j)}{\ell} \right)^2 \right).$$

4.5.4 Output Signal

Given a target signal $f(t)$, the output of the adaptive filter is the weighted sum of parallel fibre signals $\sum w_i^{\text{out}} r_i(t)$, where the values of $w_i^{\text{out}}$ are calculated using regularised mean-function tanh make take on negative as well as positive values. Neither of these differences have a significant effect on the dynamics of the network.
least-squares regression. Then the error in the approximation is:

$$e = \int \left( \sum_i w_i^{\text{out}} r_i(t) - f(t) \right)^2 dt.$$  \hspace{1cm} (4.5.1)

The mean squared error ˆe is defined as the expected value of the least-mean-squared error e where f(t) is sampled at random from the Gaussian process (4.5.3). Defining A as the matrix whose columns are the output signals r_i(t), discretised over a time step ∆t:

$$\hat{e} = \text{tr} \left( M^T \Sigma M \right),$$

where $M = A \left( A^T A \right)^{-1} A - I.$

4.5.5 Trained Recurrent Network

The trained recurrent network was constructed using the ‘FORCE’ learning procedure of [49]. The method of construction is described briefly below, for more detail see [49].

Given prescribed mossy fibre signal m(t) and target signals $f_i(t), i = 1 \ldots n$, we construct connectivity matrices $W$ and $W^{\text{in}}$ and time constants $\tau_i$, defining a network of leaky integrator neurons, a subset of which generate $f_i(t)$ when the network is stimulated by $m(t).$

We first choose a number $N$, which will determine the number of neurons in the final network. Next, we build a network consisting of $N$ network neurons and $n$ output neurons,

---

3Given an n-neuron network, let $R : \mathbb{R}^n \rightarrow L^2([0,t_{\text{max}}])$ be the linear function which maps a vector of output weights to a linear combination of output signals: $R(w)(t) = \sum_{i=1}^{n} w_i r_i(t),$ where $w = (w_1, \ldots, w_n)^T$ and $t_{\text{max}} = 1000 \text{ms}$ is the duration of the output signals. The network outputs were calculated numerically by integrating the network equations over a timestep $\Delta t = 1 \text{ms},$ so $R$ may be approximated by a $1000 \times n$ matrix $A$ whose columns are the numerically computed approximations to $r_i(t).$ Given a target signal $f,$ we calculate the vector $\mathbf{f}$ by discretising over the same timestep $\Delta t.$ Then minimising Equation 4.5.1 is equivalent to projecting $\mathbf{f}$ onto the image of $A$ by solving the normal equations $A^T A w = A^T f.$ The least-mean-squares error $e$ may then be calculated: $e = \|A w - f\|^2 = \|M f\|^2$ where $M = A (A^T A)^{-1} A^T - I.$

If $\mathbf{f} \sim \mathcal{N}(0, \Sigma)$ is sampled from a Gaussian Process with covariance matrix $\Sigma$ (see Section 4.5.3), then $Mf \sim \mathcal{N}(0, M^T \Sigma M)$ and the expected least-mean-squares error is given by $\hat{e} = \mathbb{E}[\|Mf\|^2] = \text{var}(M^T f M) = \text{tr}(M^T \Sigma M).$

4The network output is said to generate $f_i(t)$ if there is a weighted sum of output signals which closely approximates $f_i(t).$
the connectivity being defined by matrices $W^{\text{in}}(N \times 1)$, $W^{\text{net}}(N \times N)$, $W^{\text{fb}}(N \times n)$ and $W^{\text{out}}(n \times N)$. The elements of $W^{\text{in}}, W^{\text{net}}$ and $W^{\text{fb}}$ are generated randomly. We determine $W^{\text{out}}$ by carrying out the FORCE learning procedure. After training, the $n$ output neurons generate the target signals $f_i(t)$ on stimulation of the network by mossy fibre signal $m(t)$.

Finally, we construct a $W$ and $W^{\text{in}}$ as follows:\textsuperscript{5}

$$W = \begin{bmatrix} W^{\text{net}} & W^{\text{fb}} \\ W^{\text{out}} & 0 \end{bmatrix}$$

$$W^{\text{in}} = \begin{bmatrix} W^{\text{in}} \\ 0 \end{bmatrix}$$

\textsuperscript{5}A FORCE network is composed of a single input neuron, a dynamic reservoir of $N$ neurons, and $n$ output neurons, the connectivity being determined by the four weight matrices $W^{\text{in}}, W^{\text{net}}, W^{\text{out}}$ and $W^{\text{fb}}$ (see Section 3.3.2). However, in our analysis we require a single network determined by just two matrices incorporating all four of the matrices generated by the FORCE learning algorithm. This section demonstrates how to build the two matrices $W(N + n \times N + n)$ and $W^{\text{in}}(N + n \times 1)$.

Since the trained network is only tested against the case of a single input, the input matrix $W^{\text{in}}$ consists of only one column.
Sequential Pattern Formation in the Cerebellar Granular Layer

Recurrent neural networks clearly have the capability to generate diverse signals, suitable as a basis for an adaptive filter. However, the cerebellar granular layer is not only recurrent, but has a distinctive pattern of connectivity which is not accounted for by the random recurrent networks described in the previous chapter. In this chapter, I develop a more detailed model of the granular layer which incorporates the granular layer’s specific Golgi-granule cell circuit.

This chapter consists of a draft of a research paper published in *The Cerebellum*. Minor clarifications requested by the thesis examiners have been annotated as footnotes to the text.
Abstract

Here we introduce a novel mechanism for temporal recoding by the cerebellar granular layer based on three key properties: the granule cell-Golgi cell inhibitory feedback loop, bursting behaviour of granule cells and the large ratio of granule cells to Golgi cells. We propose that mutual inhibition of granule cells, mediated by Golgi cell feedback inhibition, prevents simultaneous activation. Granule cells are differentiated by firing threshold, resulting in sequential bursts of spikes. We demonstrate the plausibility of the mechanism through a computational simulation of a firing rate model, and further examine its robustness by developing a spiking model incorporating realistic postsynaptic potentials.
5.1 Introduction

The cerebellar granular layer, consisting of a dense collection of excitatory granule cells and inhibitory interneurons, forms a key component of the cerebellum’s microcircuit. Purkinje cells, the only output neurons of the cerebellar cortex, receive signals via two distinct pathways: climbing fibres arising from cerebellar deep nuclei, and parallel fibres arising from the granular layer. The computational mechanism implemented by the granular layer is therefore of fundamental importance to a system level understanding of the cerebellum itself.

