UNIVERSITY OF SOUTHAMPTON

Contributions of synaptic filters to models of synaptically stored memory.

by

Konstantinos Lagogiannis

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Physical Sciences and Engineering
School of Electronics and Computer Science

October 1,2013

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Konstantinos Lagogiannis

The question of how neural systems encode memories in one-shot without immediately disrupting previously stored information has puzzled theoretical neuroscientists for years and it is the central topic of this thesis. Previous attempts on this topic, have proposed that synapses probabilistically update in response to plasticity inducing stimuli to effectively delay the degradation of old memories in the face of ongoing memory storage. Indeed, experiments have shown that synapses do not immediately respond to plasticity-inducing stimuli, since these must be presented many times before synaptic plasticity is expressed. Such a delay could be due to the stochastic nature of synaptic plasticity or perhaps because induction signals are integrated before overt strength changes occur. The later approach has been previously applied to control fluctuations in neural development by low-pass filtering induction signals before plasticity is expressed. In this thesis we consider memory dynamics in a mathematical model with synapses that integrate plasticity induction signals to a threshold before expressing plasticity. We report novel recall dynamics and considerable improvements in memory lifetimes against a prominent model of synaptically stored memory. With integrating synapses the memory trace initially rises before reaching a maximum and then falls. The memory signal dissociates into separate oblivescence and reminiscence components, with reminiscence initially dominating recall. Furthermore, we find that integrating synapses possess natural timescales that can be used to consider the transition to late-phase plasticity under spaced repetition patterns known to lead to optimal storage conditions. We find that threshold crossing statistics differentiate between massed and spaced memory repetition patterns. However, isolated integrative synapses obtain an insufficient statistical sample to detect the stimulation pattern within a few memory repetitions. We extend the model to consider the cooperation of well-known intracellular signalling pathways in detecting storage conditions by utilizing the profile of postsynaptic depolarization. We find that neuron wide signalling and local synaptic signals can be combined to detect optimal storage conditions that lead to stable forms of plasticity in a synapse specific manner. These models can be further extended to consider heterosynaptic and neuromodulatory interactions for late-phase plasticity.

Declaration of authorship

I, Konstantinos Lagogiannis declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Contributions of synaptic filters to models of synaptically stored memory.

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work have been published as:

  – T. Elliott and K. Lagogiannis. The Rise and Fall of Memory in a Model of Synaptic Integration. *Neural Computation*,24,(10)2604–2654,2012.

Signed:.....
Date:.....

Ithaka

As you set out for Ithaka,
hope your road is a long one,
full of adventure, full of discovery.

Laistrygonians, Cyclops,
angry Poseidon-don't be afraid of them:
you' ll never find things like that on your way
as long as you keep your thoughts raised high,
as long as a rare excitement
stirs your spirit and your body.

Laistrygonians, Cyclops,
wild Poseidon-you won't encounter them
unless you bring them along inside your soul,
unless your soul sets them up in front of you.
Hope your road is a long one.

May there be many summer mornings when,
with what pleasure, what joy,
you enter harbours you're seeing for the first time:
may you stop at Phoenician trading stations
to buy fine things,

mother of pearl and coral, amber and ebony,
sensual perfume of every kind-
as many sensual perfumes as you can;
and may you visit many Egyptian cities
to learn and go on learning from their scholars.

Keep Ithaka always in your mind.
Arriving there is what you're destined for.
But don't hurry the journey at all.
Better if it lasts for years,
so you're old by the time you reach the island,

wealthy with all you've gained on the way,
not expecting Ithaka to make you rich.
Ithaka gave you the marvellous journey.
Without her you wouldn't have set out.
She has nothing left to give you now.

And if you find her poor, Ithaka won't have fooled you.
Wise as you will have become, so full of experience,
you'll have understood by then what these Ithakas mean.

(Cavafy, 1911) .

Acknowledgements

It has been quite a journey indeed for which I owe my gratitude to my supervisors Terry Elliott and Nigel Shadbolt. My interest in neuroscience was sparked by Terry's lectures and grew to a pursuit of a PhD. He suggested working on synaptically stored memory and I wish to thank him for the mentoring I received at multiple levels; from wider-ranging subjects of science to the technical level on the analysis of memory lifetime as a stochastic process. I want to thank Nigel, who made this PhD possible and gave me invaluable support, for which I am grateful. I will always hold fond memories from our day out sailing, he has shown to be a skipper with a steady hand you can trust under any weather conditions. Further, I wish to acknowledge the valuable feedback I received from Mark van Rossum.

Lastly, I want to thank my friends who made Southampton feel home and those who offered kind support with proofreading. I have dedicated this thesis to Nikoletta, who has been there for me during the difficult but also during the fun times in so many ways that I could not possibly mention here.

Contents

List of Figures

List of Tables

Introduction

The brain contains an immense structured neuron cell connectivity network. This network mediates high-level cognitive functions and it is where nature keeps one of its well guarded secrets, the mechanisms that mediate memory. Memory can be studied at multiple levels from psychology to molecular biology but also by theoretical computational models.

At the cellular level, it is known that neurons emit short electrical pulses that are transmitted to other neurons through connection points called synapses. The prominent view is that memories are formed by activity induced adaptations of the synaptic efficacy to transmit electrical pulses (see Eichenbaum and Cohen, 2004). These adaptations are collectively called *synaptic plasticity*. Molecular biology has made remarkable technological advances that have probed the role of molecular pathways of synaptic plasticity with some success (see Kandel, 2009). Cognitive psychology has mapped multiple expressions of memory and the various associated brain structures. Through theoretical studies abstract neural network models have been developed that exhibit memory properties relying on the strength state of the network's synapses. Undoubtedly, there has been significant progress in all these fields over the last century, but we have yet to obtain an understanding of memory such that we could traverse from memory phenomena to neural networks down to their biological substrates.

This thesis focuses on a theoretical study of models of synaptically stored memories with a focus on how individual synapses decide when to encode new information and thus express a change in their strength state. This is a critical problem for models of online memory processes that are required to continuously encode a stream of everyday facts and events; either all information is abruptly lost (French, 1999) in some models or in other models, changing the state of a synapse could mean that some previous memory is altered to encode a new one. Consequently it is a reflection of the balance between forgetting to make room for new learned material and retention, this balance is more commonly known as the *stability versus plasticity dilemma* (Grossberg, 1980, Abraham and Robins, 2005). A common approach to adjust this balance in synaptic memory models has been to stochastically select a fraction of synapses that will encode new memories. Each synapse would respond to a new incoming stimulus with some probability, and hence the rate of forgetting old information is bound with the rate of

new learning. Revisions to this approach have introduced progressive synaptic stabilization (Fusi et al., 2005) so the rate at which a synapse encodes a new memory is not necessarily the same as the rate this memory will be forgotten under the encoding of some future memory. Nevertheless, these models delay the ultimate demise of memories by prolonging the decay of the memory trace through synaptic stability while at the same time affecting the ability of a memory to encode new memories in one-shot.

The spearhead of this thesis will be to show that the stability versus plasticity dilemma can be addressed by a synapse model that filters stimuli before expressing plasticity. These synaptic filter models are not bound to the same assumptions on the progressive stabilization of plasticity and allow an expansion to consider how repetitive encoding of a particular memory may lead its long-term storage.

To support these ideas, this thesis begins by providing a foundation on memory related research across contexts in three separate chapters. Chapter 1 focuses on our current understanding of the taxonomy of memory systems in the brain and discusses the role and dynamics of oblivescence. Chapter 2 reviews the biological nature of learning by examining synaptic plasticity across three prominent animal-models of learning. In Chapter 3 we introduce the main mathematical methods we will be using in our analyses of synapse models for memory. Subsequently, in Chapter 4 we review typical formal models of memory and the development of ideas that led to the framework of synaptically stored memories we adopt in this thesis. Since we will propose a model for the transition to stable long-term plasticity we have also included a short review of the relevant plasticity models. In Chapter 5 we proceed to review a prominent model of synaptically stored memory and present a method to analyse the dynamics of this model's memory trace. Chapter 6 is dedicated to our proposal of utilizing filter synapse models for ongoing memory and to an analysis of their memory dynamics within a standard framework of synaptically stored memory. The next Chapter compares stochastically updating synapses in cascades against filter models on signal dynamics and memory capacity. In the Chapters that follow we extend our investigation of synaptically stored memory to consider stable forms of plasticity with filter synapses. In Chapter 8 we examine if individual filter synapses can detect optimal stimulation patterns known to lead to stable phases of plasticity and long-term memory. However, these are not easily detected within a few memory repetitions. For this reason the next Chapter examines how the the profile of postsynaptic depolarization could be utilized to assist in the detection of optimal storage conditions by individual synapses. Finally, in Chapter 10 we review and discuss results and issues surrounding our proposal of filter synapses for synaptically stored memory and suggest extensions and experimental tests.

Chapter 1

Memory phenomena

One of the most challenging problems in neuroscience is to identify the mechanisms underlying memory. The essence of memory and recall has troubled brilliant minds throughout centuries and it can be traced back to Aristotle's essay on "Memory and reminiscence" in 350 B.C. Although a great deal of progress has been made in the past few decades at the cellular level, the understanding of memory processes remains elusive. This chapter will focus on a current perspective on the organization of memory. The important principle in this organization is that there are multiple parallel memory systems. This perspective will be useful when we come to examine models of memory in later sections. A short background covering early accounts of memory function will be presented and then how the view of multiple memory systems in the brain emerged. A taxonomy of these memory systems according to the theory of declarative memory then follows. The systems are defined in terms of both function and the associated brain systems. Finally, we will shortly review research on forgetting and discuss issues surrounding its form and potential function.

1.1 A short history of connectionism

The study of mental processes was a subject of philosophy up until the emergence of experimental psychology at the beginning of the 19th century. By the turn of the 20th century rigorous methods were applied to the study of behaviour, giving rise to *behaviourism*.

The school of behaviourism believed that behaviour should be examined like phenomena in natural sciences using precise experiments and that learning can be fully understood by observation of behaviour alone. Behaviourist had the ambition that every behaviour can be reduced to a response for a given stimulus and memory is an elaboration of *stimulus and response* (S-R) pairs. Successful rigorous examinations were shown such as

the famous experiments conducted by Pavlov (1927) on the conditioned reflex response. Pavlov showed that an arbitrary stimulus which does not cause any behavioural response could be associated with a reflex response through learning, effectively associating a stimulus to a response.

S-R learning was believed to be the nature of habit and the basis of memory, collectively known as the S-R psychology. Tolman (1949) resisted the S-R account of memory. He believed that animals such as rats are able to form "cognitive maps" and make inferential judgments that guide their behaviour. For Tolman, memory was not a collection of habits formed by multiple S-R connections, an account which he likened to a *telephone switchboard*. Tolman devised experiments using mazes showing that rats obtain knowledge about the environment. Previous exposure to a maze without a particular goal would facilitate problem solving on a subsequent goal orientated task (Tolman, 1949). He showed that animals do not just reinforce connections linked to rewards, but acquire knowledge about the environment in a "cognitive map" and use it to solve problems. A further challenge to S-R learning came later in experiments which attempted to disconnect stimulus sensory areas of the cortex to motor areas. Lashley (1950) attempted to localize a learned behaviour by systematically cutting through the cortex of rats that had been trained to find an exit to a maze. After an extensive search, he failed to localize the learned behaviour and he is famously known to have stated that on the face of his data one could not persist on the S-R account of learning as it appears that memory is diffused in the brain. The S-R connection theory and the diffused memory appeared to contradict until Hebb (1949) managed to join them under the theory of connectionism. He postulated that connections between two neurons A and B are strengthened if neuron A persistently takes part in firing neuron B. This activity dependent modification would form assemblies of neurons representing percepts that are distributed over large areas of the cortex (Hebb, 1949). This insight did not answer the question of whether the engram of each memory is distributed throughout the brain. The supporting theory was that memories are encoded like a hologram, a paradigm that emerged from the invention of holographic microscopy (Gabor, 1949). Thus the holographic hypothesis was that memories are distributed throughout the brain and a small part of the memory engram can be used to reconstruct the whole memory (Pribram, 1969). However, the holography theory became obsolete after evidence, which we go through in the following sections, towards the existence of multiple memory systems specializing in particular mental faculties combined with the finding of a memory area dissociated from perceptual representation acting as a rather general memory store.

1.2 Multiple memory systems

1.2.1 A short history

The idea of multiple memory systems can be traced at least as far Gall (1835) who believed that each mental faculty has its own memory and it is a fundamental aspect of the distinct processing function of cortical processors. But how are we to go about identifying individual memory faculties? Which properties of a memory identify a separate "system"? In 1704 Liebnitz identified a dichotomy between memories based on conscious awareness. This dichotomy is on memories we are consciously aware of and others we cannot explicitly remember but leave remaining effects (Leibniz, 1916). Memories that are revealed by facilitated performance on a task due to previous experiences, without requiring conscious or intentional recollection of those experiences, are called implicit memories. The classification of implicit-explicit is based on psychological characterization at the time of retrieval. It does not make any statements on whether distinct memory systems are required for each memory expression (Schacter, 1987).

Other memory system dichotomies are based on qualitative criteria of the memory contents. Ryle (1949) a philosopher, identified two types of memories, ones that can hold information related to "*knowing how*" and others to "*knowing that*". "*Knowing how*", involves an implicit type of memory with a capacity for skill learning and learning motor procedures. The memory for "*Knowing that*" involves the subjective experience of having an explicit memory expression in mind which can be consciously recollected. This type of memory is the one that enables us to learn and consciously recollect facts and events by modelling the external world and maintaining a historic record of past events. The distinctive terms "*Knowing how*" and "*Knowing that*" define two abstract forms of memory that were not attributed to separate memory systems of the brain at the time.

During the 1960s, important progress was made from studies reporting selective memory deficits in patients with bilateral medial temporal lobe brain damage (see Scoville, 1954, for example). These patients had severe anterograde amnesia and a graded retrograde amnesia extending up to a period before they had brain damage (see Milner et al., 1998). Their memory impairment was selective to a particular type of memory, which involved facts and events, or otherwise a memory for "*knowing that*" in terms of the earlier description by Ryle. They could register perceptual information normally showing an intact immediate or short-term memory, but the information would cease to be available within about 30-40 seconds, when the subjects would lose focus or get distracted (Prisko, 1963, Sidman et al., 1968). Unilateral left or right medial temporal lobe lesions on right handed subjects produced deficits on recall and recognition that was material-specific. Depending on the side of the lesion, deficits were shown for verbal or nonverbal information on particular categories of objects (Milner, 1971).

Studies for spared learning and memory in amnesia, due to bilateral temporal lobe lesions, showed that these patients could learn new visual-motor skills like drawing by looking through a mirror (Milner, 1962).

Another fully intact learning in amnesics is *repetition priming* (Warrington, 1968). Priming involves presentation of stimuli samples (word list, pictures, objects) and then subsequent completion tests using fragments of the samples or very brief presentation of the whole item. Priming causes increases in recognition speed at the presentation of a fragmented item, or it facilitates partial word completion to match a learned sample word. Amnesics could learn through priming and showed retention across days although they could not explicitly recall the samples or had any idea that they had been training for the task.

Further investigations on spared learning revealed a capacity to acquire habits and permanent assignments of cues to appropriate responses (simple conditioning). But conditioning in amnesic patients was shown to be very sensitive to interference (Cohen, 1984); new learning would easily disrupt previous learned assignments. An extended review of spared memory in amnesia can be found in Eichenbaum and Cohen (2004) and also Gabrieli (1998).

To explain spared learning in amnesia, Cohen and Squire (1980) identified the dichotomy of memory expression using the terms *procedural* for the earlier *knowing how* distinction and the term *declarative* to replace the earlier *knowing that* characterizations of Ryle (1949). *Procedural* describes a memory for skilled performance without a record of how these skills were acquired. It was later shown that the spared learning involved various types of learning for which the term *procedural* was inadequate. There were no common features that could allow a single term to describe the spared learning capacity in amnesia. Procedural learning was replaced by the term *non-declarative* which was chosen to serve as an umbrella term for several memory systems (Squire et al., 1988, Squire and Zola-Morgan, 1991). *Declarative* is a memory with a capacity for encoding, storing and retrieving on demand facts and events.

Identifying multiple memory systems soon followed (Cohen, 1984). Anatomical dissociations were made using neurologically impaired patients with selective memory deficits, but also by functional dissociations studies on normal subjects. *Functional dissociation* is based on the assumption that if two memory tasks are served by separate memory systems then they should be affected differently by some test variable.

The spared non-declarative learning found in amnesia is an implicit form of memory served by anatomically distinct memory systems. A taxonomy of these multiple memory systems based on the theory of declarative memory is shown on Figure 1.1.

The gradual organization of memory processes has evolved out of conceptual dichotomies on the type of recall, behavioural function and temporal characteristics of memory.

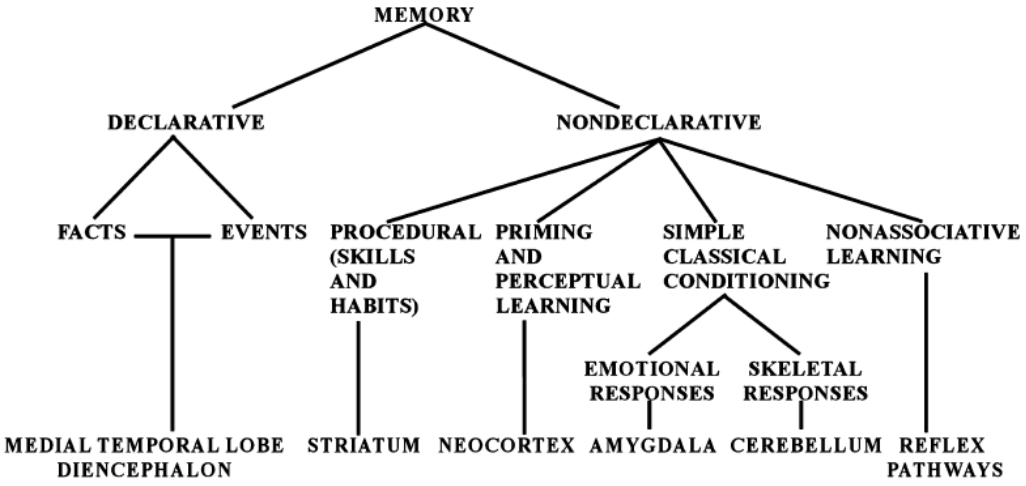


FIGURE 1.1: A taxonomy of mammalian long-term memory systems. The taxonomy lists the brain structures thought to be especially important for each form of declarative and non-declarative memory. In addition to its central role in emotional learning, the amygdala is able to modulate the strength of both declarative and nondeclarative memory. The diagram does not show any interdependence between systems which is known to exist but is yet unclear. (Source: Squire, 2004)

But taxonomies emerging from dichotomies could not adequately classify all observed memory phenomena (Tulving et al., 1982). The nervous system appears to honour the distinction of the declarative memory process since declarative representations can be selectively impaired and are bound to particular brain structures (the medial temporal lobe and the diencephalon).

A consensus has not been reached on the number of memory systems or the categorization. Tulving (1995) combined the conceptual characterizations into five main categories of human memory shown on Table 1.1.

The next sections will briefly cover simple forms of learning and then focus on the long-term memory systems described by the prominent theory of declarative memory (Cohen and Squire, 1980). This theory identifies multiple memory systems by mapping psychological characterizations of distinct memory faculties to neural systems, which have been identified using anatomical and physiological criteria (Squire et al., 1993). Overall, there are three major pathways originating from the cerebral cortex that provide input to the neostriatum, the amygdala and the medial temporal lobe. As we will discuss in sections that follow, these brain areas have been associated with distinct memory expressions.

1.2.2 Non-associative learning

Different nerve cells can adapt their response characteristics depending on their role and the stimulation received. For particular nerve types, a simple change in the gain of their

System	Other terms	Subsystems	Retrieval
Procedural	Nondeclarative	Motor skills Cognitive skills Simple conditioning Simple associative learning	Implicit
P.R.S.	Priming	Structural description Visual word form Auditory word form	Implicit
Semantic	Generic Factual Knowledge	Spatial Relational	Implicit
Primary	Working Short-term	Visual Auditory	Explicit
Episodic	Personal Autobiographical Event memory		Explicit

TABLE 1.1: Major categories of human learning and memory according to Tulving (1995). Some of the declarative memory theory terms are shown on the second column under other terms. P.R.S. = Perceptual Representation System.

response can significantly enhance an organism's survival skills. For example, adapting to respond quickly to the next stimulation keeps an animal alert in a dangerous environment. On the other hand, ignoring harmless stimuli may allow it to move around in a new unknown environment, and this behavioural change can be directly mediated by sensory and reflex neurons that have their response attenuated due to continued or repetitive stimulation. Since these behavioural changes have been studied in reflexes mediated by the peripheral nervous system, it is difficult to suggest that all reflex mediated behavioural changes constitute a memory and are not merely adaptations. Nevertheless, persistent behavioural changes have been observed and the mechanisms underlying these forms of learning may well be the same for other forms of memory that take part in shaping the behaviour of the whole organism.

Two simple forms of such learning will be presented here which have been extensively studied in the famous experiments of Eric Kandel and his colleagues on the sea-mollusc *Aplysia*. This line of research has succeeded to partly uncover the molecular mechanisms of learning. We begin with the first form of simple reflex learning known as habituation.

Habituation Habituation is defined as the decrease in behavioural response that occurs during repeated presentation of a stimulus. This does not affect the nature of the response and only refers to the intensity. An organism that adapts to a new environment

can benefit from such a mechanism by allowing it to ignore unknown non-threatening sensations by attenuating the response these sensations would normally produce.

Sensitization This behaviour describes the increase in the alertness of an animal by a strong noxious stimulus. The increase in alertness is seen by exaggerated responses to stimuli that previously evoked weak responses. It is thus *an enhancement of a response by the introduction of a strong or noxious stimulus.* This learning is not the opposite of habituation as the two involve different types of stimulus to elicit the learned behaviour. Sensitization involves the activation of general arousal systems that affect the intensity of a response. This enhancement has been observed in the reflex responses of both vertebrates and invertebrates. It has adaptive value for an organism as it places it in a state of alertness of potentially harmful stimuli by increasing its readiness to respond to stimuli that in other situations would be harmless.

1.2.3 Conditioning

Conditioning is achieved through a training session where a stimulus that produces a reflex unconditionally (US) is repeatedly paired with an arbitrary stimulus called the conditioned stimulus (CS). After training, the CS is associated with the US and will produce a response which is *identical to the reflex response* it was paired with during training. This response is not to be confused with reflexes which are wired during development, the response is identical to an innate reflex but the stimulus causing it can be arbitrary and the response can be unlearned. Pavlov and Anrep called this type of associative learning the *conditioned response* (CR) (Pavlov and Anrep, 1960). Thompson (1988) examined the eye-blink reflex response in rabbits; this is induced by an air-puff in the eye that is paired with the sound of tone. The locus of this conditioned response memory was identified to exist in the cerebellum and its associated brain stem circuitry by conducting post-training lesions. Appropriate lesions of the cerebellum totally and permanently abolished the learned responses while they had no effect on the unconditioned reflex responses (see Thompson and Krupa, 1994, for a review)

The evidence favoured the hypothesis that the essential memory trace is formed and stored in the cerebellum. Significant experimental evidence in support of the hypothesis emerged and it is accepted that the cerebellum mediates such memory, but it appears that there is redundancy in how such learning behaviour may be mediated. There are multiple parallel pathways that can mediate the formation of the unconditional response (UR) link to the conditional stimulus (the tone), and even the hippocampus is involved in a variant of the task when there is a delay of 500ms in the presentation of US (Beylin et al., 2001, Moyer et al., 1990).

1.2.4 Priming

Priming is a type of learning that can appear as a form of recognition memory that facilitates the detection or identification of test perceptual stimuli due to previous exposure to them. Thus it is an expression of an implicit form of memory.

Amnesic patients have been shown to be severely impaired in recognition of pictures they have seen once, yet priming lasting as long as seven days has been shown to be intact in amnesic patients (Cave and Squire, 1992) when studying memory under a protocol of evaluating the speed at naming pictures. Further it has been shown that it is possible to train amnesic patients to implicitly recall words, for example when asked to list words of a particular semantic category they select the ones they have previously been primed to although they cannot explicitly recall the words they have learned (Graf et al., 1985).

However, even with evidence of a dissociation between priming and an impaired explicit recall it is not clear if the implicit memory of priming is a subsystem of the declarative memory system that remains functional in amnesic patients or whether memory expression somehow combines the output of independent implicit and explicit separate memory systems. In the first case, it can be argued that implicit recall is a rather easier task that a subsystem of the damaged declarative memory system can still handle.

Gabrieli et al. 1995 presented the case of a patient known as M.S., who had intact declarative memory but showed an impairment at visual priming for recognition of words, which was expressed as no improvement in the efficiency of reprocessing the same words. As a result of an accident M.S. had lost the right occipital area of the visual cortex, which is implicated in object recognition, and due to his specific impairment to visual priming he supported the case of a dissociation between normal explicit and impaired implicit visual memory. Combined with the previous results, the double dissociation of priming and explicit memory therefore suggests that independent memory systems are processing and expressing these two forms of memory on a visual recognition task. Further studies using triple dissociation of memory function among the hippocampus, visual cortex and the caudate nucleus in rats (Kesner et al., 1993) verified this dissociation between visual recognition and recall.

The dissociation of declarative and visual recognition is an example of a distinct memory system mediating only part of the memory process which we conceive as a single experience. It is important to note therefore, that visual recognition or an increase in the efficiency at a task can be mediated by a brain system different to the one responsible for recalling facts about an object. Thus faced with a memory task, we do not use a unitary memory faculty and consequently when observing performance at a memory task the different aspects of memory expression may well belong to separate memory systems working in parallel.

1.2.5 Habit and skill memory

One of the three major pathways from the cerebral cortex involves two structures the caudate nucleus and the putamen (see Figure 1.4) collectively known as the neostriatum. These are implicated in associating cortical representations to behavioural responses that are believed to be the basis of habit. Triple dissociation experiments on rats have been used to identify the individual role of the hippocampus, amygdala and the neostriatum (caudate nucleus and putamen) in learning tasks of finding rewards in a maze. The neostriatum was shown to be critical for stimulus approach learning, where rats learn to associate specific stimuli with an approach response (McDonald and White, 1993, Kesner et al., 1993). These and related experiments showed that declarative and non-declarative systems can be used to achieve a goal and the involvement of each is based on the strategy used by the subject. Packard and McGaugh showed that rats in a cross shaped maze could find a reward by learning the location of the reward (place learning) using extra maze cues or by simply learning a turn-left response (response learning) on the first junction (Packard and McGaugh, 1996). Response learning was associated with the neostriatum and place learning with the hippocampus. Further, the role of the neostriatum in habit learning was shown in human subjects using double dissociation experiments between patients with Parkinson's disease and amnesic patients. Parkinson's disease is known to cause striatal dysfunction. By comparing these patients to amnesics the neostriatum was shown to be essential for gradual, incremental learning of associations that is characteristic of habit learning (Knowlton et al., 1996). The mechanisms of habit learning are believed to be part of a broader ability to learn categories by implicit associations between items. In category learning the subject has to extract information about the properties of an implied item by the presentation of a series of items that when averaged together describe the implied item (see Squire et al., 1993). Deficits in such tasks have been found in patients with Parkinson's disease but these experimental results are sensitive to the discrimination rule used for categorisation of items (see Filoteo et al., 2007).

The neostriatum has also been linked with the learning of skills which also requires gradual incremental learning. Patients with Huntington's disease which causes degenerative changes to the neostriatum, were found to be deficient in many skill-based tasks including mirror reading (Martone et al., 1984).

1.2.6 Declarative memory

The declarative memory system supports flexible representation of facts and events that we are consciously aware of. The capacity for flexible representations allows the comparison of facts among learning events and the items involved in the learning process (Cohen, 1984). This memory system is relational in nature and under conscious control (Eichenbaum and Cohen, 2004). According to Tulving (1983), declarative memory

can be divided into *semantic memory*, a memory for facts and general knowledge and an *episodic memory* for autobiographical information (Tulving, 1983), for a review see Tulving (2002).

The declarative memory system is a target of one of the pathways originating from the cerebral cortex and it is served by the medial temporal lobe and the set of structures called the diencephalon. The diencephalon is comprised of the thalamus, hypothalamus and the medial forebrain bundle (Squire and Zola-Morgan, 1991). There are extensive interconnections between the temporal lobe and the diencephalon suggesting a tight functional interaction. Indeed, diagnosing which region is damaged of the two is difficult because the pathology of the two regions is very similar (Richard and Saunders, 1997). That is not to say they mediate the same functions in relation to declarative memory but that declarative memory requires both structures to be intact. Both the diencephalon and the medial temporal lobe project to the frontal lobe. Declarative memory however does not depend on the frontal lobe because it has been shown that frontal lobe damage does not cause amnesia (Janowsky et al., 1989). Instead, patients with damage to the frontal lobe show intact learning of new semantic knowledge but selective loss of when and where (*source memory*) this knowledge was acquired (Shimamura and Squire, 1987, Evans and Thorn, 1966).

Among the two declarative memory structures we will focus on the medial temporal lobe because it consists of structures whose involvement in declarative memory has been extensively researched. The medial temporal lobe memory system consists of the hippocampal formation and the adjacent perirhinal and parahippocampal cortices, see Figure 1.2 (Squire and Zola-Morgan, 1991). The hippocampal formation components are the hippocampus proper (CA1-3), the dentate gyrus (DG), the subicular complex (SB) and the entorhinal cortex (EC). The formation has a largely serial (trisynaptic) unidirectional connection among its components: EC $\rightarrow$ DG, DG $\rightarrow$ CA3, CA3 $\rightarrow$CA1, see Figure 1.3 (right). Of special interest is the CA3 field which consists of pyramidal neurons with collateral axonal projections. These connections form a *recurrent network* that projects to the CA1 (Li et al., 1994). Recurrent networks have received considerable attention as a model for memory and we will discuss how such a networks could store memories in Chapter 4.

The hippocampus receives cortical inputs through the parahippocampal region and projects back through the fornix which connects to the mammillary bodies and then to the cortex, see Figures 1.3 and 1.4.

The declarative memory pathway is from the cortex via the parahippocampal region to the hippocampus (see figure 1.3). This pathway's involvement in memory processing is linked to the ability to form new long-term memories. Patients with medial temporal lobe damage have anterograde and graded retrograde amnesia for facts, events and recognition memory (Scoville and Milner, 1957), see also Squire and Zola-Morgan (1991).

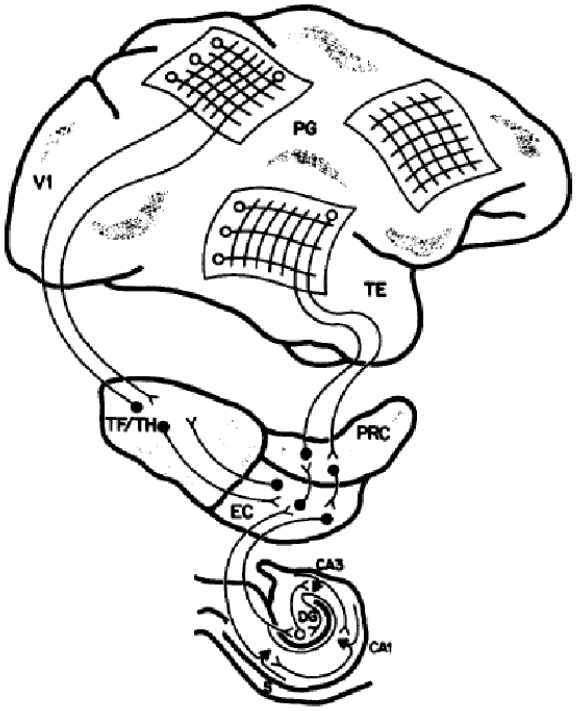


FIGURE 1.2: Schematic drawing of primate neocortex together with the structures and connections in the medial temporal lobe memory system believed to be important for establishing long-term memory. The projection back to the cortex from the TF/TH is through the fornix. The relative positioning of the structures is shown on Figure 1.4. The caudate nucleus along with the putamen form the neostriatum. (PG = parietal cortex; TE = inferotemporal cortex; TF/TH = parahippocampal cortex; PRC = perirhinal cortex; EC = entorhinal cortex; DG = dentate gyrus; S = subiculum; CA1 and CA3 are fields of the hippocampus proper. (Source: Squire, 1990)

A model of this type of amnesia has been replicated in monkeys by surgically removing the hippocampal formation and performing postoperative memory tests. These monkeys received preoperative training in an object discrimination task and two weeks after their operation they were subjected to memory tests. The test results revealed a gradual retrograde amnesia, material learned far in the past was recalled as well as control monkeys but recall was for recent material was severely impaired (Zola-Morgan and Squire, 1990). It appears that initially, memory is dependent on the medial temporal lobe and then it gradually becomes independent of it.

The hippocampus is not the site of long-term storage of declarative memories but is a major component in this pathway although its memory processing function is not clear. The evidence suggests a role in supporting flexible memory representations that are required to complete tasks where the recalled item is conditioned upon the context. In these tasks, the reward and cue association is not direct but conditionally based on the history or spatial cues (McDonald and White, 1993, Eichenbaum et al., 1989). It has also been suggested that the hippocampus is important for learning sequences of events (Fortin et al., 2002) or in a more general class of problems were discontinuous items in time or space need to be associated (Wallenstein et al., 1998). There is also a

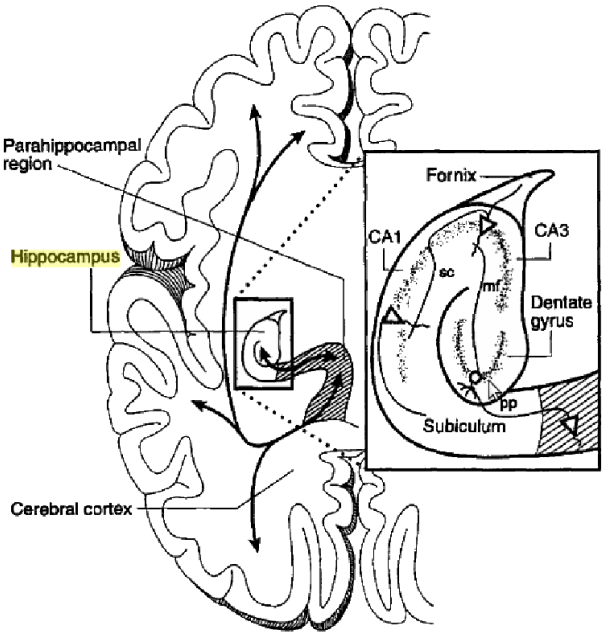
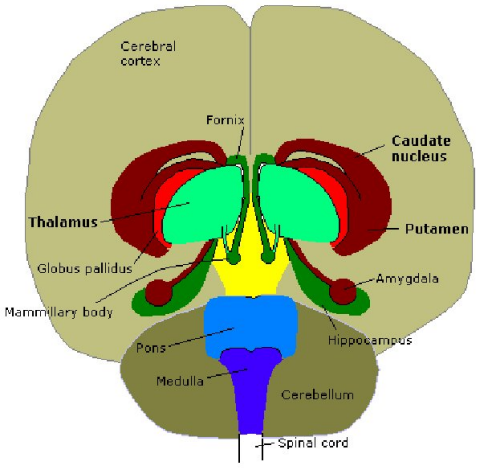


FIGURE 1.3: Major pathways of the hippocampus. Left: illustration of a horizontal section through the human brain showing major pathways by which the hippocampus is connected with cortical areas. Inset: Schematic diagram of major intrahippocampal connections. (Source: Eichenbaum and Cohen, 2004)

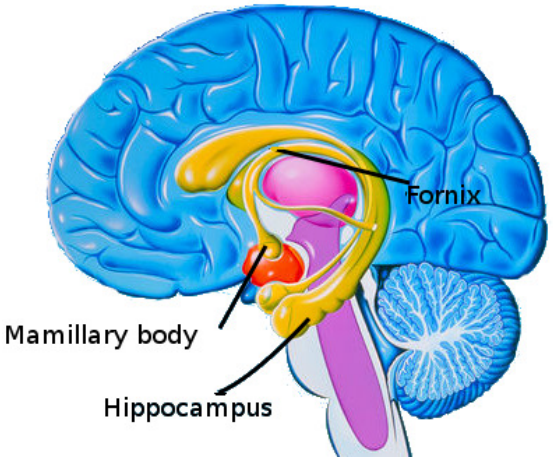
long history of research in rodents and primates showing that an intact hippocampus is required in order for the animals to complete certain spatial navigation tasks (Olton et al., 1978). This relation of the hippocampus to navigation has been reinforced by evidence of hippocampal neurons firing in relation to an animal's location in the test apparatus (O'Keefe and Dostrovsky, 1971), see Eichenbaum and Cohen (2004) for a review.

A specific proposal on the role of the hippocampal formation is that it forms memories by initially binding neocortical representations that together represent the memory of a whole event. This is a view of the hippocampus as a generic binder of disparate information and it can be used to explain how this could facilitate spatial navigation by binding sequences of environmental cues. However, the hippocampus appears to mediate many functions and it is not clear if the binder model is an appropriate representation. There is evidence contrasting the magnitude of hippocampal contribution to memory against the contribution of the surrounding cortex. Experiments on monkeys inducing excitotoxic lesions of the hippocampus, which avoids affecting the surrounding cortical areas, revealed that its contribution in memory retention tasks appears small compared to the perirhinal and parahippocampal cortical areas (Murray and Mishkin, 1998).

Nevertheless, the binding idea provides a conceptual framework on which to begin to consider the organization of memory. According to this conceptual model, the binding of the neocortical representations gradually becomes independent of the hippocampus



(a) Rostral



(b) Lateral

FIGURE 1.4: Rostral lower view of the human brain, important structures have been coloured individually. The thalamus and corpus striatum (Putamen, caudate and amygdala) have been splayed out to show detail. Also shown is a lateral view of the limbic system, showing hippocampus and fornix (yellow),mamillary body, thalamus(pink), hypothalamus (red) (Source: The sciencephotolibrary).

through a process known as *consolidation*. This process binds the neocortical representations that constitute a single memory so that the link established via the hippocampus is no longer required and the memory becomes stable (Squire, 1992). The time required for consolidation varies dramatically spanning from weeks to months. These time-frames however should not be taken as concrete constraints on consolidation as recent experiments on rats showed that rapid systems consolidation within hours is also possible if new memories relate to prior knowledge (Tse et al., 2007). The standard theory of consolidation proposes that memories are initially captured by the hippocampus and are slowly transferred away to the neocortex and finally become hippocampus independent (Squire and Alvarez, 1995). A different line in consolidation theories suggests that multiple copies of a memory trace are progressively created in the hippocampus and the

surrounding cortex. According to this theory, older memories appear more resistant to hippocampal damage because older memories would have had more traces than newer memories. Consequently, the hippocampus is permanently involved in the maintenance of episodic and spatial memory (Nadel and Moscovitch, 1997).

Regardless of the organization of memory traces, we expect the declarative memory system to selectively consolidate relevant information. In the next Section we examine a brain system involved with associating emotional context to declarative memories that can enhance their chances of being consolidated.

1.2.7 Emotional memory

The last of the three major pathways from the cerebral cortex involves the amygdala nuclei, a structurally and functionally heterogeneous region of the cerebral hemispheres located adjacent to the hippocampal formation (see Swanson and Petrovich, 1998). The amygdala is involved in both the acquisition and the expression of fear conditioning (see LeDoux, 1995), and in humans it is believed to be the system specializing in emotional memory; where our fears, stress and phobias are formed. The amygdala outputs affect the hypothalamic-pituitary axis controlling hormonal release, the autonomic nervous system, as well as widespread brain areas. A schematic of amygdala connections to other brain areas is shown on figure 1.5. The amygdala is important for emotional learning, but also responsible for the enhancement of declarative memory through emotional arousal (Adolphs et al., 1997, Packard and Cahill, 2001, McGaugh, 2000). Once an emotional stimulus activates the amygdala, it can in turn impact cognitive processes organized in the neocortex and the hippocampus. The hippocampus in turn is important in adding context to emotional situations, and the interconnections between the amygdala and the hippocampus may play a role in making a connection of the context to an emotional stimulus.

Studies on animals have shown that post-training lesions to the amygdala can interfere with behaviours associated with emotional situations such as conditioned fear, even after extensive training (see LeDoux, 1995), but also with associating a stimulus to emotions (McDonald and White, 1993).

Other studies focused on how the amygdala affect the hippocampal declarative memories under emotional or stressful situations. Stress hormones such as noradrenaline and corticosterone, are known to be released after an emotional experience. The modulation of declarative memory by these hormones is mediated via the basolateral amygdala nucleus (BLA) that expresses receptors of stress hormones (see Akirav and Richter-Levin, 2002, 1999). The link between the hormone induced enhancement of memory and the BLA was shown by selectively infusing a β-Adrenergic receptors agonists into the BLA. Lesions of the BLA or infusion of a β-adrenergic receptor antagonists into the

BLA blocked the memory-enhancing effects that result from a systemically administered glucocorticoid (Roozendaal and McGaugh, 1996). Thus, suggesting that the BLA is required for the hormone mediated memory enhancement.

The amygdala does not only enhance but it can also impair hippocampal memory due to stress (see Diamond et al., 2001, Kim et al., 2001). Evidence suggest that the effect of stress on memory depends on the time difference between the emotional experience and the time of memory formation. If an emotional or stressful experience occurs close in time with a memory then the memory is enhanced while later memories are impaired (Akirav and Richter-Levin, 2002).

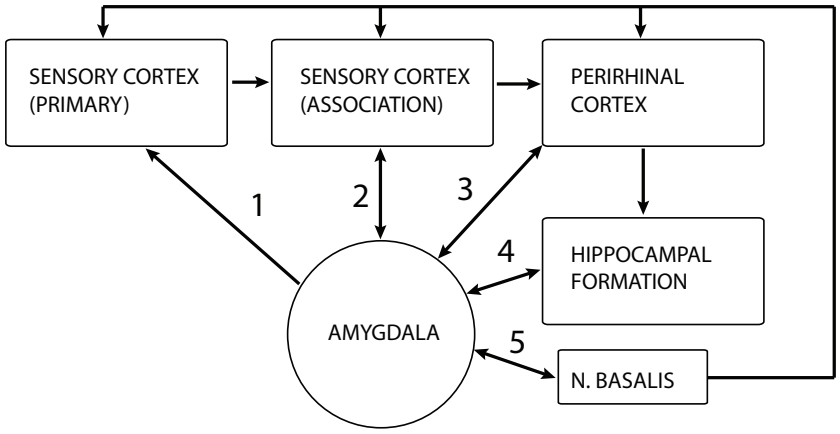


FIGURE 1.5: Amygdala influences on cortical cognitive processing. Once an emotional stimulus activates the amygdala, it can in turn impact cognitive processes organized in the neocortex. The amygdala receives inputs from sensory association areas but not from the primary sensory cortex. However, it appears to project back to primary sensory cortex (1) and to association areas (2). These projections allow the amygdala and its coding of emotional significance to control the ongoing flow of sensory information and may represent channels by which emotional processing can influence perception. The amygdala receives inputs from and projects to the perirhinal cortex and hippocampal formation (3, 4). These structures have been implicated in explicit or declarative memory processing and the interconnections may account for emotional influences on memory processing. The hippocampus is also important in adding context to emotional situations, and the interconnections between the amygdala and the hippocampus may play a role in making a connection of the context to an emotional stimulus. The nucleus basalis (N. Basalis) is the source of cholinergic inputs to widespread areas of the cortex (5) and plays an important role in cortical arousal and attention. Projections from the amygdala to this region may be important in attention and arousal processes . (Source: LeDoux, 1995)

.

In the above system we find an interaction with declarative memory system that can assist in attaching further context to the memories being formed through signals of emotional content and arousal. Studying the mechanisms of memory evaluation in neural system to identify or label memories for consolidation is a very interesting topic in itself. The processes that underlie this selection involve the mechanisms of forgetting and thus examining the dynamics of memory retention may provide partial answers to understand the relationship between memory encoding protocols and consolidation. In the next

Section we discuss the dynamics of memory retention in relation to the prominent view that forgetting is a passive monotonic process.

1.3 The two forces of memory - oblivescence and reminiscence

Obliviscence refers to the process that does not allow the recall of something that once was accessible from memory, or otherwise known as the process of forgetting. The experience of forgetting is very familiar to us all and it has received significant research focus. For a study of memory process, the time course and the mode of forgetting is important because it may help identify the forces that govern it. Another familiar experience is *reminiscence*, the ability to recall something that appeared inaccessible at a previous recall attempt. A particularly intriguing version of reminiscence involves an augmentation in memory content called *hypermnesia*. Hypermnesia is the phenomenon by which more detail may be recalled from memory at a later recall test compared to an earlier recall test. This appears as an increase in recall performance as retention time increases.

In this section we discuss research on oblivescence and reminiscence in order to appreciate that these are two processes exerting their effects on memory simultaneously. This view challenges the way we normally perceive of our capacity to remember, which is a unidirectional route towards oblivescence. We require this paradigm shift according to which memory lies in the balance of two processes in order to appreciate memory dynamics that will be discussed later in this thesis. As mentioned, the common conception for the time course of memory accuracy is believed to be a decaying function. According to this view, the accuracy of the first immediate recall after a learning session is maximum and later attempts to recall score lower due to oblivescence. Also, common belief has it that the reason we forget has to do with limitations on the amount of information that can possibly be stored in the brain.

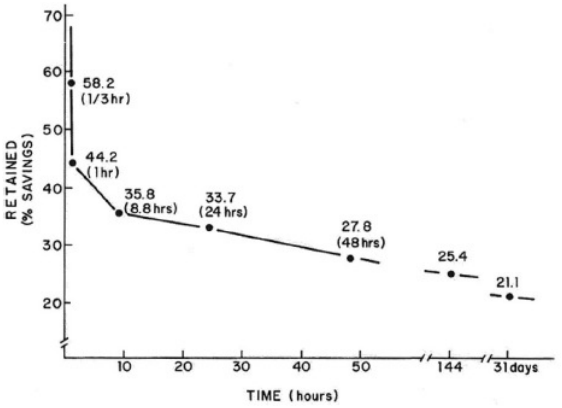
Is oblivescence a *failure* or *limitation of the memory system* or one more of its functions? Limitations could be posed by the number of neurons available in a memory system or by failures caused by fatigue of the mechanisms underlying memory. If limitations are imposed then oblivescence comes as consequence of enabling the storage of new memories. The mechanism of oblivescence could operate passively due to the effects of time or due to storage of new memories that need to overwrite previous ones due to a limitation on the resources. There is evidence to suggest that the inability to recall some information can be due to an executive control suppressing interfering memories with the aim of enhancing the recall of relevant memories (Anderson, 2003). Executive control would constitute an active mechanism for oblivescence that suppresses the recall of some memories without permanently erasing them. Finding such a mechanism would suggest

that this is a function of some *adaptive value* and not a limitation of the declarative memory system. Given that oblivescence has adaptive value it can be argued that multiple processes can exist that attempt to enhance this function in order to enhance the relevance of the assimilated information. We can then interpret the time-course of retention as an expression of a memory system's prediction of the time-period that this information is beneficial. Nevertheless, an executive control of forgetting does not occlude that one source of oblivescence may due to resource limitations of a memory system. These could limit retention at some level since old memories need to make room for new ones, but the effects of these limitations may be meliorated by other processes that determine which memories can be overwritten. Thus, given the complex function such memory systems need to perform we should expect that the dynamics of a retention curve are due to multiple processes interacting simultaneously.

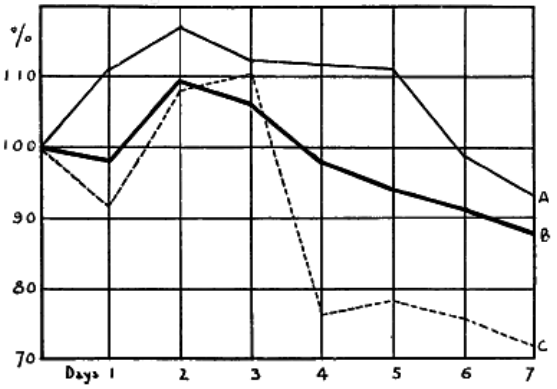
The time-course of retention was originally studied by Ebbinghaus in 1885 who subjected himself in numerous experiments of memorizing 13 nonsense syllables. Nonsense syllables were used to control for enhancement of memory due to meaning or link to previous knowledge. Memory retention was measured by measuring the time required to relearn the studied items after a particular retention interval. This measure is known as the *savings function* and it assumes that the amount of retained memory after a waiting interval facilitates relearning and thus a savings of 50% means that the second learning took only half the time of the original learning episode. During the retention intervals rehearsing or thinking of the material was discouraged. This seminal work produced the first ever forgetting curve and showed that the amount of relearning required increases when increasing the time between initial learning and relearning, see figure 1.6(a).

However, experiments measuring retention times for memorizing semantically linked material or material that is similar to previously studied items reveal that Ebbinghaus's forgetting curve is very pessimistic (Gilliland, 1948). From the early retention experiments of Ebbinghaus it was shown that spacing the relearning at sufficient time intervals significantly enhances the memory savings measured in future learning sessions. This is known as the *spacing effect* of memory and it is now a well known phenomenon of learning in many organisms. The spacing effect is revealed as a significant enhancement in memory savings when compared across the same number of learning trials delivered over short time intervals.

Initial research on the causes of oblivescence showed it was due to *retroactive interference* of new memories with previous ones (Miiller and Pilzecker, 1900). In this study, subjects learned a series of nonsense syllables and it was found that if a second learning episode occurred soon after, their memory of the initial material was impaired compared to just resting for the same amount of time. The researchers' account of this was based on the theory of *consolidation*, according to which memories require a time period for the neural changes to occur and any further learning during this period can disrupt this process erasing a previously stored memory.



(a) Ebbinghaus Forgetting Curve (1885)



(b) Ballard's Hypermnesia (1913)

FIGURE 1.6: a) The percentage of memory savings through time after memorizing non-sense syllables as measured by Ebbinghaus (1885) by conducting an experiment on himself. Memory savings measures the amount of time required to relearn the material that has been forgotten in relation to the amount of time that was required to learn it in the first place. (Source: Erdelyi, 1996). b) The original curves from Ballard's experiments on memory retention on 5192 12 year old children. Two recall tests for verses of a poem that was learned in a single 10-15 minute trial were conducted for each child. One memory recall test was immediately after memorizing and a second was unexpectedly given between 1 and 7 days later. The curves show averages of memory performance measured using the number of verses recalled immediately after learning as the basis on which the percentage is calculated. Curve A and B were on two different poems unknown to the children while curve C used nonsense verses. (Source: Ballard, 1913)

As an alternative, a passive decay account of oblivescence was proposed by Thorndike (1913) according to which unless memories are re-used, their representation decays. This account was rejected by McGeoch in (1932) who showed that the extend of memory decay is relevant to the activities taking place between test recalls. Activities during the retention time that are relevant to the memorized material reduce the decay. It appears that a failure to recall information does not always mean that it is lost from memory but it becomes inaccessible except under the right circumstances (McGeoch, 1932). The author also proposed that *interference* is what causes oblivescence and it can be *retroactive* where new memories suppress previous ones, or *proactive* where a

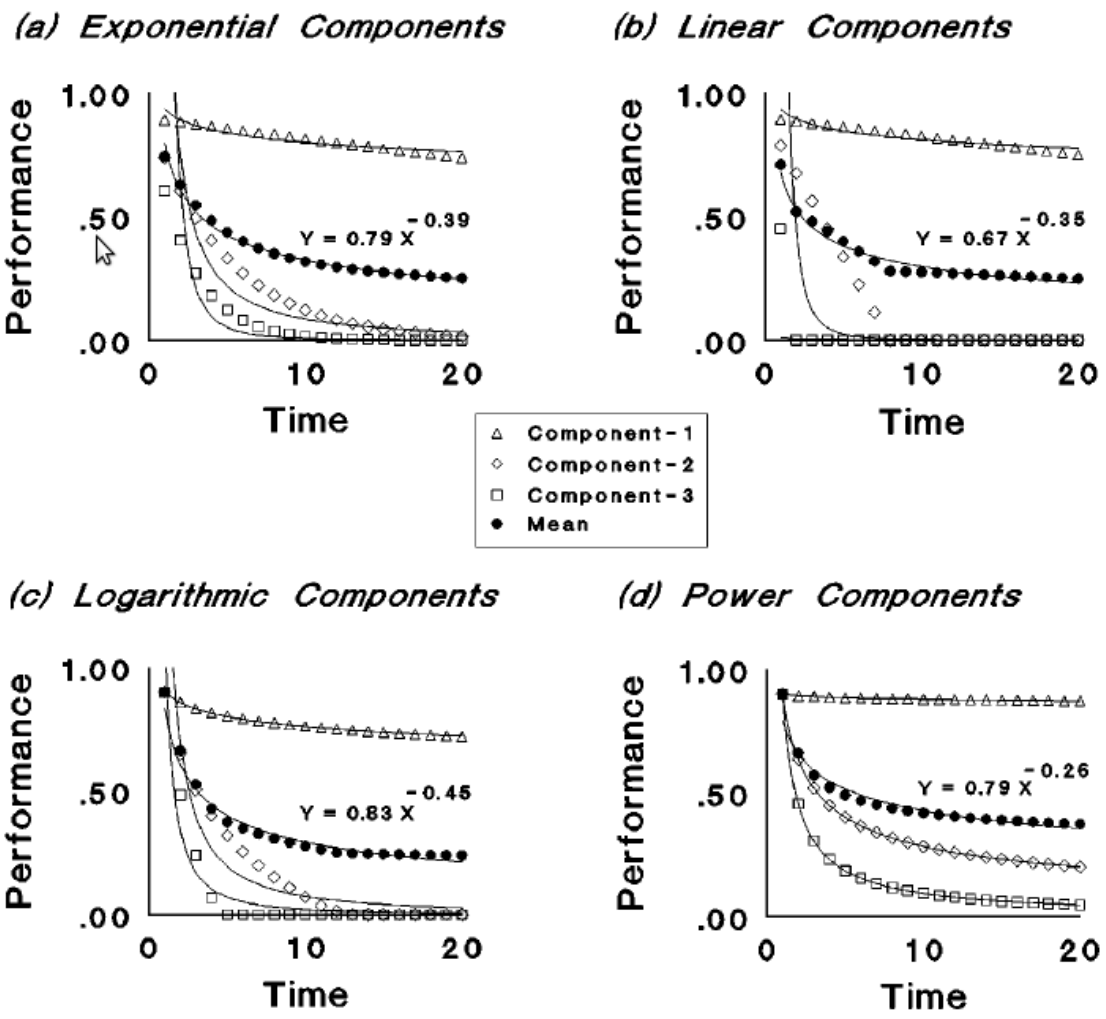


FIGURE 1.7: Aggregate forgetting curves derived by arithmetic averaging of three hypothetical component curves. Each panel shows a different type of component curves: exponential,range-limited linear, range-limited logarithmic, and power. Solid lines show best-fitting power functions. Combined basic mathematical function curves with highly variable slopes tend to produce power curve approximations at the aggregate level (Anderson, 2001).

past memory can impair recall of a more recent memory (McGeoch, 1932, 1942).

Current research on oblivescence suggests that it cannot directly be attributed to a failure or weakness of the memory system arising from the interference of memories, but that there is also an executive control function of the memory system to keep retrieved information updated (Anderson, 2003). An inhibitory action suppressing the response of neurons could be mediating the executive control on what is retrieved in response to a given cue. The inhibited memories are initially not lost or overwritten but rather suppressed and they can be readily recognized and relearned. Whatever the mechanism, oblivescence can be of adaptive value to filter out-of-date information (Bjork, 1989). For simple organisms, oblivescence is still required in order to retain the most recent relevant learning and in such organisms we do not expect to find an executive control

structure. Rather what we expect is a built-in mechanism in the learning rule that gives rise to forgetting. This can be implemented by selectively learning only when necessary. When the conditions for new learning are met then the new learned behaviour directly overrides some previous one. Given that quick access to relevant information can be of great adaptive significance, an executive control mechanism could have evolved for higher organisms to enhance the contrast of related memories. This executive control can be operating in parallel to a simpler process that signals when new learning should occur which has been retained through evolution.

Ebbinghaus's contribution to experimental psychology has been influential and his forgetting function exhibits a power-law relationship of retention against time that has been reproduced by later research under recall, recognition and savings measures on larger samples of subjects (see Wixted and Ebbesen, 1991, Wixted and Carpenter, 2007, Rubin et al., 1999). Wixted and Ebbesen (1991) studied retention in experiments on recognition memory of words or pictures and also found power-law forgetting dynamics. Indeed, a well known function in the literature is the Wickelgren (1974) power-law which is cited as predicting the time course of retention (Wixted and Carpenter, 2007). However, a curve fitting analysis on over 210 of the best available datasets on recall recognition, sensorimotor tasks on people and DMS tasks on humans and animals, against 105 different two parameter functions, showed that four basic functions performed equally well (Rubin and Wenzel, 1996). Further work, used stricter criteria on the precision of the dataset from retention data across five conditions, three continuous recall and two continuous recognition showed that the best fit is a sum of exponentials. The authors proposed the use of the following function to fit the data on the probability of recall after time t:

$$y = \alpha_1 \exp(-t/T_1) + \alpha_2 \exp(-t/T_2) + \alpha_3.$$

They interpret the first exponential as being due to a *short-term memory* process with a decay rate of $T_1 \approx 1$ days, while the second is due to an intermediate or long-term memory process with rate $T_2 \approx 27$ days, the third term α_3 is a constant. The third term is not an exponential since the decay is considered to be too slow to observe experimentally. In their account, temporal differences in memory retention are due to distinct memory systems operating in parallel (Rubin et al., 1999). But how do the Ebbinghaus and the Wixted and Ebbesen power-law forgetting dynamics emerge? As we shall see in a future chapter, power-law forgetting dynamics have been adopted by modellers as representative of biological memory. These are reproduced by combining multiple exponentials with wide ranging time scales. This mixture is known to produce a power-law approximation (Anderson, 2001, Sikstrom, 1999). But power-law like curves can emerge by combining (averaging) a variety of basic functions with slope variability as shown on Figure 1.7. Thus, irrespective of the mathematical form of the function, combining multiple functions of different rates gives power-law like curves and this could explain the ubiquity of the power law in natural processes (Anderson, 2001). Consequently, we

can argue that the observed power-law forgetting dynamics, as measured by a savings function, can be due to multiple memory systems. Each memory system could have its own characteristic time-constant and forgetting curve and thus a power-law curve would not represent any particular isolated neural memory circuit since the memory retention task is using multiple neural memory circuits.

1.4 Limitations of the savings function

The protocol to measure memory retention used initially by Ebbinghaus can obscure another memory phenomenon occurring in parallel that may augment the contents of memory. The savings index measures the relearning effort to reach perfect recall assuming that prior knowledge of the material facilitates relearning. Under a single relearning trial for each material no upward trend can be shown since the best that can be achieved is no relearning being required (100% savings) (Erdelyi, 2010). Wixted and Ebbesen (1991) did not use the savings function, instead they measured the amount of recalled and recognized material for faces and word-list. This line of research measured single encoding episodes against a single recall at a varying retention time interval and recorded the number of recalled items. Therefore, it did not record how the recalled items changed in subsequent recall trials but it also used distracting stimuli during the retention time which may have interfered with the memorized items being recalled.

A different picture emerges if memory is measured after a retention interval and then compared relative to the performance when memory retention is tested immediately after initial encoding without any interfering tasks in between. Such a protocol allows to measure not just the number of recalled items between a set of two trials, but also a comparison of the contents between the recall attempts.

Ballard (1913) examined the *relative retention* between two recall attempts after learning a poem in a single 10-15 minute session. The performance was measured by counting the number of correctly recalled verses independent of order. Ballard observed that the memory of children after learning a poem initially improves with time before it decays. The number of verses correctly recalled two days after learning increased in relation to a test given immediately after learning. On the later trial some of initially remembered verses may have been forgotten, but new initially not recalled verses had become available increasing the overall performance. He concluded that memory exhibits both oblivescence and reminiscence and thus has a tendency to decay as well as increase, these two forces combined define the overall retention. This increase in the detail of memory was later known as hypermnesia (Erdelyi and Becker, 1974) to distinguish it from reminiscence which can imply the remembering of something that was forgotten without an increase in overall performance. As poems carry meaning he also tested a smaller sample of children using nonsense text with a metric that resembled a poem.

The results for all three texts tested are shown on figure 1.6(b). The nonsense verses retention curve C shows comparatively lower overall retention but also displays a characteristic increase in the overall material remembered. Although the experiments used large sample sizes there are methodological concerns in regard to rehearsal of the poems between children of the same school given that text with a metric lends itself to rehearsing. The children were not aware of a subsequent test but they could have been reciting and assisting each other on remembering forgotten fragments for the joy of it. Nevertheless, even if the overall memory enhancement in Ballard's experiments could have been assisted by the nature of material it is well known that whenever subjects are tested repeatedly on material without intervening opportunities to study the material, subjects recall new material on later tests that could not be recalled on earlier tests (e.g. Brown, 1923, Tulving, 1967, Campbell et al., 2011).

Hypermesia has been reliably replicated in learning of pictures (Erdelyi and Becker, 1974) and it is has also been shown in autobiographical memory (Nadel et al., 2007, Campbell et al., 2011) but replication of the phenomenon has not been as direct in learning of word-list (see L. Roediger and A. Thorpe, 1978, Payne and Roediger, 1987). Evidence of hypermnesia in word-lists learning exists but it is less prominent when compared against hypermnesia in picture learning (Brown, 1923, L. Roediger and A. Thorpe, 1978, Payne and Roediger, 1987). Factors such as the number of recall trials (Roediger and Payne, 1982) and the duration of the recall test have been implicated in the extend of hypermnesia (L. Roediger and A. Thorpe, 1978). The timecourse of overall memory retention can be described by differences in the balance between the rate of reminiscence and the rate of obliviscence. If the rate of reminiscence exceeds the rate obliviscence during some time interval then hypemnesia is experienced. The rates of the two processes vary with the learning material and the relevance of the material to the subject (see Payne, 1987, for a review).

The universality of the memory augmentation phenomenon is revealed across learning material, it seen as a change in the recall contents of a target memory. Later recall trials may reveal material that was not originally recalled on the first trial although the overall recall performance may decay (see Payne and Roediger, 1987, Erdelyi, 2010, for a review).

1.5 Summary

In this chapter we presented the manner in which multiple types of learning and memory have been identified and how these have been associated to brain structures. The influence of each system to the performance at particular tasks is related to the strategy chosen to achieve a goal. Memory systems may overlap or substitute one another to perform these tasks but what is learned by each one differs qualitatively. Moreover, we

discussed a form of interaction among memory systems were one system modulates the retention of memories in another. A well known example of modulator interaction is between the amygdala and the hippocampus. This is not the only mode of interaction however and there is evidence to support that there exist memory systems that also interact in other modes that can be competitive or synergistic (Kim and Baxter, 2001).

The amygdala nuclei is involved in the modulation of memory persistence in response to emotions and arousal. This upregulation of learning is believed to be a form of memory allocation and has been described as emotional memory. A similar type of interaction has been shown to occur between the midbrain ventral tegmental area (VTA) and the hippocampus. The VTA has been involved in the modulation of encoding hippocampal memory in response to detection of novel experiences that stimulate the interest of an animal (Lisman and Grace, 2005). This concept of modulation can be traced down to single neurons that receive multiple inputs under distinct chemical signalling pathways and it will presented in the next chapter.

There is also an interaction between the neocortical and the hippocampal system that involves the consolidation and the updating or extinction of memories. Under certain conditions recent declarative/associative memories appear to only temporarily depend on the hippocampus as damage to the hippocampus has been shown to produce retrograde amnesia. However, this temporary dependence can be interpreted in two ways, it could be an expression of multiple traces of a single memory or a gradual formation of a neocortical memory through consolidation (see Wang and Morris, 2010).

Whatever the mechanism of consolidation acting to stabilise memory traces, the opposing action of oblivescence in healthy individuals does not appear to directly reflect the state of some memory trace which fades through time. At least part of forgetting due to oblivescence appears to act as a control mechanism filtering which memories are readily recalled by inhibiting the expression of interfering memories. This introduces oblivescence as a function rather than a limitation of the memory system. This function of oblivescence could come in different forms and it may be necessary to overwrite previous memories that may be interfering. At least in simpler organisms forgetting can be a result of selectively gating when new learning is to occur to replace a previous memory. These mechanisms would have also been inherited in higher organisms through evolution and may operate at particular memory systems. Studies examining the time course of declarative memory retention between two points in time reveal that oblivescence and reminiscence appear to act simultaneously. The balance between the rates of these two processes may reveal hypermnesia, where the memory recalled at a later time point rises instead of decaying contrary to the standard belief that memory decays through time.

The concept of multiple memory systems is currently widely accepted (for reviews see Eichenbaum and Cohen, 2004, Squire, 2004, White and McDonald, 2002, Gabrieli, 1998). According to this theory all systems receive information from the learning environment,

but each system is specialized to learning specific attributes and relationships between the items involved (White and McDonald, 2002) and the systems may interact in various modes (Kim and Baxter, 2001). In this chapter we presented a prominent taxonomy of these memory systems, but constructing a taxonomy can only offer a useful organization scheme and not an understanding of the principles that underlie the organization and specialization of memory systems (Willingham and Goedert, 2001). To further understand memory we have to go deeper to examine how experience can impinge upon networks of neurons. For this purpose in the next Chapter we review the biological mechanisms believed to support memory formation and maintenance in neural networks.

Chapter 2

Synaptic plasticity

In the previous Chapter we presented simple forms of learning and discussed the theories and evidence to support the existence of distinct memory expressions and the existence of multiple memory systems. We summarized a standard taxonomy of long-term memory that identifies memory systems by functional and anatomical dissociations and discussed how these systems interact using emotional memory as an example. A systems' level approach to memory was taken which did not go into the biological mechanisms behind learning in any particular memory system. In this chapter we will be reviewing evidence that link the mechanisms behind learning and recall to cellular-level biological phenomena of the nervous systems.

2.1 Introduction

A nervous system relies on the ability of neuron-cells to form networks and propagate signals in order to process information perceived by the outside world. Within these networks, information is represented by spatio-temporal patterns of neural-cell activity. By adjusting the network connections, nervous systems are able to invoke a particular neural-cell activity pattern in response to a given input pattern, and this is believed to be the basis of memory recall.

Neurons form networks by establishing connections through structures called synapses. Synapses allow the electrical activity of an afferent neuron to affect the electrical state of an efferent neuron by a transmission process that can be chemical or electrical. The ability to induce an observable *change in transmission efficacy* between two neurons is referred to as *synaptic plasticity*. As we will see in this chapter, the change in synaptic efficacy can be brought about by neural activity. These activity-induced changes are believed to underlie behavioural learning and memory.

The duration of a learned behavioural response can take short or long-term forms that may even appear as permanent. As we will see later in this chapter, the duration of some of these learned responses can be correlated with the duration of certain expression of synaptic plasticity. This makes synaptic plasticity a prime candidate mechanism for memory to support the maintenance of behavioural alterations. Assuming that learning and memory is due to synaptic plasticity, we then have to examine the conditions under which plasticity occurs to generate a new memory in response to a learning episode.

Learning via conditioning allows the formation of new association between a stimulus and a response. Although the associative nature of some forms of learning was identified early, the same is not true of the underlying cellular mechanism. Hebb (1949) proposed a theory that places the detection and formation of associations at the level of a synapse by means of a coincidence detection rule (Hebb, 1949). The same synaptic principles are believed to underlie long-term learning and memory in the central nervous system and consequently they constitute the physiological model of memory. According to this theory, memory requires activity induced synaptic modification in brain structures that have been associated with memory. The hypothesis arising is that memories are stored entirely synaptically. Proving this requires establishing a direct link between synaptic plasticity and memory. We will refer to this hypothesis by the initials *SPM* taken from *synaptic plasticity & memory* (Martin et al., 2000).

In this chapter we initially review the SPM hypothesis within a single memory system and extend to a broader view to cover the interactions between memory systems of the brain. According to SPM, the duration of a memory within a single memory system depends on the lifetime of the changes in synaptic efficacy brought about by plasticity during memory formation. These lifetimes along with the requirements to induce synaptic plasticity are examined in three model systems of learning. The aim is to appreciate the common mechanisms of learning between these systems although they belong to very different animals. Also, to recognize that there are various timescales for the duration of plasticity and different requirements for *inducing* each one. First, we find that the duration of plasticity depends on the history of synaptic stimulation. We review evidence showing that the *spacing effect* of memory repetition is also honoured by long-term plasticity. Second, that plasticity is modulated via converging signals believed to convey motivational significance from other brain centres or sensory organs. These modulatory inputs heterosynaptically interact to facilitate or inhibit the induction of persistent forms of plasticity on a stimulated synapse.

In our discussion of the above phenomena we attempt to elucidate the respective cellular mechanism involved. Cells can transfer information all the way from their membrane to the nucleus by using networks of molecular interactions as signalling channels. Neurons have evolved molecular signalling-channel machinery able to detect the conditions for plasticity and to initiate the processes that bring about changes in synaptic efficacy. We will review some of the components engaged in the process in order to reach a broad

understanding of the molecular interactions required for persistent forms of plasticity. The purpose of this endeavour is to appreciate that this molecular network constitutes a signal processing system able to integrate multiple signalling sources in a state dependent manner.

We begin with an introductory discussion on establishing a link between synaptic plasticity and memory.

2.2 From plasticity to behaviour

Previously, we reviewed evidence supporting that declarative memories rely on an intact hippocampus. If these memories are stored synaptically then the hippocampus must support a persistent form of plasticity. This hypothesis of synaptically stored memories has been summarized by Martin, Grimwood, and Morris (2000) as:
Activity-dependent synaptic plasticity is induced at appropriate synapses during memory formation, and is both necessary and sufficient for the information storage underlying the type of memory mediated by the brain area in which that plasticity is observed.

Proving declarative memories are stored synaptically in the hippocampus would signify that the localized trace of the memory *engram* Lashley had been looking for had been found (Lashley, 1950). Moreover, if those experiences that are memorized must induce plasticity and if expressing plasticity can create memories that were not previously there, then we are essentially making plasticity the only mechanism for a particular type of learning (Bear, 2003). Thus, if memory relies on plasticity alone then the characteristics of a particular memory system may be reflecting those of the underlying plasticity (Martin et al., 2000).

The relevant characteristics of plasticity correspond to stimulation requirements that act to *induce* it and the temporal properties of its maintenance after it has been expressed as a change in synaptic transmission efficacy. We will seek to understand synaptic plasticity processes at the level of a single synapse that determine when a new memory is to be encoded and what factors determine the lifetime of a memory. A note of caution should be added though, if the synapse is not the computational unit but rather clusters of synapses or whole dendritic branches are relevant (Govindarajan et al., 2011), then focusing on the single synapse may reveal processes that are secondary for long-term memory.

There are various forms of synaptic plasticity which differ with respect to duration and their underlying mechanism of induction and expression. In the case of SPM, a persistent form of plasticity is required to map to hippocampal long-term memory. As we will present later in this chapter, such a persistent form of plasticity has been found in the hippocampus and it is known as long-term potentiation (LTP) (see Bliss and

Collingridge, 1993). Establishing the link between LTP and memory however is not simple.

Using rats and mice has become an animal model of declarative memory because these rodents have a medial temporal lobe and an anatomically similar hippocampus to humans. Behavioural assays on LTP-impaired mutant rodents show hippocampal plasticity implicated in the formation of high-order memories (reviewed in Mayford et al., 1995, Eichenbaum and Cohen, 2004). Establishing the sufficiency of plasticity for memory in this animal model would require to artificially produce memories via LTP, something which is not currently feasible due to practical issues and the lack of understanding of how memory is encoded. Beyond behaviour assays, the rodent hippocampal slice preparation has become a standard in vitro system for studying use-dependent, long-lasting forms of synaptic plasticity. Nevertheless, although the rodent model has removed many of the constraints faced when experimenting on other mammals, understanding declarative memory is still very difficult due to the complexity of the system. Research using *simpler invertebrates models* has been decisive in understanding the link between learning and plasticity as well as elucidating its molecular underpinnings.

In simple organisms, it has been possible to map particular behavioural learning to modifications of neural connections due to synaptic plasticity. Here, the duration of modified behavioural response is expected to correspond to the temporal aspects of synaptic plasticity. If plasticity mediates memory in higher organisms too, then the same arguments apply and memory lifetime should correspond to the duration of plasticity, at least within a single memory system.

It is certainly conceivable, in more complex organisms, that information storage can outlast the mechanism of maintenance of plasticity (Abraham and Robins, 2005). If memories are transferred or refreshed then the stability of plasticity of particular synapse can be decoupled to the stability of a memory. In this case it is necessary to evaluate what a particular memory system attends to and examine the role of plasticity in that context.

For example lets evaluate the hippocampal memory system, which is required for the recall of recent declarative memories. In order to successfully guide future behaviour we need significant memories to be readily available when needed. Detecting the significance of a memory requires taking into account multiple criteria some of which may be available at the time of encoding. For example, significance of an event could be indicated by the emotional state at the time of memory acquisition. Indeed, we saw in the previous Chapter that emotionally charged events form lasting memories. The duration of declarative memories appears to be modulated by other brain structures that respond to different behavioural or emotional states. These emotional states may be defining the relative significance of a particular memory by signalling positive or negative rewards, and thus one of the functions of the memory system would be to

integrate these modulatory inputs (Shohamy and Adcock, 2010). As we will see later in this chapter, these brain structures influence declarative memory encoding through neuromodulatory afferents exerting heterosynaptic effects on synaptic plasticity. The action of neuromodulators appears to have a decisive role in inducing long-term plasticity and we will be reviewing evidence for the role of neurotransmitter dopamine (DA) in signalling motivationally important events.

Studying behavioural memory and plasticity in the mammalian brain poses many difficulties due to the complexity but also due to the lack of understanding on the exact function that particular brain structures mediate (Martin et al., 2000). Again, simpler organisms are valuable in understanding the basic mechanism. In this chapter we will review experimental evidence from simpler organisms showing that the same principles of neuromodulation have been identified in circuits that mediate a particular behaviour. These organisms lend themselves to study the mechanisms of learning and we shall present two of them before we return to the rather complex declarative memory system.

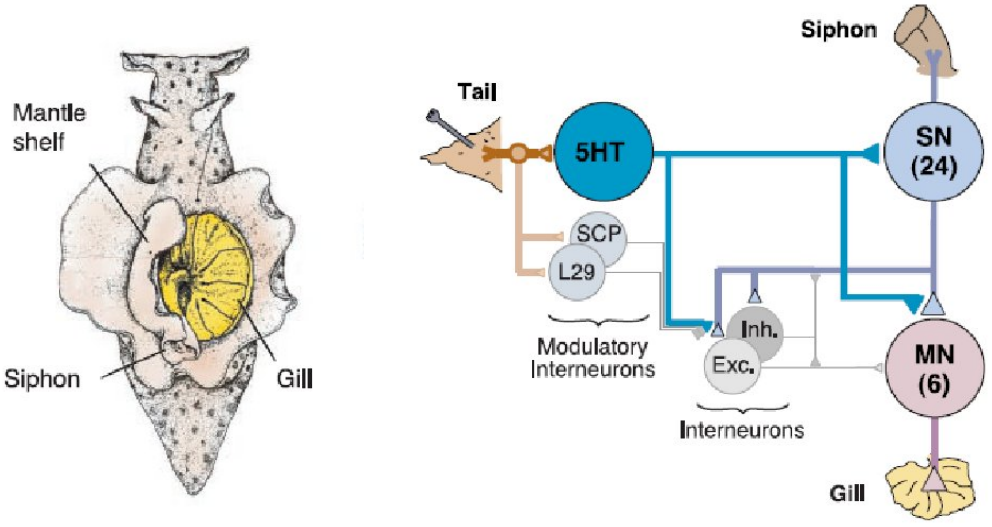
2.3 Learning in Aplysia

The mechanisms underlying simple learning have been elegantly shown in the well known experiments of Eric Kandel and his colleagues on the defensive withdrawal reflex of the siphon and gill in the large marine mollusk *Aplysia*, see Figure 2.1(a). The reflex can be elicited by a light touch to the siphon with a probe which will cause the siphon to contract and the gill to withdraw. Their research examined the neural circuit mediating *sensitization*, *habituation* and *classical conditioning* (see Section 1.2.2) of this reflex response.

Sensitization in Aplysia was elicited by receiving an aversive electric shock on the tail. The shock resulted in an enhanced defensive gill-withdrawal reflex response to a variety of subsequent stimuli even if they were innocuous. In this way a learned sensitization response was elicited making Aplysia react to any tail stimulation as potentially dangerous. A single noxious stimulus produces short-term sensitization that lasts several minutes while repeated strong noxious stimuli, delivered with intervals of pause, produce long-term sensitization that lasts days to weeks (Pinsker et al., 1973).

If instead of a strong noxious stimulus to the tail a weak tactile stimulus to the siphon or mantle shelf was repeatedly used to elicit the gill-withdrawal reflex, *habituation* was observed. Habituation appears as a decrease in the amplitude of the reflex response after repeated weak stimulation, changing the behaviour of Aplysia to not respond to this touch as a threat anymore (Pinsker et al., 1970).

A *classical conditioning* protocol in Aplysia uses a noxious stimulation of the tail to facilitate the defensive reflex response seen by stimulation of the siphon. A light tactile



(a) Aplysia (b) Diagram of the gill-withdrawal-reflex circuit

FIGURE 2.1: **a** A dorsal view of Aplysia showing the gill, the animals respiratory organ. A light touch to the siphon with a one probe causes the siphon to contract and the gill to withdraw. Here, the mantle shelf is retracted for a better view of the gill. Sensitization of the gill-withdrawal reflex, by applying a noxious stimulus to another part of the body, such as the tail, enhances the withdrawal reflex of both the siphon and the gill. **b** Diagram of the circuit of the gill-withdrawal reflex. The siphon is innervated by 24 sensory neurons that connect directly with the six motor neurons. The sensory neurons also connect to populations of excitatory and inhibitory interneurons that in turn connect with the motor neurons. Stimulating the tail activates three classes of modulatory interneurons (serotonergic neurons, neurons that release the small cardioactive peptide, and the L29 cells) that act on the terminals of the sensory neurons as well as on those of the excitatory interneurons. The serotonergic modulatory action is the most important; blocking the action of these cells blocks the effects of sensitizing stimuli (Source Kandel, 2001)

stimulus to the siphon is the conditioned stimulus (CS), which normally produces weak siphon and gill withdrawal. The unconditioned stimulus (US) is a strong electric shock to the tail, which produces a massive withdrawal reflex. Specific temporal pairing of the CS and US facilitates the subsequent withdrawal response of both the siphon and the gill due to stimulation by the CS alone (Carew et al., 1981, Hawkins et al., 1983). Significant associative synaptic facilitation at the sensor-motor connections is achieved when the CS precedes the US than if the temporal order is reversed (Hawkins et al., 1983).

The gill-withdrawal reflex behaviour was found to be mediated by an abdominal ganglion that has 24 sensory neurons that innervate the siphon and make monosynaptic connections with 6 gill motor neurons which mediate the withdraw response (Pinsker et al., 1973, Castellucci et al., 1970, Byrne et al., 1974, 1978). This relatively simple circuit also has indirect connections of the sensory neuron to the motor neurons via groups of inhibitory and excitatory interneurons, see Figure 2.1(b).

The usual experimental preparation required reducing the animal to the required experimental organs by dissecting the mantle, tail and the central nervous system away from the rest of the body and transferring them to an experimental chamber. This setup allowed recording the changes in neural responses using physiological stimulation of the sensory organs. The motor cells are conveniently large allowing microelectrode recordings of the synaptic potentials arriving from the sensory neuron in the behaving or anaesthetized animal. To measure changes in synaptic efficacy the sensory neuron was electrically stimulated using an intracellular electrode that injected short electrical pulses that could be measured at the postsynaptic motor neuron (Kandel, 2001).

Sensitization of the gill-withdrawal reflex is also mediated by neurons in the abdominal ganglion. A single noxious (sensitizing) stimulus to the tail leads to the activation of modulatory neurons that release *serotonin* (5-HT) to the sensory neuron's soma and to the presynaptic terminal of the sensorimotor synapse (Brunelli et al., 1976). In general, blocking the action of these cells results in blocking all learning behaviours that involve the stimulation of the tail to gill-withdrawal response (Glanzman et al., 1989).

Thus, with these experimental preparations the behavioural changes following sensitization of the tail-shock elicited gill-withdrawal reflex could be now paralleled to the duration of synaptic modification in vitro in order to link plasticity and behaviour. However, the in-vitro animal preparation restricted the duration of the recording period to a few hours. To extend the recording period and to directly reduce the model to its basic components a model neural circuit had to be developed. This was achieved by isolating the required neurons from a developing Aplysia and reconstituting the connection patterns by culturing the dissociated cells (Rayport and Schacher, 1986). This provided an *in-vitro* model system reproducing the observed neuronal changes due to behavioural training by directly stimulating the neurons. In its simple form, a single sensory neuron (SN) would connect to a single motor neuron and the modulator neuron would connect to the SN's presynaptic terminal. Later the neuromodulator neuron was removed as it was recognized that in this system tail shocks could be replaced by direct application of serotonin (5-HT) through a pipette. The released serotonin interacted with serotonin receptors on the membrane of the sensory neuron. A single puff of 5-HT on the sensorimotor connection produced short-term facilitation which lasted minutes, while five applications of serotonin spaced over 1.5 hours produced long-term facilitation lasting over 24 hours (Montarolo et al., 1986).

Using these experimental setups, the neuronal changes that underlie three simple forms of learning could be investigated by either measuring the extend of the gill-withdrawal response in the reduced preparation or by recording changes in the excitation of the motor neuron in the isolated neurons. It was revealed that learning has distinct forms ranging from changes that last for a short-term period to ones that last for long-term. The locus of short-term and long-term changes responsible for habituation and sensitization is the same (Frost et al., 1985).

Hence, we have a model organism of learning whose behavioural changes have been reduced to synaptic plasticity, and we can now go deeper by asking what the cellular mechanisms of learning are. Next, we review the mechanism for short-term learning before we go into the mechanisms and requirements that produce long-term learned behaviours.

2.3.1 Short-term learning

Short-term habituation reduces the response of the motor neurons to stimulation from the sensory neurons. The reduction in response is due to depression of the sensorimotor synapse. Evidence from a quantal analysis of short-term synaptic depression at the sensorimotor synapse suggest that depression is due to a decrease in the number of transmitter quanta released per impulse (Castellucci and Kandel, 1974). The sensitivity of the postsynaptic receptor that mediates synaptic transmission does not change and thus short-term depression has a presynaptic locus. This is also supported by research showing that synaptic depression is expressed in the presence of an antagonist of the postsynaptic receptors (Armitage and Siegelbaum, 1998). A morphological study of the changes after short-term habituation revealed a reduction in the number of neurotransmitter synaptic vesicles that are available close to the presynaptic active release sites (Bailey and Chen, 1988b). In this section we will focus on the molecular mechanisms of facilitation whose molecular mechanisms are better understood than the less characterized mechanisms underlying short-term and long-term depression (see Glanzman, 2009).

Facilitation of the sensorimotor synapse results after conditioning or sensitization. Short-term facilitation can be expressed in response to single tail-shock training trial or two trials spaced 15 minutes apart (Sutton et al., 2002). This short-term form of facilitation is accommodated by an increase in the number of transmitter quanta released. The sequence of events, which lead to the increased transmitter release, begin after a tail-shock causes 5-HT to be released by interneurons on specific receptors of the *presynaptic terminals* of the sensory neuron.

These serotonin receptors belong to a family of *G-protein-coupled* receptors that mediate signal transduction from the extracellular space to intracellular molecular signalling cascades. When the receptor is activated by binding 5-HT it activates G-protein that mediates the activation of adenylyl cyclase (AC). AC is also a transmembrane protein that acts as an enzyme to catalyze the conversion of Adenosine-5'-triphosphate (ATP) to the second messenger 3',5'cyclic adenosine-monophosphate (cyclic AMP). The catalysis of the conversion process results in raising the levels of cAMP (Brunelli et al., 1976), which is used for intracellular signal transduction. Thus, the action of binding 5-HT on the extracellular part of the receptor raises the intracellular production of cAMP for a few minutes.

The increased cAMP levels activate an enzyme called protein kinase A (PKA) by releasing its regulatory domain from its catalytic domain. The free catalytic units of PKA then phosphorylate K^+ channels resulting to a slow depolarizing potential and spike broadening, among other actions, and thus also increased calcium Ca^{2+} entry. In parallel, an independent mechanism is also activated by PKA and protein kinase C (PKC) that enhances transmitter release by synaptic vesicles (Byrne and Kandel, 1996).

The increased Ca^{2+} levels are part of an important signalling mechanism of neural activity. The intracellular basal concentration of Ca^{2+} is around 10^4 times lower than the extracellular concentration when neurons are at their resting membrane potential. This gives the potential to calcium to act as a clear signalling mechanism by increasing its concentration either through the import of calcium through membrane ion channels or by the release through intracellular stores. Here Ca^{2+} acts as an indicator of spike activity that is evoked by the CS. Spike activity causes an increase in the concentration of Ca^{2+} by enabling voltage gated calcium channels (VGCCs) leading to the release of further Ca^{2+} by intracellular stores (see Ghosh and Greenberg, 1995, for a review). This Ca^{2+} signal is only briefly available as Ca^{2+} binding proteins and ion pumps act to either harvest it back into intracellular stores or transport it outside of the cell. This brief signal can be sensed by a Ca^{2+}-dependent calmodulin (Ca^{2+}/CaM), a protein that undergoes a conformation change in response to a rise in the concentration in Ca^{2+}. Critically, it requires four calcium molecules to bind to it and thus it acts as a non-linear sensor of calcium levels. Once activated, CaM may then further propagate the signal by acting on its substrates, which can be various downstream proteins or ion channels that become potent through CaM mediated phosphorylation (for a review see Chin and Means, 2000).

Learning a conditioned response requires a mechanism to act as a stimulus convergence detector able to detect both the calcium signal due to sensory neuron activation after the CS and the serotonin signal due to US. The response of the detector when the signals coincide will be to produce a rise in cAMP levels to initiate the cAMP-pathway. Studies using calmodulin inhibitors showed a reduction in AC activation in response to Ca^{2+} waves induced by neurotransmitter. When AC was activated via the 5-HT G-protein receptors, application of a Ca^{2+} showed amplified activity and this could be removed by calmodulin inhibitors (Abrams et al., 1991). The organizing principle of the system appears to be that a robust cAMP signal is produced by the AC due to the supralinear addition of coincidental signals from both the unconditioned stimulus (US) and the conditioned stimulus (CS). The US affects the activation level of the presynaptic AC via *G-protein-coupled serotonin receptors* and the CS must act by Ca^{2+}/calmodulin (Ca^{2+}/CaM) due to a rise in Ca^{2+} concentration from sensory neuron activity (Abrams et al., 1991, Yovell et al., 1992).

Further research on the structure of AC proteins revealed a particular type of AC attached to the G-protein coupled 5-HT receptors. This was named AC-AplA and it

has a Ca^{2+}/CaM binding domain that corresponds to the anticipated Ca^{2+} response mechanism. It is activated in response to a rise in Ca^{2+} concentration or due to 5-HT, thus exhibiting *dual regulation*. Genetic knockdown experiments revealed that the Ca^{2+}/CaM-sensitive AC-AplA isoform is responsible for the great majority of 5-HT-induced cAMP-mediated plasticity in sensory neurons' somata (Lin et al., 2010). Therefore, the AC-AplA mediates most of the actions of 5-HT while it can be also activated by Ca^{2+}.

There is also a variant of AC known as AC-AplC that inhibits cAMP production in response to Ca^{2+}. AC-AplC can counteract the upregulation of cAMP from AC-AplA when the AC-AplA is stimulated by Ca^{2+} alone. With the 5-HT signal combined, the cAMP production from the AC-AplA is further boosted to the point of exceeding the losses incurred by the AC-AplC. This dual regulation of AC-AplA has been hypothesized to function as an associative integrator. Coincidental activation by the US and CS produces an enhanced activation of the AC-AplA boosting the levels of cAMP production through the combined action of a rise of Ca^{2+} and activation of G-protein (Abrams et al., 1991, Lin et al., 2010).

Given sufficient AC activation, a rise in cAMP levels triggers the molecular cascade of the cAMP-pathway and the downstream signalling of PKA and PKC. As we saw earlier, these kinases eventually lead to facilitation at the presynaptic terminal to the motor neuron by spike broadening and enhanced transmitter release. The relative contribution of PKA and PKC depends on the state of the synapse, whether it has been depressed in the past or not (Manseau et al., 2001, Ghirardi et al., 1992, Byrne and Kandel, 1996).

2.3.2 Long-term learning

Obtaining behavioural long-term sensitization requires *repeated* tail-shocks delivered *spaced* in time (Sutton et al., 2002). This protocol can be fully replicated in-vitro, using Aplysia neurons in culture, by five applications of 5-HT at the sensorimotor synapse spaced over 1.5 hours. Critically, it would appear that 5-HT alone can induce long-term sensitization but there may be background activity in the sensorimotor synapse jointly stimulating with 5-HT since synapses were not silenced during these experiments.

The mechanism underlying short-term facilitation in Aplysia has been attributed to changes of pre-existing synaptic connections supported by sustained kinase activity. To maintain longer-lasting changes in efficacy it is more likely that stable structural changes occur. These can be morphological changes at existing synapses or the growth of new synapses to strengthen the connection. Indeed, morphological changes showing an increase in the number and the size of active presynaptic neurotransmitter release zones between facilitated and control synapses can be observed by using intracellular labelling combined with the examination of synaptic serial sections under microscope (Bailey and

Chen, 1983). Also, massive structural changes following repeated stimulation can be seen as either growing or retracting synapses. Bailey and Chen (1988a) reported that under a long-term sensitization protocol the number of sensorimotor synapses doubled, showing presynaptic and postsynaptic structural changes. In contrast, under a habituation protocol, which reduces the magnitude of the reflex response, sensory neurons retracted around 35% of synapses to motor neurons.

These morphological and structural synaptic changes require the synthesis of new proteins either for constructing new synapses or for modifying the structure of existing ones. Using protein synthesis inhibitors can specifically block the *long-term facilitation* of the sensorimotor synapse (Schwartz et al., 1971, Montarolo et al., 1986, Castellucci et al., 1989). However, using agents that block protein synthesis is a crude method which does not specifically block protein synthesis for plasticity. In effect protein synthesis is blocked throughout the cell and thus its normal function may be compromised in such experiments. Plasticity related protein synthesis is signalled through the cAMP-pathway, this same pathway is also recruited in short-term plasticity to exert synapse-local effects, but for long-term memory it appears to play a crucial role in signalling the neuron's nucleus. Applying repeated synaptic activation to induce long-term facilitation causes the released catalytic units of PKA by cAMP to migrate to the nucleus. During that movement, PKA also recruits a mitogen activated protein kinase (MAPK) that, together with PKA, activate gene transcription (Martin et al., 1997b).

Genes themselves are known to interact by forming networks of protein interactions. The product of one gene can be a transcription factor that promotes or suppresses the transcription of another gene. A well-known transcription factor is the cAMP response element binding protein (CREB). CREB is a nuclear protein that modulates the rate of transcription of genes that have cAMP responsive elements in their promoters. The transcriptional efficacy of CREB is affected by its phosphorylation state (Yamamoto et al., 1988) and kinases like PKA and PKC can phosphorylate it and set it in an active state. A link between CREB and long-term synaptic changes in Aplysia was shown by testing the transcription of a CREB-regulated reporter gene under the repeated application of 5-HT (Kaang et al., 1993). The reporter gene expression was not induced following a single pulse of 5-HT, but its induction became progressively more effective following two or more pulses. Therefore, we find that the repeated application of 5-HT, which triggers the formation of long-term facilitation (LTF), also activates CREB. In principle, CREB could also be somehow involved in the process of LTF formation. Injecting an inhibitor of CREB in a cultured neuron preparation showed that CREB activation was necessary for long-term but not for short-term facilitation (Dash et al., 1990). To summarize the process, repetitive spaced stimulation of Aplysia sensory neurons allows PKA to activate CREB which in turn leads to an upregulation of the gene-transcription rate required for the synthesis of plasticity related proteins (PRPs) used for structural changes that support long-term facilitation. The relationship between CREB and long-term memory

has been replicated in neurons of other animal models of learning and is now widely accepted (see Silva et al., 1998).

Gene transcriptional control does not have only expression upregulators but also expression repressors. It is common that the expression of a gene is restrained by both mechanisms. To enable gene transcription, the upregulating transcription factor needs to be activated and the repressor transcription factor deactivated. Both of these should be addressed if we are to understand how stimulus repetition leads to gene-transcription and long-term memory. In Aplysia, the transcription repressor (CREB-2) is deactivated by MAPK that translocates to the nucleus (Bartsch et al., 1995). The effect of combining CREB-2 inhibition and the activation of CREB is an overall upregulation of the rate of gene transcription for PRPs.

To control the activation of CREB-2, the Aplysia CREB-2 was cloned and a specific antibody was produced that can be delivered in an antiserum to inhibit CREB-2. After injection of a CREB-2 antiserum in Aplysia sensory neurons, a single pulse of 5-HT that normally produces only short-term now produced long-term facilitation (Bartsch et al., 1995). These discoveries suggest that memory has positive and negative regulators. Under normal circumstances the balance of these two should provide a minimum activity threshold high enough to ensure only persistent stimuli are able to form long-term memories. However, in situations where particular memory should be encoded under one memory encoding trial (one-shot), the suppressors can be removed and this will allow for the long-term retention of these memories.

2.3.3 Synaptic tagging and capture

In Aplysia, the protein requirement of LTF is believed to be directly served by the activation of gene expression. The assumption is that the stimulation protocol for long-term facilitation activates gene transcription that generates mRNA to synthesize plasticity related proteins for LTF. Gene transcription is a cell-wide process occurring in the cell's nucleus but the products of transcription and the proteins produced thereafter are required for long-term plasticity at a particular synapse. But if synaptogenesis underlies LTF, then synapse specificity should not be implied in any strict sense here, because new synapses will be formed close to the activated site and not from within the activated synapse. Neurons may have hundreds or thousands of synapses and the question arises how do these gene products only act on or near specific synapses. Synapse specificity was demonstrated by culturing a single Aplysia sensory neuron with a bifurcating axon projecting to two motor neurons. Repeated application of 5-HT on one sensorimotor synapse produced long-term facilitation restricted to this synapse alone, implying that the protein products were delivered only to the synapse that was activated by 5-HT (Martin et al., 1997a). Thus, despite the activation of nuclear processes, long-term changes in synaptic function and structure are restricted to those synapses that have

been marked by prior stimulation by 5-HT. How is synapse specificity retained if long-term facilitation requires gene-transcription at the nucleus?

Long-term facilitation requires the activation of gene transcription to produce mRNA that can be translated to PRPs either at the cell body or close to the synapse due to local protein synthesis at the dendrite. Polyribosomes are the machinery that translate mRNA into proteins and it has been recognized that these can be found not only in the soma of neurons, but also preferentially located beneath synapses (Steward and Levy, 1982). By severing the cell body Martin et al. (1997a) showed that LTF induced by five pulses of 5-HT relied on local protein synthesis using locally available mRNA at the presynaptic terminal. Repeating presynaptic exposure to 5-HT five times increased the basal level of local protein synthesis 3-fold. It should be noted here that in fact there are a large number of individual synaptic contacts formed between the sensory and the motor neuron and that the observed facilitation under localized serotonin application is in fact *branch specific*.

One hypothesis to explain synapse specificity assumes that the trafficking of the gene-products is specifically altered. Some transport mechanism is modulated to preferentially deliver gene products to the activated dendritic branch and there is in fact evidence to support this hypothesis showing that mRNA delivery can be targeted to branches that receive 5-HT stimulation (Schacher et al., 1999). However, the mechanisms of long-term facilitation and the relationship with gene-transcription appear to be more complex.

In the bifurcating axon experiments discussed earlier, short-term facilitation induced by single pulse of 5-HT could be converted to long-term if five pulses of 5-HT were delivered at the other sensorimotor synapse branch within the last hour (Martin et al., 1997a). Thus, the weakly stimulated synapse could *capture* the gene-products of the strongly stimulated synapse. The amplitude of facilitation of the weakly stimulated synapses was significantly lower however. Nevertheless, these results demonstrate that a single pulse of 5-HT, which induces short-term changes, marks a synapse in some way that it is able to transform this transient plasticity into long-lasting one if strong stimulation is given at some other synapse. Surprisingly and in contrast to standard view, the transformation of the weakly stimulated synapse did not require local protein synthesis since local perfusion of a protein synthesis inhibitor did not block it. On the contrary, the synapse that received five 5-HT pulses did require local protein synthesis to express long-term facilitation. These last results are somewhat confusing, it becomes suspicious that we obtain a weak form of LTF that does not require protein synthesis as a characteristic of LTF is the protein synthesis component.

How these synapses are marked is yet an unresolved question, for now according to this theory of synaptic capturing, these synapses are known as “tagged”. This phenomenon of interaction of two synaptic connections is not Aplysia specific and a similar interaction between two stimulated pathways in the hippocampus has been found that requires the

use of a "tag", the working hypothesis behind these phenomena is known as *synaptic tagging and capture* (Frey and Morris, 1997).

The processes that enable the *capturing* of proteins by activated synapses appear to engage after a synapse receives weak stimulation that induces short-term facilitation. Another in-vitro experiment, using again a single sensory neuron innervating two motor neurons, showed that short-term facilitation could initiate CREB dependent transcription and be converted to long-term if five pulses of 5-HT were delivered anywhere on the neuron (Casadio et al., 1999). If the synapse to the other motor neuron was primed by a single pulse of 5-HT then long-term facilitation could be expressed on that synapse too.

Efforts to explain the function of this phenomenon in the behaving animal argue that this is a mechanism to allow the interaction of two synaptic events distant in time because the "tag" on the weakly stimulated synapse can persist for hours. According to this theory, a significant event like a fearful experience occurring at some point in time can be associated with events that followed or preceded this event within an interval of a few hours, breaking away from the close contingency in time required for plasticity induction in other protocols (Martin et al., 1997a).

The idea that a particular molecule acts as a "tag" to signpost synapses for capturing proteins has received attention and although the identity of the tag/tags is unknown certain criteria for them have been set. The possibilities are broad and any class of synaptic molecules that can be spatially restricted and persistently mark previous synapse activity by having either its phosphorylation, configuration or concentration altered in response to synaptic activity, can be a candidate tag (Martin and Kosik, 2002, Barco et al., 2002). The proposed tag for active synapses in *Aplysia* sensory neurons is the *cytoplasmic polyadenylation element binding protein* (CPEB and ApCPEB for Aplysia specific) which is found in low concentration in naive synapses and in an inactive or repressive state (Si et al., 2003b). The function of activated CPEB is to act on dormant mRNA. Dormant mRNA is in a state that does not allow its translation by polyribosomes into proteins. Active CPEB changes this mRNA into a state that enables its translation and through this mechanism local protein synthesis is regulated. Thus, the state of CPEB can act as a switch to activate mRNA for translation at the synapse. The lifetime of the switch's active state will determine the duration that the protein synthesis process remains enabled. For long-term memory the assumption is that CPEB's state should be actively maintained for as long as the memory is required (Kandel, 2009), thus it would act as a "memory molecule". For a protein to remain in a state for that long it needs to exhibit *prion*-like properties.

Prions can switch between two conformational states, the active one is self-perpetuating with the capability to transform neighbouring inactive molecules into the active state. Once activated, ApCPEB allows the translation of dormant mRNA at synapses. ApCPEB's

concentration can be increased by a single pulse of the neuromodulator 5-HT while depleting ApCPEB results in specifically inhibiting the maintenance of late long-term facilitation and not its early expression (Si et al., 2003a). Hence, without ApCPEB long-term memory processes that depend on protein synthesis are initiated and can last days before they are blocked. To maintain the self-perpetuating ApCPEB activation there needs to be a sufficient concentration and thus its properties are concentration dependent but still they are consistent with the requirements for a synaptic tag (Si et al. 2003b, see Barco et al. 2008, for a review).

The ApCPEB tag hypothesis has the appealing factor that its conformation state may self-perpetuate to outlive the protein lifetime restriction. However, it has been shown that the weakly activated synapse that "captures" the gene products, due to the strong activation of the other synapse, does not initially rely on local protein synthesis (Martin et al., 1997a) while other experiments show that the role of CPEB is significant for the maintenance of facilitation beyond days but not for its initiation, and thus although it could act as the tag its role seems restricted to late-phase processes (Si et al., 2003a).

The "memory molecule": The hypothesis that ApCPEB could act as a tag whose lifetime is linked to the lifetime of the memory may be appealing, especially to molecular biologists, but it is not new but rather part of a conceptual framework claiming that the duration of a memory is linked to the state-lifetime of a particular molecule, which we would analogous to a "holy grail" for molecular biologists working on memory. This has been the case for type II CaM kinase (CaMKII), whose role in LTP will be discussed in a later section. The hypothesis was raised that the kinase, which is localized at synapses, can sustain its active state for long-term due to its ability to autophoshorylate and thereby it may act to maintain the potentiated state of a synapse for long-term (Lisman, 1994, Lisman et al., 1997, Graupner and Brunel, 2007). Yet more recent experimental evidence report that the activation of CaMKII is rather transient lasting around a minute (Lee et al., 2009) Similarly, a self-perpetuating isoform of protein kinase C (PKC) known as PKM-ζ has received significant attention. This enzyme is believed to be postsynaptically localized in order to maintain synaptic enhancement specifically to synapses that have gone through a learning experience by affecting the trafficking of AMPAR to the postsynaptic membrane active site (Sacktor, 2010). Pharmacological inhibition of the enzymes activity through infusing the zeta inhibitory peptide (ZIP) in the hippocampus has been shown to erase all memory. These impressive findings did not withstand the complexity of synaptic plasticity however. Soon, counter-evidence followed from research that used genetic manipulation to disrupt the PKM-ζ and PKC-ζ genes either constitutively or conditionally in order to show that the memory deficit was caused by the specific actions of ZIP and not the disruption of enzyme's activity (Volk et al., 2013, Lee et al., 2013). Nevertheless, genetic manipulations can also lead to ambiguity since it is possible that the silencing of an enzyme's gene can lead to

the compensatory upregulation of a replacement enzyme, which could rescue memory function (see Glanzman, 2013, for a review). In contrast, after days of training a rat on a taste-aversion task, acute disruption of PKM-ζ by inducing expression of an inhibitory form of the enzyme was shown sufficient to erase an established memory (Shema et al., 2011), thus providing evidence towards memory being maintained by the active state of PKM-ζ.

Arguably, structural changes and the growth of synaptic spines can be more stable for storing a memory than in the conformational state of a large molecule that is subject to molecular turnover and stochastic fluctuations. We note that the idea of a memory molecule is not being discussed in the shadow of structural plasticity as seen in the expression of long-term memory which is accompanied with dendritic morphological changes but also growth of new synapses (see Lamprecht and LeDoux, 2004, for a review). Thus, we anticipate that molecular pathways are initiated due to learning experiences but maintenance of stable long-term memory may associated with the lifetime of the active maintenance of synapses and the activation of whole pathways. Such processes will involve molecular pathways involving numerous molecules and it is unlikely that long-term learning and memory of an animal would rely in the activity of a single type of molecule but rather we expect a system giving some form of redundancy and stability against noise. In these active processes some molecules may be more important than others, and since memory maintenance is probably a dynamic process, acute disruption of one key player may not allow the system to compensate for the loss and thus in such experimental protocols a memory deficit is observed (see Frankland and Josselyn, 2013).

The molecular view of memory assumes that existing synapses retain the synaptic weight by increasing the potency and the number of receptors, with possibly structural changes in the size of the synaptic sites. On the other hand, if the growth of new synapses maintains late-phase potentiation, as it has been already discussed above for Aplysia LTF, this would abolish the need for a maintaining memory in the state of a molecule acting as tag of potentiation for the long-term.

A molecular tag, whether it relates to simply capturing plasticity factors or it extends to the maintenance of long-term plasticity at synapses, may be just of a convenient way to think of the phenomena but probably not related to the actual processes that take place. Thus, the idea of a marked synapse might be just conceptual tool to represent a whole process that is initiated at weakly stimulated synapses. Perhaps weak stimulation initiates growth processes and local protein synthesis but the expression of the processes may be restricted by the availability of particular gene products (Martin and Kosik, 2002). For example, weak stimulation may initiate the process for synapse growth but this process is limited by the availability of plasticity-factors such as proteins and related mRNA. A strong 5-HT stimulation could upregulate mRNA and protein production and this could serve all synapses that have the demand and thus weakly stimulated synapses too. This later hypothesis is supported by experimental evidence showing that five 5-HT

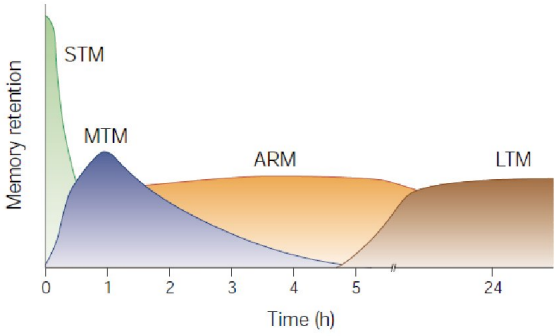
pulses to the soma were sufficient to produce LTF at a weakly stimulated synapse too. On the other hand, if the stimulus was not strong enough so as to produce sufficient plasticity-factors then the growth process cannot occur. In fact there is evidence demonstrating competition for capturing the plasticity-factors among synapses by a study that limited the protein-synthesis in hippocampal slices between two independent pathways, one weakly and the another strongly stimulated (Fonseca et al., 2004). In the control case stimulating both pathways simultaneously results in LTP lasting over 4 hours, but if subsequently the weakly stimulated pathway is reactivated under the action of a protein synthesis inhibitor, the reactivated pathway exhibits additional e-LTP but at the expense of the previously strongly activated pathway, whose previous enhancement is simultaneously lost. Such evidence would be consistent with a role of an active process behind the initiation of plasticity at activated synapses without the need of a tagging molecule, and therefore also consistent with a view of a dynamic maintenance of plasticity.

2.4 Learning in Drosophila

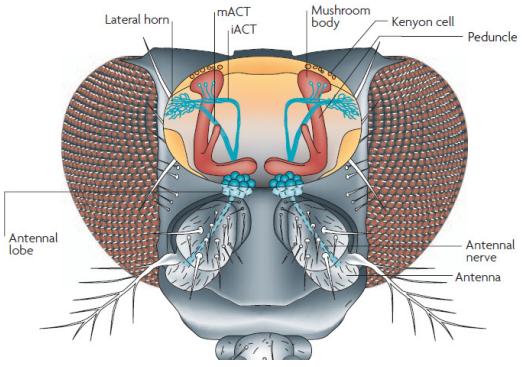
Drosophila Melanogaster is a fruitfly model of learning able to associate a particular odour with an unconditional stimulus (US) of either a positive or negative reward using a classical conditioning protocol (Quinn et al., 1974). After training, a memory induced behavioural change is observed as Drosophila will show a relative avoidance or preference to the learned odour.

The behavioural memory has four distinct phases that have been identified on the basis of genetic manipulation and the requirement for protein synthesis (for reviews see Keene and Waddell, 2007, Heisenberg, 2003, Davis, 2005). The short-term (STM) phase characterizes the initial phase of memory retention and it decays in less than an hour. The next is called middle-term memory (MTM) which lasts from one to three hours and is then followed by a gradual shift to longer lasting memory. These more persistent forms are anaesthesia resistant memory (ARM) and long-term memory (LTM), see Figure 2.2(a). The ARM component of memory is not protein synthesis dependent while the LTM requires protein synthesis. Forming aversive olfactory LTM requires multiple training sessions pairing odour and electric shock with 15 minute rest intervals (Tully et al., 1994, Yin et al., 1995). If the rest periods are omitted only ARM is expressed, and thus aversive olfactory LTM exhibits a spacing requirement for its induction. However, expression of a CREB activator isoform allowed a single training to induce long-term memory in Drosophila (Yin et al., 1995) suggesting that CREB activity acts as "switch" for the allocation of memory. These interpretations however ignore the role of local-protein synthesis for LTM, which we will be discussing in later sections. Instead, appetitive LTM requires a sucrose reward and can be induced with a single training session if flies are motivated by prior starvation (Krashes and Waddell,

2008). This motivational state dependence for LTM induction suggests that the memory circuit receives signals for hunger or satiety. The motivational state must arrive through some motivation signals that are integrated by the odour memory system to decide on whether a new odour-reward association should be made that will guide long-term behaviour (Krashes et al., 2009). To summarize, we find that the stimulation requirements for LTM seem to vary across learning modes but also that motivation signals appear to modulate the process.



(a) Drosophila Olfactory memory phases



(b) Drosophila head showing brain

FIGURE 2.2: a) The four phases of Drosophila memory. Only long-term memory (LTM) requires protein synthesis. There is an overlap among the different phases and LTM rises during a period of Anaesthesia-resistant memory (ARM). ARM cannot be erased by anaesthesia treatment once it has consolidated and there is evidence to support that LTM and ARM are actually mutually exclusive forms of consolidated memory (Isabel et al., 2004, Plaçais et al., 2012). The ARM phase is preceded by middle-term-memory (MTM) and occurs between 1 and 3 hours post-training. The initial phase of memory is short-term (STM) and decays in less than an hour. The distinctions between the early memory phases is based on genetic manipulations which showed that particular gene knock-out mutants lacked specific memory phases. b) Dorsal view of Drosophila melanogaster head showing a representation of the olfactory pathway and the mushroom bodies. (Source: Heisenberg, 2003).

The brain structure that is crucial for the formation of long-term olfactory memory is called the mushroom body (MB), this has been identified by genetically engineering MB deficient mutants who turned out to be also LTM-impaired (Heisenberg et al., 1985). There are two symmetrical sets of MBs. Each contains a network of 2,500 Kenyon cells that receive multimodal sensory input, preferentially from the antennal lobe where

olfactory stimuli are sensed. The olfactory stimuli are processed in two levels before it is sent to the MBs, see Figure 2.2(b).

In a manner analogous to Aplysia, the US activates modulatory neurons (Schwaerzel et al., 2003). But in this system we have two separate US pathways to represent negative and positive rewards. Dopaminergic (DA) pathway is the negative or pain reward, while the neuromodulator octopamine (OA) appears to convey a positive reward signal associated with appetitive motivation (Schwaerzel et al., 2003). An electric shock following an odour activates DA neurons to reinforce learning of an *aversive* behaviour linked with the particular odour (Schwaerzel et al., 2003, Riemensperger et al., 2005).

These neuromodulators are linked to the respective motivational systems by their specific modulatory pathways and they are integrated in the MB along with olfactory neuron activation. The exact circuit mediating this learned response is not known, the data suggests that a coincidence detection of CS and US occurs at the presynaptic terminals of Kenyon cells onto MB extrinsic neurons that determine the fly's behaviour (Heisenberg, 2003, Gerber et al., 2004). Indeed, a set of MB output neurons called MB-V2 has an essential role in eliciting the conditioned aversive response associated with odour and it does not take part in innate aversive odour responses (Séjourné et al., 2011).

The punishment US arrives from a motivational signalling circuit identified as a set of 12 dopaminergic neurons. An interesting experiment was devised to show its role in gating long-term aversive memory in behaving Drosophila. The twelve dopaminergic neurons were genetically modified to express a membrane bound light-activated channel found in the retina. This channel allows activity of the DA neurons to be optically controlled by allowing ion influx in response to particular colours of light. By changing the colour of light in the test chamber the researchers could evoke neural activity, effectively replacing the US, to programme aversive olfactory memory on demand (Schroll et al., 2006, Claridge-Chang et al., 2009). The same experiment was repeated targeting OA neurons believed to convey the US reward signal induced by feeding sucrose. The experiment demonstrated that appetitive behaviour could be programmed by replacing the reward stimuli by light-controlled activation of the OA neurons (Schroll et al., 2006).

Having established that neuromodulators gate particular forms of learning we now seek to understand if underlying molecular mechanisms are similar to the ones found in Aplysia. Therefore we need to examine how the OA and DA receptors exert their effects through G-protein coupled receptors. These are believed to operate via a Drosophila AC homolog (Schwaerzel et al., 2003, Davis, 2005). As in Aplysia, AC activity generates cAMP and initiates the signalling via the cAMP pathway.

Single gene mutations have been identified that produce learning impaired flies in the aversive odour classical conditioning task (Dudai et al., 1976). The normal functioning of the cAMP-pathway is compromised by particular gene-mutations, specifically the *dunce* gene encodes for an enzyme that can degrade cAMP. These enzymes are known

as phosphodiesterases (PDEs) and due to their action being much faster than cAMP synthesis they are able to regulate the localization, amplitude and duration of cAMP signalling within subcellular domains (Nikolaev and Lohse, 2006). Hence, mutant flies of the *dunce* (dnc) gene exhibit very high levels of cAMP due to a deficiency of the cyclic AMP-dependent phosphodiesterase (PDE) encoded by the dnc gene (Byers et al., 1981, Davis and Kiger, 1981). The bnc mutant flies can learn but appear to learn less well than normal flies and forget much more rapidly (Dudai et al., 1976).

Another learn deficient mutant has been identified through mutations of the *rutabaga* gene, which been identified to encode a Ca^{2+}/CaM AC. Mutations on this gene produced learning deficient flies (Levin et al., 1992).

When interpreting the above results we should also take into account that the learning defects found on mutant flies can be due to anomalies during the development of their nervous system and not due to targeted molecular malfunctions only affecting learning in the mature fly. Improved techniques have been devised to address the developmental anomalies that genetic manipulation may pose. Researchers can now produce mutants on which the expression of a particular gene is spatially and temporary controlled.

With these improved techniques, a controlled expression of the *rutabaga* gene showed that the Ca^{2+}/calmodulin-dependent AC is necessary for memory encoding in the MB of the mature fly (McGuire et al., 2003). The rutabaga-encoded AC activity is modulated by the G-protein-coupled receptors for DA or OA. This may lead to elevations in intracellular cAMP and activation of PKA that leads to STM but also to transcription required by LTM (Davis, 2005).

Consistent with the synergistic activation of AC to activate the cAMP-pathway the functional principle that is potentially in operation is that both the CS and US converge to the activation of AC. Indeed, the first cloned receptor responding to OA and tyrosine (Tyr) was shown to inhibit AC (Arakawa et al., 1990, Robb et al., 1994). A functional imaging for the levels of cAMP revealed a differential modulation by OA and DA. Pairing neuronal depolarizations with DA showed an increase in the levels of cAMP in the mushroom body lobes, while pairing with OA had the reverse effect on cAMP (Tomchik and Davis, 2009). Therefore, the converging signals are integrated at the level of cAMP, and these are differentially modulated by two reward centres in a manner that controls future behaviour upon sensing the memorized odour.

As in the Aplysia sensorimotor synapse, coincidental activation of reward and sensory systems for encoding memory is required and the mechanism relies on molecular signal detection at the level of an enzyme (Schwaerzel et al., 2003).

2.5 Hippocampal plasticity as a model of memory

In the previous Chapter we summarized evidence that justify the pivotal role of the hippocampus in the research for the mechanism underlying declarative memory. For the SPM hypothesis to be true in declarative memory, the hippocampus needs to support synaptic plasticity that is sufficient and necessary for the formation of declarative memories.

Initial evidence for a persistent form of hippocampal synaptic plasticity came from experiments stimulating the connections between the perforant path (PP) and the dentate gyrus (DG) granule cells of an anesthetized rabbit, see Figure 2.3. Inducing stable synaptic modifications required high frequency (tetanic) stimulation of the PP in the hippocampus (Bliss and Lømo, 1973). This activity dependent synaptic modifications of efficacy persisted without decay for over 30-60 minutes and it was called long-term potentiation (LTP). Using an extracellular recording technique that measures the field excitatory postsynaptic potential (fEPSP) on groups of DG nerve cells they were able to measure a rise in fEPSP within 30-60sec after tetanus was administrated to the PP.

The discovery of LTP provided a biological mechanism to support the theory of Hebbian learning (Hebb, 1949). However, for LTP to be Hebbian the synapse needs to detect correlated presynaptic and postsynaptic activity and therefore LTP needs to be *associative*. There are various LTP phenomena, but we will focus on an associative form of LTP that can be induced by pairing stimulation of an afferent with strong stimulation of a target neuron. This type of LTP shows that persistent presynaptic activity that activates a postsynaptic neuron may increase the connection strength between the two neurons, just as Hebb had postulated for the growth of a "cell assembly". We focus on LTP/D protocols initially and then we turn briefly to discuss protocols that consider spike-interactions in a spike-timing dependent form of plasticity (STDP) in a later Section.

There is consensus that the mechanism underlying associative LTP is mediated by the N-methyl-D-aspartate (NMDA) receptor, a ligand gated ion channel permeable by Ca^{2+} (see Bliss and Collingridge, 1993, Malenka et al., 1999). This receptor has the critical property that it is ligand activated by glutamate but its conductance is blocked by magnesium (Mg^{2+}) while the membrane voltage is at the resting potential. NMDA receptors allow the entry of Ca^{2+} only if an extracellular Mg^{2+} block is removed by sufficient membrane depolarization ($V_m > -45\text{mV}$) while the ligand glutamate is bound extracellularly (Nowak et al., 1984). Because the NMDA receptor is blocked at resting potential, the initial ion flux needed to provide the rise in membrane potential is initiated by a different glutamate gated ion channel called AMPA after the name of its specific agonist the α-Amino-3-Hydroxy-5-Methyl-4-Isoxazole Propionic Acid. After glutamate binds to the AMPA receptor a flow of cations (Na^+) through this channel depolarize the membrane. Once the membrane is sufficiently depolarized the NMDA receptor mediates a flow of

Ca^{2+} that can be taken as a signal of concurrent postsynaptic depolarization and presynaptic neurotransmitter signalling. The ability to detect presynaptic and postsynaptic activation has given rise to the hypothesis that the NMDA receptor is the substrate of the coincidence detector required for Hebbian learning.

The demonstration that LTP is able to persistently increase synaptic efficacy motivated the search for the reverse process. The expectation was that a mechanism should exist that is able to depotentiate by reversing LTP or by depressing naive synapses. Indeed, a form of the later was revealed with the use of low frequency stimulation (LFS) on the rat PP in-vitro (Dunwiddie and Lynch, 1978) that is known as long-term depression (LTD). Further reports of LTD followed showing LFS can induce homosynaptic LTD on Schaffer collateral projections to CA1 that is also NMDA receptor dependent (Dudek and Bear, 1992). With the addition of LTD, there was now evidence that the hippocampus supports bidirectional plasticity that is NMDA receptor depended while the direction of plasticity depends on the stimulation frequency. Indeed, using high or low frequency stimulation resulted in graded bidirectional synaptic modification with depression shown to erase the effects of potentiation (Dudek and Bear, 1993).

However, in the literature on LTD we may distinguish LTD from the reversal of LTP, referred to as *depotentiation*, as there have been reports that only depotentiation but no LTD could be induced in the CA1 hippocampal region of anaesthetized rats using prolonged LFS (Doyle et al., 1997, Staubli and Lynch, 1990). These two forms of depression may serve different functions in a memory system and are supported by different molecular mechanisms (see Collingridge et al., 2010, for a review of the forms of LTD).

In general we will be referring to the collection of processes responsible for bringing about the changes that lead to enhancement or depression of synaptic efficacy as the *expression* of LTP/LTD. Synapses expressing LTP/LTD undergo changes that seem to involve several components presynaptically and/or postsynaptically (see Malenka et al., 1999). Possible changes include increasing the probability or the quantity of neurotransmitter release, changing the potency of neurotransmitter receptors or their number, modifying the conductance of membrane ion channels, activating silent synapses and structural changes involving the growth of new synapses (Chang and Greenough, 1984, Engert and Bonhoeffer, 1991). The expression of LTP/LTD arises from complex cell biological mechanism interacting with network activity. We shall refer to the collection of events in the network activity that lead to the expression of LPT/LTD as the *induction* of LTP/LTD.

Both, LTP and LTD result in NMDA mediated Ca^{2+} entry and a generally accepted hypothesis has emerged that the postsynaptic level of Ca^{2+} is the critical variable that determines the direction of plasticity (Lisman, 1989). According to this *calcium control hypothesis*, a large transient in the levels of Ca^{2+} in response to strong depolarization after HFS leads to LTP while a modest sustained level of Ca^{2+} whose influx can be due

to LFS leads to LTD. If the levels of Ca^{2+} are below a threshold no plasticity occurs. There are experimental evidence to support this hypothesis, directly modifying the levels of Ca^{2+} in CA1 using a caged calcium compound showed that the level and the transient nature of the Ca^{2+} signal can induce bidirectional plasticity (Malenka et al., 1988, Yang et al., 1999). Also, experiments using either reporters of calcium levels or calcium buffer substances, which are able to reduce the levels of postsynaptic Ca^{2+}, have further supported the link between the levels of calcium and the direction of plasticity (Malenka et al., 1988, Cormier et al., 2001, Cho et al., 2001). According to this view, the levels of Ca^{2+} appear to control the direction of plasticity and thus we have to assume that its concentration level is able to somehow differentially activate the appropriate protein kinases and phosphatases to mediate the synaptic modifications leading to LTP or LTD respectively. However, by examining the levels of Ca^{2+} signals using imaging techniques on basal dendritic spines of the somatosensory cortex during STDP protocols, Nevian and Sakmann 2006 found that Ca^{2+} elevations correlate with both LTP and LTD. The induction of LTD required the activation of metabotropic glutamate receptors. Thus, the calcium levels per se do not seem to dictate the direction of plasticity.

Although LTP/LTD are the primary models of hippocampal long-term memory (Bliss and Collingridge, 1993), establishing a link of these phenomena to memory is not straightforward. Under memory guided behavioural tests, blocking LTP/LTD should disrupt learning and therefore behaviour. However, LTP and LTD are laboratory phenomena produced by highly unnatural patterns of stimulation using high frequency and stimulating many afferent fibres simultaneously. Experiments have attempted to pharmacologically block NMDAR dependent LTP (e.g. Morris et al., 1986, Davis et al., 1992, Steele and Morris, 1999) to test the effects on learning in behaving animals, but the effects of these manipulations on other mental processes are not clear. NMDA antagonist in humans, such as ketamines, have hallucinogenic effects while in mice some sensorimotor side-effects have been observed during the training phase in a water maze task (see Andersen, 2007, pg. 432).

The task known as the Morris water-maze has become standard in evaluating hippocampal LTP dependent contextual map learning in rodents (Morris et al., 1986). The water in the maze is opaque and hides a platform where a rodent can stand safely without having to swim to keep buoyant, once the platform is found the animal is removed from the arena. Learning where the location of the hidden platform is believed to involve place-learning and it is measured as a reduction in the escape latency of trained animals compared to naive ones. Hippocampal infusion of aminophosphonovaleric acid (AP-5), an NMDA antagonist shown to block LTP in vivo, results in impaired learning in this spatial task. Such results offered promising evidence between a link of plasticity to memory but have nevertheless been challenged by arguments towards there being a sensorimotor deficit that correlated with the learning deficit (Cain et al., 1996), but also by showing that pre-training in the absence of AP-5 would abolish the drug-induced

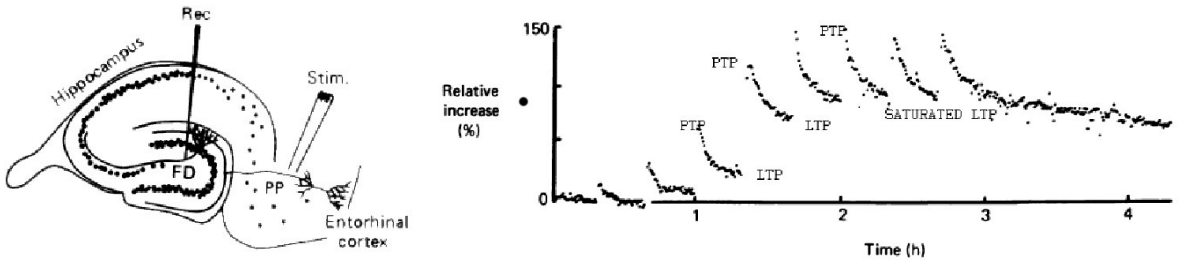
learning deficit under later training (Saucier and Cain, 1995) suggesting that NMDARs may contribute to learning but may otherwise not be essential. Nevertheless, later data suggesting that memory for specific episodes of spatial learning remains dependent on NMDAR and LTP, even after the animals have learned the environment and the general rules of the spatial task emerged by using a delayed match to place (DMP) protocol (Steele and Morris, 1999). Further, genetic knockout of NMDA receptors specifically from CA1 of the mouse hippocampus showed impaired spatial learning and a lack of LTP in CA1 synapses (Tsien et al., 1996). A reconciliation of the above evidence could be achieved if we were to accept that NMDAR plasticity is necessary for some forms of spatial learning while others remain intact (see Eichenbaum and Cohen, 2004, Chapter 3).

Activity driven LTP of synaptic contacts has also been found outside the hippocampus and research has investigated the role of NMDA dependent plasticity in the development of receptive fields in the visual cortex by infusing an NMDA antagonist agent in the behaving animal (see Bear, 2003, for a review). Receptive fields define the area in the sensory space to which a neuron responds to and in the visual cortex these change in response to changes in the input received between the two eyes during a critical period for the development of the visual system. Monocular deprivation causes a shift of receptive fields in the kitten visual cortex which is due to synaptic plasticity on the afferent synapses carrying signals from the two eyes. Infusion of an NMDA receptor antagonist in the visual cortex of a kitten under monocular deprivation showed that these shifts could be prevented (Bear et al., 1990), thus providing evidence towards a critical role of the NMDA receptor in experience driven plasticity.

Nevertheless, a central point of criticism has been that the complicated evidence emerging from research on spatial learning is due to the fact that we are faced with an overall much more sophisticated system than assumed, that uses various forms of plasticity but also with multiple memory systems working in parallel and thus blocking one form of LTP, the NMDAR dependent, cannot clearly reveal the critical role of LTP for memory (see Eichenbaum and Cohen, 2004). Further scepticism stems from the consideration that except of a few exceptions (Good and Bannerman, 1997, Steele and Morris, 1999) with most pharmacological interventions, which attempt to show the link learning and hippocampal LTP, it is difficult to assess if these result in overall hippocampal malfunction or just selectively blockade of learning related plasticity. The learning deficits between LTP- blocking and complete functional shutdown via ablation look very similar (Eichenbaum and Cohen, 2004) thus raising a question on whether the yet unknown normal information processing of the hippocampus is still online with only learning being switched off.

Although these manipulations have given some optimistic indications that hippocampal NMDA dependent LTP is required for learning something about space, they cannot be used as definite evidence for a link between plasticity and memory, specially so

given that there are multiple forms of plasticity and these preparations concentrate on a specific one but also due to the fact that we do not know if plasticity is required for the normal information processing of the hippocampus. Further supporting evidence for SPM may obtained by examining whether the duration of plasticity is sufficient to support behavioural memory expressions, something we examine in the next Section.



(a) Perforant path stimulation

(b) PTP and LTP

FIGURE 2.3: (Left) Schematic transverse section through the hippocampal formation showing the location of stimulating and recording electrodes. (Right)Relative increase of amplitudes of the population EPSPs expressed as percentages of the preconditioning mean amplitude. The dots are test shocks calibrated to measure synaptic efficacy. Every peak occurs after tetanic stimulation which consists 5 bursts of 10 pulses at 400Hz spaced at 1 second, the tetani intensity increases by increasing the duration of the pulses for each stimulation from 30μs-250μs, tetanic stimulations occur every 20min. The first two stimulations show no LTP but a PTP which decays back to baseline within a few minutes. Notice the baseline represents the LTP component which reaches saturation on which PTP is still possible. The LTP component persists until the end of the recording period. (Modified from: Bliss et al., 1983)

2.5.1 Temporal discriminations of synaptic plasticity

Behavioural memory has been characterized by distinct phases, with each phase displaying a characteristic overall lifetime. According to the SPM hypothesis, if LTP/LTD underlie memory then we expect to find distinct phases of these plasticity phenomena too. Indeed, the duration of LTP can vary dramatically with the time-scales of plasticity persistence varying depending on the stimulation protocol used.

Pharmacological manipulations reveal that there are overlapping distinct molecular processes that are required for the maintenance of plasticity during different phases in time. Each molecular mechanism has a distinct time course for activation and persistence and this principle can be seen across plasticity in Aplysia, Drosophila and the mammalian hippocampus (see Huang, 1998, Kandel, 2001, Bliss and Collingridge, 1993).

However, obtaining the full *in-vitro lifetime* of LTP in order to make an association with behavioural memory, which is known to last weeks or years, is difficult. Monitoring the duration of LTP/LTD hits technical boundaries because in-vitro studies can only support a limited recording time of synaptic strength down to a few hours. Obtaining demonstrations of prolonged maintenance of LTP beyond this limit would expand the

temporal range of memories that could, in principle, be supported by it (Abraham, 2003). *In-vivo* studies using rats with chronically implanted electrodes in the DG have shown that LTP generally decays to baseline over a period of several days but potentiation lasting a year has been reported (Abraham et al., 2002).

In this section, we reviewed the elementary facts about the distinct phases of hippocampal plasticity and the terms used to identify them. Although in the experimental literature these terms are not universally referred to, we may infer the appropriate term by examining the molecular mechanisms the expressed plasticity depends on.

2.5.1.1 Transient LTP

The initial phase of NMDA receptor dependent plasticity can last at least 30-60 minutes and it is known as short-term potentiation (STP) or transient LTP (t-LTP). The nature of t-LTP and the relationships between t-LTP and LTP are still under debate. t-LTP could be a premature form of LTP or each can be independent having different mechanisms of induction and maintenance (Volianskis and Jensen, 2003).

Evidence towards identifying t-LTP as a distinct phase were produced by infusing protein kinase C (PKC) inhibitors and showing that this only affects later forms of plasticity and not t-LTP. In these experiments potentiation decayed back to baseline within 60 minutes and therefore t-LTP was shown to be a distinct early form of plasticity that does not depend on PKC (Lovinger et al., 1987).

t-LTP is believed to rely on modifications of NMDA and AMPA receptor proteins. With t-LTP each receptor's potency is modified and the duration of this change relies on each receptors' distinct time constant for t-LTP (Xie et al., 1996). According to this view the locus of t-LTP is purely postsynaptic and based on the state of the neurotransmitter receptors (but see Davies and Collingridge, 1989, Volianskis and Jensen, 2003, for evidence towards a presynaptic locus of t-LTP).

The later phases of plasticity are usually known under the umbrella term LTP, but three distinct phases can be further identified based on pharmacological and genetic manipulations.

2.5.1.2 Early phase LTP

The initial early phase of LTP (e-LTP) can be in the CA1 or DG induced by a single HFS (100Hz). Using inhibitors of protein synthesis in the DG showed that e-LTP cannot exceed 3-6 hours (Krug et al., 1984) and thus initially LTP is protein synthesis independent and probably relies on kinase activity. However, there is evidence suggesting the dependence of protein-synthesis may be modulated by activity; repeated HFS stimuli

spaced at various intervals increasing from a minute to 12 seconds showed an increasing effect of protein synthesis inhibitors on fEPSP measured 30 minutes after induction in CA1 (Fonseca et al., 2006). Thus, if protein-synthesis dependence correlates with the activity dictated by the stimulus then it is not clear whether a distinction between protein-synthesis-dependent and protein-synthesis-independent LTP can be made, since it would appear that synaptic activity increases the turnover of proteins thus modulating the decay of e-LTP (Fonseca et al., 2006).

There is a significant body of evidence indicating that the type II CaM kinase (CaMKII) plays a crucial role in e-LTP of the CA1 area (see Fink and Meyer, 2002, Lisman et al., 2002, for a review). For example, in rat hippocampal slices it was shown that e-LTP induction in CA1 depends on the CaMKII and PKC. Inhibitors of CaMKII or PKC blocked the induction but not the maintenance of e-LTP (Huang and Kandel, 1994, Malinow et al., 1989) while inhibitors of PKA had no effect on e-LTP (Huang and Kandel, 1994, Frey et al., 1993). This early form of LTP is also known as LTP1 and lasts less than 3-6 hours. To maintain LTP beyond that interval it appears protein synthesis is required.

The CaMKII is highly expressed in the nervous system and it may provide the mechanism for the calcium control of plasticity mentioned earlier, and can serve as a link between stimuli and synaptic plasticity (Lisman et al., 2002, Hanson and Schulman, 1992). There are around 20 isoforms of this oligomeric kinase consisting of 12 identical subunits as shown on Figure 2.5(a). Each subunit consists of a catalytic and regulatory domain whose activation depends on cumulative autophoshorylation in a highly cooperative way due to the ability of one subunit to autophosphorylate neighbouring ones. Its subunits are phosphorylated in response to sufficient levels of Ca^{2+} concentration by binding Ca^{2+} and CaM. Under sufficient Ca^{2+} levels, the kinase may become autophosphorylated at which point the kinase is partially autonomous allowing it to retain its state and to act on its substrate, believed to be the phosphorylation of AMPA during LTP (Barria et al., 1997), for prolonged intervals in the absence of a sustained Ca^{2+} signal (see Lisman et al., 1997). Computational studies have suggested that the lifetime of the active state could extend to years (Miller et al., 2005), and therefore as previously discussed serve as a "memory molecule". Nevertheless, more recent evidence using glutamate uncaging and two-photon microscopy to monitor the spatio-temporal dynamics of CaMKII on individual spines show that the activation of CAMKII is of transient nature (1 min), while it remains confined to the stimulated spine (Lee et al., 2009).

The subunits can also be dephosphorylated by the PP1 phospatase and we can expect that the rate of dephosphorylation depends on the concentration of PP1. With this bidirectional control, a bistable autophosphorylation state in relation to Ca^{2+} concentration can exist that ranges from below the resting value of the intracellular Ca^{2+} to the threshold concentration for the induction of long-term potentiation (LTP) (Zhabotinsky, 2000).

2.5.1.3 Late phase LTP

Using a stronger stimulation protocol by repeatedly delivering HFS can elicit a late phase of LTP (l-LTP). Three trains of 100Hz delivered 10 minutes apart in area CA1 of the hippocampus can initiate a slowly rising late phase LTP potentiation in 1-3 hours (Huang and Kandel, 1994, Frey et al., 1993, Reymann et al., 1988). For its expression l-LTP requires protein synthesis, this was first demonstrated in hippocampus by Krug et al. (1984). Infusing a reversible protein synthesis inhibitor in the DG of the anaesthetized rat blocked the maintenance of LTP beyond 3-6 hours. Similar results have been obtained in CA1 neurons in vitro using a protein synthesis inhibitor that successfully blocked l-LTP (Frey et al., 1988). Plasticity related protein (PRP) synthesis occurs within 15 minutes after tetanisation (Otani et al., 1989). Initially, protein synthesis relies on existing mRNA found locally at the dendritic compartments because inhibitors of mRNA synthesis did not affect the maintenance of l-LTP lasting more than 3 hours (Otani and Abraham, 1989, Otani et al., 1989). This initial l-LTP phase is also known as LTP2 and it is yet unclear if PRP synthesis is required postsynaptically only (Otani and Abraham, 1989), because any structural changes that may occur due to long-term plasticity will eventually involve morphological changes at the presynaptic site too.

At the molecular level protein synthesis initiation relies on the cAMP-pathway signalling cascade, see Figure 2.8. Inhibitors of PKA block l-LTP in CA1, while infusing a cAMP analog that activates PKA resulted in triggering l-LTP thus revealing that the cAMP pathway signalling is important for l-LTP (Frey et al., 1993, Huang and Kandel, 1994). Also, genetic manipulation that produced less potent PKA had a negative effect on l-LTP but not on e-LTP (Abel et al., 1997). The levels of cAMP in CA1 are increased by activated AC, and activation of AC could be mediated by the activation of the CAMKII kinase by Ca^{2+} alone. However, given the effects of neuromodulators we have seen up to now with Aplysia and Drosophila, stimulation of AC could be synergistic, requiring the combined action of a biogenic amine such as with the neuromodulator DA. Indeed, pharmacologically blocking DA receptors prior to induction blocked l-LTP (Frey et al., 1991, 1993)

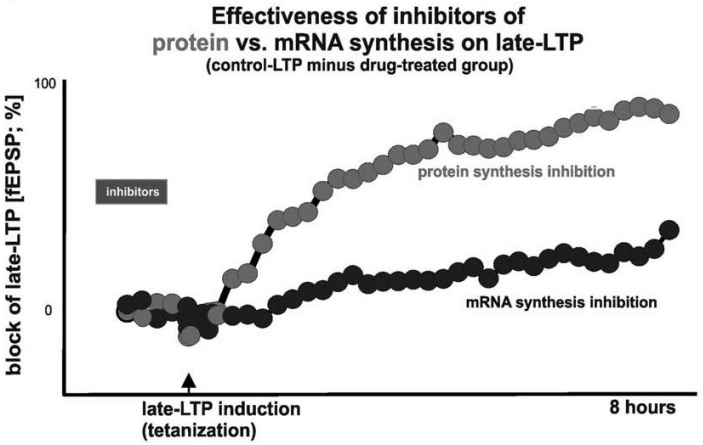
Further, studies using protein-synthesis inhibitors in CA1 showed a requirement for protein synthesis during the induction of l-LTP in what would appear as a transient requirement for protein-synthesis during induction for obtaining l-LTP that lasts hours (Fonseca et al., 2006). But this requirement was shown to be dependent on whether the pathway was stimulated in general, as even test pulses during the application of the protein synthesis inhibitor would affect the decay of l-LTP, thus bringing into question whether protein synthesis is only required during induction. There may, however, be two separate processes running, where the re-activation triggers a form of re-consolidation that results in the decay of expressed LTP in the absence of available protein and it should not be seen as protein-synthesis requirement for the maintenance of l-LTP.

To express l-LTP that lasts several days to weeks further synthesis of proteins is needed and this depends on gene transcription (Nguyen et al., 1994). The transition to a requirement of gene-transcription against an initial requirement for protein synthesis alone has been examined by looking at the relative contribution of each one on the maintenance of l-LTP by measuring the fEPSP . The relative effect of inhibitors, applied prior to stimulation, of protein synthesis against inhibitors of mRNA transcription at various time points after l-LTP induction are shown on Figure 2.4. According to this data, there is an increasing requirement for transcription at later stages of l-LTP, but over the course of eight hours monitored here it appears that it is the availability of protein-synthesis is the strong requirement in relation to transcription. The phase of l-LTP that is protein synthesis and gene-transcription dependent is also known as LTP3 and its decay is a slow process with a time-constant of days (Bliss and Collingridge, 1993). In Aplysia we reviewed evidence supporting that activation of the transcription factor CREB and release of the repressor CREB2 is required for LTF. Removing the repressor CREB2 had the effect of lowering the induction threshold for LTF to a single pulse of 5-HT (Bartsch et al., 1995). The same interplay between genes and memory has been demonstrated in the nervous systems of mammals. Decreasing the analogous CREB repressor in mice also lowered the stimulation threshold for gene-expression dependent l-LTP (Costa-Mattioli et al., 2005) while viral-vector-mediated increases in CREB expression within the rat amygdala had the same effect for long-term memory in a fear conditioning paradigm (Josselyn et al., 2001).

The evidence suggest that each individual LTP phase goes through distinct phases of intracellular signalling mechanisms and each successive stage appears as more stable. These stages should manifest as distinct decay rates of synaptic strength when examining the time course of all phases of potentiation. Curve fitting of LTP decay data using recordings from animals that had chronically implanted electrodes, showed three distinct decay functions (for a review, see Abraham, 2003). These could be characterized by three distinct negative exponential functions which could fit LTP decay curves. The terms used are LTP1, LTP2 and LTP3 and their average decay time constants are 2.1 hours, 3.5 days and 20.3 days respectively (Racine and Milgram, 1983, Abraham and Otani, 1991). These data are only indicative as LTP induced in behaving animals can be influenced by many unknown factors and directly comparing these to the recordings obtained in-vitro is difficult. On the other hand, in-vitro slice preparations have the limitation that they cannot be preserved indefinitely and recordings are usually only obtained for the first 3-8 hours.

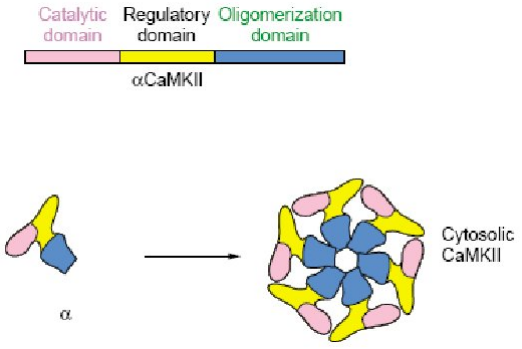
Finally, for the induction of l-LTP we again observe that spacing the induction stimuli is important. Stimuli trains spaced at 10 min intervals generate more persistent LTP than massed delivery (Reymann et al., 1988, Huang and Kandel, 1994). This is another expression of a cellular equivalent of the spacing requirement for long-term learning which seems to operate on all three models of learning we have examined. Although it

would be convenient to link this cellular phenomenon to the prolongation of memory by spacing the training sessions we presented in the previous chapter, the mapping is not direct. The persistence of plasticity has also been shown to be affected by the type of experimental preparation and the region of the hippocampus under study and thus linking cellular phenomena to requirements for learning is complicated (Abraham, 2003). However the ubiquity of the spacing requirement for the induction of long-term plasticity appears as a significant component in the decision process of synapses to express long-term plasticity and we will further discuss this phenomenon in Section 2.6.



(a) The relative requirements for mRNA and protein synthesis in LTP

FIGURE 2.4: The relative effectiveness of protein synthesis inhibitors against mRNA synthesis inhibitors on LTP measured as the change in fEPSP related to the fEPSP level before tetanic stimulation. The inhibitors were infused before delivery of l-LTP inducing tetani in hippocampal CA1. Protein synthesis inhibitors prevented LTP after about 4h, i.e prevented l-LTP, whereas mRNA-inhibitors exerted their action only at later time-points of l-LTP. (Source: Frey and Frey, 2008).



(a) Structure of CaMKII

FIGURE 2.5: The structure of the calcium calmodulin protein kinase II (CaMKII). This kinase consist of 12 subunits that oligomerize into a double hexamer structure (Source: Fink and Meyer, 2002).

2.5.2 Dopamine modulation of LTP/LTD and memory

In the previous Chapter we reviewed evidence to support that multiple parallel memory systems operate in the mammalian brain that serve qualitatively different functions. There, we discussed the modulatory action that emotional states have on the acquisition of declarative memories. In this section we will discuss the mechanisms that appear to underlie this modulation of memory.

At the cellular level, the induction of late-phase plasticity during learning in Aplysia and Drosophila requires the convergence of two transmitter systems on the AC kinase to gate late-phase plasticity. It is then natural to ask if the modulation of hippocampal memory is supported by similar mechanisms in the induction of l-LTP. The experimental protocols for l-LTP we reviewed in the previous Section, required only repeated HFS to induce stable long-term plasticity and did not appear to require the action of neuromodulators. However, this strong stimulation of hippocampal pathways for l-LTP could engage the activation of different transmitter systems due to the simultaneous stimulation of multiple fibres by the experimental protocol. Therefore, under these protocols the induction of LTP should not be viewed as the isolated action of a single neurotransmitter because neuromodulatory pathways could be engaged in cooperative action (see Frey and Frey, 2008).

Tracing techniques have revealed significant DA innervation of parts of the rat's amygdala and the hippocampal areas SB and CA1 from dopaminergic pathways arising from the midbrain's ventral tagmental area (VTA) (Scatton et al., 1980), but see also (Gasbarri et al., 1997). A prominent theory on the function of the VTA is that it controls the entry of new information to the hippocampus via the modulatory action of DA on plasticity (Lisman and Grace, 2005). The evidence we will be reviewing suggest that the action of the neuromodulator DA is to influence hippocampal memory by gating or modulating the prime model for memory LTP/LTD (see Bethus et al., 2010, Rossato et al., 2009).

DA is known to act through a certain class of neurotransmitter and hormone receptors that are linked to their signal transduction pathways through guanine nucleotide binding regulatory proteins (G-proteins), see Figure 2.6. There are currently five known DA receptors (D1-D5) and they have been classified into two categories depending on whether their G-protein activates or inhibits the enzyme adenylyl cyclase (AC) (Kebabian and Calne, 1979).

The role of AC in hippocampal plasticity appears critical (see Hanoune and Defer, 2001, for a review). Supporting evidence show that two abundantly expressed isoforms of AC in the hippocampus, AC1 and AC8, are necessary for signalling the formation of l-LTP and LTM through the cAMP pathway. Specifically, experiments on mice with a genetic knockout of both AC1 and AC8 showed that the mutant mice did not exhibit l-LTP or

LTM. If either AC1 or AC8 was expressed or if cAMP production was increased in other AC isoforms by infusion of forskolin then l-LTP was recovered and so was LTM (Wong et al., 1999). Thus, given the importance of AC for l-LTP and LTM, a key action of DA could be to gate LTM through the production of cAMP via its receptor coupling to the AC enzyme.

The first category of DA receptors contains the D1 and D5 receptor types known to be positively coupled to AC (Kebabian and Greengard, 1971). D1 and D5 have been shown by immunohistochemical staining techniques to be highly expressed along pyramidal cells of the CA1 (Huang et al., 1992a). In contrast, activation of D2-like receptors (D2,D3,D4) results in various responses one of which is the inhibition of AC activity (Vallar and Meldolesi, 1989). To summarize, there is evidence to support a critical role of AC for l-LTP, through its ability to initiate the cAMP pathway, and there is a class of abundantly expressed DA receptors that is coupled to AC. Through these coupled receptors cAMP production can be enhanced or inhibited and given that cAMP is necessary for l-LTP, then we may infer that DA is able to modulate the requirements for the induction of l-LTP.

The effects of DA on LTP appear to form a consistent picture with behavioural assays showing exposure to novel stimuli facilitating LTP by lowering the threshold for its induction. These assays have tested the role of DA receptors in the storage and retention of new memories in a rat spatial learning tasks. Li, Cullen, Anwyl, and Rowan (2003) report a link between exposure to a novel spatial environment and a facilitation for LTP induction in a D1/D5 receptors dependent manner. Blocking D1/D5 receptors by infusing an antagonist into the hippocampi bilaterally prevented the facilitation of LTP in awake animals exposed to novel stimuli, while introducing an agonist in anaesthetized animals facilitated the induction of LTP (Li et al., 2003). Hence, by exposing rats to novel environments the threshold for inducing LTP at CA1 synapses was lowered, while pharmacological blockade of DA receptors suggested that this facilitatory effect was DA dependent.

Wang, Redondo, and Morris (2010) used an event arena and allowed for a training trial so the rats learn the location of food rewards by exploration before later testing recall. They reported that blockade of the hippocampal dopamine D1/D5 receptor or protein synthesis inhibition within 15 minutes of exploration prevented persistent place memory and l-LTP expression. The reinforcement of LTP due to neuromodulators was also examined in the rat CA1 in-vivo by electrically activating the VTA at various times after tetanization. The results showed that e-LTP could be consistently transformed to l-LTP if the VTA was stimulated within 30 minutes after CA1 stimulation (Frey, 2001, Frey and Frey, 2008).

Other than being a requirement for l-LTP, DA also exerts effects on the magnitude of e-LTP through activation of D1/D5 receptors (Otmakhova and Lisman, 1996). However,

these effects are secondary as they appear to be region specific (Roggenhofer et al., 2010) and thus gating l-LTP seems to be the primary function of the D1/D5 receptors.

2.5.2.1 Synergistic action of DA and GLU

The ability of DA to gate l-LTP was demonstrated by delivering tetanization, which would normally produce l-LTP, along with infusion of a D1 antagonist. In the presence of the D1 antagonist the tetanic stimulations failed to produce l-LTP (Frey et al., 1991, 1990). Conversely, introducing an D1/D5 receptor agonist alone without tetanization was shown to give rise to a delayed (after 50-60 minutes) slow onset of protein-synthesis-dependent l-LTP that peaked after 3-4hrs (Huang and Kandel, 1995). This would imply that DA alone is sufficient to induce l-LTP, but in this DA induced l-LTP experiment however, the researchers did not control for the activation of NMDA receptors due to the neurotransmitter glutamate. Although the protocol used did not induce afferent pathway stimulation, noisy background activity may still activate NMDA receptors. Indeed later research showed that applying the D1/D5 agonist in the presence of a NMDA antagonist was inadequate to produce activity independent LTP in CA1 or the subiculum (SB) (Roggenhofer et al., 2010). These results suggest that the activation of *both* neuromodulator and neurotransmitter receptors is required for l-LTP and that there is a synaptic mechanism to detect when both signals are present.

One detection mechanism is a form of direct protein-protein interaction between NMDARs and the D1 receptor, which allows the function of NMDARs to be directly affected by D1 receptors (Lee et al., 2002). On a further level, it has been suggested that the interaction of glutamatergic and heterosynaptic processes, including neuromodulation, is synergistic and occurs at downstream molecular signalling cascades in order to transform early phase to late phase plasticity (Sajikumar and Frey, 2004a, Navakkode et al., 2007).

The signalling molecular cascades involved include the familiar cAMP-pathway previously described in Aplysia and Drosophila learning. Frey et al. (1993) report that they were able to induce a slow onset l-LTP by infusing a cAMP-dependent protein kinase (PKA) agonist. Hence, expressing l-LTP by infusing this cAMP analog suggests that high levels of PKA constitute a sufficient signal to initiate the processes for l-LTP. The researchers postulated that because the l-LTP produced by PKA is identical to the one from D1/D5 activation alone, then the D1/D5 receptors probably share a cellular signalling pathway with PKA that leads to l-LTP.

Another mechanism for synergistic activation of the signalling cascades is through the now familiar AC enzyme that is activated by the combined action of Ca^{2+}/calmodulin

and G-protein. In fact, the D1/D5 receptors are members of the G-protein-coupled receptor superfamily that are coupled to AC that once activated can stimulate the cAMP-PKA signalling cascade; activation of PKA through the calcium-stimulated adenylyl cyclase has been shown to be necessary for the formation of LTM after contextual fear conditioning in mice (Sindreu et al., 2007). This D1/D5 coupling suggests that DA acts mainly through activation of PKA and the cAMP-pathway, a signalling process which is intimately involved with the cell's nucleus. In Aplysia this direct signalling from synapse to nucleus via the cAMP-pathway is responsible for enabling gene-transcription. Gene transcription appears to be a requirement for the expression of late phase LTF and so the synapse seems to signal the immediate demand for new mRNA and PRPs synthesis. In the hippocampal pyramidal neurons however, addressing an immediate demand of synapses may be unrealistic, because there are can be thousands of synapses along a very intricate and large dendrite and thus we expect that structural changes may be frequent and this would impose an ongoing demand for PRPs. Thus, the role of the cAMP-pathway and signalling the cell's nucleus may, in this case, be rather to adjust the rate of gene-transcription in manner that is modulated by the actual dendritic activity. In this manner D1/D5 receptor activation combined with NMDA activity can signal an upregulation of local protein synthesis and affect the rate of gene transcription by activation of the cAMP response element binding protein (CREB) and deactivation of the CREB repressor (see Kandel, 2001, for a review of the PKA molecular cascade).

The effects of D1/D5 receptor activation are not specific to LTP but rather both LTP and LTD have been shown to be dependent on DA and protein synthesis in rat hippocampal slices in vitro (see Lemon and Manahan-Vaughan, 2006, Sajikumar and Frey, 2004a). Chen, Fujii, Ito, Kato, Kaneko, and Miyakawa (1995) showed that if a DA agonist is infused while delivering LFS in CA1 slices then the magnitude of LTD is enhanced. If however a DA antagonist is used to to block DA receptors along with LFS then an late phase LTD component is preferentially blocked, revealing an early LTD component that lasts around 160 minutes (Sajikumar and Frey, 2004a).

2.5.2.2 Relative timing of DA signals to induce LTM

Taking the view that a DA signal and an NMDA mediated Ca^{2+} signal converge to interact in a meaningful way to gate l-LTP we therefore need to examine the effect that the time interval between their activation has on LTM. The experimental evidence do not seem to suggests a very clear temporal window of DA interaction with memory encoding. Behavioural assays have attempted to elucidate the relative time requirements for DA to facilitate memory by either blocking or stimulating DA receptors through the infusion of agents in the hippocampus that act as DA receptor antagonists or agonist. Infusion of an antagonist prior to training rats in a spatial memory task showed an impairment of memory 6 hours later but not immediately after encoding (Carroll et al.,

2006). Infusing the antagonist after training had no effect on the retention of memory, suggesting that DA is required prior or during learning in order to maintain memory beyond 6 hours (see also Lemon and Manahan-Vaughan, 2006, Bethus et al., 2010).

Rossato, Bevilaqua, Izquierdo, Medina, and Cammarota (2009) identified a time-window for the later activation of D1/D5 to interact with a memory encoded in the past and stabilize the trace. They injected a D1/D5 antagonist 12 hours after a fearful experience into the dorsal hippocampus that resulted in the rapid in-vivo extinction of a long-lasting fear LTM, while immediate injection or in 9 hours did not have the same effect. The late interaction of DA with a memory could be due to a late consolidation period using a reward signal, but this very particular timing of 12 hours after a memory event makes the interpretation very difficult and unclear.

Wang, Redondo, and Morris (2010) used an event arena in which rats had to learn the location of a food reward and recall its location 24 hours later. Using a weak reward during training showed that rats where unable to recall the location after 24 hours. Exposure to a novel environment within 30 minutes after food-place learning showed that 24 hour memory was rescued due this novelty signal. The effect of the novelty signal was blocked by a D1/D5 antagonist infused during novelty exploration suggesting that rescue of LTM can been attributed to the activation of VTA dopaminergic neurons after food-place learning. However, memory enhancement was also possible by exposure to novelty 1 hour prior to the learning session but infusing a D1/D5 antagonist during learning did not affect the memory rescue 24 hours later.

It appears that activation of D1/D5 receptor can interact with a memory proactively or retroactively giving a wider time frame for the enhancement of memory encoding but yet the effects of unknown system-wide functions that go into memory consolidation make these results particularly unclear for understanding the timing of the cellular processes.

2.5.2.3 A molecular mechanism of memory allocation

We can not say with certainty if DA acts as an initiator or modulator of LTP but it does appear to be critical for the formation of persistent memory traces. A synergistic action of glutamatergic and dopaminergic activation of CA1 synapses appears to gate long-term plasticity through the cAMP-pathway, but not all the effects on the induction of LTP/LTD are clear as well as the relative timing requirements for DA action against the timing of neurotransmitter stimulation.

Both of the neuromodulators we have discussed, dopamine and serotonin, act through adenylate cyclase (AC) which is part of the cAMP-dependent signalling pathway. We also know that calcium influx through the NMDA receptor is the trigger for NMDAR-mediated LTP and the calcium signal also converges to AC through Ca^{2+}/calmodulin kinase II (see Waltereit and Weller, 2003a). The experimental evidence presented above

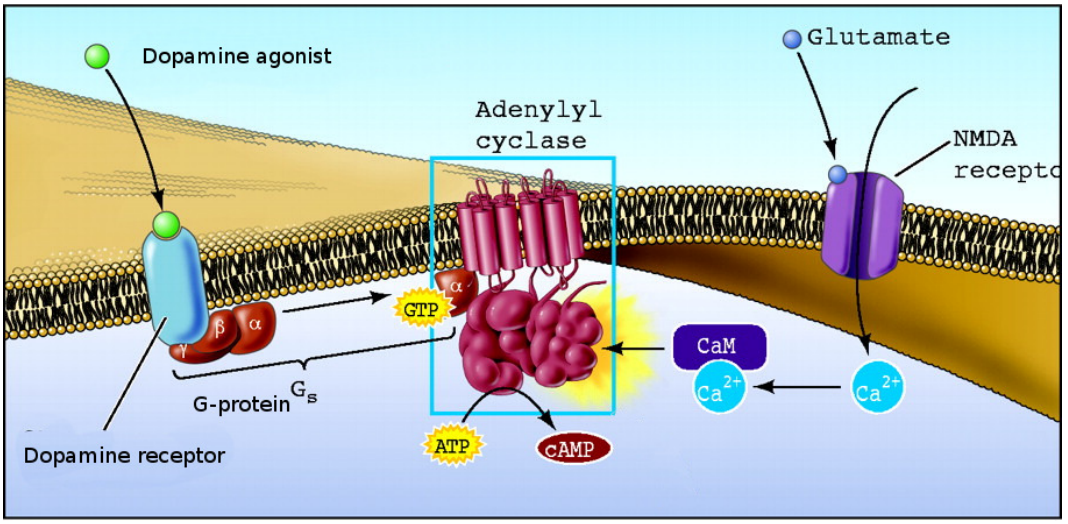


FIGURE 2.6: Model of the activation of adenylyl cyclases by Dopamine through G-protein-coupled receptors and calcium via Ca^{2+}/calmodulin kinase. The combined sensitivity allows the detection of neuromodulator DA and neurotransmitter glutamate. Glutamate binds to the NMDA receptor and with sufficient depolarization it allows Ca^{2+} to enter the neuron. Calcium binds to CaM to stimulated the AC and produce cAMP. The AC1 type can be synergistically activated by G protein coupled receptors like the D1/D5. Activation of G protein–coupled receptor results in a direct association of GTP-bound Gα subunits with adenylyl cyclases in the membrane. (Adapted from: Ferguson and Storm, 2004)

suggest that AC can be used as the integration site of converging signalling sources. The interaction of multiple signalling sources can synergistically activate or inhibit AC. Figure 2.6 shows a model of how the complex of G-protein coupled receptor, AC and NMDAR is believed to interact.

AC activation affects the levels of cAMP which can initiate the processes required for synaptic plasticity. The effects of activating the cAMP signal-transduction pathway range from transient local synaptic changes, due to changes in kinase activity, to long-lasting consequences through its interaction with the transcription factors that affect the rate of expression of specific genes. The mode of interaction between the converging neurotransmitter and neuromodulator signals can be synergistic implementing a requirement for mutual activation to initiate late-phase LTP/LTD.

In this section we have argued that this requirement for mutual signalling may reflect a key requirement for *allocating cellular resources* for the encoding of memories that carry behavioural significance. The converging signals could be mediating a signal from motivational centres indicating to the target memory system the significance of a memory. The target memory system is then able to interpret the signals by implementing integration rules at the cellular level that dictate the requirements for allocating long-term memory encoding.

2.5.3 Spike timing dependent plasticity (STDP)

The plasticity experiments we have presented up to now use the typical plasticity inducing protocol of a high-frequency (HF) or low-frequency (LF) stimulation on a pathway to induce LTP or LTD. The train of HF or LF spikes produced drive the postsynaptic neurons' firing rate thus creating a causal relationship between presynaptic and postsynaptic activity. This correlated activity drives synaptic plasticity in a manner reflecting Hebb's postulate. In this section we review a form of plasticity that responds to the timing between individual spike pairs and is believed to be the basic unit behind how synapses respond to a causal relationship between presynaptic and postsynaptic stimulation to determine the sign and magnitude of plasticity.

First indications of a timing relationship between presynaptic and postsynaptic activity came from associative LTP experiments on anaesthetized rats' by pairing stimulation of two converging pathways to the DG (Levy and Steward, 1983). One pathway is sparsely connected and stimulating it gives only a weak response to the DG neurons with no measurable potentiation. The other pathway is well connected with excitatory synapses and thus forms a strong pathway to DG. By concurrent activation of these two pathways it was shown that stimulation of the strong pathway would enable the weak pathway to exhibit plasticity. The order of activation determined the sign of plasticity and there was a time window of around 20ms for the two pathways to interact. By first stimulating the weak and then the strong pathway, within the given time window, the weak pathway was potentiated. When the weak pathway stimulation preceded the strong by more than 200ms then the weak pathway was depressed instead of potentiated.

Further, a time window of 100ms for the interaction of presynaptic and postsynaptic activity was shown using burst activity between two neurons on guinea pig hippocampal slices. LTP was expressed after repeatedly pairing presynaptic pulses followed by postsynaptic depolarizing current pulses delivered using a microelectrode (Gustafsson et al., 1987). The magnitude of LTP was shown to be inversely related to the time difference between the presynaptic and postsynaptic stimulation events.

The time window of interaction was later studied for LTD using LFS presynaptically paired with single postsynaptic depolarizations. The results showed that associative LTD was expressed when presynaptic activity follows postsynaptic depolarization within a circumscribed time window and that LTD may even be produced on synapses where LTP had been previously expressed (Debanne et al., 1994).

The causal relationship was better shown using single presynaptic and postsynaptic stimulations at cortical pyramidal cell synapses. These experiments showed that potentiation or depression depended on the relative timing between the pairing of a presynaptic and a postsynaptic spike. Maximal modification was achieved at close to zero time interval with the amount declining up to a separation of 20ms after which it was not detected

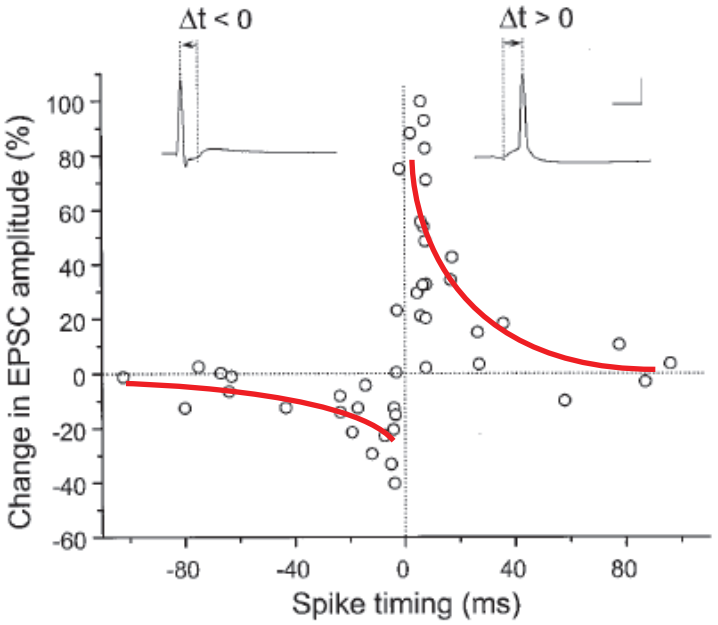


FIGURE 2.7: STDP between pairs of hippocampal pyramidal neurons showing a critical window for the induction of synaptic potentiation and depression. Each circle represents the percentage change in the EPSC amplitude at 2030 min after the repetitive correlated spiking (60 pulses at 1 Hz) was plotted against the spike timing. Spike timing was defined by the time interval (Dt) between the onset of the EPSP and the peak of the postsynaptic action potential during each cycle of repetitive stimulation, as illustrated by the traces above (Bi and Poo, 1998). Points on LTP and LTD phase can be approximated by an negative exponential function each. Figure modified from (Bi and Poo, 1998)

(Markram et al., 1997). Positive time differences induced LTP and negative induced LTD, the sign of the change in synaptic strength is a function of the relative timing (Δt) of single presynaptic spikes paired with a postsynaptic action potential.

Detailed studies of this timing dependence followed and the phenomenon was reproduced in the hippocampus (Debanne et al., 1998, Bi and Poo, 1998). The relation between magnitude and direction is shown on Figure 2.7, revealing an antisymmetric curve where the time windows for LTD and LTP are not equal (Bi and Poo, 1998). The typical asymmetric plasticity curve shown is not universal and different ones may be obtained in other brain areas, for example in the cerebellum the relative spike timings express the plasticity of the reverse sign.

Using the above findings one cannot directly interpret the events that occur during a typical LTP tetanic stimulation protocol as sequence of spike pairs. The effect of multiple spike pairings is not a simple superposition of the effects of each spike pair. Using spike triplets of presynaptic (π) and postsynaptic spikes (p) in a sequence as p-π-p with interspike intervals less than 20ms we need to understand how synapses interpret this sequence of induction stimuli. If the first pair alone p-π is repeatedly presented

it is known to induce depression (DEP) while the second pair π-p is known to induce potentiation (POT). Froemke and Dan (2002) for example, showed that the initial spike-pairs interact with further spike pairings in a manner that appears to exert a suppressive influence.

Wang, Gerkin, Nauen, and Bi (2005) extended to a quadrupled spike-pairing protocol on hippocampal slices to understand how a potentiating (POT) spike-pair interacts with a depression inducing spike-pair when presented close in time. By varying the time between the POT and DEP they found that when POT is followed by DEP (POT $\rightarrow$ DEP), a form of integration occurs that cancels the initial POT stimulus. On the other hand if DEP is followed by POT (DEP $\rightarrow$ POT) then a time difference larger than 70ms is required to achieve a neutralising integration between the two induction stimuli. If DEP $\rightarrow$ POT is presented within less that 70ms they found that POT dominates. The dominance of POT over DEP was also apparent in spike-triplets using p- π $-p$ while the ability for DEP to cancel the previous POT was also displayed using π-p-π. Thus, a form of *integration of the induction stimuli* appears to take place and at least for spike-pairs delivered after 70ms long intervals it appears that POT and DEP signals compete and no plasticity is expressed.

The mechanisms supporting the spike-timing dependence of plasticity have not yet been completely understood (Bi and Rubin, 2005), but it is known that STDP depends on NMDA receptor activation (Wigström and Gustafsson, 1985, Bi and Poo, 1998). As we saw previously NMDA receptors can act as coincidence detectors of presynaptic and postsynaptic activity and produce a calcium signal. The mechanism for detecting the postsynaptic activity is believed to involve the interaction of these channels with the potentials backpropagating through the dendrites back to the synapses that where stimulated, leading to increases in postsynaptic Ca^{2+} (Magee and Johnston, 1997).

Further, there is evidence to support a strong role of neuromodulation on the direction and the gating of STDP. Neuromodulators can act to decrease or increase the threshold for plasticity while there are examples where different neuromodulators cooperate (see Pawlak et al., 2010, for a review). The aforementioned theme of the synergistic action of AC enzyme and neurotransmitter stimulation is also manifested in STDP, but the findings of Seol et al. (2007) support a more radical role of neuromodulators in the pyramidal neurons of the rodent visual cortex. Activating β-adrenergic receptors coupled to AC during spike-pairing resulted always in LTP regardless of the order of stimulation pairing while activating receptors coupled to the phospholipase cascade (PLC) always induced LTD regardless of the stimulus pairing. With conjoint activation of both cascades the normal STDP window appeared with bidirectional plasticity set by the ordering of the stimulus pairs. Thus, the cAMP pathway is linked to LTP while the PLC to LTD opening the possibility that external reward-reinforcement systems dictate the direction of plasticity. Similar findings for STDP have been reported in insect olfactory learning, where Cassenaer and Laurent (2012) showed that π-p pairing followed by

the local delivery of a DA can specify the synapses that will undergo plasticity between MB neurons and their targets in the locust.

Overall, the relationship between spike-timing and plasticity could be used to explain the stimulation frequency dependence of plasticity induced under HFS or LFS protocols in terms of multiple spike pairs interacting. Models have been built that attempt to explain how spike-to-spike interactions could give rise to rate based Hebbian plasticity rules believed to describe what has been observed by tetanic stimulation protocols of pathways (see Cooper, 2010, Morrison et al., 2008). Having narrowed plasticity to spike-pairings leaves us to wonder whether spikes or rates are the relevant neural language for the formation of memories.

2.5.4 Induction history dependence of LTP/LTD

The capacity for plasticity is not constant under a fixed protocol of induction stimuli. Access to each of the previously described plasticity states depends on the history of induction stimuli as well on the current state of a synapse. For example, experimental evidence shows that the induction of e-LTP in hippocampal synapses occludes further e-LTP induction for up to 4 hours until the initially induced e-LTP has converted to l-LTP (Frey et al., 1995). During this interval only short-term potentiation can be induced that decays back to the baseline of the previously induced e-LTP. Such evidence may be interpreted by assuming that the intracellular cascades necessary for the maintenance of e-LTP have to become available again to generate new e-LTP (Frey et al., 1995). On the other hand, it may be that additional LTP requires the growth of new inactive synaptic buttons which can become potentiated after further LTP induction. Thus, single synapses may have access to a limited number strength states and the ability to further potentiate relies on the accumulation of synapses between a connection of two neurons.