The idea that the function of the granular layer is to transform or recode a set of contextual input signals has been central to most theories of the cerebellum ever since early theories proposed by Marr and Albus. According to the Marr-Albus theory, the granular layer increases the information content of signals conveyed by parallel fibres [25, 26]. Purkinje cells, whose output encodes signals related to motor control, select from amongst the many thousands of parallel fibre signals, guided by an error signal conveyed by climbing fibres. Later work introduced the idea of an adaptive filter [27], which, through a simple anti-Hebbian learning rule at parallel fibre synapses, acts to decorrelate mossy fibre and climbing fibre inputs.

The adaptive filter theory, being supported by a range of experimental and computational evidence, remains the most widely accepted system-level theory of the cerebellum [4]. A key requirement of the adaptive filter is a temporally diverse set of parallel fibre signals from which the Purkinje cell constructs an output consisting of a weighted linear sum. Computational studies have demonstrated the influence of the set of basis functions on the performance of the adaptive filter [6].

Much of the evidence for the adaptive filter theory of cerebellar function has been drawn from classical conditioning experiments such as the eyeblink reflex [31] as well as studies of reafference suppression by cerebellum-like sensory processing nuclei in fish. The adaptive filter function of such cerebellum-like structures has been clearly demonstrated in sharks [57, 61] and weakly electric fish [35]. These experimental results demonstrate a core requirement of the granular layer recoding: to increase the temporal diversity of the signals conveyed by parallel fibres.

For example, in eyeblink experiments, the relationship between mossy fibre input and Purkinje cell output is straightforward. Mossy fibres deliver a signal representing the unconditioned stimulus, which after a fixed delay is followed by a pause in Purkinje cell
activity coinciding with the conditioned response. Any set of parallel fibre signals which encodes activity at a range of delays with respect to the mossy fibre input would be suitable as an adaptive filter basis [54].

As yet there is no consensus on how such a recoding might be implemented by the granular layer, although a number of mechanisms have been proposed. Mechanisms based on granule cells with varied response characteristics [67] suffer from the disparity between the timescales associated with cerebellar function (hundreds of milliseconds) and physiologically realistic neuron response times (perhaps tens of milliseconds). Delay line theories suggest that cells connected in series result in their sequential activation [36], but such connectivity has not been found in the granular layer.

Perhaps more promising are theories based on the idea of an ‘echo state network’ or ‘liquid state machine,’ in which recurrent inhibition within a randomly connected network results in a complex sequence of activity which outlasts the input signal [31, 38, 70]. Using neural network training algorithms, it is possible to build a network which generates biologically useful parallel fibre signals, such as a set of Gaussian waveforms [6].

While such theories do account for the granular layer’s recurrent architecture, they do not necessarily result in biologically realistic circuits. For example, echo state networks typically assume that neurons may make both excitatory and inhibitory connections. Furthermore, they do not account for other distinctive properties of the granular layer, such as the curiously high ratio of excitatory granule cells to inhibitory Golgi cells, and the tendency of granule cells to generate high frequency bursts of action potentials.

In this paper, we introduce and critically examine a novel mechanism for the formation of temporally diverse patterns of activity in the granular layer. The mechanism, based on a modified winner-take-all network, relies on mutual inhibition of granule cells, mediated by the GrC-GoC-GrC recurrent loop, to prevent simultaneous activation. Granule cells are differentiated by firing threshold, resulting in sequential activation of granule cells as each firing threshold is exceeded. Hysteresis-loop bursting causes granule cells to fire a uniform burst of spikes, truncated once input current drops sufficiently. A process of activity-dependent inhibition prevents cells from firing more than one burst. The proposed mechanism generates a sequence of bursts of parallel fibre activity, with each granule cell firing a burst of activity at a predictable delay with respect to a step function stimulus.
**Figure 5.1.** Schematic of the mechanism. A mossy fibre delivers a step input to a large number of granule cells. Granule cells, which have a range of firing thresholds, start firing once their input is sufficiently large. Inhibition via a GrC-Go-GrC recurrent loop prevents further granule cells from firing, until the granule cell input current falls below the ‘off’ firing threshold. A slow process, such as depletion of neurotransmitter, prevents granule cells from firing more than one burst.

### 5.2 Sequential Winner-Take-All Mechanism

Consider a simplified view of the granular layer consisting of the following (Fig. 5.1):

i. A large number of bistable granule cells. Cells change from silent to spiking once input current exceeds an ‘on’ threshold, and returns to silent when current falls below a lower ‘off’ threshold. The on threshold for granule cells varies across a range of values.

ii. A single mossy fibre, which branches to deliver an identical input signal to all of the granule cells.

iii. A single Golgi cell, which receives excitatory input from all granule cells, and in turn sends inhibitory output to all granule cells. The spike rate of the Golgi cell is proportional to input current.

iv. A process of activity-dependent inhibition that reduces the excitability of each granule cell after spiking activity.
Now consider the response of the system to mossy fibre input consisting of a ‘step’ waveform. That is, the mossy fibre is silent until $t = 0$ whereupon it starts spiking at a high rate. (This type of input signal is equivalent to the conditioned stimulus in classical conditioning experiments, where the stimulus is a continuous tone).

Each Mo-GrC synapse starts to inject current into its target granule cell, resulting in an increasing postsynaptic potential. The granule cells are assumed to have a variety of firing thresholds, so that a handful of cells with low thresholds start firing. Once these granule cells start firing, they begin stimulating the Golgi cell, which in turn inhibits the granule cells via recurrent pf-GrC connections. This inhibitory current prevents the remaining granule cells from firing, but due to bistability, the currently active granule cells remain active.

Once the inhibitory current reaches a level sufficient to bring the active granule cells below their ‘off’ threshold, they stop firing. This results in a reduction in Golgi cell activity and a consequent reduction in inhibition of the granule cells, leading to an increase in granule cell input current.

In the absence of any further mechanisms, this process would repeat, resulting in a subset of low-threshold granule cells generating a periodic sequence of bursts for the duration of the mossy fibre input. However, we postulate a process of activity-dependent inhibition which causes the excitability of granule cells to decrease during firing, effectively increasing their firing thresholds above the thresholds of the remaining cells. The result is a sequence of bursts of granule cells, each burst consisting of the lowest threshold granule cells which are yet to fire. The process continues until all cells have generated a burst of cells.

The constituent elements of this mechanism are derived from the known connectivity of the granular layer and based on realistic biophysics. The details of the relevant neurophysiology are outlined in the following sections.

### 5.2.1 The Granular Layer Microcircuit

The granular layer is composed primarily of a very large number of excitatory granule cells, a smaller number of inhibitory Golgi cells, and small masses of nerve fibre terminals termed glomeruli [32]. Other cells, such as Lugaro cells and unipolar brush cells, exist in smaller numbers. Ascending granule cell axons, the ‘T’ shaped branches of which are termed parallel fibres, form the primary output of the granular layer. Parallel fibre signals, considered to be basis signals for the adaptive filter, are transmitted to Purkinje
cells whose axons form the output of the cerebellar cortex.

Input to the granular layer is received from mossy fibres which form synapses onto granule cells and Golgi cells. Golgi cells make connections with granule cells, and receive input from mossy fibres as well as recurrent connections from parallel fibres. Golgi cells therefore form both feedforward and feedback connections onto granule cells. The connections between mossy fibres, granule cells and Golgi cells are made in glomeruli, which consist of mossy fibre axon terminals, granule cell dendrites and Golgi cell dendrites wrapped in an electrically insulating glial sheath.

Mossy fibres convey diverse signals representing sensory and motor information [53]. Each mossy fibre contacts several hundred granule cells, and each granule cell receives connections from dozens of mossy fibres [32]. Thus, there is significant divergence from mossy fibres to granule cells. Each Golgi cell receives connections from thousands of granule cells, but granule cells are inhibited by at most a handful of Golgi cells.

The sequential winner-take-all mechanism relies primarily on two distinctive properties of the granular layer microcircuit: recurrent inhibition from Golgi cells onto granule cells, and the large ratio of excitatory cells to inhibitory cells.

### 5.2.2 Hysteresis-Loop Bursting

Recent advances have allowed direct recordings of the activity of granule cells in vivo, revealing that granule cells fire at a high frequency (up to 1000 Hz) in discrete bursts of around 50 ms [75]. Burst firing, where periods of high frequency action potentials are followed by periods of quiescence, is a highly prevalent phenomenon in the nervous system.

Burst firing emerges naturally from many biophysical conductance-based models. Such models predict a sharp transition between repetitive firing and resting states as input current crosses a threshold dividing steady-state and limit-cycle behaviour. Under certain conditions, the parameter ranges of the two types of solution overlap, resulting in a region of bistability where both steady-state and limit-cycle solutions are stable. Movement around the resulting hysteresis loop results in a system which oscillates between resting and spiking states, a phenomenon termed hysteresis loop bursting [76].

The upper and lower limits of the region of bistability determine two firing thresholds corresponding to bifurcations in the underlying dynamical system. This can be modelled using a multivalued activation function, as illustrated in Fig. 5.4 (b). An increasing
input current results in the firing rate remaining at zero until it crosses the upper firing threshold, whereupon it transitions to a non-zero value. Conversely, a decreasing input current results in a non-zero firing rate until the input current crosses the lower firing threshold.

5.2.3 Activity-Dependent Inhibition

The phenomenon where neurons’ excitability depends on past activity is a prominent feature of the nervous system. The cerebellum itself provides one of the most well-studied examples: long term depression (LTD) and long term potentiation (LTP) at synapses between parallel fibres and Purkinje cells [77].

Activity-dependent inhibition occurs where elevated activity of a neuron leads to decreased sensitivity to subsequent stimulation. Such plasticity may be caused by an association between pre- and postsynaptic activity (spike-timing dependent plasticity) [78] or postsynaptic activity alone, and may be mediated by a synaptic [79] or intracellular mechanism (such as a calcium-activated potassium channel) [80].

To prevent granule cells from firing multiple bursts, we propose a mechanism of activity-dependent inhibition of granule cells, acting on short time scales of the order of tens or hundreds of milliseconds.

5.2.4 Parallel Fibre Signals

According to the adaptive filter theory, the function of the granular layer is to perform a spatio-temporal recoding of mossy fibre input, giving rise to a set of basis signals which are subsequently weighted and summed by Purkinje cells. A requirement of the theory is that the signals be sufficiently diverse to form any desired output signal by a linear sum.

The ‘delay’ eyeblink reflex paradigm, a type of associative conditioning, provides an instructive example. Mossy fibre input consists of a signal whose onset is triggered by the conditioned stimulus. The desired output is a motor command delayed with respect to the conditioned stimulus. Theoretical studies have demonstrated that a set of basis signals consisting of discrete pulses offset by varying delays with respect to the stimulus is sufficient for this scenario [54, 67]. Other studies have demonstrated how such signals can form a basis for a wide variety of output signals [6].
Figure 5.2. Firing rate model. Neuron $i$ receives synaptic connections from neurons delivering synaptic drives $z_j(t)$. These are weighted by synaptic strengths $w_j$ and combined to form synaptic current $I_i(t) = \sum_j w_j z_j(t)$. Current is converted to a spike rate $r_i(t) = F_i(I_i(t))$, where $F_i$ is a nonlinear activation function. Finally synaptic drive $z_i(t)$ is calculated according to the equation $\tau_i \frac{dz_i}{dt} = -z_i(t) + r_i(t - \Delta_i)$ where $\tau_i$ is the time constant associated largely with synaptic currents and $\Delta_i$ is the delay caused by axonal action potential transmission.