Another line of experiments suggests that the activity threshold required for the induction of LTP and LTD changes in response to prior activity. For example, priming synapses with weak activity on the Schaffer collateral could inhibit later induction of e-LTP for at least 30 minutes. This inhibition by LFS priming could be overcome by administering stronger tetanic stimulation and thus it appears that priming had raised the threshold of the requirements for the induction of e-LTP (Huang et al., 1992b). In agreement with Huang, Colino, Selig, and Malenka, short trains of 5-30Hz *priming* stimulation that are subthreshold for inducing LTP can still facilitate subsequent LTD induction in the DG and in CA1 area hippocampal areas (Christie and Abraham, 1992, Wexler and Stanton, 1993).

It has also been demonstrated that priming with LFS may modify distinct phases of LTP. Using cellular electrophysiological recording methods in mouse hippocampal slices, LFS

before HFS did not modify e-LTP, but it selectively inhibited the expression of l-LTP. The inhibition of l-LTP expression was NMDA receptor dependent and inhibitors of protein phospatase1 (PP1) and phosphatase 2A (PP2A) abolished the effects of LFS on l-LTP (Woo and Nguyen, 2002). Thus, priming by LFS suppresses the expression of l-LTP by activating PP1 and PP2A, but these have no effect on e-LTP. PP1 is known to be a CREB-inactivating phospatase (Genoux et al., 2002) and, as we have seen previously, CREB regulates the transcription rate of genes required for the expression of long-term plasticity, see Figure 2.9. This priming phenomenon is not symmetric however, because priming with HFS appeared to facilitate subsequent induction of LTD instead of inhibiting it (Dudek and Bear, 1992).

Taken together the data supports the idea that the threshold for synaptic plasticity is not static and can vary dynamically according to recent history of synaptic activity. These effects are synapse specific, require the activation of NDMA receptors and can alter the balance for the induction threshold of LTP against LTD and also affect the transition to l-LTP by regulating transcription factor activity.

Phenomena showing the capacity for plasticity to depend on the history of synaptic activity or neuromodulatory activity are termed *metaplasticity* (Abraham and Bear, 1996). This is a broad term that refers to the facility to induce LTP/LTD, it can be thought of as the plasticity of synaptic plasticity. We have already presented evidence that the threshold for LTP/LTD induction is affected by neuromodulators. Since neuromodulators in general can lead to a decrease of the threshold for LTP induction then they must also exert influence on metaplasticiy (see Abraham and Tate, 1997). From the evidence we have discussed, a cellular mechanism of a common mode of action for the history dependence of plasticity and neuromodulation has not emerged. It is difficult to assume that the mechanisms of the two do not interact. We anticipate that metaplasticity due to neurotransmitter activity and other processes, such as neuromodulation, jointly set the threshold for plasticity.

2.6 The spacing effect for the formation of long-term memories

Learning episodes that are repeated appropriately spaced in time appear to be most effective in producing long-term memory than if the same number of learning episodes occurred close in time. In this thesis we have already come across this phenomenon in the behavioural experiments of the psychologist Herman Ebbinghaus (1885) who was the first to experimentally discover the that repeated study with intervening rest periods are needed for LTM retention. The phenomenon of a *spacing requirement* has been previously discussed in the context of the experiments on the conditioned reflex of

Aplysia, as appropriately spaced repeated stimulation was required for inducing long-term facilitation and the same was true for olfactory learning in Drosophila. These model animals can be valuable in the search of the cellular mechanism underlying the spacing requirement because the learned behaviour has been attributed to plasticity at certain neural circuits. Therefore, this behavioural spacing requirement finds its correlate at the cellular level. By studying the biomolecular interactions engaged to express long-term plasticity we may come to understand the mechanism that give rise to the spacing requirement and the factors that influence it.

One crucial signalling pathway for LTF in Aplysia is the MAPK (Sharma et al., 2003). MAPK activation follows 45 minutes after a single bout (tail shock) of a sensitization protocol and if a second stimulation is delivered during MAPK activation then LTF can be induced with only two stimuli repetitions instead of the classic spaced protocol of four repetitions delivered 15 minutes apart (Philips et al., 2007). Indeed, the stimulation protocol appears to be less specific to the actual pattern of stimulation, if three massed training sessions are followed by a single one spaced at 45 minutes then again LTF was expressed while reversing the order between massed and spaced stimulation did not affect the results. These results suggest that synapses do not rely on counting the number of repetitions to induce LTF but rather training stimuli have to be timed according to dynamic process that is initiated after the first stimulation.

The fact that MAPK peaks at the optimal time for repetition (45 minutes) does not imply that memory repetition interacts with MAPK per se. It could be that MAPK activation is synchronized with another process that enhances the effect of memory repetition so that it exceeds the cellular signal threshold required for initiating LTF. Indeed, in physiological conditions the gene-transcription supressor CREB-2 is inhibited when activated MAPK translocates to the nucleus (Bartsch et al., 1995). The role of CREB-2 in LTF has already been discussed in the section on Aplysia learning, the spacing requirement for long-term facilitation can be abolished by inhibiting the suppressor of gene transcription CREB-2 (Bartsch et al., 1995). It remains unknown however, how CREB-2-mediated disinhibition of transcription interacts with memory repetition.

Activation of the MAPK pathway is also observed after experiments on rat's spatial learning or in contextual fear conditioning, but also in electrophysiology experiments after the induction of LTP (see Waltereit and Weller, 2003b, Sweatt, 2004). Accordingly, Wu, Deisseroth, and Tsien 2001 report that persistent activation of MAPK was achieved by spaced stimulation and this resulted in the growth of new dendritic filopodia in cultured hippocampal neurons. Given that it is known that spaced stimulation can induce l-LTP, this growth of dendritic protrusions after spaced stimulation supports the hypothesis of structural changes mediating long-term memory.

Signalling the nucleus for CREB activation is also crucial for the formation of LTM in Drosophila. Transgenic flies with a heat-sock induced form of a CREB *repressor*

abolished the 7 day LTM normally produced by a spaced training protocol (Yin et al., 1995). On the other hand, transgenic flies expressing a heat-sock inducible *activator* of CREB abolished the spacing requirement and in fact a single training session was sufficient to produce LTM. The above results suggest that the spacing requirement for long-term memory in Drosophila is honoured somewhere upstream of MAPK activation by gene transcription networks and suggest that the level of active CREB is important in determining the training requirements for the formation of long-term memory.

Research in aversive long-term olfactory memory of Drosophila by Pagani et al. (2009) reported that a cycle of activation and inactivation of MAPK was required to induce long-term olfactory memory. This is in line with the idea that transcription networks found upstream of MAPK control the processes of LTM. A surprising effect of spacing training events was that an initial training event would activate MAPK and a second training delivered at particularly spaced time would inactivate MAPK. This spacing was controlled by the state of activity of *corkscrew*, the fly homolog of SHP2 protein tyrosine phospatase.

SHP2 is a phosporylation detector of tyrosine and is recruited to many receptor tyrosine kinases upon activation while it is generally a positive regulator of Ras/MAPK signaling, see Figure 2.9. Thus the attachment of SHP2 to a receptor after activation controls the length of time the MAPK pathway remains active (Pagani et al., 2009). These mechanisms could be specific to the Droshophila MB responsible for olfactory memory but such mechanisms could be evolutionary conserved.

Pagani, Oishi, Gelb, and Zhong (2009) showed that a mutant fly that overexpressed an active form of corkscrew extended the minimum required repetition interval to produce LTM from 15 to 40 minutes. On the other hand, significantly shorter inter-training intervals from 15 to 2.5 minutes were achieved if mutant flies overexpressed a wild-type corkscrew.

What function the spacing effect is serving is yet not clear. It may be assumed that impressions or memories that are presented spaced in time reflect a constant property of the environment an animal has to adapt to, and thus it is beneficial to remember it in a manner that guides future long-term behaviour. Nevertheless, if the spacing effect serves as yet another gatekeeper to long-term memory we expect the spacing intervals to be adaptive to suit the purpose of the particular memory system or situation.

2.7 Summary

In this chapter we presented the primary experimental models for synaptically stored memory across three different animal models of learning. In Aplysia there is a bidirectional persistent form of plasticity, which can be evoked in response to behavioural

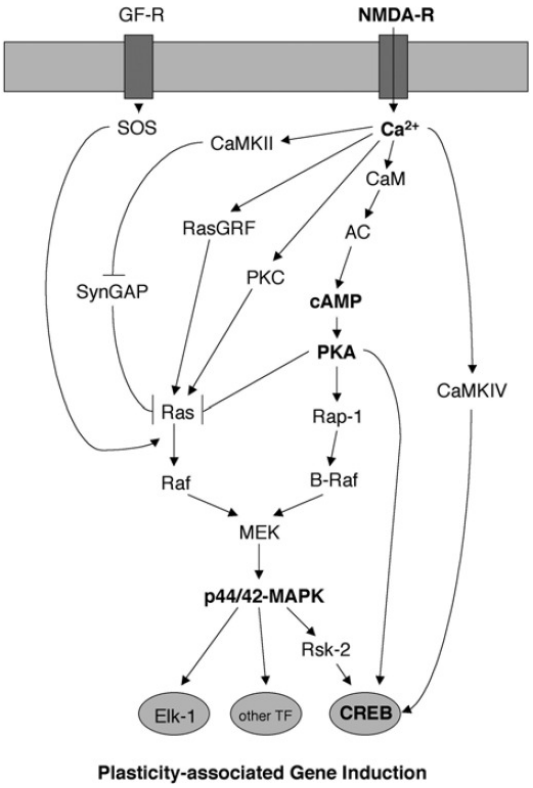


FIGURE 2.8: Signaling pathways activated by calcium influx through the NMDA receptor to the transcriptional activation of plasticity-associated genes. The key molecular events during synaptic plasticity are highlighted. cAMP and the MAPK pathway are simultaneously activated and it is believed that competition for the activation of the two pathways occurs since it is known that PKA inactivates Ras/MAPK (Waltereit and Weller, 2003a). Unidentified transcription (TF) factors in addition to CREB are believed to contribute to MAPK-regulated transcription of plasticity-associated genes. Dysregulation of the MAPK signal transduction mechanism has been linked to learning disabilities (Sweatt, 2004). (Source: Waltereit and Weller, 2003a).

learning. We also reviewed odour discrimination learning in Drosophila and the respective mechanism believed to underlie motivated long-term learning. The mammalian hippocampus is able to express LTP and LTD in response to persistent stimulation of the PP and Schaffer collateral pathways. This bidirectional form of synaptic plasticity is the primary model for declarative memory and learning. Its associative nature make it a good candidate for modifying the effective connection structure of neural assemblies as the Hebbian hypothesis predicts.

These phenomena also exhibit differences in lifetime, and they can be separated into short-term, early and late long-term. The short-term processes are protein kinase independent and do not require protein synthesis. The next phase, early long-term plasticity requires kinase activity to phoshorylate target proteins and enhance synaptic efficacy but does not depend on protein-synthesis. Stable long-term plasticity requires protein synthesis and later the transcription of genes. This whole framework reflects a concept of progressive stabilization of a memory.

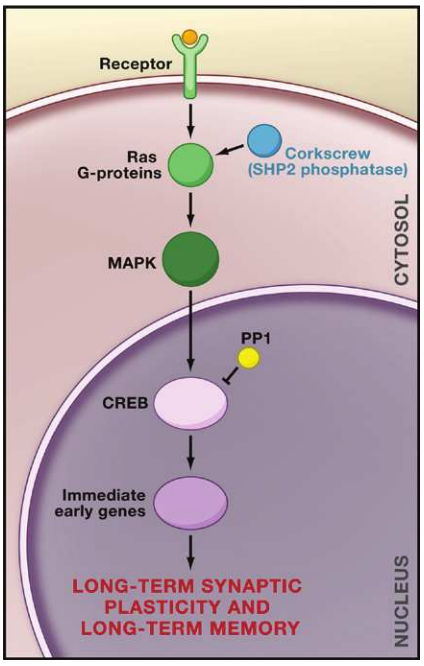


FIGURE 2.9: Candidate nuclear signalling elements recruited during the induction of long-term synaptic plasticity and memory. G-protein coupled receptor signaling promotes MAPK activation , whose activity can support the activation of the CREB transcription factor. CREB activated transcription leads to the expression of immediate early genes and the production of plasticity related proteins for long-term synaptic plasticity and long-term memory formation. Massed training sessions can recruit inhibitory phosphatases, including the CREB-inactivating protein phosphatase 1 (PP1), which inhibits long-term synaptic plasticity and memory. The spacing effect for long term memory is regulated by *corkscrew* (blue), the Drosophila homolog of the MAPK activating protein SHP2. The optimal spacing interval for induction of long-term memory can be modulated by the levels of active SHP2. (Source: Philips and Carew, 2009).

The transcription of genes requires the activation of the CREB transcription factor which is implicated with the formation of long-term memory in all three of the animal models examined (Deisseroth et al., 1996). The conditions and mechanisms that lead to its activation, if understood, will elucidate the signalling requirements to initiate long-term memory and thus memory allocation. CREB activation comes in response to either cAMP or Ca^{2+} signalling. However, CREB activation in response to Ca^{2+} is rapid and the brief Ca^{2+} signal is due to membrane depolarizations under synaptic activity. Thus, cAMP signalling may be a regulator of CREB activity over longer timescales than rapid Ca^{2+} signals.

Second order properties were also reviewed according to which the capacity for synaptic plasticity appears to depend on the history of synaptic activity or neuromodulatory activity. Across the three animal models we found that acquisition of stable memory traces required signalling by modulatory neurons or networks. In the hippocampus, there is evidence that DA gates long-term memory and behavioural experiments show a prolongation of memory retention associated with activation of the VTA. Therefore, the

VTA appears to give a modulatory or instructive signal for the formation of long-term memories. Evidence from an experiment showing that direct control of dopaminergic neurons in Drosophila can program aversive olfactory memory on demand suggest that there are common principles in the organization of memory between phyla.

The facility to induce plasticity bidirectionally can also be modified due to prior activity, a phenomenon known as metaplasticity. Of particular interest is the fact that priming stimulation can affect distinct phases of LTP/LTD, where LFS could affect the induction of l-LTP but not e-LTP or HFS facilitates the induction of LTD. Such data indicate that induction of LTD and LTP can be differentially modulated. Also, the facility to induce late phase plasticity is affected by prior activity and neuromodulators acting presynaptically or postsynaptically. The threshold for inducing late phases of plasticity is partly modulated by upregulating phoshatase activity (PP1) which is known to inhibit the transcription factor of plasticity related proteins CREB.

Presynaptic neuromodulation was found in the simple gill-withdrawal reflex circuit in Aplysia, but such organization is not seen in vertebrates and thus neuromodulation initially affects the post synaptic site. Nevertheless, in both cases neuromodulators are known to intracellularly interact with afferent spike activity to gate plasticity and form behaviourally relevant memories. Two well known examples involve classical conditioning in Drosophila and Aplysia.

It has been proposed that the conditioned stimuli (CSs) and unconditioned stimuli (USs) converge at the level of the transmembrane enzyme AC, one is signalled via Ca^{2+} and calmodulin and the other via G-protein coupled receptors (Dudai et al., 1988). The levels of Ca^{2+} are increased by spike activity on the CS pathway and the G-proteins are stimulated by the neuromodulatory inputs of the US stimuli. Yovell, Kandel, Dudai, and Abrams (1992) conducted a study between the three models, Aplysia, Drosophila and the rat to determine the sensitivity of AC to Ca^{2+} concentration, knowing that AC can also be regulated by G-protein coupled receptors. They report that sensitivity of AC to Ca^{2+} concentration was dependent on the concentration of calmodulin and Mg^{2+}. Increasing the concentration of Mg^{2+} raised the threshold required for the concentration of Ca^{2+} to activate AC. Thus, a common mechanism appears to be responsible for detecting stimulus convergence and its effectiveness depends on the concentration of calmodulin which can be under the control of cell-wide processes.

The current experimental evidence suggest that metaplasticity defines the conditions for expressing long-lasting changes in synaptic efficacy in response to prior activity and this faculty can be modulated by mechanisms operating at the level of molecular interactions from receptors to the neuron's nucleus. The effects of metaplasticity on models of memory is not yet clear and we shall explore it in future chapters where we examine the effect that a synapse model of progressive stabilization has on the capacity of memory.

Chapter 3

Mathematical Methods

In subsequent chapters we are going to analyse the dynamics of memory and synapse models governed by the storage of random uncorrelated memories. Consequently a random element is inherent in the way we examine memory systems and thus our mathematical analysis requires the use of probability. Systems which develop in time or space in accordance with probabilistic laws are dealt in mathematics by the theory of stochastic processes.

In this chapter we will introduce elements from the theory of stochastic processes relevant to the analytical methods used in further chapters. The methods described here can be found in Cox and Miller (1977), Grimmett and Stirzaker (1992) and Gillespie (1992). We begin by a short clarification on notation used in this thesis and elementary measures of mean and variance for random processes. Further, we introduce the general framework of Markov chains that will enable us to formulate and analyse state-based random processes. The analysis of these processes is greatly simplified with the use of generating functions by enabling us to represent and manipulate long sequences of numbers rather simply. For this reason a the basics of generating functions are given in Section 3.3. Their use is then demonstrated on the analysis of a fundamental random process known as the *simple random walk*. This random process along with the Poisson process and its generalization the renewal process which is introduced next, form a basis on which we will elaborate to develop the analysis of models in subsequent chapters. Finally, we present methods that will be required for the analysis of Markov chains that require the solution of a linear recurrence relations and the factorization of matrices.

3.1 Basic definitions

Probabilities will be denoted with the bold capital letter and angled brackets $\mathbb{P}[X]$. When the bold type $\mathbb{X}$ is not accompanied by angle brackets then this should be interpreted as referring to a matrix named $\mathbb{P}$. We denote an $n \times n$ matrix of elements p_{ji}

found at the jth row and ith column as $\mathbb{P} = (p_{ji})$. It is common to examine the likelihood of combinations of events, the joint probability of both events A and B occurring is denoted by $\mathbb{P}[A \cap B]$. Further, we may wish to examine the conditional probability of event A occurring given that event B has occurred, without implying causality. This is denoted as $\mathbb{P}[A|B]$ and it is related to joint probability by

$$\mathbb{P}[A \cap B] = \mathbb{P}[A|B]\mathbb{P}[B] = \mathbb{P}[B|A]\mathbb{P}[A]$$

We can represent random quantities by variables denoted in capital letters such as X,Y and Z whose value is subject to chance. In more formal terms, letting the set of all elementary outcomes of our process be Ω we may define a random variable through a function $X : \Omega \to \mathbb{R}$ mapping elementary events $\omega \in \Omega$ to the real line. These mappings describe the uncertainty in value when sampling this variable of a system by defining the probability as the frequency with which X takes the value ω when sampled multiple times.

When the values taken by random variables are discrete, i.e the variable take values from a countable subset $\{x_1, x_2, \cdots, x_n\} \in \mathbb{R}$, we define the *probability mass function* $g : \mathbb{R} \to [0, 1]$ given by $g(x) = \mathbb{P}[X = x]$ where lower case x represents a particular value. When considering continuous or non-denumerable random variables then variable X may be a random real number. In this case we cannot distribute probability to each of infinitely many values lying in the real number line and thus we resort to define a probability density that $\mathbb{P}[x]\Delta x = \mathbb{P}[X \in \{x, x + \Delta x\}]$. Through a limit argument by taking $\Delta x \to 0$ we are led to the definition of the probability distribution :

$$F(x) = \int_{-\infty}^{x} f(u)du \qquad x \in \mathbb{R}, \tag{3.1}$$

where $f(u)$ describes the *probability density function* (PDF) of the random variable X, which can also be written as $f_X(u)$ for clarification. The PDF has $\int_{-\infty}^{\infty} f(u)du = 1$ and it can be obtained from $f(u) = F'(u)$. The expectation of a continuous random variable is:

$$\mathrm{E}[X] =< X >= \int_{x=-\infty}^{+\infty} xf(x)dx \tag{3.2}$$

and for a discrete random variable we change the integral with sum over the region of definition of x. From here on we shall not differentiate between the continuous and discrete case and refer to $f(x)$ as the distribution of X, taking into account that integration in the discrete case is a sum operation.

Following the definition of expectation we generalize to the definition of *moments* of a distribution. The **nth moment** m_n of X is defined to be $m_n = \mathrm{E}[X^k]$ where n is a positive integer. Thus, by this definition the expectation $\mathrm{E}[X]$ defined above is also the

first moment of X and in general we may define the n^{th} *moment of* X as :

$$\mathrm{E}[X^n] = \int_{-\infty}^{\infty} x^n f(x) dx \tag{3.3}$$

Depending on how fast $f(x) \to 0$ as $|x| \to \infty$ these higher moments may be insignificant or non-existent. If all moments of X exist, then they will be required in determining the average of function of the random variable.

Further, we may be interested in calculating the spread of a random variable around its mean value. This is given by the *variance* of X defined to be :

$$\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] = \int_{-\infty}^{\infty} (x - \mathrm{E}[X])^2 f(x) dx \tag{3.4}$$

and by expansion we see how it relates to the first and second moments:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x^2 - 2\mathrm{E}[X]x + \mathrm{E}[X]^2) f(x) dx \tag{3.5}$$

$$= \mathrm{E}[X^2] - 2\mathrm{E}[X]\mathbb{E}[X] + \mathrm{E}[X^2] = \mathrm{E} \tag{3.6}$$

$$= [X^2] - \mathrm{E}[X]^2 \tag{3.7}$$

Thus, the variance of X can be expressed rather simply in terms of the first and second moments of X and we can see it is always positive with $\mathrm{E}[X^2] \geq \mathrm{E}[X]^2$. Naturally, it follows that when $\mathrm{E}[X^2] = \mathrm{E}[X]^2$ then X is not random but assumes some constant value. The square root of the variance of X is called the *standard deviation* of X denoted by σ and it measures the expected difference between the sample values or the *fluctuations* of X about the mean of X.

Finally, we have that the expectation of r identical independent random variables being equal to the sum of their expectations $\mathrm{E}[\sum_i^r X_i] = r\mathrm{E}[X]$ and for independent X we have $\mathrm{Var}(rX) = r^2\mathrm{Var}(X)$.

3.2 Markov chains

In many situations we may be dealing with systems that can exist in one of a countable set of states at any particular point in time. At each discrete time-step some process drives a change in the system that is observed as a movement from one state to another. Modelling such systems as Markov chains allows a rigorous treatment analysing the relationships and the properties of the state transitions of a system. If the transitions between states are probabilistic then the state of the system at time n can be denoted by a random variable X_n representing the potential state of the system taking values from say $X_n \in \{\alpha_0, \alpha_1 \cdots \alpha_i\}$. The values α_0 of the random variable X do imply that X is some numerical quantity of our system but could simply represent numerical labels of

system states. If the next state of a process only depends on its current state then we say that the system has the *Markov property*. According to this property the transition probabilities to the next state are not altered from prior state transitions but only depend on the current state. We may therefore define a probabilistic transition matrix from the set of transition probabilities that describe the process. Depending on whether the transition matrix is applied from left or right we write $\mathbb{M} = (p_{ji})$ or $\mathbb{M} = (p_{ij})$ respectively. The probability p_{ij} of moving from state i to state j in one timestep is written :

$$p_{ji}(n+1, n) = \mathbb{P}[X_{n+1} = j | X_n = i]. \tag{3.8}$$

For $n - step$ transition probabilities we write $p_{ji}(m+n, m) = \mathbb{P}(X_{m+n} = j | X_m = i)$ and thus the matrix is $\mathbb{M}(m, m+n) = (p_{ji}(m+n, m))$. This is a stochastic matrix that describes the changes in probability density at each timestep and assumes that the overall probability of all states being occupied is conserved. For a **right** transition matrix each row sums to one $\sum_j p_{ij} = 1$, while for a **left** each column sum is unity $\sum_j p_{ji} = 1$. In this thesis we adopt convention of using left transition matrices. If we let the initial mass function be $f_i^{(n)} = \mathbb{P}[X_n = i]$ and write $\boldsymbol{f}^{(n)}$ for a column vector with entries $(\boldsymbol{f}^{(n)} : i \in S)$ we have that $\boldsymbol{f}^{(n)} = \mathbb{M}^n \boldsymbol{f}^{(0)}$. It then naturally follows that the random evolution of the chain is determined by the transition matrix $\mathbb{M}$ and the initial mass function $\boldsymbol{f}^{(0)}$.

The continuous-time Markov chains are described in similar terms, but instead of writing the n-step probabilities in a matrix $\mathbb{M}$ we use a *generator* matrix $\mathbb{G}$ that contains transition rates (derivatives of probabilities) g_{ij}. The theory behind moving from discrete to continuous time is related to the theory supporting the formulation of a Poisson process, which we look at a later Section 3.5). In either case the necessary elements to formulate a Markov chain in continuous time requires the examination of events in the limit of small timesteps. In this case we consider the transition probabilities during small time intervals $\Delta t \to 0$ assuming that the probability of more than two transitions in the interval $(t, t + \Delta t)$ is very small. We take the probability of transition from $j \to i$ in Δt as approximately linear :

$$p_{ij}(\Delta t) \simeq g_{ij} \Delta t, \tag{3.9}$$

for some constant g_{ij} when $i \neq j$. Accordingly we write the transition probability that no change occurred during time Δt:

$$p_{ii}(\Delta t) \simeq 1 + g_{ii} \Delta t, \tag{3.10}$$

where here g_{ii} has to be $g_{ii} \leq 0$. We write the constants g_{ij} in matrix form as $\mathbb{G}$ and from the above equation notice that this can be written as $\mathbb{G} = (\mathbb{M} - \mathbb{I})$. The evolution

of probabilities in small increments is :

$$p_{ij}(t+\Delta t) = \sum_k p_{kj}(t)p_{ik}(\Delta t) \tag{3.11}$$

using (3.9) we rewrite as :

$$p_{ij}(t+\Delta t) = p_{ij}(1+g_{ii}\Delta t) + \sum_{k\neq i} p_{kj}(t)g_{ik}\Delta t \tag{3.12}$$

$$= p_{ij}(t) + \Delta t \sum_k p_{kj}(t)g_{ik} \tag{3.13}$$

and therefore we obtain a derivative of p_{ij} as $\Delta t \to 0$:

$$\frac{1}{\Delta t}\left(p_{ij}(t+\Delta t) - p_{ij}(t)\right) \simeq \sum_k p_{kj}(t)g_{ik}, \tag{3.14}$$

therefore we find the derivative of $p_{ij}(t)$. We define the matrix $\mathbb{F}(t)$ containing the transition probabilities $p_{ij}(t)$ and then write the above as :

$$\mathbb{F}'(t) = \mathbb{F}(t)\mathbb{G} \tag{3.15}$$

this is known as the *forward equation* with solution :

$$\mathbb{F}(t) = \exp\left(\mathbb{G}t\right) \tag{3.16}$$

As in discrete time, the elements of this transition matrix, $p_{ij}(t)$ represent the probability of a transition from state j to state i with the addition of the dimension of time t. Thus, given an initial distribution among states as a row vector $\boldsymbol{f}^0$ the distribution at time t is :

$$\boldsymbol{f}(t) = \mathbb{F}(t)\boldsymbol{f}^0. \tag{3.17}$$

In a later section (3.8) we look at algebraic properties of $\mathbb{G}$ that allow as to manipulate the above expression and obtain the evolution of the stochastic system.

Equilibrium distributions: In the long-term we expect a physical system to approach an equilibrium state. Thus after applying a transition matrix n times on on a state vector the state will eventually reach an equilibrium distribution $\boldsymbol{\pi}$ regardless of what the initial state was at $n = 0$. The equilibrium distribution $\boldsymbol{\pi}$ would then have the property :

$$\mathbb{M}\boldsymbol{\pi} = \boldsymbol{\pi} \tag{3.18}$$

making it a *stationary distribution*. If a stationary distribution is given as the initial distribution $\boldsymbol{f^{(0)}} = \boldsymbol{\pi}$ then it will not be modified for all iterations over n, $\boldsymbol{f^{(n)}} = \boldsymbol{\pi}$. Consequently, the limiting probability of finding a system in state k is independent of the initial distribution. The above relationship can also be seen as $\boldsymbol{\pi}$ being an eigenvector

of the $\mathbb{M}$ with eigenvalue $\lambda = 1$. Consequently systems like this may be analysed by studying the algebraic properties of $\mathbb{M}$. There are cases where unique equilibrium distributions do not exist, but in broad terms for non-negative square matrices the *Frobenius-Perron* theorem states that conditions for equilibrium apply and such matrices have a maximal non-negative eigenvalue that corresponds to a non-negative eigenvector (see Cox and Miller, 1977). Thus we may interpret $\boldsymbol{\pi}$ in Equation (3.18) as an eigenvector with eigenvalue of unity. From here on we assume that the relevant biological systems of study will exhibit an equilibrium point and the structure of the matrices satisfy the above conditions.

3.3 Generating functions

Our analysis will frequently evolve evaluating probabilities that are defined by recurrence relations. These recurrence relations, if expanded, result in sums of sequences of terms usually involving other probabilities.

A powerful way to represent a sequence of numbers $\alpha_0, \alpha_1, \alpha_2 ... \alpha_n$ is to attach them as coefficients to a power series. Once formulated as a power series we wish to describe entire sequences by simple formulas which will allow us to manipulate, analyse or combine them with other series. Such transforms are suited to the study of sums of independent random variables which can be used to study random walks and branching processes. *Generating functions* empower us to do just that by working with the sums of power series whose coefficients are the terms we are interested in, for example :

$$G_\alpha(s) = \sum_{n \geq 0} \alpha_n s^n$$

By letting $s \to 1$ we can evaluate the above sum of coefficients since :

$$\lim_{s \to 1} G_\alpha(s) = \sum_{i=0}^{n} \alpha_i \tag{3.19}$$

We will use a generating function transforms to simplify the analysis towards obtaining the underlying probability density function.

The convolution of two sequences $\alpha = \{a_i : i \geq 0\}$ and $\beta = \{\beta_i : i \geq 0\}$ is the sequence

$$\gamma_n = \alpha_0 \beta_n + \alpha_1 \beta_{n-1} + ... + \alpha_n \beta_0$$

which is the convolution of the two written as $\gamma = \alpha * \beta$. Using generating functions we represent the convolved series as :

$$G_\gamma(s) = \sum_{k=0}^{n} \gamma_k s^k = \sum_{k=0}^{n} \left(\sum_{i=0}^{k} \alpha_i \beta_{n-i} \right) s^k \tag{3.20}$$

$$= \sum_{i=0}^{n} \alpha_i s^i \sum_{k=i}^{n} \beta_{k-i} s^{k-i} = G_\alpha(s) G_\beta(s) \tag{3.21}$$

We can now use generating functions to deal with sums of random variables that have a distribution defined by $f(i) = \mathbb{P}(X = i)$, writing :

$$G(s) = \mathrm{E}[s^X] = \sum_i s^i \mathbb{P}[X = i] = \sum_i s^i f(i) \tag{3.22}$$

to represent the probability distribution.

The ability to represent sums by generating functions can be used to obtain moments of the PDF. We have already seen in Section 3.1 that the mean of the random variable is given by 1st moment of the process. With generating functions the first derivative of $G(s)$ gives :

$$G'(s) = \sum_{n}^{\infty} n s^{n-1} f(n) \tag{3.23}$$

and thus $\mathrm{E}[X] = G'(1)$ where $G'(1)$ is to be interpreted as $\lim_{s \to 1} G(s)$. In general :

$$\mathbb{E}[X(X-1)\cdots(X-k+1)] = G^k(1) \tag{3.24}$$

gives the kth *factorial moment* of X. Accordingly the variance of X can be calculated :

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= G''(1) + G'(1) - G'(1)^2 \end{aligned} \tag{3.25}$$

When we are specifically interested in the moments we can work with the moment generating function (MGF). The moment generating function (in continuous time) is defined as a two-sided Laplace transform of the PDF of a random variable X as :

$$\hat{f}(s) = \int_{-\infty}^{+\infty} \exp(-sx) f(x) dx \tag{3.26}$$

$$= \int_{-\infty}^{+\infty} \left(1 + sx + \frac{s^2 x^2}{2!} + \cdots + \frac{s^n x^n}{n!} + \cdots \right) f(x) dx \tag{3.27}$$

$$= 1 + s m_1 + \frac{s^2 m_2}{2!} + \cdots + \frac{s^n m_n}{n!} \cdots , \tag{3.28}$$

where m_n denote the nth moment, so we may obtain :

$$\mathrm{E}[X] = \{\frac{d}{ds}\hat{f}(s)\}_{s=0} \tag{3.29}$$

3.4 A simple random walk

A random walk on the integers $\mathbb{Z}$ is stochastic process defined by random independent steps from an initial state $X_0 \in \mathbb{Z}$. The processes can be described as a sum X_n of independent identically distributed random steps ξ_i drawn from a step distribution F :

$$X_n = X_0 + \sum_{i=1}^{n} \xi_i \tag{3.30}$$

or a recurrence relation :

$$X_{n+1} = X_n + \xi_{n+1} \tag{3.31}$$

In the simple case the steps are of fixed size $\xi_{\in}\{+1, -1\}$ and occur with probability q and p respectively, with $q = 1 - p$. The above single dimensional definition can be extended to d-dimensional spaces $\mathbb{Z}^d$ but we will only be concerned with the single dimensional form in order to show how methods using generating functions can be applied to study this problem.

The simple random walk is spatially homogeneous:

$$\mathbb{P}(X_n = j|X_0 = \chi) = \mathbb{P}(X_n = j + b|X_0 = \chi + b)$$

and temporally homogeneous:

$$\mathbb{P}(X_n = j|X_0 = \chi) = \mathbb{P}(X_{n+m} = j + b|X_{0+m} = \chi)$$

thus the distribution of expected covered distance in n steps is not modified by the starting point but also for a given starting point the time of observation does not influence the process. This brings us to recognize the Markov property:

$$\mathbb{P}(X_{m+n} = j|X_0, X_1, X_2, ..., X_m) = \mathbb{P}(X_{m+n} = j|X_m) \text{ for } n \leq 0$$

The values of the random walk after the mth step are independent of the values before it and only depend on the value X_m. For it to reach the point k at the nth step the walk needs to make r_1 positive jumps, r_2 negative jumps and r_s zero jumps :

$$r_1 - r_2 = k, \text{ and } r_s = n - r_1 - r_2$$

The probability generating function (PGF) of each step ξ_r is:

$$\mathrm{E}[z^{\Xi_r}] = F(z) = pz + (1-p-q) + qz^{-1}. \tag{3.32}$$

The mean step size is given by $\mu = F'(1)$, giving $\mu = p - q$. The standard deviation is:

$$\sigma^2 = F''(1) + F'(1) - (F'(1))^2 = p + q - (p-q)^2 \tag{3.33}$$

Sums of independent random variables have the important property that their joint density function is a convolution of the PDF of each variable. Since the steps in a random walk are independent and given that generating functions simplify convolutions to products then we may write the PGF of the sum as an elegant product :

$$\mathrm{E}[s^{X_n}] = F(z)^n \tag{3.34}$$

We assume the walk starts from $X_0 = 0$ and thus $F_0(z) = 1$ and now we define a generating function :

$$G(z,s) = \sum_{n=0}^{\infty} s^n [F(z)]^n = \frac{1}{1 - sF(z)} \tag{3.35}$$

$$= \frac{z}{-spz^2 + z\{1 - s(1-p-q)\} - sq} \tag{3.36}$$

With $G(z,s)$ we have obtained a generating function of the pgfs $F(z)$. $G(z,s)$ contains all the information of the process, the coefficient of s^n is the $F(z)$ pgf, while the $s^n z^k$ contains the $\mathbb{P}[X_n = k]$ probability information. We may sum the steps through the $G(z,s)$ generating function, effectively summing $F(z)$, to obtain the well known results:

$$\mathrm{E}[X_n] = n\mu \text{ and } \mathrm{Var}(X_n) = n\sigma^2. \tag{3.37}$$

showing that the mean of n independent steps is the sum of the mean step size μ, while the variance grows linearly with n and thus the standard deviation grows as the $\sqrt{n}$. A classic applications of random walks is to represent the random motion of particles in some medium, the distance travelled in time t is characteristic of a the diffusion process.

3.5 Poisson process

In the problems we will be considering we require to model the random occurrence of event in time t. The Poisson process serves as mathematical model for various empirical phenomena where the inter-arrival X times are random events that can be formulated as a Markov process with discrete states in continuous time. The arrival times X_1 are exponentially distributed, and since all X_n are identical and independent then they are

all *exponentially distributed* with parameter ρ and PDF:

$$f_X(t) = \rho \exp(-\rho t). \tag{3.38}$$

Assuming $\rho > 0$, the mean time until the next event is :

$$\mathrm{E}[X] = \int_0^\infty t\rho \exp(-\rho t) dt = \frac{1}{\rho} \tag{3.39}$$

and the variance :

$$\mathrm{Var}(X) = (\mathrm{E}[X^2] - \mathrm{E}[X]^2) = \left(\frac{2}{\rho^2} - \frac{1}{\rho^2}\right) = \frac{1}{\rho^2} \tag{3.40}$$

For r independent events we have $\mathrm{E}[X_1 + \cdots + X_r] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_r] = r\mathrm{E}[X]$ and $\mathrm{Var}(X_1 + \cdots + X_r) = r^2\mathrm{Var}(X)$. The probability of arrival of j events by time t is :

$$p_j(t) = \frac{(\rho t)^j}{j!} e^{-\rho t} \tag{3.41}$$

with mean and variance $\rho = \mathrm{E}[X] = \mathrm{Var}[X]$.

3.6 The renewal process

In a Poisson process, the intervals between events are independently exponentially distributed. We want to generalize and consider processes that are renewed after an event occurs. These would still be independently and identically distributed events but they would be drawn from a given PDF $f(x)$.

We denote the count of events at time t as $N(t)$ and we let X_r denote the random variable of the time between the $r-1$ and the rth event having the PDF $f_r(x)$. If $f_r(x)$ remains fixed for all r then we have *ordinary renewal process* with $f_r(x) = f(x)$ and distributions $F_r(x) = F(x)$. It follows that the time of the rth event is a sum of random variables $T_r = \sum_{r=0}^{n} X_r$, and consequently $T_{r+1} = T_r + X_{r+1}$. Conditioning on the arrival time of the first event, we may write the expected number of events as :

$$\mathrm{E}[N(t)|X_1 = x] = \begin{cases} 0 & t < x \\ 1 + \mathrm{E}[N(t-x)] & t \geq x \text{ the process restarts} \end{cases} \tag{3.42}$$

since the first arrival occurs after time x. After each event arrival the process is identical and therefore restarts. Commonly with renewal processes we wish to examine the number of events in $(0, t]$, denoted $H(t) = \mathrm{E}[N(t)]$. The probability of r events occurring by time t is given by $\mathbb{P}[N(t) = r] = F_r(t) - F_{r+1}(t)$, and the $h(t)$ gives the probability density of events, which depend on $f(x)$ for small t. We formulate $h(t)$ the probability

density of events by conditioning on previous renewal events at some time $(t - t')$:

$$h(t) = f(t) + \int_0^t f(t')h(t - t')dt' \tag{3.43}$$

The first term on the RHS gives the probability that an event occurred near time t , and the second term considers the probability of an event over the remaining interval t' given that the process has been renewed at time $t - t'$.

A Laplace transform on the above expression will give the MGF for the whole process and at the same time simplify the expression by converting the convolution structure on the RHS to a product:

$$\hat{h}(s) = \hat{f}(s) + \hat{h}(s)\hat{f}(s) \tag{3.44}$$

that can be simplified to obtain :

$$\hat{h}(s)\frac{\hat{f}(s)}{1 - \hat{f}(s)} \tag{3.45}$$

and evaluated through inverse Laplace transform, given $\hat{f}(s)$ are known.

3.7 Solving linear recurrence relations

Linear recurrence relations have the general form :

$$x_n = \alpha_1 x_{n-1} + \alpha_2 x_{n-2} + \cdot + \alpha_k x_{n-k} + \beta_n \tag{3.46}$$

The β_n is a constant term making the recurrence relation non-homogeneous when $\beta_n \neq 0$. Here we will cover only the homogeneous case which is simpler and a generic method may be applied. The coefficients α_n above have been taken to be constant but they could depend on n and thus write $\alpha_n(n)$. We wish to obtain a closed-form expression that will give a way to obtain the nth term.

Here we will summarize two main methods used in this thesis, first the method of characteristic equation and then follows the method of generating functions.

3.7.1 The characteristic equation

The basic approach for solving a linear homogeneous recurrence relation of degree k is to look for solutions of the form $x_n = z^n$ where z is constant. Once we substitute in z^n we divide by z^{n-k} and obtain what is known as the *characteristic equation*. The solutions of this equation are called the characteristic roots of the recurrence relation. If the characteristic equation has k distinct roots $r_1, r_2, \cdots, r_k$, then the general solution of the recurrence is :

$$x_n = \alpha_1 r_1^n + \alpha_2 r_2^n \cdots \alpha_k r_k^n$$

where $n = 1, 2, \cdots$, and α_k are constants. Otherwise, if we have a root with multiplicity m then the solution contains $m, r^n, nr^n, n^2 r^n \cdots n^{m-1}$ Given the k boundary conditions, we can uniquely the α_k and determine the values of the sequence by solving k simultaneous equations.

For example, consider the following second order recurrence equation with $k = 2$:

$$\beta z^{n-1} + \alpha z^n + \beta z^{n+1} = \lambda z^n$$
$$\beta z^{n+1} + (\alpha - \lambda) z^n + \beta z^{n-1} = 0$$

the last step by dividing through by z^{n+1-k} because here we are taking the highest power to be $n + 1$.

We are then left with solving the quadratic:

$$\beta z^2 + (\alpha - \lambda) z^1 + \beta z^0 = 0 \tag{3.47}$$

and then form a solution using the two roots r_1,r_2:

$$x_n = A r_1^n + B r_2^n$$

3.7.2 The generating function method

Generating functions may be employed as an alternative way in solving recurrence relations. This method follows four main steps after a generating function has been defined as above

$$G(s) = \sum_{n \geq 0} \alpha_n s^n$$

We aim to express recurrence relations explicitly in terms of a generating function $G(s)$. The first step requires multiplying both sides by the variable s^n and summing over n over the range that the recurrence holds so:

$$x_{n+1} = \alpha x_n + \beta x_{n-1}$$

with α, β being constant coefficients becomes :

$$\sum_{n=1}^{N-1} x_{n+1} s^n = \alpha \sum_{n=1}^{N-1} x_n s^n + \beta \sum_{n=1}^{N-1} x_{n-1} s^n$$

The next step is to manipulate the power series to express them in terms of $G(s)$. To do this we need to make every sum to be over $n \geq 0$:

$$\begin{aligned} \frac{1}{s}\sum_{n=1}^{N-1} x_{n+1}s^{n+1} &= \alpha\sum_{n=1}^{N-1} x_n s^n + \beta s\sum_{n=1}^{N-1} x_{n-1}s^{n-1} \\ \frac{1}{s}\left(\sum_{n=0}^{N} x_n s^n - x_1 s\right) &= \alpha\sum_{n=0}^{N} x_n s^n + \beta s\left(\sum_{n=0}^{N} x_n s^n - x_{N-1}s^{N-1}\right) \\ G(s) &= x_1 s + \alpha s G(s) + \beta s^2\left(G(s) - x_{N-1}s^{N-1}\right) \end{aligned} \tag{3.48}$$

Then solving for the unknown function $G(s)$ we obtain a simplified formula to represent the sequence of terms generated by the recurrence relation :

$$G(s) = \frac{s(x_1 - x_{N-1}s^N)}{(1 - \alpha s - \beta s^2)} \tag{3.49}$$

For the final step we aim to get an exact expression for the terms of the sequence defined by the recurrence relation. Recognizing that successive differentiation may yield the required term :

$$x_n = \frac{1}{n!}G_x^{(n)}(0) = \frac{1}{n!}(n!x_n + (n+1)!x_{n+1}s^1 \cdots)$$

we see that it may be possible to obtain an exact expression for the nth term in the sequence of coefficients if $G(s)$ can be expanded into a power series by some method.

When $G(s)$ is a quotient of two polynomials then using *partial fractions* we may be able to obtain a closed-form expression for the n-th term x_n in the unknown sequence by expanding and handling each quotient separately. This is an easy way to find the Taylor series through the method of partial fractions. The idea is that we can express this rational function as a sum of fractions of the form $A/(1 - bx)$ for some values of A and b, but these fractions are functions whose Taylor series we already know :

$$\frac{A}{1 - bx} = \sum_{n=0}^{\infty} Ab^n x^n \tag{3.50}$$

Then we rewrite equation 3.49 in partial fractions :

$$G(s) = \frac{A}{(1 - as)} + \frac{B}{(1 - bs)}$$

where we let a and b be the inverse of the roots of the polynomial at the denominator of equation 3.49. Then as usual we proceed to find the constants A and B and end up with two fractions which we can identify as being of the form of equation 3.50. We therefore

rewrite as:

$$G(s) = \sum_{n=0}^{\infty} (Aa^n + Bb^n)s^n.$$

Using the fact that

$$G(s) = \sum_{n \geq 0} \alpha_n s^n$$

and we can clearly now see that the coefficient of $G(s)$ we are interested in is :

$$x_n = Aa^n + Bb^n$$

3.8 Matrix analysis

Previously we saw that Markov chains are usually formulated using a transition matrix and that long-term properties of the random evolution of a chain can be determined analysing the algebraic properties of transition matrix $\mathbb{M}$.

We have already seen how an probability distribution $\boldsymbol{f}$ evolves under a transition matrix $\mathbb{M}$ in discrete time, at each timestep $\boldsymbol{f}^{(0)}$ evolves as :

$$\boldsymbol{f} \Rightarrow \mathbb{M}\boldsymbol{f}$$

After multiple timesteps t we have :

$$\boldsymbol{f} \Rightarrow \mathbb{M}^t\boldsymbol{f}$$

This is a description of the mean process in discrete time. We would like to decompose the dynamics of the transition matrix $\mathbb{M}$ into n independent vector equations.

In the next two sections we will cover methods of factorizing $\mathbb{M}$ and then move on to show how to factorize a generic size $N \times N$ matrix of a particular form known as the tridiagonal which we shall be using in the analysis of our models in this thesis.

3.8.1 Spectral decomposition of a matrix

Usually $\mathbb{M}$ is not diagonal and thus raising it to the power n is not simple. We wish to find a matrix $\mathbb{Q}$ such that:

$$\mathbb{M} = \mathbb{Q}\Lambda\mathbb{Q}^{-1} \tag{3.51}$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a diagonal matrix of eigenvalues. This factorization process will simplify the calculation of the raising to matrix power n down to raising the elements of a diagonal matrix to n.

Spectral decomposition recasts a matrix in terms of its eigenvalues and eigenvectors. If a matrix is *normal* then we can be sure that spectral decomposition applies. A normal matrix is diagonalizable and has the property $\mathbb{M}^T\mathbb{M} = \mathbb{M}\mathbb{M}^T$ giving an orthonormal basis of eigenvectors $\left|\left|\boldsymbol{e}^k\right|\right|_k^n = 1$ for $\mathbb{R}^k$.

Such a matrix $\mathbb{M}$ can be decomposed to the sum of products of its eigenvectors and eigenvalues. We denote the ith right eigenvector as $\boldsymbol{e}_r^i$ and the respective left eigenvectors as $\boldsymbol{e}_l^i$ where $\boldsymbol{e}_r^i.\boldsymbol{e}_l^j = \delta_{ij}$ and λ_i is the ith eigenvalue. The left and right eigenvectors are multiplied as outer products for each component of the $\mathbb{M}$:

$$\mathbb{M} = \lambda_1 \boldsymbol{e}_r^1 (\boldsymbol{e}_l^1)^T + \lambda_2 \boldsymbol{e}_r^2 (\boldsymbol{e}_l^2)^T + ... + \lambda_i \boldsymbol{e}_r^i (\boldsymbol{e}_l^i)^T. \tag{3.52}$$

This result can be rewritten in matrix form with Q denoting the matrix of right eigenvectors arranged as columns and Λ the diagonal matrix of eigenvalues of $\mathbb{M}$ as: $\mathbb{M}\mathbb{Q} = \mathbb{Q}\Lambda$. Multiplying both sides by $\mathbb{Q}^{-1}$ we can rewrite as in equation 3.51 and it then follows that raising M to any power n results to

$$\mathbb{M}^n = \mathbb{Q}\Lambda^n\mathbb{Q}^{-1} \tag{3.53}$$

by observing that

$$\mathbb{M}^2 = \mathbb{Q}\Lambda\mathbb{Q}^{-1}\mathbb{Q}\Lambda\mathbb{Q}^{-1} = \mathbb{Q}\Lambda^2\mathbb{Q}^{-1}$$

and that the same applies for any power n.

Further, this result can be used to decompose the matrix exponential $\exp(\mathbb{M})$ through a Taylor expansion as:

$$\exp(\mathbb{M}) = \sum_{n=0}^{\infty} \frac{\mathbb{M}^n}{n!} = \mathbb{Q} \sum_{n=0}^{\infty} \frac{\Lambda^n}{n!} \mathbb{Q}^{-1} = \mathbb{Q} \exp(\Lambda) \mathbb{Q}^{-1} \tag{3.54}$$

which results conveniently to :

$$\exp(\mathbb{M}) = \mathbb{Q} \exp(\Lambda) \mathbb{Q}^{-1} \tag{3.55}$$

therefore allowing us to examine the dynamics towards the long-term ($t \to \infty$ steady-state) distribution evolution of a stochastic system described by the stochastic matrix $\mathbb{M}$.

Nevertheless, in some cases the matrix $\mathbb{M}$ may have degenerate eigenvalues, where by some λ_i occurs m times and is associated with more than one eigenvector. In that case the matrix is not diagonalizable as above but can be "block-diagonalizable" requiring the use of the Jordan canonical form according to which the diagonal entries contain eigenvalues and the superdiagonal entries are 1's. The Jordan canonical form consists of Jordan blocks :

$$\mathbb{J}_{n_m} = \mathbb{S}_m + \Lambda_m$$

on the diagonal where $\mathbb{S}_m$ is a nilpotent matrix [1] and Λ_m is the matrix of eigenvalue λ_m in the diagonal of each block. Thus now the matrix Λ does not necessarily contain distinct eigenvalues. Without going into details of the Jordan form we limit the discussion in saying that the matrix exponential in this case becomes :

$$\exp(\mathbb{M}) = \mathbb{Q}\exp(\Lambda)\exp(\mathbb{S})\mathbb{Q}^{-1}, \tag{3.56}$$

with the respective Taylor series expansion $\exp(\mathbb{S}) = \sum_{n=0}^{\infty} \frac{\mathbb{S}^n}{n!}$ giving some extra terms from the nilpotent matrix whose contribution in practice can be small due to the small values of eigenvalues obtained for the stochastic matrices studied in this thesis. Thus we may even use the decomposition of the non-degenerative matrices and obtain relevant results.

3.8.2 Eigenvalues of tridiagonal matrix

We will be dealing with systems that exhibit nearest-neighbour interaction :

$$x_n = \beta_n x_{n+1}(x) + \alpha_n x_n(x) + \gamma_n x_{n-1}(x)$$

These can be described by matrices whose only elements are arranged around the diagonal at a distance of one cell. Such matrices are called *tridiagonal* and have many applications on pure and applied mathematics, engineering and physics. For this reason they have been thoroughly studied and we will here recite the elementary results of obtaining the eigendecomposition of such a matrix. Let such a matrix be defined as follows :

$$\mathbb{A}_n = \begin{pmatrix} \alpha_0 & \gamma_0 & 0 & \cdots & \cdots & 0 \\ \beta_0 & \alpha_1 & \gamma_1 & 0 & \cdots & 0 \\ 0 & \beta_1 & \alpha_2 & \gamma_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \ddots & \ddots & \gamma_{n-1} \\ 0 & \cdots & \cdots & 0 & \beta_{n-1} & \alpha_n \end{pmatrix}$$

For such matrices we may define the determinants

$$D_1 = \begin{vmatrix} \alpha_0 & \gamma_0 \\ \beta_0 & \alpha_1 \end{vmatrix}$$

in a recursive form :

[1] A nilpotent matrix is a square matrix such that $\mathbb{S}^n = 0$ for some positive integer matrix power n.

$$D_n = \alpha_n D_{n-1} - \beta_{n-1}\gamma_{n-1} D_{n-2} \tag{3.57}$$

Determinants of this form where all elements are zero except the diagonal elements and the adjacent lines are called *continuants*. Overall, for some matrix size n we have $D_0 = 1$, $D_1 =$, $D_2 = \alpha^2 - \beta\gamma$ and so on. We use a characteristic equation method to solve the above recurrence relation:

$$\theta^2 - \alpha\theta + \beta\gamma = 0 \tag{3.58}$$

with roots:

$$\theta_\pm = \frac{\alpha \pm \sqrt{\alpha^2 + 4\beta\gamma}}{2} \tag{3.59}$$

and solutions given as:

$$D_j = A\theta_+^j + B\theta_-^j \tag{3.60}$$

where $A + B = 1$, since $D_0 = 1$ and $A\theta_+ + B\theta_- = \alpha$ since $D_1 = \alpha$ we solve for the unknown constants :

$$\begin{bmatrix} 1 & 1 \\ \theta_+ & \theta_- \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix} \tag{3.61}$$

Inverting the matrix we write :

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{\theta_- - \theta_+} \begin{bmatrix} \theta_- & -1 \\ -\theta_+ & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha \end{bmatrix} \tag{3.62}$$

$$= \frac{1}{\theta_- - \theta_+} \begin{bmatrix} \theta_- - \alpha \\ -\theta_+ + \alpha \end{bmatrix} \tag{3.63}$$

Expanding $(\theta - \theta_+)(\theta - \theta_-) = 0$ and comparing to Equation (3.58) we obtain $\theta_- + \theta_+ = \alpha$ and $\theta_+\theta_- = \beta\gamma$. The structure of the problems we will be studying in this thesis have $\beta = \gamma = -1$, so to keep our results relevant we shall assume these values. So $-\theta_+ = \theta_- - \alpha$ and $\theta_- = -\theta_+ + \alpha$ which allows us to write the solutions of A and B as:

$$A = \frac{\theta_+}{\theta_+ - \theta_-} \qquad \text{and} \qquad B = \frac{-\theta_-}{\theta_+ - \theta_-} \tag{3.64}$$

giving solutions of the form:

$$D_j = \frac{\theta_+^{j+1} - \theta_-^{j+1}}{\theta_+ - \theta_-}. \tag{3.65}$$

Looking for the eigenvalues we have $D_n = \det(\mathbb{A} - \lambda\mathbb{I}) = 0$ we have $\theta_+^{j+1} - \theta_-^{j+1} = 0$ and thus $\left(\frac{\theta_+}{\theta_-}\right)^{j+1} = 1$ and $\theta_+ = \theta_-^{-1}$. We can then write $\theta_+^{2(n+1)} = 1$ solutions with a root

of unity:

$$\theta_+ = \exp(im\pi\Delta x) \quad m = 1, \cdots, n \tag{3.66}$$

and with $\theta_+ = \theta_- = \alpha$ we have $\theta_\pm = \exp(\pm im\pi/(n+1))$, and substituting into (3.65) :

$$D_n = \frac{\exp(im\pi\Delta x)^{n+1} - \exp(-im\pi\Delta x)^{n+1}}{\exp(im\pi\Delta x) - \exp(-im\pi\Delta x)}. \tag{3.67}$$

The n solutions of $D_n = \det(\mathbb{A} - \lambda\mathbb{I}) = 0$ determine the eigenvalues $\lambda_m, m = \{1, \cdots, n\}$ since $\lambda = -(\theta_+ + \theta_-)$, the n eigenvalues are : From $\alpha = \theta_+ + \theta_- = 2\cos(m\pi\Delta x)$ the eigenvalues are:

$$\lambda_m = 2cos(m\pi\Delta x), \qquad \mathrm{m} = 1, \cdots, n \tag{3.68}$$

We can now proceed to determine the corresponding eigenvectors $\boldsymbol{e}^m$ using a similar recurrence relations. We proceed to consider the eigendecomposition as the following problem :

$$A\boldsymbol{e}^m = \lambda_m \boldsymbol{e}^m,$$

where $\boldsymbol{e}^m = (e_1^m, \cdots, e_l^m) \in R^l$ with $k \in \{1, \cdots, l\}$, we can then formulate the following recurrence relation for some m :

$$\beta e_{j-1} + \alpha e_j + \gamma e_{j+1} = \lambda e_j, \ j \in \{1, \cdots, l\}.$$

Taking $m \in \{1, \cdots, n\}$ gives n distinct eigenvectors, using a similar procedure as above we obtain roots $z_\pm = -\frac{1}{2}\exp[\pm im\pi\Delta x]$, and the lth entry of $\boldsymbol{e}^m$ is :

$$e_l^m = \frac{\sin ml\pi\Delta x}{\sin m\pi\Delta x} \tag{3.69}$$

As before $z_+ z_- = 1$, and using this fact the norm of the eigenvector $\boldsymbol{e}^m$ is obtained by:

$$||e_n||^2 = \sum_{l=0}^{n-1} \frac{(z_+^{l+1} - z_-^{l+1})^2}{z_+ - z_-} \tag{3.70}$$

$$= \frac{1}{(z_+^2 - 2 + z_-^2)} \sum_{l=0}^{n-1} \left(z_+^{2(l+1)} - 2 + z_-^{2(l+1)} \right) \tag{3.71}$$

$$= \frac{1}{(z_+^2 - 2 + z_-^2)} \sum_{l=0}^{n-1} \left(z^2 \left(\frac{z^{2n-1}}{z_+^2 - 1} - 2n + z^2 \left(\frac{z^{2n}}{z_-^2 - 1} \right) \right) \right) \tag{3.72}$$

with $z_+^{2(n+1)} = 1$ these products become:

$$||e_n||^2 = \frac{1}{(z^2 - 2 + z_-^2)} \left[\frac{1 - z_+^2}{z_+^2 - 1} - 2n + \frac{1 - z_-^2}{z_-^2 - 1} \right] \tag{3.73}$$

$$= \frac{-2(n+1)}{(z_+^2 - 2 + z_-^2)} \tag{3.74}$$

$$= \frac{-2(n+1)}{(2i \sin \frac{m\pi}{n+1})^2} \tag{3.75}$$

$$= \frac{(n+1)}{2 \sin^2 \frac{m\pi}{n+1}} \tag{3.76}$$

So we can write the normalized eigenvector as :

$$\hat{\boldsymbol{e}}^m = \sqrt{2\Delta x} \left(\sin(m\pi\Delta x), \sin(2m\pi\Delta x), \cdots, \sin(nm\pi\Delta x) \right)^T \tag{3.77}$$

These results will be used in Chapter 6 when we come to decompose a stochastic matrix.

Chapter 4

Formal models of memory

In previous Chapters we reviewed neuroscience research outcomes from cognitive and biological studies on memory. Biologists look for models of a system by searching for animals that have an accessible neural network for study composed of few neurons that can be mapped, excited and recorded from. Aplysia and the Drosophila Melanogaster are two such example model systems that we presented earlier in Chapter 2 on synaptic plasticity. On the other hand, a theoreticians approach to models of memory stems from the methods applied in theoretical physics. These methods require stripping of complexity to the bare essentials required for evaluating particular hypothesis of a system. Although detailed models of neurons are also considered in computational neuroscience, simple models are required in order to obtain understanding through formal mathematical treatment. In this sense, theoreticians of neuroscience, like physicists, approach problems using *toy models* one of which is the *formal neuron*. A formal neuron is an artificial entity far abstracted from real neurons that attempts to capture the behaviour of a neuron under a particular setting or context. In this chapter we focus on the theoretical study of memory that examines computational aspects of *artificial neural networks*.

With formal neurons we attempt to obtain a model of learning within the framework of the SPM hypothesis. These formal models use Hebbian type learning integrated with research results from the synaptic plasticity literature to show that synaptic plasticity enables memory and learning. The motivation of introducing the models of memory in this chapter is to appreciate the framework in which the synaptic models of memory have been developed leading to the mechanisms for plasticity we will be working with in this thesis.

4.1 The McCulloch-Pitts neuron

Inspired by neurobiology and digital computers a simple neural computational unit was first put forward by McCulloch and Pitts (1943). They attempted to capture the

notation of symbolic logic or propositions relating these with the "all-or-none" response of a model neuron. Neurons were modelled as simple threshold functions of a sum of the inputs that operated synchronously as simple logical switches. Their internal activation is:

$$h = \sum_i x_i^{in} \tag{4.1}$$

The output of the neuron is described by a simple threshold function of h, 1 if $h > \theta$ and 0 otherwise. The threshold θ ensures that a certain fixed number of synapses must be excited simultaneously, within the period being considered, to excite a neuron and produce an output. The inputs are binary and represent either excitatory or inhibitory inputs. There are two types of inhibition that can be employed by the model. *Absolute inhibition* corresponds to the original model implementation according to which an inhibitory synapse completely prevents excitation of the neuron. *Relative inhibition* allows an inhibitory input synapse to raise the threshold for activation by one. We may assume that a single unit can project multiple synapses to another unit that can result in increasing its excitation influence on the target neuron but also converting relative inhibition to absolute, if the number of inhibitory synapses exceeds the total possible level of excitation allowed by the given θ.

McCulloch-Pitts units divide the input space into two half-spaces. For a given input (x_1, x_2) and a threshold θ the condition $x_1 + x_2 < \theta$ is tested, which is true for all points to one side of the line described by the equation $x_1 + x_2 = \theta$ and false for all points on the other side. This separation of the input space can be used to identify groups of patterns and create what is known as a *classifier*.

Although very simple and binary, 14 out 16 possible logical functions of two inputs can be constructed with a single unit if we allow the use of relative inhibition. In fact McCulloch and Pitts (1943) proved that all those processes which can be described with a finite number of symbolic expressions, like simple arithmetic classifying or recursive application of logical rules, can be constructed by networks of such units.

Three simple two input logic gates are shown on Figure 4.2(a). The set NOT, AND and OR gates form a logical basis for the construction of all other logic functions, consequently all functions not implemented by a single unit can be constructed by networks of other units. To construct the AND gate we set $\theta = 2$, for two binary inputs $(x_1 + x_2)$ to activate the neuron, both of them have to be on simultaneously to exceed the threshold and in this way the neuron replicates the truth table of an AND gate, see table 4.1.

4.2 The Perceptron

Classifiers can be thought of as memories because they can respond in a set way to an input pattern that resembles a previously learned one. Perceptrons where introduced by

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

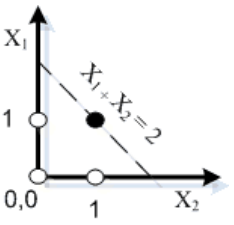
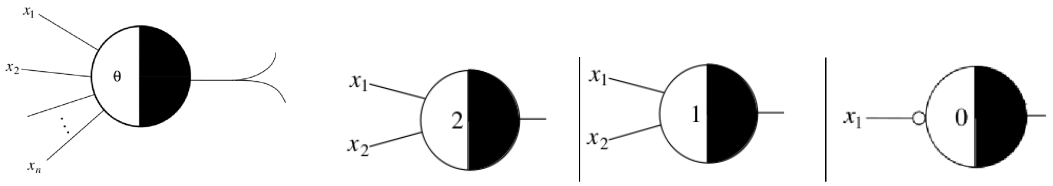


FIGURE 4.1: AND gate truth table and a geometric interpretation of a trained classifier to reproduce it. The AND logical gate truth table has x_n input and y as output. The diagram on the right shown the linear constraint set by the threshold function of the McCulloch-Pitts neuron which enables the separation of the input space into two classes, the white dots and the black dots. There is a single black dot representing the output 1 of the neuron.



(a) McCulloch-Pitts Neuron (b) AND (c) OR (d) NOT

FIGURE 4.2: **a**. A simple schematic of a McCulloch-Pitts unit, with input $x_1 \cdots x_n$ and an activation threshold θ. **b**. This neuron has a $\theta = 2$ that requires the activation of both inputs in order for the unit to respond, this models an AND gate, see Figure 4.1. **c**. OR gate **d**. A NOT gate, the small circle at the input denotes inhibition and thus an input of 1 is conveyed as -1

Rosenblatt in 1958 and came as an extension of the McCulloch-Pitts classifiers, by adding real-valued weights to inputs and introducing a learning algorithm. Such a model can learn the mapping of a set of input vectors to a particular output under a supervised learning scheme (Rosenblatt, 1958). A perceptron receives input from a single input bias b with an associated weight w_0 and N synapses each with an associated weight w_i, see Figure 4.3(a). The output of the perceptron neuron is determined by a non-linear function, which implements the neuron's activation function. A commonly used function is the hard limiter step-function :

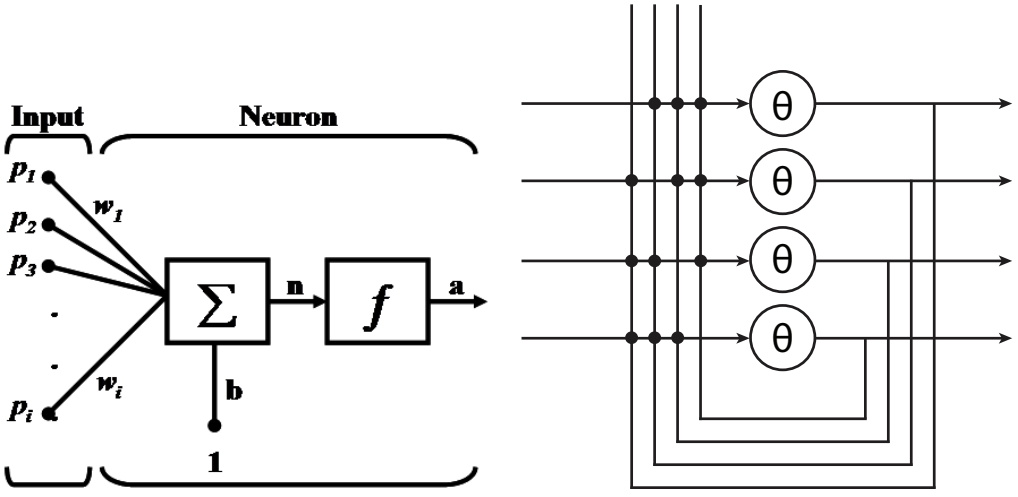
$$f(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

The hard limit function enables the perceptron to classify the input space into two regions. The output of the neuron is 1 if the net input to the transfer function is greater than zero. The outputs are interpreted as classification of the input pattern presented at the synapses:

$$\text{if } \boldsymbol{wx} + b > 0 \text{ then assign } \boldsymbol{x} \text{ to class A}$$
$$\text{if } \boldsymbol{wx} + b \leq 0 \text{ then assign } \boldsymbol{x} \text{ to class B}$$

Networks of perceptrons can learn complex input-output association functions, with the complexity of the learned function increasing with the number of layers. Multilayer perceptrons consist of a cascade of layers with feed-forward connections between layers. A single layer can learn to classify by modelling a separation plane in the N dimensional input space; having two layers (input, output) it can learn a surface of separation, while with three (input, hidden, output) it can possibly learn a surface that encloses all the points of a class (closed surface) in the N dimensional space but for enclosing the points it requires that the number of hidden units is more than the number of input units (Gori and Scarselli, 1998). The learned function defines a mapping of the input space onto a neighbourhood on the output space that can be used to learn lookup tables such as the one on Table 4.1.

The "learning" process is achieved by adjusting the Perceptron's weight vector $\boldsymbol{w}$ and bias b. The process uses a training set of input patterns and the associated desired outputs, let $\boldsymbol{D} = \{\{p_1, t_1\}, ..., \{p_m, t_m\}\}$ be a training set with a $\boldsymbol{p}_i$ an input pattern, t_i the associated target output and y the output of the perceptron. Training is an optimization process that adjusts the parameters of each perceptron embedded in a network until the presented $\boldsymbol{p}$ input gives the desired output t, thus an error signal is defined $e = t - y$. For the training to converge in a single layer perceptron the input patterns need to be linearly separable.



(a) The Perceptron (b) Hopfield network architecture

FIGURE 4.3: (a) The components of a perceptron neuron. Each input is weighted by a weight w_j, and the sum of the weighted inputs is sent to a threshold function f, usually f is a hard-limit function. There is also a single bias input which is added to the hard limit function input. (b) Typical Hopfield network connectivity. Each neuron has a threshold of activation θ and is connected with bidirectional synapses to every other neuron with no feedback connection.

The standard Perceptron learning algorithm requires the sequential presentation of the training input patterns. Each time a pattern is presented the error $e = t - y$ is evaluated

and an adjustment Δw is made to the weight vector $\boldsymbol{w} = \boldsymbol{w} + \Delta\boldsymbol{w}$ according to:

$$\Delta\boldsymbol{w} = (t - y)\boldsymbol{p} \tag{4.2}$$

The target output t and the Perceptron output y take values of 0 or 1. The error $e = t - y$ can take values $e \in \{-1, 01\}$. At every iteration the weights are adjusted by an amount equal to the input $\boldsymbol{p}$ in a direction set by the error signal e or not adjusted at all if $e = 0$. The algorithm is proven to converge in a finite number of steps over the training set for linearly separable classes of inputs or otherwise it will not stop if the problem cannot be learned. In that case the algorithm can be modified to measure the average error over the presentation of a complete training set and the algorithm made to stop when the average error falls below a set point.

4.3 Associative memory

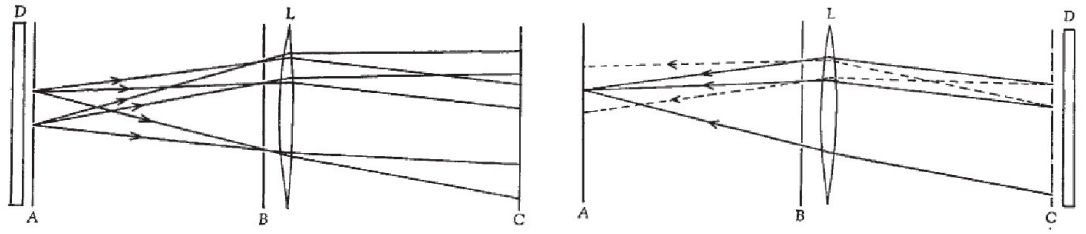
In this section we review another class of systems that is able to *associate* an input vector and its neighbourhood, as defined by some metric, with a given output vector. The function of an associative memory is to recognize a previously learned input vector, even if the input is not completely faithful to the originally learned one.

We examine two types of systems, *Heteroassociative networks* and *Autoassociative networks*. Heteroassociative networks map each member $\boldsymbol{p}^m$ of a set of n dimensional input vectors to one of m output vectors from the set $\boldsymbol{t}^m$. A process of learning associates $\boldsymbol{p}^i \rightarrow \boldsymbol{t}^i$, while recall is able to reproduce vector $\boldsymbol{t}^i$ at the output in response to a cue pattern $\boldsymbol{\acute{p}^i}$ whose difference from the learned $\boldsymbol{p}^i$ one is not larger than some maximum error value $|\boldsymbol{\acute{p}} - \boldsymbol{p}^i|^2 < \epsilon$.

Autoassociative networks operate in similar principle to heteroassociative networks only this time each vector is associated with itself. The function of such a network is to enable recall of a complete pattern from a noisy cue pattern where still the constraint $|\boldsymbol{\acute{p}} - \boldsymbol{p}^i|^2 < \epsilon$ holds. In the next sections we present an prominent examples of an heteroassociative and an autoassociative network. In both networks increasing the number of stored patterns decreases the allowed error ϵ in the cue pattern.

4.3.1 Willshaw cue-recall memory

Willshaw, Buneman, and Longuet-Higgins in 1969 inspired by holography and building upon a parallel to a simple correlogram proposed a heteroassociative memory system, see Figure 4.4.



(a) Constructing a correlogram (b) Reconstructing a pattern

FIGURE 4.4: D is a diffuse light source, A is a plane with two pinholes, B is the second pattern of pinholes, L is the lens and C is the correlogram of A and B. **a**. D projects the rays of light through two pinholes on plane A. Each of the two rays is then projected through each of the three pinholes of plane B and then through the lens onto plane C, the projected 2*3 points is the correlogram between the pattern at A and that of B. **b**. shows the reconstruction. The dotted lines indicate rays that did not exist on during the construction of the pattern. The filled lines that converge to points will create points of higher luminosity than the spurious rays. The point of convergence finds the position of one of the original pinholes on plane A by opening pinholes on C where the projections from the A pinhole where on **a**. Thus, using the correlations of the patterns of A and B projected on C, we can use B and C to find A by setting a threshold on the minimum brightness of the spots on A to ignore spurious projections.
Source: (Willshaw et al., 1969)

In the neural equivalent we assume a two layer architecture with feedforward connections where every input neuron projects an axon to every output neuron. The connections create a grid of synaptic connections and each synapse can either be active or inactive. The network model uses stored correlations on a matrix W between patterns $\boldsymbol{p}$ and $\boldsymbol{t}$ to reconstruct one of the patterns ($\boldsymbol{t}$) by using the other pattern ($\boldsymbol{p}$) and the correlation matrix W. In matrix form, we let P be the $m \times n$ matrix whose rows are each one of the input vectors and T an $m \times n$ matrix whose rows are the output vectors. We require a weight matrix W that gives :

$$PW = T \tag{4.3}$$

We assume that $\boldsymbol{p}$ and $\boldsymbol{t}$ are random uncorrelated vectors of length N. Willshaw et al. (1969) in their original implementation used N length vectors of positive integers with sparse coding, meaning that only a percentage f of the inputs contained values to be associated with a fraction f of the output vector. Thus, decreasing f reduces the density of the encoded information by decreasing the number of activated neurons from the patterns :

$$\sum_{i=1}^{N} p_i^{\alpha} = \sum_{i=1}^{N} t_i^{\alpha} = fN. \tag{4.4}$$

The components of W, w_{ij} represent the synaptic efficacy between presynaptic neuron j and the postsynaptic neuron i. Learning is achieved by imposing the input and output patterns and changing the value of w_{ij} to 1 if input i and output j are active

simultaneously or 0 otherwise :

$$W_{ij} = \begin{cases} 1 & \text{if } p_i^\alpha t_j^\alpha \text{ for any } \alpha \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

The authors showed that the density in bits per synapse with which the associative network stores information in the limit of large N is 0.69 bits (Willshaw et al., 1969), but reaching these limits requires specific values on the coding f to be set per network size N ($f = \ln N/N$) away from which capacity is significantly reduced. Thus the model works well under sparse coding. We will not dwell on information capacity in this model as such analysis does not serve the purposes of this thesis, we direct the reader to authors who have studied this model and its capacity Meunier et al. (1991), Nadal and Toulouse (1990), Nadal (1991), Palm (1980).

4.3.2 The Hopfield network

The Willshaw associator had no feedback connections, adding feedback creates recurrent networks that can act as an autoassociative memory. A Hopfield network (Hopfield, 1982) is such a recurrent neural network. Instead of correlating two patterns $\boldsymbol{p}$ and $\boldsymbol{t}$, it can associate an input pattern with itself by introducing feedback connections in the neural network. The feedback connections serve to sustain output activity in the absence of input. The recurrent dynamics progressively move the output closer to a stored pattern. These dynamics are due to attractor states of neural activation stored in the patterns of the network's synaptic connections. Such dynamics allow recall by using incomplete input patterns as a partial recall cue. Presenting an incomplete version of a stored pattern, will produce an output that will be fed back to the input. Through multiple iterations, we expect input states that are sufficiently similar to the memory stored states to settle at the stored pattern that closely matches the cue. To present cue pattern we assume that each neuron receives a direct input from a source external to the network along with the input from other neurons in the network. Given an initial condition of external inputs the network activity is expected to reach a stable equilibrium (attractor) state within a given number of iterations. The updates in the Hopfield occur asynchronously and therefore we expect that each unit i preserves its individual state until it is randomly selected for an update.

We define the weight of a synapse from neuron j to neuron i as w_{ij} and the externally applied direct input to neuron i as I_i. Each weight is symmetric with $w_{ij} = w_{ji}$. The original implementation used binary output neurons in discrete time whose state was

defined by the activation function:

$$s_i(t+1) = \begin{cases} 1 & \text{if } \sum_k w_{ik} s_k(t) + I_i > \theta \\ 0 & \text{otherwise} \end{cases}$$

Each unit outputs +1 if the total level of excitation exceeds the threshold θ. The threshold for activation can be set to zero. Assuming $\theta = 0$, we may now write:

$$s_i(t+1) = \text{sign}[\sum_k w_{ik} s_k(t) + I_i]$$

The equilibrium points of neuron activation represent stored memories that can be retrieved by partial activation cues delivered through the external inputs to the network. The network can learn new memories by modifying the synaptic matrix $\mathbf{W}$ using Hebbian learning rules.

Adding new memories involves forming new attractor states in the state space. Linear accumulation of new patterns is possible by loading the selected n-dimensional stable states $\xi^1, \xi^2, \cdots, \xi^\mu$ on the network one at a time, with $\mu = 1 \cdots P$. The patterns $\boldsymbol{\xi}^\mu$ are n-dimensional vectors randomly generated by setting ξ_i^μ to values $\{0, 1\}$ with equal probability. The original implementation used continuous unbounded variables to represent synaptic efficacies. The synaptic matrix is assumed to contain all zeros initially and it accumulates patterns incrementally using Hebbian learning:

$$w_{ij} = \sum_\mu (2\xi_i^\mu - 1)(2\xi_j^\mu - 1) \text{ for } i \neq j$$

The above rule assumes there is a mean firing rate of 1 and when pre- and postsynaptic activity is sitting at the same side of the mean activity synaptic weight increases, otherwise it weakens. This is also known as the covariance learning rule according to which the weight increases if presynaptic and postsynaptic activity are positively correlated. If we recast ξ_i^μ to take values $\{-1, 1\}$ and normalize the weight matrix by the number of patterns N we may rewrite the above incremental learning rule as:

$$w_{ij}^\mu = w_{ij}^{\mu-1} + \frac{1}{N} \xi_i^\mu \xi_j^\mu \; i \neq j.$$

From now on we will use bimodal states $\{-1, 1\}$ to indicate the activation patterns, this does not change the overall behaviour of the model while it makes it similar to a well studied model of magnetic spin in physics (Amit et al., 1985a). Such ± 1 patterns raise concerns as these can be interpreted as switching synapses between inhibitory and excitatory, but their interpretation is generalized to weak and strong states by considering the recasting under the covariance learning rule above.

With linear accumulation the order that the patterns are presented does not matter and all patterns can be loaded in the weight matrix at once:

$$w_{ij} = \frac{1}{N} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu}, \ i, j = 1, \cdots, n \text{ and } i \neq j.$$

In matrix notation we simply write the outer product of each of the input vector making sure there are no self connections by substracting the identity matrix $\mathbf{I}$:

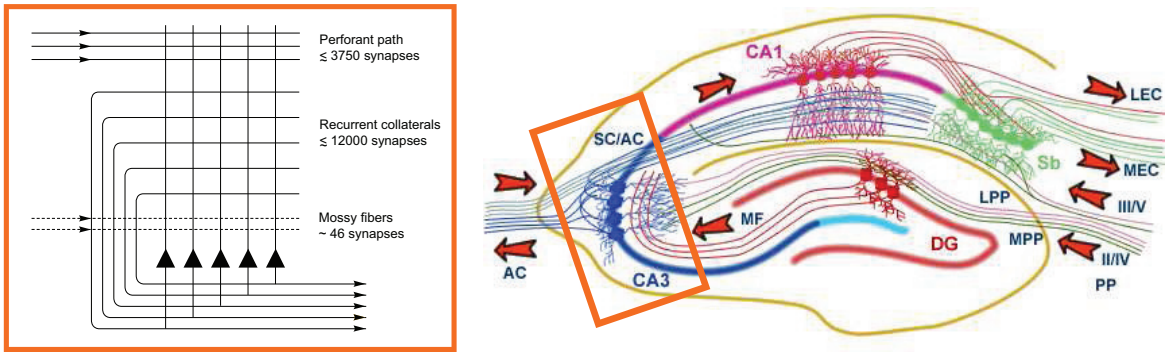
$$W = (\boldsymbol{\xi^1}^T \boldsymbol{\xi^1} - I) + ... + (\boldsymbol{\xi^\mu}^T \boldsymbol{x^\mu} - I)$$

The state of the network is described by the collection of the activities of the neurons. We consider a discrete model of neural activity with neurons acting as bimodal output units, the units are either firing or not. It should be noted at this point that the collective properties of the Hopfield network are retained with graded response neurons also (Hopfield, 1984). The state of the system can be described by a binary n-tuple of n neurons with the model neurons being simple bistable elements each being capable of assuming two values, $s_i = -1$ for basal activity and $s_i = 1$ for high firing activity. Making the input patterns ξ^μ fixed points of the dynamics requires the definition of a global energy function:

$$E = -\frac{1}{2} \sum_{i,j}^{N} w_{ij} s_i s_j \tag{4.6}$$

The minima of the energy function define stored attractor states representing the stored vectors. A key issue that determines the capacity of such networks is locating as many attractors as possible in the input space and each one of them should define a region of influence. This region will result in similar vectors eventually being drawn to the same attractor state. These attractors are guaranteed to be stable if the connections in a recurrent neural network are symmetric and there are no self- connections (Cohen and Grossberg, 1983), meaning that the diagonal of the $\mathbf{W}$ matrix is zero and $w_{ij} = w_{ji}$. When the stored $\boldsymbol{\xi}$ vectors are orthogonal then the distribution of attractor states is evenly spaced, but finding a set of such bipolar orthogonal vectors requires constructing what is known a Hadamart matrix for which there is no general construction rule for any size matrix.

Of main interest is measuring the number of patterns P that can be stored and recalled in a Hopfield network and then how the model synapse influences this performance measure. Analysis of the original Hopfield model attractor dynamics and various extensions have studied by many authors McEliece et al. (1987), Amit et al. (1987), Mézard et al. (1986), Amit et al. (1985a,b), Amit and Fusi (1994). Here, we will be presenting a basic form of analysis to understand the stability of a single neuron and we also present key results from research on the collective dynamics of this recurrent network.



(a) Model of rat Hippocampal associative network

(b) Hippocampal section showing neural networks

FIGURE 4.5: Recurrent network in hippocampus and abstract model of connection pattern. In a labelled hippocampal cross section we find that in the CA3 area there are neurons with recurrent collateral connection.

4.3.3 The capacity of the Hopfield network

The measure of capacity has become a central benchmark in evaluating models of memory, mostly because hippocampal cross sections show that CA1 and CA3 areas have recurrent network topology, see Figure 4.5. Therefore, since recurrent connections are characteristic of a Hopfield model this raises the possibility that areas such as CA3 operate as autoassociative networks.

If P fundamental memory patterns are chosen at random, the maximum asymptotic value of P in order that most of the P original memories are exactly recoverable defines the memory capacity of the network. Recall is assumed to be successful if a pattern can be reproduced faithfully (3% error is usual) by the network's output neurons after cue pattern has been imposed. The cue pattern is enforced to the output neurons for a few update cycles and then once removed the dynamical system is expected stabilize on the attractor of the recalled memory. Using Equation (4.3.2) we may substitute w_{ij} and rewrite the the state of each neuron at time t as:

$$s_i(t+1) = \text{sign}[\frac{1}{N}\sum_j \sum_\mu \xi_j^\mu \xi_i^\mu s_j(t)]$$

We may perform a signal to noise analysis by picking the first pattern $\mu = 1$ as the stored state and tracking how the storage of subsequent patterns may effect the state of a single output neuron. The simple single neuron method reproduced here is standard and can be found in Trappenberg (2010), Peretto (1992). We assume that the desired output state of the neuron is the first pattern that stored $\boldsymbol{\xi}^1.\boldsymbol{s}(t) = N$. The expression for the activation of the neuron is split into two parts, the first is the activation due to the first tracked pattern, and the second is the contributions due to the subsequent

stored patterns:

$$s_i(t+1) = \text{sign}[\frac{1}{N} x_i^1 \sum_j^N \xi_j^1 s_j(t) + \frac{1}{N} \sum_j^N \sum_{\mu=2}^P \xi_j^\mu \xi_i^\mu s_j(t)]$$

We assume that all other neurons are in the correct state and examine when this particular neuron loses its output. Thus, the first term can be simplified since the first pattern stored is also the activation pattern we are tracking, ie. $s_j(t) = \xi_j^1$. The product of the first term is always 1 and it sums to $N-1$ because it can be rewritten as $\sum_{j=1}^{N-1}(\xi_j^1)^2$, since there is no self connection between neurons the sum does not evaluate at $j \neq i$. We ignore this last point and simplify by taking the sum as N. The normalization term of the Hebbian learning is then simplified. We may then rewrite as:

$$s_i(t+1) = \text{sign}[\underbrace{\xi_i^1}_{signal} + \underbrace{\frac{1}{N} \sum_j^N \sum_{\mu=2}^P \xi_j^\mu \xi_i^\mu \xi_j^1}_{noise}]$$

We find that there are two contributions to the activation function. The first term of the sum is called the *signal* and it represents the fixed memorized value we wish to retain. The second term is called the *noise* because it represents the contributions of the patterns stored after the tracked first one which may destabilize the output. This condition is met when ξ_i^1 becomes smaller than the noise term as more patterns are being memorized. As a side note, if the cue recall pattern given is $s_j = -\xi_j^1$ then this would also be a stable pattern because the sum $\sum_j^N \xi_j^1 s_j(t)$ would simply evaluate to -1 instead of 1. Thus, when storing a pattern $\boldsymbol{\xi}^\mu$, $-\boldsymbol{\xi}^\mu$ also becomes an attractor state.