The mechanism outlined in this paper generates parallel fibre signals consisting of discrete bursts of activity at predictable delays with respect to mossy fibre input, much like the examples cited above.

5.3 Firing Rate Model

In order to demonstrate the plausibility of the sequential winner-take-all mechanism, we developed a simple simulation based on a spike rate model. Each neuron is characterised by two variables; $r_i(t)$ represents instantaneous mean firing rate, and $z_i(t)$ represents synaptic drive, which broadly corresponds to the current delivered to its postsynaptic neurons (Fig. 5.2).

Given presynaptic neurons delivering synaptic drives $z_j(t)$, the synaptic current $I_i(t)$, spike rate $r_i(t)$ and synaptic drive $z_i(t)$ of postsynaptic neuron $i$ are calculated as follows:

$$I_i(t) = \sum_j w_{ij} z_j(t),$$  \hspace{1cm} (5.3.1)

$$r_i(t) = F_i(I_i(t)), \hspace{1cm} (5.3.2)$$

$$\tau_i \frac{dz_i}{dt} = -z_i(t) + r_i(t - \Delta_i), \hspace{1cm} (5.3.3)$$

where $w_j$ represents the synaptic weight of the connection between neurons $j$ and $i$, $F_i$ is
Figure 5.3. Mathematical model of the sequential winner-take-all mechanism. A mossy fibre delivers a prescribed signal \( r_{mo}(t) \), resulting in synaptic drive \( z_{mo}(t) \), which stimulates a number of granule cells (for clarity, only one shown here). Granule cells spike at a rate given by \( r_{gr}(t) \), generating synaptic drive \( z_{gr}(t) \) incorporating axonal transmission delay \( \Delta_{gr} \). Golgi cells spike at a rate \( r_{go}(t) \) and in turn inhibit granule cells via synaptic drive \( z_{go}(t) \).
**Figure 5.4.**

- **a** Mossy fibre input signal

  ![Mossy fibre input signal diagram](image)

  - Spike rate
  - 1 kHz
  - $t = 0$
  - $r_{mo}(t)$
  - Time

- **b** Granule cell activation function

  ![Granule cell activation function diagram](image)

  - Spike rate
  - 1 kHz
  - $I_{off}$, $I_{on}^i$
  - Synaptic current
  - $F_{gr}$

- **c** Golgi cell activation function

  ![Golgi cell activation function diagram](image)

  - Spike rate
  - 1 kHz
  - Synaptic current
  - $F_{go}$

**Figure 5.4.** **a** The Mossy fibre input signal is a step function starting at $t = 0$. **b** The granule cell activation function $F_{gr}$ is bistable, resulting in bursting behaviour. Inactive cells start firing when current reaches threshold $I_{on}^i$; active cells stop firing when current drops below $I_{off}$. Values for $I_{on}^i$ vary between cells. **c** The Golgi cell activation function is a monotonic, saturating function $F_{go}(I) = \tanh(I)$. 
an activation function which translates input current into spike rate, $\tau_i$ is a time constant dominated by the decay of synaptic currents, and $\Delta_i$ is the time taken for action potentials to travel along the axon of the postsynaptic neuron. See Fig. 5.2 for an illustration of the model, which can be derived by assuming that discrete postsynaptic potentials may be approximated by a continuous mean spike rate [41].

The granular layer simulation consists of a single mossy fibre, a number of granule cells, and a single Golgi cell connected according to the network illustrated in Figure 5.3.

5.3.1 Mossy Fibre

A single mossy fibre spikes at a rate $r_{mo}(t)$ of 1 kHz beginning at $t = 0$ (see Figure 5.4 (a)), which generates mossy fibre synaptic drive $z_{mo}(t)$:

$$\tau_{mo} \frac{dz_{mo}(t)}{dt} = -z_{mo}(t) + r_{mo}(t),$$

(5.3.4)

where $\tau_{mo}$ is the time constant of Mo-GrC synapses.

5.3.2 Granule Cells

Granule cell $i$ receives synaptic current $I_{gr}^i(t)$ composed of an excitatory mossy fibre component and an inhibitory Golgi cell component:

$$I_{gr}^i(t) = w_{in}^i(t)z_{mo}(t) - w_{go}z_{go}(t).$$

(5.3.5)

Synaptic weight $w_{go}$ is set to a constant value, whereas $w_{in}^i(t)$ is a time-dependent variable whose values diminishes during spiking (see Section 5.3.4).

Granule cell input currents are translated to firing rates $r_{gr}^i(t)$ by activation functions $F_{gr}^i$:

$$r_{gr}^i(t) = F_{gr}^i \left( I_{gr}^i(t) \right).$$

(5.3.6)

$F_{gr}^i$ is characterised by thresholds $I_{off}$ and $I_{on}^i$ which determine the transition between spiking and non-spiking states (Fig.5.4 (b)). For values of input current below $I_{off}$, the spike rate is zero; for values above $I_{on}$, the spike rate is 1 kHz. For intermediate currents, the spike rate maintains its present value. On-thresholds $I_{on}^i$ are prescribed separately for each granule cell.
Synaptic drive $z_{gr}^{i}(t)$ is determined by the following equation:

$$\tau_{gr} \frac{d}{dt} z_{gr}^{i}(t) = -z_{gr}^{i}(t) + r_{gr}^{i}(t - \Delta_{gr}), \quad (5.3.7)$$

where $\tau_{gr}$ represents the GrC-GoC synaptic time constant.

$\Delta_{gr}$ represents the delay associated with action potential conduction along parallel fibres. Due to the compact nature of the cerebellar cortex, delays associated with action potential conduction are generally small. However, due to their narrow, unmyelinated axons, the conduction time along parallel fibres can be significant. We estimate that the time between generation of granule cell spike and its arrival at the GrC-GoC synapse is of the order of 5–10 ms. We use a value of 5 ms in our simulation. We assume that other axon conduction delays are insignificant.

### 5.3.3 Golgi Cell

The Golgi cell receives synaptic input from each granule cell, resulting in input current $I_{go}(t)$:

$$I_{go}(t) = \sum_{i} w_{gr} z_{gr}^{i}(t), \quad (5.3.8)$$

where $w_{gr}$ is the synaptic weight of GrC-GoC synapses.

Golgi cell input current is translated to firing rate $r_{go}(t)$ by activation function $F_{go}$:

$$r_{go}(t) = F_{go} (I_{go}(t)) = \tanh(I_{go}(t)). \quad (5.3.9)$$

$F_{go}$ is an increasing function which saturates at 1 kHz (Figure 5.4).

Synaptic drive is given by:

$$\tau_{go} \frac{d}{dt} z_{go}(t) = -z_{go}(t) + r_{go}(t), \quad (5.3.10)$$

where $\tau_{go}$ is the GoC-GrC synaptic time constant.
\[\tau_{\text{mo}} \quad \text{Mo-GrC time constant} \quad 10 \text{ ms}\]
\[\tau_{\text{gr}} \quad \text{GrC-GoC time constant} \quad 10 \text{ ms}\]
\[\tau_{\text{go}} \quad \text{GoC-GrC time constant} \quad 10 \text{ ms}\]
\[\Delta_{\text{gr}} \quad \text{granule cell axon conduction delay} \quad 5 \text{ ms}\]
\[\tau_{\text{slow}} \quad \text{synaptic depression time constant} \quad 125 \text{ ms}\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau_{\text{mo}})</td>
<td>10 ms</td>
</tr>
<tr>
<td>(\tau_{\text{gr}})</td>
<td>10 ms</td>
</tr>
<tr>
<td>(\tau_{\text{go}})</td>
<td>10 ms</td>
</tr>
<tr>
<td>(\Delta_{\text{gr}})</td>
<td>5 ms</td>
</tr>
<tr>
<td>(\tau_{\text{slow}})</td>
<td>125 ms</td>
</tr>
</tbody>
</table>

**Table 5.1.** Parameters for the firing rate model.

### 5.3.4 Activity-Dependent Inhibition

We model activity-dependent inhibition of granule cells by introducing a spike-rate dependent decrease in synaptic weight \(w_{\text{in}}^{i}(t)\):

\[\tau_{\text{slow}} \frac{d}{dt} w_{\text{in}}^{i}(t) = -w_{\text{in}}^{i}(t) r_{\text{gr}}^{i}(t), \quad (5.3.11)\]

When the granule cell is inactive, the value of \(w_{\text{in}}^{i}\) remains constant; during spiking, its value decays at a rate determined by the time constant \(\tau_{\text{slow}}\). Note that we are not making the assumption that the inhibition is caused by a synaptic mechanism; this formulation is merely a convenient means of making a durable reduction in the excitability of granule cells. Later we discuss possible mechanisms for this process.\(^{1}\)

### 5.3.5 Time Constants

The viability of the mechanism described above depends critically on the time delays related to the transmission of signals through the network. Typically delays are associated with membrane time constants, the temporal profile of postsynaptic currents, passive dendritic conduction and the transmission of action potentials along axons.

In our model (Figure 5.2), the first three are incorporated into the time constant \(\tau_{i}\) which is dominated by the decay of the postsynaptic current. A typical value for this parameter, as measured in electrophysiological experiments, is about 10 ms at Mo-GrC, GrC-GoC [81] and GoC-GrC synapses [82].

The values for synaptic time constants are shown in Table 5.1.

---

\(^{1}\)This activity-dependent inhibition mechanism only allows for a decrease in the synaptic weight \(w_{\text{in}}^{i}(t)\). Although not discussed here, a more realistic model would include a further mechanism, acting on an even slower timescale, which resets the synaptic weight to its initial value.
5.3.6 Synaptic Weights

All synaptic weights are set to a value of 1, except for the Mo-GrC synaptic weights $w_{in}^i$ which are set to a value that results in an initial spike time of $\Delta t = 10$ ms:

$$w_{in}^i(0) = \frac{1}{1 - e^{-\frac{\Delta t}{\tau_{mo}}}}, \quad (5.3.12)$$

where we assume that at least one granule cell has threshold $I_{on}^i = 1.2$.

5.4 Results

We constructed a firing rate model containing four granule cells with distinct firing thresholds $I_{on}^i = 1.00, 1.40, 1.50, 1.55$ and $I_{off} = 0.50$ (Figure 5.5). These values were carefully selected to ensure that only one cell fires during each cycle: the firing threshold for each granule cell is low enough such that the mossy fibre input current is sufficient to trigger firing, and on the other hand, high enough to prevent two granule cells firing simultaneously. This results in a sequence of granule cells bursts with each cell firing once. Granule cell bursts were of a duration of roughly 70 ms, with the onset of bursts separated by around 100 ms.

Next, we increased the number of granule cells, and set $I_{on}^i$ to the minimum value to avoid firing synchronously with the previous cell (Figure 5.6). This resulted in a sequence of thresholds increasing towards a maximum value given by

$$I_{max} = w_{in}^i(0) \approx 1.6, \quad (5.4.1)$$

which represents the maximum (steady-state) current delivered by the mossy fibre. The difference between consecutive thresholds decreases towards zero. In principle, arbitrarily many bursts could be generated in this manner, although there are physical limitations of this scenario, which we discuss later.

Figure 5.7 shows the results of the model where thresholds are generated at random. Thresholds were sampled from a uniform distribution between 1.0 and $I_{max}$. In this scenario, granule cells burst concurrently in groups, where each group contains granule cell with similar values for $I_{on}^i$. The number of cells in each group diminishes with time.

---

$^2$The initial spike time is defined as the time that the first granule cell exceeds its firing threshold, calculated by solving Equation 5.3.4. $w_{in}^i(0)$ is then calculated by substituting $z_{mo}(t) = I_{on}^i / w_{in}^i(0)$. 