The noise term consists of the random sequence of further memorized patterns $\xi_i = \pm 1$, forming a sequence of Bernoulli trials. Noise is then the average over N of sum of $N \times P$ of contributions of either $+1$ or -1. This creates a random walk of the noise term and we can treat it as a random variable with a binomial distribution of zero mean and $\sigma = \sqrt{P-1}$. For large N we can take the distribution to be Gaussian, with a standard deviation $\sigma = \sqrt{(P-1)N}$ and adding the normalization factor N yields:

$$\sigma = \sqrt{\frac{P-1}{N}} \approx \sqrt{\frac{P}{N}} \tag{4.7}$$

The fraction P/N represents the load of the network as a ratio of the number of patterns over the number of neurons, and we shall denote it as α. Knowing the distribution of the noise term is Gaussian we may estimate the probability that it will reach the value of $\xi_i^1 \in \{+1, -1\}$ and interfere with the signal term. We shall take the signal term to be always 1, as the problem is symmetric for both $\xi_i^1 = -1$ and $\xi_i^1 = +1$. Since the noise term has a Gaussian distribution, we use the cumulative distribution function to find

the probability that a random variable falls outside of the range of the signal value:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-(x-\mu)^2/(2\sigma^2)} dx \tag{4.8}$$

We define this probability as :

$$\mathbb{P}[\xi_i^1 + \sigma < 0] = \mathbb{P}[\xi_i^1 < -\sigma] = \frac{1}{2}[1 - erf(\frac{1}{\sqrt{2}\sigma})] \tag{4.9}$$

The error function erf gives the probability that a Gaussian distributed variable is within some range $(0, x)$. Here, by evaluating the integral of the Gaussian PDF of Equation (4.8), the erf is used to obtain the probability that noise has exceed the signal. We can now estimate the probability that the recalled vector is correct within some error margin for a given load α. As the load increases the probability that the noise will exceed the signal increases and thus a particular output neuron will wrongly recall the opposite state. This extends to all neurons in the network with the number of incorrectly recalled components of $s_j(t)$ being approximately Poisson distributed (see McEliece et al., 1987). As an example, to ensure that correct recall with a probability above 99% we may set $\alpha = 0.18$ which gives a probability :

$$\mathbb{P}[error] = \frac{1}{2}[1 - \text{erf}(\frac{1}{\sqrt{2 * 0.01}})] \simeq 0.01$$

The above analysis assumes that storing random patterns in a Hopfield network increases the retrieval error as a result of an increase in the noise term, but the network also exhibits collective dynamics due to the recurrent connections which have not been taken into account. If an output flips state due to a retrieval error then this error will propagate to other neurons on the next time-step adding further noise. The course of this effect may amplify or attenuate and thus strongly affects the dynamics of the Hopfield network. As new patterns are stored the probability of retrieval error due to noise rises and the probability of the error propagating rises simultaneously but with different dynamics.

Specifically for this network the storage capacity p for random uncorrelated patterns is proportional to the number of neurons $p = \alpha N$ (Amit et al., 1985a). Crossing a critical threshold on the number of memories stored $\alpha_c \approx 0.14$, takes the network in a disrupted state and the system looses the ability to retrieve any of the stored patterns (Parisi, 1986, Nadal et al., 1986, Amit et al., 1985a,b, van Hemmen, 1987). But more generally, van Hemmen (1987) showed that there is a universal function $F(x)$ that determines the threshold for catastrophic forgetting α_c for any inner product learning rule and the highest capacity comes with the standard Hopfield linear rule with $\alpha_c = 0.138$ (see also Amit et al., 1985b).

Thus, learning needs to stop before hitting the critical threshold on capacity, otherwise none of the stored patterns can be recalled. This phenomenon has been termed

catastrophic forgetting or *catastrophic interference* and in the Hopfield network it is manifested as a sudden loss of all attractor states simultaneously as a result of its collective dynamics. The catastrophic interference phenomenon is a general problem of connectionist networks (McCloskey and Cohen, 1989, Ratcliff, 1990) and it reflects a requirement for a balance in making a memory system that is simultaneously sensitive to encode new memories but is not radically disrupted by these new patterns (see Abraham and Robins, 2005, French, 1999).

4.4 Palimpsest memories

In the previous section, we reviewed a method that reveals an upper limit in memory capacity for a Hopfield network of a fixed number of resources. The fact that all memories fade simultaneously leading up to catastrophic forgetting may perhaps render us with an unfamiliar experience compared with modern storage media we are accustomed, but the existence of an upper limit in capacity should not be surprising. In any case the fact that the Hopfield memory is completely disrupted after reaching this threshold is a major limitation of the model. How are we to proceed with a requirement for continuous memory storage in the face of limited resources? Solutions proposed for resolving the issue will be presented in this section leading to the development of the synaptic plasticity framework we will be adopting in this thesis. This framework takes on an ancient old tactic of writing-material reuse to introduce automatic forgetting due to new memory storage. This solution was used in the past to deal with scarcity in writing material and has resulted in many marvel ancient texts, such as a mathematical treatise by Archimedes, to be overwritten only to be rediscovered using x-ray imaging (see Netz and Noel, 2011). Material that could be reused to write new information on-top of old is known as a *palimpsest*; the origin of the word stems from Greek "παλίμψηστος" and its etymology is the combination of the words "πάλιν" translated as "again" or "repeat" and the noun derived from the verb "ψάω" that translates into "I scrape/polish".

4.4.1 Marginalist learning, Nadal et al. (1986)

Early attempts to resolve the overloading issue relied on a hypothesis that this was due the *uniform acquisition* of all memories. To circumvent this issue the method of *marginalist learning* was suggested. This method breaks the uniformity in the acquisition of memories pattern by enhancing the variation in the synaptic strength induced due the storage of the most recent pattern M. The kth pattern increased its synaptic strength exponentially as $\exp(\gamma k)$, where γ is a constant of embedding strength (Nadal et al., 1986). Since the plasticity step grows with k the synaptic strength essentially depends only on the last input patterns and the old patterns are automatically attenuated and forgotten (Nadal et al., 1986). By choosing an appropriate value of γ the state of

total confusion is avoided and the memory still keeps the capability of storing a number of patterns proportional to N, although now capacity is halved (Nadal et al., 1986) Although the original Hopfield implementation used continuous unbounded synapses that could grow up to the critical limit, the marginalist solution assumes that synapses continuously grow without bounds and exponentially fast. Thus, the method may provide for automatic forgetting but its assumptions on synaptic strength are unrealistic.

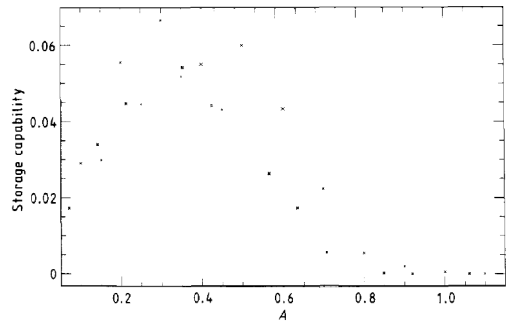
4.4.2 Bounded strength

A different solution to the marginalist learning was independently developed by Parisi, Nadal, Toulouse, Changeux, and Dehaene, that provides realistic synapses through an upper bound on synaptic efficacy (Parisi, 1986, Nadal et al., 1986, Sompolinsky, 1987, van Hemmen, 1987) but also introduces automatic forgetting. As we will see below, this idea was then further developed to synapses of discrete states of strength.

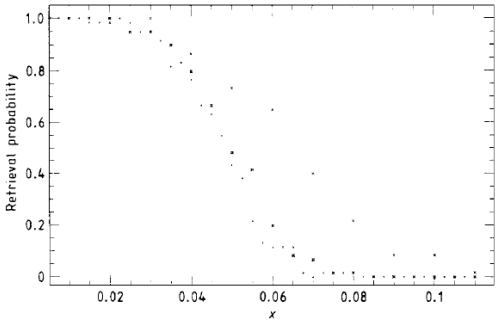
Parisi (1986)

Parisi in 1986 performed simulations to examine the limit in weight space after which catastrophic forgetting occurs. He then proposed to constrict the range of synaptic weights by imposing two hard bounds to constraint efficacy within limits before hitting the catastrophic forgetting critical threshold. The bounds were implemented by truncating any transition that would take the weight value out of an allowed range A. The range was $-A \leq w_{ij} \leq A$, allowing synapses to change between inhibition and excitation. To find the bounding value A, he argued that A should not be larger than the distribution variance of the synaptic weight w_ij. With unbounded synapses every synaptic modification is superimposed over older distributions of weights set by previous learned patterns. The distribution of w_{ij} is a Gaussian with variance $\sqrt{\alpha}$. Given that catastrophic forgetting occurs at $\alpha = \alpha_c \approx 0.14$, then the distribution of w_{ij} should be modified at this point so $A < \sqrt{\alpha_c}$.

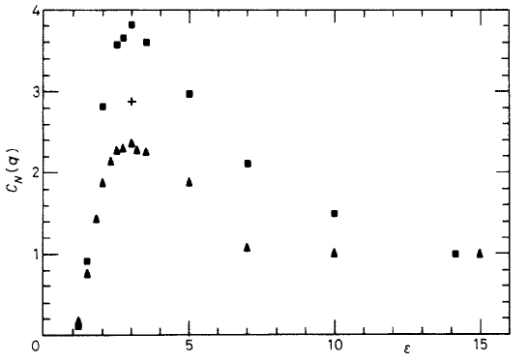
Bounding synapses introduces a non-linear constraint in the distribution of synaptic weights. This allows for an interference property by which old memories decay by being overwritten by new ones (Parisi, 1986). The retrieval probability of a stored memory drops as k new memories are added, see Figure 4.6(a). The forgetting induced was from the storage of new memories and it had the effect of not allowing the Hopfield memory to enter into a state of catastrophic interference where all memories are lost simultaneously (see also Amit et al., 1987). Using simulations Parisi found that an optimum value for A was around 0.35, while for $A \geq 0.7$ catastrophic forgetting occurred, the original results of this research are shown on Figure 4.6(a).



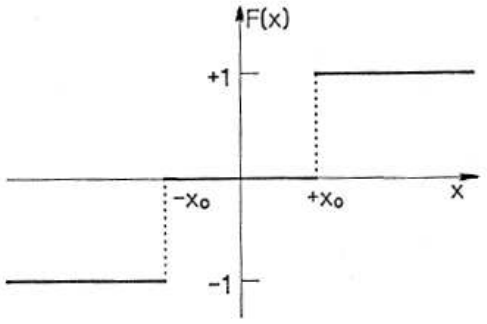
(a) Capacity against bounding value A



(b) Retrieval probability



(c) Capacity against embedding strength



(d) The 3 state non-linear clipping function

FIGURE 4.6: Three values for network size are shown N=100 (×), N=200 (◇),N=400(●). **a**. The retrieval probability of a given pattern as a function of $x = k/N$ after k patterns have been subsequently stored in the network. **b**. Storage capability as a function of A, the storage capacity appears proportional to N at a fixed A. A maximum exists around 0.35 which amounts for $0.05N$. Source : (Parisi, 1986) **c**. Asymptotic capacity: C_N as a function of ϵ(see text), for $N = 100$ (△) and $N = 200$ (⋄, +). For comparison with the other models, C_N is the mean number of patterns with retrieval better than 97% (Nadal et al., 1986) **d**. After training the synapses are clipped according to the activation function $F(x)$ between three values +,0,-1. The x_0 shows the threshold point for mapping synaptic strength among the 3 states. Source: Sompolinsky (1987).

Nadal et al. (1986)

Other early attempts to introduce bounded synapses aimed to improve the biological realism of model synapses and were not directed towards solving the catastrophic forgetting issue. It was pointed out that because cortical synapses are either inhibitory or excitatory then model synapses should be bound at zero and only allowed to grow indefinitely in one direction (Toulouse et al., 1986).

Around the same time, Nadal et al. 1986 modified the earlier sign preserving constraint and introduced bounds on strength also, where $-A \leq w_{ij} \leq 0$ or $0 \leq w_{ij} \leq A$, depending on the sign of the synapse. These bounded synapses gave qualitatively similar results to Parisi (1986), see Figure 4.6(c). New memory encoding increases the mean squared

strength K and a study of the minimum threshold strength s required to encode a new memory against K showed that safe encoding of a new memories requires s to grow as :

$$s = \epsilon^2 K/N,$$

where ϵ is a constant that sets the encoding strength. Large values for s preserved only the last encoded memory and small s exhibited catastrophic interference after a few memories where encoded. Using simulations the authors found an optimal value at $\epsilon = 3$ where catastrophic forgetting was abolished and the maximum capacity was achieved, see Figure 4.6(c). Dropping the sign preserving constraint did not affect their results qualitatively, thus effectively reproducing Parisi's results.

The networks exhibited forgetting under new memory storage. This property has been likened to a *palimpsest* and is characteristic of networks using bounded synapses (Nadal et al., 1986). Palimpsest memories have an approximately fixed window of retained memories where the oldest memories are forgotten to make room for new memories.

Sompolinsky (1986)

Instead of bounding the strength during sequential learning, Sompolinsky (1986) considered clipping synapses using a non-linear sign function after all k patterns had been presented:

$$w_{ij} = (\sqrt{k}/N)\text{sign}(\sum_{\mu}^{k} \xi_i^\mu \xi_j^\mu) \tag{4.10}$$

The result showed that all memories were faithfully recalled as long as the storage limit $k = \alpha N$ memories was $\alpha < 1/(\pi \ln(N)$ (see also McEliece et al., 1987, for a derivation). This learning scheme required synapses to retain their intermediate values formed by small step changes of the scale $1/N$ for every pattern learned, until the end of the learning period at which point the strengths are clipped.

Such non-linear realizations of the Hebb rule were also extended to three-state clipping. As a side note, analysis of optimal state count in a communication channel assuming a cost $K = N_{channels} * N_{states}$ reveals that the optimal is e, and 3 states per channel (see Rojas, 1996, Wiener, 1948). Three-state clipping operates by modifying the weight value of a synapse, after learning, to one of three values {+1,0,-1 } depending on whether the value of w_{ij} is above or below two thresholds $-x_0$ and $+x_0$ (Sompolinsky, 1987), see Figure 4.6(d). The zero-strength achieves a dilution of the network by pruning synapses in an activity dependent manner. For certain values of $x_0 \cong 0.82$ (and $x_0 = 0$ naturally) the performance of the three-state network is equivalent to the fully connected 2-state clipped model, while it increases the critical storage point to $\alpha_c \approx 0.12$ at values around $x_0 \approx 0.62$ (Sompolinsky, 1987).

Amit and Fusi (1992)

The clipping mechanism received further support because it provided solutions to practical issues in analog neuromorphic chip design (Morgenstern, 1987, Amit and Fusi, 1992). Implementing synapses of infinite analogue depth on an electronic chip requires the use of capacitors to store the weight value at any point in time, but as potential leaks through time then it becomes necessary to periodically refresh synaptic values. However, it is not possible to refresh the potential to the exact original value so it would have to be updated to predetermined one, thus effectively manifesting the clipping method of Sompolinsky (1987).

The original clipping mechanism was applied after the learning period was over. For online memory however, synaptic weight clipping has to be periodically applied to refresh the potential stored in the capacitors (Amit and Fusi, 1992). Given this fixed cycle, the effects on learning were found to depend on the presentation speed of the sequence of patterns to be learned. Presenting more memories per refresh cycle improved memory capacity. However, long-term memory should rely on the stable synaptic states, and thus mass presentation needs to be ignored to examine learning of a single pattern per refresh cycle.

At this point, with one memory per refresh cycle the clipping model has made a transition from incremental learning in infinitesimal steps to discrete states of strength. Each memory is strongly encoded and only a maximum of $k = \ln N$ memories can be stored in the weight matrix with three-state synapses (Amit and Fusi, 1992). To overcome the drastic drop in capacity with discrete states a new mechanism of *stochastic learning* was devised. According to this mechanism every presentation of new pattern only affects a fraction of the synapses, effectively slowing down learning. This line of research contemplated the idea that storage between refresh cycles in chips should be analogue, but discrete synaptic states are required for long-term retention while these are obtained via the periodic clipping mechanism. These ideas where then transferred to models of biological synapses to consider the hypothesis that long-term memory in real synapses may also rely on a few stable states and stochastic learning (Amit and Fusi, 1992, 1994).

Binary synapses Tsodyks (1990) and Amit and Fusi 1994

Strong encoding with binary strength synapses leads to large capacity decreases compared to palimpsests with analogue (real valued) strength synapses. Tsodyks (1990) and Amit and Fusi 1994 examined binary strength synapses with stochastic transitions between states, see Figure 4.7 (left). Binary synapses switch between potentiated and depressed states in response to a set of uncorrelated binary stimuli continuously presented with some rate rt. The two states set the bounds of the synaptic strength and

the probability of transition q sets the learning rate. Given a POT stimulus the transition from weak state to the strong state takes place with probability q, the reverse requires a DEP stimulus, thus every plasticity stimulus is accepted with probability q. This model allows new patterns to be continuously written on-top of previous ones, with q setting both the speed of learning and the erasure of old patterns simultaneously.

Binary synapses with stochastic learning have been used to model associative memory (Tsodyks, 1990) and successfully applied on perceptron classification of linearly separable patterns (Senn and Fusi, 2005). It has also been adopted in neuromorphic chip design as it is favoured over other models due to its low component count per electronic synapse (Fusi et al., 2000). Fusi (2002) suggested that such a local stochastic update rule exists in biological synapses and that synapses exploit the stochasticity of the naturally occurring interspike interval recorded from cortical neurons to achieve stochastic transitions.

Extensions to multistate synapses have also been proposed with a sequential arrangement of strength states having equally spaced values between 0 and 1. Further revisions, have introduced hidden states that modify the transition probability to model metaplasticity (Abraham and Bear, 1996), in order to break the symmetry between the speed of learning and forgetting.

4.4.3 Metaplastic synapse models

Synaptic models of metaplasticity have a set of hidden states corresponding to each level of efficacy. These hidden states change the degree of plasticity that a synapse will express at the next memory encoding and therefore they also constitute models of metaplasticity (Abraham and Bear, 1996). We briefly mention two such models here since this thesis will focus on one of these models in more detail in the next Chapter.

4.4.3.1 Multistate Synapse

Multistate models have a number of states connected in sequence, see Figure 4.7 (middle) thus enforcing a serial transversal of states that is stopped by one terminal state at each end. Different versions of serial synapse has appeared in Amit and Fusi (1994), Rubin and Fusi (2007), Leibold and Kempter (2008), with the last two assuming binary strength by letting an equal number of serially connected states be associated with one of two synaptic efficacy values. In this model the induction stimulus determines the direction of motion between the connected states deterministically. There is one terminal state associated with each of the two synaptic strengths beyond which the synapse state can only change by reversing induction stimulus to move in the other direction. In effect the hidden states are implementing a delay, a synapse in one of the deep states will have to wait for a number of stimuli to arrive before it can express plasticity. There are no

explicitly set of transition probabilities but assuming an equilibrium distribution of state occupancy only the top two among n states can encode the initial signal of memory.

Other versions of a multistate synapse (Amit and Fusi, 1994) associate each state with a small discrete step change of efficacy between $s \in -1 \cdots +1$ where the step size is determined by the total number of states n as $s_m = 2m/(n-1) - 1$. This last version is also referred to as the linear synapse and transitions between state can be made stochastic associating a probability q of accepting the transition indicated by an arriving induction stimulus. If the number of states reaches the limit $\sqrt{N}$ then the number of memories stored goes back to being proportional to the number of synapses N as was the case with non-discrete synapses in Hopfield networks.

With equal transition probabilities among states we obtain a uniform distribution of synapses and thus increasing the model size reduces the initial signal by a factor of $1/n$. Even with equal state transition probabilities the model expresses $2n-1$ timescales of decay that give rise to approximate power-law forgetting dynamics (Leibold and Kempter, 2008). Lastly, Barrett and Van Rossum 2008 have shown that a band diagonal matrix, which would belong to a multistate synapse, is optimal in the sense that it maximizes Shannon information, but the optimal transition probabilities at the terminal states are somewhat modified resulting in a non-uniform distribution that is peaked at both ends and flat in the middle.

4.4.3.2 Cascade Synapse (Fusi et al., 2005)

The cascade model consists of a tower of hidden states attached below each strength state, see Figure 4.7. Transitions between states are stochastic and the cascade of states predict progressive stabilization by arranging a geometric progression leading to smaller transition probabilities towards the bottom of the cascade. The model attempts to prolong memory retention in palimpsest memories by exhibiting power-law decay dynamics. The serial arrangements of metaplastic states attempts to capture correlations in the plasticity stimuli experienced each synapse. We will be describing this model in more detail and conduct an analysis of its mean memory signal in the next Chapter, so we won't go into any more details here.

The cascade model predicts the progressive stabilization of synaptic changes may prolong memory lifetimes, but it assumes synapses may have access to a number of stability states. However, as we have discussed in Chapter 2, synaptic plasticity exhibits only a few stable states with the main two being e-LTP and l-LTP. We will be considering the transition to late-phase forms of plasticity in the second part of this thesis. We give a short review of existing models of late-phase plasticity after we introduce the general framework used to evaluate memory lifetimes with state-based synapse models.

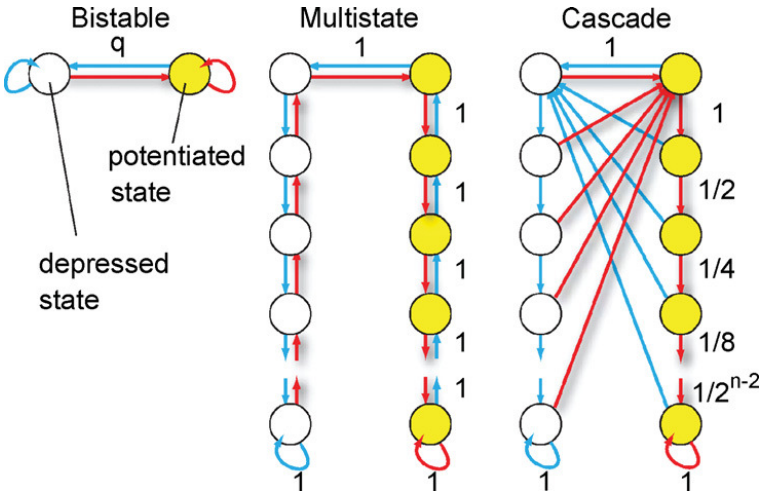


FIGURE 4.7: **Forms of discrete state synapses, the bistable, multistate and the cascade model.** The yellow (right) circle for each model represents a strong or potentiated synapse while the empty circles are the depressed or weak efficacy state of a synapse. Red arrows show transition where neural activity tends to potentiate the synapse. Blue arrows correspond to synaptic depressions and the numbers represent the conditional transition probabilities given the appropriate induction stimulus. The states at the bottom are the most resistant to plasticity. Source: Rubin and Fusi (2007)

4.5 Measuring synaptically stored memories

4.5.1 The general approach

Within the cascade framework the memory signal is measured by taking the view of an ideal observer able to read each synapse's state. Memory dynamics have been abstracted away from any particular neural network architecture. It is assumed that regardless of the network, the memories are stored at the synapse and no protective mechanism or re-storage exists. Thus, the information content in synapses is examined by looking at the number of memories that can be stored before the synaptic strength configuration that existed after initial memory storage is significantly different. The details of the required network dynamics that would allow for the synaptic state to be read are ignored and it is assumed that these can somehow be fully recovered.Therefore, memory lifetimes are measured by taking the view of an ideal observer that is able to "read" each synapse individually and track how initial traces of memory degrade due to storage of new memories or random plasticity events. The assumption is that synapses are embedded in an ongoing palimpsest memory of N_{syn} synapses that encodes memories on equilibrium synaptic state distributions. We therefore need a method to gauge the duration of a memory against the ravages of background activity caused by the storage of new memories.

To examine the memory dynamics, the method presented here tracks synapses depending on the stimulus they received upon encoding of a memory pattern. The tracked synapses belong to a memory which is no different to any of the patterns occurring in

the background. We assume the distribution of synaptic states is in equilibrium and choose the first encoded memory to track the changes in the synapses encoding it in order to gauge memory lifetimes. A fraction f_+ of the synapses experience a candidate potentiating (POT) stimulus and a fraction f_- a candidate depression (DEP) stimulus, $f_\pm$ being the probability of a POT or DEP stimulus occurring at any particular synapse with $f_- = 1 - f_+$.

The tracked memory signal is defined as the average distance separating the potentiated and depressed synapses from their equilibrium values (Fusi and Abbott, 2007). We let w_i denote the synaptic weight and $\bar{w}$ the average weight at equilibrium. A generic signal measure $S(t)$ for N_{syn} synapses can be defined as in (Fusi and Abbott, 2007) :

$$S(t) = S_{POT}(t) - S_{DEP}(t) \tag{4.11}$$

$$\tag{4.12}$$

where each group can be evaluated via the distance of the weight vector from the average weight $\bar{w}$:

$$S_{POT/DEP}(t) = \frac{1}{N_{syn}} \sum_{i \in POT/DEP} (w_i(t) - \bar{w}). \tag{4.13}$$

POT denotes the group of synapses that received a potentiating induction stimulus and DEP the group of all synapses that where depressed under the storage of a tracked memory. A specific version of such a signal will be examined in a later section for the cascade model, the key point is that the signal is a measure of a distance of a weight distribution from its equilibrium and to observe this fluctuation in the distribution the potentiated $S_{POT}(t)$ and the depressed group $S_{DEP}(t)$ of synapses need to be observed separately, which from now on we refer to as the P group and the D group respectively.

4.5.2 Signal to noise ratio

Memory lifetimes are computed by comparing a mean tracked memory signal against a general level of "noise". The signal can defined as the distance between the synaptic strength groups P and D as above. Noise for a stochastically updating synapse arises due to the inherent probabilistic nature of synaptic state changes, which are driven by random uncorrelated input stimuli and stochastic synaptic responses. The lifetime of a tracked memory signal relies on a subset of N synapses retaining their synaptic state in the face of ongoing plasticity. The signal needs to be detected despite fluctuations caused by background plasticity on the distributions of synapses retaining the signal, and therefore the lifetime of the signal is evaluated in relation this noise. A standard method is to evaluate the signal to noise ratio (SNR) to determine the detectability of the signal in due to its variability appearing as noise.

The "noise" due to fluctuations in synaptic strength caused by ongoing plasticity is proportional to $\sqrt{N_{syn}}$ and equal to the standard deviation of the signal (Fusi and Abbott, 2007) :

$$\sigma = \sqrt{\frac{1}{n}\sum_{i\in POT}(w_i - \bar{w})^2 - S^2_{POT} + \frac{1}{n}\sum_{i\in DEP}(w_i - \bar{w})^2 - S^2_{DEP}} \tag{4.14}$$

An approximate noise term can be estimated if we assume that the sequence of synaptic updates is independent. These can be modelled as a sum of N_{syn} independent Bernoulli trials. This sum gives rise to a binomial distribution with a standard deviation

$$\sigma = \sqrt{N_{syn}q(1-q)} \tag{4.15}$$

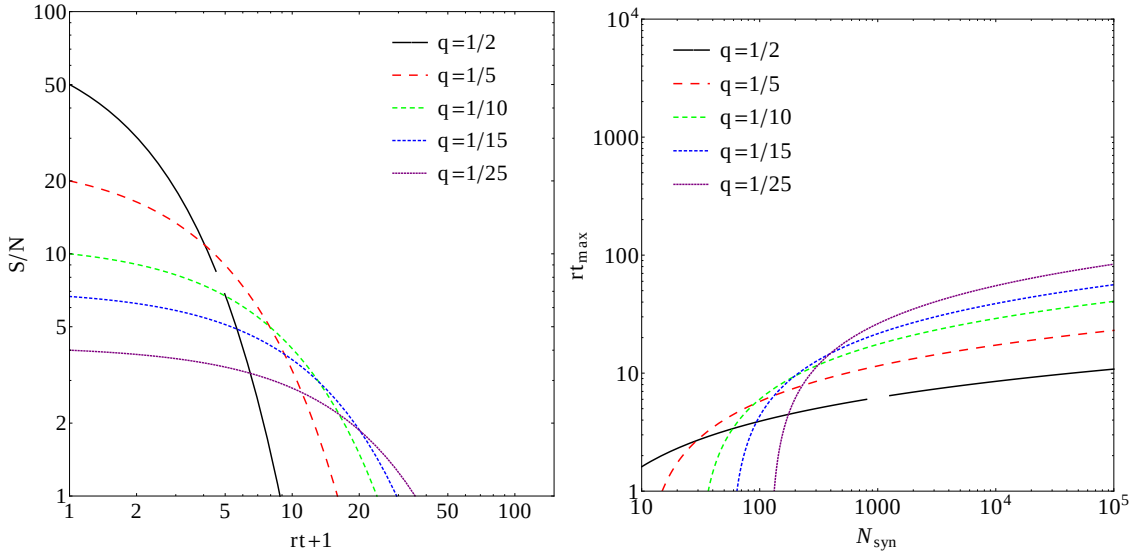
which means that $\sqrt{N_{syn}}$ could give an approximate noise term.

A lower signal-to-noise (SNR) threshold can be chosen under which the memory is assumed to be no longer retrievable. This is usually taken to be the point at which SNR drops to unity and the time when this occurs is taken as the memory lifetime t_{max} (Fusi et al., 2005).At the point when SNR= $mu(t)\sigma = 1$ the signal is one standard deviation away from a zero signal. As in Section 4.3.3, we take that at the limit of large population of synapses the distribution of the signal around the mean follows a Gaussian, then the probability that the signal is below zero is given by (4.9). When the erf is evaluated at $x/(\sigma\sqrt{2})$ then we obtain the probability that the signal is not further than x from the mean value. For SNR ¡= 1 we the signal needs to be at least σ distance away from the mean, and this occurs with a probability given by the reciprocal of the error function :

$$\mathbb{P}[error] = \frac{1}{2}[1 - \mathrm{erf}(\frac{\sigma}{\sqrt{2}\sigma})] \simeq 0.16,$$

at the time point when the SNR drops to unity then in approximately 16% of the times the memory signal would be found on or below zero. The general SNR analysis framework described above applies to to discrete state synapse models in general. A simple example would be to examine the memory lifetime of the binary synapse (Tsodyks, 1990) we summarized in an earlier section. Each synaptic strength is represented by an independent binary random variable and thus the sum of all synapses follows the binomial distribution with a standard deviation as in Equation (4.15). At the time of each memory storage the mean number of synapses changing state is $T_{syn} = qN_{syn}$. Under further memory storage with a mean rate of arrival rt, T_{syn} will exponentially decay with a rate qrt.

$$\mu(t) = qN_{syn}e^{-qrt} \tag{4.16}$$



(a) Signal against transition probability

(b) Maximum capacity against number of synapses N

FIGURE 4.8: The lifetime of the signal of a stochastic updater synapses increases as transition probabilities q become smaller. The initial signal however drops. This is also seen on the maximum capacity figure where the minimum N is increased as q becomes smaller. From the same figure we see that the capacity grows very slowly ($\approx \ln(\sqrt{N_{syn}})$) as the number of synapses increase resulting in very little gain from increasing the network size.

The time when the mean signal is equal to noise is :

$$q\exp(-qrt)N_{syn} = \sqrt{N_{syn}q(1-q)} \tag{4.17}$$

$$-qrt = \ln(\sqrt{N_{syn}q(1-q)}) - \ln(N_{syn}q), \tag{4.18}$$

solving for t gives the maximum mean memory lifetime :

$$t_{max} = \frac{\ln(N_{syn}q) - \ln(\sqrt{N_{syn}q(1-q)})}{qr} \tag{4.19}$$

$$rt_{max} = \frac{\ln(\sqrt{N_{syn}})}{\sqrt{(1-q)}}. \tag{4.20}$$

Showing that capacity is low, with a fixed probability q a pattern will not be retrievable after approximately $\ln(N_{syn})$ new patterns are subsequently stored.

Figure 4.8(a) shows the relationship between initial signal and memory lifetime as q becomes smaller. Fast learning rates result in large initial signal but shorter memory lifetimes. Increasing the size of the network would increase the capacity but in a manner proportional to the square root of the number of synapses $\sqrt{N_{syn}}$, see Figure 4.8(b). This last figure also shows that for low q a minimum network size is required, below which the memory lifetime is zero. The SNR memory-lifetime shows that to achieve long memory lifetimes the probability q needs to be accurately adjusted as a function

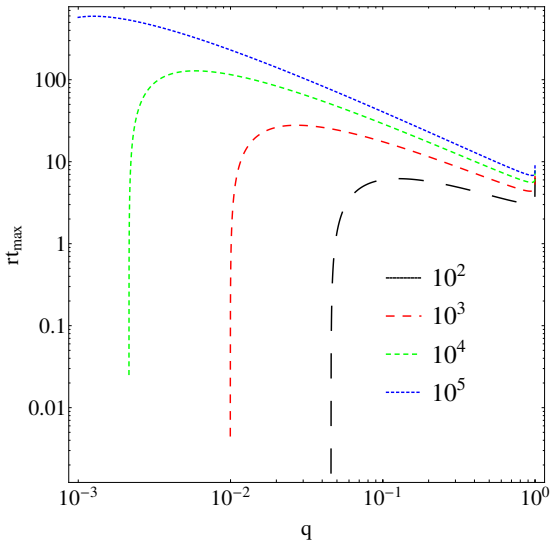


FIGURE 4.9: Maximum memory lifetime under a range of transition probabilities q for different network size $N_{syn} = 10^2, \cdots, 10^5$. Below a critical value where the initial signal qN is below the noise memory lifetime is zero, while the optimum q value sits close to the minimum value. Away from q optimum towards one we find that the network size has almost no effect on memory lifetimes. Therefore q needs to be tuned to network size.

of the number of synapses N_{syn} such that the signal at least stands above the noise (see Amit and Fusi, 1994). Figure 4.9 plots t_{max} against q for typical network sizes N_{syn} showing that optimum value is close to the minimum q. The implication of the simple SNR analysis above is that within biological networks q should be tuned to the number of synapses in the network.

4.5.3 The Perceptron signal

In this section we describe the standard signal that we will be using to gauge stored memory without explicitly having to track synaptic distributions but by reading the postsynaptic response of single formal neuron. Our approach was inspired by considering a standard Hopfield (Hopfield, 1982) network of N neurons and isolating a single neuron from that network (Elliott and Lagogiannis, 2012). Typically, such a network is required to store $\boldsymbol{\xi}^{\alpha}$ random uncorrelated memory patterns with $\xi_i^{\alpha} \in \{-1, +1\}$, where α is an index of a memory pattern in a sequence. Each ξ_i^{α} takes values of ± 1 with a probability set by $f_+ = 1 - f_-$, taken as $f_+ = 1/2$ unless otherwise stated. A value $\xi_i^{\alpha} = -1$ can be taken as a low input activity and $\xi_i^{\alpha} = +1$ as a strong firing activity. We assume that the output of this neuron during memory encoding is fixed to $+1$, thereby learning to classify all input patterns to a high output. Eventually we end up with a method that has also been previously used in Barrett and Van Rossum (2008) and in van Rossum et al. (2012), where the plasticity rule induces depotentiation to synapses receiving a low input $\xi_i = -1$ and potentiation to synapses receiving $\xi_i = +1$. The signal from memory

α is gauged by the weighted sum that defines the neuron's output :

$$h_x(t) = \frac{1}{N}\sum_{i=1}^{N} \xi_i^{\alpha} S_i(t), \tag{4.21}$$

where $S_i(t)$ denotes the strength of a synapse and since we will be considering only binary synapses $S_i(t) \in \{-1, 1\}$. Where ± 1 represents strong/weak synaptic strength.

In practice, memory lifetimes will be gauged by measuring the fidelity of recall to the first pattern stored ξ^1 as the $h(t)$ output. The signal will be measured at various time points after initial memory storage while new memories are being encoded with rate rt. This response to the tracked memory in particular can be written as $h_{\xi^1}(t)$ but for simplicity we will be referring to perceptron signal as the P signal whose measure is simply $h(t)$. For a pattern that has not been learned before, the output will be the sum of N random uncorrelated binary values $\{-1, 1\}$, which on average would be zero. If however a pattern is recognized then correlations between $\boldsymbol{\xi}^{\alpha}$ and $\boldsymbol{S}(t)$ would give rise to an output signal. Writing $\tilde{S}_i(t) = \xi_i S_i(t)$, we may denote the contribution of each synapse to the memory signal, since positive values would mean $\xi_i = S_i(t)$ and thus synapse is storing the memory and $\xi_i \neq S_i(t)$ would contribute negatively. Because all synapses are the same and will experience the same probabilistic induction stimuli during on-going memory storage, the $\tilde{S}_i(t)$ are identically distributed random variables. We write $\tilde{S}(t)$ to represent any one of these variables. Rewriting the perceptron output with $\tilde{S}_i(t)$ shows that the signal is just an average over N of these random variables:

$$h_x(t) = \frac{1}{N}\sum_{i=1}^{N} \tilde{S}_i(t). \tag{4.22}$$

For a sum of random variables, the signal's mean $\mu(t)$ and variance is defined as :

$$\mu(t) = \mathrm{E}[\tilde{S}(t)] \tag{4.23}$$

$$\sigma(t)^2 = \frac{1}{N}\mathrm{Var}[\tilde{S}(t)] + \left(1 - \frac{1}{N}\right)\mathrm{Cov}(t), \tag{4.24}$$

where $\mathrm{E}[\tilde{S}(t)]$ and $\mathrm{Var}[\tilde{S}(t)] = 1 - \mu(t)^2$ denote mean and variance and $\mathrm{Cov}(t)$ denotes the covariance between any pair of synapses. In simulations of N synapses the variance is computed as standard by $\mathrm{Var}(t) = \mathrm{E}[h(t)^2] - \mathrm{E}[h(t)]^2$ and the covariance is computed by $\mathrm{Cov}(t) = \mathrm{Var}(t) - (1 + \mathrm{E}[h(t)]^2)/N$. In a later section we will be presenting an analysis of the mean signal dynamics. Analytically calculating the covariance requires to consider any pair of synapses undergoing synchronous updates to obtain $\mathrm{Cov}(t) = \mathrm{E}[\tilde{S}_1(t)\tilde{S}_2(t)] - \mathrm{E}[\tilde{S}_1(t)]\mathrm{E}[\tilde{S}_2(t)]$. For two synapses undergoing updates $\mathrm{E}[\tilde{S}_1(t)\tilde{S}_2(t)]$ the transition matrix is obtained via the tensor product $\mathbb{M}_2 = \mathbb{M}_1 \otimes \mathbb{M}_1$. In continuous time with Poisson distributed encoding events, the generating matrix for a single synapse is $\mathbb{G}_1 = \mathbb{M}_1 - \mathbb{I}$, and thus the probability of synaptic updates goes as $P(t) = \exp(rt(\mathbb{G})P(0))$

while for a pair of synapses we need to obtain a matrix for the synchronous updates of two synapses via the tensor product of the transition matrices $\mathbb{G}_2 = \mathbb{M}_1 \otimes \mathbb{M}_1 - \mathbb{I} \otimes \mathbb{I}$, noting that $\mathbb{G}_2 \neq \mathbb{G}_1 \otimes \mathbb{G}_1$.

In this section, we have discussed a generic memory signal to use in order to evaluate mean memory lifetimes instead of specifically tracking synapses to measure distances between distributions as discussed in the previous section. Nevertheless, the underlying principles governing the evolution of this signal also rely on distances of distributions from equilibrium and we will also use such methods in our analysis of the mean signal $\mu(t)$.

4.6 Models of late-phase plasticity

In our earlier review of synaptic plasticity across three animal models of memory we identified two main forms of plasticity that differed in the stimulation requirements and on the basis of a protein-synthesis requirement. The early form (e-LTP) lasts hours while the expression of stable late-phase form of plasticity (l-LTP) lasts from hours to days in vitro and months in vivo and requires protein synthesis (Krug et al., 1984, Manahan-Vaughan et al., 2000). The late-phase requires specific stimulation conditions implicating neuromodulatory processes (Frey et al., 1990, Huang and Kandel, 1995, Swanson-Park et al., 1999, Sajikumar and Frey, 2004a). Repetition protocols with sufficient pause intervals along with neuromodulatory stimulation provide optimal storage conditions leading to late-phase plasticity. At the cellular level neuromodulatorary stimulation has been shown to synergistically activate the necessary signalling pathways for l-LTP. For memory systems with afferent neuromodulatory input these can be signals carrying emotional or reward significance controlling how synapses process incoming plasticity stimuli (Schwaerzel et al., 2003, Lisman and Grace, 2005, Frey and Frey, 2008, Seol et al., 2007). Since the conditions for the induction of long-term plasticity are different to the earlier forms then models of long-term memory should take the transition to this stable state under special consideration.

Existing models that explicitly consider the transition to late-phase plasticity have focused on formal descriptions of the STC hypothesis (Clopath et al., 2008, Barrett et al., 2009, Päpper et al., 2011) addressing the interaction between weak and strong stimulations. These models vary in the amount of detail at which they describe input patterns but do not explicitly deal with the problem of detecting stimulus patterns. On the other hand, most models and theories of temporal pattern selectivity have proposed explanations that rely on molecular pathway dynamics whose activation timing correlates with the optimal repetition intervals for l-LTP without explicitly addressing the underlying mechanisms (Ajay and Bhalla, 2004, Pagani et al., 2009, Philips et al., 2007, Kim et al., 2010). In a later part of this thesis we will be addressing the question of how synapses

detect stimulation patterns that lead to late-phase plasticity. Here, we present existing models of late-phase plasticity but since these models do not explicitly consider the detection of stimulation patterns we confine this section to a short overview of existing late-phase plasticity models.

4.6.1 The TagTriC-Model

On STC specifically, the first abstract model proposed is the *TagTriC-Model* (Clopath et al., 2008); it uses Hebbian learning to induce bidirectional plasticity in a postsynaptic voltage dependent manner. According to this model LTD is expressed when a recent time average postsynaptic activity exceeds a lower threshold while LTP is expressed when the momentary value of the postsynaptic potential exceeds an upper threshold. In both cases internal state variables indicating a "tag" are set and protein synthesis is initiated when the total number of set tags exceeds a threshold. This threshold is modulated by the levels of DA in this model. Reaching a sufficient level of protein synthesis can potentially allow a stability variable to change so that synapses enter a late-phase of plasticity. This model responds differently to strong and weak tetanic stimulations and was shown to account for cross-tagging (Sajikumar and Frey, 2004a) but the strength of the protocol is determined solely by the number of induction stimuli interacting close in time. If stimuli are too far out in time they will not interact to sufficiently elevate activity to threshold but such mechanism pays no heed to optimal temporal patterns for strong stimulation protocols.

4.6.2 Barrett, Billings, Morris, and Van Rossum (2009)

Barrett, Billings, Morris, and Van Rossum (2009) proposed a six-state continuous time model to investigate the effects of STC in network modelling. Although the model does not explicitly focus on detecting stimulus patterns it does take into consideration that l-LTP is induced after a strong stimulation pattern consisting of regularly spaced bursts of induction stimuli. The synapse model consists of a set of three states attached to each strength containing a *basal*, an *early* and a *late-phase* plasticity state. The e-LTP/D states are entwined with the "tag" and therefore e-LTP/D expression and the setting of a tag are linked (see however Redondo and Morris, 2010, for a review suggesting this link should be revised). The stimulation protocol modifies the probabilities of stochastic transitions towards stable states and these can subsequently passively occur in the absence of any further stimuli. For LTP two types of stimuli are considered, a weak HFS representing a single burst of high frequency stimulation and a *strong* HFS, which consists of three *weak* HFS separated by ten minute intervals. The HFS stimulus is predefined and directly results in modulating transition probabilities, the stimulus frequency or strength is not explicitly modelled. The transition probability from basal

to early/"tagged" state is modelled with alpha function dynamics with a peak-time tuned to 10 minutes taken from relevant experimental evidence (Sajikumar and Frey, 2004b). These dynamics are initiated in response to weak HFS in a synapse specific manner. The alpha function time-to-peak also matches the interval time between the three stimulus repetitions that define a strong HFS protocol in this model. A strong HFS results in initiating the dynamics of the transition probability from early to late plasticity states but in contrast to a weak HFS this probability applies over all synapses of a cell. This global modification of the transition probability towards late-phase plasticity-states allows this model to exhibit heterosynaptic interactions in accordance with STC hypothesis.

Here the detection of the strong protocol is taken for granted. There is no explicit mechanism to determine when sufficient repetitions of a weak HFS have occurred that would constitute a strong HFS. Nevertheless, the authors discuss the issue and propose that integrating the induction stimuli up to threshold would provide a detection mechanism of a minimum number of repetitions allowing the model to independently detect when sufficient repetitions have occurred that constitute a strong HFS protocol. Critically however, with this model repetition protocols with small inter-repetition time-intervals (massed) would be as efficient to spaced protocols. Therefore, both massed and spaced protocols could trigger processes of protein synthesis that initiate cell-wide transitions to late-phase plasticity.

4.6.3 Päpper, Kempter, and Leibold (2011)

Päpper, Kempter, and Leibold (2011) combined ideas from the two previous models in a discrete-time state based STC model. The model synapse has five states representing e-LTP/D and l-LTP/D and a neutral state acting as a reservoir of available synapses while transitions among states are conditioned on the stimulation protocol. Here also, unimportant and important memory patterns are assumed to result from weak and strong stimulation protocols and no explicit mechanism to detect them exists.

4.6.4 Smolen, Baxter, and Byrne (2006)

A biophysical model involving more than ten differential equations was proposed by Smolen et al. (2006) to address the main molecular interactions believed to underlie l-LTP and STC. In contrast to earlier abstract models of plasticity, the potential of this biophysical model to reproduce a spacing effect was also evaluated. The authors report that spaced (10-60 minute intervals) protocols produced only a slight enhancement of l-LTP compared to massed protocols (1 minute intervals). However, when spaced stimuli where grouped into two bursts delivered with pauses of 16 hours the amount of expressed l-LTP after 24 hours was doubled. This enhancement was due to the long

pauses overcoming a depletion of a necessary *precursor protein* which occurs beyond four burst in this model. Essentially, depletion of one pathway component implements a mechanism of a saturated response by assuming a lower bound of *precursor protein* in this model. A later revision simplified the above model to propose a mechanism of how activity may contribute to the long-term maintenance of l-LTP in the shadow of molecular turnover (Smolen, 2007) but did not address temporal patterns of stimulation.

We therefore conclude that existing theoretical models of l-LTP focused on the STC hypothesis do not simultaneously address the question of how synapses detect temporal patterns of stimulation that optimize storage conditions.

4.6.5 Zhang, Liu, Heberton, Smolen, Baxter, Cleary, and Byrne 2012

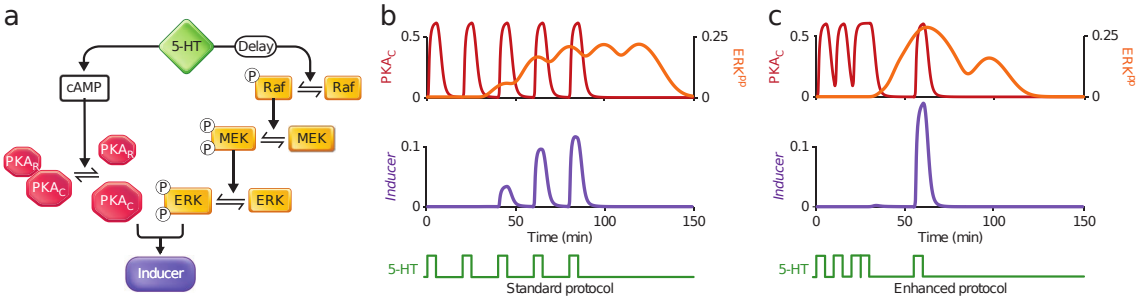


FIGURE 4.10: **a.** Schematic of interacting components in a signalling cascade starting from 5-HT that stimulates two pathways one leading to production of cAMP that activates the various isoforms of PKA and another resulting in stimulating ERK **b.** Standard stimulation protocol and the dynamics of the PKA and resulting inducer. **c.** The enhanced protocol consists of five repetitions with pause intervals of 10,10,5 and 30 minutes. Source: (Zhang et al., 2012)

A further biophysical model of the dynamics of the early biochemical cascades believed to underlie the induction of late-phase plasticity in Aplysia sensory neurons is presented. The model focuses on the dynamics of PKA and the extracellular signal-regulated kinase (ERK) cascades whose involvement in LTM across animal models of learning we discussed in Chapter 2. Its aim is to use the dynamics of the signalling pathway to provide a mechanism as to why spaced protocols are more effective than massed repetition protocols at producing LTM.

There are 13 equations describing the dynamics of the various biochemical interactions following stimulation by 5-HT that lead to the upregulation of variable named *inducer*, see Figure 4.10 (a). The inducer is taken to be a proxy for the interaction between PKA and ERK that lead to late-phase plasticity. The model's dynamics are used to evaluate the efficacy of stimulation protocols by comparing how the timing between stimulus repetition intervals modify the peak values of inducer.

This publication evaluates a standard spaced stimulation protocol (Figure 4.10 (b)), known to lead to stable late-phase plasticity in vitro. The protocol consists of five 5-HT pulses (lasting 5 minutes each). These are delivered to sensory neurons in culture with a pause interval of 20 minutes and have been shown to enhance the activation of CREB but also the synaptic efficacy measured using electrophysiology on the sensorimotor synapse. The authors then compared these results against the dynamics of the model shown on Figure 4.10 (b), and found agreement in the sense that such a protocol could stimulate the inducer. Further, the dynamics of the model are used to discover a more efficient protocol, which is named *enhanced*, that instead of fixed repetition time intervals it has variable timing between stimulations and results in enhancing the inducer's peak, as shown on 4.10 (c), but was also shown to enhance LTF and the levels of CREB activation. A delay in the activation of the Raf is an essential component in making the enhanced pattern of activation more effective as it delays the activation of ERK; an initial strong stimulation would make the ERK rise and then a final repetition that needs to wait long enough to hit the peak of the ERK is required for strong inducer activation.

A consensus on the *final cause* of the spacing effect of memory (the purpose of that the spacing is serving) could perhaps be that it attempts to capture stable properties of the environment. Models of molecular signalling pathways, offer what could be classified as a *material cause* that could give an explanation as to what properties of the substrate are responsible for the memory spacing phenomenon but however do not provide explanations on the *formal cause*, as to what function or process is this cascade of signals performing. In a later part of this thesis, we aim to develop a model of the spacing effect for late-phase plasticity that is developed so as to overcome limitations faced by individual synapses in determining the form of the stimulation protocol.

4.7 Summary

In this chapter we gave an overview of prominent artificial neural networks used to model memory. The review was directed towards presenting the issues that led to the development of the synaptically stored memory framework we will be working with in this thesis. Single formal neurons have been initially used to model logic and then as classifiers of input vectors with the perceptron learning. Further developments considered associative memory in larger networks where one pattern could be used as cue to invoke a learned output response pattern. The Hopfield network can operate as an autoassociative memory, where an incomplete input pattern can be used as a cue to bring the network in a self-sustaining firing pattern with the respective output pattern at its output. The output pattern recalled is the one closest matching the cue pattern from the patterns stored in weight matrix since each stored pattern forms an attractor state in the recall dynamics of the network. The original Hopfield learning rule allows

unconstrained growth of synaptic strength because each learning event produces just a linear change in the synaptic weight matrix. The memory however collapses after a critical limit on the number of memories stored is reached due to overloading. One solution for overcoming the overloading issue was found by bounding synapses to constrict their growth below a particular value. However, the capacity of the network is dramatically reduced but bounding creates an automatic forgetting mechanism. Indeed, if we assume that synapses cannot take arbitrary small steps in synaptic efficacy and introduce bounds in synaptic strength then a population of synapses is constrained to lie within a distribution of a limited number of states. When we bound synapses the linear learning rule that kept each learning event independent in the Hopfield network is broken. Automatic forgetting with bounded synapses then occurs because each memory introduces a fluctuation in the equilibrium distribution of synaptic states that interferes with the fluctuations produced by previous memories (see Fusi and Senn, 2006). Binary synapses are an extreme case of bounded synaptic strength existing only in two strength states. Bounding synapses within particular limits appears as a simple solution to the problem of catastrophic forgetting, but reduces capacity dramatically. Stochastic learning has been proposed as a way to extend memory lifetimes. Stochastic learning achieves a reduction on the interference between new memories and previous ones by randomly selecting synapses to encode a new memory. By reducing the probability q that a synapse will be selected for an update we are effectively slowing down the rate of learning while keeping the strength of a learning step the same. This mechanism then presents a new dilemma. A slow learning rate by reducing q increases the capacity and thus the lifetime in a palimpsest memory, but it reduces the initial strength of the trace since the number of synapses encoding the new trace is on average qN (Fusi and Abbott, 2007). Therefore, we have a manifestation of a stability-plasticity dilemma, slow learning increases stability but reduces plasticity and vice versa. To circumvent this issue models of synapses have been extended to include "hidden" states that do not affect synaptic efficacy but rather modify the probability of plasticity. Still synapses are confined to exist within a limited number of strength-states and thus the same issues on on the lifetimes of fluctuations on weight distributions apply. However, hidden states represent metaplasticity that modifies the probability of plasticity in order to break the symmetry between the speed of learning and forgetting to address the stability-plasticity dilemma (Abraham and Bear, 1996). A prominent probabilistic synapse-model of metaplasticity is the cascade model (Fusi et al., 2005) and it will form the main focus of the next chapter where we describe its main properties and analyse its dynamics.

Finally, in this chapter we reviewed models that consider the transition to late-phases of plasticity and found that these models do not explicitly address the question of how synapses detect optimal storage conditions known to lead to late-phase plasticity.

Chapter 5

Cascade models of synaptically stored memory

In the previous chapter we discussed how memory systems are formally modelled, the arising issue of catastrophic forgetting and how plasticity models have attempted to circumvent it by bounding synapses to create memory palimpsest (Nadal et al., 1986). These memory palimpsest however face a dilemma between expressing plasticity to encode new memories or remaining stable to retain previously stored ones (Abraham and Robins, 2005) and we reviewed synapse models that attempt to resolve this dilemma. In this chapter we focus on the cascade model for synaptically stored memories proposed by Fusi et al. (2005). The cascade model extends the method of stochastically updating synapses to resolve the plasticity-stability dilemma by assuming synapses have access to progressively lesser degrees plasticity in order to provide for both fast learning and slow forgetting.

This ability is believed to be biophysically supported by the various biochemical synaptic pathways; a hypothesis that has appealed to biologists as it relates to the observed progressive stabilization of changes in synaptic efficacy (Kandel, 2009) and the related metaplastic phenomena.

Although the cascade has been shown not to be optimal in terms of memory capacity (Barrett and Van Rossum, 2008, Leibold and Kempter, 2008) we focus on it in this thesis because we aim to establish the merits of computation before plasticity expression against a stochastic view of synaptic plasticity. The cascade in particular has become a prominent metaplasticity model that attempts to prolong memory lifetimes under a stochastic view of synaptic plasticity. It extends the ideas of the simpler stochastic synapse by relying on synapses having access to progressively lower transition probabilities. We begin with a short history on the development of this model to show that it follows as an extension to the stochastic view of synaptic plasticity and proceed to analyse its structure and function in detail.

5.1 Introduction

As we discussed in the previous chapter, memory encoding on palimpsest memories with discrete-state synapses requires changing the strength states of the synapses to increase their correlation with the memory vector being encoded. In the limit of two strength states we may separate synapses between those that are strong and weak and between those that a tracked memory requires them to be strong P or weak D. For perfect correlation of the weight to the tracked memory all the distribution of synapses in P must be all in the strong state, while for the D they must be in the weak state, but before memory storage these are equally distributed between strong and weak synapses. Thus, encoding a memory requires the introduction of a fluctuation that would increase the occupancy of strong synapses in P bringing the distribution of synapses that should be strong away from equilibrium and respective fluctuation in the D group. The size of the initial fluctuation determines how strongly a memory has been encoded or as otherwise stated the *memory signal* $\mu(t)$. In summary, for each encoded memory we have a distribution of synapses that should be weak D and one for the synapses that should be strong P. When these distributions return to equilibrium the memory is lost and the signal has decayed to zero.

Stochastically updating synapses (Tsodyks, 1990, Amit and Fusi, 1992, 1994) express a fraction of the plasticity inducing stimuli they receive according to a probability q; we will refer to these synapses as the *stochastic updater* (SU). Tuning the probability of synaptic plasticity expression q is actually handling the dilemma of *stability versus plasticity* (Abraham and Robins, 2005) but in a rather naive way by letting each synapse randomly select which plasticity stimulus it will respond to, and therefore to the encoding of which memory it will participate in. In a population of synapses with low q, each memory encoding affects a small fraction of synapses. Reducing q decreases the size of the fluctuation encoding any particular memory but it also increases the time it takes for the palimpsest distributions to return to equilibrium. To recover the size of the initial fluctuation the memory would have to be repeatedly encoded and thus a reduced q effectively slows the learning rate and the forgetting rate simultaneously by dilating time. Conversely, a large initial signal would result in fast decay but also in a shorter memory lifetime and therefore q sets both the learning and the forgetting rate simultaneously. We have therefore a trade-off between initial signal and signal decay rate (Fusi and Abbott, 2007, Barrett and Van Rossum, 2008). Balancing new memory encoding over erasure of older memories ultimately determines the capacity of the memory (see Abraham and Robins, 2005, Fusi and Senn, 2006). How can the symmetry between fast learning and fast forgetting be broken to introduce fast learning and slow forgetting?

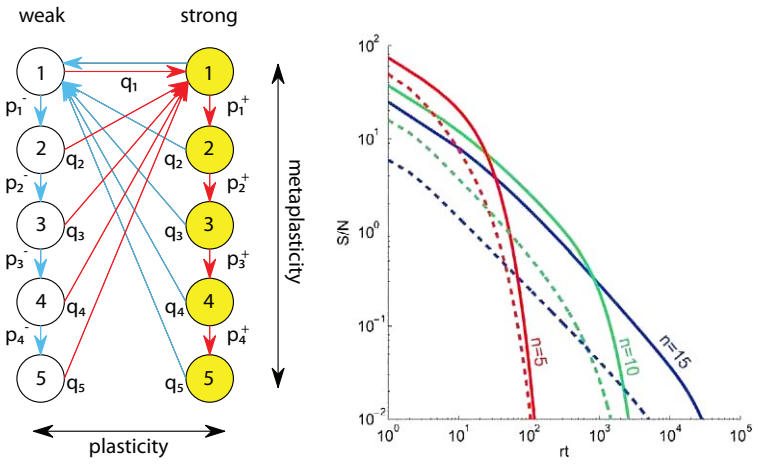
The cascade model emerges as a natural continuation in the development of those earlier ideas. The cascade synapse proposes a model by which a binary strength synapse may progressively move to states with smaller transition probabilities. It aims to break the

symmetry between fast learning and fast forgetting by introducing a variety of transition probabilities. The states with high q give a high initial fluctuation in the distribution while the states further down the cascade make synapses more resistant to changes in efficacy, see Figure 4.7. As synapses move between these states the rate at which plasticity is expressed is affected. At each state the cascade model uses stochastic updating which means that transitions between states is probabilistic and conditioned upon the induction stimulus. The states down the cascade attempt to model synaptic metaplasticity according to which the ability of synapses to express plasticity depends on the history of induction events (Abraham and Bear, 1996). To this end, the cascade structure is believed to represent the biochemical cascades found in synapses that are responsible for metaplasticity (Fusi et al., 2005, Kandel, 2009). Consequently, the cascade replaces the exponential forgetting dynamics of stochastic updating synapses that have access to a single transition probability q by proposing that memory performance relies on the complexity of synapses to introduce multiple timescales for learning and forgetting q_i that result in power-law memory retention dynamics. The progressive stabilization of plasticity attempts to capture correlations in the perceived train of induction stimuli by each synapse so persistent stimuli are retained longer.

We proceed to describe the cascade model and the method proposed to measure the signal from its synaptically stored memory, followed by a spectral decomposition of the cascade signal. We then present a method from the original authors for obtaining the cascade's mean signal numerically, a method which loses the noise information and finally, we present an exact mean signal analysis from our publication Elliott and Lagogiannis (2012) that uses a master equation approach that retains the fluctuations information and thus allows an investigation of the noise.

5.2 Overview of the model

The model is shown on Figure 5.1(a), it consists of n hidden states attached to each of the two strength states with the allowed transitions between states shown with arrows. We sometimes refer to the set of states that belong to the strong strength as the strong cascade and the other set of n states as the weak cascade. These hidden states associated to each synaptic strength model metaplasticity; moving down the cascade does not alter synaptic strength but only the associated transition probabilities. These transitions have associated probabilities q_i and p_i for moving horizontally and vertically between the states respectively. Metaplastic transitions from state $i \to i{+}1$ occur with probability p_i conditioned on the direction of the plasticity induction. The plasticity stimuli need to agree with the strength state, either a strong synapse receiving a potentiation (POT) plasticity stimulus or a weak synapse a depression stimulus (DEP). In the opposite case, when a plasticity stimulus directs a synapse towards the opposite strength, then with probability q_i a cascade synapse in state i may express plasticity by moving to the top of



(a) (b)

FIGURE 5.1: **a**. A sample cascade of n=5 size showing metaplasticity vertically and conventional synaptic plasticity horizontally between two strengths. Weak states are denoted with − and strong with +. Transitions from the weak cascade to the strong occur given a POT stimulus with probability q_i. If the synapse is in a state in of the strong cascade, a DEP stimulus will cause a transition to the weak cascade with probability q_i. Transitions $p_i^{\pm}$ link the states within the weak (-) and strong (+) cascade and correspond to metaplastic transitions. **b**. Tracking the signal to noise ratio of a memory signal stored at time t=0 in the cascade (solid line). These are compared against a bistable synapse (dashed line) with a heterogeneous transitions probabilities q equal to each of the q_i for each the cascade state. The plot shows that memory lifetime improvement is not just from having multiple timescales of plasticity but also due to the structure of the cascade. Source: (Fusi et al., 2005).

the opposite cascade $i \rightarrow 1$. Notably, plasticity expression is accompanied by a change in stability. The synapse will change from stable to labile as it moves from the bottom to the top of the opposite cascade. In fact, all plasticity expressing transitions via q_i end up at the top of the opposite cascade (see Figure 5.1(a).

The cascade model aims to break the symmetry between the speed of acquisition of memory signal and its degradation by introducing a range of q_i and p_i transition probabilities. The top of the cascade is responsible for the bulk of the initial signal as it contains the most labile synapses, while the states deep in the cascade primarily increase memory lifetimes by being resistant to plasticity. To obtain both fast and slow rates in a finite number of states the transition probabilities need to quickly drop, and for this reason the range of transition probabilities decreases exponentially x^i fast among states. Clearly, at any one time a single synapse expresses the set of q_i and p_i depending on the cascade state it exists in. For the full range of transitions probabilities to be equally available then a population of synapses should be uniformly distributed among all cascade states. To maintain a uniform distribution the rate of entering a cascade state needs to match the rate of exit. Starting from the highest rate at the top of the cascade $q_i = 1$ the transition probabilities are arranged in a geometric sequence with

$x = 1/2$ so rates half as we go down the cascade. For plastic transitions we have:

$$q_i = x^{i-1} \text{ for } i \in \{1, \cdots, n-1\} \tag{5.1}$$

$$q_n = x^{n-1}/(1-x), \tag{5.2}$$

where at q_n the probabilities are set to so as to compensate for the boundary effect in the absence of p_i transitions. For metaplastic transitions we have :

$$p_i^{\pm} = x^i/(1-x) \text{ for } i \in \{1, \cdots, n-1\}. \tag{5.3}$$

Throughout, the superscript $\pm$ refers to strong/weak synapse strength and its equivalent cascade, thus p_i^+ refers to metaplastic transitions on the right of Figure 5.1(a).

Each memory trace consists of synapses associated with a range of transition probabilities distributed across states. The synapses at the top of the cascade have lifetimes similar to an equivalent bistable synapse with probability q_1, these allow for fast learning. The synapses occupying deeper states in the cascade have slower dynamics, since $q_i < q_1$, and these aim to prolong memory lifetimes.

It is assumed that memories arrive with a rate r and each memory encoding delivers plasticity inducing stimuli at each synapse; either potentiation with probability f_+ or depression $f_- = 1 - f_+$. Therefore, synaptic strengthening is expressed with rate rf_+q and weakening at a rate rf_-q. Naturally, modifying the stimulus rates $1/f_\pm$ under a fixed structure of cascade transition probabilities would affect the distribution of synapses, with $f_+ > f_-$ potentiation events would push more synapses over the strong cascade and vice versa. Since all of the model's states need to be occupied in a population N_{syn}, retaining uniform distribution so that the full range of transition probabilities is equally available requires adjusting the transition rates. Transitions down the cascade only occur through p_i transitions and therefore adjusting these (5.3) in a manner that reflects the balance between stimuli would maintain the uniform occupancy across states. Thus, to remedy the effects of stimulus imbalance on the state occupancy, the transition probabilities must be calculated in a rate depended manner as:

$$p_i^{\pm} = f_{\mp}x^i/f_{\pm}(1-x). \tag{5.4}$$

Such mechanisms imply that synapses are able to somehow adapt their progression of probabilities in response to balance in plasticity stimuli. Although, Fusi et al. (2005) have shown via Monte Carlo simulations that vary the cascades p_i and q_i that the precise progression is not critical, stimuli imbalances do effect the model's capacity and thus synapses need to somehow adjust apparently all cascade probabilities simultaneously. In the next section, we discuss the framework proposed by Fusi et al. (2005) to obtain a signal measure representing the strength of the trace in synaptically stored memories.

5.3 "Mean Field" analysis and results

The method reproduced in this section is described in Fusi et al. (2005), it permits the evaluation of memory dynamics in continuous-time by evaluating the evolution of mean occupancies numerically, thus avoiding the requirement to average over many trials of a full simulation that considers each synapse individually. Although stated as a mean field, the approach essentially describes the continuous-time dynamics of the mean signal using occupation numbers, the description of the mean dynamics is exact but all information about fluctuations around these mean dynamics is lost. In a later section we use a master equation approach to analyse the signal in a manner that retains the probability distribution of the stochastic variables and thus the statistics of the signal.

Here, the memory dynamics are evaluated as averages of state occupancies and not by evaluating transitions of each synapse individually. With the default transition probabilities, the equilibrium cascade state occupancies are uniform with $1/2n$ for each state. These occupancies are in a dynamic equilibrium, the distribution remains constant but individual synapses constantly change state in response to new memories. The equilibrium state occupancies for the strong and weak cascade are denoted by :

$$\sum_{i=1}^{n} F_i^{\pm} = \sum_{i=1}^{n} F_i^{+} + \sum_{i=1}^{n} F_i^{-} = 1. \tag{5.5}$$

Memory patterns are represented by random uncorrelated plasticity events occurring at all synapses simultaneously. These encoding events may modify individual state occupancies but there is no overall observed change as the system is in dynamic equilibrium. The dynamics of the system will be described using transition equations. For the lowest cascade state $i = 1$ where all q_i transitions converge we have:

$$\frac{dF_1^{\pm}}{dt} = r\left(f_{\pm}\sum_{j=1}^{N} q_j F_j - (f_{\pm}p_1^{\pm} + f_{\mp}q_1)F_1^{\pm}\right), \tag{5.6}$$

the intermediate state occupancies evolve as :

$$\frac{dF_i^{\pm}}{dt} = r\left(f_{\pm}p_{i-1}F_{i-1} - (f_{\pm}p_i^{\pm} + f_{\mp}q_i)F_i^{\pm}\right), \tag{5.7}$$

while in the last state where we have no further p transitions we write:

$$\frac{dF_n^{\pm}}{dt} = r\left(f_{\pm}p_{n-1}F_{n-1}^{\pm} - f_{\mp}q_n F_n^{\pm}\right) \tag{5.8}$$

We cannot track the disturbance in the distribution of $\boldsymbol{F}^{\pm}$ because both stimuli with rates f_+ and f_- are applied and thus the occupancies always appear in equilibrium. We may track the effects of f_+ and f_- stimuli individually by splitting the occupancy

$\sum_{i=1}^{n} F_i^{\pm} = 1$ in two subgroups.

$$F_i^{\pm} = P_i^{\pm} + D_i^{\pm},$$

$P_i^{\pm}$ represents those synapses that during memory storage where potentiated and $D_i^{\pm}$ those that where depressed. We let $P_i^{\pm}$ and $D_i^{\pm}$ be components of vectors $\boldsymbol{P}^{\pm}$ and $\boldsymbol{D}^{\pm}$ respectively. We denote the joint distribution that includes the synaptic strength as $\boldsymbol{P} = \left((\boldsymbol{P}^{-})^T|(\boldsymbol{P}^{+})^T\right)$ and $\boldsymbol{D} = \left((\boldsymbol{D}^{-})^T|(\boldsymbol{D}^{+})^T\right)$ for the potentiated and depressed group respectively.

When the stimuli rates are balanced $f_+ = f_-$ then the distributions of each subgroup $\boldsymbol{P}^{\pm}$ and $\boldsymbol{D}^{\pm}$ will be mirror images of each other and will evolve in the same way. Each subgroup consists of synapses in the strong and the weak cascade. The $P^{\pm}$ group represents the fraction f_+ of synapses that were potentiated and $D^{\pm}$ the fraction f_- that where depressed by memory storage (received a POT or DEP induction stimulus):

$$\sum_{i=1}^{n}(P_i^{+} + P_i^{-}) = f_+$$
$$\sum_{i=1}^{n}(D_i^{+} + D_i^{-}) = f_-$$

$F_i^{\pm}$ is the distribution of occupancies of the whole synaptic population N_{syn}. Before memory storage the $D_i^{\pm}$ and $P_i^{\pm}$ subgroup occupancies are in equilibrium, $\boldsymbol{P}^{+} = \boldsymbol{P}^{-}$ and $\boldsymbol{D}^{+} = \boldsymbol{D}^{-}$. After memory storage, the distribution of each subgroup will be moved away from equilibrium. By dividing the occupancy into two subgroups we can track the change in the equilibrium distributions caused by a particular memory storage event. Storing the memory requires inducing a potentiating stimulus to the $P^{\pm}$ subgroup under the following transition equations:

$$P_1^{+} \to P_1^{+} + \sum_{j=1}^{n} q_j P_j^{-} - p_1^{+} P_1^{+} \tag{5.9}$$
$$P_1^{-} \to P_1^{-} - q_i P_1^{-} \tag{5.10}$$

and for other cascade indexes $1 < i \leq n$ we have

$$P_i^{+} \to P_i^{+} + p_{i-1}^{+} P_{i-1}^{+} - p_i^{+} P_i^{+} \tag{5.11}$$
$$P_i^{-} \to P_i^{-} - q_i P_i^{-} \tag{5.12}$$
$$P_n^{+} \to P_n^{+} + p_{n-1}^{+} P_{n-1}^{+} \tag{5.13}$$

and a depression stimulus to the $D_i^{\pm}$ subgroup with:

$$D_1^- \rightarrow D_1^- + \sum_{i=1}^{n} q_j D_j^+ - p_1^- D_1^- \tag{5.14}$$

$$D_1^+ \rightarrow D_1^+ - q_i D_1^+ \tag{5.15}$$

while for other cascade indexes $1 < i \leq n$ we have

$$D_i^- \rightarrow D_i^- + p_{i-1}^- D_{i-1}^- - p_i^- D_i^+ \tag{5.16}$$

$$D_i^+ \rightarrow D_i^+ - q_i D_i^+ \tag{5.17}$$

$$D_n^- \rightarrow D_n^- + p_{n-1}^- D_{n-1}^- \tag{5.18}$$

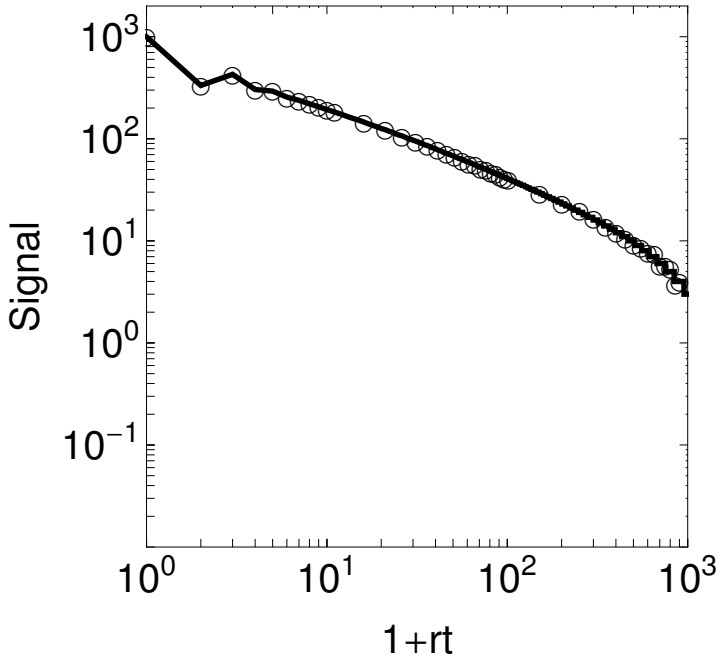
The initial memory storage step will take the $\boldsymbol{P}^{\pm}$ and the $\boldsymbol{D}^{\pm}$ away from the uniform equilibrium distribution. The $\boldsymbol{P}^+$ subgroup contains all synapses that were potentiated and metapotentiated after memory storage. Because the subgroup $\boldsymbol{P}^{\pm}$ had only POT induced, $\boldsymbol{P}^+$ (the strong state) will have an excess number of occupied states compared to equilibrium, see Equation (5.9). The mirror argument applies to subgroup $\boldsymbol{D}^{\pm}$, $\boldsymbol{D}^-$ has higher occupancy than in equilibrium due to DEP only induction stimuli at the memory storage step, see Equation (5.14). Let $p_{\infty}^{\pm}$ denote the fraction of strong (+) and the weak (-) synapses in equilibrium $p_{\infty}^{\pm} = \sum_{i=1}^{n} F_i^{\pm}$. The memory signal is then measured by comparing the distribution of occupancies on each group against the equilibrium :

$$Signal = N_{syn}(\sum_{i=1}^{n} P_i^+ - f_+ p_{\infty}^+ + \sum_{i=1}^{n} D_i^- - f_- p_{\infty}^-), \tag{5.19}$$

where $f_+ p_{\infty}^+$ gives the fraction of strong synapses that would have been in the P^+ group prior to memory storage state and $f_- p_{\infty}^-$ the respective fraction in the $\boldsymbol{D}^-$ group. After memory storage the subgroups are subjected to random new patterns of activity that result in signal decay as each subgroup is drawn back to equilibrium. Strictly speaking , $\boldsymbol{P}^+$ contains synapses that had received a potentiating stimulus and may become strong at any point in time after memory storage. Critically however, although the equations are not explicitly confined to track the synapses that initially stored the memory, the authors of the model explicitly specify they are tracing the memory signal contained in the initial changes of synaptic strength. This is because memory storage causes a fluctuation in the tracked vectors $\boldsymbol{P}^+$ and $\boldsymbol{D}^-$ whose ultimate demise the cascade is trying to prolong as there is no mechanism to revive or augment the fluctuation. We will refer to a signal measured as above that only tracks the decay of the initial signal as the F-signal. Here, the $\boldsymbol{P}^+$ and $\boldsymbol{D}^-$ can only be occupied by synapses that formed the initial signal and are tracked thereafter.

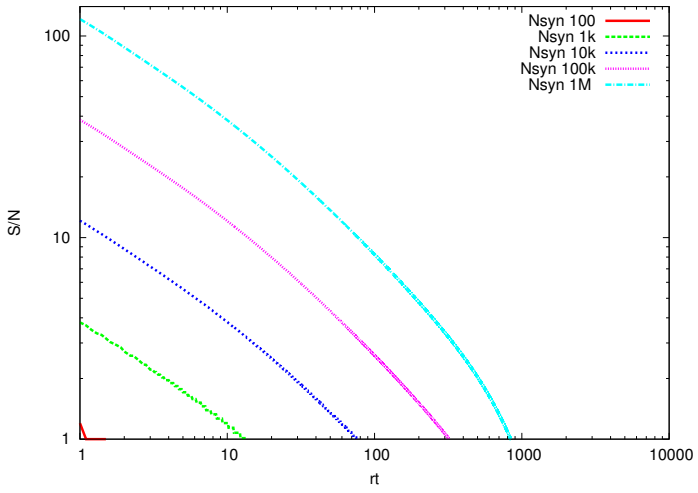
The above system was evaluated numerically in discrete-time and compared against a full simulation which measures the signal in the same way as above but averaged over

$T = 10^3$ trials. In Appendix 2 we outline the simulation algorithm for a cascade in continuous-time from which the discrete-time time simulation can be derived letting the timesteps be fixed increments. Figure 5.2 shows results of "mean field" simulations (in discrete-time) compared against the mean signal from a full simulation of $N_{syn} = 10^4$ synapses, showing that the two signals match. Using this method we have essentially numerically computed the evolution of the signal stored as fluctuations in the $\boldsymbol{P}$ and $\boldsymbol{D}$ distributions without having to average over many trials of a full simulation. On the same figure we observe a small kink in the early time evolution of the signal. This kink is apparent in discrete-time evaluation of the signal and is not a numerical artefact. The discrete-time simulation can be checked against a repeated application of the stochastic matrix $\mathbb{M}$ on the vector of the initial signal distributions $\boldsymbol{P}^+$ and $\boldsymbol{D}^-$ to find that this kink is due to the structure of the matrix. Checking the eigenvalues of $\mathbb{M}$ we find that indeed there is the eigenvalue λ_{2n-1} changes sign from negative for odd powers of $\mathbb{M}^n$ to positive on even powers giving rise to an up and down trend. However, and in agreement with the Frobenius-Perron theorem (see Section 3.2), we find a highest eigenvalue of unity that is retained with increasing powers n while the rest of the eiqenvalues in the spectrum decrease as the distributions return to equilibrium.

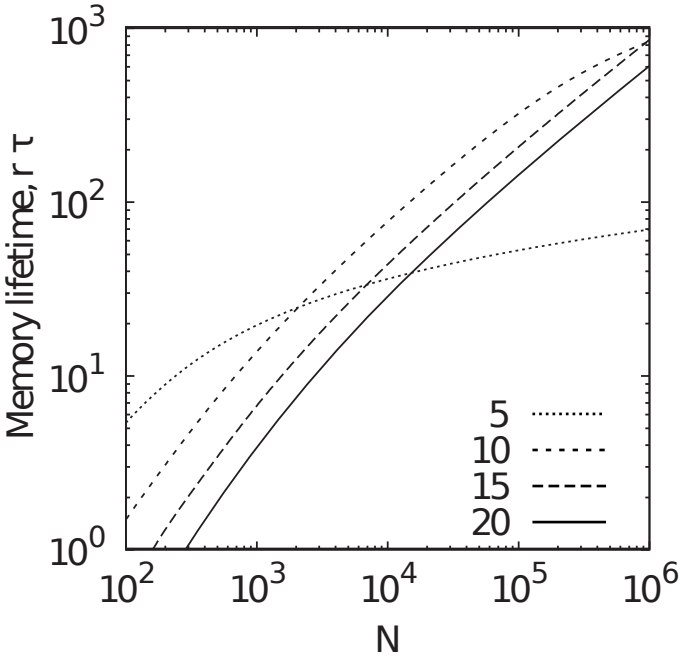


(a) n=10

FIGURE 5.2: Comparing "mean-field" memory signal to mean signal (F) obtained via simulation for a $n = 10$ cascade with $N = 10^4$ and $T = 10^3$ trials in discrete-time. The open circles are from simulation results showing exact agreement to "mean-field" signal evaluation method. The kink in the signal is only apparent in discrete-time and is not a numerical artefact but due to the structure of the stochastic matrix (see text) .



(a) Memory signal $n = 10$ per N_{syn}



(b) Mean memory lifetime per cascade size n and network size N

FIGURE 5.3: **a**. The effect of population size N_{syn} on cascade's initial SNR and memory lifetime for $n = 10$ (F signal). N_{syn} used for the plots are 10^1 (bottom left),$10^2, 10^3, 10^4, 10^5, 10^6$. Initial signal and memory lifetime increase with the number of synapses N_{syn}. **b**. Dependence of memory lifetime on number of synapses for a few indicative cascade sizes. The rt_{max} is the point where SNR $= 1$ on plot 5.3(a), the units are in $1/r$. There appears to be an optimal cascade size per network size.

Since the mean dynamics do not contain the fluctuations, the SNRs is approximately obtained by taking the noise as the standard deviation of the binomial distribution:

$$N = \sqrt{N_{syn} p_\infty^+ p_\infty^-}. \tag{5.20}$$

Increasing the number of synapses has the effect of reducing the noise term and therefore results in increasing SNRs. Figure 5.3(a) shows how SNRs lifetimes of an $n = 10$ cascade extend when increasing N_{syn} from 10^2 to 10^6. We may define the maximum memory lifetime rt_{max} at the point when SNRs=1. Figure 5.3(b) plots rt_{max} against N_{syn} for different cascade sizes n. This figure reveals that there is an optimal cascade size for different values of N_{syn}. For example, small cascades $n = 5$ give higher capacity over networks as small as $N_{syn} < 10^3$, beyond which point $n = 10$ becomes optimal. Nevertheless, the authors find that there is a weak dependence and no precise tuning is required for optimality (Fusi et al., 2005). The cascade size n alters the dynamics of the SNR as increasing the size of the size of the cascade extends the period of the power law dynamics over rt but reduces the initial signal, see Figure 5.1(b), exhibiting the known trade-off between initial SNR and memory lifetimes. But this extension is not only due to the collection set of available timescales but the arrangement of states in the cascade structure seems to also come into play. On Figure 5.1(b) the SNR for various cascade sizes is compared against a collection of heterogeneous bistable synapses that exhibit the same range of transition probabilities as a cascade. We observe that indeed significant differences begin to emerge as the cascade size increases and so memory extension in cascades in not simply due to the availability of multiple learning/forgetting rates but the sequential structure of stabilization operates to further extend memory lifetimes.

These results assume that the cascade is under a regime of balanced excitation and the cascade distribution is uniform on average. If $f_+ \neq f_-$ and the $p_i^{\pm}$ rates are not adjusted to the stimuli balance then the uniform distribution among cascade states is lost and memory lifetimes are reduced, see Figure **??**.