Figure 5.5. A firing rate model containing four granule cells with firing thresholds $I_{on} = 1.00, 1.40, 1.50$ and $1.55$ in response to a constant mossy fibre input beginning at $t = 0$. The synaptic current received by each granule cell (black lines) increases until the firing threshold (dotted black lines) of the first granule cell $I_{on}$ is reached, whereupon it starts firing at a constant rate (grey bars, top). After a delay of $\Delta_{gr} = 5$ ms (due to axonal conduction time), the Golgi cell starts firing at an increasing rate (solid red line) and inhibitory feedback causes the synaptic current received by the granule cells to decrease. Activity-dependent inhibition causes the input weight $w_{in}^i$ of the active granule cell to decrease resulting in a more significant drop in synaptic input current compared to the inactive granule cells. Once its input current drops below $I_{off}$, it stops firing. The Golgi cell firing rate starts decreasing and granule cell activity increases until the second granule cell’s firing threshold is reached. This process repeats until all four granule cells have generated a burst of spikes.
Figure 5.6. A firing rate model containing 10 granule cells. The firing threshold $I_{\text{on}}$ of each cell is set to the minimum value to avoid firing concurrently with the previous cell. Arbitrarily many granule cell bursts may be generated in this manner, although the difference between consecutive spiking thresholds becomes vanishingly small.

Figure 5.7. A firing rate model containing 20 granule cells. The firing thresholds $I_{\text{on}}^i$ are random values between 1.0 and $I_{\text{max}} \approx 1.6$. The granule cells fire in groups, with the number of cells in each group decreasing with time.
Figure 5.8. A firing rate model containing 20 granule cells. The firing thresholds $I_{on}^i$ are set uniformly to 1.0, while initial synaptic weights are random values between $1/I_{max} \approx 0.6$ and 1.0.

As an alternative to varying the granule cells’ firing thresholds, it is possible to keep the firing thresholds constant, and vary the initial values of the synaptic weights, $w_{in}^i(0)$. For all granule cells, we set $I_{on}^i = 1$. Values for $w_{in}^i(0)$ were sampled at random between $1/I_{max} \approx 0.6$ and 1.0.

Figure 5.8 shows the results of the simulation. The results are very similar to Figure 5.7, with cells with similar values for $w_{in}^i(0)$ bursting concurrently.

5.4.1 Adaptive Filter Basis

A limitation of the parallel fibre signals generated by the sequential winner-take-all mechanism is their coarse granularity. For example, in Figure 5.5, only four bursts of activity are generated across a duration of 500 ms; an adaptive filter using these four signals as a basis set would only be able to assemble an output signal whose change in value coincides with the onset of one of the bursts. In the case of the eyeblink reflex, a conditioned response would be possible at $t = 10, 100, 200$ and 300 ms but not at intermediate times.

It is possible to decrease the granularity of the set of basis signals by decreasing the value of the model time constants, resulting in briefer and more frequent bursts. However, there
**Figure 5.9.** Three sequential winner-take-all networks combined in parallel. Each network receives the same mossy fibre input, but has different time constants (top). Medium network: $\tau_{mo}, \tau_{gr}, \tau_{go} = 10$ ms; $\tau_{slow} = 125$ ms; $\Delta_{gr} = 5$ ms. Fast network: $\tau_{mo}, \tau_{gr}, \tau_{go} = 5$ ms; $\tau_{slow} = 62.5$ ms; $\Delta_{gr} = 2.5$ ms; Slow network: $\tau_{mo}, \tau_{gr}, \tau_{go} = 20$ ms; $\tau_{slow} = 250$ ms; $\Delta_{gr} = 10$ ms. By combining three networks in parallel, the temporal resolution of the output signals is improved while maintaining the total timespan (bottom).
Figure 5.10. A presynaptic spike generates a postsynaptic current consisting of the sum of two exponentials with rise time $\tau_r$ and decay time $\tau_d$.

is a trade-off between granularity of the basis set and the total timespan: decreasing the time constants improves the temporal resolution, at the expense of the length of time over which bursts are generated.

This problem may be solved by constructing several networks, each with different values for the time constants, and combining them in parallel. Each network receives an identical mossy fibre signal, with parallel fibre output being combined at the Purkinje cells. In Figure 5.9, the parallel fibre signals generated by three distinct networks are shown. The ‘fast’ and ‘slow’ networks have time constants twice and half the values of the ‘medium’ network respectively. As before, thresholds $I_{on}$ were generated at random. The resulting network generates bursts of output with a good temporal resolution for the full timespan (Figure 5.9, bottom).

5.4.2 Spiking Model

The firing rate model underlying these simulations (Equations 5.3.1-5.3.3) relies on the assumption that the incoming spike train received at presynaptic terminals translates smoothly to a synaptic current. In this section, we examine this assumption by constructing a spiking model equivalent to the firing rate model described in the previous section. We replace the continuous equation for synaptic drive (5.3.3) with a process which generates a sequence of discrete spikes $t_j$ at a the frequency given by the rate variable $r_i(t)$.³

Each presynaptic spike generates a postsynaptic current described by a waveform consisting of the sum of two exponentials with a rise time $\tau_r$ and decay time $\tau_d$ (Figure 5.10). A postsynaptic current corresponding to a spike arriving at the presynaptic terminal at

³See Appendix 5.6.1.
Figure 5.11. Spiking model of the sequential winner-take-all network. Despite an increased variation in the input currents, the mechanism is still capable of generating discrete granule cell bursts, in a manner similar to the firing rate model.

\[ t = 0 \text{ is given by:} \]
\[
I_{\text{PSP}}(t) = \frac{T_{\text{ISI}}}{\tau_d - \tau_r} \left( e^{-t/\tau_d} - e^{-t/\tau_r} \right),
\]

\hspace{1cm} (5.4.2)

where \( T_{\text{ISI}} \) is the minimum interspike interval (i.e. the interval between spikes when \( r = 1 \)) and the normalisation factor \( T_{\text{ISI}} / (\tau_d - \tau_r) \) results in an average current independent of time constants and minimum interspike interval [83]. Postsynaptic current \( I(t) \) is then calculated by summing PSPs:

\[ I(t) = \sum_{t_j < t} I_{\text{PSP}}(t - t_j), \]

\hspace{1cm} (5.4.3)

where \( t_j \) are the times of spikes received from all presynaptic neurons. In the limit where \( T_{\text{ISI}} \rightarrow 0 \), the model reverts to the continuous firing rate model.