5.4 Spectral decomposition of mean dynamics

Here, we wish to decompose the cascade signal dynamics by considering the evolution of the state occupancies as a *Markov process*. Let $\mathbb{M}$ describe a discrete time transition matrix and $\boldsymbol{A}$ the joint distribution of strength and cascade state occupancies. Whenever a new memory is stored we have :

$$\boldsymbol{A} \to \mathbb{M}\boldsymbol{A}.$$

After multiple memories n we have :

$$\boldsymbol{A} \to \mathbb{M}^n\boldsymbol{A}.$$

This is a description of the mean process in discrete time. In continuous time the memories are presented at a fixed Poisson rate r and the evolution of cascade states can

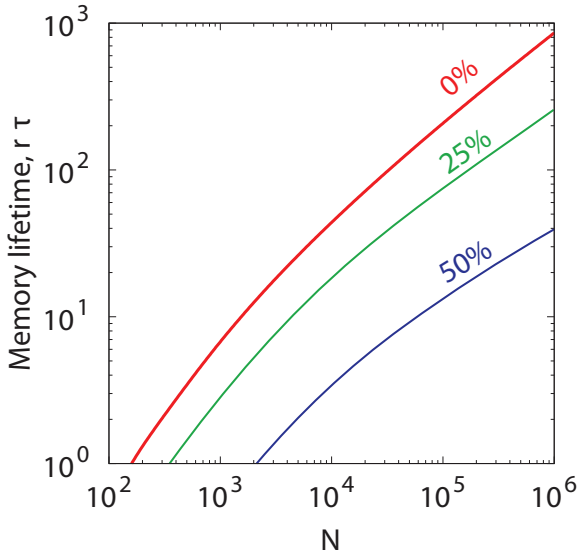


FIGURE 5.4: Unbalanced excitation of a cascade of size $n = 15$ reduces memory lifetimes if the transition probabilities are not adjusted accordingly. The percentages represent the percentage difference between f_+ and f_-. Adapted from :Fusi et al. (2005).

be studied by the master equation :

$$\frac{dA}{dt} = r(\mathbb{M} - \mathbb{I})\boldsymbol{A}, \tag{5.21}$$

while the transition matrix can be written :

$$\mathbb{A}(t) = \exp[(\mathbb{M} - \mathbb{I})rt]. \tag{5.22}$$

The components of $\mathbb{M}$ follow the same transitions as described in equations (5.6), (5.7) and (5.8), the transition probabilities are :

$$p^{\pm}_{i+1|i} = f_{\pm} p^{\pm}_i \tag{5.23}$$

$$p^{\pm}_{i|i} = 1 - f_{\mp} q_i - f_{\pm} p^{\pm}_i \tag{5.24}$$

$$p^{\pm}_{1|i} = f_{\pm} \sum_{i=1}^{n} q_i \tag{5.25}$$

$$p^{\pm}_{n|n} = -f \mp q_n \tag{5.26}$$

$$p^{\pm}_{j|i} = 0 \text{ otherwise} \tag{5.27}$$

where $i \in \{1, \cdots, n\}$ denoting states of an n size cascade with the positive superscript denoting the strong cascade and the $(-)$ the weak cascade. The positive states $(+)$

occupy the left side of the $\mathbb{M}$ and the negative $(-)$ side is on the right half, following the indexes of the state occupancy $\boldsymbol{A} = (\boldsymbol{A}^-|\boldsymbol{A}^+)$:

$$\boldsymbol{A} = \left(A_n^-, A_{n-1}^-, \cdots, A_1^-, A_1^+, \ldots, A_{n-1}^+, A_n^+\right)^T. \tag{5.28}$$

We write $\mathbb{G} = (\mathbb{M} - \mathbb{I})$ as the generating matrix of this process.

We know that the signal from bistable synapse gives an exponential decay while cascade synapses with $n > 1$ give power-law decay dynamics. In this analysis we would like to decompose the dynamics of the $\mathbb{G}$ to examine the components of the mean signal for some cascade size $n > 1$. We take a cascade of $n = 5$ states and denote state occupancy A_i starting from the bottom right of the cascade which is the strong and moving anti-clockwise, see Figure 5.1(a).

We denote $\frac{dA_i^\pm}{dt}$ as $\dot{A}_i^\pm$ for a strong $\dot{A}_i^+$ and a weak synapse $\dot{A}_i^-$ respectively. For a cascade of size $n = 1$ without metaplastic states and thus no p_i transitions we may directly write the coefficients of the generating matrix $\mathbb{G}$ as:

$$\mathbb{G}\boldsymbol{A} = \begin{pmatrix} \dot{A}_1^- \\ \dot{A}_1^+ \end{pmatrix} = \begin{pmatrix} -f_+q_1 & f_-q_1 \\ f_+q_1 & -f_-q_1 \end{pmatrix} \begin{pmatrix} A_1^- \\ A_1^+ \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} A_1^- \\ A_1^+ \end{pmatrix}, \tag{5.29}$$

since $q_1 = q_n = x^{n-1}/(1-x)$ this essentially gives bistable synapse.

For larger matrices, it helps to separate the transition matrix in potentiating and depressing transitions :

$$\mathbb{M} = \mathbb{M}^+ + \mathbb{M}^-,$$

where $\mathbb{M}^+$ is the original matrix with only the potentiating transitions ($f_- = 0$) and $\mathbb{M}^-$ has only the depression transition ($f_+ = 0$). Thus, we may write down the potentiating transition matrix $\mathbb{M}^+$ for a $n = 5$ cascade as:

$$f_+ \begin{pmatrix} 1-q_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-q_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-q_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-q_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-q_1 & 0 & 0 & 0 & 0 & 0 \\ q_5 & q_4 & q_3 & q_2 & q_1 & 1-p_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_1 & 1-p_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_2 & 1-p_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_3 & 1-p_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_4 & 1 \end{pmatrix} \tag{5.30}$$

and take $\mathbb{M}^-$ to be the mirror of the above along with replacing f_+ with f_-.

The probability distribution of cascade state occupancy for the potentiated and depressed group of synapses immediately after the storage of the tracked memory at time $t = 0$ is obtained by :

$$\boldsymbol{P} = \mathbb{M}^{+}\boldsymbol{A} \tag{5.31}$$

$$\boldsymbol{D} = \mathbb{M}^{-}\boldsymbol{A}, \tag{5.32}$$

where $\boldsymbol{D} = (\boldsymbol{D}^{-}|\boldsymbol{D}^{+})$ denotes the joint distribution weak/strong synapses that received a depression stimulus immediately after memory storage and $\boldsymbol{P} = (\boldsymbol{P}^{-}|\boldsymbol{P}^{+})$ those that received a potentiation stimulus. The memory signal is encoded as a fluctuation on $\boldsymbol{P}$ and $\boldsymbol{D}$. The above two distributions $\boldsymbol{P}$ and $\boldsymbol{D}$ are mirror images of each other, $\xi_i^1 = -1$ stimuli shift synapses of $\boldsymbol{D}$ over to the weak strength $S = -1$ and $\xi_i^1 = +1$ stimuli push over to the strong $S = +1$. The F-signal measures the signal as the sum of the distance of $\boldsymbol{P}^{+}$ and the $\boldsymbol{D}^{-}$ from equilibrium (5.19). With the P- signal however, the memory signal is measured via :

$$h(t) = \frac{1}{N}\sum_{1}^{N}\xi_i^1 S_i(t) \equiv \frac{1}{N}\sum_{1}^{N}\tilde{S}_i(t),$$

but we can still relate this signal to the $\boldsymbol{P}$ and $\boldsymbol{D}$. The size of the $h(t)$ output signal relates to the difference in the balance of positive and negative terms $\tilde{S}_i(t) = \pm 1$. All strengths $S_i(t)$ are viewed via a product with the tracked memory ξ_i^1 and thus weak synapses $S_i(t) = -1$ in $\boldsymbol{D}$ contribute positively to the memory signal because they are multiplied by $\xi_i^1 = -1$ when summed, but strong synapses in $\boldsymbol{D}$ would contribute negatively. The converse is true for $\boldsymbol{P}$. This conversion of the summed $S_i(t)$ effectively makes the right half of the $\boldsymbol{P}$ vector and the left half of the $\boldsymbol{D}$ vector contribute positively to the signal. At the same time, the left half of $\boldsymbol{P}$ and the right half of $\boldsymbol{D}$ contribute negatively. Since the $h(t)$ output is a difference between positive and negative contributions then the difference between the vectors $\boldsymbol{P} - \boldsymbol{D}$ is sufficient to measure the signal out of $h(t)$ exactly. Since, the F-signal measures only positive contributions, if we were to normalize it by the number of synapses N then it would give a scaled version of the mean P-signal $\mu(t)$ by a half.

Next, we perform a spectral decomposition of the generator matrix $\mathbb{G}$ into its eigenvectors and eigenvalues as described in the methods of Section 3.8.1, with $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_2 n)$ being the diagonal matrix of eigenvalues which for $n = 5$ there are $2n$ states with eigenvalues :

$$\boldsymbol{\lambda} = (-1.371, -0.815, -0.593, -0.365, -0.277, -0.130, -0.125, -0.125, -0.069, 0)\,.$$

There is a repeated eigenvalue of $\lambda = -0.125$, which means that the eigenvalue is degenerate. In this case we may still proceed with the decomposition assuming (3.54) still holds with minor side-effects expecting that some of the decay timescales will not

be visible (see Section 3.8.1). The largest eigenvalue is the zero one, whose associated eigenvector gives the steady-state distribution; since this eigenvalue is zero this eigenvectors contributions do not change over time according to Equation (5.34) and thus the associated eigenvector gives the long-term distribution as $t \to \infty$.

Next, with $\mathbb{Q}$ the matrix of right eigenvectors, the evolution of state occupancies of the potentiated group of synapses is decomposed as :

$$\boldsymbol{P}(t) = \mathbb{Q} \exp(\Lambda r t) \mathbb{Q}^{-1} \boldsymbol{P}(0), \tag{5.33}$$

where $\boldsymbol{P}(t)$ describes the evolution of the probability distribution of the potentiated group, these dynamics are shown on Figure 5.5(b).

We can now proceed to analyse the spectrum of the total signal. Memory is encoded as a fluctuation in these distributions which fades under ongoing memory as synapses eventually leave the potentiated or depressed group.

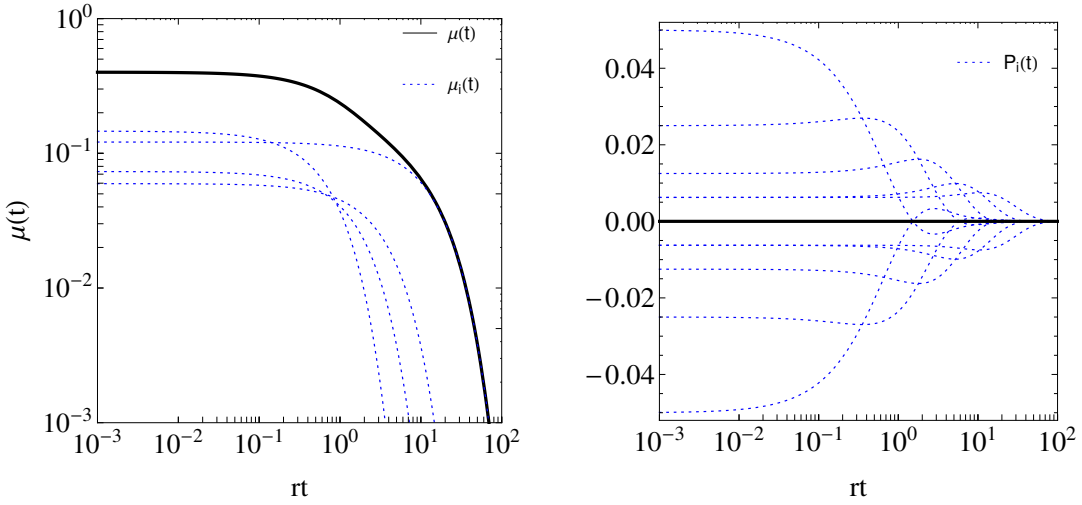
We use the eigendecomposition of the generator matrix (3.52) and the decomposition of the matrix exponential above to examine how each component contributes to the total signal. We let $\boldsymbol{s} = (-1, -1, \cdots, +1, +1)$ be an auxiliary vector of $2n$ components representing the synaptic strength associated with each cascade state. Each eigenvalue and vector pair determines a single exponential component i of the total signal $\mu(t)$:

$$\mu_i(t) = \exp(\lambda_i t) \boldsymbol{e}_r^i (\boldsymbol{e}_l^i)^T . (\boldsymbol{P} - \boldsymbol{D}) . \boldsymbol{s}. \tag{5.34}$$

Figure 5.5(a) shows the decomposition of the signal in $\mu_i(t)$ exponential signal components for an $n = 5$ cascade. There are four exponential components that combine to give the total signal and result in power-law decay dynamics. The figure only reflects the fact that the cascade combines multiple timescales due to the different transition probabilities of each embedded stochastic updater. During the early phases of the signal slow and fast exponentials combine to give the cascade its power-law like decay dynamics, while for longer $rt > 10$ the slowest component dominates the signal decay dynamics and these become a decaying exponential.

5.5 Analysis of mean signal

The cascade model's authors have analysed the evolution of the mean dynamics giving a description of the mean F-signal as discussed in Section 5.3. We will reproduce a derivation which has been published in Elliott and Lagogiannis (2012) for an exact analytical solution for the mean memory signal $\mu(t)$ of a cascade model using a master equation approach and renewal arguments (see Section 3.6). This analysis will permit to capture all the statistics of the signal by integrating out internal synaptic states to



(a) Spectral decomposition of total signal

(b) Evolution of $\boldsymbol{P}$ state occupancies

FIGURE 5.5: **a**. Spectral decomposition of a P signal in a cascade model of size $n = 5$ synapses. Four exponential components contribute to the total signal, their superposition gives a small region of power-law decay. **b**. Memory encoding causes the synapses in the two groups P and D to move away from the equilibrium uniform cascade distribution, here we observe the evolution of cascade occupancies of the P group. Memory encoding increases the occupancy of strong synapses in the $\boldsymbol{P}$ and decreases the occupancy of weak ones. The occupancy of the most labile states experiences the highest changes while they decay faster.

obtain the transitions between strengths. Such methods have been previously developed in Elliott (2010b) (2010b) and used to analyse mean expression times of synaptic filter models in (Elliott, 2011a) (2011a).

Here, when referring to *cascade* we mean the sequence of n states attached to each strength state. Thus, there are two cascades, one associated with a strong $(+)$ synapse and one with weak $(-)$, each having $i \in \{1, \cdots, n\}$ states. It follows that in a cascade transitions are strength state dependent. This means that the induction stimulus does not solely determine the direction of movement in state space $i \in \{1, \cdots, n\}$. The cascade state i of a synapse may increment only if synapse's strength state agrees with the induction stimulus received. Assuming transitions are deterministic ($p = q = 1$) then the state i of synapse j would increment when :

$$\begin{aligned} \text{with } \xi_j = +1 \text{ and } S_j(t) = +1, i \to i+1 \\ \xi_j = -1 \text{ and } S_j(t) = -1, i \to i+1 \end{aligned}$$

, otherwise it could escape towards the opposite strength cascade :

$$\begin{aligned} \xi_j = -1 \text{ and } S_j(t) = +1, i \to 1 \text{ and } S_j(t) \to -1 \\ \xi_j = +1 \text{ and } S_j(t) = -1, i \to 1 \text{ and } S_j(t) \to +1 \end{aligned}$$

then its state is reset $i \to 1$ and the strength switched.

As already discussed with the default transition probabilities the equilibrium distribution of state occupancies is uniform. We will use the vector $\boldsymbol{A}$ to denote the equilibrium distribution with elements $1/(2n)$.

We will now formulate *renewal process* as a description of the random synaptic transitions which will be later used to determine the evolution of the mean memory signal. We let $p_{i|j}^{A|B}(t)$ denote this probability density for a change of strength from B to A. We note here that $p_{i|j}^{A|B}(t)$ does not say anything of what happens during the transition from B to A in time t, the process may make any intermediate change. In this framework we then divide all possible paths that lead from j to i initially into two types, those that moved to state i without an intermediate change of strength and those that moved within the same cascade without any strength change.

For the first type, we let $f_{i|j}(t)$ denote the probability of moving from cascade state j to state i in time t without a change in strength. Then for the second type, we use $G_i^{\pm}(t)$ to denote the probability of the event of a first escape from an internal cascade state $i \in \{1, \cdots, n\}$ of the opposite strength (strong $+$/weak $-$) cascade by time t. After switching strength these synapses move to the target state with probability given again from the master equation but now the remaining time has been reduced by τ and thus $p_{i|j}^{A|B}(t-\tau)$.

By combining the above we can discuss the probability of moving from state j to state i and the accompanying strength change in continuous time t. We write down the following renewal equations governing transitions between cascade strength states:

$$p_{i|j}^{+|+}(t) = f_{i|j}(t) + \int_0^t p_{i|1}^{+|-}(t-\tau) G_j(\tau) d\tau, \tag{5.35}$$

$$p_{i|j}^{+|-}(t) = \int_0^t p_{i|1}^{+|+}(t-\tau) G_j(\tau) d\tau, \tag{5.36}$$

$$p_{i|j}^{-|+}(t) = \int_0^t p_{i|1}^{-|-}(t-\tau) G_j(\tau) d\tau, \tag{5.37}$$

$$p_{i|j}^{-|-}(t) = f_{i|j}(t) + \int_0^t p_{i|1}^{-|+}(t-\tau) G_j(\tau) d\tau. \tag{5.38}$$

$p_{i|j}^{+|+}(t)$ and $p_{i|j}^{-|-}(t)$ express the fact that to move from $j \to i$ a synapse can either directly move given $j < i$ without a strength change or it may escape to the opposite cascade only to return at some point and move to state i. Once a synapse switches strength then its state is set to the top of the cascade $i = 1$. From there the return process is identical to one of switching strength and moving from $1 \to i$ described by $p_{i|1}^{+|-}(t)$ or $p_{i|1}^{-|+}(t)$ and thus the process is renewed after each strength change. Consequently the probability density of moving from $j \to i$ of the same cascade either strong or weak introduce the inhomogeneous term $f_{i|j}(t)$ in the integral equations above. Then under the integral we

calculate the probability of escape from one cascade for the first time at any time τ with density $G_j^{\pm}(\tau)$ after which the process is renewed with time $t-\tau$ remaining to move from 1 to its final state i.

Because the integral equations above contain simple convolutions we choose to perform a Laplace transform on the above equations. We take s to be the transformed variable and denote the transformed functions with a hat over them :

$$\hat{p}_{i|j}^{+|+}(s) = \hat{f}_{i|j}(s) + \hat{p}_{i|1}^{+|-}(s)\hat{G}_j(s), \tag{5.39}$$

$$\hat{p}_{i|j}^{+|-}(s) = \hat{p}_{i|1}^{+|+}(s)\hat{G}_j(s), \tag{5.40}$$

$$\hat{p}_{i|j}^{-|+}(s) = \hat{p}_{i|1}^{-|-}(s)\hat{G}_j(s), \tag{5.41}$$

$$\hat{p}_{i|j}^{-|-}(s) = \hat{f}_{i|j}(s) + \hat{p}_{i|1}^{-|+}(s)\hat{G}_j(s). \tag{5.42}$$

The above equations are coupled in pairs, 5.39 with 5.40 and 5.41 with 5.42. We may determine the $\hat{p}_{i|j}^{+|\pm}$ couple by setting $j=1$:

$$\hat{p}_{i|1}^{+|+}(s) = \hat{f}_{i|1}(s) + \hat{p}_{i|1}^{+|-}(s)\hat{G}_1(s), \tag{5.43}$$

$$\hat{p}_{i|1}^{+|-}(s) = \hat{p}_{i|1}^{+|+}(s)\hat{G}_1(s), \tag{5.44}$$

$$\tag{5.45}$$

and the writing in matrix form

$$\begin{bmatrix} 1 & -\hat{G}_1(s) \\ -\hat{G}_1(s) & 1 \end{bmatrix} \begin{bmatrix} \hat{p}_{i|1}^{+|+}(s) \\ \hat{p}_{i|1}^{+|-}(s) \end{bmatrix} = \begin{bmatrix} \hat{f}_{i|1}(s) \\ 0 \end{bmatrix} \tag{5.46}$$

and solving for $\hat{p}_{i|1}^{+|\pm}(s)$:

$$\begin{bmatrix} \hat{p}_{i|1}^{+|+}(s) \\ \hat{p}_{i|1}^{+|-}(s) \end{bmatrix} = \frac{1}{1-\hat{G}_1^2(s)} \begin{bmatrix} 1 & \hat{G}_1(s) \\ \hat{G}_1(s) & 1 \end{bmatrix} \begin{bmatrix} \hat{f}_{i|1}(s) \\ 0 \end{bmatrix} \tag{5.47}$$

we obtain:

$$\hat{p}_{i|1}^{+|+}(s) = \frac{\hat{f}_{i|1}(s)}{1-\hat{G}_1^2(s)} \tag{5.48}$$

$$\hat{p}_{i|1}^{+|-}(s) = \frac{\hat{f}_{i|1}(s)\hat{G}_1(s)}{1-\hat{G}_1^2(s)}. \tag{5.49}$$

Accordingly we obtain the other couple of equations using the same steps as above :

$$\hat{p}_{i|1}^{-|-}(s) = \frac{\hat{f}_{i|1}(s)}{1 - \hat{G}_1^2(s)} \tag{5.50}$$

$$\hat{p}_{i|1}^{-|+}(s) = \frac{\hat{f}_{i|1}(s)\hat{G}_1(s)}{1 - \hat{G}_1^2(s)}, \tag{5.51}$$

which are identical to the $\hat{p}_{i|1}^{+|\pm}(s)$. Substituting the above equations for $j = 1$ into Equations (5.39), (5.40), (5.41) and (5.42) we obtain:

$$\hat{p}_{i|j}^{+|+}(s) = \hat{f}_{i|j}(s) + \frac{\hat{f}_{i|1}(s)\hat{G}_1(s)}{1 - \hat{G}_1^2(s)}\hat{G}_j(s) \tag{5.52}$$

$$\hat{p}_{i|j}^{+|-}(s) = \frac{\hat{f}_{i|1}(s)}{1 - \hat{G}_1^2(s)}\hat{G}_j(s) \tag{5.53}$$

$$\hat{p}_{i|j}^{-|+}(s) = \frac{\hat{f}_{i|1}(s)}{1 - \hat{G}_1^2(s)}\hat{G}_j(s) \tag{5.54}$$

$$\hat{p}_{i|j}^{-|-}(s) = \hat{f}_{i|j}(s) + \frac{\hat{f}_{i|1}(s)\hat{G}_1(s)}{1 - \hat{G}_1^2(s)}\hat{G}_j(s). \tag{5.55}$$

We define the probability of not changing strength in time t from cascade state j as:

$$H_j(t) = 1 - \int_0^t G_j(\tau)d\tau$$

and its Laplace transform :

$$\hat{H}_j(t) = \frac{1}{s}\left[1 - \hat{G}_j(s)\right], \tag{5.56}$$

but this probability is also defined as the sum of probabilities of no escape in time t starting from state j as :

$$H_j(t) = \sum_{i=1}^{n} f_{i|j}(t). \tag{5.57}$$

Using these results we can now derive an expression for the mean memory signal $\mu(t)$ of a tracked memory ξ^1 stored at time $t = 0$. We write the full transition matrix using submatrix notation where the submatrix $\mathbb{P}^{\mathbb{A}|\mathbb{B}}(s)$ contains the Laplace transformed expression determined above $\hat{p}_{i|j}^{A|B}(s)$:

$$\mathbb{P}(s) = \left[\begin{array}{c|c} \hat{\mathbb{P}}^{-|-}(s) & \hat{\mathbb{P}}^{-|+}(s) \\ \hline \hat{\mathbb{P}}^{+|-}(s) & \hat{\mathbb{P}}^{+|+}(s) \end{array}\right]$$

The right half of this matrix contains the probability of transitions from a synapse at the strong cascade and the left half the respective transitions starting from the weak cascade.

The mean signal arises as a fluctuation in each of the equilibrium distributions of

synapses that received a potentiating $\xi = +1$ stimulus or a $\xi = -1$ depressing stimulus. We let the joint probability distribution of strength and occupancy be the vector $\boldsymbol{P}$ for synapses that experienced $\xi = +1$ at time $t = 0$ and vector $\boldsymbol{D}$ for synapses that experienced $\xi = -1$. So the joint distribution of synapses that received $\xi = +1$ evolves as $\mathbb{P}(t)\boldsymbol{P}$. We can obtain the mean strength of these synapses by summing the distribution with the use of an auxiliary n-th dimensional vector $\boldsymbol{n}$ of $(1...1)^T$ as $(-\boldsymbol{n}|\boldsymbol{n})\mathbb{P}(t)\boldsymbol{P}$, where two $\boldsymbol{n}$ are concatenated in one with the sign denoting the strength of the synapse. When evaluating $(-\boldsymbol{n}|\boldsymbol{n})\hat{\mathbb{P}}(s)$ we find that in essence we are performing the following operations on the submatrices :

$$(-\boldsymbol{n}^T|\boldsymbol{n}^T)\mathbb{P}(s) = (\boldsymbol{n}^T\hat{\mathbb{P}}^{+|-}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|-}(s)|\boldsymbol{n}^T\hat{\mathbb{P}}^{+|+}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|+}(s)), \tag{5.58}$$

where the differences can be written component-wise as:

$$\sum_{i=1}^{n}\left[\hat{p}^{+|-}_{i|j} - \hat{p}^{-|-}_{i|j}\right] = \sum_{i=1}^{n}\left[\frac{\hat{f}_{i|1}(s) - \hat{f}_{i|1}(s)\hat{G}_1(s)}{1-\hat{G}_1^2(s)}\right]\hat{G}_j(s) - \sum_{i=1}^{n} f_{i|j}(t) \tag{5.59}$$

$$= \sum_{i=1}^{n}\left[\frac{\hat{f}_{i|1}(s)(1-\hat{G}_1(s))}{(1-\hat{G}_1(s))(1+\hat{G}_1(s))}\right]\hat{G}_j(s) - \hat{H}_j(s) \tag{5.60}$$

$$\sum_{i=1}^{n}\left[\hat{p}^{+|+}_{i|j} - \hat{p}^{-|+}_{i|j}\right] = \sum_{i=1}^{n}\left[\frac{-\hat{f}_{i|1}(s)(1-\hat{G}_1(s))}{(1-\hat{G}_1(s))(1+\hat{G}_1(s))}\right]\hat{G}_j(s) + \hat{H}_j(s) \tag{5.61}$$

Writing $\boldsymbol{G}(s)$ and $\boldsymbol{H}(s)$ as vectors with components $\hat{G}_j(s)$ and $\hat{H}_j(s)$, and by taking into account (5.57) we can write the above differences in a compact vector form :

$$\boldsymbol{n}^T\hat{\mathbb{P}}^{+|-}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|-}(s) = +\frac{\hat{H}_1(s)}{1+\hat{G}_1(s)}\hat{\boldsymbol{G}}(s) - \hat{\boldsymbol{H}}(s), \tag{5.62}$$

$$\boldsymbol{n}^T\hat{\mathbb{P}}^{+|+}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|+}(s) = -\frac{\hat{H}_1(s)}{1+\hat{G}_1(s)}\hat{\boldsymbol{G}}(s) + \hat{\boldsymbol{H}}(s). \tag{5.63}$$

We are now in position to define the Laplace transformed signal in terms of the initial synapse distributions after storage in vector notation. Note that the signal is stored on a fluctuation on the distribution of the potentiated $\boldsymbol{P}$ and depressed $\boldsymbol{D}$ synapses. We can measure the distance of each of these distributions from equilibrium as the difference in occupancy between strong/weak cascade but since $\boldsymbol{P}$ and $\boldsymbol{D}$ distributions are mirror images (see Section 5.3) and thus contribute to the signal $\hat{\mu}(s)$ equally we may calculate the signal using only the distance of $\boldsymbol{P}$ from equilibrium as :

$$\hat{\mu}(s) = \left(\boldsymbol{P}^+ - \boldsymbol{P}^-\right)\cdot\left[\hat{\boldsymbol{H}}(s) - \frac{\hat{H}_1(s)}{1+\hat{G}_1(s)}\hat{\boldsymbol{G}}(s)\right]. \tag{5.64}$$

This is because we assume equal rates of POT and DEP $f_+ = f_-$ while the $\tilde{S}_i(t)$ random variables are identically distributed. The q_i retain their usual meaning of being

the probabilities of changing strength when the synapse is in cascade state i. We use the fact that the metaplastic transition probabilities p_i^+ and p_i^- are equal where $p_i^{\pm} = q_i$ to rewrite Equation (5.64) from vector-form to an equation involving sums taking the default values for these as in the original cascade where $q_i = \frac{1}{2}^{i-1}$ for $i = 1, \cdots, n-1$ and $q_n = \frac{1}{2}^{n-2}$ (Fusi et al., 2005). The P group according to Equation (5.31) requires the $\mathbb{M}^+$ matrix being applied onto the equilibrium distribution $1/2n(\boldsymbol{n}|\boldsymbol{n})$, which could be also seen as summing of the rows of that matrix and scaling by $1/2n$. Equivalently, if we consider the difference $\boldsymbol{P}^+ - \boldsymbol{P}^-$ for an $n > 2$ and let $p_i^{\pm} = q_i$, we can write it as :

$$\boldsymbol{P}^+ - \boldsymbol{P}^- = \frac{1}{2n}\left(\sum_{i=1}^{n} q_i, \underbrace{q_1, \cdots, q_{n-2}}_{\text{n-2 terms}}, q_{n-1} + q_n\right)^T \tag{5.65}$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} 1, \underbrace{\frac{1}{2}, \cdots, \frac{1}{2}^{n-2}}_{\text{n-2 terms}}, \frac{1}{2}^{n-2}\right)^T \tag{5.66}$$

and thus we may write Equation (5.64) in component form as :

$$\begin{aligned}\hat{\mu}(s) =& \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2}^{(i-1)}\left[\hat{H}_i(s)\hat{G}_i(s)\frac{\hat{H}_1(s)}{1+\hat{G}_1(s)}\right] \\ &+ \frac{1}{n}\frac{1}{2}^{n-1}\left[\hat{H}_n(s) - \hat{G}_n(s)\frac{\hat{H}_1(s)}{1+\hat{G}_1(s)}\right].\end{aligned} \tag{5.67}$$

The last part of this equation considers only the escape from state n, evolving according to Equation (5.63) and showing that the return to equilibrium at the terminal state is governed by the difference in the rate of remaining in the strong cascade and the one escaping to the weak cascade. The contribution to the signal from all other states evolves in an identical manner as shown in the sum term containing a product of the Laplace transformed no-escape probability $\hat{H}_i(s)$ and the probability of first escape $\hat{G}_i(s)$ for each state i; these two events are mutually exclusive while the product defines a convolution in the time-domain that defines the decay of the signal from each of these states $1 < i < n$.

However, up to now we have used $\hat{G}_i(s)$ to denote the Laplace transformed densities of escape from state i towards the opposite cascade in response to an induction stimulus, but we have not yet derived $G_i(t)$. To do so we consider the elemental transitions occurring between states, beginning with the ones that lead to escape towards the opposite strength cascade, a transition which we refer to as *plastic.*A synapse in a state i receives a number of induction stimuli arriving at random time intervals. Depending on the strength state of the synapse and the direction of the induction stimulus the transition could result in an escape from a either state of the Weak/Strong cascade. This would require the arrival of an induction stimulus that will make a synapse move to the cascade of opposite strength. The arrival of these induction events are modelled as a Poisson

process having exponentially distributed time intervals with a mean rate $\frac{1}{2}r$ (assuming an equal rate of potentiating and depressing stimuli). Synapses are viewed here as stochastic devices that out of all arriving plasticity stimuli they select one at random. A state transition may occur with probability q_i, after a candidate plasticity stimulus has arrived. If selected with probability q_i then a transition of the synaptic state to the top of the opposite strength cascade $i \to 1$ occurs. We define the two processes of stimulus arrival and selection by a single probability density which takes into account the rate of stimulus arrival and the probability of responding to this stimulus :

$$\frac{dK_i^p(t)}{dt} = -rq_iK_i^p(t), \tag{5.68}$$

where the superscript p denotes that its the PDF of a plastic transition. We can solve the above equation using the boundary condition at $t = 0$ where the probability of a stimulus arrival and that this stimulus will be selected is $K_i^p(0) = rq_i$. Hence, the probability density of a plasticity event for a synapse in state $i < n$ is :

$$K_i^p(t) = \frac{1}{2}rq_i \exp(-rq_it). \tag{5.69}$$

However, because $p_i^{\pm} = q_i$ the densities of either accepting a plastic or a metaplastic transition are in cascade state $i < n$ are both equal and thus :

$$K_i^p(t) = K_i^m(t) = K_i(t). \tag{5.70}$$

For synapses in terminal states $i = n$, there are no metaplastic transitions and thus $K_n^p(t) = K_n(t) = \frac{1}{2}rq_n \exp(-rq_nt)$.

We can now write the probability density of changing strength for the first time starting from state i in time t denoted by $G_i(t)$, in terms of the probability density of selecting an incoming stimulus $K_i(t)$ as a recurrence relation :

$$G_i(t) = K_i^p(t) + \int_0^t K_i^m(\tau)G_{i+1}(t-\tau)d\tau. \tag{5.71}$$

The above relationship reflects that a synapse in state i may have changed strength at t by accepting an induction stimulus with probability $K_i^p(t)$ or it may have accepted a metaplastic stimulus with probability $K_i^m(t)$ and moved to state $i+1$ beyond which point unless $i = n$, the same argument on the probability of strength change is repeated in the remaining time $t - \tau$ as an identical process $G_{i+1}(t)$. For $i = n$, only plastic transitions are possible and so

$$G_n(t) = K_n^p(t). \tag{5.72}$$

By performing a Laplace transform of (5.71), we obtain :

$$\hat{G}_i = \hat{K}_i^p(s) + \hat{K}_i^m(s)\hat{G}_{i+1}(s) \tag{5.73}$$

giving a recurrence relation in the Laplace transforms of $\hat{G}_i(s)$, $i < n$ that can be rewritten :

$$\hat{G}_i(s) = \hat{K}_i(s)[1 + \hat{G}_{i+1}(s)], \tag{5.74}$$

by using the equality (5.70). The Laplace transform of (5.69) gives :

$$\hat{K}_i(s) = \begin{cases} \frac{r}{(2r+2^i s)} & \text{for } i < n \\ \frac{r}{(r+2^{n-1}s)} & \text{for } i = n \end{cases} \tag{5.75}$$

To simplify the above expressions we may write $\hat{G}_i(s) = X_i(s)/Y_i(s)$ and obtain the recurrence relation :

$$Y_i(s) = (2r + 2^i s)Y_{i+1}(s) \tag{5.76}$$

$$X_i(s) = r[X_{i+1}(s) + Y_{i+1}(s)] \tag{5.77}$$

subject to $Y_n(s) = r + 2^{2n-1}s$ and $X_n(s) = r$. Iterating $i \to n$ for $Y_i(s)$ we obtain products of the form $\prod_{k=i}^{n}(2r + 2^k s)$ that can be compactly written with the use of q-Pochhammer symbol $(\alpha; q)_n = \prod_{k=0}^{n-1}(1 - \alpha q^k)$ by taking the $2r$ term out of the parentheses and writing :

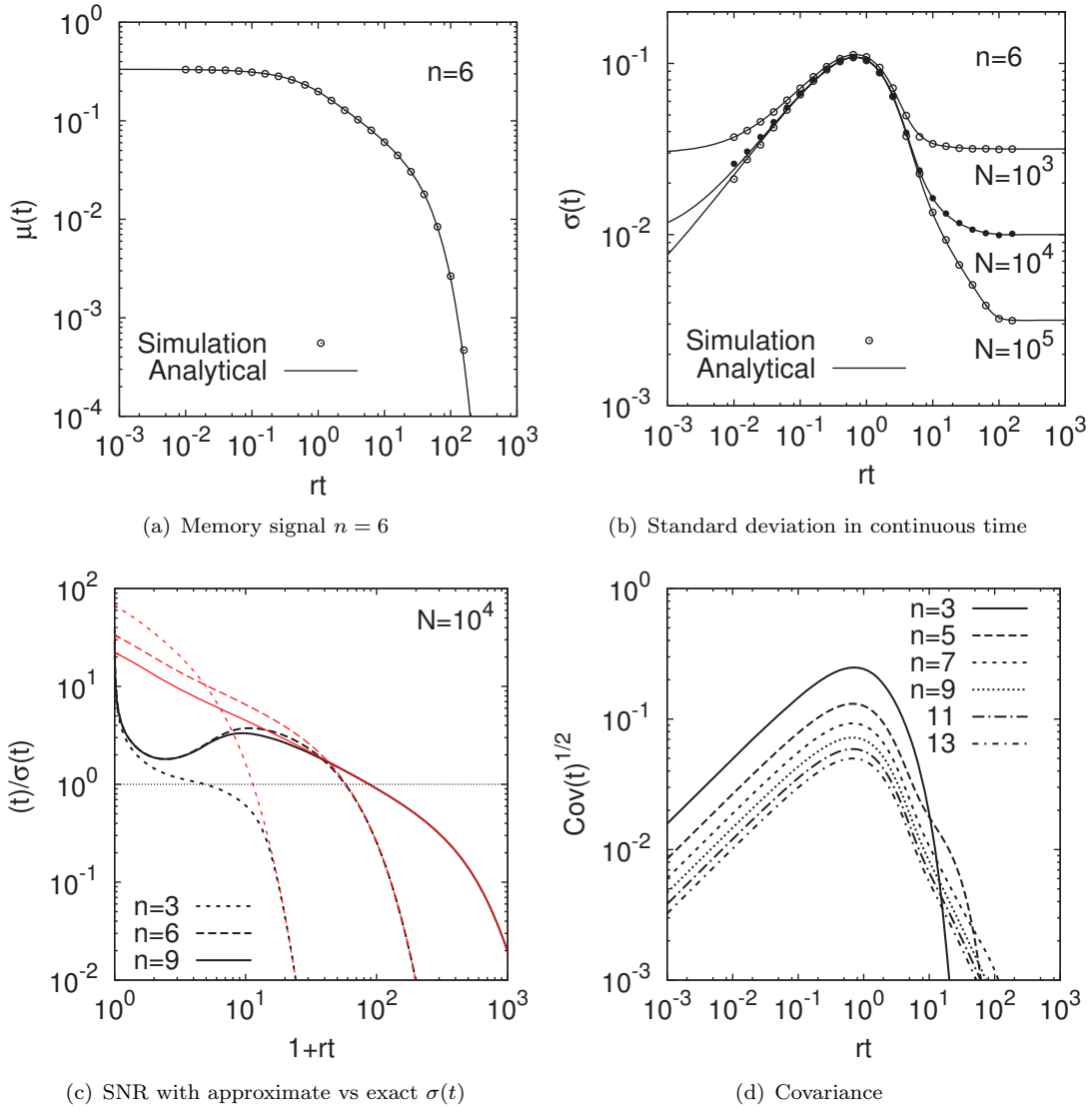
$$Y_i(s) = (2r)^{n-i}\frac{(-\frac{s}{2r}; 2)_n}{(-\frac{s}{2r}; 2)_i}(r + 2^{n-1}s), \tag{5.78}$$

where the q-Pochhammer symbol which assumes $(\alpha; q)_0 = 1$. Iterating $X_i(s)$ we obtain:

$$X_i(s) = r^{n+1-i} + \sum_{j=0}^{n-1-i} r^{n-i-j}Y_{n-j}(s). \tag{5.79}$$

The Laplace transformed signal of (5.67) reduces to a ratio of polynomials in s. The roots can be evaluated numerically for Laplace inversion with numerical inversions obtained in reasonable time for cascade sizes up to $n < 15$. Plotting analytical and simulation results shows an exact match in mean signal obtained in simulation, see Figure 5.6(a).

Figure 5.7(a) shows that the signal in discrete-time matches the continuous time results over most of the memory lifetime. Differences exist during the early steps after initial encoding, as discussed earlier due to the characteristic step-wise decay in discrete time signals as most labile states switch between cascades on every encoding step. But this matching is on mean signals only without considering differences in fluctuations in the signal between continuous and discrete time. The above methods retain the statistical structure allowing us to obtain further moments and examine the fluctuations in the mean signal, as variance and covariance. The strength states for each synapse are identically distributed random variables, and we may obtain covariance by examining covariance between any two synapses. To do so we require a transition matrix $\mathbb{P}_2$ for the synchronous updates of any two synapses. As mentioned in Section 4.5.3, in discrete time



(a) Memory signal $n = 6$

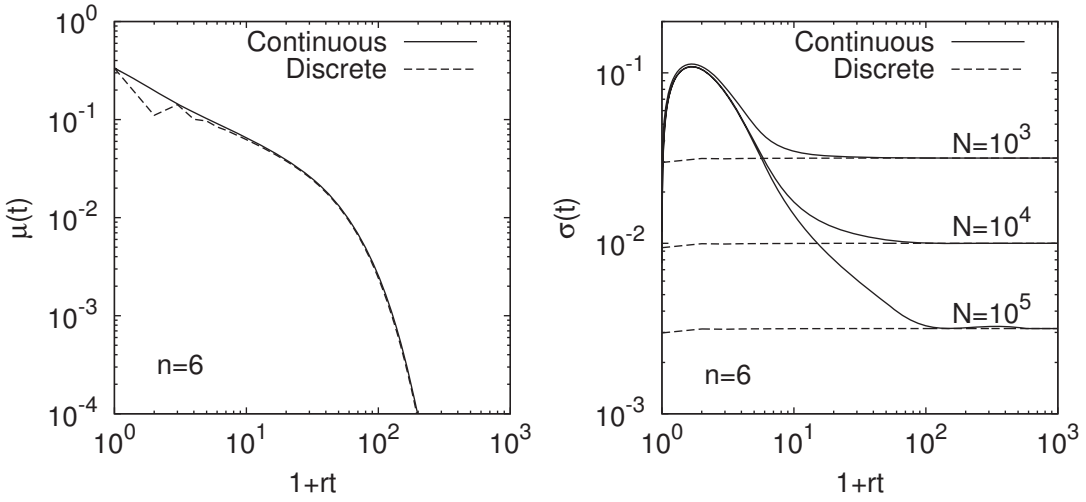
(b) Standard deviation in continuous time

(c) SNR with approximate vs exact $\sigma(t)$

(d) Covariance

FIGURE 5.6: Comparison between analytical and simulation results for $n = 6$ cascade mean memory signal **a** and standard deviation **b** $\sigma(t)$ show exact agreement. The simulation was averaged over $N = 10^5$ and $T = 10^4$ trials. **c**. SNR $\mu(t)/\sigma(t)$ for cascade synapses with cascade size n for $\sigma(t)$ exact (black thick lines) and $\sigma(t) \approx N^{-1/2}$ (red thin lines). **d**. Covariance between cascade synapses for different cascade sizes. Covariance reduces with size n but it reduces to a core that does not get much smaller with increasing n.

we obtain $\mathbb{M}_2$ which factorizes as $\mathbb{M}_2^n = \mathbb{M}_1^n \otimes \mathbb{M}_1^n = (\mathbb{M}_1 \otimes \mathbb{M}_1)^n$. In continuous time the generating matrix for a single synapse is $\mathbb{G}_1 = \mathbb{M}_1 - \mathbb{I}$, and the probability of updates goes as $P(t) = \exp(rt(\mathbb{G}_1)P(0))$. For a pair of synapses we need to obtain a matrix for the synchronous updates of two synapses via the tensor product of the transition matrices $\exp(rt(\mathbb{G}_{\not\models}) = \exp(rt(\mathbb{G}_{\not\models}) \otimes \exp(rt(\mathbb{G}_{\not\models})$ which does not factorize with matrix exponentials. We therefore consider $\mathbb{G}_2 = \mathbb{M}_1 \otimes \mathbb{M}_1 - \mathbb{I} \otimes \mathbb{I}$ to obtain the mean $\mu_2(t)$ signal and calculate covariance. Critically discrete time matrices $\mathbb{M}$ factorize while $\mathbb{G}s$ do not, showing that the two processes behave differently.



(a) Mean signal (b) Standard deviation

FIGURE 5.7: Comparison between discrete-time and continuous-time analytical results for the cascade's mean signal **a** and standard deviation **b**. The mean signals match except for early on in discrete-time but standard deviations do not match because the synchronous presentation of encoded memory patterns drives a covariance term. The peak of the covariance occurs has soon after the memory is encoded.

Indeed, plotting the variance between discrete and continuous time we find significant differences, see Figure 5.7(b). In discrete time, as expected, the synaptic states are independent and thus the covariance is zero. However, in continuous time although the memory encoding events intervals are exponentially distributed during each encoding event all synapses are updated synchronously giving rise to a covariance term. The peak of the covariance is located early on near the time when it is expected that the most labile states of the cascade switch strength. This rise of covariance, could compromise the SNRs in continuous time. Figure 5.6(c) plots SNRs using exact $\sigma(t)$ and approximate $\sigma(t) = \sqrt{N}$ showing that the impact of covariance on the initial SNR can be significant. Therefore, although the mean signals between discrete and continuous time seen in Figure 5.7(a) match, discrete time results are not generic.

5.6 Summary

This chapter has focused on the cascade model of synaptically stored memories. The model achieves an extension of lifetimes of the standard bistable stochastic updating synapse by introducing hidden states to model metaplasticity (Abraham and Bear, 1996). A cascade of n hidden states is attached to each strength. The hidden states have decreasing transition probabilities, which in the proposed configuration attempt to decrease the trade-off between the speed of learning and forgetting. The cascade memory

operates as a palimpsest, new memories interfere and erase the oldest ones. We discussed the general methods by which memories encoded in palimpsest can be measured and analysed. Palimpsest memories burn a fluctuation on equilibrium state distributions and by tracking the evolution of these distributions we may measure the memory signal. The cascade's memory dynamics exhibit a phase of power-law forgetting, which is extended as the cascade size increases. Nevertheless, the cascade is still tightly bound by the trade-off between initial signal and memory lifetime and it exhibits a relationship between optimal cascade size and network size N_{syn}. We presented an analysis of the mean dynamics and an exact master equation calculation for the mean signal and variance. The variance in continuous time showed that there is a rising covariance term which affects SNRs making the discrete time results non-generic.

In the next Chapter we propose models for synaptically stored memory that attempt to resolve the stability versus plasticity dilemma but also address the trade-off between initial signal and memory lifetimes.

Chapter 6

Filter synapses

In the previous chapter we presented an influential model of synaptic plasticity for memory which consists of a complex cascade of stochastic synapses that progressively locks synapses in increasing stability. This model extends the memory lifetimes of stochastically updating synapses by introducing multiple memory signal decay time constants that result in power-law forgetting dynamics when superimposed. In this chapter we extend a model of synaptic plasticity filters, which has been first applied in neural development (Elliott, 2008, Elliott and Lagogiannis, 2009), to consider the stability versus plasticity dilemma (Grossberg, 1980, Abraham and Robins, 2005) faced by ongoing memory processes. This framework separates the processes of plasticity induction and expression to specifically filter the induction stimuli before plasticity expression. Thus in contrast to the cascade's stochastic view of synaptic plasticity, filter synapses do not rely on intrinsic random responses to induction stimuli but offer a mechanism by which plasticity expression can depend on the history recent stimuli. The material in this chapter is mainly based on our publication Elliott and Lagogiannis (2012).

6.1 Introduction

The aforementioned models of synaptic plasticity for ongoing memory processes we discussed in Chapter 4 have suggested that neural systems solve the stability versus plasticity dilemma (Grossberg, 1980, Abraham and Robins, 2005) by effectively adjusting the rate that new memory engrams erase previous ones. This is achieved by stochastically expressing a fraction of the induced plasticity at the level of single synapses (Tsodyks, 1990, Amit and Fusi, 1994, Brader et al., 2007), which has then been extended by adding history dependence through metaplasticity (Fusi et al., 2005, Amit and Fusi, 1994, Leibold and Kempter, 2008), or via assuming sparse neural codes where each memory only induces plasticity at a fraction of synapses (Tsodyks and Feigel'Man, 1988, Rubin and

Fusi, 2007, Leibold and Kempter, 2008). Such methods allow for one-shot memory encoding in the states of synaptic strength. The fidelity of the new engram relates to the fraction of synapses that expressed plasticity in response to the encoding event. Recall fidelity is maximum immediately after memory storage and it degrades under further memory encoding.

Therefore, from the perspective of individual synapses, whose strength state is already part of some memory engram, the plasticity inducing stimuli under new memory encoding constitute destabilizing fluctuations. Nevertheless, to encode new memory engrams some synapses would need to express plasticity and thus synapses are bestowed with the task of deciding between stability and plasticity. Within the context of neural development synapses are faced with a similar dilemma where they need to decide if an incoming stimulus is part of a developing trend and should be expressed or noise and therefore it should be suppressed. In that context, it has previously been proposed that plasticity induction and expression needs to be separated so synapses process the stream of induction stimuli to average out fluctuations before expressing plasticity (Elliott, 2008). If these processes are separated then synapses would require a time period over which they would be processing induction signals before expressing any plasticity. Indications for a processing period before expression can be drawn from LTP experiments where typically plasticity expression can be delayed up to one minute after its induction (Gustafsson et al., 1989, Petersen et al., 1998, O'Connor et al., 2005), or from STDP experiments where multiple spike-pairs need to be presented before any overt plasticity is expressed (Bi and Poo, 1998).

In an analogy to integrate-and-fire neurons, models for developmental plasticity called "integrate-and-express" (Elliott and Lagogiannis, 2009) have been developed that separate processes of induction from expression and specifically seek to suppress fluctuations by employing low-pass filters in the induction process (Elliott, 2008, Elliott and Lagogiannis, 2009, Elliott, 2010b, 2011a,b). These filters attenuate induction stimuli that frequently switch the direction of plasticity as high-frequency noise while they perceive low-frequency changes in the direction of plasticity as signals that need to be expressed.

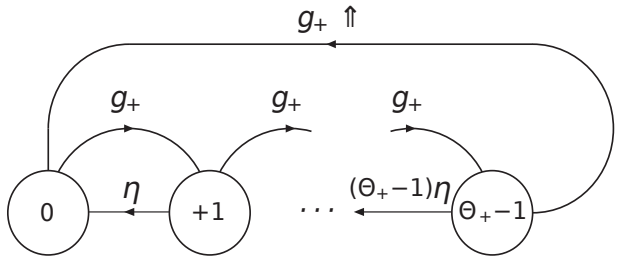
In this chapter we adopt the integrate-and-express synapse models as a solution to the stability versus plasticity dilemma faced by synapses in ongoing memory processes. Thus, the process of plasticity induction is separated from its expression. Synapses integrate plasticity stimuli in the induction process and express plasticity according to certain criteria. Initially, we consider an induction model that integrates potentiating and depressing stimuli separately only to express plasticity when one of the integrators becomes full. We then unify the integrating process so potentiating and depressing stimuli compete for reaching the threshold of expression $\pm\Theta$. This unified filter model essentially implements a low-pass (LP) filtering processes in the sequence of induction stimuli. Expressing plasticity in this model requires a sufficiently large fluctuation in the train of plasticity inducing stimuli to drive synapses to threshold.

We examine the dynamics of memory storage and recall under these filter synapses and find that they exhibit novel signal dynamics. In contrast to the monotonically decaying signal trace of earlier models, filter synapses exhibit both a rise and a fall in the signal-to-noise ratio. The location and amplitude of the signal peak depends on the filter size which is defined by the upper and lower thresholds $\pm\Theta$. This model may be extended to a more generic form through the addition of a time dependent state decay process that returns the internal filter state to the reset position, such an extension has also been previously examined in the context of neural development (Elliott, 2011a). The decay process can be used to constraint the interaction of memory encoding stimuli in time but it also broadens the range of mean expression times available to standard unified filters of the same size. This extension of expression times will be used in later chapter when we compare models.

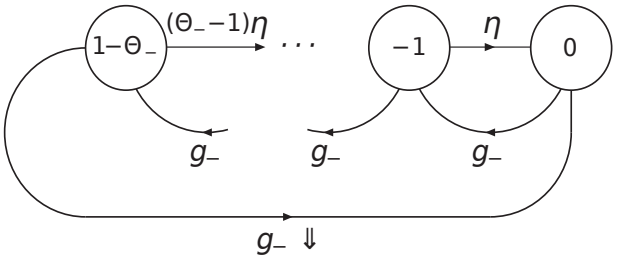
It should be noted, that the framework used here examines the encoding of a stream of equally important memories, we are assuming that these processes rely on early phases of plasticity and not the late protein-synthesis dependent phases (Krug et al., 1984, Manahan-Vaughan et al., 2000).

6.2 The dual-filter with a decay

With integrate-and-express synapses plasticity induction stimuli $\xi_i^\alpha h^\alpha$ drive changes in hidden synaptic states of a synapse i that may trigger the expression of plasticity. In this section we describe a synapse model that integrates plasticity stimuli using two integrators, one for potentiating stimuli and a separate one for depression stimuli. Each integrator consists of a set of Θ discrete states that specifically filter one type of induction stimuli to only allow expression when either one of the integrators associated with a filter reaches Θ. The filter integrating potentiating stimuli is called the p-filter and the other one for the depression stimuli is called the q-filter. Both integrators start from a state $i = 0$ and have their state incremented $i + 1$ when an appropriate stimulus arrives, a $\xi_i h^\alpha = +1$ increments the p-filter and a $\xi_i h^\alpha = -1$ increments the q-filter only. Since these filters only increment towards threshold we may consider to add a state decay process to each filter so states also decrement passively with a rate set by a parameter η. Once the integrator of either filter reaches threshold then the respective plasticity direction is expressed and both of them are reset to zero. Thus, effectively this synapse introduces a race towards reaching Θ between induction stimuli, the one that wins the race gets expressed. We refer to this type of synapses as *integrate-and-express* synapse while this particular type we call the *dual-filter* because it consists of two identical filters that each integrate induction stimuli. On Figure 6.1 we depict a state diagram that represents the abstract form of the p-filter and q-filter with the decay process extension that may underlies the biophysical mechanism. Here, each state may represent the number of activated molecules due to the induction history. The decay rate of each



(a) p-filter



(b) q-filter

FIGURE 6.1: The dual-filter has two such threshold processes running simultaneously one is called the p-filter and the other the q-filter. Each p or q filter has one holding boundary at state zero and an absorbing boundary at Θ. Each state independently decays as a Poisson process of rate η, and when either one of the occupied states decays then the current state is decremented by one. Thus in the state diagram the decay rate increases with state number as the *combined rate* of all previous states, which for Poisson processes this is the sum of the rates and therefore for the near threshold state $i = \Theta - 1$ the decay rate is $(\Theta - 1)\eta$. In discrete time we consider the decay process as the probability p_i^n that n out of i have decayed during a time step. Thus, for each filter two simultaneous processes are involved: a decay process p_i^n, which may decrement the state i by n steps towards zero, and the other increments towards Θ due to the arrival of an induction stimulus. The q-filter and the p-filter run independently, one increments due to a DEP stimulus and the other by a POT stimulus. When one of the two processes reaches threshold both p-q filters are reset to zero .

state is η and each independently decays as a Poisson process of rate η. Since the state occupied in the model represents the number of activated substrates that maintain the filter state thus larger filter states will have a decay rate that is the *combined rate* of all previous states, which for Poisson processes this is the sum of the rates and therefore for the near threshold state $i = \Theta - 1$ the decay rate is $(\Theta - 1)\eta$ for one step decay.

This race towards expression can be formulated by a two dimensional random walk in the integers, where one dimension is the steps taken by the p-filter and the other represents the steps taken by the q-filter see Figure 6.2. Ignoring any boundaries and counting the paths that lead to each vertex in the grid gives the Pascal triangle of binomial coefficients and thus the distribution is a binomial centred along the diagonal. However, the threshold process on each filter implements a boundary and thus the 2-d walk is bounded in a quadrant set to the right by Θ_+ and above by Θ_-. In the absence of decay the maximum number of induction steps that can occur before expression assuming filters start from the zero state is $2\Theta - 1$, while the minimum number of steps

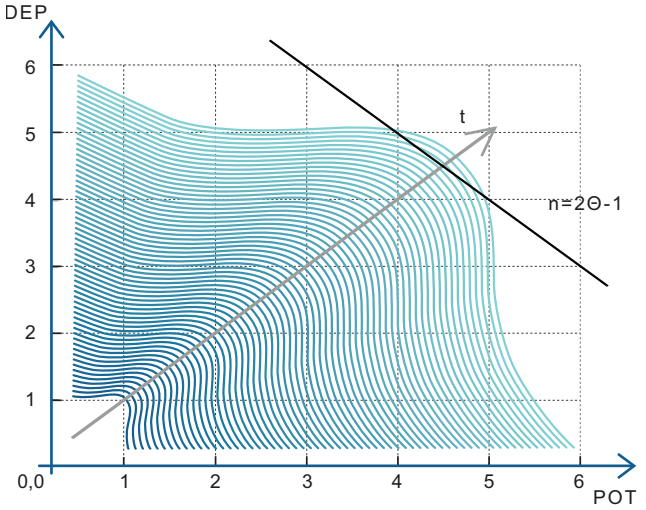


FIGURE 6.2: The graph is an illustration of the dual-filter state occupancy in two dimensions. Each axis represents the state of one of the two filters q-filter or p-filter of the dual filter. When combined it appears as 2-D random walk that begins from state 0,0 and evolves towards an upper boundary around the axis $x = y$. The distribution is a binomial centred at the $x = y$ and evolving towards this diagonal up to an upper boundary. The upper boundary represents the maximum number of steps before the dual-filter expresses a transition. Without decay, the maximum number of induction stimuli before expression is $2\Theta - 1$, at that point the filter has to respond to the next induction stimulus by expressing a transition.

to expression would occur only when a single type of stimulus is applied thus filling one of the filters in at least Θ steps. With decay however, a stochastic decay process occurs between timesteps with rate η on both p- and q-filter simultaneously.

We may calculate discrete transitions between states but require a way to calculate the decay steps that may have happened between induction steps. This simultaneous process can be calculated by counting the number of decay steps between the last and the next induction step. Assuming the lifetime of each state filter state decays independently the number of states that have decayed in a fixed interval is then binomially distributed. The probability that one decay step occurs in time t is given by $p = 1 - e^{-\eta t}$. The probability that m decay steps occur in time t is given by:

$$p_m^n = \binom{n}{m}[1 - e^{-\eta t}]^m e^{-\eta t(n-m)} = \binom{n}{m} p^m (1-p)^{n-m} \text{ where } p = 1 - e^{-\eta t} \tag{6.1}$$

This way the model remains discrete, jumping between induction stimuli but at every time interval we calculate the number of decay steps that probabilistically occurred during the interval. In practice, we assume the decay steps occurred before the current induction step and thus the decay steps first reduce the occupied state j by k steps before the new inductions step are accounted for.

To understand how this filter model works lets assume we are examining a strong

synapse. We split the transitions into two types, those that involve an increment on the p-filter with probability g_+ due to a potentiating stimulus and those that increment the q-filter with probability g_- under a depression stimulus.

To examine the transition probabilities would require to formulate them in 2-d index notation. We let $P_{i,j|k,l}(n)$ denote the transition from vertex $(k,l) \mapsto (i,j)$ in n discrete steps. Without decay we have :

$$P_{i,j|k,l}(n+1) = (1-g_- - g_+)P_{i,j|k,l}(n) + g_+ P_{i,j|k+1,l}(n) + g_- P_{i,j|k,l+1}(n) \tag{6.2}$$

However, with the addition of decay k' steps may occur in between step $n+1$ and n on either of the two filters and thus we need to sum these probabilities also :

$$\begin{aligned} P_{i,j|k,l}(n+1) = (1-g_- - g_+) &\sum_{k'=0}^{k}\sum_{l'=0}^{l} p_{k'}^k p_{l'}^l P_{i,j|k-k',l-l'}(n) \\ &+ g_+ \sum_{k'=0}^{k}\sum_{l'=0}^{l} p_{k'}^k p_{l'}^l P_{i,j|k+1-k',l-l'}(n) \\ &+ g_- \sum_{k'}^{k}\sum_{l'=0}^{l} p_{k'}^k p_{l'}^l P_{i,j|k-k',l+1-l'}(n) \end{aligned} \tag{6.3}$$

Explicitly calculating transition probabilities would be rather complicated. We may instead directly consider expression times $\tau_{i,j}$ starting from state i,j and escaping from the bounded quadrant. The relationship to escape times is similar to the above. We let $g_n = (1-g_- - g_+)$ and collect all terms under sums to shorten the expression into:

$$\tau_{i,j} = 1 + \sum_{k'=0}^{k}\sum_{l'=0}^{l} p_{k'}^k p_{l'}^l \left[g_n \tau_{(i-k)',(j-l')} + g_+ \tau_{(i+1-k'),(j-l')} + g_- \tau_{(i-k'),(j+1-l')} \right] \tag{6.4}$$

Looking at a single dimension for one filter, from each state j one of the filters can move to $j+1-k'$ given an induction stimulus g_+ and k' decays events. Given the required induction has not occurred then k decays occurring with probability $p_{k'}^k$ will move the filter to state $j-k$. At the boundary $j=\Theta$ escape occurs and thus the waiting time is $\tau_\Theta = 0$. Here we need to define two dimensional boundary condition along the quadrant, but the principle is the same. Obtaining expression times requires solving this system of equations, and thus resorting to matrix inversion.

By obtaining a few samples of escape times against Θ, assuming symmetric q-filter and p-filter sizes with $\eta = 0.0$ and $g_\pm = 1/2$, we find that the escape time grows linearly with Θ in and a linear fit can be :

$$\tau(\Theta) = -\frac{9}{8} + \frac{7}{4}\Theta \tag{6.5}$$

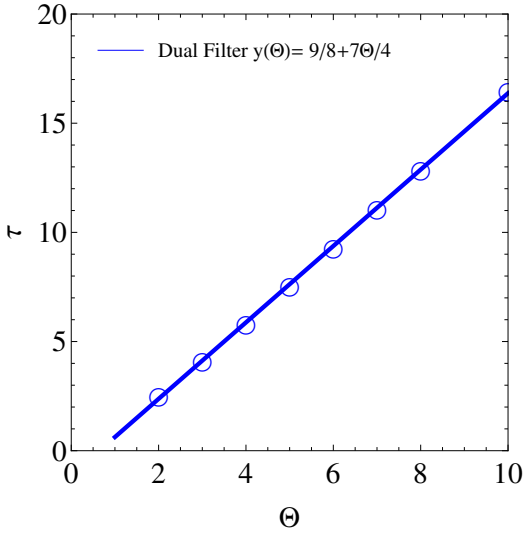


FIGURE 6.3: Calculated dual-filter escape times for $\eta = 0.0$ and $g_{\pm} = 1/2$ (open circles) and the linear fit function of Equation (6.5).

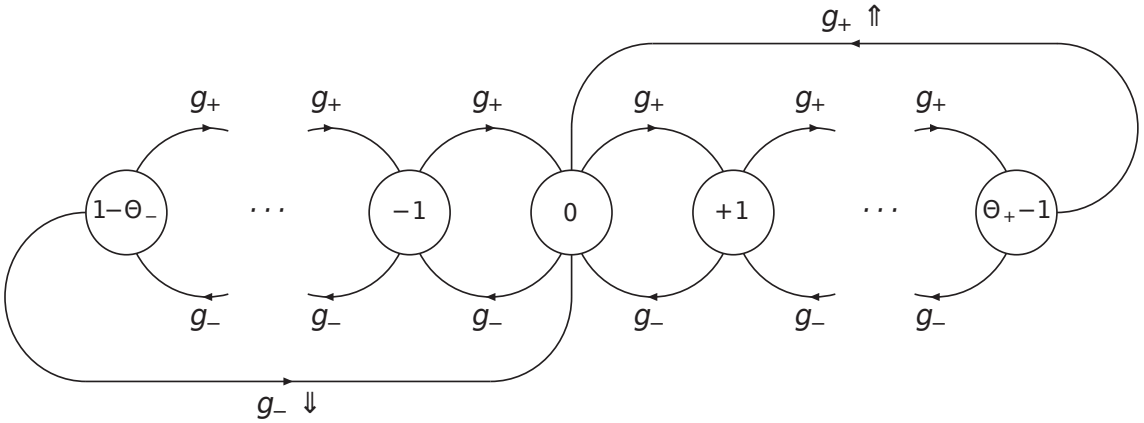


FIGURE 6.4: A unified (LP) filter with two symmetric thresholds Θ_+ and Θ_-. This filter has the POT and DEP induction stimuli compete for expression. POT induction events occur with probability g_+ and lead to an increment in filter state and DEP with probability $g_- = 1 - g_+$ lead to a decrement in filter state. The schematic also shows that upon a threshold event the filter state is reset to zero ,and either a potentiation plasticity step is expressed ($\Uparrow$) or a depression step ($\Downarrow$). The mean expression time starting from the zero state is Θ^2. Sufficiently large fluctuations in the balance between the two stimuli can drive towards escape from either threshold $\Theta_{\pm}$ and consequently to plasticity expression depending on the threshold crossed. In the framework we are considering binary strength synapses and thus strong synapses expressing via the upper threshold would not exhibit any change in strength. However, the synaptic filters can be extended to a framework having access to more discrete strength states.

We will not dwell on the dual-filter as it is only an initial approach that will be later used only for comparison against another filter model. Instead, we proceed to consider a simpler integration method which is also more analytically tractable.

6.3 A single unified filter

We extend our approach to include competition between potentiation, denoted $\Uparrow$, and depression ($\Downarrow$) inducing signals. This is achieved by unifying the two separate filters, discussed previously, and attaching them back- to-back at the zero state. Initially we may remove the decay process and examine this unified filter in discrete time model assuming synapses of binary strength. Here, we will be focusing on balanced excitation $g_- = g_+ = \frac{1}{2}$ and thus we consider symmetric thresholds $\Theta = \Theta_\pm$ in order to keep mean expression times equal between potentiation ($\Uparrow$) and depression ($\Downarrow$). As before, we let h^α denote the neuron output to the tracked memory α and ξ_k^α denote the induction stimulus delivered to synapse k by the tracked memory. The transitions in the state i of a filter synapse upon receiving an induction stimulus can be summarized as follows :

$$\begin{aligned}
\xi_k^\alpha h^\alpha = +1 &\begin{cases} i \mapsto i+1 & \text{for } i < +(\Theta-1) \\ i \mapsto 0 \ \& \Uparrow & \text{for } i = +(\Theta-1) \end{cases} \\
\xi_k^\alpha h^\alpha = -1 &\begin{cases} i \mapsto i-1 & \text{for } i > -(\Theta-1) \\ i \mapsto 0 \ \& \Downarrow & \text{for } i = -(\Theta-1) \end{cases}
\end{aligned} \tag{6.6}$$

This filter uses a single integrator for both types of stimuli, but in this case POT stimuli increment the integrator while DEP stimuli decrement it. Therefore, unifying the two processes introduces competition between induction stimuli; when operating far from either threshold $\pm\Theta$ a POT ($+1$) stimulus can be cancelled by a subsequent DEP (-1) stimulus and vice versa. Thus, reaching either threshold would require a sufficiently large fluctuation in the balance between POT/DEP stimuli, while small fluctuations would be insufficient to express plasticity and thus would be suppressed. This makes the unified filter act as a low-pass (LP) filter in the sequence of plasticity inducing stimuli and we will refer to this model with both terms interchangeably.

Further, since induction and expression are separated the filter model is not bound to binary strength synapses and it can be generalized to discrete strength synapses. When synaptic strength becomes saturated, the filter may still perform the standard threshold transition with the associated reset to zero-state indicating that plasticity should be expressed but yet no further change in strength would occur.

Before we proceed with an analysis of this model, we discuss an extension to include stochastic decay (Elliott, 2011a) which makes the filter model more general. An analysis of mean expression times is then presented on this extended filter model which also applies to the simpler no-decay case.

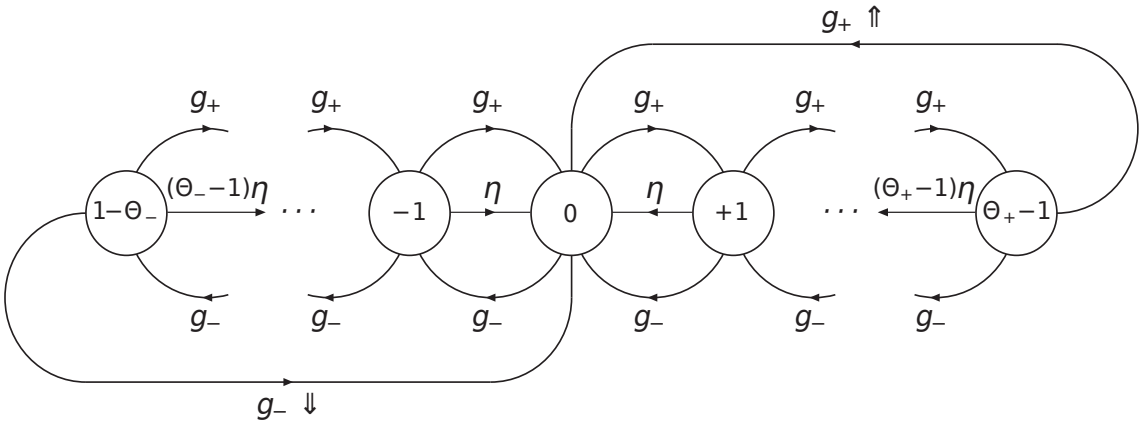


FIGURE 6.5: A filter with two symmetric thresholds Θ_+ and Θ_-. Each state of the filter may decay back towards the zero state during the time between the arrival of induction stimuli. This filter also has the POT and DEP induction stimuli compete for expression but the decay process adds a dimension of time requirement between their interaction. Fluctuations in the balance between the two stimuli drive expression and setting the thresholds and decay to appropriate values can give a desired mean expression time to match one of the states of the cascade model.

6.3.1 Adding stochastic decay to the discrete unified filter

The unified filter is able to identify a trend in the structure of plasticity induction history by keeping track of the balance between DEP and POT stimuli. However, the filter has no sense of time between induction events. Without a time variable the structure of induction stimuli appears flat and plasticity events with an arbitrary distance in time may trigger synaptic modifications in a history dependent manner. A limit to the amount of time between the interaction of two induction stimuli in the filter can be introduced by making the filter states decay. Such a model has been previously proposed in the context of controlling fluctuations in neural development (Elliott, 2011a). The extension introduces a stochastic lifetime of filter states in the same manner to the decay in the dual-filter presented previously in Equation (6.1).

The lifetime of a state is effectively a decay term which may return a filter's internal sum to zero if no plasticity stimuli arrive for a given time. For a filter at state j, and before the arrival of the next induction stimulus, m states out of j may decay which will change the state of the filter to $j - m$. The probability of m decay steps is given by p_m^j. Thus, the changes brought to the filter state by an induction stimuli at some point in time may expire through decay and may not interact with the next stimuli arriving at the same synapse.

6.3.2 Mean escape times for unified filter with decay in discrete time

This model has been previously proposed as a low-pass filter for the control of fluctuations in developmental plasticity (Elliott and Lagogiannis, 2009) and an analysis of mean

escape times for this model in continuous time has first been published in Elliott (2011a), here we present the simpler derivation of the mean escape times in discrete time. We consider a two sided process for the probability $P_{i|\pm j}(n)$ of moving from state j to state i in n steps. The state diagram of this filter is shown on Figure 6.5, with probability g_+ for POT signal and g_- for DEP, while we may leave open that neither stimuli arrive $g_n = 1 - g_+ - g_-$. This filter operates in identical manner to the unified-filter.

As with the two filter process, the paradigm of radio active decay is adopted for each filter state j that has been reached due to plasticity induction. The probability density of the the decay of a single filter state is $p(t) = 1 - e^{-\eta r t}$. However, in discrete time since the decay event can occur between timesteps we need to consider instead the number of decay steps that occur between induction steps. We denote the probability that m out of n possible states have decayed as:

$$p_m^n = \binom{n}{m}(1 - e^{-\eta})^m e^{-(n-m)\eta} \tag{6.7}$$

The first equation corresponds to the positive filter states $0 \to +(\Theta - 1)$:

$$\begin{aligned} P_{i|+j}(n+1) = \sum_{k=0}^{j} p_k^j [&(1 - g_+ - g_-) P_{i|+j-k}(n) \\ &+ g_+ P_{i|+j-k+1}(n) + g_- P_{i|+j-k-1}(n)], \end{aligned} \tag{6.8}$$

and similarly for the negative states :

$$\begin{aligned} P_{i|-j}(n+1) = \sum_{k=0}^{j} p_k^j [&(1 - g_+ - g_-) P_{i|-j-k}(n) \\ &+ g_+ P_{i|-j-k+1}(n) + g_- P_{i|-j-k-1}(n)]. \end{aligned} \tag{6.9}$$

Now we may generalize the above expressions to include starting from point j and moving to any point i within some interval I, which is given by :

$$f_j(n) = \sum_{i \in I} P_{i|j}(n) \tag{6.10}$$

and its reciprocal of not being in the interval I in n steps starting from j is

$$s_j(n) = 1 - f_j(n).$$

Combining the two cases in equations (6.8) we generalize to $i \in \{-(\Theta - 1), \cdots, +(\Theta - 1)\}$ and proceed to define the probability of not being in the interval I at step $n + 1$

conditioning on step n as before:

$$s_{\pm j}(n+1) = \sum_{k=0}^{j} p_k^j \left[g_n s_{\pm(j-k)}(n) + g_+ s_{\pm(j-k)+1}(n) + g_- s_{\pm(j-k)-1}(n) \right]. \tag{6.11}$$

The above equation accounts for j possible decay steps having occurred between steps. We can now examine the escape through threshold Θ_+ or Θ_- in an interval $I \in \{-\Theta_-, +\Theta_+\}$, for example escape through Θ_+ defines the boundary conditions:

$$\begin{aligned} &s_{+\Theta_+}(n) = \delta_{n0} && \Rightarrow \text{Escape in } n = 0 \text{ if we start from threshold.} \\ &s_{-\Theta_-}(n) = 0 && \Rightarrow \text{Never escapes from } \Theta_+ \text{ since it crossed } \Theta_- \text{ already.} \end{aligned}$$

We will use a probability generating function (PGF) to describe escape in n steps starting from filter state j and define:

$$G_j(z) = \sum_{n=0}^{\infty} z^n s_j(n) \tag{6.12}$$

The PGF defines an infinite sum over all n of the probability $g_j(n)$ of not being in the interval I. In the limit $z \to 1$ the PGF returns the probability of escape $G_j(1) = \pi_j$ through $\Theta_\pm$ in any number of steps starting from j. Given the above boundary conditions we would be only considering escape through the upper threshold Θ_+, $G_j^+(1) = \pi_j^+$.

Using moments of the PGF we may further define expressions to obtain other useful properties such as the mean number of steps $\mathrm{E}_j[n]$ for filter to reach threshold starting from point j and its variance.

We may define a generating function for $s(n+1)$ by rewriting :

$$\sum_{n=0}^{\infty} z^n s_{\pm j}(n+1) = \frac{1}{z} \sum_{n=0}^{\infty} z^{n+1} s_{\pm j}(n+1), \tag{6.13}$$

using Equation (6.11) we write the PGF :

$$G_{\pm j}(z) = z \sum_{k=0}^{j} p_k^j [g_n G_{\pm(j-k)}(z) + g_+ G_{\pm(j-k)+1}(z) + g_- G_{\pm(j-k)-1}(z)]. \tag{6.14}$$

We then proceed to obtain 1st and 2nd order moments by differentiating with respect to z:

$$\begin{aligned} G'_{\pm j}(z) = z\sum_{k=0}^{j} p_k^j [g_n G'_{\pm(j-k)}(z) + g_+ G'_{\pm(j-k)+1}(z) + g_- G'_{\pm(j-k)-1}(z)] \\ + \sum_{k=0}^{j} p_k^j [g_n G_{\pm(j-k)}(z) + g_+ G_{\pm(j-k)+1}(z) + g_- G_{\pm(j-k)-1}(z)] \end{aligned} \tag{6.15}$$

and the second order moment:

$$\begin{aligned} G''_{\pm j}(z) = z\sum_{k=0}^{j} p_k^j [g_n G''_{\pm(j-k)}(z) + g_+ G''_{\pm(j-k)+1}(z) + g_- G''_{\pm(j-k)-1}(z)] \\ + 2\sum_{k=0}^{j} p_k^j [g_n G'_{\pm(j-k)}(z) + g_+ G'_{\pm(j-k)+1}(z) + g_- G'_{\pm(j-k)-1}(z)]. \end{aligned} \tag{6.16}$$

Letting $z \to 1$ transforms the generating functions into sums of their coefficients. In this case the PGF $G_{\pm j}(1)$ gives the probability of escape $\pi_{\pm j}$ in any number of steps :

$$G_{\pm j}(1) = \pi_{\pm j} \tag{6.17}$$

Accordingly the first order moments:

$$G'_{\pm j}(1) = \sum_{n=0}^{\infty} n s_{\pm j}(n) z^{n-1} = \tau_{\pm j} \pi^+_{\pm j}, \tag{6.18}$$

give the mean number of steps to escape $\mathrm{E}_{\pm j}[n]$, with τ_j the conditional escape time through the upper threshold Θ_+. Finally :

$$G''_{\pm j}(1) = \sum_{n=0}^{\infty} n(n-1) s_{\pm j}(n) z^{n-2} \tag{6.19}$$

gives :

$$G''_{\pm j}(1) = E[N(N-1)] = E[N^2] - \pi_{\pm j}\tau_{\pm j} = \Phi_{\pm j} \tag{6.20}$$

Taking the limit as $z \to 1$ in equations (6.14), (6.15) and (6.16) and substituting $G_{\pm j}(1)$, $G'_{\pm j}(1)$, $G''_{\pm j}(1)$ according to equations (6.17),(6.18) and (6.20) we obtain :

$$\pi_{\pm j} = \sum_{k=0}^{j} p_k^j [g_n \pi_{\pm(j-k)} + g_+ \pi_{\pm(j-k)+1} + g_- \pi_{\pm(j-k)-1}] \tag{6.21}$$

$$\begin{aligned} \pi_{\pm j}\tau_{\pm j} =& \sum_{k=0}^{j} p_k^j [g_n \pi_{\pm(j-k)} \tau_{(j-k)}] \\ &+ \sum_{k=0}^{j} p_k^j [g_+ \pi_{\pm(j-k)+1} \tau_{\pm(j-k)+1} + g_- \pi_{\pm(j-k)-1} \tau_{\pm(j-k)-1}] + \pi_{\pm j} \end{aligned} \tag{6.22}$$

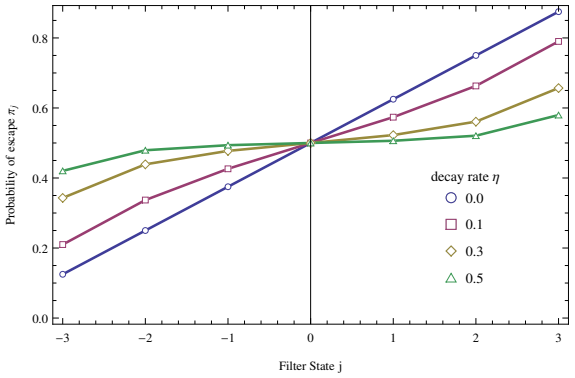
$$\Phi_{\pm j} = 2(\pi_{\pm j}\tau_{\pm j} - \pi_{\pm j}) + \sum_{k=0}^{j} p_k^j [g_n \Phi_{\pm(j-k)} + g_+ \Phi_{\pm(j-k)+1} + g_- \Phi_{\pm(j-k)-1}] \tag{6.23}$$

We can obtain the $\pi_{\pm j}$ for all j by writing all $\pi_{\pm j}$ equations and solving the system of equations using the boundary conditions of $\pi_{+\Theta} = 1$ for escape from the upper threshold or $\pi_{-\Theta} = 1$ from the lower. We find that for a filter without decay $\eta = 0$ the probability of escape conditioned on the threshold (i.e π_j^+ or π_j^-) increases linearly with the filter state j. Figure 6.6(a) shows π_j^+ and examines the effects of decay η on the escape probability from a symmetric filter ($\Theta_\pm = \pm 4$). Increasing the rate of decay η decreases the gradient in the escape probability. With high enough η the probability of escape becomes almost constant across most internal states. Under a high decay rate the filter effectively loses its history depend expression property as the escape probability from most state is the same (see Figure 6.6(a)).

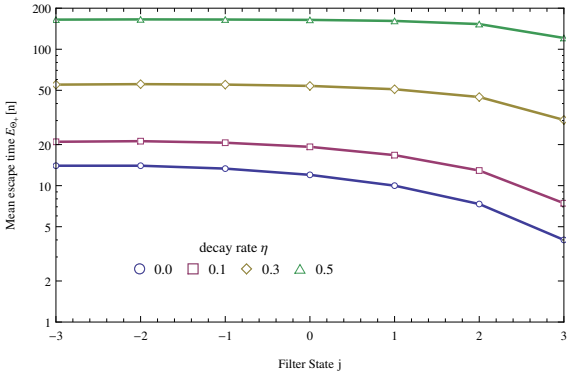
We may proceed to obtain the conditional mean escape time by solving the system of equations in (6.22) for $\tau_{\pm j}$ with $\tau_{(\pm\Theta)} = 0$ using the results obtained for $\pi_{\pm j}$. Figure 6.6(b) examines how escape times over the Θ_+ threshold increase with η for a filter $\Theta = 4$ starting from state $\pm j$. Under no decay $\eta = 0.0$ starting from state $j = 0$ the escape time is $\tau_0 \approx 12$, with increasing η the escape times shift upwards reaching $\tau_0 \approx 164$ for $\eta = 0.5$.

Further, we may obtain the variance of the escape time process by recognizing the relationship between variance and the moments of the PGF. We can rewrite the expression for $\mathrm{Var}_{\pm j}[n]$ in a form that we can use the 1st and 2nd order moments of the PGF:

$$\begin{aligned} \mathrm{Var}_{\pm j}[n] &= E_{\pm j}[n^2] - E_{\pm j}[n]^2 \\ &= E_{\pm j}[n(n-1) + n] - E_{\pm j}[n]^2 \\ &= E_{\pm j}[n(n-1)] - E_{\pm j}[n] + E_{\pm j}[n]^2 \\ &= G''_{\pm j}(1) + G'_{\pm j}(1) - G'_{\pm j}(1)^2 \end{aligned} \tag{6.24}$$



(a) Probability of escape for different decay rates



(b) Escape time from Θ_+ for different decay rates

FIGURE 6.6: The effect of increasing the decay rate on the escape time and the probability of escape from the upper boundary of a unified filter $\Theta_\pm = 4$ for four different decay rates : $\eta \in 0.0, 0.1, 0.3, 0.5$. For a symmetric filter, the results obtained for the lower boundary would be a horizontal reflection of the above. **a.** Increasing the decay rate increases the unconditional escape time from all states simultaneously. **b.** For $\eta = 0$ the probability of escape through the upper threshold is an increasing linear function of filter state towards unity. For a $\eta = 0.5$ the probability of escape is almost constant among states. Thus, increasing η reduces the relative differences in the probability of escape among states, effectively making the escape process independent of filter state and therefore the filter loses its memory of the induction history.

To obtain the variance we need an expression for $G'_{\pm j}(1) = \pi_{\pm j}\tau_{\pm j}$ and $G_{\pm j}(1) = \pi_{\pm j}$. We then proceed to solve the final system by equating $\Phi_{\pm j}$ with the expression $2(\pi_{\pm j}\tau_{\pm j} - \pi_{\pm j})$:

$$\begin{aligned} -2(\pi_{\pm j}\tau_{\pm j} - \pi_{\pm j}) = &- \Phi_{\pm j} \\ &+ \sum_{k=0}^{j} p_k^j g_n \Phi_{\pm(j-k)} \\ &+ \sum_{k=0}^{j} p_k^j \left[g_+ \Phi_{\pm(j-k)+1}(z) + g_- \Phi_{\pm(j-k)-1} \right] \end{aligned} \tag{6.25}$$

Here, we have reproduced a preliminary method for the analysis of the LP (unified) filter with decay to obtain mean escape times $\mathrm{E}_{\pm j}[n]$ and their variance $\mathrm{Var}_{\pm j}[n]$ in discrete

time. These methods will be later used to calculate parameters $\Theta_{\pm}$ and η such that these filters express at desired unconditional mean escape times τ_0.

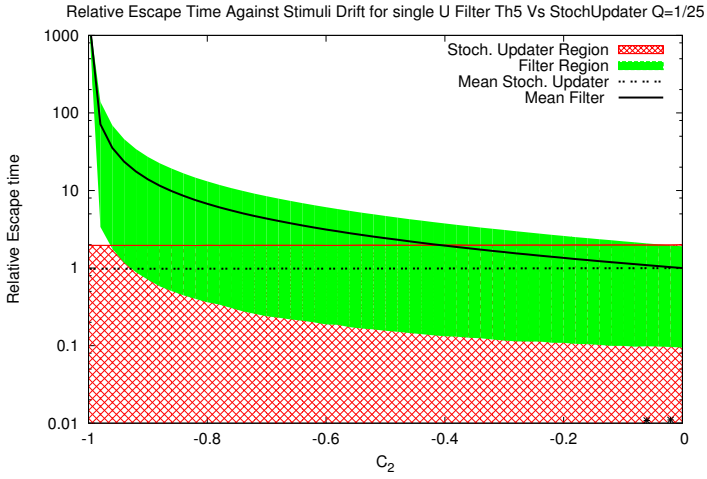
One interesting property of this model is the ability to filter stimuli by allowing the mean escape time to change in response to the statistical properties of the recent history of the induction stimuli. We may define the statistic of the train of stimuli by letting C_2 denote the probability that the next stimulus (POT or DEP) is identical to the previous one $\text{Prob}[I_2 = \pm I_1] = (1 \pm C_2)/2$. When $C_2 = -1$ the train of stimuli will toggle deterministically between POT and DEP.

A characteristic response at the level of a single synapse if either an LP or dual-filter is in operation will be that the number of stimuli for plasticity expression will vary depending on the train of stimuli. For a homogeneous train of stimuli expression will occur sooner than if a mixture of induction stimuli is used because every single induction step will be directed towards the same threshold. Conversely, assuming that the filters detect all induction signals without losses, a driftless train of stimuli will extend the expression time of filter synapses. In the case of an LP filter synapse driftless induction stimuli will produce on average no change in synaptic strength because the synapse will alternate between two internal filter states and thus the mean time to reach threshold will be infinite.

For comparison we also consider an SU synapse, according to which the decision to express plasticity given each induction stimulus is based on a Bernoulli trial. In this case the number of induction stimuli for expression is not influenced by drift in the direction of induction stimuli because an SU probabilistically delays the expression of plasticity and does not have any memory of the recent stimuli. Therefore, a stochastic synapse would show a constant expression time regardless of the statistics in the stimulus train. The results of a simulation comparing the mean expression times between an LP filter and a matched SU synapse ($q = 1/\Theta^2$) under varying drift is shown on Figure 6.7(a). If each successive induction stimulus is anti-correlated to the previous one, the mean expression time of an LP filter synapse goes to infinity while for an SU synapse it remains constant. However, as the correlation between induction stimuli goes to zero the mean expression times between LP filter and an SU synapse converge as required. Thus, filter synapses would be suitable to evaluate the storage of stimuli based on volatility as indicative of uninteresting stimuli. We will return to this issue in the final chapter where we propose an experimental protocol that uses expression times to identify between cascade, filter and stochastic updating synapses.

6.3.3 Unified filter escape densities in continuous time

In the previous section we considered the mean expression time of the unified filter with a decay term η but our analysis was confined to discrete time. It is standard to



(a) Stochastic updater Vs Filter synapse on relative escape times

FIGURE 6.7: The escape time of a filter size $\Theta = 5$ is compared against a stochastic updating synapse under various stimulus drift regimes. C_2 denotes the probability that the next stimulus (POT or DEP) is identical to the previous one $Prob[I_2 = \pm I_1] = (1 \pm C_2)/2$. When $C_2 = -1$ the train of stimuli toggles deterministically between POT and DEP. The plotted escape times are relative to the $C_2 = 0$ case where POT or DEP occur with equal probability regardless of the previous stimulus. A clear difference between filter is seen when compared against a matched stochastic updater which has a constant escape time regardless the induction stimulus. This result is trivial, the stochastic updater expresses 1/25 stimuli randomly without any memory of previous events.

examine memory storage in discrete-time because one of the main assumptions in the SPM framework is that the major cause of memory degradation is ongoing memory storage and thus time is irrelevant and events can be fixed to occur at regular time intervals. Typically the tracked memory pattern ξ^α is presented at time $t = \alpha - 1$, which can be $t = 0$ if we assume the synaptic states are in equilibrium due to a long history of memory storage.

Off course fixed time intervals in memory arrival is rather unrealistic, a more realistic approximation to the arrival times of memories in an ongoing memory system would be to assume that storage times follow a continuous-time Poisson process at some fixed rate r. We do not expect the mean memory signal to be affected by this change, but we find that the variance is. Although, the time intervals of memory storage vary in the continuous time framework when a storage event occurs it drives simultaneous changes to all synapses. These synchronous updates give rise to a covariance term even though the synaptic inputs are statistically independent and uncorrelated.

In this section we aim to derive the probability densities $G_j^{\pm}(t)$ for reaching filter thresholds starting from a filter state j in continuous time t for the simpler unified filter without decay. A representation of this model as a continuous-time Markov process can be obtained if the probabilities $g_\pm$ of Figure 6.4 are replaced by rates $rg_\pm$.

To obtain the escape densities we need first to derive the probabilities of internal state transitions from filter state j to filter state i $f_{i|j}(t)$. As we will see later in this analysis, the absence of the decay term makes the structure of the transition matrix tridiagonal allowing us to consider state transitions over filters of arbitrary length Θ. The escape probability densities $G_j^{\pm}$ will then be used to obtain the filter distributions which will eventually lead to consider the mean memory P signal $\mu(t)$ in continuous-time. The derivation presented here has been previously published in Elliott and Lagogiannis (2012).

We will be considering only random uncorrelated memories without any bias between POT/DEP stimuli and thus the probabilities $g_+ = g_- = \frac{1}{2}$. Consequently, we also assume symmetric filters with $\Theta_\pm = \Theta$ so strong and weak synapses are equally present. Initially, we assume $\Theta_\pm$ are absorbing states that once reached do not reset the filter to the zero state. This way we may specifically study the transition probabilities between filter states before any threshold event occurs.

We let the transition matrix $\mathbb{M}^+$ represent a potentiating induction event without considering the re-injection to the zero $+(\Theta-1) \to 0$ state upon a threshold event. Since potentiating events are conventionally taken to conduct a single increment to the filter state then this matrix has zero everywhere and unity along the lower diagonal, $(\mathbb{M}^+)_{I,J} = \delta_{I-1,J}$. Accordingly the depression stimuli matrix has unity on the upper diagonal $\mathbb{M}^- = (\mathbb{M}^+)^T$, also excluding the $-(\Theta-1) \to 0$ transitions due to threshold events. Thus, with equiprobable induction stimuli the overall transition matrix is $\mathbb{M} = \frac{1}{2}(\mathbb{M}^- + \mathbb{M}^+)$.

We let $\mathbf{\Pi} = \left(\pi_{-(\Theta-1)}, \cdots, \pi_{(\Theta-1)}\right)^T$ denote the probability distribution of internal filter states that also excludes threshold events which would return the filter state to zero. The master equation of the evolution of the $\mathbf{\Pi}$ distribution under induction events presented with a Poisson rate r are :

$$\frac{d}{dt}\mathbf{\Pi} = r(\mathbb{M} - \mathbb{I})\mathbf{\Pi}, \tag{6.26}$$

leading to the definition of the matrix $\mathbb{F}$ of the transition probabilities $f_{i|j}(t)$ defined above :

$$\mathbb{F}(t) = \exp\left[(\mathbb{M} - \mathbb{I})rt\right], \tag{6.27}$$

noting that transition probabilities $f_{i|j}(t)$ give the probability of moving from state j to state i in time t without considering any thresholds processes. Computing these elements requires conducting an eigendecomposition of $\mathbb{M}$. The elements of the $\mathbb{M} = \frac{1}{2}(\mathbb{M}^- + \mathbb{M}^+)$ matrix are arranged around the diagonal since $\mathbb{M}^-$ has unity on the upper diagonal and zero elsewhere, while $\mathbb{M}^+$ has unity on the lower diagonal and zero elsewhere. The eigenvalues of $\mathbb{M}$ are the obtained by solving $\det(\mathbb{M} - \lambda\mathbb{I}) = 0$. Since the $\mathbb{M} - \lambda\mathbb{I}$ matrix is symmetric tridiagonal then its eigendecomposition can be obtained by making use of

the fact that its determinants are *continuants* (see Section 3.8.2). We may directly use the results of (3.68) to write down the eigenvalues λ_m by taking into account that here $\beta = \gamma = g_\pm = 1/2$:

$$\lambda_m = \cos\left(\frac{m\pi}{n+1}\right), \text{ for } m = 1 \cdots n \tag{6.28}$$

which can be rewritten in relation to Θ and filter indices as:

$$\lambda_L = \cos\left(\frac{\Theta + L}{2\Theta}\right) \equiv -\sin\frac{L\pi}{2\Theta}, \text{ for } L = -(\Theta), \cdots, +(\Theta - 1). \tag{6.29}$$

We use Equation (3.77) to write down the components of the normalized eigenvectors $\hat{\boldsymbol{e}}^m$ that correspond to λ_m:

$$\hat{\boldsymbol{e}}^m = \sqrt{\frac{2}{n+1}}\left(\sin\frac{m\pi}{n+1}, \sin\frac{2m\pi}{n+1}, \cdots, \sin\frac{nm\pi}{n+1}\right)^T \tag{6.30}$$

where together $\{\hat{\boldsymbol{e}}^1, \cdots, \hat{\boldsymbol{e}}^n\}$ these form an orthonormal basis, $\hat{\boldsymbol{e}}^l \cdot \hat{\boldsymbol{e}}^m = \delta^{l,m}$. The above equation written in terms of filter indices gives :

$$\hat{\boldsymbol{e}}^m = \sqrt{\frac{1}{\Theta}}\left(\sin\frac{(\Theta + L)\pi}{2\Theta}, \sin\frac{2(\Theta + L)\pi}{2\Theta}, \cdots, \sin\frac{(2\Theta - L)(\Theta + L)\pi}{2\Theta}\right)^T, \tag{6.31}$$

for $L = -(\Theta - 1), \cdots, +(\Theta - 1)$. The eigendecomposition allows us to obtain an expression to evaluate $\mathbb{F}(t)$:

$$\mathbb{F}(t) = \sum_{L=-(\Theta-1)}^{+(\Theta-1)} \hat{\boldsymbol{e}}^L (\hat{\boldsymbol{e}}^L)^T \exp\left[-rt\left(1 + \sin\frac{L\pi}{2\Theta}\right)\right] \tag{6.32}$$

with the components,

$$f_{i|j}(t) = \frac{1}{\Theta}\sum_{L=-(\Theta-1)}^{+(\Theta-1)} \sin\frac{(\Theta + i)(\Theta + L)\pi}{2\Theta}\sin\frac{(\Theta + j)(\Theta + L)\pi}{2\Theta} \times \exp\left[-rt\left(1 + \sin\frac{L\pi}{2\Theta}\right)\right] \tag{6.33}$$

We can use the transitions probabilities between internal filter states $j \to i$ to obtain the probabilities of reaching threshold by conditioning on the near threshold probability. When a synapse is next to threshold $j = \pm(\Theta - 1)$ it takes the arrival of a single induction event in the correct direction to lead to a threshold event. Induction events arrive at a Poisson rate r, and these maybe potentiation or depression events with probability $g_+ = g_m = 1/2$ (balanced excitation assumed). In the small time interval δt, a potentiation or depression induction event occurs with probability $1/2r\delta t$, and this could lead a filter to threshold if the filter occupies state $\pm(\Theta - 1)$. Using the probability of being near threshold $f_{+(\Theta-1)|j}$ we may define the probability densities of threshold

events having started from j as :

$$G_j^{\pm}(t) = \frac{1}{2} r f_{\pm(\Theta-1)|j}(t), \tag{6.34}$$

we may write the probability of not reaching threshold in terms of the sum of probabilities of purely internal transitions:

$$H_j(t) = \sum_{i=-(\Theta-1)}^{+(\Theta-1)} f_{i|j}(t), \tag{6.35}$$

or in relation to the probabilities of threshold events:

$$H_j(t) = 1 - \int_0^t d\tau [G_j^+(\tau) + G_j^-(\tau)] \tag{6.36}$$

, which will be used later on.

6.3.4 Equilibrium filter distributions

The transition matrix $\mathbb{M}$ describes the neighbour interactions of internal filter states without any wrap-around or transitions that would return synapses that go over the boundary states in the matrix. Thus, our previous definitions had excluded the threshold events that re-inject to the zero state by assuming that probability leaks through the absorbing states $\Theta_{\pm}$. To conserve probability and obtain the equilibrium distribution we need to include the reset-to -zero events that occur upon transitions to Θ and result to possible changes in strength state depending on the threshold crossed. These reset transitions are to be added to the master Equation (6.26) which we now write $\mathbf{\Pi}$ component wise :

$$\frac{d}{dt}\pi_{\pm(\Theta-1)} = \frac{r}{2}\left[\pi_{\pm(\Theta-2)} - 2\pi_{\pm(\Theta-1)}\right] \tag{6.37}$$

$$\frac{d}{dt}\pi_i = \frac{r}{2}\left[\pi_{i-1} + \pi_{i+1} - 2\pi_i\right], \text{ for } i \neq 0 \;\&\; i \neq \pm(\Theta-1) \tag{6.38}$$

$$\frac{d}{dt}\pi_0 = \frac{r}{2}[\pi_{+1} + \pi_{-1} + \underbrace{\pi_{+(\Theta-1)} + \pi_{-(\Theta-1)}}_{\text{from threshold processes}} - 2\pi_0], \tag{6.39}$$

explicitly adding the contributions of threshold events to the zero state. Solving for

$$\frac{d}{dt}\mathbf{\Pi}_i = 0$$

gives the equilibrium distribution $\boldsymbol{A}$ with components :

$$\pi_i = A_i = \frac{(\Theta - |i|)}{\Theta^2}. \tag{6.40}$$

Figure 6.14(a) shows a scaled equilibrium distribution for a small filter $\Theta = 6$.

A memory $\boldsymbol{\xi^1}$ is encoded on the equilibrium distribution $\boldsymbol{A}$ by applying $\mathbb{M}^{\pm}$ at $t = 0$, after which point the distribution is modified accordingly. Applying the potentiating transition matrix only $\mathbb{M}^{+}\boldsymbol{A}$ will shift the whole distribution to the right towards the upper threshold Θ_{+}, but since $\mathbb{M}^{+}$ does describe threshold events the occupancy of the zero states will not be updated correctly.

We let the vector Δ represent the a filter state distribution where all synapses are in the zero state, namely $\Delta = 0, \cdots, 0, 1, 0, \cdots, 0)^T$. We will use this as an auxiliary vector to represent the re-injection of synapses from threshold to zero filter state. On the example above, after a potentiating stimulus there will be $1/\Theta^2$ synapses that perform a threshold transition and get re-injected to zero from the strong synapses and another $1/\Theta^2$ injected from weak synapses. These weak synapses will then become strong and move to the probability distribution of strong filter states also. Thus, the synapses that receive potentiating stimulus have $2/\Theta^2$ synapses re-injected to zero from the two threshold events. The distribution of filter states for strong synapses then becomes:

$$\boldsymbol{P}^{+} = \mathbb{M}^{+}\boldsymbol{A} + 2\boldsymbol{\Delta}/\Theta^2 \tag{6.41}$$

while the joint probability distribution of synaptic strength and filter states $\boldsymbol{P} = \left(\boldsymbol{P}^{-}|\boldsymbol{P}^{+}\right)$ is described by :

$$\boldsymbol{P} = \frac{1}{2}\left(\underbrace{(\mathbb{M}^{+}\boldsymbol{A})^T}_{S=-1} | \underbrace{(\mathbb{M}^{+}\boldsymbol{A} + 2\boldsymbol{\Delta}/\Theta^2)^T}_{S=+1}\right)^T, \tag{6.42}$$

where weak synapses experiencing a $\xi_i^{\alpha} = +1$ at $t = 0$ perform a shift by $\mathbb{M}^{+}$. This shift also affects strong synapses who additionally receive $2/\Theta^2$ synapses in the zero filter state from the opposite strength. The factor of $1/2$ normalizes the $\boldsymbol{P}$ joint distribution of $\boldsymbol{P}^{-}$ and $\boldsymbol{P}^{+}$ by assuming that the population of synapses is equally distributed between strong and weak synapses in equilibrium. Accordingly, for synapses that received a depression induction stimulus we have :

$$\boldsymbol{D} = \frac{1}{2}\left(\underbrace{(\mathbb{M}^{-}\boldsymbol{A} + 2\boldsymbol{\Delta}/\Theta^2)^T}_{S=-1} | \underbrace{(\mathbb{M}^{-}\boldsymbol{A})^T}_{S=+1}\right)^T. \tag{6.43}$$

The above two distributions are mirror images of each other, $\xi_i^1 = -1$ stimuli shift synapses of $\boldsymbol{D}$ over to the weak strength $S = -1$ and $\xi_i^1 = +1$ stimuli push over to the strong $S = +1$. However we need to keep in mind that the memory signal is measured via :

$$h(t) = \frac{1}{N}\sum_{1}^{N} \xi_i^1 S_i(t) \equiv \frac{1}{N}\sum_{1}^{N} \tilde{S}_i(t).$$

Thus, the size of the signal relates to the difference in the balance of positive and negative terms $\tilde{S}_i(t) = \pm$. All strengths $S_i(t)$ are viewed via a product with the tracked memory ξ_i^1 and thus weak synapses $S_i(t) = -1$ in $\boldsymbol{D}$ contribute positively to the memory signal because they are multiplied by $\xi_i^1 = -1$ when summed, but strong synapses in $\boldsymbol{D}$ would contribute negatively. The converse is true for $\boldsymbol{P}$. This conversion of the summed $S_i(t)$ effectively makes the right half of the $\boldsymbol{P}$ vector and the left half of the $\boldsymbol{D}$ vector contribute positively to the signal. At the same time, the left half of $\boldsymbol{P}$ and the right half of $\boldsymbol{D}$ contribute negatively. Since the signal is a difference between positive and negative contributions then the difference between the vectors $\boldsymbol{P} - \boldsymbol{D}$ is sufficient to measure the signal. Their difference $\boldsymbol{P} - \boldsymbol{D}$ can be written as :

$$\boldsymbol{P} - \boldsymbol{D} = \left(\underbrace{\left(\frac{1}{2}(\mathbb{M}^+ - \mathbb{M}^-)\boldsymbol{A} - \frac{1}{\Theta^2}\boldsymbol{\Delta}\right)^T}_{S=-1} \mid \underbrace{\left(\frac{1}{2}(\mathbb{M}^+ - \mathbb{M}^-)\boldsymbol{A} + \frac{1}{\Theta^2}\boldsymbol{\Delta}\right)^T}_{S=+1} \right)^T \tag{6.44}$$

which captures the initial signal arising from both types of stimuli $\xi^1 = \pm 1$, and will be used below to obtain the mean signal dynamics.

6.3.5 Signal analysis of the unified filter in continuous time

The same methods employed in the analysis of the cascade mean signal of Section 5.5 can be also applied here to integrate out internal filter states thus focusing on strength state transitions. Similarly to the cascade, the P signal is also analysed here and thus the earlier conventions over the relation of stimulus patterns and induction stimuli ξ_i apply. Specifically we let the ξ_i stimuli directly translate to plasticity inducing stimuli of POT (+1) or DEP (-1).

These stimuli result in internal filter state transitions only when the current state j of the filter lies within threshold $-(\Theta-1) < j < +(\Theta-1)$. When j is near either threshold $\pm(\Theta-1)$ then an appropriate stimulus $\xi_i^\alpha = \pm 1$ at time t may generates threshold event that changes the strength state or a $\xi_i^\alpha = \mp 1$ results in an internal filter transition away from the threshold.

$f_{i|j}(t)$ retains its meaning from the previous Sections as the probability of transition from j to i without escape, meaning purely internal filter transitions. Similarly, $G_j^\pm(t)$ represents the escape densities through upper (+)/lower (-) boundaries in time t starting from state j. Using these, we describe the probability of each possible path between a initial internal state j of a synapse under strength state B to a state i of synapse in strength A in time t by $p_{i|j}^{A|B}(t)$. The path from the source to target state may require a threshold crossing event or it may be reached through internal transitions only. The process is renewed after each threshold event at time τ and so we may condition on the time that these events $G^\pm(\tau)$ occur and restart the process from the $j = 0$ state over

the remaining time $p_{i|0}^{\pm|\pm}(t-\tau)$. Initially, we write down the equations that require at least one threshold event in order to move from source state $j^{\pm}$ to target state $i^{\mp}$:

$$p_{i|j}^{-|+}(t) = \int_0^t d\tau G_j^+(\tau) p_{i|0}^{-|+}(t-\tau) + \int_0^t d\tau G_j^-(\tau) p_{i|0}^{-|-}(t-\tau) \tag{6.45}$$

$$p_{i|j}^{+|-}(t) = \int_0^t d\tau G_j^+(\tau) p_{i|0}^{+|+}(t-\tau) + \int_0^t d\tau G_j^-(\tau) p_{i|0}^{+|-}(t-\tau) \tag{6.46}$$

For transitions over the same strength state we have to include the inhomogeneous term $f_{i|j}(t)$, which represent the probability of a purely internal transitions path, in addition to paths that include possible threshold events :

$$p_{i|j}^{-|-}(t) = f_{i|j}(t) + \int_0^t d\tau G_j^+(\tau) p_{i|0}^{-|+}(t-\tau) + \int_0^t d\tau G_j^-(\tau) p_{i|0}^{-|-}(t-\tau) \tag{6.47}$$

$$p_{i|j}^{+|+}(t) = f_{i|j}(t) + \int_0^t d\tau G_j^+(\tau) p_{i|0}^{+|+}(t-\tau) + \int_0^t d\tau G_j^-(\tau) p_{i|0}^{+|-}(t-\tau) \tag{6.48}$$

We note that transitions to the strong state are considered by the pair of equations (6.46) and (6.48), while transitions to the weak strength state by (6.45) and (6.47). The above integral equations contain convolutions that become simple products under a Laplace transform. We treat the above equations in pairs and perform a Laplace transform taking s to be the transformed variable :

$$\hat{p}_{i|j}^{+|+}(s) = \hat{f}_{i|j}(s) + \hat{p}_{i|0}^{+|+}(s)\hat{G}_j^+(s) + \hat{p}_{i|0}^{+|-}(s)\hat{G}_j^-(s), \tag{6.49}$$

$$\hat{p}_{i|j}^{+|-}(s) = \hat{p}_{i|0}^{+|+}(s)\hat{G}_j^+(s) + \hat{p}_{i|0}^{+|-}(s)\hat{G}_j^-(s). \tag{6.50}$$

This pair allows us to determine $\hat{p}_{i|j}^{+|-}(s)$ by setting $j = 0$ and the writing in matrix form:

$$\begin{bmatrix} 1-\hat{G}_0^+(s) & -\hat{G}_0^-(s) \\ -\hat{G}_0^+(s) & 1-\hat{G}_0^-(s) \end{bmatrix} \begin{bmatrix} \hat{p}_{i|0}^{+|+}(s) \\ \hat{p}_{i|0}^{+|-}(s) \end{bmatrix} = \begin{bmatrix} \hat{f}_{i|0}(s) \\ 0 \end{bmatrix} \tag{6.51}$$

and solving for $\hat{p}_{i|0}^{+|\pm}(s)$ we obtain:

$$\begin{bmatrix} \hat{p}_{i|0}^{+|+}(s) \\ \hat{p}_{i|0}^{+|-}(s) \end{bmatrix} = \frac{1}{1-\hat{G}_0^+(s)-\hat{G}_0^-(s)} \begin{bmatrix} 1-\hat{G}_0^-(s) & \hat{G}_0^-(s) \\ \hat{G}_0^+(s) & 1-\hat{G}_0^+(s) \end{bmatrix} \begin{bmatrix} \hat{f}_{i|0}(s) \\ 0 \end{bmatrix} \tag{6.52}$$

, which can be written as:

$$\hat{p}_{i|0}^{+|+}(s) = \frac{1-\hat{G}_0^-(s)}{1-\hat{G}_0^+(s)-\hat{G}_0^-(s)} \hat{f}_{i|0}(s) \tag{6.53}$$

$$\hat{p}_{i|0}^{+|-}(s) = \frac{\hat{G}_0^+(s)}{1-\hat{G}_0^+(s)-\hat{G}_0^-(s)} \hat{f}_{i|0}(s). \tag{6.54}$$

Accordingly, for the other couple of equations over the weak strength :

$$\hat{p}^{-|-}_{i|j}(s) = \hat{f}_{i|j}(s) + \hat{p}^{-|+}_{i|0}(s)\hat{G}^{+}_{j}(s) + \hat{p}^{-|-}_{i|0}(s)\hat{G}^{-}_{j}(s), \tag{6.55}$$

$$\hat{p}^{-|+}_{i|j}(s) = \hat{p}^{-|+}_{i|0}(s)\hat{G}^{+}_{j}(s) + \hat{p}^{-|-}_{i|0}(s)\hat{G}^{-}_{j}(s), \tag{6.56}$$

we obtain solutions going through the same steps as above :

$$\hat{p}^{-|-}_{i|0}(s) = \frac{1-\hat{G}^{+}_{0}(s)}{1-\hat{G}^{+}_{0}(s)-\hat{G}^{-}_{0}(s)}\hat{f}_{i|0}(s) \tag{6.57}$$

$$\hat{p}^{-|+}_{i|0}(s) = \frac{\hat{G}^{-}_{0}(s)}{1-\hat{G}^{+}_{0}(s)-\hat{G}^{-}_{0}(s)}\hat{f}_{i|0}(s). \tag{6.58}$$

We can simplify the above expression if we consider symmetric filters with $\Theta_{\pm} = \Theta$. This symmetry also results in equal escape probabilities from either threshold $G^{+}_{0}(t) = G^{-}_{0}(t)$ and this carries on across states $G^{+}_{+i}(t) = G^{-}_{-i}(t)$. Using Equation (6.36) we may rewrite escape probabilities in terms of the reciprocal non-escape ones as $\hat{G}^{\pm}_{0}(s) = \frac{1}{2}[1 - s\hat{H}_0(s)]$. This substitution simplifies the expressions of our earlier results :

$$\hat{p}^{+|\pm}_{i|0}(s) = \frac{1 \pm s\hat{H}_0(s)}{2s\hat{H}_0(s)}\hat{f}_{i|0}(s). \tag{6.59}$$

We can now substitute this expression for $\hat{p}^{+|\pm}_{i|0}(s)$ back into the Laplace transformed equations for probabilities of transitions between strength dependent filter states. To obtain $\hat{p}^{+|\pm}_{i|j}(s)$ we substitute into (6.49) and (6.50) equations (6.53),(6.54) with $\hat{G}^{\pm}_{0}(s) = \frac{1}{2}[1 - s\hat{H}_0(s)]$, and similarly for $\hat{p}^{-|\pm}_{i|j}(s)$ we substitute into equations (6.56),(6.55) the equations (6.53) and (6.54) :

$$\hat{p}^{+|+}_{i|j}(s) = \hat{f}_{i|j}(s) + \frac{1}{2}\left\{\frac{1-s\hat{H}_j(s)}{s\hat{H}_0(s)} + [\hat{G}^{+}_{j}(s) - \hat{G}^{-}_{j}(s)]\right\}\hat{f}_{i|0}(s) \tag{6.60}$$

$$\hat{p}^{+|-}_{i|j}(s) = \frac{1}{2}\left\{\frac{1-s\hat{H}_j(s)}{s\hat{H}_0(s)} + [\hat{G}^{+}_{j}(s) - \hat{G}^{-}_{j}(s)]\right\}\hat{f}_{i|0}(s) \tag{6.61}$$

$$\hat{p}^{-|+}_{i|j}(s) = \frac{1}{2}\left\{\frac{1-s\hat{H}_j(s)}{s\hat{H}_0(s)} - [\hat{G}^{+}_{j}(s) - \hat{G}^{-}_{j}(s)]\right\}\hat{f}_{i|0}(s) \tag{6.62}$$

$$\hat{p}^{-|-}_{i|j}(s) = \hat{f}_{i|j}(s) + \frac{1}{2}\left\{\frac{1-s\hat{H}_j(s)}{s\hat{H}_0(s)} - [\hat{G}^{+}_{j}(s) - \hat{G}^{-}_{j}(s)]\right\}\hat{f}_{i|0}(s) \tag{6.63}$$

The above Laplace transformed expression describes all possible transitions in our system. These expressions can be embedded in the full transition matrix, which jointly characterizes both strength and filter states. This transition matrix will then be used to derive an expression for the mean signal $\mu(s)$ after the storage of the tracked memory

at $t = 0$:

$$\hat{\mathbb{P}}(s) = \left[\begin{array}{c|c} \hat{\mathbb{P}}^{-|-}(s) & \hat{\mathbb{P}}^{-|+}(s) \\ \hline \hat{\mathbb{P}}^{+|-}(s) & \hat{\mathbb{P}}^{+|+}(s), \end{array}\right]$$

assuming we recover the respective $\mathbb{P}(t)$ in the time domain. The right half of this matrix contains the transition probabilities starting from a strong synapse and the left half the respective transitions starting from a weak synapse.

The initial encoding induced $\xi = +1$ stimuli or a $\xi = -1$ depressing stimulus with equal probability $\text{Prob}[\xi_i^1 = \pm 1] = \frac{1}{2}$. For synapses that experienced $\xi = +1$ at time $t = 0$ we let the joint probability distribution of strength and occupancy be the vector $\boldsymbol{P}$, while the vector $\boldsymbol{D}$ represents the joint distribution of synapses that experienced $\xi = -1$. Beyond initial encoding, the joint distribution of synapses that received $\xi = +1$ evolves as $\mathbb{P}(t)\boldsymbol{P}$. Here, we can obtain the mean strength of the $\boldsymbol{P}$ synapses by summing the distribution with the use of an auxiliary n-th dimensional ($n = 2\Theta - 1$) vector $\boldsymbol{n}$ of $(1...1)^T$ as $(-\boldsymbol{n}|\boldsymbol{n})\mathbb{P}(t)\boldsymbol{P}$, where two $\boldsymbol{n}$ are concatenated in one with the sign denoting the strength of the synapse. When evaluating $(-\boldsymbol{n}|\boldsymbol{n})\hat{\mathbb{P}}(s)$ we find that in essence we are performing the following operations on the submatrices :

$$(-\boldsymbol{n}^T|\boldsymbol{n}^T)\mathbb{P}(s) = (\boldsymbol{n}^T\hat{\mathbb{P}}^{+|-}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|-}(s)|\boldsymbol{n}^T\hat{\mathbb{P}}^{+|+}(s) - \boldsymbol{n}^T\hat{\mathbb{P}}^{-|+}(s)). \tag{6.64}$$

The component-wise differences $[\hat{\mathbb{P}}^{+|-}(s) - \hat{\mathbb{P}}^{-|-}(s)]$ and $[\hat{\mathbb{P}}^{+|+}(s) - \hat{\mathbb{P}}^{-|+}(s)]$ translate to:

$$\sum_{I=-(\Theta-1)}^{+(\Theta-1)} [\hat{p}_{i|j}^{+|-} - \hat{p}_{i|j}^{-|-}] = [\hat{G}_j^+(s) - \hat{G}_j^-(s)]\hat{H}_0(s) - \hat{H}_j(s) \tag{6.65}$$

$$\sum_{I=-(\Theta-1)}^{+(\Theta-1)} [\hat{p}_{i|j}^{+|+} - \hat{p}_{i|j}^{-|+}] = [\hat{G}_j^+(s) - \hat{G}_j^-(s)]\hat{H}_0(s) + \hat{H}_j(s). \tag{6.66}$$

The $G_j^{\pm}(s)$ and $H_j(s)$ components are now written in vector form as $\boldsymbol{G}^{\pm}$ and $\boldsymbol{H}$. Taking into account that $\text{Prob}[\xi_i^1 = \pm 1] = \frac{1}{2}$ and thus synapses are shared between $\boldsymbol{P}$ and $\boldsymbol{D}$ we write the Laplace transform of the mean memory signal as :

$$\hat{\mu}(s) = \frac{1}{2}(-\boldsymbol{n}^T| + \boldsymbol{n}^T)\mathbb{P}(s)(\boldsymbol{P} - \boldsymbol{D}) \tag{6.67}$$

where the difference $\boldsymbol{P} - \boldsymbol{D}$ has been previously expanded in (6.44). Taking into account that

$$((-\boldsymbol{n}^T| + \boldsymbol{n}^T)) \cdot (\boldsymbol{P} - \boldsymbol{D}) = (\mathbb{M}^+ - \mathbb{M}^-)\boldsymbol{A} - \frac{1}{\Theta^2}\boldsymbol{\Delta}$$

where $\boldsymbol{A}$ denotes the equilibrium distribution with components $A_i = (\Theta - |i|)/\Theta^2$ (6.40), we may rewrite the Laplace transform of the mean signal in (6.67) as

$$\hat{\mu}(s) = \frac{1}{2}\hat{H}_0(s)[\hat{\boldsymbol{G}}^+(s) - \hat{\boldsymbol{G}}^-(s)] \cdot (\mathbb{M}^+ - \mathbb{M}^-)\boldsymbol{A} + \frac{1}{\Theta^2}\boldsymbol{\Delta} \cdot \hat{\boldsymbol{H}}(s) \tag{6.68}$$

If the tracked vector pattern consisted of only $\xi_i^\alpha = +1$ induction signals then only the $\boldsymbol{P}$ distribution would contribute to the signal. Yet, with balanced excitation the contribution to the $h(t)$ signal is shared between $\boldsymbol{P}$ and $\boldsymbol{D}$, due to synapses that received either $\xi_i^\alpha = +1$ or $\xi_i^\alpha = -1$ respectively during encoding of the tracked pattern. We may then take the $\boldsymbol{P}$ contribution twice and remove the $\boldsymbol{D}$ to simplify the expression for the signal and writing it as :

$$\hat{\mu}(s) = \hat{H}_0(s)[\hat{\boldsymbol{G}}^+(s) - \hat{\boldsymbol{G}}^-(s)] \cdot \mathbb{M}^+\boldsymbol{A} + \frac{1}{\Theta^2}\hat{H}_0(s), \tag{6.69}$$

noting we have removed the vector notation from the re-injected to zero components. This equation may be further simplified by taking into account that the $G^+_{+j} = G^-_{-j}$ components of $\hat{\boldsymbol{G}}^+(s) - \hat{\boldsymbol{G}}^-(s)$ reverse sign when reflected about the zero state $j = 0$:

$$\hat{G}^+_{+j}(s) - \hat{G}^-_{+j}(s) = \hat{G}^-_{-j}(s) - \hat{G}^+_{-j}(s) = -[\hat{G}^+_{-j}(s) - \hat{G}^-_{-j}(s)], \tag{6.70}$$

the expansion of the dot product can then be written in terms of a sum from $j = 1 \cdots (\Theta - 1)$ to become :

$$\left[\boldsymbol{G}^+(t) - \boldsymbol{G}^-(t)\right] . \mathbb{M}^+\boldsymbol{A} = \frac{2}{\Theta^2}\sum_{j=1}^{\Theta-1}[\boldsymbol{G}^+_j(t) - \boldsymbol{G}^-_j(t)] \tag{6.71}$$

and used to write the Laplace transformed signal :

$$\hat{\mu}(s) = \frac{1}{\Theta^2}\hat{H}_0(s)\left\{1 + 2\sum_{j=1}^{\Theta-1}[\boldsymbol{G}^+_j(t) - \boldsymbol{G}^-_j(t)]\right\} \tag{6.72}$$

In principle, this form allows a Laplace inversion, by recognizing that the above products become convolutions in the time domain :

$$\mu(t) = \frac{1}{\Theta^2}\left\{H_0(t) + 2\sum_{j=1}^{\Theta-1}\left[\int_0^t d\tau H_0(t-\tau)[G^+_j(\tau) - G^-_j(\tau)]\right]\right\} \tag{6.73}$$

6.4 Filter mean signal dynamics

The above analysis gives us an expression for the mean signal (6.73) in the time domain. However, evaluating such integrals using expressions for $H_0(t)$ and $G^\pm_j(t)$ from equations (6.33), (6.34) and (6.36) can be very difficult and can be avoided by computing the inverse Laplace transform of Equation (6.72). To do so we need first to compute the inverse Laplace transforms of $\hat{H}_0(s)$ and $\hat{G}^\pm_j(s)$. From Equation (6.36) we see that the probability of not reaching threshold $\hat{H}_0(s)$ is the reciprocal of the escape densities from reaching either threshold $\hat{G}^\pm_j(s)$ and thus we need to focus on calculating these terms. Previously, when we examined the filter with decay in discrete time, we defined escape probabilities in Equation (6.11) and used generating functions to solve the recurrence relations for escape from an interval I. In Elliott (2011a) equation (3.5), a similar

recurrence relation is also defined in continuous-time, for the filter with no decay using the Laplace transformed $\hat{G}_j(s)$, that satisfies :

$$s\hat{G}_j(s) = \frac{1}{2}r[\hat{G}_{j+1}(s) - \hat{G}_j(s)] + \frac{1}{2}r[\hat{G}_{j-1}(s) - \hat{G}_j(s)] \tag{6.74}$$

and in a similar method to the discrete time, we set boundary conditions in order to calculate $\hat{G}_j^+(s)$ or $\hat{G}_j^-(s)$, by letting $\hat{G}_{-\Theta}(s) = 0$ and $\hat{G}_{+\Theta}(s) = 1$ or $\hat{G}_{-\Theta}(s) = 1$ and $\hat{G}_{+\Theta}(s) = 0$.

Using a characteristic equation approach to solve (6.74), we write :

$$s\Phi(s) - \frac{1}{2}r[\Phi(s)^2 - \Phi(s)] - \frac{1}{2}r[1 - \Phi(s)] = 0 \tag{6.75}$$

$$-\frac{1}{2}r\Phi(s)^2 + \Phi(s)[s + r] - \frac{1}{2}r = 0 \tag{6.76}$$

giving roots :

$$\Phi_\pm(s) = \frac{(s + r) \pm \sqrt{(s + r)^2 - r^2}}{r} \tag{6.77}$$

giving $\Phi_+(s) + \Phi_-(s) = 2(s/r + 1)$ and $\Phi_+(s)\Phi_-(s) = 1$. In general :

$$G_j(s) = A\Phi_+(s)^j + B\Phi_-(s)^j \tag{6.78}$$

and solving given boundary conditions $\hat{G}_{\pm\Theta}(s)$:

$$\begin{bmatrix} \Phi_+(s)^{(+\Theta)} & \Phi_-(s)^{(+\Theta)} \\ \Phi_+(s)^{(-\Theta)} & \Phi_-(s)^{(-\Theta)} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \hat{G}_{+\Theta} \\ \hat{G}_{-\Theta} \end{bmatrix} \tag{6.79}$$

Inverting the matrix and using the fact stated earlier that $\Phi_+ = \Phi_-^{-1}$ we write :

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \begin{bmatrix} \Phi_-(s)^{(-\Theta)} & -\Phi_-(s)^{(+\Theta)} \\ -\Phi_+(s)^{(-\Theta)} & \Phi_+(s)^{(+\Theta)} \end{bmatrix} \begin{bmatrix} \hat{G}_{+\Theta} \\ \hat{G}_{-\Theta} \end{bmatrix} \tag{6.80}$$

we calculate A and B for each boundary condition :

$$A = \frac{\Phi_-(s)^{(-\Theta)}}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \quad B = \frac{\Phi_-(s)^{(-\Theta)}}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \text{ with } \begin{cases} \hat{G}_{+\Theta}(s) = 1 \\ \hat{G}_{-\Theta}(s) = 0 \end{cases} \tag{6.81}$$

$$A = \frac{-\Phi_-(s)^{(+\Theta)}}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \quad B = \frac{\Phi_-(s)^{(+\Theta)}}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \text{ with } \begin{cases} \hat{G}_{+\Theta}(s) = 0 \\ \hat{G}_{-\Theta}(s) = 1 \end{cases} \tag{6.82}$$

the solution Equation (6.78) can now be written :

$$\hat{G}_j^{\pm}(s) = \frac{\Phi_+(s)^{(\Theta \pm j)} - \Phi_-(s)^{(\Theta \pm j)}}{\Phi_+(s)^{(2\Theta)} - \Phi_-(s)^{(2\Theta)}} \tag{6.83}$$

and we may now put this expression for $\hat{G}_j^{\pm}(s)$ and $\hat{H}_0(s) = [1 - \hat{G}_0^{+}(s) - \hat{G}_0^{-}(s)]/s$ into Equation (6.72) and evaluating the resulting geometric series we obtain:

$$\hat{\mu}(s) = \frac{2}{r\Theta^2} \left\{ \frac{\Phi_+(s)^{(\Theta)} - 1}{\Phi_+(s) - 1} \right\}^3 \frac{\Phi_+(s)[1 + \Phi_+(s)]}{\{1 + [\Phi_+(s)]^{2\Theta}\}\{1 + [\Phi_+(s)]^{\Theta}\}}, \tag{6.84}$$

which after computing the inverse Laplace transform gives :

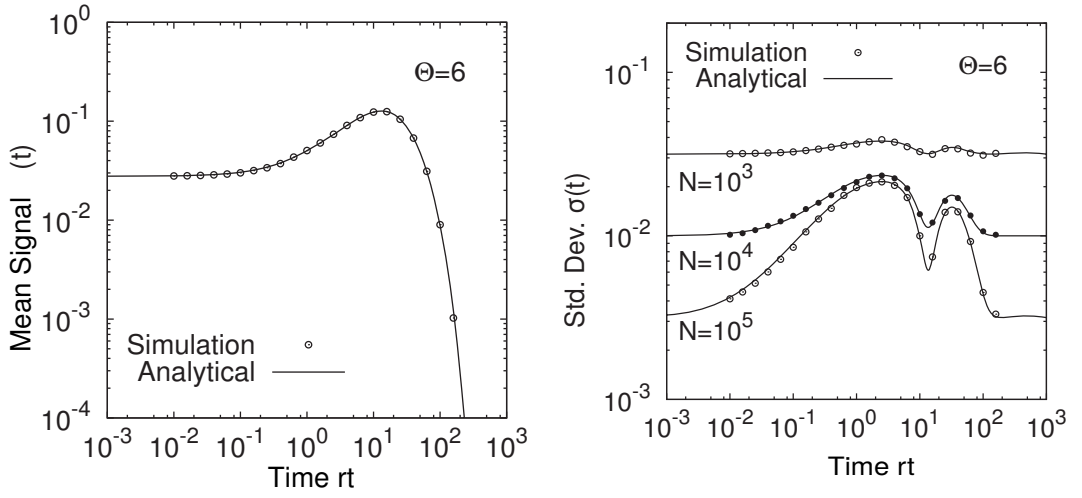
$$\begin{aligned} \mu(t) = \frac{1}{\Theta^3} \sum_{L=0}^{\Theta-1} \cot\left(\frac{2L+1}{4\Theta}\pi\right) \exp\left\{-rt\left[1 - \cos\left(\frac{2L+1}{4\Theta}\pi\right)\right]\right\} \\ -\frac{4}{\Theta^3} \sum_{L=0}^{q_\Theta} \cot^2\left(\frac{2L+1}{2\Theta}\pi\right) \exp\left\{-rt\left[1 - \cos\left(\frac{2L+1}{4\Theta}\pi\right)\right]\right\} \end{aligned} \tag{6.85}$$

where $q_\Theta = (\Theta - 2)/2$ for Θ even and $(\Theta - 1)/2$ for odd. Instead of direct substitution in the time-domain, an alternative method that results in directly evaluating the inverse Laplace transforms of (6.69) can also be found in Elliott and Lagogiannis (2012), first described in (Elliott, 2011a).

As discussed earlier in Section 4.5.3, the perceptron's output $h(t)$ consists of an average over N of the $\tilde{S}_i(t)$ identically distributed random variables, the variance therefore is given by:

$$\sigma(t)^2 = \frac{1}{N}\mathrm{Var}[\tilde{S}_i(t)] + \left(1 - \frac{1}{N}\right)\mathrm{Cov}(t) \tag{6.86}$$

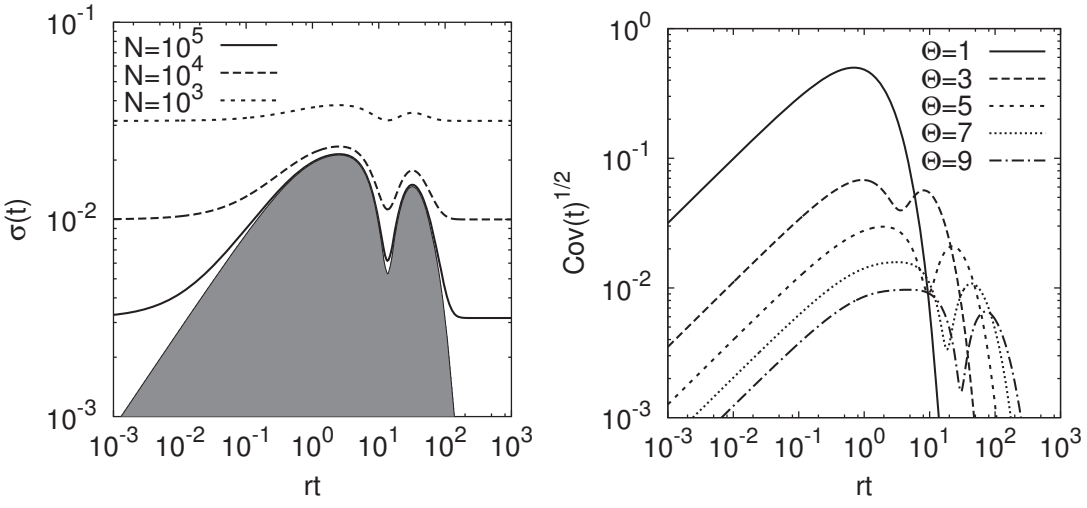
In any case, having obtained an expression $\mu(t)$ we may plot the signal and compare it against simulations of the P signal $h(t)$ with filter synapses. We use the same simulation methods as with the cascade model; a pseudo-code of the simulation algorithm for a filter synapse can be found in Appendix 1. In summary, filter simulations begin by initializing a population of N filter synapses randomly setting the state of each synaptic filter according to the equilibrium distribution $\boldsymbol{A}$ for internal states and randomly choosing synaptic strength between strong and weak with equal probability. Depending on whether simulations are run in discrete-time or continuous time, simulated time is advanced either in fixed or in exponentially distributed intervals. A memory storage event takes place at every time point using randomly generated stimuli $\xi_i = \pm 1$ to represent stored patterns $\boldsymbol{\xi}^{\alpha}$. The first pattern stored is assumed to be the tracked pattern $\boldsymbol{\xi}^{\mathbf{1}}$ and its recall $h(t)$ is tested at each time point after memory storage. The measured



(a) Mean signal continuous-time

(b) Verification std. dev. in continuous-time simulation

FIGURE 6.8: Comparing mean $\Theta = 6$ U. filter signal without decay ($\eta = 0.0$) between analysis and simulation of $N = 10^5$ and averaged over $T = 10^3$ in continuous time. The transitions in the filter occur deterministically conditioned to the arrival of POT/DEP stimuli rates, thus the $g_\pm$ are given by the rate of plasticity induction, here set to $rt = 1$.



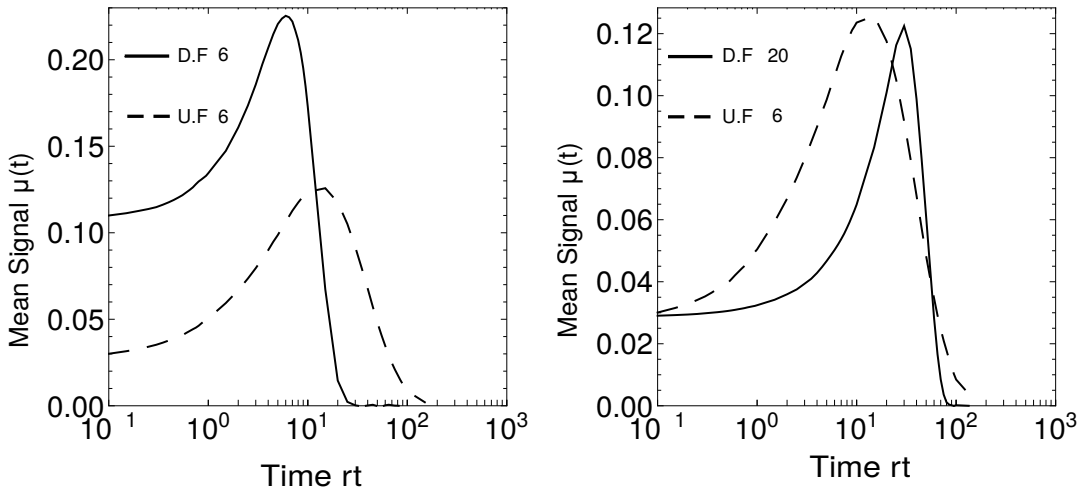
(a) Convergence to core of Cov. in continuous-time

(b) Convergence to core of Cov. in continuous-time

FIGURE 6.9: **a.** Convergence of the covariance of a $\Theta = 6$ filter to an irreducible core as the number of synapses N grows. **b.** Dependence of covariance on filter-size Θ, showing overall that it decreases in overall magnitude as filters become larger but this decrease slows down with Θ.

$h(t)$ signal from a single trial is expected to be noisy and thus we average over T independent trials to obtain a mean signal along with its variance. For the number of trials T is chosen according to the product $N \times T = 10^8$, as measuring a signal over larger sets of synapses N improves the averaging on every trial and thus less trials are required to obtain good estimates of the mean signal.

The memory dynamics with a symmetric synaptic filter (LP-filter) of size $\Theta = 6$ are



(a) Matched $\Theta = 6$

(b) Matched expression time $\tau \approx 36$

FIGURE 6.10: Comparison between unified-filter signals and dual-filter signals for $\Theta = 6$ in simulation using $N = 10^4$ synapses. The dual-filter integrator also exhibits a signal rising phase. **a**. When the two models are matched on $\Theta = 6$, the dual-filter has higher initial signal as the expression time is shorter, also the peak signal for the dual filter is around $rt \approx 6$ while the U.filter gives $rt \approx 13$. **b**. We require $\Theta = 20$ in the dual-filter to match the expression time of a $\Theta = 6$ U.filter. Although the peaks are similar, we notice that the area under the dual-filter curve is smaller than the U.filter and the peak is relatively delayed followed by a sharp signal decay.

shown on Figure 6.8 where we also compare analytical and simulated $\mu(t)$ signal in continuous time. We find exact agreement for both mean signal and standard deviation $\sigma(t)$. Most notably, we find that the filter signal dynamics *rise* before they begin to decay. Such dynamics come in contrast to the classic belief that memory can only decay under new memory storage, here the recall ability of the tracked memory initially improves with further memory storage. A plot of the simulated dual-filter signal on Figure 6.10(a) shows that it also exhibits a rising signal but its profile is different to an LP-filter of the same size $\Theta = 6$. The initial signal of the dual-filter $\Theta = 6$ signal is higher while its lifetime is shorter. Although matched on Θ, it appears we are comparing two very different models as these utilize the internal filter states very differently. Perhaps matching the models on escape time is more relevant in comparing signal dynamics, as such matching (as close as possible) would give at least equal initial signal. Figure 6.10(b), compares a dual-filter $\Theta = 20$ against an LP-filter $\Theta = 6$ showing that these two different filter sizes give approximately the same initial signal but the rising dynamics are different, with the LP-filter dominating over the area of recall.

The rising mean signal may be very promising but the nature of this signal is stochastic and therefore subject to variance. When evaluating recall dynamics therefore, the variance in the signal may be high enough so in some realizations the signal actually falls below zero. Measuring SNRs ($\mu(t)/\sigma(t)$) may therefore be a more relevant quantity for evaluating recall, an SNR =1 would mean that the signal is one standard deviation

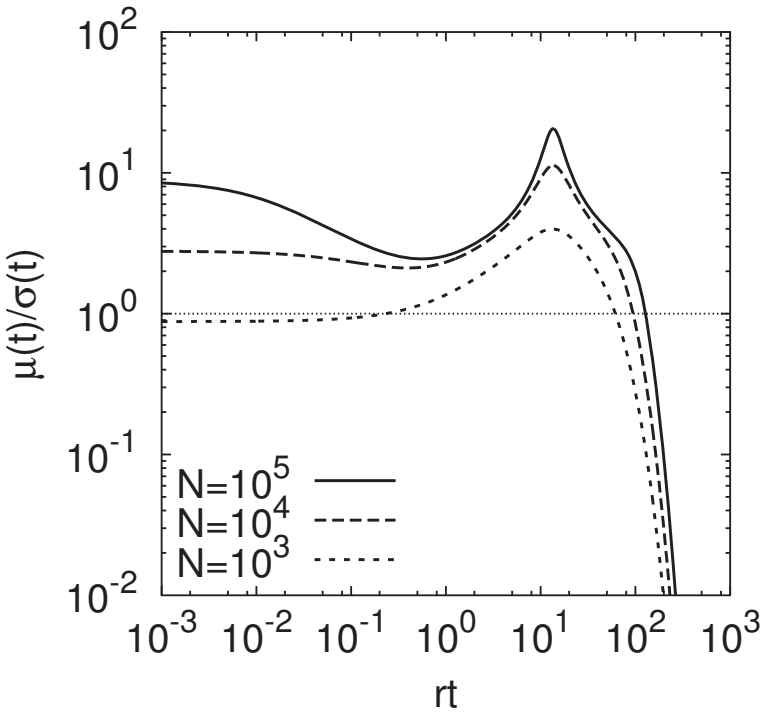


FIGURE 6.11: The SNR of unified filter using exact $\sigma(t)$, although covariance rises the filters still exhibit a signal peak

above the zero signal and thus it is likely that on any particular realization there will not be any signal. The rising recall dynamics could be compromised under SNR if the variance of the signal also increases. Although plotting SNR in continuous time shows that the covariance also exhibits a peak near the time of the peak mean signal, we find that the SNR ($\mu(t)/\sigma(t)$) dynamics retain the rising dynamics (see Figure 6.11). Further examination of $\sigma(t)$ under increasing synapse numbers N shows that it drops with increasing N up to a remaining core that belongs to the covariance, see Figure 6.9(a). Increasing the filter size Θ also decreases the noise, revealing the covariance-core for large $\Theta > 9$, see Figure 6.9(b).

But why is there a rising signal with synaptic filters? The key to understanding these dynamics is to recognize that the overall signal is composed of two separate parts. The first one exists due to the immediate encoding of the tracked memory by synapses that where near threshold. According to the $\boldsymbol{A}$ distribution there are $1/\Theta^2$ synapses near threshold, and on average these would be equally divided between strong and weak strength. Upon encoding of the tracked memory approximately N/Θ^2 synapses would cross threshold and reset to the zero state. Half of them would become or remain strong because they received a potentiating stimulus and the other half would become or remain weak because they were pushed through threshold by a depression stimulus. This pool of synapses, we refer to it as Δ, forms the initial signal and as with non-integrative synapses, the initial signal is produced by an immediate change of strength under the encoding of the tracked memory. Once encoded, the signal ($\mu_\Delta(t)$) monotonically decays in time under further memory storage. The decay occurs because these Δ synapses move

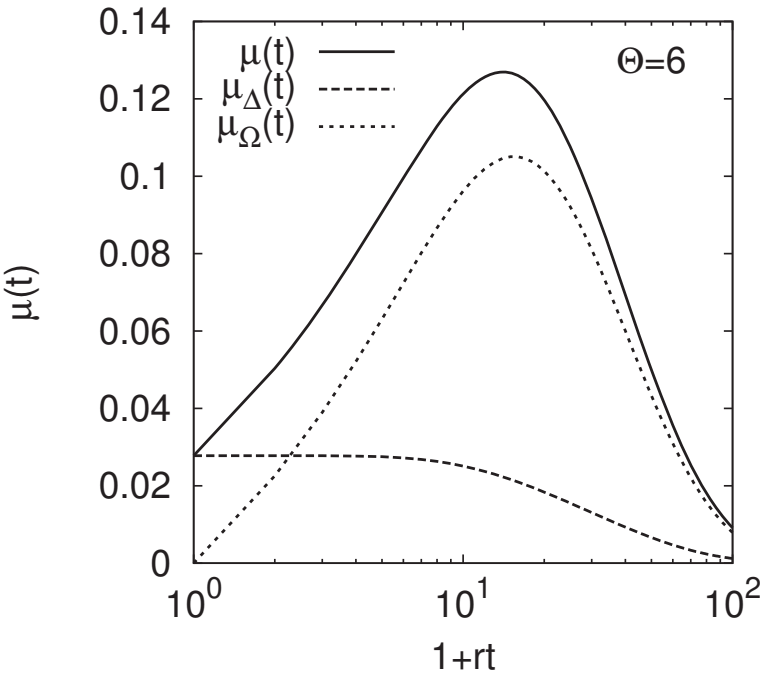


FIGURE 6.12: Analysis of filter signal $\mu(t)$ into the distinct components of initial encoding $\mu_\Delta(t)$ and the latent rising signal $\mu_\Omega(t)$. $\mu_\Delta(t)$ results from synapses that where near threshold and where pushed towards threshold during the encoding of the tracked memory. The state of these synapses is reset to zero and after a natural refraction period relating to Θ they begin to monotonically decay. The $\mu_\Omega(t)$ component is from synapses that did not reach threshold upon the encoding of the tracked memory but had an internal filter state change that biased these synapses to express towards the desired threshold at a later time, thus giving this rising signal up to a peak.

from the zero state randomly towards either threshold expressing either strength. There exists a small refractory period before a synapse can first reach threshold from the zero state which is equal to Θ. Beyond that period, the signal $\mu_\Delta(t)$ monotonically decays in a similar fashion to non-integrating synapses but not exponentially (see Figure 6.12).

The other subset of synapses, which did not reach threshold when the tracked memory was encoded at $t = 0$, nevertheless conducted an internal filter state transition. This internal transition formed a "memory" of the induction history in an $N(1 - \Theta^2)$ pool of synapses we call Ω. Without any plasticity expressed the signal $\mu_\Omega(t)$ arising from this pool of synapses is zero on average. This is because the Ω group of synapses that received a potentiating stimulus is equally divided between strength states and this is also true for the synapses that received a depression stimulus. Nevertheless, all potentiated synapses in Ω, we denote these by Ω^+ regardless of strength, had their internal filter state incremented moving them closer towards expressing through the upper boundary. Similarly, synapses that received a depression stimulus, which we denote by Ω^-, moved towards the lower boundary. Each step taken closer to threshold however, increases the probability of escape towards that threshold. This increase in the probability biases synapses to express plasticity over the threshold dictated by the tracked memory.

We may quantify this biasing using the distribution $\boldsymbol{A}$ and the probability of escape towards one of the thresholds, the upper one for example π_j^+. The escape probability rises linearly towards the associated threshold as $\pi_j^{\pm} \equiv (\Theta \pm j)/(2\Theta)$ (see Figure 6.13). The probability that all Ω^+ synapses reach the upper threshold is :

$$\frac{\sum_{j=-(\Theta-1)}^{+(\Theta-1)} \pi_j^+ A_{j-1}}{\sum_{j=-(\Theta-1)}^{+(\Theta-1)} A_{j-1}} = \frac{1}{2} + \frac{1}{2(1+\Theta)} > \frac{1}{2} \tag{6.87}$$

Figure 6.13 shows an example of a biased distribution $\boldsymbol{A}$ for a filter size $\Theta = 6$ and the probability of escape π_j^+ overlayed. A larger fraction of the Ω^+ population will

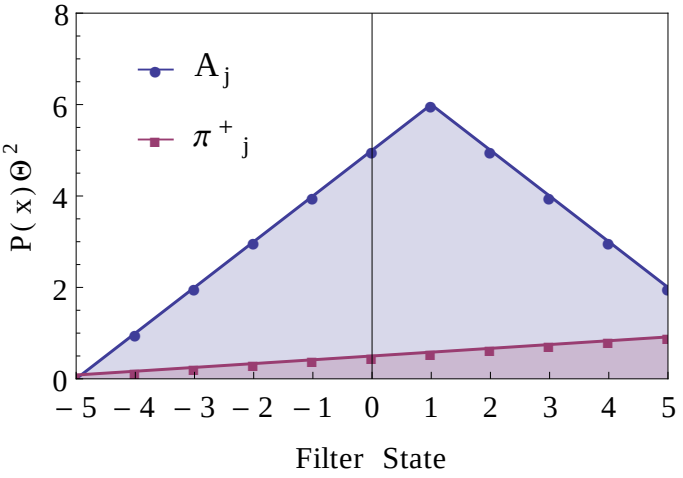


FIGURE 6.13: Biased filter distribution and escape probability over Θ_+ for a filter of size $\Theta = 6$. Vector $\boldsymbol{A}$ represents a biased filter state distribution (scaled by Θ^2) after a single encoding of memory. We see that the distribution has shifted one position towards the right. The shifting causes $1/\Theta^2$ synapses that crossed threshold after encoding to be re-injected to zero state (not-shown). The probability of escape π_j^+ towards the desired threshold is also shown as a linear function ranging from 0 on the lower threshold to 1 on the Θ_+ threshold.

cross the designated Θ_+ threshold rather the Θ_- because the probability of expression towards the upper threshold is more than $1/2$. As Ω^+ synapses reach the designated threshold, the $\mu_\Omega(t)$ signal rises. The $\mu_\Omega(t)$ component is shown on Figure 6.12 as a lagged rising response along with the decaying $\mu_\Delta(t)$ component and the total memory signal $\mu(t) = \mu_\Delta(t) + \mu_\Omega(t)$.

These two separate components can also be seen in the analytical expression for $\mu(t)$. We may extract the separate signals by decomposing (6.73) into two parts :

$$\mu_\Delta(t) = \frac{1}{\Theta^2} H_0(t) \equiv \frac{1}{\Theta^2} \left\{ 1 - \int_0^t d\tau [G_0^+(\tau) - G_0^-(\tau)] \right\} \tag{6.88}$$

$$\mu_\Omega(t) = \frac{2}{\Theta^2} \sum_{j=1}^{\Theta-1} \int_0^t d\tau H_0(t-\tau)[G_j^+(\tau) - G_j^-(\tau)] \tag{6.89}$$

Overall, we have an initial signal Δ that arises due to an asymmetry in the occupancies between strengths in both distributions, those that received potentiation $\boldsymbol{P}$ and those that received depression $\boldsymbol{D}$ induction stimuli by the tracked memory pattern ξ^1. The same encoding event also introduces an asymmetry in the internal filter distributions giving rise to the biasing of the Ω^+ and Ω^- population of synapses to express over the Θ_+ and Θ_- respectively. The biasing actually enhances the strength asymmetry in time, as biased synapses reach the respective threshold. A further implication of this is that a signal *rise can be cancelled* by reversing the induction signals caused by the previous memory the biasing can be immediately be removed.

The rising dynamics of memory can only be seen as the expression of a particular memory expression mechanism and not the one observed under a repertoire of potential molecular pathways for each memory phase. In that sense although our filter's dynamics may look surprising, we find that particular phases of memory arise in separate waves that exhibit both a rise and a falling phase; while during the transition from the falling phase of one to the rising phase of the next memory recall may exhibit lapses and be susceptible to disruption only during the initial rising-phase (Marra et al., 2013), a behaviour that is analogous to our filter synapses dynamics if one considers that the susceptibility to remove the state-bias causing the signal rise is most sensitive immediately after initial encoding. The molecular pathways could operate in series and thus the initial encoding for some memory phase could be taken to be the point after which one pathway stimulates the next.

Finally, the bulk of the biased synapses reach the threshold and become re-injected to the zero-state, beyond which point the distribution of internal filter states slowly resymmetrizes and thus the biasing is reduced. Once the internal distributions completely resymmetrize then biasing is lost and the signal no longer rises, thus exhibiting a peak beyond which point the signal only decays. The decay of the signal represents a phase over which occupancy differences between strength states of $\boldsymbol{P}$ and $\boldsymbol{D}$ resymmetrize. A few snapshots of the dynamics of the internal distributions for weak and strong synapses of $\boldsymbol{P}$ are shown on Figure 6.14 for a $\Theta = 6$ filter. The effects of storing a tracked pattern on internal and between strength states in distributions of $\boldsymbol{P}$ are shown. The internal distributions resymmetrize for this filter around $t_p \approx 13$ and by $t = 64$ we observe that the strength states have resymmetrized as well.

The location of the $\mu(t)$ signal's peak can be found approximately by taking the two slowest-decaying exponentials terms from the positive and negative parts of the signal in the limit of large Θ. The peak occurs at $rt_{peak} \approx \alpha\Theta^2$, where $\alpha = 8\log_e 4/3\pi^2 = \approx 0.375$, and has an amplitude $\mu_{peak} \approx \beta\Theta^{-1}$ with $\beta = 6\sqrt[3]{2}/\pi^2 \approx 0.766$. The initial signal given by $\mu_\Delta(t)$ simply relates to the average number of synapses near threshold and thus according to the equilibrium A_Θ it is Θ^{-2}. Therefore the signal rises from Θ^{-2} to $\beta\Theta^{-1}$ approximately Θ times.

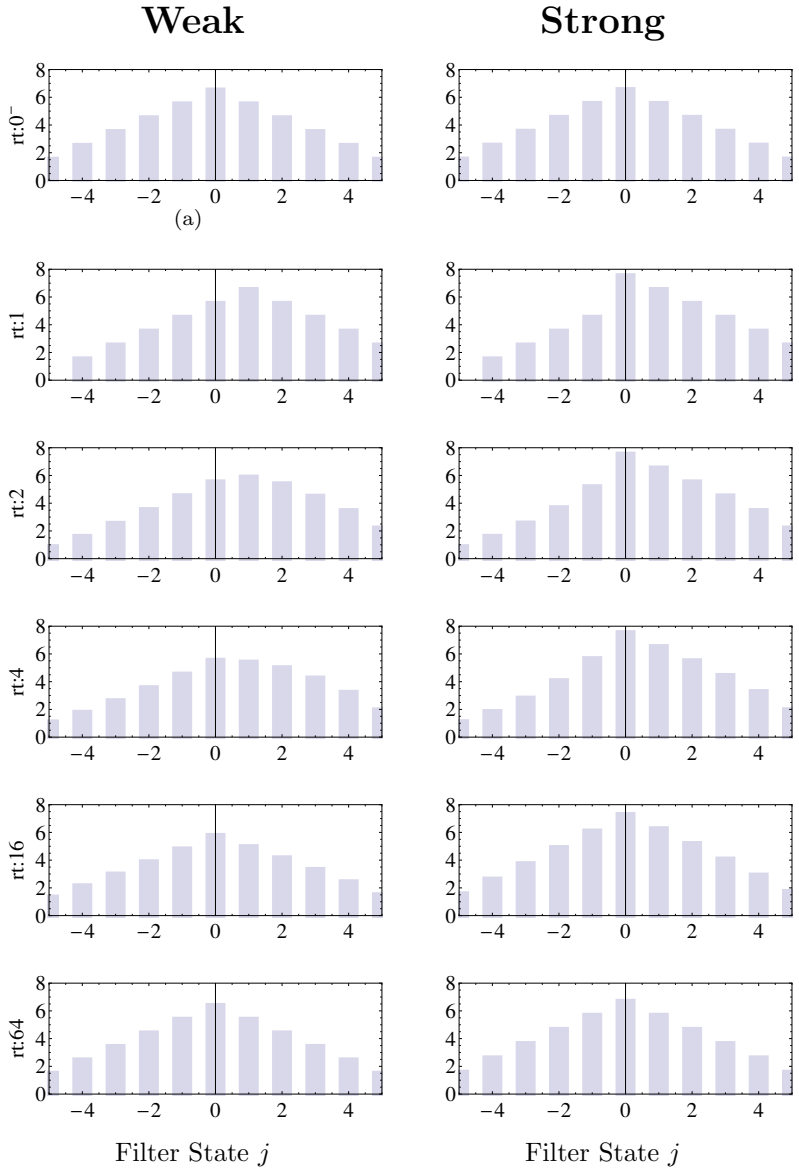


FIGURE 6.14: Equilibrium distribution of filter states is shown before memory storage at $rt : 0$ for weak and strong synapses. Following the storage of the 1st memory ξ^1 right after time $t = 0s$, the sequence of graphs going down shows at each rt the evolution of probability distributions for a filter of size $\Theta = 6$. In the example shown, all synapses receive a potentiation signal after the storage of the tracked pattern $\xi_i^1 = +1$ at time $t = 0 + \epsilon$. The effect of storage is seen at time $t = 1$, where all state occupancies have been shifted towards the upper filter threshold. The weak synapses that were at $(\Theta - 1)$ are forced through the upper threshold to change strength and get injected to the zero state of the strong synapse distribution. A strong synapse that crosses the upper threshold gets its filter state reset to zero but there is no change of strength. The opposite events would have occurred if the storage of the 1st memory consisted of all depression signals $\xi_i^1 = -1$. In time under the effect from the storage of other random memories causing threshold events, the distributions resymmetrize. Here, resymmetrization of internal filter states and strengths is reached around $t = 64$.

6.5 Summary

In this chapter we have argued that synapses would need to process the history of plasticity inducing stimuli in order to efficiently solve the stability versus plasticity dilemma (Abraham and Robins, 2005). For this reason we have extended models of integrate-and-express plasticity to a synaptically stored memory framework. Such models have first been introduced to control fluctuations in developmental plasticity (Elliott, 2008, Elliott and Lagogiannis, 2009, Elliott, 2011a). In these models, the processes of plasticity induction and expression are separated allowing synapses to integrate plasticity stimuli before expressing any changes in strength. The expressed plasticity is therefore dependent on the history of plasticity induction. We have initially considered a model that integrates potentiating and depressing stimuli separately in two filters, the q-filter and p-filter, running in parallel. These filters are each composed of a number of discrete states that essentially count the number of stimuli received. We called this model the dual-filter, the plasticity expressed depends on which filter reached threshold first. Once plasticity is expressed both filters are reset to the zero-state, therefore restarting the race between the plasticity inducing stimuli. In the absence of a decay regressing the synaptic state to zero, the maximum number of stimuli this filter can "remember" is $2\Theta - 1$, after which point both integrators are full and necessarily express the next incoming stimulus depending on which filter it belongs to.

We next considered a single integrator with a lower and an upper threshold $\Theta_{\pm}$ that integrates opposing stimuli effectively introducing competition for expression between them. Such models have been previously employed to implement a low-pass filter of induction stimuli for the control of fluctuations in a developmental plasticity (Elliott and Lagogiannis, 2009, Elliott, 2011a). We focused on a version of this model in the absence of a decay processes and presented an exact analysis of its mean signal in a perceptron framework. We find remarkable signal dynamics exhibiting both a rise and a fall in SNR. The signal can be analysed into two separate components, a $\mu_{\Delta}(t)$ component represents the initial signal expressed by N/Θ^{-2} synapses on average that were near threshold before the storage of the tracked memory and crossed threshold due to its encoding. These synapses are reset to the zero state over the new strength and the signal trace they hold decays as these synapses begin to randomly move towards either threshold under further memory storage. However, there is a second signal component μ_{Ω} that is only latently expressed. This signal arises from $N(1 - \Theta^{-2})$ synapses that were away from threshold when the tracked memory was being encoded. Although these synapses did not express plasticity immediately, plasticity induction drove transitions in the internal filter states moving them on state closer to the threshold dictated by the tracked memory, thereby biasing their future expression to give a latent signal rise. The time of the signal peak is $t_p \approx 0.375\Theta^2$ and the amplitude $\mu_{peak} \approx 0.766\Theta^{-1}$. The synchronous updates of all synapses in the continuous time memory framework gives rise

to a covariance term that nevertheless does not comprise the observed signal dynamics in SNR.

The signal rises Θ-fold from initial signal to peak offering promising dynamics to resolve constraints in the minimum learning rate. If the learning rate is too slow then the initial signal is too low to allow recall, however lowering the learning rate prolongs memory lifetimes as the forgetting rate is also reduced. In other models where the signal exhibits only decaying dynamics this relation between initial signal and future recall poses tight constrains on the minimum learning/forgetting rate that can be used. However, the rising peak of filters may soften these requirements. We discuss the biophysical relevance of our model and an experimental protocol that could reveal if such filters operate in synapses in final chapter of this thesis. In the next chapter we compare filter synapses against stochastic updating (Tsodyks, 1990) and cascade synapses (Fusi et al., 2005).

Chapter 7

Comparing cascade to filter models

In the previous chapter we showed the memory signal dynamics of synapses that integrate plasticity-inducing stimuli before expressing any overt strength changes. The recall dynamics of these synapses exhibit both a rising and a decaying phase. In this chapter we compare the cascade model (Fusi et al., 2005) against filter synapse models on memory capacity.

The transitions down the cascade occur stochastically in response to induction stimuli and thus the model advocates a stochastic view of the nature of the synapse. In this chapter we aim to challenge this view of synaptic plasticity against our filter synapses that express plasticity through a simple computation on the recent history of induction stimuli.

However, behind the cascade's ability to prolong memory lifetimes is also its ability to capture correlations in the pattern of potentiation and depression events, as a simple population of stochastic updating synapses with the identical repertoire of transition probabilities has lower capacity (Fusi et al., 2005). Thus the model not only relies on expressing multiple timescales for forgetting but also in storing the history of the induction stimuli in the serial arrangement of states going down the cascade. Yet, stochastic transitions are not well suited to detect drifts in the sequence of induction stimuli and changing the cascade to deterministic transitions would severely impact performance, as any single stimulus would be able switch a synapse to the opposite cascade. We argue that filters are naturally suited to this task as these read the induction history and express persistent trends in stimulation patterns without relying on stochastic transitions. Thus, we examine if the mean-memory lifetimes improve by replacing stochastic transitions in cascades with filters of a matched mean-expression-time. For this reason we establish a filter cascade-equivalent that utilizes the cascade structure in order to exhibit progressive stabilization but instead of stochastic transitions each cascade-state

expresses plasticity according to a filter process. By matching the expression times of each filter-cascade state to the original cascade we can examine the merits of filters to stochastic transitions by comparing models, of the same cascade size n, on mean memory lifetimes.

Next, we directly compare single filter synapses, not embedded in a cascade structure, against cascade synapses with the same number of states $2n = 2\Theta - 1$. Again, capacity is gauged by measuring the lifetime of the mean signal $\mu(t)$ but here we extend our approach to consider mean memory lifetimes. These are obtained by measuring the mean first passage time (MFPT) of the stochastic signal below a set threshold. Finally, we present indicative capacity results using Hopfield (Hopfield, 1982) network to show that our earlier mean signal lifetime results transfer to capacity measured in a network that exhibits recall dynamics. We begin with a discussion on how filter-cascades can be constructed to match the original cascade model and then proceed to compare memory lifetimes between synapse models.

7.1 Matching filter cascades

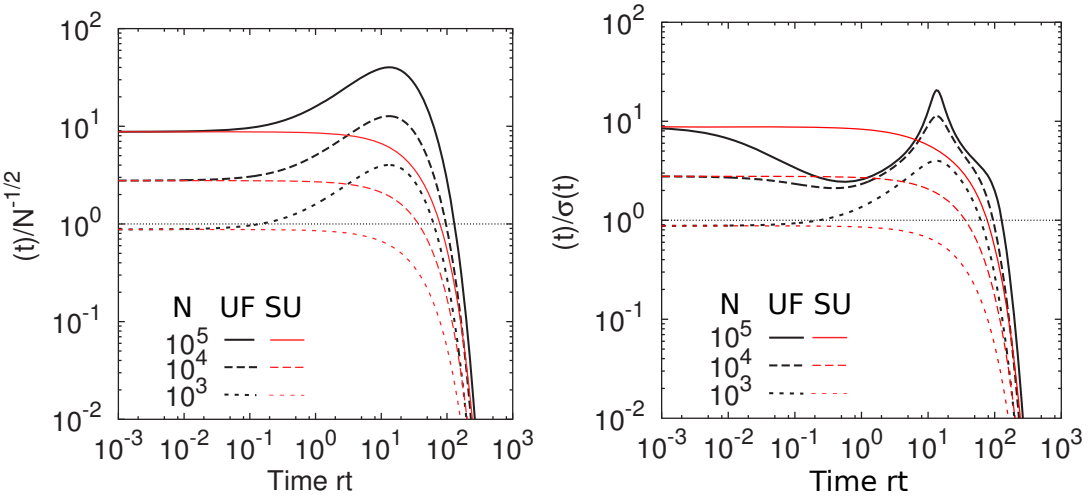
The structure of the cascade model can be seen as two towers of stochastic updater synapses (see Chapter 5) with progressively lower transition probabilities. In this section we construct a filter equivalent of a cascade model by replacing each stochastic updater embedded in each cascade state with either a dual-filter or with the unified filter model we described in the previous chapter. To achieve equivalence, the mean escape time from each cascade state needs to be matched with the escape time of a filter embedded in the same state. We interpret probabilities as rates and match them against mean escape times of filters. As we saw in the previous chapter, for symmetric thresholds, the dual-filter escape times rise linearly with Θ (6.5), while for the unified filter (LP filter) these are $\tau(\Theta) = \Theta^2$. We may associate p_i metaplastic transitions with escape over one threshold, say Θ_+ for a strong synapse, and q_i with plastic transitions associated with Θ_-. For a weak synapse, the role of each threshold is reversed, Θ_- express metaplastic transitions and Θ_+ switch a synapse to the strong strength. However, the default transition probabilities of the cascade dictate that escape times double as synapses progress down the cascade while the threshold dependent escape times described above cannot directly match this progression. To tune escape times more accurately we can utilize the filters with a stochastic decay component and adjust the decay rate to closely match each cascade state. Characteristic, of perfectly matched escape times to the original cascade's default transition probabilities is the uniform distribution of synapses among cascade states.

To find the filter parameters that match the required escape time we first choose a set of thresholds that gives an expression time that is shorter than the required value. We

can then extend this escape time by increasing the decay rate η. Using the calculations presented in Sections 6.2 and 6.3.2 we obtain an expression of escape time for a given set of thresholds as a function of decay rate. We then solve this expression for a required mean escape time to obtain the decay rates. For small escape times the choice of thresholds is restricted, for example for an escape time in one step there is only a single choice for thresholds for both models $\Theta_+ = 1$ and $\Theta_- = -1$. For larger escape times there is some freedom in the set of possible thresholds we could choose from. Matching long escape times, found deeper in the cascade, with a small threshold values would require compensation by increasing the decay rate. As we will see, it is not always possible to match the escape times exactly and thus the parameters of each the filter-cascade state have been chosen so they match the expression time of each respective cascade state as close as possible.

The matching is performed for both the dual-filter and the U. filter with decay. We set the parameters $\Theta_\pm$ and the decay η accordingly for each cascade state i. When escape times are matched, to an external observer the cascade states would appear the same, but in fact the state among the two cascade models operate in a qualitatively different manner. In the original cascade, each state i is a model of a stochastic updater that can randomly express a given induction stimulus with a probability p_i or q_i depending on whether POT or DEP stimulus is seen. Each presentation of a stimulus is then a Bernoulli trial with probability p_i or q_i and the two transition directions, metaplastic (p) or plastic (q) can be seen as independent processes conditioned on the induction stimulus. For example, when a POT stimulus arrives at a strong synapse a metaplastic transition can be described by a Bernoulli trial with probability p. Assuming a sequence of POT stimuli are seen then a metaplastic transition will on average occur after $1/p$ POT stimuli. For the same reasons the mean time to express a DEP stimulus, if looking at the incoming DEP stimuli only, is $1/q$. In the stochastic updater the two different Bernoulli processes for p and q do not interact and by nature are memoryless, each trial is not influenced by previous ones but which process is used on each timestep is conditioned on the induction stimulus, POT arriving with rate f_+ and DEP with f_-. Each stochastic transition of the cascade can be seen as a process which determines transitions by sampling the induction stimuli of $f_+ r$ (POT) or $f_- r$ (DEP) at a random time determined by probabilities p_i or q_i, the transition is completely determined by a single sample taken at some random time step t.

Therefore, we may examine each cascade state as a simple stochastic transition and calculate the mean time before either a plastic or a metaplastic transition occurs.However, as we discussed in Chapter 6, filter synapses operate in a radically different way, according to which plasticity expression depends on the history of induction stimuli. These differences can also be seen in the memory dynamics. Figure 7.1 compares the SNR dynamics of a single U. filter against a stochastic updater of matched expression time $\Theta^2 = 1/p$. We observe that both synapses give the same mean initial signal, but filter



(a) Approximate Noise

(b) Exact Noise

FIGURE 7.1: Comparing analytically obtained P-signal between a $\Theta = 6$ U. filter and a stochastic updater of matched mean expression time with $p = 1/36$. Both models have equal initial signals but the filter signal initially rises after memory storage while SU exhibit signal decay only. **a**. SNR $\mu(t)/\sigma(t)$ using approximate noise $\sigma(t) \approx N^{-1/2}$. **b**. SNR calculated using exact $\sigma(t)$ for U.filter and SU shown in the background

dynamics give a rising signal that prolongs memory lifetime, at least for values of N up to $N = 10^5$ tested here.

Using the default transition probabilities p_i and q_i of the cascade model and assuming balanced excitation we get $q_i = p_i$ (see Section 5.2). Using the methods discussed in Section 6.3.2 we match the escape time from either threshold of a filter $\text{Esc}(\Theta, \eta)$ against the escape of the stochastic updater $1/(p_i + q_i)$ at cascade state i by taking Equation (6.22) fixing τ to the desired conditional escape time and solving for η. An exception to this rule needs to be dealt with at terminal cascade states where there are no p_n transitions. The details of the dual-filter cascade and the U. filter cascade are discussed next.

7.1.1 Dual filter with decay cascade

For the dual-filter the mean time to change state in the cascade model is matched against the mean escape time from this filter. Thus, we look for the mean expression time through either the q-filter or p-filter to match the combined escape rate from a cascade state $1/(f_{\pm}(p_i+q_i))$. We associate the q-filter with plasticity expression equivalent to the cascade's q transitions and the p metaplastic transitions occur when the p-filter reaches threshold.

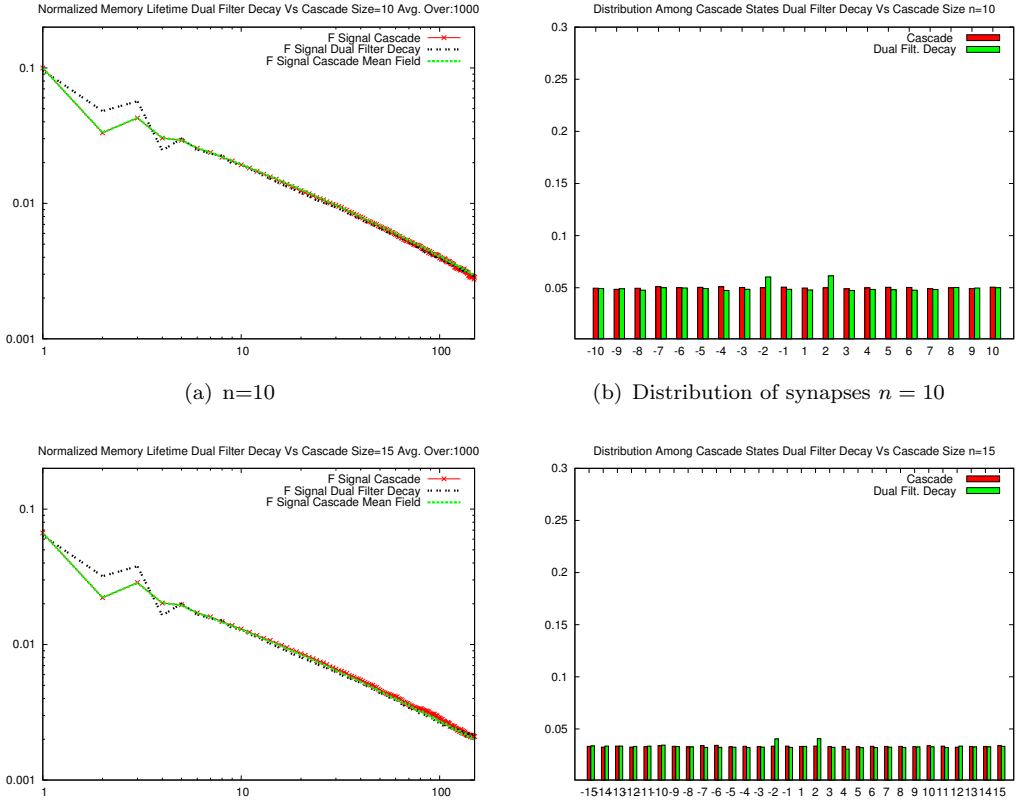
Matching the filter escape times to each cascade state i requires first choosing a set of thresholds such that without decay ($\eta = 0.0$) give an expression time either equal or less

than the desired escape time and then adjusting the decay rate so the escape time is increased to the required value. At the terminal state we simply remove the p-filter and thus ignore meta-potentiating stimuli. There is more liberty in choosing thresholds when matching deeper cascade states because the expression times increase. In general, we have used small threshold sizes, not larger than $\Theta = 4$, because with small Θ solutions on escape times and matching η can be quickly obtained. Thus, we compare against a minimum filter-cascade model, but its performance will be indicative to the benefits of filters as the overall state count will not be much larger than the original cascade model. In principle, for progressive stabilization using filters we could easily image a molecular instantiation with Θ growing from $\Theta = 3$ to $\Theta = 15$ as the filter moves down the cascade. Instead of matching against the original cascade we could apply simple more natural rules on how decay and thresholds change after each transition to increase expression times and model progressive stabilization. Consequently, the values used for Θ and η are artificial just so as to match performance. With the small Θ values we used we obtain minimal integrative ability to each cascade state but yet these still operate differently to SUs, at least for low values of η.