Figure 5.11 shows the results of a spiking model equivalent to the firing rate model shown in Figure 5.7. The minimum interspike interval \( T_{\text{ISI}} = 4 \text{ ms} \) for mossy fibres, granule cells and Golgi cells, equivalent to a maximum firing rate of 250 Hz. The time constants are set to \( \tau_d = 10 \text{ ms} \) and \( \tau_r = 1 \text{ ms} \). The firing thresholds \( I_{\text{on}} \) were sampled randomly between 1.0 and \( I_{\text{max}} \).
Despite an increased variation in the input currents, the mechanism is still capable of generating discrete granule cell bursts, in a manner similar to the firing rate model.

### 5.5 Discussion and Conclusions

In this paper, we described a biologically plausible mechanism for the generation of predictable sequences of parallel fibre activity, suitable as a basis for an adaptive filter. The mechanism incorporated three characteristic properties of the granular layer: the large number of granule cells compared to Golgi cells, the bursting behaviour of granule cells and the inhibitory feedback connections from Golgi cells onto granule cells. A computational simulation, based on a model derived from genuine biophysical principles, demonstrated the efficacy of the mechanism.

Using realistic values for parameters such as synaptic time constants, the model generated outputs consisting of short bursts of activity delayed with respect to the onset of the mossy fibre input signal. Burst firing is a common feature of neurons in general, and recent \textit{in vivo} recordings in rabbits showed that, in response to vestibular stimulation, granule cells generate distinct high frequency bursts of action potentials [84].

According to the simulation, Golgi cell activity oscillates in antiphase with respect to the granule cell bursts. In other words, Golgi cell spike rates increase during periods of granule cell inactivity, and decrease between granule cell bursts. In principle, this is a testable prediction, although present techniques are unable to reliably record from Golgi and granule cells simultaneously.

The model relies on a process of activity-dependent inhibition, operating at timescales of tens or hundreds of milliseconds, which reduces the excitability of granule cells after firing. It is not known whether such a phenomenon exists in the granular layer, although mechanisms which regulate neuronal excitability are widespread throughout the nervous system. Intracellular mechanisms, such as a calcium-activated potassium current [85, 86], are plausible, as are mechanisms based on synaptic depression. However the details of the mechanism are not important to the implementation of the model, which merely requires that activity causes granule cells to become less excitable.

Bistability, in turn leading to hysteresis-loop bursting, underlies the ability of each granule cell to generate a sustained burst of spikes. Without such a mechanism, burst duration would be limited to the time course of the GrC-GoC feedback loop, consisting of at most
one or two spikes. While this behaviour could plausibly generate useful parallel fibre signals, and has been predicted by detailed modelling studies [87], it does not conform with known granule cell recordings, which have demonstrated bursts lasting roughly 50 ms. A number of biophysical mechanisms have been proposed for the bursting behaviour of granule cells, including intracellular mechanisms based on slow K$^+$ channels [88].

In addition to lateral inhibition, most winner-take-all networks incorporate a mechanism of self-excitation, a property which underlies the ability of such networks to sustain stable states of activity. In contrast, the sequential winner-take-all network relies on two mechanisms in order to sustain network activity by two properties: bistability of granule cells, and the prolonged nature of the input signal. The prolonged mossy fibre input signal used in our simulations (Figure 5.4 (a)) is necessary in order to provide excitation to the granule cells for the duration of their output. This type of input signal is suggestive of the conditioned stimulus commonly used in ‘delay’ classical conditioning experiments, which often consists of a continuous tone. However, in general, we would not expect mossy fibres to deliver such a sustained stimulus to the granule cells. In such cases, another source of excitation would be required, perhaps intrinsic to the granule cells themselves, and possibly balanced by tonic inhibition from Golgi cells [89].

The model developed in this paper is entirely deterministic, and does not incorporate any noise or stochasticity in spike trains. A commonly suggested solution to the problem of noise within neuronal circuits is the idea of population coding. By encoding a signal as the mean activity of a large number of neurons, the signal variance is greatly reduced. The large number of granule cells in the cerebellum, as well as the large number of mossy fibre inputs, does indicate that some form of population coding may be taking place.

The 300 ms over which our model is capable of reliably generating granule cell bursts is in accordance with the timescales with which the cerebellum is typically associated. For example, the duration of signals associated with complex motor commands and typical classical conditioning experiments are of the order of half a second. However, the adaptive filter model of the cerebellum also requires that the set of basis filters is sufficiently varied so that Purkinje cells may synthesise all required output signals. For example, in ‘delay’ classical conditioning experiments, the output signal is a motor command which triggers the eyeblink reflex. A basis consisting of a small number discrete bursts (such as Figure 5.5) would only be capable of generating an motor command which coincided with the onset of one of the bursts.

However, the time of onset of each of the bursts is dependent on various parameters, such as the time constant of synaptic decay. By connecting in parallel a number of Golgi/-
granule cell complexes (Figure 5.1), each with slightly different parameters, a continuum of granule cell burst times could be achieved, as in Figure 5.9.

Echo state networks, consisting of excitatory and inhibitory cells connected at random, are known to be capable of generating complex sequences of activity [45], and it has been suggested that such networks are suitable as generators of parallel fibre signals in models of the cerebellum [31]. Previous work has identified the importance of the connectivity of the network [6], and identifying the optimal structure of such networks is an important matter for future theoretical investigation. However, such networks are not constrained by the known connectivity of the cerebellar granular layer, and do not in general result in plausible patterns of connectivity.

Winner-take-all networks, in which neurons compete for activation, are a class of recurrent inhibitory network which have been shown to be effective at a range of tasks including pattern recognition, decision making and signal processing [90]. Such neural networks, consisting of a group of excitatory cells along with a process of lateral inhibition, have also been shown to generate complex dynamics, including oscillations, short term memory [91] and the generation of patterns. The mechanism presented here implements a modified form of winner-take-all, in which the level of inhibition generated by the ‘winner’ diminishes with time. It is this feature which enables the network to generate sequential patterns of activity, as each cell takes its turn to outcompete its neighbours.