The set of values chosen is shown on Table 1. Note that it is not possible to match the second cascade state with dual-filters. In the cascade model combining the escape rates p_2&q_2 gives an escape time through either transition in $1/f_{\pm}(p_2 + q_2) = 2$ steps. This escape-time cannot be matched by a dual-filter even with a threshold set as low as $\Theta = 2$ and decay $\eta = 0.0$. In this case, the combined mean escape time through either p-filter or q-filter at this index cannot go lower than 2.5 because the minimum number of steps to reach threshold from either q-filter or p-filter is 2 and the maximum is 3. As the paths to either escape time are equally likely, we obtain a mean escape time of 2.5. Changing the thresholds to $\Theta = 1$ values only makes the escape time equal to 1 and thus further away from the desired value of 2.

Using standard discrete-time simulations of binary synapses with the matched dual-filter cascades we plot the mean memory signal and the distribution among cascade states on Figure 7.2. The Figure compares the F-signal, which measures mean signal from synapses that initially encoded the memory only, of both models. A signal obtained via the mean-dynamics method of cascade memory lifetimes is also shown for comparison. The plots show good overall agreement between filter-cascade and original model on the F-signal, with a small difference in the "kink" found in discrete-time simulation as labile state synapse toggle between top cascade states. This discrepancy at $rt = 2$ is due to the mismatching on cascade state $i = 2$ that results in a transient increase in the signal of the dual-filter. The mismatching is also evident when looking at the distribution of synapses among cascade states and it seen as a increased occupancy on the 2nd index (-1,1) in the distribution because escape time is $rt = 2.5$ instead of $rt = 2$ (see Figure 7.2(b)). Nevertheless, the overall lifetime of the memory as measured by the cascade model F-signal very close; the two signals converge to a common trajectory under F-signals but

looking at Figure 7.3 we find there are differences between the F-signal and P-signal. This difference is due to the latent signal of filter synapses that is not captured by the tracked synapse population used by F-signals. In this case, memory performance relates to the P-signal, which is not confined to tracked synapses but the F-signal shows that we have managed to get a close matching of filter-cascades to original cascades under standard signal-decay dynamics. In the next section we proceed to obtain a matched unified filter cascade.



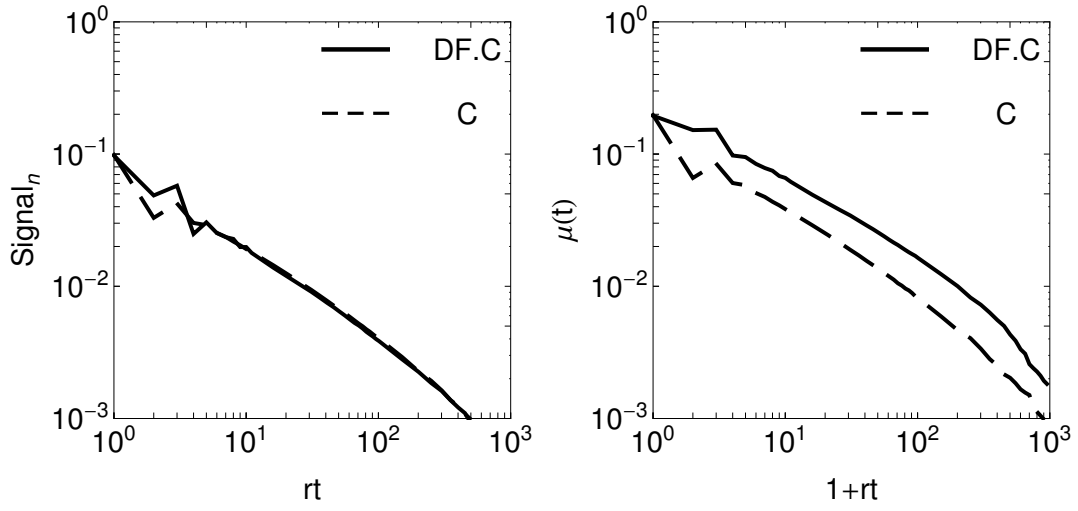
(a) n=10

(b) Distribution of synapses $n = 10$

(c) n=15

(d) Distribution of synapses n=15

FIGURE 7.2: Simulation results comparing memory lifetime and distribution of cascade model and dual filter equivalent after matching escape times for cascades of sizes $n = 10$ and $n = 15$. The memory signals are in discrete-time averaged over $T = 10^3$ trials. The size of the F-signal has been normalized by the number of synapses $N = 10^4$ in these simulations. The two models display a match on the overall lifetime with the dynamics retaining the overall power-law like decay with some minor differences. A higher signal is obtain at step $rt = 2$ and a lower at step $rt = 3$ for the dual-filter cascade. The distribution reveals a key difference between the models at cascade state $i = \pm 2$, where the dual filter has a higher occupancy because the required escape time of 2 can only be approximated by the filter with an escape time of 2.5. This changes the dynamics slightly but as it can be seen the overall course of the signal has been matched. The values used for the dual-filter cascade matching are given in Table 1.



(a) F-Signal normalized (b) P-signal

FIGURE 7.3: Simulation results comparing memory signals of original cascade model (C) and matched dual-filter cascade (DF.C) of sizes $n = 10$ between F-signal on **a** and P-signal on **b**. The P-signal reveals that memory lifetimes are higher in the dual-filter cascade. This is not shown on F-signal which only tracks the memory stored in the initial signal ignoring the latent signal rise in filters. The memory signals are in discrete-time averaged over $T = 10^3$ trials. The F-signal has been normalized by the number of synapses $N = 10^4$.

7.1.2 Unified filter with decay cascade

Similarly to dual-filters, matching mean escape times requires setting appropriate values for the lower Θ_- and the upper Θ_+ thresholds as well as the decay rate η. We take $1/f_\pm p_i$ to be the rate for metaplastic transitions at cascade index i, while $1/f_\pm q_i$ to be the rate of plastic transitions at each cascade index. Each filter-cascade state expression time is matched with the rate of transition from either q_i or p_i for every cascade index i as $1/(p_i^\pm f_\pm + q_i f_\mp)$ against escape from either filter threshold.

The filter has distinct lower Θ_- and Θ_+ thresholds. When these thresholds are reached the synapse will either switch strength state (q transition) for the Θ_- threshold or be meta-potentiated (p transition) by reaching the Θ_+ if it is in the strong state. The q or p transitions are reversed for a weak synapse, crossing Θ_- will cause meta-potentiation and Θ_+ will cause a q transition changing synaptic efficacy to strong.

The lowest cascade state n does not allow p transitions and the filter matching this particular state must have its upper absorbing boundary Θ_+ replaced by a reflecting boundary. Consequently we are required to treat the last cascade index separately in our analysis. Each time a transition is made we assume the filter has to match the next equivalent cascade state. This is achieved by the adjustment of the filter thresholds. When the set of thresholds is asymmetric (i.e $\Theta_+ \neq \Theta_-$) it should be assumed that

each synapse randomly allocates which threshold is for p and q transitions, except in the terminal states which will be treated separately. Plastic transitions (q) will reset the filter thresholds to an original state equivalent to the cascade index $i = 1$.

Matching the filters expression times for each cascade state requires first choosing values for thresholds, keeping in mind that in the absence of decay the escape time grows as Θ^2. The equivalent rates for the U.filter are found using Equation (6.22) to calculate escape times through either thresholds, having first used Equation (6.21) with escape through either Θ as boundary conditions. The equation obtained for the set of thresholds is then solved for η for the required escape times. Evidently, for a given expression time τ, changing the set of thresholds and adjusting η also affects the signal dynamics, see Figure 7.4. As previously discussed, adding decay minimizes the probability gradient among filter states and thus reduces any biasing effect of a transition in the filter states, which is responsible for the signal rise.

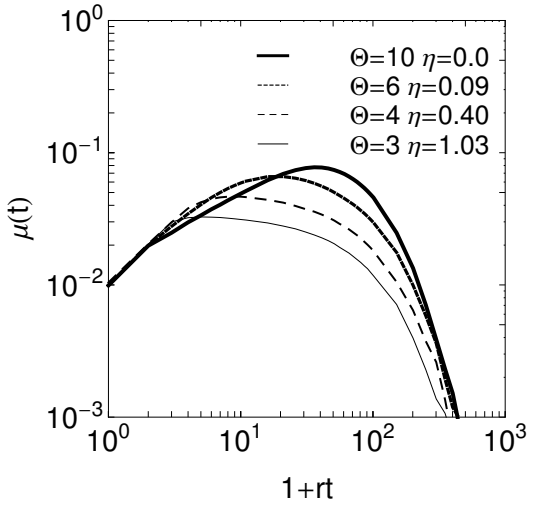


FIGURE 7.4: P-signal of different Θ unified filters that maintain equal expression time $\mathrm{Esc}[n] = 100$ by increasing the decay rates η. The filter with $\eta = 0.0$ and the largest threshold $\Theta_{\pm} = \pm 10$ shows the largest rise in the signal. The peak of the signal drops as the decay rates are increased and the thresholds are reduced. The signals initially follow the same trajectory but begin to decay at different time-points. Larger decay rates show earlier decay onset points in the signal.

On the last cascade index, there is no escape through a p transition. To model this with a filter we have the option of artificially resetting the running sum to zero once the boundary that does not emit is hit. Alternatively, a more natural way would be to construct filters with a single absorbing boundary for q transition as before and a reflecting boundary on the p transitions' end. The internal sum of the filter cannot increase further than the reflecting threshold, once reached the sum is returned to the previous value. The opposite threshold remains absorbing and crossing it represents a q transition. The escape rate for the cascade is $1/f_{\pm}q_i$ where $q_i = q_{i-1}$ at the last

index and the equivalent to the conditional mean escape time through a filter with a reflecting boundary is $\tau_n\Theta_- = 1/fq_n = 1/fq_{n-1}$ for strong synapses. We then need to modify the unified filter transition probabilities to add a reflecting boundary and using the same methods, use the PGF to solve for a given escape time and obtain the required parameters. We have again constrained our set of available thresholds to small filter sizes ($\Theta < 4$) that progressively increase the decay rate for stability, the values chosen for intermediate and terminal cascade states along the respective escape times are listed on Table 2.

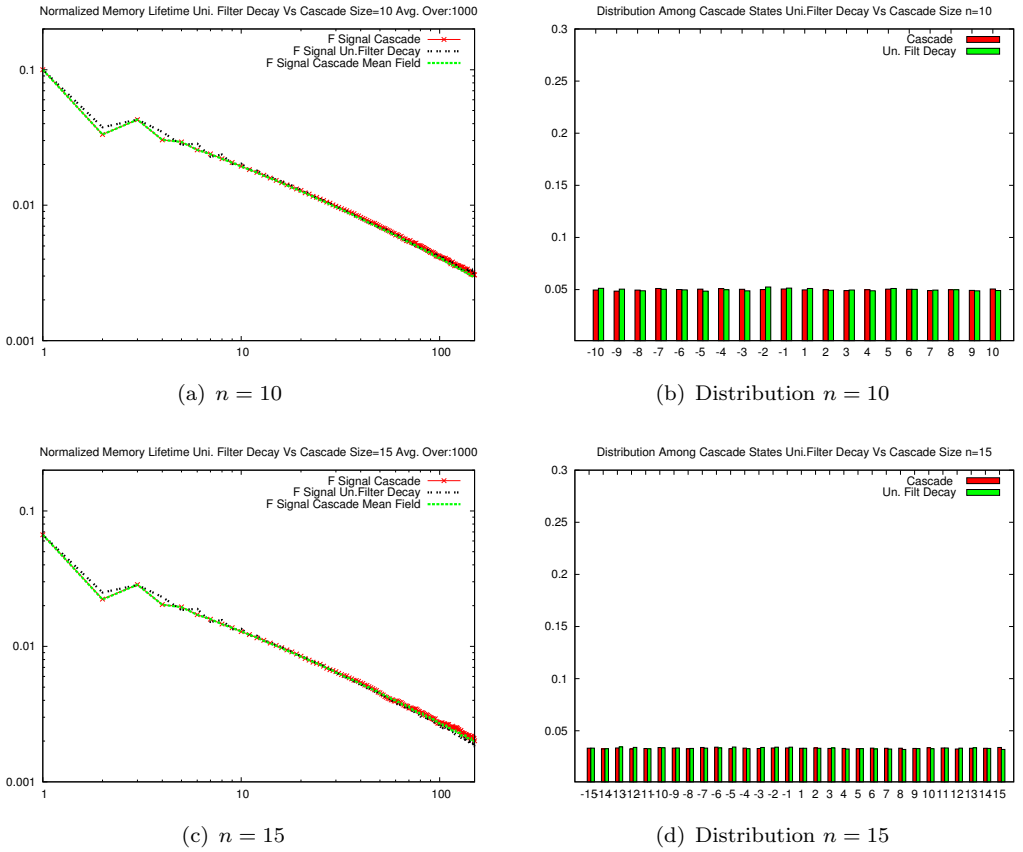
Once transition probabilities and conditional mean escape times have been matched, we expect the two systems to behave equivalently. As before, we establish similarity by comparing F-signals between filter matched cascade and original cascade in discrete time, see Figure 7.5. The figure shows that the cascade state distribution for both models is uniform and the F-signal dynamics match very closely as required giving power-law like decay dynamics to the matched filter-cascades also. These power-law like dynamics do not depend on the specific choices for decay rates and filter sizes used to match the cascade; as long as the expression times between cascade states are matched we expect decay dynamics. In the next section we proceed to obtain memory lifetime results comparing memory lifetimes between original cascades and filter cascades. Initially, we examine these on a fixed network size N across cascades of size n and in later sections a more thorough mapping of capacity over different N is presented.

7.2 Cascade against matched filter-cascades

The F-signal assumes that a memory is stored in the set of synapses that formed the initial signal at the encoding step and therefore it measures the memory signal arising from these synapses only. Here, we use the P-signal $\mu(t)$, which is based on the mean activation of a neuron $h(t)$, to measure memory lifetimes (see Section 4.5.3). In principle, we could split the P-signal into a tracked and a non-tracked pool of synapses and examine the contribution of each set of synapses to the total signal and obtain a scaled version of the F-signal dynamics, but in this section we are interested in actual recall potential and thus focus on the overall $\mu(t)$.

In Chapter 5 we presented a standard method to measure memory lifetimes by obtaining the time when the mean signal $\mu(t)$ falls below a set minimum threshold. We refer to memory lifetime under this method as τ_{max}. Clearly, if the low-threshold point is set equal to the size of the signal noise (σ) then we are actually measuring the time at which signal over noise equals to unity SNR = 1.

Figure 7.6 plots discrete-time simulation results of memory lifetimes for the cascade model and the matched filter-cascades across cascade size n and various low-signal thresholds. Our findings verify that there is an optimal cascade size n^{opt} for the number



(a) $n = 10$

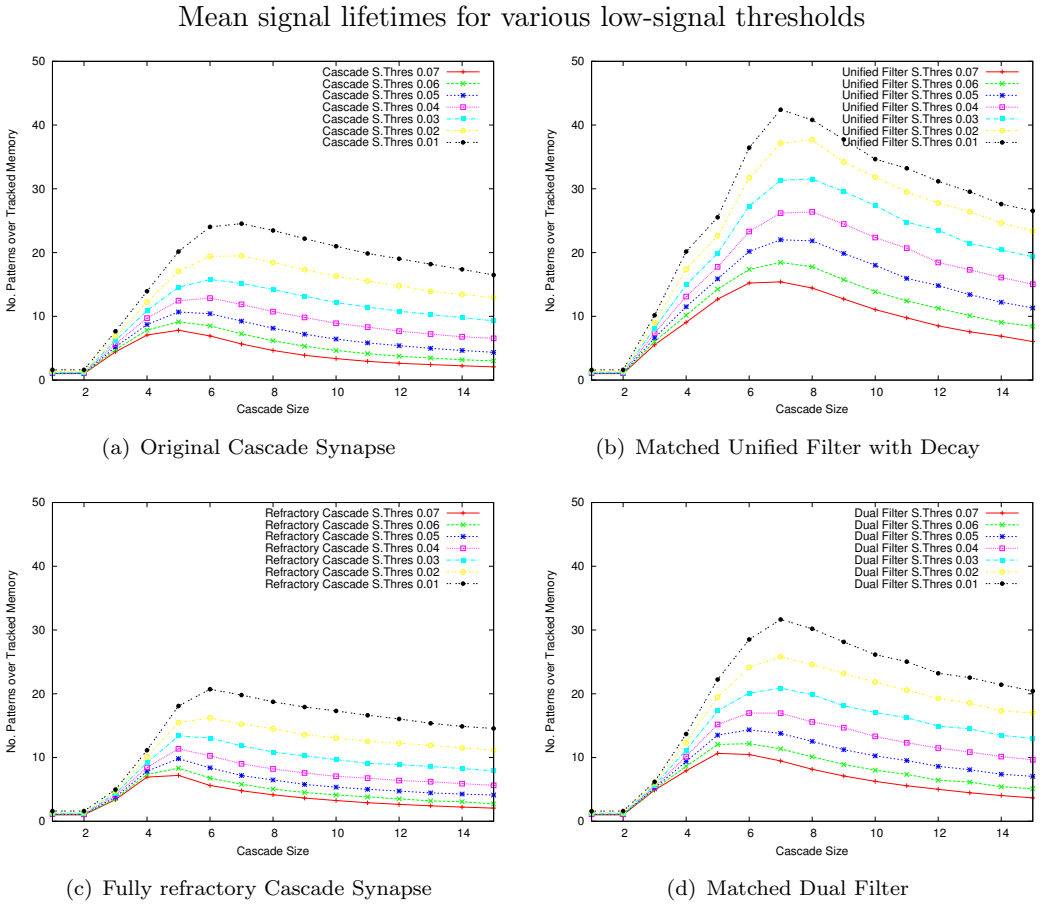
(b) Distribution $n = 10$

(c) $n = 15$

(d) Distribution $n = 15$

FIGURE 7.5: **a,c.** Memory lifetimes plots of a matched U. filter with decay cascade against the original cascade. Plots compare cascade model against a unified filter with decay for $n = 10$ and $n = 15$ cascade sizes. The dynamics of the matched filter cascade retain the power-law like decay of the cascade and the characteristic kink found in discrete-time dynamics. The signal obtain via the mean-field method for cascade model is also shown for comparison giving a close match between all three signals. Memory lifetime plots are averaged over $T = 10^3$ trials. **b,d.** Distribution of synapses among cascade states are shown to be uniform for both models.

of synapses for $N = 10^3$ for each threshold value chosen. Varying the low signal threshold shows that this optimal cascade size peak moves toward small cascade sizes for high threshold values and larger n^{opt} for small threshold values. The low-signal condition is usually taken to be at SNR=1 and so the low-signal threshold is set by the noise as $\approx \sqrt{N}/N$. The issue of optimal number of states has implications on selecting a suitable synapse model for memory encoding beyond the capacity criterion. Leibold and Kempter 2008 showed that although the cascade's capacity is lower than a multi-state synapse (see Section 4.4.3.1), the optimal number of states n^{opt} grows much slower for the cascade against changes in network size N or the connectivity of the network, which can define the coding ratio in the patterns. Therefore, although the cascade capacity may be lower than the capacity of a multi-state synapse (Leibold and Kempter, 2008, Barrett and Van Rossum, 2008), cascade models offer a synapse model less sensitive to changing network parameters under continuous network re-organization.



(a) Original Cascade Synapse

(b) Matched Unified Filter with Decay

(c) Fully refractory Cascade Synapse

(d) Matched Dual Filter

FIGURE 7.6: Memory lifetimes measured using discrete-time simulations to find the point where $\mu(t)$ crosses a lower threshold. The U. filter cascade has the highest capacity across low-signal thresholds and n. There is a "wobble" on U. filter cascade around $n = 5$ which becomes evident as the low signal threshold is lowered. This is not a numerical artefact, it exists for U. filter cascades on the $n = 5$ due to a sudden onset of decay on both $i = 4$ and the terminal state $n = 5$ that are high $\eta = 1.09$. The effect of this would be that suddenly as the filters grow from $n = 4$ to $n = 5$ two states lose their integrative ability and effectively express randomly losing the μ_Ω rising component. Beyond $n = 5$ the decay rates are reduced again by increasing the filter size and this sudden change does not occur again in filter growth (see Table 2). Besides these differences, the overall power-law like signal decay dynamics are preserved in the matched-filter cascades as expression times between cascade states remain in defining a repertoire of decay rates. Simulations run using $N = 10^3$ with low-signal threshold values ranging from 0.07 down to 0.01, with the signal at each point averaged over $T = 10^4$ trials.

Figures 7.6(b) and 7.6(b) show the memory lifetimes of the matched dual and unified filter-cascades examined in the previous section. The results indicate that the filter-cascades exhibit higher capacities to the original cascade at least for $N = 10^3$, while we also obtain an optimal n size. The U. filter cascade has the highest capacity in comparison to the three models examined so far. Besides these differences, the overall power-law like signal decay dynamics are preserved in the matched-filter cascades as expression times between cascade states remain giving a repertoire of decay rates.

We have already examined results showing that single filters offer superior memory lifetimes to SU of matched escape times (see Figure 7.1). Beyond integrating synaptic stimuli however, a filter also exhibits a refractory period. Once a synapse moves to a new filter-cascade state it is initialized to the zero filter-state. From there, a minimum number of induction steps, at least equal to the smallest $\Theta_{\pm}$, is required before any change can possibly be expressed again. The original cascade does not have refractory periods between cascade states. In the next section we examine if refractory periods could be partly responsible for the observed capacity gains under filter-cascades.

7.2.1 The role of refraction in cascade's capacity

We chose two tests to indicate the effect of refraction on the mean memory lifetimes. First, we introduced a refractory period k to the original cascade. The refraction period k introduces a delay during which all induction stimuli DEP or POT are ignored. After the k induction stimuli have occurred then a stochastic transition is allowed to take place. The probabilities q_i and p_i where adjusted so to keep the mean expression times $E[n]$ of each cascade state matched to the original values.

Adding a refractory period makes the $E[n] = k + 1/\rho$ for each of the q or p transitions, where ρ represents the modified probability of p or q. Increasing k decreases the variance of the escape times as a deterministic component is added and the stochastic waiting time for expression is decreased via increasing the transition probabilities p_i or q_i so the overall the escape time remains unaltered. The modified probability ρ_p for a p transition is given by $\rho_p = \frac{p_i}{1-p_i k}$ and $\rho_q = \frac{q_i}{1-q_i k}$ for a q transition. Keeping in mind that the refractory delay is shared between both transition probabilities p and q. A single counter k is decremented regardless of the stimuli for $i < n$ cascade states and the transition probabilities are set to $p_i = q_i = \rho_i$.

The size of refraction can be chosen to make the cascade transitions completely deterministic by setting $k_i = E[n]_i - 1$ for each cascade state. This makes each $q_i = p_i = 1$ and can reveal the effect of refraction to the extreme. The effects of refraction to the capacity of the cascade model can be seen on Figure 7.6(c) by comparing against the original non-refractory cascade model of Figure 7.6(a). Our results showed that refraction in the cascade slightly decreases capacity and this was true for smaller refractory periods as well.

Another approach to testing the effect of refraction was to remove the minimum refraction period from the filters by not "injecting" to filter state zero after a threshold event. We used the probability distribution of a internal filter states to randomly inject synapses within internal states after a threshold event instead of injecting to the zero state. Effectively, the refractory period was removed on average while the equilibrium

filter distribution was retained. Our results showed that re-injecting synapses according to the PDF on the U.filter-cascade also resulted in a minor decrease in capacity.

In conclusion refraction does not seem to be causing the differences in capacity seen between filter implementations of the cascade and the original cascade model. Adding refraction to the cascade or removing refraction from the filter-cascade had a negative impact on memory lifetimes.

The results examined so far have only been indicative of the capacity differences between original and matched cascades as they only tested $N = 10^3$ over mean signal memory lifetimes. Next, we examine a direct method to obtain mean memory lifetimes and then proceed to compare lifetimes between cascades and filter synapses of fixed size.

7.2.2 Using mean first-passage times to measure memory lifetimes

Measuring mean-signal lifetimes is common in a synaptically stored memory framework, with the low-threshold usually taken to be the size of the standard deviation of the mean signal (SNR=1). However, the mean signal does not necessary reflect any particular realization of an encoding episode as the course of a signal from a tracked memory is a stochastic process. The mean-time until the signal drops below a fixed threshold does not necessarily equal the time that the mean signal crosses the lower threshold. Here, we will be introducing a memory lifetime measure taken to be the mean-time until the signal drops below a predefined threshold. This statistic is known as the mean-first passage time (MFPT) in stochastic processes and we will denote these memory lifetimes as $< \tau >_{max}$. Indeed, we argue here that a mean first passage time measures may be more relevant to a synaptically stored memory framework. This is because such measures directly estimate the mean duration of memory before recall fails and thus obtain the expected lifetime of a memory, while measuring the duration of the mean signal may only offer an approximation to the mean memory lifetime as the mean dynamics may not reflect the time-course of any particular realization of the process. Simulations of MFPT, can be rather slow however, since each trial has to wait for a variable indeterminate time before the stochastic signal reaches zero.

We proceed to examine both τ_{max} of the mean signal and $< \tau >_{max}$ of MFPT. Only a single low-threshold will be used $h(t) = 0$ for the MFPT and $\mu(t) = \sigma$ to measure τ_{max}. Figure 7.7 contains simulation results using both capacity measures for the original and the filter-cascade against cascade size n for a range of N. A few outcomes can be drawn from these results. First, the filter-cascade capacity is always higher than the original cascade, a direct comparison between τ_{max} of the two models is shown on Figure 7.8(a) confirming this outcome. Second, the MFPT lifetimes confirm the differences in capacity among models, while the $< \tau >_{max}$ seem to follow the τ_{max} up to some cascade size n, after which point τ_{max} begins to decay while $< \tau >_{max}$ plateaus. This is seen for

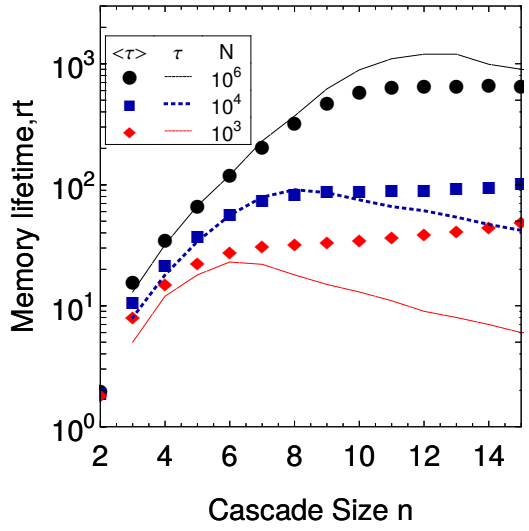
both cascade models, the MFPT increases to a plateau as n increases while the τ_{max} exhibits particular optimal n tuning. Thus, changing to measuring MFPT removes this presumed requirement for optimal n tuning. Figures 7.7(d) and 7.7(c) plot $< \tau >_{max}$ with errorbars showing the standard deviation of the MFTP. We notice that beyond the point where the profiles of τ_{max} and $< \tau >_{max}$ begin to deviate the standard deviation of $< \tau >_{max}$ quickly increases. For the cascade this tuning reflects a compromise between initial signal and memory lifetime. This is reflected as an increased variance in MFPT. A low initial signal increases initial encoding failures contributing to short memory lifetimes, while successful initial encoding under large n, where signal decay is slow, would prolong the memory duration and thus result in a higher contribution to mean memory duration. Therefore, as n increases the distance between the memory lifetimes samples obtained increases giving a higher variance in the mean memory lifetime.

Finally, we confirm the superior memory lifetimes of filter-cascades by comparing $< \tau >_{max}$ across a range of N on Figure 7.9(a) for $n = 10$ with the cascade that utilizes small U. filters with decay. Matching without decay would require U. filters that grow to very large thresholds. Table 3 list the progression of filters for such a matched filter-cascade and Figure 7.8(b) compares the memory lifetimes against a cascade showing that these filter-cascades give even higher capacity gains. In summary, in this section we have shown that memory lifetimes for a cascade model are inferior to a U. filter-cascade equivalent for two matched filter cascades. In the next section, we proceed to compare these models on the basis of the number of states they consist.

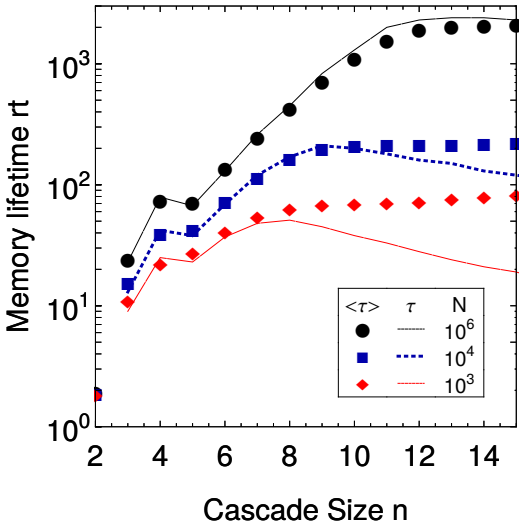
7.3 Filters against cascade over the same state-count

The cascade model's hidden states represent a memory of metaplastic history. Maintenance of these states requires molecular machinery able to represent each of the synaptic states. This adds a computational burden to synapses to support a type of molecular state machine. Similarly, the proposed filters require synapses to maintain discrete filter states in order to process the history of induction stimuli and therefore would also require the relative molecular resources to perform this processing. It therefore seems appropriate to compare memory lifetimes between models based on the number of states they consist as these may reflect some synaptic resource requirement.

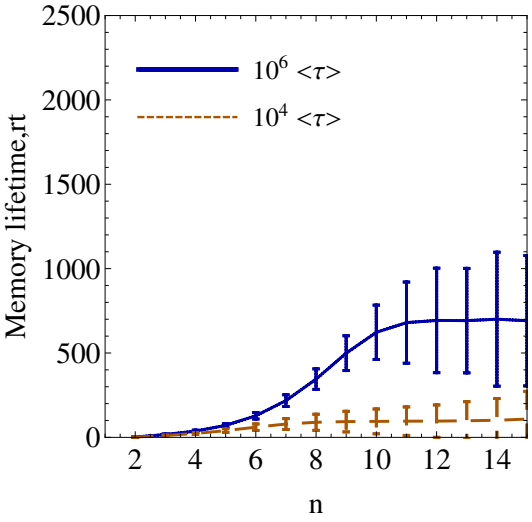
A cascade of size n maintains $2n$ hidden states as there is one cascade of n states attached to each synaptic strength of a binary strength synapse. The strength state of a synapse is embedded into the cascade model as each side of the cascade is associated with a particular strength. In the filter however, induction and expression processes are separated and thus an extra state is implied which maintains the strength state of the synapse. Overall, a symmetric filter of size Θ maintains $2(\Theta - 1)$ internal states (excluding thresholds at Θ as these are not holding states), plus one extra state for the



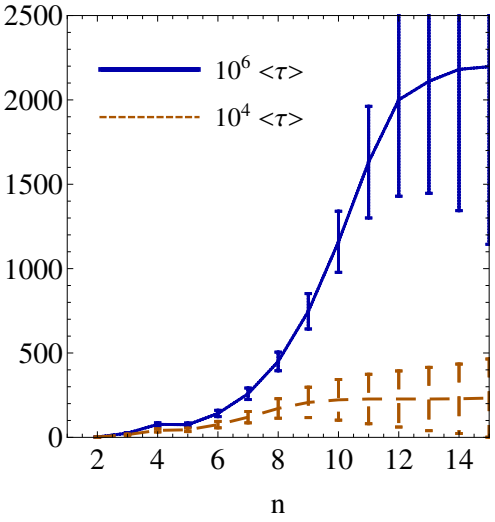
(a) MFPT vs MLT of Original cascade



(b) MFPT vs MLT of matched U.filter cascade

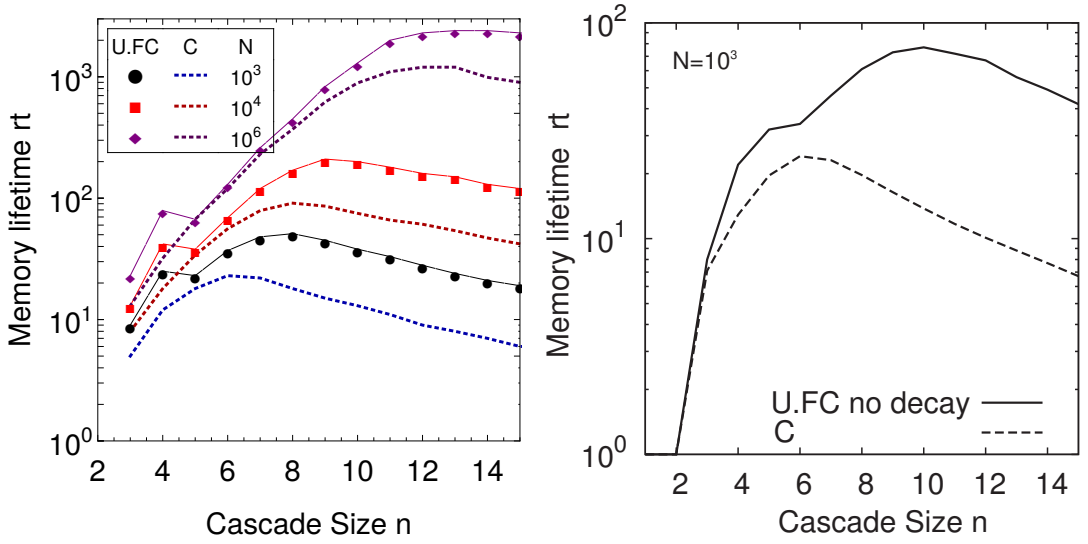


(c) MFPT and noise of original cascade



(d) MFPT and noise of matched U.filter cascade

FIGURE 7.7: **a,b** Comparing simulation results of mean first passage time signal $< \tau >$ against the time that the mean signal takes to drop below the noise τ. Results shown for the original cascade and the matched cascade of unified filters with decay. The filter cascade capacity appears at least twice that of the original cascade, for both capacity measures $< \tau >$ and τ but both capacities have a $\log N$ dependency. The mean signal capacity has a point of optimal cascade size n against network size, requiring tuning to network size. However when measuring $< \tau >$ mean first passage times, this requirement dependency vanishes showing a plateau of memory lifetimes. The "wobble" in the filter cascade is not a numerical artifact but is an effect of our choice of filter growth parameters going from $n = 4 \to 5$. Nevertheless, a matched filter process is always superior to a cascade for any choice of N. **c,d** Plotting the MFPT $< \tau >$ with noise we find that although the MFPT does not show an optimal cascade size beyond which the memory lifetimes drop, the variance increases beyond the point which capacity plateaus.



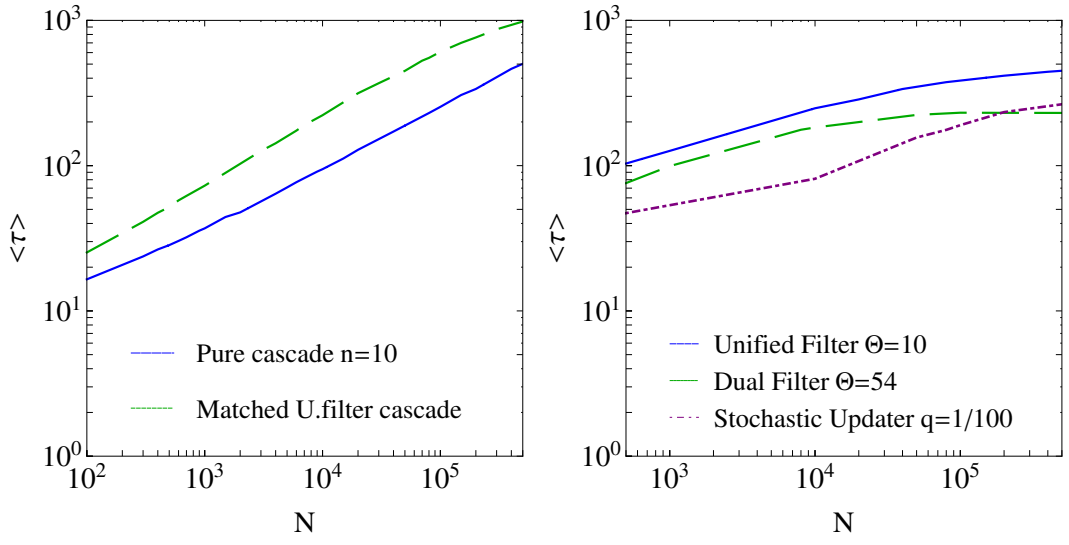
(a) Matched filter-cascade with decay

(b) Full filter-cascade no decay

FIGURE 7.8: Comparing mean memory lifetimes τ between original cascade and matched U. filter cascade (assuming small filters by using a set of small theta values). **a**. The results shown combine results from Figures 7.7(a) and 7.7(b). The U. filter capacity is consistently higher across the different $N \in \{10^3, 10^4, 10^6\}$ examined. **b**. In the absence of decay this filter-cascade does will not precisely match on escape times on all states but the differences are minor. The issue with non-decay is that we need rather large filters to match cascade escape times. The choice of thresholds is shown on Table 3. The "wobble" in the filter cascade is not a numerical artifact but is an effect of our choice of filter growth parameters. For example in **a** going from $n = 4 \rightarrow 5$ we have a drop in performance in order to achieve matching. At those steps performance drops close to the original cascade because the filter-cascade states $i = 4, 5$ with a high $\eta \approx 1$ effectively express stochastically just like a cascade state. Nevertheless, our matched filter process is always superior to a cascade for any choice of N.

zero filter state and an additional state to indicate the strength of the synapse, giving $2(\Theta - 1) + 2 = 2\Theta$ states in total. A cascade of size n is then equivalent to a filter of size Θ as both maintain $2n = 2\Theta$ states overall.

Figure 7.10 shows τ_{max} and $< \tau >_{max}$ signals of a U. filter synapse along the equivalent figure from a cascade synapse repeated here to aid comparison. Note, that since U. filter synapses exhibit a rising signal, the initial signal may be below the noise but subsequently it may rise to allow recall before it drops below the noise again. For memory lifetimes reported here we record the point where the signal drops below the noise ignoring any initial low-signal condition. We find that even single U. filters, removed from the complex cascade structure, can exhibit higher memory lifetimes over a large range of the parameter space. Interestingly, single filter synapses do not exhibit an optimal Θ constraint per N. Overall, both τ_{max} and $< \tau >_{max}$ follow the same trajectory so a mean-signal approximation to memory lifetimes would be accurate. An exception to this rule occurs when mean filter signals never rise above the noise and thus $\tau_{max} = 0$,



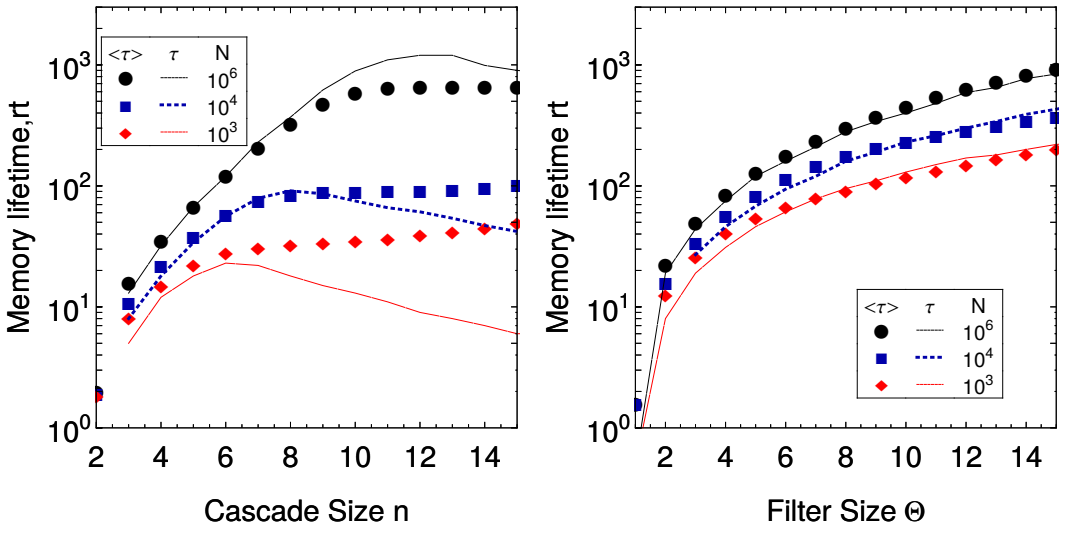
(a) cascade models

(b) Non-cascade synapses

FIGURE 7.9: Simulation results comparing mean first passage time signal $< \tau >$ against network size N for original cascade (C) and matched unified filter cascade (UF.C). **a**. The capacities shown are for cascade size $n = 10$. **b**. Capacity between models of matched on a fixed mean expression time to $\text{Esc}[n] = 100$, with U. filters exhibiting the highest capacity across N. Notice how the rate of SU capacity increase goes up as it exceeds $N = 10^4$, where the noise goes below the mean initial signal of a synapse with $q = 1/100$.

seen here at $\Theta = 8$ for $N = 10^2$. Figure 7.11 compares lifetimes of the two models over a wide region of parameter space, while Figure 7.11(b) indicates the areas in the parameter space where filter lifetimes exceed cascades. Filter synapses dominate recall over the parameter space except in two small regions. For small N in which filter synapses' SNRs do not reach unity and a second region with high $N > 10^5$ and large n, which is biologically irrelevant. Thus, single U. filter synapses can outperform cascades of progressive stabilization over a biologically relevant range of synaptic counts.

Comparing a U. filter synapse against a dual-filter and a stochastic updater synapse with matched mean expression times we find that U.filter synapses provide the highest capacity, (see Figure 7.9). Also, for the same mean expression time, a dual-filter requires $\Theta = 54$ while U. filters need $\Theta = 10$ and thus according to memory lifetimes they utilize these smaller number of states more efficiently and therefore dual-filters are inferior to U. filters. For stable synapses requiring long mean expression times, a U. filter process could be based on minimal molecular substrates that implement a few states to give long expression times. It is easy to imagine that Θ up to 15 and higher would not be excessive for a molecular instantiation of a filter. In contrast, a cascade with $n = 15$ would require states with probabilities as low as $p \leq 2^{-14} \approx 6{\times}10^{-5}$ and it is questionable how synapses could biophysically tune to obtain such small probabilities.



(a) Original cascade

(b) Single U.filter

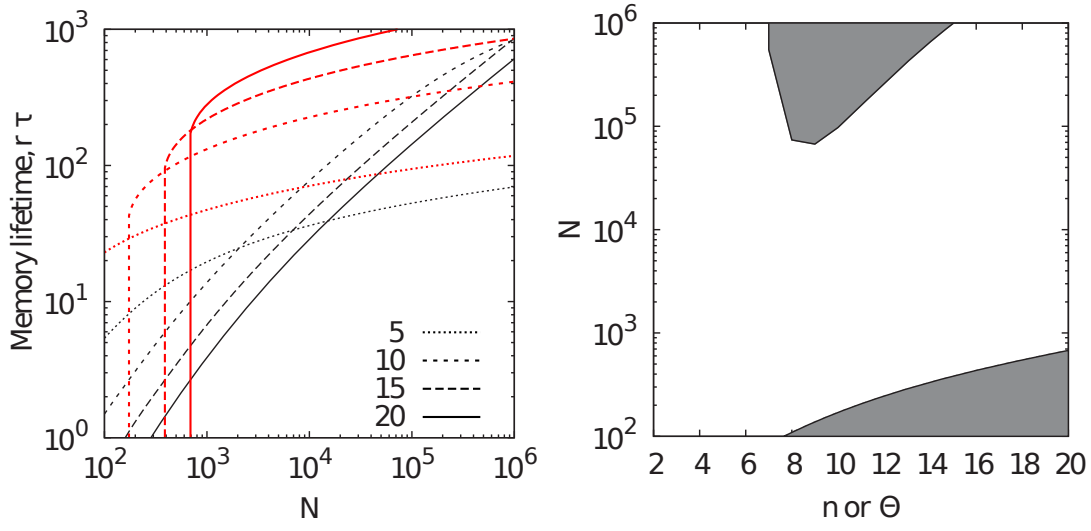
FIGURE 7.10: Simulation results for cascade and single U. filter over mean memory lifetime τ(thin lines) and mean first passage time $< \tau >$ (points). Because the U. filter signal has a rising phase we take the memory lifetime to be the point beyond the peak signal where it drops below threshold, a method which ignores if the initial signal is below threshold. Overall for the U. filter the MFPT follows the trajectory of τ as the filter grows. An exception occurs for low $N = 10^2$ where beyond $\Theta = 7$ the signal never crosses above noise (peak signal $\beta/\Theta \approx 0.95$ when $\sigma = 0.1$). The cascade figure is repeated here so two figures can be seen side-by-side. Over most of parameter space a single U. filter exhibits higher capacity to a cascade of an equivalent number of states $n = \Theta$.

In this section we have shown that U. filters can give higher memory lifetimes compared to cascade models over the same number of states $n = \Theta$. We have measured capacities using both a MFPT and a mean signal lifetimes. In the next section we test capacity within a Hopfield (Hopfield, 1982) network to establish if our results actually transfer over to memory recall under recurrent network dynamics.

7.4 Memory capacity in a Hopfield network

Here, we compare the models by a direct measure of the actual memory capacity of Hopfield associative memory model (Hopfield, 1982) we reviewed in Chapter 4. The motivation is to evaluate the relevance of our previous single neuron results to a memory network that involves network dynamics in recall.

We present results for single filter synapses and matched filter-cascades from simulations using a Hopfield network of $N = 10^3$ neurons with binary synapses. In this setting, the stored and tracked patterns are random vectors having values of $\xi_i = \pm 1$ and the binary neurons produce outputs $h_i = \pm 1$. Each simulation trial involves four steps.



(a) (b)

FIGURE 7.11: **a**. Comparing single U. filter against cascade on the number of states using mean signal memory lifetimes simulations to obtain the point where SNR=1. **b**. Comparing regions of model lifetimes in the n–N plane we find two small regions of parameter space (grey areas) in which cascade memory lifetimes exceed the single U. filter memory lifetimes. The upper region exists only for $N > 10^4$ and is thus biologically irrelevant. The lower region arises because for small N, filter synapses' SNRs do not reach unity. Encoding failure also occurs with cascade synapses, but at smaller values of N.

First, synapses are randomly initialized and random patterns are loaded from a file and shuffled, these are then sequentially given as input to the particular synapse model tested and at the end of a training sequence a weight matrix is formed by reading the strength state of each synapse. Once the network has been trained, plasticity is locked and no further changes can occur to the weight matrix. The next two steps involve testing capacity by attempting to recall a particular tracked pattern that has been learned. This pattern is no different to the other random patterns learned. To test recall the tracked pattern is imposed as a probe vector to the output neurons and the network is allowed to update its state for a few cycles. To emulate asynchronous network updates the update algorithm randomly selects an output neuron and calculates its output according to the state of all other output neurons at that time. Finally, the probe vector is removed and the network is allowed to settle to an attractor state, once settled we then compare the output of the neuron with the tracked pattern.

To assess memory recall the Hamming distance of network output is compared to the tracked memory. For all comparisons against patterns a level of acceptable error has been set at $e = 5\%$. Thus, this error margin is used to assess when the network output settles to the stable state but it is also used as the accepted error in comparisons between the tracked memory pattern and the network stable output state to asses recall. We estimate recall probability of n patterns by running $T = 10^2$ trials of a simulation storing

n patterns and testing recall on each trial. Recall probability of the first pattern stored is estimated by counting the ratio of successful recall trials over failed ones. The pattern is considered recoverable if the stable output pattern and the tracked pattern differ within the accepted error e. When recall probability is $p < 0.5$ (here estimated using $n = 10^2$ samples) for L consecutive patterns the algorithm stops and records the number of patterns stored $n - L$ before recall began to fail, otherwise it proceeds to measure recall probability under $n + 1$ patterns.

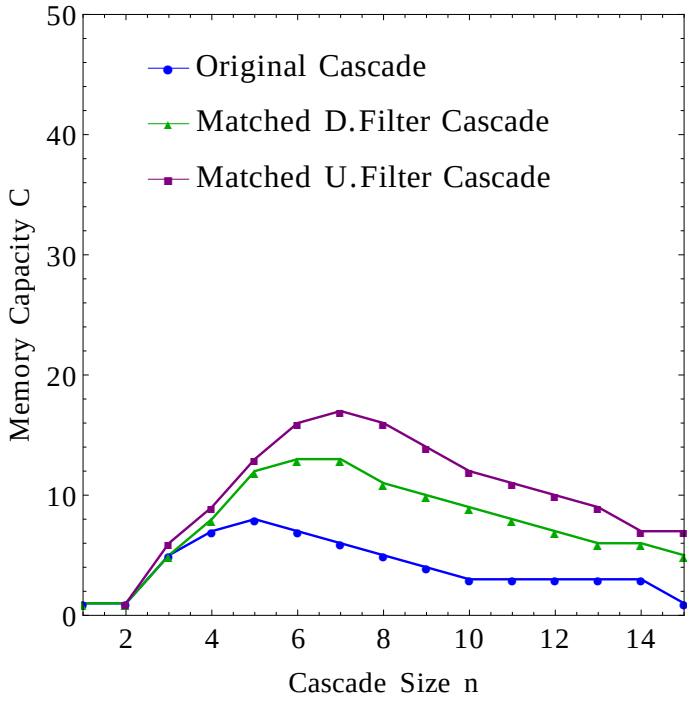
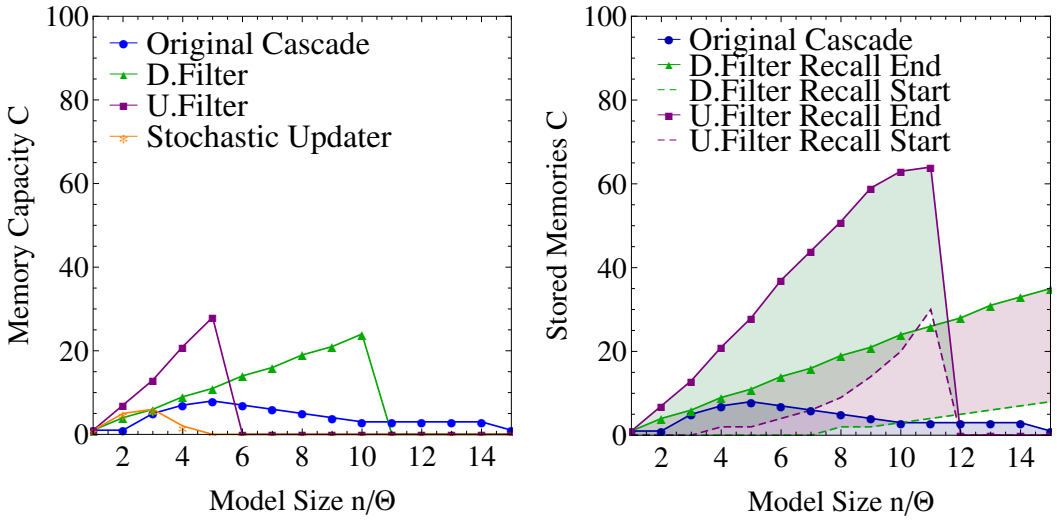


FIGURE 7.12: Comparing the capacity of the cascade model against filter-cascades matched on escape time, as close as each model allows, in Hopfield network with $N = 10^3$. Capacity counted as the number of memories stored in a Hopfield before recall probability drops to $1/2$ across $T = 100$ trials. The allowed error in pattern recall is 5% error.

The capacities of the cascades against matched filter-cascades are shown on Figure 7.12, confirming that U. filter cascades give higher capacities to cascades across cascade size n. The dual-filter cascade also shows a capacity improvement compared to the original cascade, its capacity sits between cascade model and U. filter cascade. Therefore, filter-cascades give significant capacity increases that exceed a doubling of capacity for the network size tested here $N = 10^3$. Next, on Figure 7.13 we compare the capacity of a cascade model against single fixed size filters of an equivalent number of states. As previously discussed, single filters have rising-signal dynamics and thus a memory may not be initially recalled after it has been stored. If we impose strict constraints on initial recall ability when measuring capacity then we obtain filter capacities that collapse beyond $n = 5$ for the U. filter and $n = 10$ for the dual-filter, see Figure 7.13(a). The cascade however is able recall across n although it exhibits a strong preference for

optimal $n = 5$. Nevertheless, if we take into account the rising signal we may extend the range over which filters operate.

To account for the rising signal we modify the stop condition of $L = 3$ failed recall patterns to $L = 0.375 * \Theta^2$, which gives a sufficient waiting time to allow for enough patterns to be stored until the peak filter signal. Figure 7.13(b) verifies that indeed later recall is possible and we find remarkable recall dynamics that give a region of recall (shaded area). With the rising signal, the capacity of the Hopfield network is markedly increased further and the range over which filters operate extended from $n = 5$ to $n = 11$. Here, the dual-filter exhibits an advantage over the U. filters by operating across all n tested with a consistently growing region of recall. Considering the shaded areas of recall, U. filters provide interesting new recall dynamics along with marked capacity increases even if we are to solely consider the shaded areas, where memory recall is possible.



(a) Initial signal constraint

(b) Waiting for signal peak

FIGURE 7.13: **a**. Capacity of Hopfield network with 10^3 neurons using cascade compared to filter synapses of equal number of states. Capacity counted as the number of memories stored in a Hopfield network allowing recall of the first pattern. Recall is considered possible if the first pattern stored can be recalled with less than 5% error in at least 50% of the memory test trials. Cascade and S.U give similar maximum capacities while filters show superior performance, which however collapses beyond some Θ because the initial signal becomes too low. **b** Same as before but here ignoring a low initial signal. The shaded areas represent the region in the sequence of stored patterns during which the tracked pattern can be recalled. U. filters shows markable capacity improvements to cascades. We find that a cascade structure is required in order to secure an initial signal for immediate recall across n.)

7.5 Summary

This chapter aimed to compare filters to the cascade model on the basis of memory lifetimes. The cascade model introduces hidden states to allow correlations in the pattern of potentiation and depression events at a single synapse to affect the degree of plasticity (Fusi et al., 2005). However, the model uses stochastic updating synapses within each cascade state that are agnostic of the induction history. In contrast filter synapses respond to the history of plasticity induction. By embedding filters to a cascade we may produce a filter based cascade model that allows a direct comparison of stochastic updating and filter approaches. We used minimal filters with very small threshold values to match our filter-cascades to the original cascade and showed that even these basic filter-cascades can give significant improvements in memory lifetimes. Filter-cascades may represent the continuous adjustment of thresholds and decay rates to ongoing plasticity so synapses become progressively more stable when detecting correlations in the induction stimuli. We then removed the cascade structure from the filters and attempted to compare a fixed filter against a stochastic cascade by comparing between models of an equal number of state. We take the number of states a model maintains to reflect a resource requirement in a biomolecular instantiation and thus memory lifetime comparisons between models of the same state-count would indicate the relative efficiency between models.

We find that our LP filter synapse (U. filter) give significant capacity increases compared to cascades within biologically relevant synaptic counts $N < 10^5$ while utilizing the same number of states $n = \Theta$ as cascade. These capacity gains are shown when measuring mean signal memory lifetimes, or measuring the mean memory lifetimes (MFPT) we introduced in this chapter. We also report that testing capacities in a Hopfield network of binary neurons also showed that LP filters give markable capacity gains and new recall dynamics, with a region of recall as new patterns are being learned. Specifically for MFPT, we find that a previously reported optimal tuning of cascade size n to network size is removed from mean memory lifetimes but it gets manifested in the variance of the mean lifetimes.

One of the motivations behind the cascade model's invention is that multiple timescales would allow synapse to obtain high initial signal from fast synapses and long memory lifetimes from slower synapses in the cascade. With filters, a low initial signal of a large Θ for the network size N may still allow later recall when the signal rises above the noise. Thus, the filters' richer dynamics result in improving the memory capacity of simple memory models but also weaken requirements for optimal tuning of the state count in the model against the network size. We find that filters modify the meaning of memory lifetime since there can be an initial period during which recall is not possible, yet the ultimate demise of a memory signal will occur after the rise and fall of the memory signal at a later time. How are we to measure memory performance in this

case is debatable, here we have resided to measuring the time point when the signal falls below the noise.

With single filters we have simpler synapses that give higher memory capacities to cascades but there is no stabilization of memory traces. Filter-cascades may offer such a revision, by progressively growing filters to stabilize a trace, but as we discussed in Chapter 2, the transition of synapses to stable forms of plasticity requires particular stimulus patterns. These patterns are believed to be indicative of persistent features of the environment that memory systems should learn and retain for long-term. In the next chapter we examine if LP filter synapses can detect the pattern of stimulation so they can selectively stabilize synaptic traces.

Chapter 8

Filter mechanisms for memory allocation

In the previous chapter we compared the original cascade model against filter synapses. First we paralleled the original cascade model against a filter cascade equivalent by replacing each stochastic updater at each state by filters of matched mean escape times. Secondly, the two models were matched on the basis of resources by examining mean memory lifetimes of an n state cascade against a $\Theta = n$ filter. In both cases filter dynamics outperformed the original cascade, which is based on stochastic updater dynamics, when compared within biologically relevant numbers of synapses per neuron (i.e $N < 10^5$). Assuming that each synaptic state represents some biological resource, the integrative dynamics of a synaptic filter appear as a more efficient mechanism of using synaptic resources to provide an ongoing memory storage mechanism than a cascade structure of progressive stabilization.

Although multiple stable states of plasticity are not necessary for long-memory lifetimes, we require a memory system to allocate some memories of particular significance. Memory allocation within an SPM framework will require some synaptic resources to be removed from the pool of available synapses for new memory storage. Hence, we assume that filter synapses have two degrees of stability, a stable or "locked" state that represents a mechanism for LTM and an early state which has a volatile strength state subject to ongoing memory storage. We assume that the transition to the stable state is triggered in response to a strong encoding protocol involving regularly spaced repetitive memory encoding that are taken to signify salient information suited for long-term memory storage. This follows from earlier discussions according to which massed repetition protocols have been shown to be inefficient at inducing LTM and l-LTP.

If the computational unit of the memory system is the single synapse then single synapses should be able to detect such a regularly spaced strong protocol. The focus of this chapter is on the question whether individual filter synapses are able to detect a strong induction

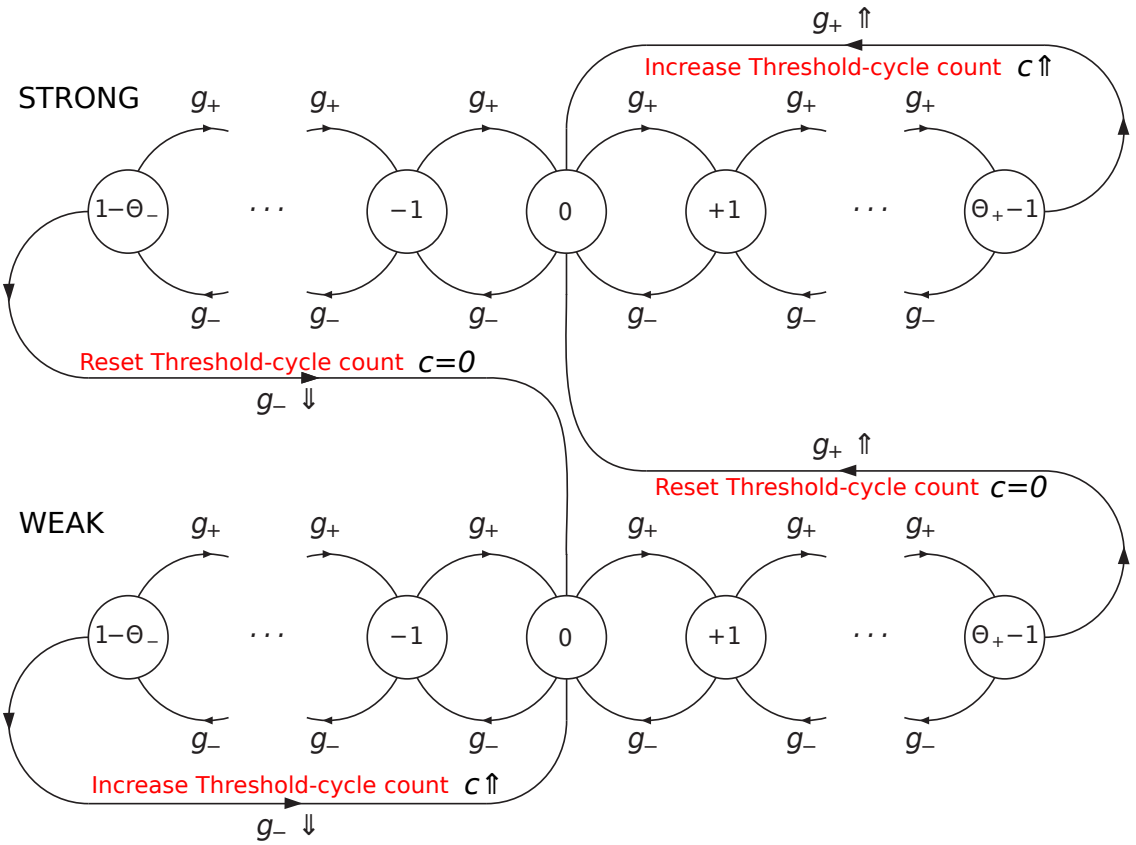


FIGURE 8.1: Markov diagram showing how changes in threshold cycle (TC) size are counted in a unified (LP) filter with two symmetric thresholds Θ_+ and Θ_-. Shown is one filter per synaptic strength (STRONG/WEAK), any threshold transition that moves from one to the other resets the threshold-crossing count c, while transitions that keep the synapse within the same strength increase c by one. The threshold transitions could be taken as LTP/LTD (or depotentiating) events or for the binary synapse case shown here there may be no associated strength change with a threshold-crossing event, these could be simply metaplastic transitions.

protocol in order to trigger memory allocation. To this end, we begin by exploring the influence that different induction protocols have on the mean memory signal and on the distribution of synapses among filter states. We define a measure that counts the number of *consecutive same threshold crossings*, which we call *threshold-cycles* (TC) and examine the effect that different encoding protocols have on the distribution of TCs. The size of a TC is determined by the number of metaplastic (same-threshold) transitions c made without any intervening opposite threshold-transition, see Figure 8.1. The threshold transitions are not necessarily associated with a change in strength and thus we avoid using the term LTP/LTD (or depotentiating) events for clarity, because for the binary synapse case we are focusing on these are purely metaplastic transitions without an associated strength change.

Specifically, we are investigating if a spaced repetitive encoding protocol gives some distinct signature in TC size distribution that could be used by synapses to detect when such a protocol is being induced. Our results show that spaced and massed protocols

result in different TC distributions that could, in principle, be utilised by synapses to characterise the protocol being induced. We examine the convergence of a distribution of TCs sampled in time towards the empirical distributions of either a massed (TC_{massed}) or spaced protocol (TC_{spaced}). We find that single synapses need to obtain a large number of repeated induction protocol applications in order to obtain sufficient TC distribution statistics. The number of protocol applications required is unable to directly interpret experimental results showing that a single trial of a spaced repetition protocol is sufficient to induce LTM. In our model the number of trials required increases with the filter size Θ. We discuss issues around the interpretation of the experimental results and how these couple onto synaptic filters to argue that strong stimulation protocols may effectively reduce the minimum TC samples required by single synapses but yet we conclude that single filters are unable to obtain sufficient statistics to reliably detect the protocol being induced within a few trials.

8.1 Introduction

As we discussed in Chapter 5, extending memory lifetimes in an abstract model of synaptically stored memory was the primary motivation in the inception of the cascade. In addition to extending memory lifetimes it has been argued that a cascade model also captures the progressive stabilization of synaptic plasticity which relies on complex neural biochemical pathways (Kandel, 2009). According to this view the various stability degrees of plasticity can be mapped to the states of a cascade model. Each state reflects a distinct signalling pathway that maintains this plasticity form (Fusi et al., 2005, Kandel, 2009). Therefore, the cascade model's terminal state reflects a long-term form of plasticity and the states before it corresponds to lesser degrees of synaptic stability. However, two forms of LTP are believed to contribute to long-term memory formation, the early phase of LTP (e-LTP) and the late phase of LTP (l-LTP) (see Reymann and Frey, 2007, for a review). l-LTP is thought to be the basis of long-term memory and requires structural changes and synaptogenesis (Malinow et al., 1989, Frey et al., 1988, 1993, Yan-You Huang et al., 1996, Kang and Schuman, 1996). Therefore, of all the cascade states it appears that only a couple of its terminal states are of interest for long-term memory. With this in mind, and the fact that single synaptic filters can outperform a cascade structure we do away with the multiple degrees plasticity framework and begin to consider a memory system of two stability states. In this system a memory is either subject to interference by the ongoing memory storage or, given a strong encoding protocol, it becomes resilient to further memory storage events by locking the state of the synapses that encode it. The state locking is a simplified model of long-term plasticity and long-term memory within a single memory system. It is only used here to represent some stable form of plasticity which lasts much longer than the earlier form. The duration of long-term synaptic plasticity can relate to timescales of memory

phenomena and as we saw in Chapter 2 its induction requires *strong* stimulation protocols. These consist of regularly spaced repetitive hippocampal pathway stimulation to induce LTP or repetitive stimulation of single invertebrate synapses to induce LTF. Behavioural training protocols have also shown that properly spaced recurring events can be preferentially encoded by a memory system showing parallels to synaptic plasticity (see Kornmeier and Sosic-Vasic, 2012, Litman and Davachi, 2008). This preference in spaced patterns of stimulation probably allows a memory system to preferentially consolidate persistent salient features in an animal's environment.

We are therefore compelled to seek if patterns of repetition can be detected in the filter memory framework in order to make particular memories further resistant to degradation through some allocation mechanism that will lock synaptic state. The locked state represents the stable form of plasticity such as l-LTP. In our simplified model of memory allocation, once synapses reach this state they are no-longer subject to changes under ongoing memory storage. In single synapse based memory allocation a key point is to examine if individual filters of synapses in the early plasticity form can detect when a regularly spaced repetitive protocol is being induced to make a transition to the stable "locked" form.

Initially, we take a naive view on the effect that l-LTP and behavioural memory inducing protocols have on synaptic filters and treat each memory encoding event equally by assuming they stimulate a single synapse with a single POT or DEP induction stimulus. Nevertheless, it is very likely that an l-LTP protocol, which consists of repeated 100Hz one second long pulses, actually delivers more induction steps at every stimulus repetition. Moreover, behavioural protocols inducing LTM also use strong stimulation methods. These usually involve an electrical shock that may also result in intense neural activity and thus to stronger memory encoding than normal. In that case, these behavioural protocols no-longer represent the repetition of a single memory encoding event among equal others but rather the repetition of very special encoding events which can, in principle, deliver more POT/DEP induction steps per repetition than other events. These considerations can lead to scepticism when interpreting the experimental data as evidence towards memory systems being able to detect salient features within a few memory repetitions and to trigger LTM in response. If the encoding protocols are not physiological but actually much stronger in respect to induction steps, then a very specific salience signal is given to a synapse at every repetition which could accelerate the processes that lead to LTM or l-LTP. Thus, the prominent requirement for at least four memory repetitions for LTM may not be a realistic lower bound and in fact it has been scaled down by the protocols induced. An animal's long-term memory most probably does not respond to a fixed number of repetitions but it's rather a combination of factors that unlock LTM, the strength and the pattern of stimulation are probably two of them.

The effects of memory repetition on filter synapses can be seen by either examining the mean memory signal or the statistics of internal filter state transitions. However, these

two quantities are related. Memory repetition will enhance the memory signal of the repeated memory simply by recruiting more synapses to the required synaptic strength state. This enhancement relates to an increase in the number of threshold crossings towards the threshold dictated by the memory being repeated and therefore the mean number of threshold crossings for a particular direction will be increased subject to memory repetition.

The filter mean memory signal consists of a μ_Δ and a μ_Ω component and therefore signal enhancement under memory repetition could result from enhancing both components. An increase in the bias of a filter to express through the same threshold would increase the size of the TCs and μ_Ω, but also directly driving synapses through a threshold under memory repetition would increase μ_Δ and TCs.

The pattern of repetitive encoding may have different effects on μ_Δ and μ_Ω and therefore modify the statistics of TCs. In the next sections we examine how the TC distribution under protocols of regularly spaced repetition intervals differs from a TC distribution measured under massed protocols involving a train of consecutive stimuli.

8.2 The effects of memory repetition

In this section we explore the effects of memory repetition over an ensemble of filter synapses at two levels, first at the level of the mean filter signal which reflects the strength state of all synapses and secondly at the level of internal filter states and threshold transitions. Initially we aim to explore how different stimulation protocols reflect on these two on average. The protocols examined differ in the number of repetitions used n_r and in the inter-repetition time-interval T_r.

Figure 8.2 shows the effect of typical repetition protocols on the mean signal of a filter size $\Theta = 7$ under a simulation of $N = 10^5$ synapses. A massed protocol involving eight memory repetitions $n_r = 8$ delivered with an interval of $T_r = 1$ results in maximum mean signal augmentation when compared against all other protocols. A spaced protocol is also shown having a repetition interval equal to the mean time until filter's peak signal $T_r = 0.375\Theta^2$, in this example $T_r \approx 18$, and we will refer to this protocol as $T_r^\star$. The regularly spaced protocol $T_r^\star$ offers limited gain in maximum signal as it asymptotes after approximately four repetitions ($n_r > 4$). Under this regularly spaced repetition protocol we also observe that the time of signal peak moves towards smaller intervals making regular intervals $T_r = 18$ miss the signal peak. Qualitatively similar results for $T_r^\star$ and $T_r = 1$ are shown for a smaller filter of $\Theta = 4$ on Figure 8.3.

A protocol that aims to repeat a memory at the apparently regressing peaks of the mean signal would involve a mixture of spaced and massed intervals. Examining how the peak moves after each repetition we find that such a protocol would require repetition

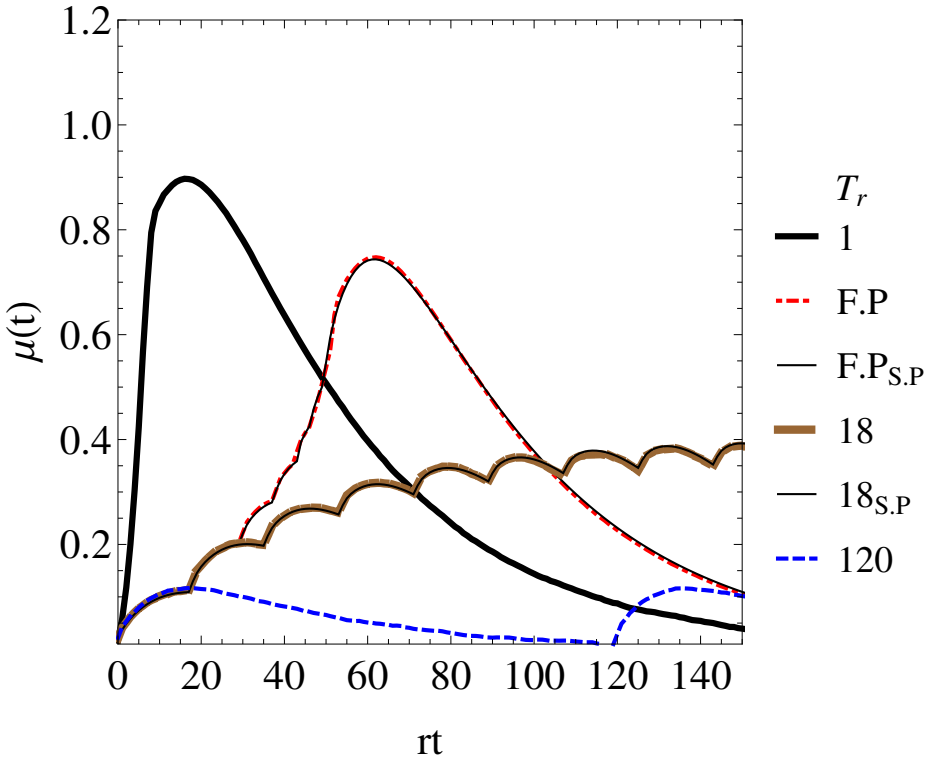


FIGURE 8.2: The effect of repetition interval on the mean signal of a $\Theta = 7$ synaptic filter under eight repetitions $n_r = 8$. Massed repetition $T_r = 1$ shows the highest peak signal overall. Fixing the repetition interval to the time until filter peak $T_r = 18$ shows that regular spaced repetition makes the maximum signal asymptote after approximately $n_r > 4$ repetitions. Under this protocol there seems to be a precession of the time of the filter's signal peak in relation to the next repetition time although the timing has been set to $T_r = t_p$ so it hits the signal peak. An approximate signal $S.P$ is made from the superposition of a single filter signals on every repetition time. The $(T_r = 18_{S.P})$ reveals that this effect of the moving peaks is not due to a change in the filter internal dynamics but a result of signal superposition by repetition. The $T_r = F.P$ and $T_r = F.P_{S.P}$ signals result from a protocol timed to repeat on the apparent mean signal peak, which progressively get smaller with a ratio 2/3. Finally, changing the repetition interval to relatively long values $T_r = 120$ shows that the maximum peak signal does not rise and thus memory repetitions do not interact.

intervals that begin with a $T_r = t_p$ and progressively shorten by a factor of 2/3. Plotting a signal with such a succession of intervals gives $T_r = F.P$. This protocol begins with spaced intervals and ends with massed intervals giving an overall signal that rises to a single peak, which is higher than the asymptotic maximum signal of a regularly spaced protocol of fixed intervals $T_r = 18$. Thus, shorter intervals appear to give higher peak signals, a massed protocol gives the highest peak signal while the F.P protocol gives a peak signal which is between the massed and spaced protocol.

In the next section, we examine the phenomenon of peak precession and whether it reflects a change in the timing of filter state resymmetrization.

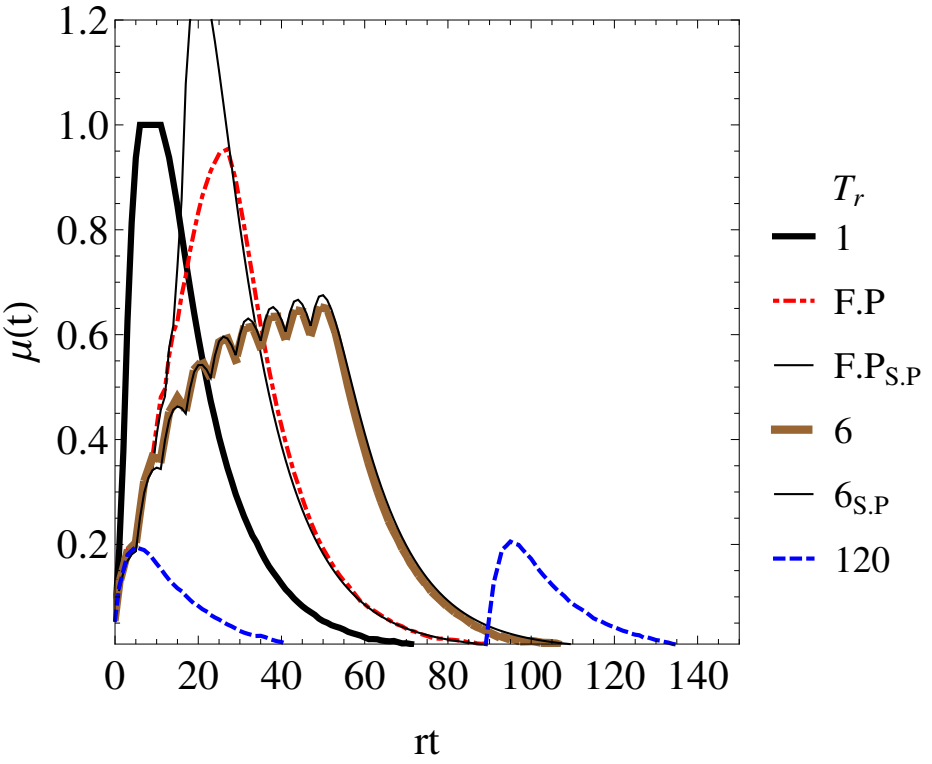


FIGURE 8.3: Applying different $n_r = 8$ repetition protocols on a single $\Theta = 4$ LP filter (no decay $\eta = 0$) under eight repetitions ($n_r = 8$). The qualitative effect of these protocols are the identical to the $\Theta = 7$ filter shown on Figure 8.2, with $T_r = 1$ giving maximum signal. Also, the on-peak $T_r = 6$ protocol gives an asymptotic maximum signal with the regressing signal peak phenomenon shown reproduced by the superimposed signal as well.

8.2.1 Spaced repetition changes the apparent timing of signal peak

In previous analysis of filter dynamics we showed that the time of the filter's peak signal (t_p) coincides with the time at which the filter's state distribution symmetrizes, see Figure 8.4. If a memory is repeated at the t_p then we would expect that the next time filter states re-symmetrize will also coincide with the next filter peak. However, Figure 8.2 shows that under a $T_r^\star$ memory repetition protocol it appears that the signal's peaks move towards shorter intervals after each repetition although repetition intervals are fixed to $T_r = t_p$, which is the mean time until filter states re-symmetrize. To examine this phenomenon, we construct an approximate mean signal under repetition through the superposition of mean signals from independent memory systems $M(t) = \mu(t) + \mu(t_1) + \mu(t_2) \cdots \mu(t - t_i)$, where $\mu(t - t_i)$ is the signal resulting from a single encoding of the tracked memory with a delay time $t_i = i * T_r$ set to the time of memory repetition i. This approximate signal is shown on both Figures 8.3 and 8.2 for two protocols, a spaced on-peak $T_r^\star$ and the advancing interval protocol ($F.P$). Overlaying this approximate signal with an exact signal for the $T_r^\star$ protocol shows that the exact and the approximate superimposed signal ($S.P.$) match. Therefore, the change in filter

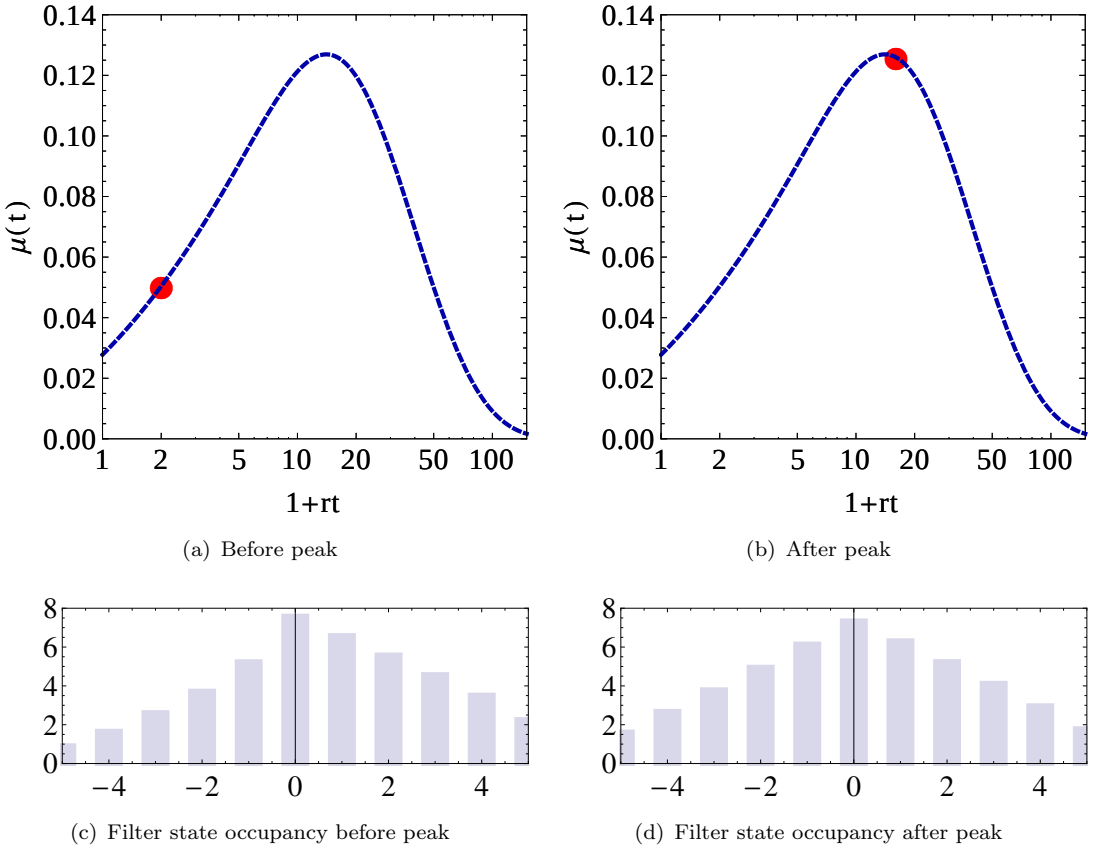
peak signal timing is only a result of the superposition of independent signals and not some change in the filter's time to re-symmetrize.

The approximate signal appears to work quite well on Figure 8.2 against the equivalent protocols using the exact filter. However, Figure 8.3 shows that the approximation breaks down for $T_r = F.P$ when the signal goes beyond $\mu(t) > 0.5$. The $F.P_{S.P}$ is shown here in order to reveal that the superposition approximation signal works under a constraint regime where the signal is far from signal saturation (< 1) while the size of the approximation error depends on the difference in the exact signal and the equilibrium filter state distribution. The first obvious constraint is that our superimposed signal does not account for an upper bound imposed by a finite population of synapses and this is shown on Figure 8.3 where the $F.P_{S.P}$ is made of an unconstrained superposition of individual (constrained) filter signals. Other than the limits imposed by the saturation of the signal, we also expect an error coming from differences in the internal filter state distribution at the time of repetition. For $T_r < t_p$ before peak-signal, the distribution has not reached equilibrium and therefore, given the triangular shape of the distribution (see Figure 8.4(d)), any repetition would cause a larger number of synapses to cross threshold than if repetition was to occur when filter states have re-symmetrized. Under this non-symmetric distribution, repetition would have the same effect on the strength of biasing. The number of synapses biased towards a given threshold is increased as the distribution is shifted further away from equilibrium and therefore the error in the approximation for some T_r and Θ also depends on the difference in the expected number of synapses biased under the exact filter distribution undergoing memory repetition and the equilibrium one used by the approximate signal. However, both of these differences are diminished as the threshold size is increased, in the limit of large Θ the asymptotic filter state distribution would approach a uniform distribution and therefore any shift in the internal distribution would cause minor differences in bias and the number of near-threshold synapses. Thus, the approximation using superposition improves as Θ increases and this is shown in the differences between the $T_r = F.P_{S.P}$ signal for $\Theta = 7$ on Figure 8.2, which works quite well, against the $T_r = F.P_{S.P}$ for $\Theta = 4$ shown on Figure 8.3 where the approximation fails as it does not follow the $T_r = F.P$ signal.