The mechanism of sequential winner-take-all presented in this paper, relies on intrinsic neuronal properties (granule cell bistability) in addition to recurrent inhibition. This contrasts with traditional echo state networks, which typically consist of simple linear neurons, or neurons with sigmoidal activation functions. We suggest that an interesting avenue for future research would be to investigate the possible improvements to such networks by the introduction by neurons with more sophisticated properties.

5.6 Appendix

5.6.1 Spiking Model

Equation 5.3.3 represents the synaptic drive generated by a continuously varying firing rate \( r_i(t) \). Here we derive an equivalent formulation based on a sequence of discrete spikes.

First, note that Equation 5.3.3 can be restated in integral form:
\[ z_i(t) = \int_0^t \frac{1}{\tau_i} \exp \left( \frac{s-t}{\tau_i} \right) r_i(s - \Delta_i) ds. \] (5.6.1)

Now, the spike rate \( r_i(t) \) represents the instantaneous average of a sequence of discrete spikes, which can be represented as follows:

\[ z_i(t) \approx \int_0^t \frac{1}{\tau_i} \exp \left( \frac{s-t}{\tau_i} \right) \left( \sum_{t_j<t} T_{ISI} \delta(s - t_j) \right) ds, \] (5.6.2)

where \( \delta \) is the delta function and the sequence of spikes \( t_j \) is calculated as follows:

\[ t_{j+1} = t_j + \frac{T_{ISI}}{r(t_j - \Delta_i)}, \] (5.6.3)

where \( T_{ISI} \) is a small arbitrary constant representing the interspike interval at \( r = 1 \).

Rearranging, we arrive at the following equation for \( z_i(t) \):

\[ z_i(t) = \sum_{t_j<t} \frac{T_{ISI}}{\tau_i} \exp \left( \frac{-(t - t_j)}{\tau_i} \right). \] (5.6.4)

Setting \( \tau_i = \tau_d \) and for \( \tau_r \ll \tau_d \), we define a postsynaptic potential:

\[ I_{PSP}(t) = \frac{T_{ISI}}{\tau_d - \tau_r} (e^{-t/\tau_d} - e^{-t/\tau_r}), \] (5.6.5)

and so:

\[ z_i(t) = \sum_{t_j<t} I_{PSP}(t - t_j). \] (5.6.6)
Connecting the structure of neuronal circuits to their behavioural and computational function is a major goal of neuroscience. The cerebellum in particular, due to its unique architecture and its independence from the rest of the brain, has stimulated much investigation of both a theoretical and experimental nature. The aim of this thesis was to discover connections between the structure of one component of the cerebellum — the granular layer — and its computational function.

Two major studies were presented. Firstly, a high level study which compared network models of the granular layer using a measure of their computational efficacy. Secondly, a more detailed model which proposed a specific mechanism based on the granular layer microcircuit. While the nature of the modelling varies between them, they both support the major conclusion that the connectivity of the granular layer gives it the capability to function as a temporal recoding device. Furthermore, the specific connectivity peculiar to the granular layer in the cerebellum, as well as cerebellum-like structures in fish, suggests an important role for recurrent feedback in generating temporally diverse output signals.

Areas of scientific research are usually characterised by one or more broad overarching theories which unify the field. On the one hand, such scientific paradigms serve as yard-
sticks against which new results may be measured; on the other, they serve as scaffolds on which to hang further theory. If there is a cerebellar paradigm, then the adaptive filter probably has the best claim, fulfilling Thomas Kuhn’s criteria of simplicity and broad scope [92]. While the adaptive filter theory is not universally accepted, few researchers reject it outright, and there are no other theories which provide a universal system-level description of cerebellar function based on its microcircuit (but see [5] for a discussion of alternative views). That the cerebellum operates as an adaptive filter is a fundamental assumption to the investigation undertaken in this thesis.

Given that the cerebellum is an adaptive filter, I make the further assumption that the granular layer functions as a recoding device which increases the temporal diversity of parallel fibre signals. This assumption is by no means universally accepted, and is mainly justified by the observation that the cerebellum is capable of generating accurately timed output signals, as well as the huge number of outputs from the granular layer.

Mathematical modelling necessarily requires making a number of simplifying assumptions. In this thesis, populations of spiking neurons are represented by networks of units whose activities are measured by a continuous variable representing a mean firing rate, and in Chapter 3 I presented a derivation which justifies these types of models. In general, because the aim was to discover the influence of network-level properties of the microcircuit, I ignored lower level details of neuronal function.

Due to the crudeness of the models, it can be challenging to interpret their output in terms of direct, falsifiable claims. In fact, the utility of this type of modelling may lie less in measurable predictions, and more in their ability to unambiguously capture hypotheses less easily described in words. The process of constructing these models serves as a tool to gain understanding of the scope and limitations of such hypotheses, and as a language to communicate them to others.

A central theme of this thesis has been the versatility of recurrent neural networks. It has long been known that feedforward networks can approximate any given function [93]; more recently it has been demonstrated that recurrent neural networks are universal approximators for dynamical systems [42]. There is no doubt that recurrent networks are of fundamental importance to the nervous system. However, implementing networks which reproduce biological behaviour using biologically realistic circuits remains a huge challenge. Particular difficulties include determining biologically plausible learning mechanisms which solve the ‘credit assignment problem’, and finding means to constrain the inherent instability brought about by output feedback.
Although the recurrent network modelling in this thesis was presented in the context of the cerebellum, it is sufficiently generic that it may have application to other parts of the nervous system. Recurrence is a ubiquitous feature of the brain, and it is possible that similar mechanisms may be in operation elsewhere, for example in the cortex or hippocampus. Temporal pattern generation may well be a computation necessary for a variety of cognitive functions such as memory or higher sensory processing, and it is a matter for future research to investigate brain regions where such processing may be taking place.

Future progress in determining the algorithm implemented by the cerebellar granular layer is likely to be led by improving experimental techniques. Methods which enable simultaneous recording from multiple neurons, and methods such as optogenetics which enable manipulation of individual neurons, show great promise in shedding light on circuit level neuronal mechanisms. However, without a theoretical framework in which new experimental results may be framed, their contribution to the understanding of the cerebellar algorithm will be limited.