8.2.2 Threshold-crossing statistics depend on repetition intervals

Having examined the influence of memory spacing T_r and the number of repetitions n_r on the mean signal we proceed to examine their effects on the statistics of the internal filter states. With respect to filter state distributions the timing of memory repetition T_r can be broken down to two basic periods, the time before state resymmetrization and the one after it. Figure 8.4 shows the mean filter state occupancy at these two distinct periods in time. At the time of the signal peak the distribution is symmetric and therefore there is no bias on expressing towards either filter threshold in particular,

in contrast before the signal has peaked the distribution is asymmetrical with a higher occupancy towards one threshold. It is therefore anticipated that the effect of memory



(a) Before peak

(b) After peak

(c) Filter state occupancy before peak

(d) Filter state occupancy after peak

FIGURE 8.4: Distribution of filter state occupancy prior and just after signal peak. The distributions here include the synapses that have been re-injected to zero after initial memory encoding and thus Figure **c** includes a peak on zero and a shift rightwards. The distribution right after the signal peak shows symmetry and thus no bias exists towards expression through either threshold. These distributions have been scaled by Θ^2.

repetition on filter distributions will depend on the interval T_r. This is because the strength of biasing and the number of synapses found near threshold depend on state of the internal filter distributions and therefore the effects of repetition on both of these factors will change depending on the repetition intervals T_r and the number of repetitions n_r.

To measure how repetition protocols change the filter state dynamics we examine the statistics of TCs by sampling the distribution of the number of consecutive same threshold crossings c before a synapse changes strength. Here, we refer to a same-threshold crossing also as a *metaplastic* transition. By definition there are two thresholds and therefore the size of TC, denoted by c, is a count of the number of threshold-crossings

over the threshold that does not change the synaptic strength. TCs are labelled according to whether they occur over a *correct* or *wrong* state, with correct being relevant to the tracked memory being repeated. Thus, if the tracked memory being repeated requires a synapse to be strong and this synapse performs x metaplastic transitions through the upper threshold before it performs a plastic transition through the opposite threshold then we let $c = x$. When the desired state and threshold transition do not match we give a negative sign to the threshold cycle $c = -x$.

Memory repetition within some constrained time interval was previously shown to augment the memory signal, this implies that an increased number of synapses was crossing through a given threshold to obtain the desired synaptic strength. Therefore, we anticipate that repetition would entrain synapses through metaplastic transitions over the correct synaptic strength either by pushing synapses through the desired boundary immediately at the time of repetition or by increasing a synapse's bias towards crossing the correct threshold. Thus, memory repetitions should enhance the number of correct metaplastic transitions and therefore increase the average size of correct TCs.

Using simulation of $N = 10^6$ synapses of a filter $\Theta = 7$ we obtain empirical distributions of TCs under massed and spaced protocols. Figure 8.6 compares the TC distribution of a massed, a spaced and an F.P protocol for $n_r \in \{4, 8\}$ repetitions. On this Figure, the spaced protocol shown has a lower occupancy for the first three cycle sizes ($c < 4$) than the massed repetitions protocol, and this is true for $n_r = 4$ and $n_r = 8$. The lower occupancy in small TCs, is due to higher occupancies distributed among larger TCs that are difficult to spot in the figures. Indeed, examining the PDF of escape through the same threshold under a single repetition, we find that the probability of same-threshold crossing is near the signal peak, see Figure 8.5. Thus any other repetition protocol than spaced should give intermediate results between the TCs distributions seen from massed and spaced protocols.

The TC distribution of the F.P protocol, which has a range of repetition intervals between massed and spaced, shares characteristics from both the massed and the spaced protocol TC distributions. For $n_r = 4$ the F.Ps intervals $T_r \in \{18, 12, 8, 5\}$ give distributions closer to a spaced protocol while for higher n_r, where T_r becomes even shorter, the F.P protocol begins to display higher occupancies in smaller c values as expected. Comparing Figure 8.6 to Figure 8.7 we see that increasing the filter size to $\Theta = 13$ makes the differences between massed and spaced protocols become less pronounced than the smaller filter $\Theta = 7$ while the F.P. protocol distribution still lies in between a massed and a spaced protocol.

We conclude that the largest differences are seen between massed ($T_r = 1$) and spaced ($T_r^\star$) repetition protocol TC distributions when considering intervals that range from the time immediately following encoding and the time until filter states re-symmetrize. Therefore we will be focusing on discriminating between these two as any other T_r in

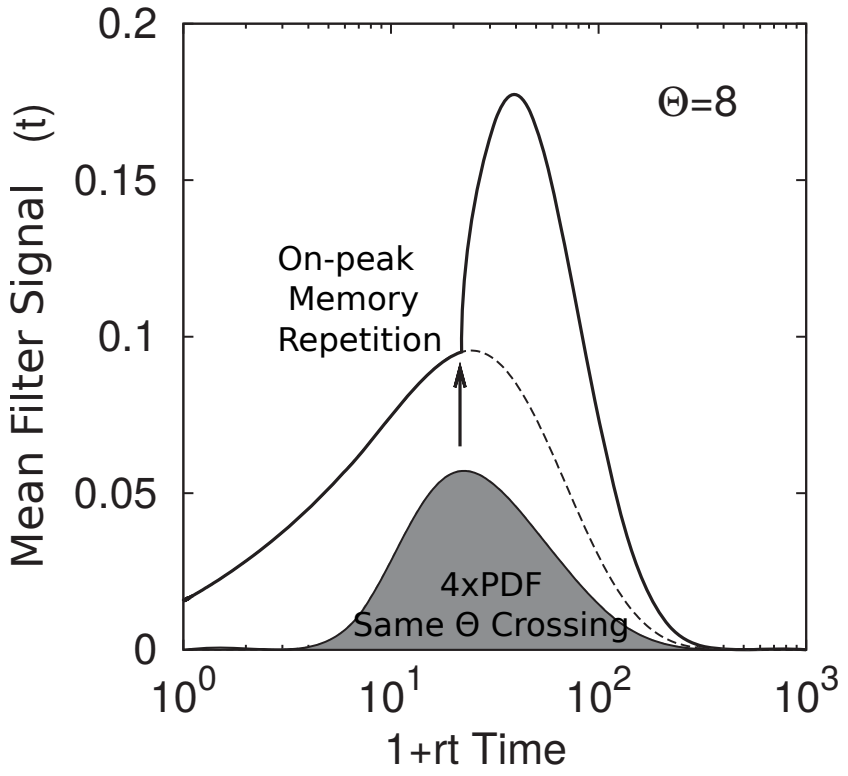
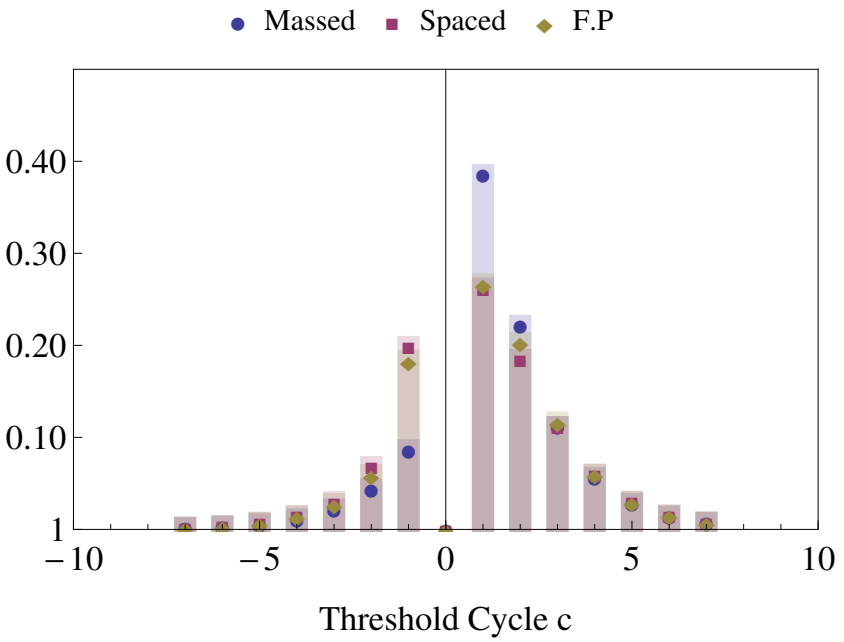


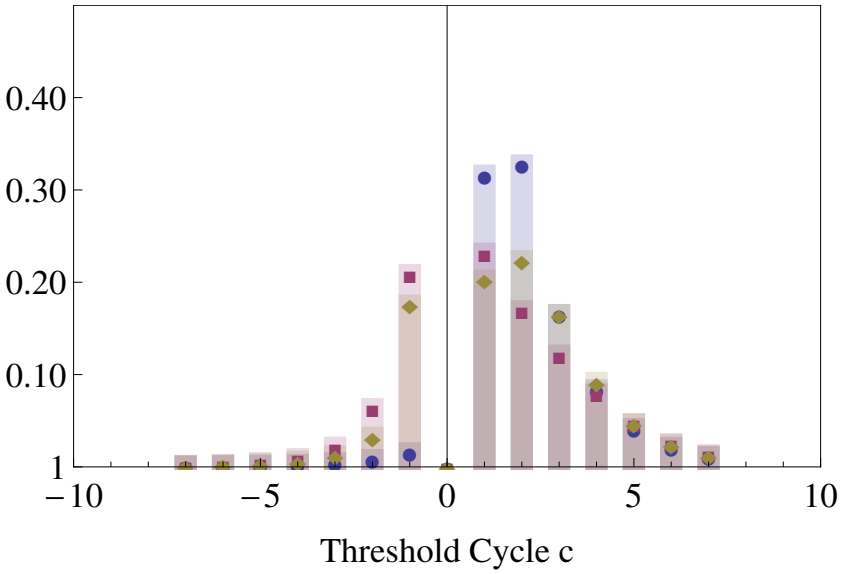
FIGURE 8.5: The probability of a synapse going through the same threshold after a memory repetition is maximized if the repetition time t_{r_1} of a memory stored at $t = 0$ is near or just before the peak. The arrow indicates the repetition time when the PDF of escape from state zero through the correct threshold, given by $G_0^+(t) + G_0^-(t)$, is maximal. (The pdf is scaled four times for visibility). The dashed line shows the course of the mean signal if there had been no repetition of the memory encoding.

the range will involve a mixture between spaced and massed intervals that bring about intermediate effects. In later results we look at whether synapses can discriminate $T_r^\star$ against protocols of comparatively very long repetition intervals such that the repetition events do not effect the TC distribution. We refer to distribution under this protocol as $\mathrm{TC}_{\mathrm{neutral}}$.

The massed protocol shifts the internal filter distributions towards the desired threshold deterministically. Once a synapse crosses the desired threshold and it is re-injected to the zero filter-state further massed repetitions will increase the biasing towards the same threshold. Consequently, a long ($n_r = 8$) massed repetition protocol will result in a concentration of TC around a few of the smaller positive c cycles. On the other hand, the spaced repetition protocol achieves an apparently small enhancement of the small cycles but it also enhances larger positive TCs. The enhancement relies on repeating a memory at a time when filter distributions have re-symmetrized to re-introduce a bias for a synapse to go through the same threshold again. With this re-biasing method spaced repetition successively pushes a subset of synapses, whose size is relative to the biasing strength, through the same threshold at each repetition and thus it achieves an increase in the size of threshold cycles c by regularly repeating memories at the time $T_r^\star$.

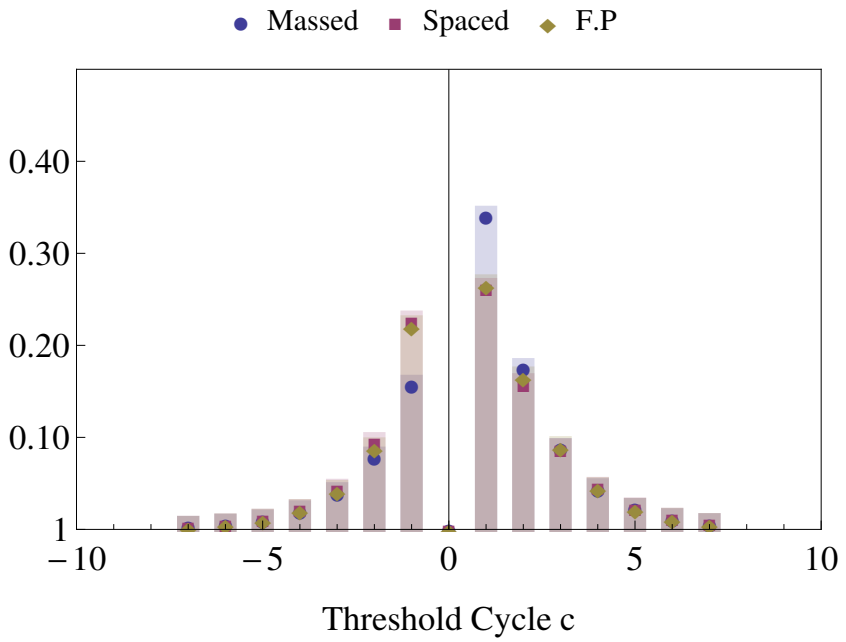


(a) $n_r = 4$

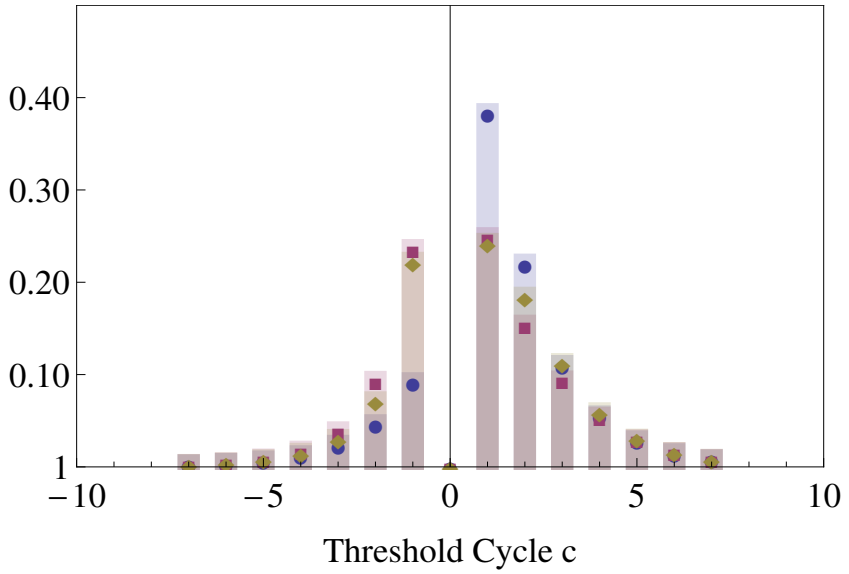


(b) $n_r = 8$

FIGURE 8.6: Threshold-cycle distribution for three repetition protocols, a massed, a spaced protocol and a mixed protocol of progressively smaller intervals ($F.P$) on a $\Theta = 7$ filter sampled over $N = 10^6$ synapses. The upper figure compares the protocols under $n_r = 4$ repetitions and the lower under $n_r = 8$. In both figures the massed protocol shows a characteristic increase in the smaller threshold cycles $c < 4$ when compared to the other two protocols. The $F.P$ protocol gives distributions that share characteristics with both massed and spaced protocols. Under $n_r = 4$ the $F.P$ is closer to spaced protocol, while for $n_r = 8$ it increases its occupancy in small c. This transition is a result of the decreasing repetition intervals $T_r \in \{18, 12, 8, 5, 4, 2, 2, 1\}$ of the $F.P$ that become massed beyond $n_r > 4$. Thus, the largest differences in threshold-cycle distributions are seen between a massed and a spaced protocol.



(a) $n_r = 4$



(b) $n_r = 8$

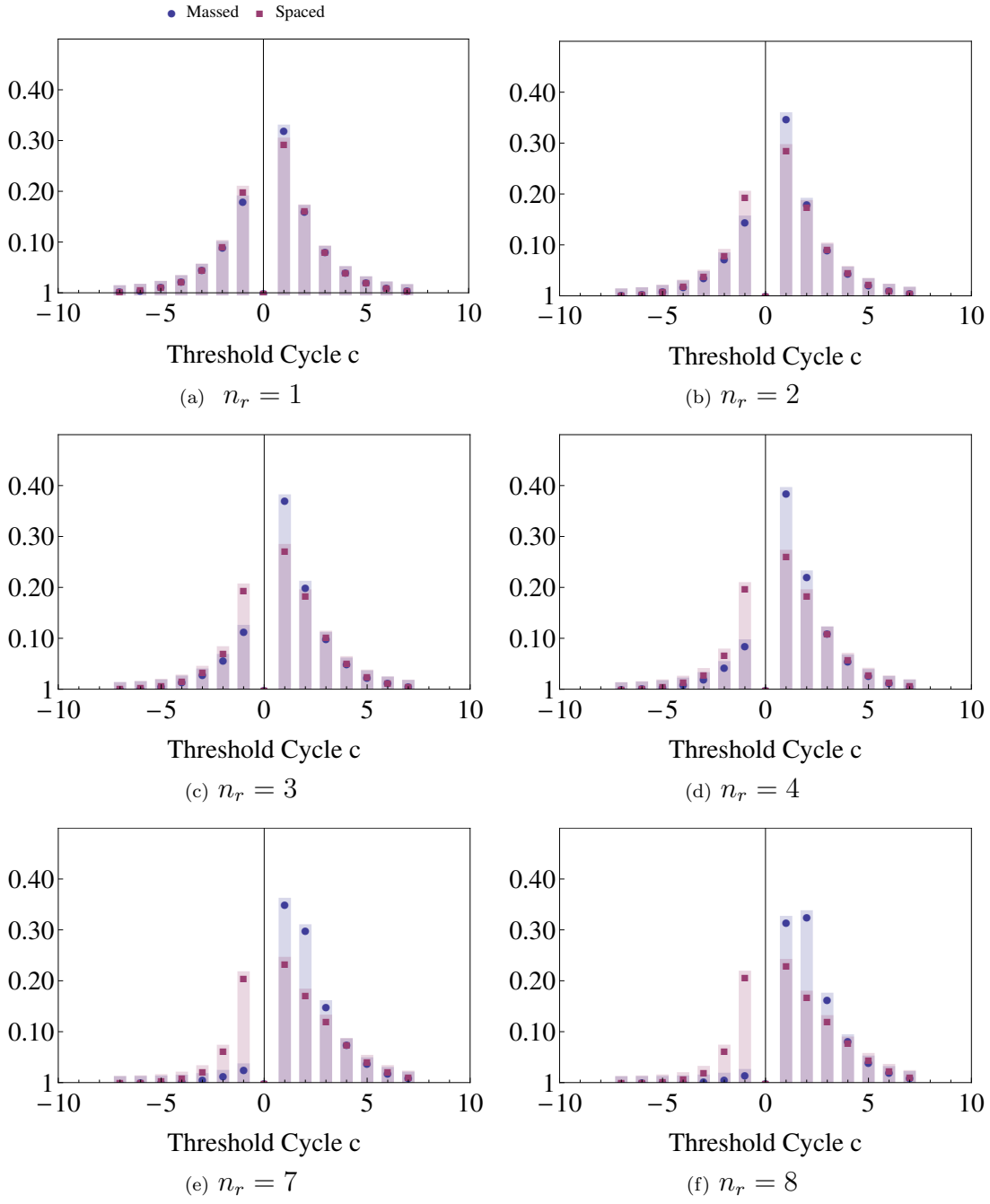
FIGURE 8.7: Threshold-cycle distribution for three repetition protocols, a massed, a spaced protocol and a mixed protocol of progressively smaller intervals ($F.P$) on a $\Theta = 13$ filter. As with $\Theta = 7$ of figure 8.6, the characteristic difference of a massed protocol to the other two protocols is the increase in the smaller threshold cycles. The spaced protocol is very similar to the $F.P$ for low $n_r = 4$ while $F.P$ moves closer to the massed protocol at $n_r = 8$. This transition of $F.P$ towards a massed distribution results from its intervals $T_r \in \{63, 42, 28, 19, 12, 8, 6, 4\}$ moving closer to massed repetition intervals beyond $n_r > 4$

Figure 8.8 shows how differences between TC distributions of massed and spaced protocol develop as n_r increases for a filter of $\Theta = 7$. At $n_r = 1$ there are minor differences

between the distributions of massed and spaced protocols with the former showing a slightly higher occupancy at $c = 1$. By $n_r = 8$ these differences develop further. Massed repetitions result in enhancing the occupancy of small threshold cycles and depleting negative c. On the other hand, spaced repetition at $n_r = 8$ shows a wider enhancement of positive c and minor higher occupancy at large threshold cycles ($c > 4$) than the $T_r = 1$ protocol. Critically, spaced protocols always occupy $c = -1$, while massed ones may deplete it under high n_r. These differences are amplified under smaller filter sizes, for example Figure 8.9 shows TC distributions of a $\Theta = 4$ filter.

The results from the smaller filter $\Theta = 4$ verify that $n_r = 4$ massed repetitions give a relatively larger increase in $c = 1$ and $c = 2$ threshold-cycles. This follows naturally because four immediate repetitions on a filter as small as $\Theta = 4$ would deterministically push synapses through the desired threshold by shifting the internal filter distribution four places to the right. Doubling the number of repetitions to $n_r = 8$ shows that massed repetition now strongly enhances $c = 3$ cycles too. Here, almost all synapses are pushed to perform one threshold crossing and then, once re-injected to zero, half of them are further pushed towards the same threshold after which the population mostly splits between synapses that close a $c = 2$ cycle and ones that perform $c = 3$ cycle. Again here, spaced repetition appears to enhance a wider range of positive threshold cycles while as shown on Figure 8.10 increasing the filter size to $\Theta = 13$ attenuates the differences between $\mathrm{TC}_{\mathrm{massed}}$ and $\mathrm{TC}_{\mathrm{spaced}}$ distributions.

Earlier we discussed the possibility that induction protocols may be such that they don't represent the encoding of a memory naturally and thus each particular repetition in a protocol may represent a strong encoding episode. Within the synaptic filter framework these strong encoding episodes may be represented by multiple POT/DEP induction steps on the filters. Thus, here each memory repetition is not equal to the encoding of any other memory in the sequence of ongoing memory encoding, which normally induces a single induction step, but rather lets assume that it causes multiple steps. For example, we let each memory encoding induce two plasticity stimuli. For a $\Theta = 7$ under a spaced protocol of $n_r = 4$, the standard sequence of repetition intervals $T = \{18, 18, 18, 18\}$ is replaced by $T = \{18, 1, 18, 1, 18, 1, 18\}$ and for the relevant massed protocol it becomes $T = \{1, 1, 1, 1, 1, 1, 1\}$. By inducing two steps at each encoding the biasing induced and the number of synapses driven through threshold are those of a filter half the size showing that such a protocol simply achieves a reduction in the filter size for the memory being strongly encoded. This theme can be extended to four induction steps per encoding episode by adding further single induction steps at each repetition time. Figure 8.11 shows TC distributions of a $\Theta = 7$ filter under massed and spaced protocol using two step and four step encoding at each repetition. We find that the $n_r = 4$ two step protocols give distributions similar to those of an $n_r = 8$ one step protocol of Figure 8.8(f). Under the even stronger four step protocol the TC distributions shift further towards larger

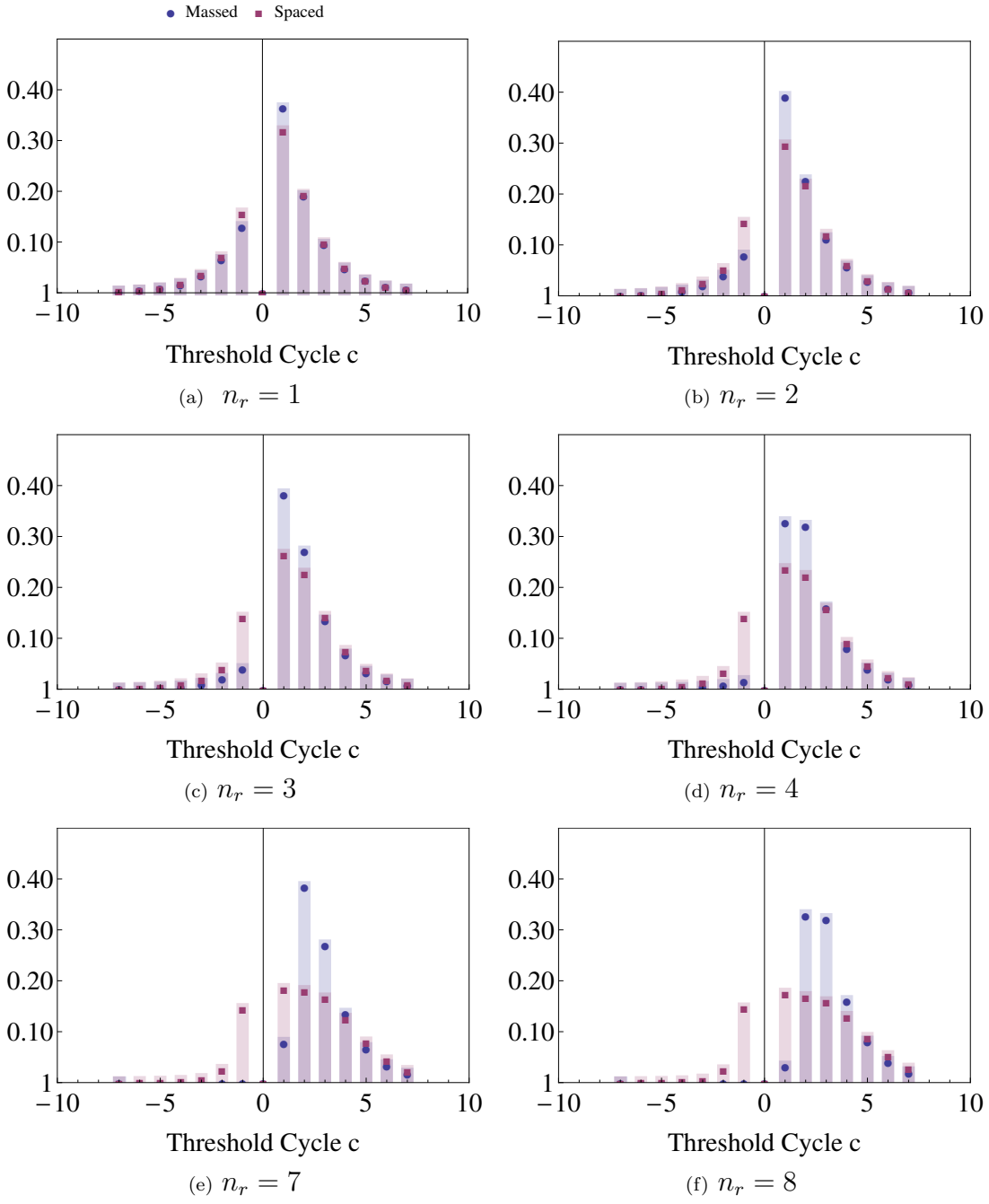


(a) $n_r = 1$ (b) $n_r = 2$ (c) $n_r = 3$ (d) $n_r = 4$ (e) $n_r = 7$ (f) $n_r = 8$

FIGURE 8.8: Comparing a massed $T_r = 1$ versus a spaced $T_r = 18$ protocol under increasing number of repetitions on the distribution of threshold cycles c using a $\Theta = 7$ filter. As n_r increases the $T_r = 1$ shows an increase in the first few cycles $c < 3$ against the $T_r = 18$. In contrast to a massed protocol, a spaced protocol always retain occupancy in $c = -1$. Results obtained over simulation of $N = 10^6$ synapses stopping when all synapses have completed one threshold cycle after initial encoding.

cycles but still retain the relative differences and overlap between a massed and spaced protocol, see Figure 8.11(b).

Given the influence that the repetition protocol has on TCs and the relative differences



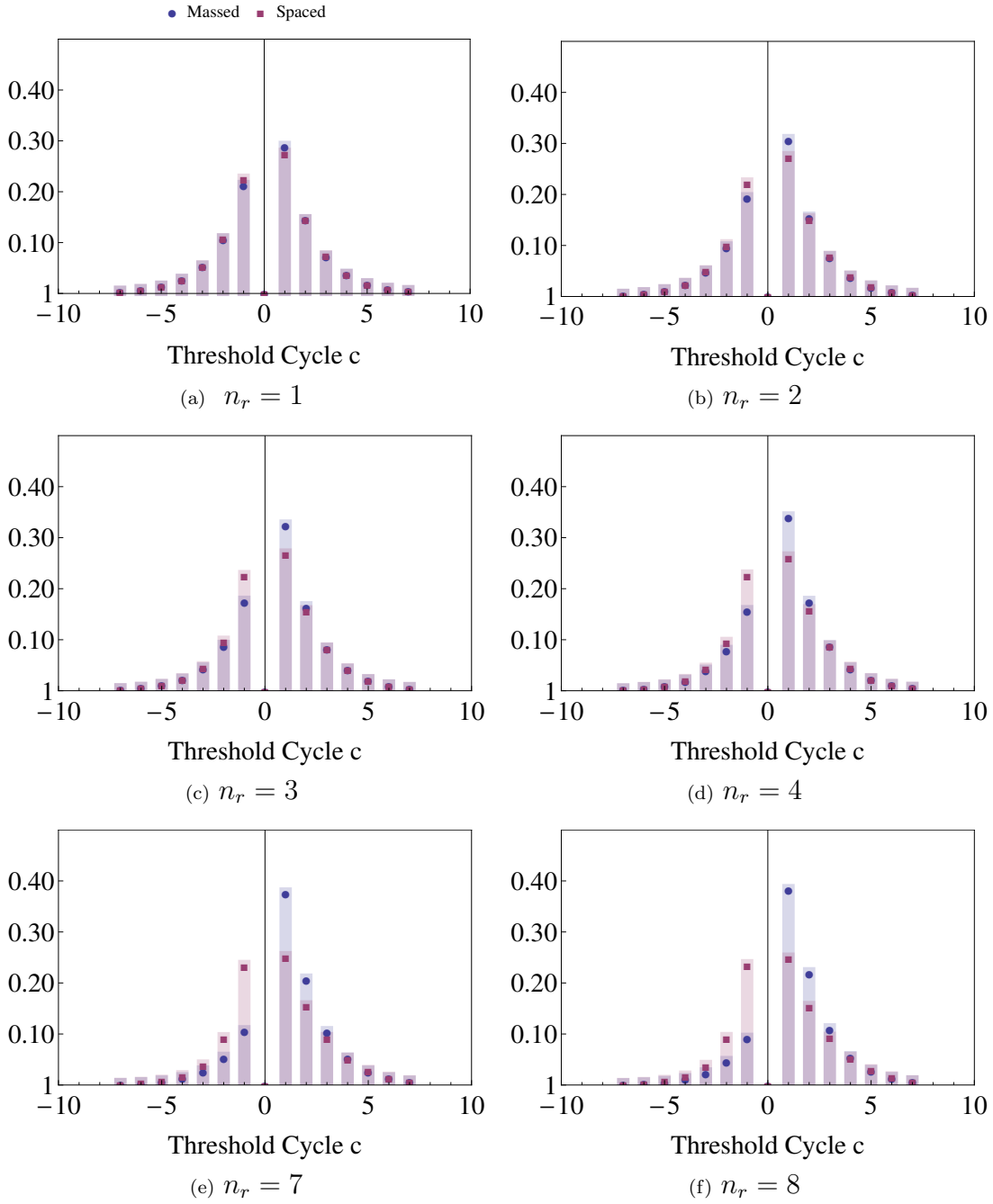
(a) $n_r = 1$ (b) $n_r = 2$

(c) $n_r = 3$ (d) $n_r = 4$

(e) $n_r = 7$ (f) $n_r = 8$

FIGURE 8.9: Using a $\Theta = 4$ filter to compare a massed $T_r = 1$ against a spaced $T_r = 6$ protocol under increasing number of repetitions on the distribution of threshold cycles c. As n_r increases the $T_r = 1$ shows an increase in the first few cycles $c < 3$ against the $T_r = 6$. Results obtained over simulation of $N = 10^6$ synapses stopping when all synapses have completed one threshold cycle after initial encoding.

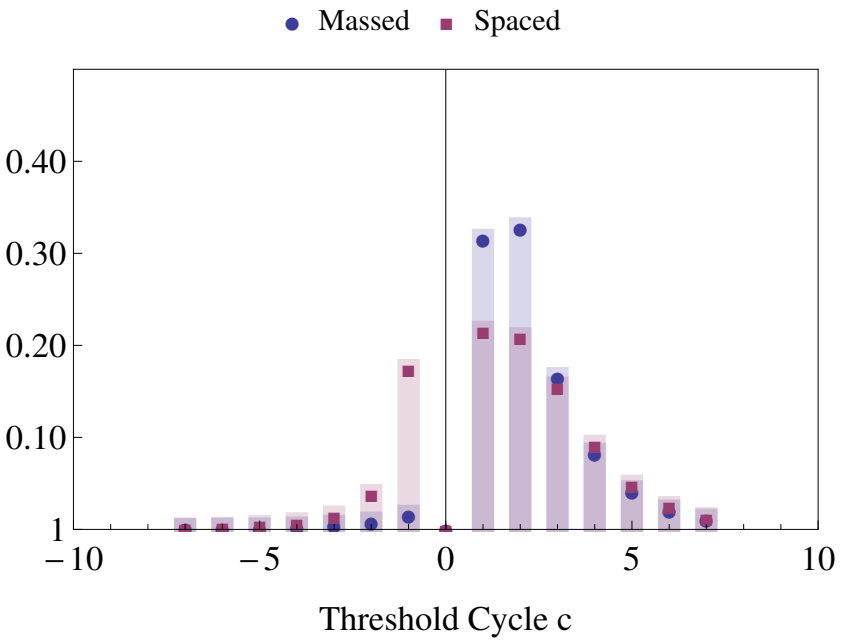
between massed and spaced protocols, it may be possible for a single synapse to estimate which protocol is being induced by examining its TC distribution accumulated in real-time. Each synapse could in principle, obtain a temporal average of the distribution



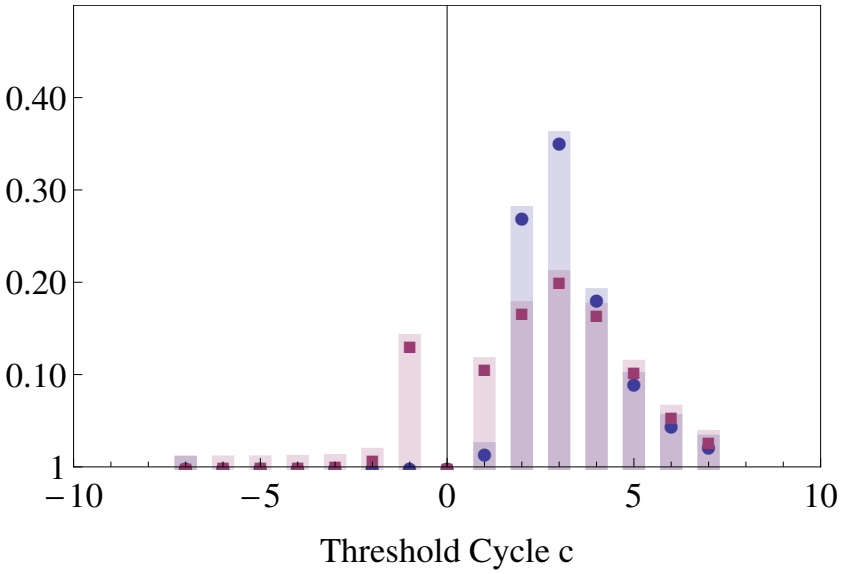
(a) $n_r = 1$ (b) $n_r = 2$ (c) $n_r = 3$ (d) $n_r = 4$ (e) $n_r = 7$ (f) $n_r = 8$

FIGURE 8.10: Comparing a massed $T_r = 1$ versus a spaced $T_r = 63$ protocol under increasing number of repetitions on the distribution of threshold cycles c using a $\Theta = 13$ filter. As n_r increases the $T_r = 1$ shows an increase in the first few cycles $c < 3$ against the $T_r = 63$. Results are over simulation of $N = 10^6$ synapses and the sample is obtained when all synapses have completed one threshold cycle after initial encoding. A characteristic of long $n_r \geq \Theta$ massed protocols is the reduction in variability shown by the diminished occupancy of negative TCs.

of TCs and use it to decide when to trigger the processes that lead to long-term plasticity. We proceed to compare a temporal averaged distribution against the empirical distributions above to establish if single synapses can obtain clear enough information



(a) 2 Step $n_r = 4$



(b) 4 Step $n_r = 4$

FIGURE 8.11: $\Theta = 7$ two step and four step encoding protocols equivalent to a $n_r = 4$. The intervals of the two- step spaced protocol are $T_r \in \{13, 1, 13, 1, 13, 1, 13, 1\}$ and every spaced interval encoding is followed by an immediate extra induction step. Multi-step encoding accelerates the affect of repetition protocols on TCs. For example the $n_r = 4$ two step protocol results in a distribution similar to the one step $n_r = 8$ of the same filter. The four step gives a large shift of both distributions toward larger cycles but retains the characteristic differences between massed and spaced protocols, where massed protocols enhance smaller size cycles than spaced protocols.

to detect the protocol being delivered.

8.3 Temporal average of threshold-cycle distribution

At this stage, we wish to examine if synapses could, in principle, estimate the form of TC distribution by accumulating samples in time under multiple repetitions (trials) of a memory encoding protocol. One TC sample is obtained by a synapse after each trial of a repetition protocol. Each sample obtained represents an independent re-iteration of the induction protocol and thus spatial and temporal averages of TC distributions are theoretically identical. We aim to estimate if a reasonably small sample size is sufficient for a synapse to discriminate whether a spaced protocol is being applied. If this is possible, then single filter synapses could, in principle, detect and trigger late-phase plasticity under a few iterations of a plasticity induction protocol. In the previous section we found that TC distributions between massed and spaced protocols have characteristic features that can be identified by looking at empirical distributions averaged over a very large sample $N = 10^6$. Specifically, massed repetition protocols when compared against those resulting from a spaced protocol have the distinctive feature of high occupancies concentrated around a couple of threshold-cycles. Here, we focus on a spaced protocol with $n_r = 4$ repetitions and a repetition interval set to the on-peak filter time. The form of the TC distribution for a filter of size $\Theta = 7$ under a spaced protocol averaged over $N = 10^6$ synapses was shown on Figure 8.6(a), while the TC distribution of a $\Theta = 13$ filter is shown on Figure 8.7(a).

To support the accumulation of a relevant distribution, a signal is required to initiate the TC counting process after the initial encoding of some memory of interest. Such a signal could be delivered at the initial encoding of an important event. Once the signal is perceived synapses initiate the threshold-cycle measuring process in response. The biological nature of this signal could come in the form of the neuromodulatory signals we covered in Chapter 2, which were necessary for stable long-term forms of memory and plasticity. Nevertheless, we are interested in the theoretical limits of the decision making process while any specific biophysical implementation constraints are initially ignored.

We examine the convergence of a sampled distribution under a spaced protocol against the empirical distributions obtained under other protocols. First, we induce a spaced protocol on $N_{syn} = n$ of our theoretical synapses and we let them accumulate a distribution of threshold-cycles after the salience signal has been received. Then, using measures from statistical goodness-of-fit tests we obtain the distance of the n sample distribution against empirical distributions obtained under a massed, a spaced and a neutral repetition protocol over m synapses. We repeat the procedure for progressively larger samples n and then look for the point where the distance between sample ($N_{syn} = n$) and empirical ($N_{syn} = m$) distributions is sufficient to safely separate between distributions of different protocols. For example, when a spaced protocol is repeatedly being administered over n trials (here taken as n synapses tested independently), the difference of

the accumulated distribution and that of the spaced empirical distribution (see Figure 8.6(a)) obtained will decrease as the number of trials n increases. At the same time, the distance to the empirical distribution from other protocols should increase.

We use a standard measure drawn from statistical goodness-of-fit tests for discrete variables called the *Pearson's test* (Pearson, 1900) to obtain the difference between distributions. Using this test we estimate the occupancy probabilities of each threshold-cycle size by combining $N = n + m$ the sample and the empirical distribution A and B to obtain the expected frequencies for each threshold-cycle.

$$\hat{\pi}_i = (A_i + B_i)/N = C_i/N \tag{8.1}$$

Then we compute the discrepancy measure using a squared difference to compare the sampled process over a sample size n against the highly averaged empirical distributions ($m = 10^6$).

$$T = \sum_{i=c_{min}}^{c_{max}} \left[\frac{(A_i - m\hat{\pi}_i)^2}{m\hat{\pi}_i} + \frac{(B_i - n\hat{\pi}_i)^2}{n\hat{\pi}_i} \right], \tag{8.2}$$

where $c_{min/max}$ define the range/width of the distribution of threshold cycle size we wish to examine (here the maximum width is from $c_{min} = -16$ to $c_{max} = 16$. With the T measure we may obtain the average distance of a distribution of n spaced protocol applications to the empirical distribution of either massed or spaced protocol. After every trial of the $T_r^\star$ protocol with $n_r = 4$ the distribution of TC is stochastically determined and so its distance to the empirical distributions T is a random variable and therefore we need to obtain a number of samples n to estimate its mean $< T >$ and variance. Figure 8.12 shows the mean distance $< T >$ of a $\Theta = 7$ filter subjected to a spaced protocol n times against the empirical distributions of a spaced protocol $\mathrm{TC}_{\mathrm{spaced}}$ as T_{SS}, the distance to $\mathrm{TC}_{\mathrm{massed}}$ as T_{SM} but also the distance of the sampled TC distribution against a protocol with a repetition time $T_r = 120$ as T_{SN}. This last distance T_{SN} is used to check if a synapse can detect an on-peak spaced protocol against one which has very long repetition intervals occurring at a time when the signal has died away. As required the sampled distribution from synapse receiving a spaced protocol converges towards zero showing that the T_{SS} distance to the spaced empirical distribution is minimized as the sample size n increases. Also, T_{SM}, the distance of the sampled distribution against $\mathrm{TC}_{\mathrm{massed}}$ shown on Figure 8.12, reflects the difference between massed $\mathrm{TC}_{\mathrm{massed}}$ and $\mathrm{TC}_{\mathrm{spaced}}$ spaced distributions seen by inspecting Figure 8.6(a). T_{SM} shows that the distance of the sampled TC converges towards a constant distance against the empirical distribution of a massed protocol ($\mathrm{TC}_{\mathrm{massed}}$). The same is true of T_{SN}, which sits between T_{SM} and T_{SS}, showing that the distribution to a protocol of very long repetition intervals can be discriminated from a $T_r^\star$ spaced protocol.

The standard deviations for each of T_{SS},T_{SM} and T_{SN} are also shown on Figure 8.12 using errorbars about the mean. We may use this information to estimate when a

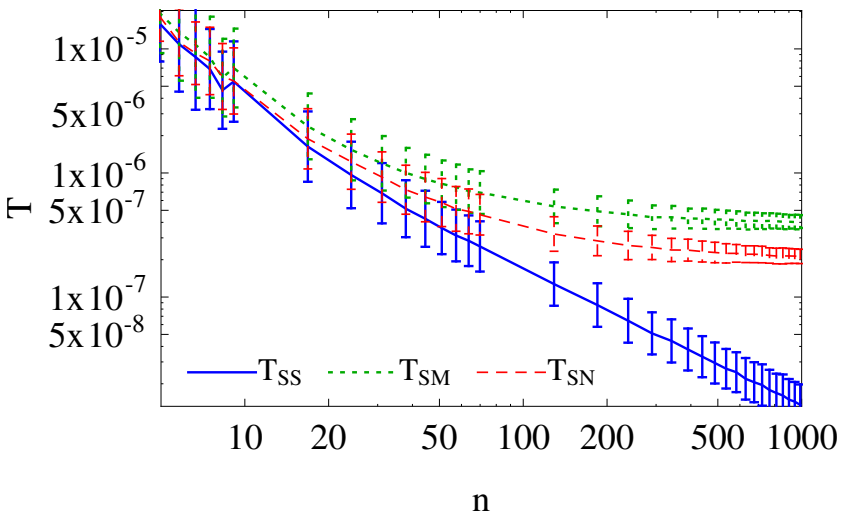


FIGURE 8.12: Convergence of threshold-cycle distribution using Pearson distance measure T with a filter $\Theta = 7$ under an $n_r = 4$ spaced $T_r^\star$ protocol repeated over n trials. Each point represents the mean and standard deviation of the random variable T averaged over 10^3 trials. Curve T_{SS} shows the distance of the sampled distribution to the asymptotic empirical distribution of a spaced protocol. T_{SM} is the distance of the sampled distribution against the $\text{TC}_{\text{massed}}$ empirical distribution and T_{SN} is the distance to an empirical distribution of a very long repetition interval ($T_r = 120$). T_{SS} separates by one standard deviation from the other two after $n > 200$ protocol samples have accumulated.

synapse can reliably detect that the sampled distribution is a $\text{TC}_{\text{spaced}}$ by looking for the point where the standard deviation of T_{SS} does not overlap the standard deviation of either T_{SM} or T_{SN}. We find that for this filter $\Theta = 7$ a synapse would require the accumulation of distribution of around 200 repetition protocol application trials before it could safely detect that a spaced protocol is being used.

Figure 8.13 shows that reducing the filter size to $\Theta = 4$ accelerates convergence to around $n = 60$ samples, beyond which the standard deviation of T_{SS} separates from T_{SN}. On the other hand, increasing the filter size to $\Theta = 13$ has the opposite effect. Figure 8.14 shows that in this case we require over a 1000 samples for the standard deviations to separate well.

Using a standard test to compare discrete distributions we have estimated the number of repeated applications of a standard $n_r = 4$ memory encoding protocol before a synapse can safely discern which TC distribution it is converging to. We found that even for filter sizes as small as $\Theta = 7$ convergence would require on average the accumulation of 200 protocol trials before a synapse could safely determine a spaced $T_r^\star$ protocol.

If LTM has a specific minimum repetition requirement then it is further worth investigating whether a protocol with n_r lower than four can be discriminated from an $n_r = 4$ protocol. Further, in a previous section we argued that perhaps strong encoding protocols are perceived by synaptic filters not as single induction steps but as multiple

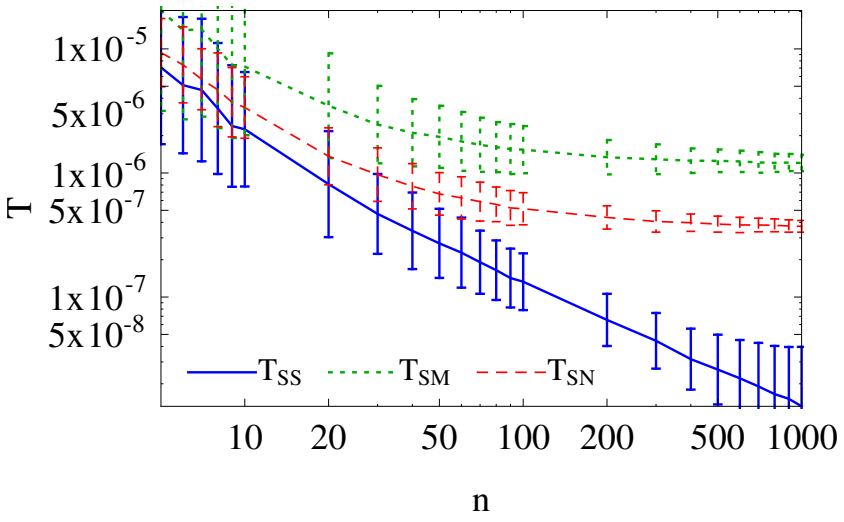


FIGURE 8.13: Using a Pearson's measure to obtain the average distance of a TC distribution under n spaced protocol repetitions against the $N = 10^6$ empirical distribution of a spaced protocol T_{SS}, massed protocol T_{SM} and a long-interval repetition protocol T_{SN}. T is averaged over 1000 trials with the error bars show the standard deviation of the random variable. For this small filter $\Theta = 4$ under an $n_r = 4$ spaced $T_r^\star$ protocol, a synapse could safely detect the spaced protocol above $n > 60$; beyond which point the standard deviation of T_{SS} does not overlap with the distance against other protocols T_{SM}, T_{SN}.

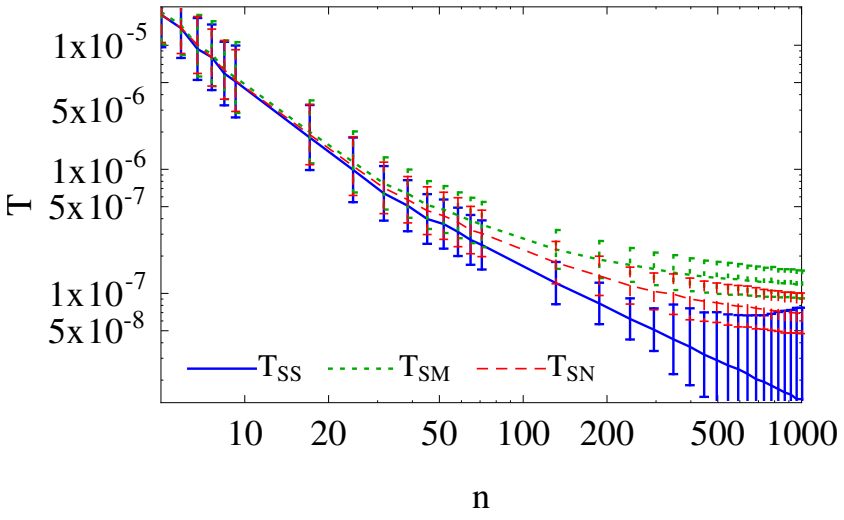


FIGURE 8.14: Large filter $\Theta = 13$ Pearson's convergence of threshold-cycle distribution to the spaced protocol distribution with $n_r = 4$. Curve T_{SS} shows the distance of the sampled distribution to the asymptotic empirical distribution of a spaced protocol. For larger filters the distance in standard deviations between a massed and a spaced protocol is smaller when compared with smaller filters. Also, for this filter a synapse cannot clearly identify an on-peak spaced protocol from a non-repetitious protocol.

induction steps on each encoding repetition. Knowing that smaller filters give quicker convergence of the sampled distribution to $\text{TC}_{\text{spaced}}$ perhaps such multi-step encoding protocols could effectively reduce the filter size and give a reasonable convergence time.

However, before we attempt to answer these questions of comparing alternative encoding protocols we are going to verify our results against a different goodness-of-fit measure.

8.3.1 The Kolmogorov-Smirnov goodness-of-fit test

We next turn to a goodness-of-fit measure known as the *Kolmogorov-Smirnov* (K-S) so we may confirm the discrepancies found with the Pearson's measure. The K-S goodness-of-fit test is simpler, it is based on the maximum difference between the two cumulative distribution curves:

$$D = sup_{-k<i<+k}[abs(F_i - G_i)] \tag{8.3}$$

where k is the number of bins in the distribution of TC, which we have limited to $-10 < k < 10$, and *sup* denotes the supremum, which returns the first element greater than or equal to all elements of a subset. If a set has a greatest element then that is the supremum of that set.

Using this measure we plot the K-S goodness-of-fit results, as before for a synapse receiving an $T_r^\star$ protocol with $n_r = 4$ for filter sizes $\Theta = 4$,$\Theta = 7$ and $\Theta = 13$ on Figures 8.15, 8.16 and 8.17 respectively. The results obtained are qualitatively similar to the

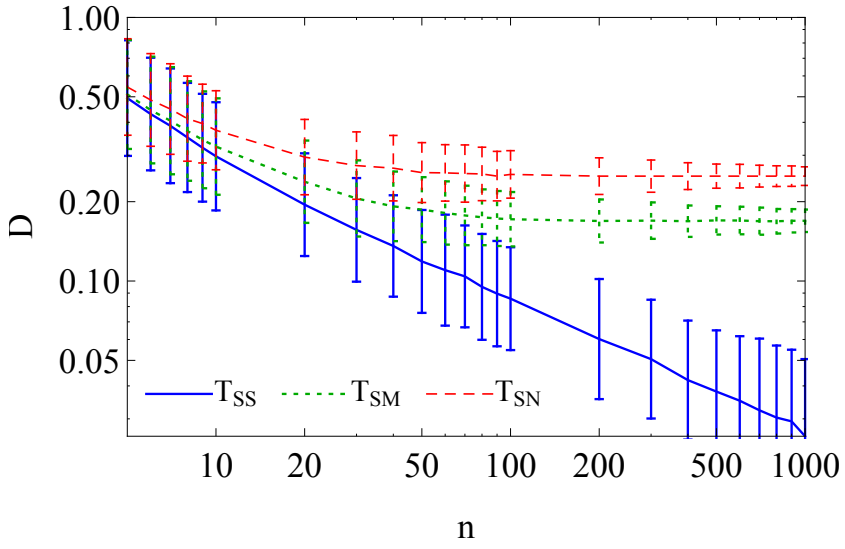


FIGURE 8.15: Using a K-S measure on a small filter $\Theta = 4$ to obtain the average distance of the TC distribution under n spaced protocol repetitions against the $N = 10^6$ empirical distribution of a spaced protocol T_{SS}, massed protocol T_{SM} and a no repetition protocol T_{SN}. D is averaged over 1000 trials and the error bars show the standard deviation of the random variable. For this small filter, the TC distributions differentiate above $n > 100$ after which point a synapse could safely detect a spaced protocol because the standard deviation of T_{SS} does not overlap with the standard deviation of the distance from other protocols T_{SM}, T_{SN}.

earlier Pearson-test results, although the absolute values of D and T measures differ. A noticeable difference to earlier results is on a small filter $\Theta = 4$ that on Figure 8.15

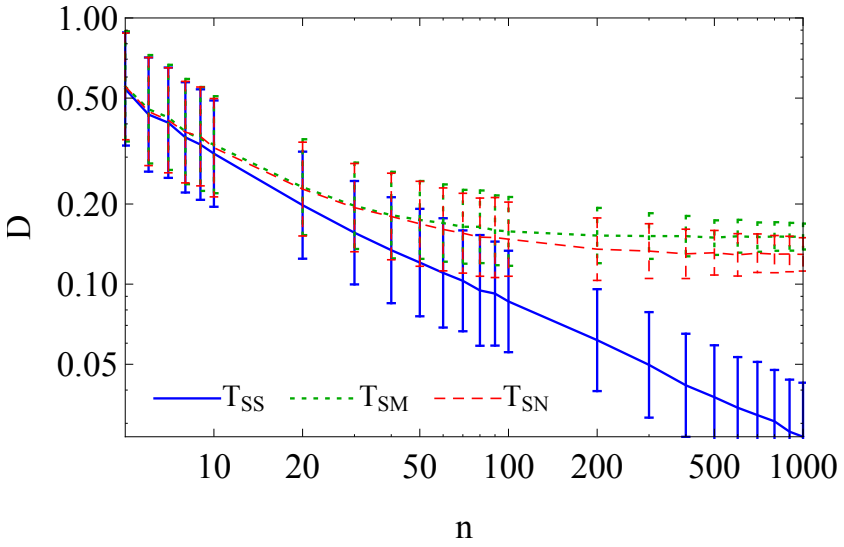


FIGURE 8.16: Using a K-S measure to obtain the average distance of a TC distribution under n spaced protocol repetitions against the $N = 10^6$ empirical distribution of a spaced protocol T_{SS}, massed protocol T_{SM} and a no repetition protocol T_{SN}. The mean of the random variable D was obtained over $N = 1000$ trials for each point and the standard deviation is shown by the error bars. For this $\Theta = 7$ filter the standard deviations separate from above $n > 200$ beyond which a synapse could safely detect a spaced protocol.

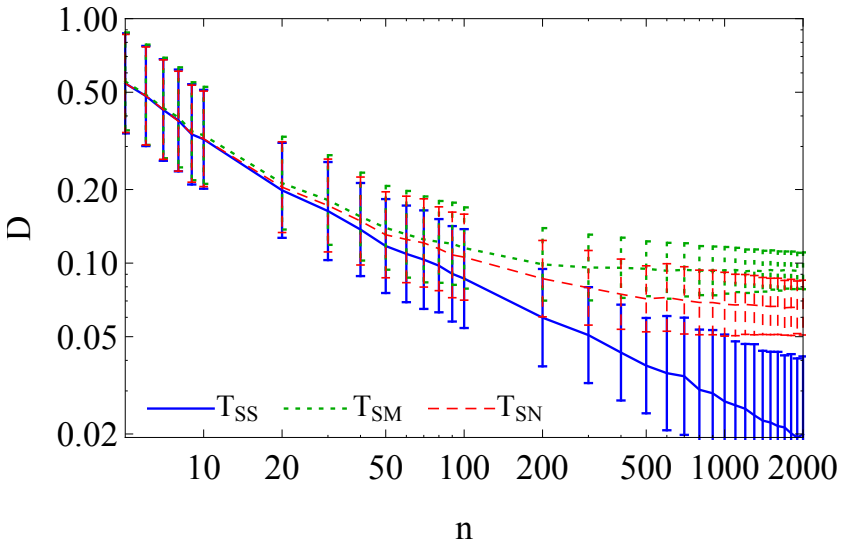


FIGURE 8.17: K-S goodness-of-fit for a large filter $\Theta = 13$ of threshold-cycle distribution to the spaced protocol distribution. Curve T_{SS} shows the distance of the n sample distribution to the asymptotic empirical distribution of a spaced protocol. For larger filters the distance in standard deviations between a massed and a spaced protocol is smaller when compared with smaller filters. Also, for this filter the synapse cannot clearly identify an on-peak spaced protocol from a non-repetitious protocol unless $n > 1000$ samples are obtained.

the distance of T_{SS} to the neutral distribution T_{SN} is now larger than the distance to T_{SN}, meaning that for the K-S measure $\text{TC}_{\text{spaced}}$ and $\text{TC}_{\text{massed}}$ appear closer than what

the Pearson measure suggests. However, for larger filter sizes this difference disappears. Also, on the larger filter $\Theta = 13$, Figure 8.17 shows that T_{SN} and T_{SM} are now close to each other with respect to their standard deviation. These differences should not be of concern however, since we are interested in detecting a spaced protocol against the other two protocols, and Figure 8.17 shows that beyond $n > 1000$ T_{SS} separates well from T_{SM} and T_{SN}. This sample size value is similar to the one obtained by the *Pearson* goodness-of-fit measure, for a filter of this size, but it appears that the K-S goodness-of-fit measure is somewhat better at classifying the sampled TC distribution if one compares standard deviation distances for $n > 1000$ with Figure 8.14.

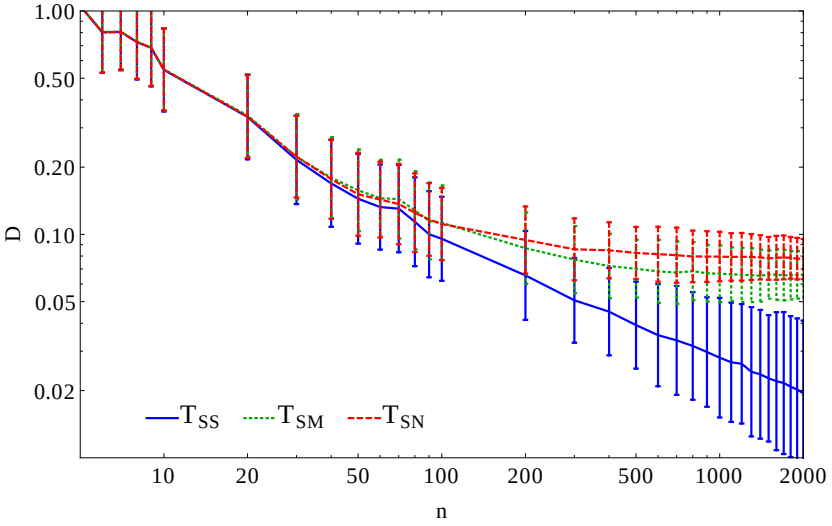


FIGURE 8.18: K-S goodness-of-fit for a filter $\Theta = 7$ of threshold-cycle distribution to the spaced protocol distribution of few $n_r = 2$. Curve T_{SS} shows the distance of the n sample distribution to the asymptotic empirical distribution of a spaced protocol with n_r. Compared to an $n_r = 4$ the convergence is slower due to the relatively smaller differences between the TC distributions between $\text{TC}_{\text{spaced}}$ and $\text{TC}_{\text{massed}}$. For this filter the synapse cannot clearly identify an on-peak spaced protocol of $n_r = 2$ from a non-repetitious protocol unless $n > 1400$ samples are obtained.

The results obtained so far have focused on discrimination between massed/spaced protocols of fixed n_r over different filter sizes Θ. If the filters are able to detect particular stimulation conditions through the TC distribution then the distributions of these conditions should be relatively unique to the protocol being applied. By default, both $\text{TC}_t extmassed$ and $\text{TC}_t extspaced$ start from a close distance to the $\text{TC}_t extneutral$ and move away as n_r is increased. For few $n_r = 2$ Figure 8.18 shows the convergence of the $\Theta = 7$, which when compared to the $n_r = 4$ Figure 8.16 of the same filter we find a seven-times increase in convergence from $n > 200$ to $n > 1400$. Further, from Figure 8.9 we observe that the TC distribution evolves differently under massed protocols than it does under spaced ones and we have discussed the clear difference of there always being an occupancy at $c = -1$ under spaced protocols which is depleted under long n_r massed protocols. Due to these differences $\text{TC}_t extmassed$ distribution will not pass

through a distribution that belongs to a TC$_t$*extspaced* as n_r is increased, and vice-versa. For example we test the case on whether a massed protocol of $n_r = 2$ passes over a TC$_t$*extspaced* of $n_r = 4$ on Figure 8.19 and find that is not the case, a distance between both spaced/massed protocols is seen when the applied $n_r = 2$ is compared against the TC$_t$*extspaced* of $n_r = 4$.

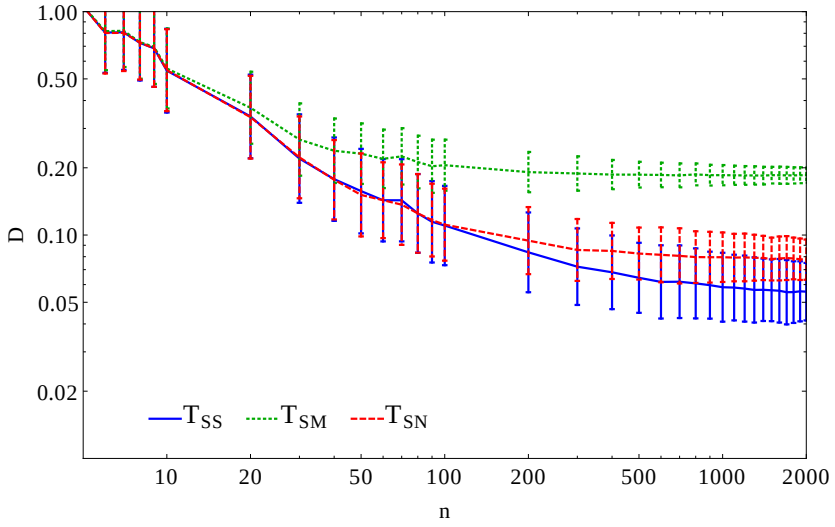


FIGURE 8.19: K-S goodness-of-fit for a filter $\Theta = 7$ of the TC distribution to spaced protocol distribution of $n_r = 4$ when the massed/spaced protocol being applied has only $n_r = 2$. Curve T_{SS} shows the distance of the n sample distribution of an $n_r = 2$ protocol against the asymptotic empirical distribution of a spaced protocol with $n_r = 4$. The distance never goes to zero as required, since the distributions of $n_r = 4$ and $n_r = 2$ protocol are different. Thus TC$_t$*extmassed* does not look-like a TC$_t$*extspaced* of $n_r = 4$ before it differentiates, as seen when comparing convergence of a TC$_t$*extmassed* of $n_r = 4$ vs TC$_t$*extspaced* of $n_r = 4$. Further, from the T_{SN} we find that the spaced distribution is closer to the neutral one (no repetition protocol applied), than the massed.

In any case, standard spaced and massed protocols require synapses to obtain a large number of TC samples, while increasing filter size only amplifies the problem. Perhaps the experimental stimulation protocols can be re-interpreted in such a way that they provide faster convergence. In the next section we examine protocols that induce multiple induction stimuli at each memory repetition.

8.3.2 Multi-step protocols reduce convergence time

As we discussed earlier in this chapter, commonly electrophysiology protocols use four electrical stimuli of 1 second duration at 100Hz to induce l-LTP. During this one second it is very possible, given the stochastic nature of a synapse, that out of the 100 current pulses a synapse experiences two or more induction stimuli.
It may even be that the number of stimuli received by a synaptic filter under these intense stimulation protocols even exceeds the number of filter states. In such a case we

may assume that each stimulus provides $2\Theta - 1$ steps and therefore shifts all synapses through the desired threshold after each stimulus is applied.

It is easy to see that four massed repetitions will induce at least six threshold crossings deterministically, as synapses will be pushed through a threshold and then once re-injected to zero they will be pushed further by the train of induction stimuli to obtain the familiar massed protocol distribution showing most TC concentrated around two c values. On the other hand, under a multi-step spaced protocol there is no guarantee that a synapse will experience the same number of threshold crossings as in the massed protocol. A synapse can make an opposite threshold crossing during the time between repetitions. Thus even under a multi-step spaced protocol the characteristic of a $\mathrm{TC}_{\mathrm{spaced}}$ distribution of a wider occupancy of threshold-cycle sizes will be retained. An example empirical distribution of this scenario can be seen on Figure 8.20 for a $\Theta = 4$ filter with seven-step encoding protocol of $n_r = 4$ repetitions. Given how TC distributions develop even under multi-step protocols it is clear there will be an overlap between T_{massed} and T_{spaced} and therefore classifying stimuli will require that synapses obtain more than a handful of distribution samples to identify which protocol is being induced. Even in the

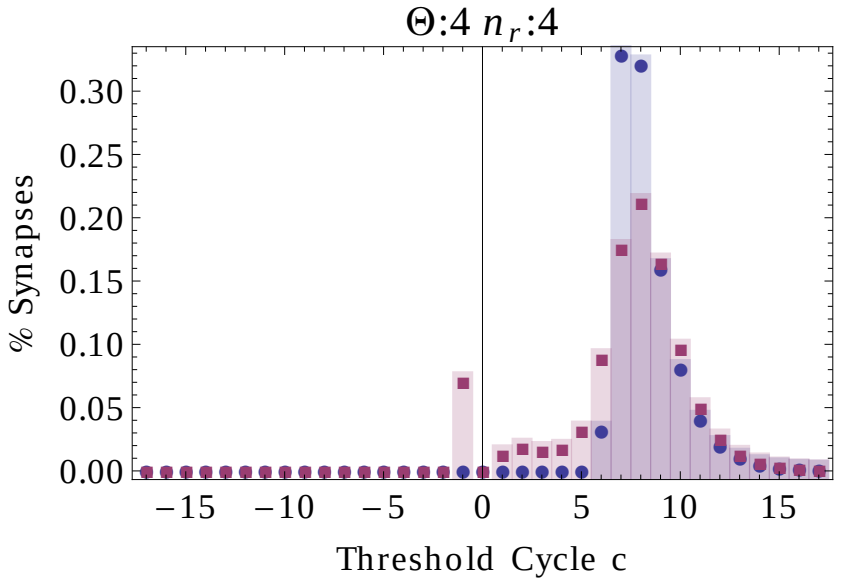


FIGURE 8.20: Under a $2\Theta - 1$ step encoding protocol with a $\Theta = 4$ filter, a massed protocol gives a continuous train of 28 induction stimuli resulting in a minimum of six c. On the other hand a synapse experiencing a spaced protocol has a distribution resembling a binomial around six threshold-cycles. In this extreme example of multi-step encoding we see the characteristic differences between massed and spaced protocol TC distributions.

limit of large n-step encoding the spaced distribution would obtain very low occupancy at the c where the massed protocols distribution peaks, yet due to this overlap, on average, a single memory encoding protocol application would not be enough for a synapse to determine which protocol is being induced.

However, multi-step protocols could possibly reduce the convergence time to reasonably low values which could relate to experimental results. For this reason we test the convergence of a standard size filter $\Theta = 7$ under a multi-step induction protocol for two and four steps and examine how the minimum required sample size changes.

The empirical distribution ($N = 10^6$) obtained from a $\Theta = 7$ filter under the influence of a two step protocol were shown earlier on Figure 8.11(a), and for a four step protocol on Figure 8.11(b). Comparing these figures to the earlier standard one-step protocols we observed that the two step protocol distributions are now closer to the ones obtained under a filter of almost half the size $\Theta = 4$ as shown on Figure 8.9(d), and for the four-step protocol the distribution appears similar to one with twice the number of repetitions $n_r = 8$ for $\Theta = 4$ as on Figure 8.9(f). Thus, multi-step protocols make the TC distribution of a filter appear as one of a smaller filter size, this comes naturally as multi-step protocols effectively reduce the size of a filter by replacing a single induction step by multiple ones and therefore adjacent filter states are combined as one during the encoding episode. Due to the differences between single step and multi-step TC distributions we take the multi-step empirical distribution as being the target distribution a synapse will have to converge to instead of testing convergence of multi-step protocol towards the single-step empirical distribution. If in reality synapses trigger allocation in response to the multi-step TC then this would mean that for a single-step protocol to converge towards the multi-step TC the number of n_r repetitions will have to increase. If multi-step protocols turn out to converge quickly then the experimentally derived low $n_r = 4$ is only an artefact of intense stimulation and in reality synapses require more stimuli before they produce l-LTP.

We have already seen that smaller filters converge faster and since the multi-step protocols effectively reduce the size of a filter for the memory being repeated then we anticipate that the number of samples required will be reduced under these protocols. Figure 8.21 shows that with two-step encoding the separation of T_{SS} from the other two distributions for a $\Theta = 7$ filter is now possible from $n > 70$. Compared to Figure 8.16 where $n > 200$ was required under single step encoding of the standard $n_r = 4$ repetition protocol we find that a multi-step protocol has indeed accelerated the convergence but still requires at least $n = 70$ samples.

Increasing the number of induction steps on each encoding to four and measuring the convergence we obtain Figure 8.22, according to which the standard deviations now separate after sample size of $n > 50$. Thus, increasing the induction steps at encoding appears to reduce the minimum number of memory encoding protocol trials required for a synapse to identify if a $T_r^\star$ protocol is being induced. Nevertheless, a protocol with more than $\Theta/2$ induction steps only reduced the minimum samples to at least $n = 50$ which is still quite high, assuming that a single trial of a spaced four-repetition $n_r = 4$ protocol could be sufficient to induce LTM.

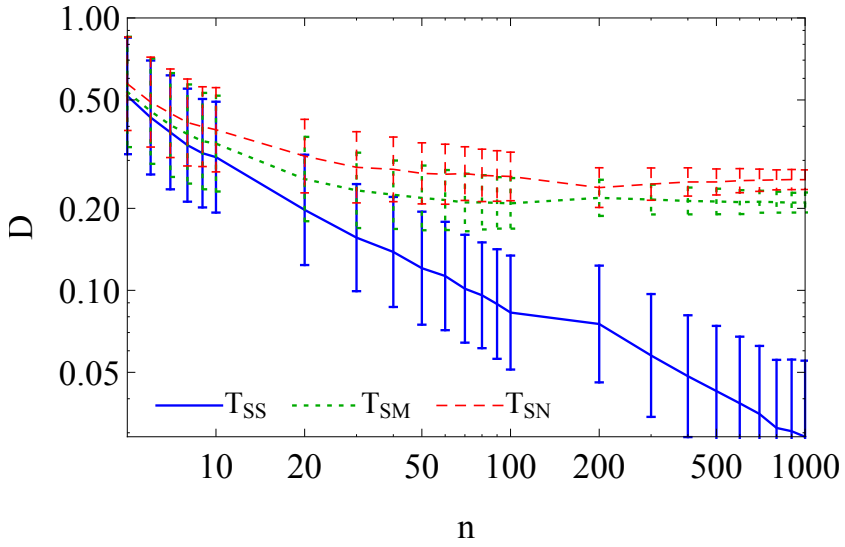


FIGURE 8.21: K-S goodness-of-fit of threshold-cycle distribution to the spaced protocol distribution for a large filter $\Theta = 7$ under a two-step spaced induction protocol. Compared against single step protocols the convergence is quicker, for this filter a synapse will be able to safely identify an on-peak spaced protocol from other protocols after $n > 70$ samples have accumulated.

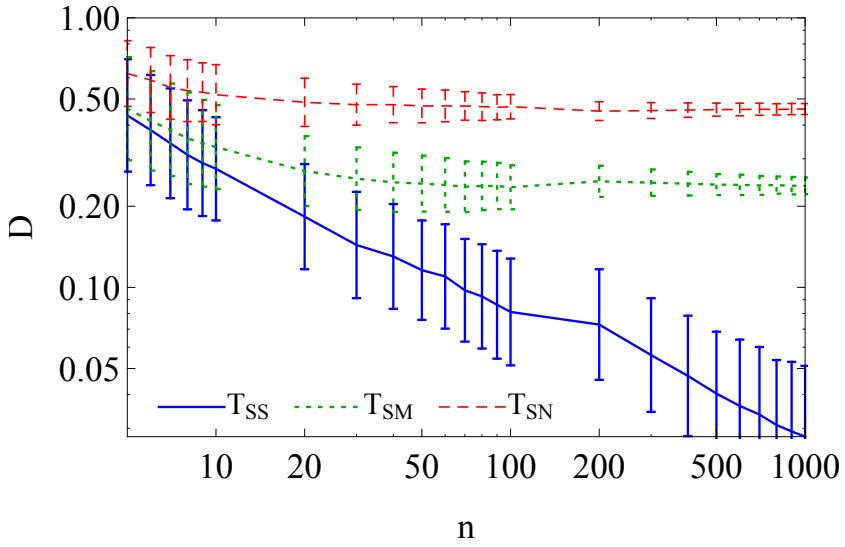


FIGURE 8.22: K-S goodness-of-fit of threshold-cycle distribution to the spaced protocol distribution for a large filter $\Theta = 7$ under a four-step spaced induction protocol. Compared to protocols of a single or two-step this four-step protocol has quicker convergence. We find that it is possible to safely identify an on-peak spaced protocol from other protocols after $n > 50$ samples have accumulated.

However, with single step encoding we required $n > 200$ samples, when we doubled the number of induction steps the samples reduced to approximately $n > 70$ and a further doubling to four resulted to $n > 50$ showing that the increase in convergence speed slows down. From a $200/70 = 2.8$ ratio in convergence time between one-step to two-step encoding it drops to $70/50 = 1.4$ between two-step and four-step and, given our

earlier discussion on how TC_{massed} and $\text{TC}_{\text{spaced}}$ distributions evolve with an overlap even under multi-step encoding (see Figure 8.20), it is reasonable to expect that the minimum n samples level off somewhere in the range of a few tens of samples.
Therefore, even under strong multi-step encoding protocols a synapse would require tens of repetition protocol trials before it can identify the protocol being induced. It appears then that synapses require alternative means to quickly obtain better statistics of the protocol being induced and we shall explore such mechanisms in the next chapter.

In the meantime, it is worth exploring if we can replace the two standard methods we used to compare TC distribution above with a simpler method. Given the differences in the form of the TC distributions we discussed above and the fact that the K-S measure uses cumulative distributions to compare distributions, perhaps a simple threshold carefully positioned on TC could be used to discriminate between $\text{TC}_{\text{massed}}$ and $\text{TC}_{\text{spaced}}$.

8.4 Testing a simple Threshold-cycle allocation criterion

As discussed in the previous section, massed repetition protocols result in TC distributions that concentrate occupancies above some $c = c_\Theta$, which depends on the filter size and the number of induction stimuli. In contrast, spaced repetition results in wider distribution of synapses among positive c as the number of n_r increases. Although we have seen that there are insufficient statistics available within a few memory repetitions, perhaps it is possible to statistically determine which protocol a synapse is likely experiencing through a simple threshold on c. The TC frequencies obtained are now seen as probabilities and the position of the threshold c can be translated to point on discrete probability distribution that gives the probability of a synapse performing $TC < c$. This interpretation allows us to replace the complexity of retaining an internal representation of the accumulated TC distribution and using a random variable D or T to statistically detect the protocol being induced by comparing the relative probability of a synapse being above c under a massed or spaced protocol. For this to work and thus offer some means of statistical detection, the total number of synapses found above some c under strong spaced encoding protocols would need to differ from those under a massed protocol. Table 8.1 summarizes the comparison of the $\text{TC}_{\text{massed}}$ and $\text{TC}_{\text{spaced}}$ by listing the ratios of the cumulative TC distributions above some $c = c_i$. The ratios are shown for different n_r for a small $\Theta = 4$ filter and reveal that spaced protocols have relatively higher occupancies over high c as n_r repetitions increase. For example, for $n_r = 4$ and with a threshold of $c = 4$ there are 10% more synapses in $\text{TC}_{\text{spaced}}$ than there in a $\text{TC}_{\text{massed}}$ distribution.

Increasing the filter size to $\Theta = 7$ we obtain qualitatively similar results but as Table 8.2 shows, the differences are attenuated compared to the smaller filter. Further, increasing the filter size to $\Theta = 13$ gives the results of Table 8.3 which confirm the attenuation by

n_r	c_1	c_2	c_3	c_4	c_5	c_6	c_7
1	0.95	0.95	1.0	1.0	1.0	1.0	1.0
4	0.82	0.82	0.87	1.1	1.1	1.2	1.2
8	0.81	0.81	0.65	0.73	0.98	1.2	1.3

TABLE 8.1: Using a threshold to compare $\mathrm{TC}_{\mathrm{spaced}}$ and $\mathrm{TC}_{\mathrm{massed}}$ distributions for a filter size $\Theta = 4$. The values listed are ratios of the sums of occupancies above c of $\mathrm{TC}_{\mathrm{spaced}}$ and a $\mathrm{TC}_{\mathrm{massed}}$. The columns represent the threshold cycle c over which we sum $\sum_{j=c}^{\infty} C_j$ the TC distributions and the rows represent the number of repetitions n_r in the protocol. In the second row $n_r = 4$ the values are from the distribution shown on Figure 8.9(d) and in the last row $n_r = 8$ those from Figure 8.9(f). We find that for this filter size a spaced protocol of $n_r = 4$ pushes synapses deeper in threshold cycles c compared against a massed protocol and this effect becomes stronger for $n_r > 4$.

looking at the $n_r = 4$ row, yet it is still possible to discriminate the two distributions under $n_r = 8$. Nevertheless, in previous sections we found that filter size also prolonged the convergence of temporal averages of TC distributions and thus this problem is not unique to the c_Θ threshold mechanism.

n_r	c_1	c_2	c_3	c_4	c_5	c_6	c_7
1	0.96	0.96	1.0	1.0	1.0	1.0	1.0
4	0.81	0.81	0.93	1.0	1.1	1.1	1.1
8	0.71	0.71	0.70	0.88	1.0	1.2	1.2

TABLE 8.2: Cumulative threshold cycles ratios between massed spaced protocols for a filter size $\Theta = 7$. The columns represent the threshold cycle c over which we sum $\sum_{j=c}^{\infty} C_j$ the distributions shown on Figure 8.6 and rows represent the number of repetitions used to produce the distribution. The values are the ratios of the sums of occupancies above c of spaced over massed. We find that for this filter size a spaced protocol of $n_r = 4$ pushes synapses deeper in threshold cycles c, but the effect goes away for $n_r > 4$.

n_r	c_1	c_2	c_3	c_4	c_5	c_6	c_7
1	0.98	0.98	1.0	1.0	1.0	1.0	1.0
4	0.86	0.86	0.96	1.0	1.0	1.0	1.0
8	0.73	0.73	0.81	0.93	1.0	1.1	1.1

TABLE 8.3: Ratio of cumulative distribution of threshold cycles of a spaced protocol over a massed protocol using a simulation of a $\Theta = 13$ filter of $N = 10^6$ synapses. We observe that spaced repetition gives almost always lower occupancies than massed repetition when summing from low threshold cycles. Spaced repetition favours larger threshold cycles when the number of repetitions is increased $n_r > 4$.

The results reveal that the distributions differentiate as the number of repetitions increase. Overall massed repetitions have higher occupancies when summing over low values of c_Θ. Under $n_r = 4$ massed repetitions give higher occupancies for $c < 4$ while spaced repetitions have higher occupancies above $c \geq 4$. Increasing the number of repetitions to $n_r = 8$ results in enhancing the contrast between massed and spaced repetitions. For a $\Theta = 7$ filter, even for this relatively large number of repetitions, under a spaced

protocol there are 20% more synapses than $\mathrm{TC}_{\mathrm{massed}}$ above $c_\theta > 4$, while for low c_θ the gain of $\mathrm{TC}_{\mathrm{massed}}$ over TC_{spaced} is 1.4. Therefore, a synapse that has conducted more than c_Θ TCs is more likely experiencing a spaced protocol than a massed protocol under sufficient repetitions $n_r > 4$. However, for larger filters and a high enough c_Θ we find that the occupancies of TC_{spaced} and $\mathrm{TC}_{\mathrm{massed}}$ are very close. According to the table above, the probability of a synapse experiencing either a massed or a spaced protocol is almost equal but a synapse which has conducted less than c_Θ same threshold-crossings is more likely experiencing a massed protocol.

Therefore, under certain conditions it is possible to preferentially allocate synapses experiencing a spaced repetitive protocol by setting a threshold-cycle criterion c_Θ. This would be a very simple mechanism for synapses to identify between the two TC distributions but it appears it offers a weak preference to a spaced protocol and therefore a synapse should utilize additional means that have access to improved statistics in order to identify the protocol being induced.

8.5 Summary

In this chapter we argued against the multiple synaptic stability states of the cascade and towards a synaptic filter with two stability states. This development was based on experimental evidence showing two main plasticity stability states but also based on our own theoretical results according to which a filter synapse of a comparable number of states to a cascade gives longer memory lifetimes. We consider the plasticity of single filter synapses we have seen up to now as representing e-LTP only and examine how transitions to stable plasticity such as l-LTP could be explicitly modelled within the filter framework. In our model a stable plasticity form is a "locked" plasticity state of synaptic filters to indicate the relatively long duration of l-LTP compared to the earlier phase where filters integrate ongoing stimuli.

Evidence from behavioural and electrophysiology experiments suggest that repetitive regularly spaced stimulation protocols are more efficient at producing LTM and l-LTP. Therefore, filter synapses are required to detect regularly spaced repetitive encoding and subsequently make a transition to the stable "locked" form of plasticity. Hence, we began to explore how different memory encoding protocols affect synaptic filters.

In terms of filter state distribution dynamics we can classify two types of repetition protocols, those that repeat stimuli before the filter states have re-symmetrized, and thus act to intensify an existing tendency in filter states, and those that repeat onto a symmetric filter state distribution to re-introduce a tendency. Initially, we examined the effects of these two types of memory repetition protocols on the mean memory signal. We found that memory repetition protocols with short repetition intervals $T_r = 1$ (massed) enhance the mean signal as the number of repetitions n_r is increased. On the other

hand, spaced repetition protocols with intervals timed to a filter's state resymmetrization ($T_r = 0.375\Theta^2$) showed that the apparent peak signal under repetition advances towards smaller intervals and the signal reaches an asymptotic peak beyond a few memory repetitions ($n_r > 4$). This asymptotic peak was lower than the maximum signal obtained if the same number of repetitions is delivered with massed intervals while we showed that the signal-peak precession phenomenon is an artefact of the superposition of filter signals. Repetition protocols with progressively smaller intervals timed to hit the apparent filter peak-signal resulted in a signal profile exhibiting a single peak which however was lower than the one obtained under a massed repetition protocol.

Using the fact that the two main types of repetition protocol act differently upon filter state dynamics we postulated that perhaps massed and spaced protocols give different threshold crossing behaviour. To measure this we defined the threshold-cycles measure as the number of consecutive same threshold crossing before a synapse conducts an opposite threshold crossing after the initial memory encoding. Indeed, empirical distributions of TCs revealed differences between massed and spaced repetition protocols and these were enhanced as the number of repetitions n_r increased. Overall, a few massed repetitions n_r strongly enhanced the occupancy of small size threshold cycles, while spaced repetitions gave a wider distribution with a somewhat increased occupancy of larger threshold-cycles. This relative gain of spaced protocols in increasing larger threshold-cycles occurs because such protocols have the effect of re-biasing the internal filter state distribution towards the desired threshold after it has symmetrized. On the other hand, the train of stimuli under a massed protocol deterministically pushes synapses to conduct a fixed minimum number of TCs and therefore distributions have a high occupancy at that point. Hence, under low n_r synapses are concentrated at small TCs.

Given these differences in TC distributions between the two main protocols, we then asked whether single filter synapses could identify a spaced protocol being induced within a few trials so that they make a transition to the stable state. We delivered repeated applications of a spaced memory encoding protocol and let synapses sample the TCs they conducted to retain an internal distribution of TCs. Then we measured the convergence of the sampled distribution towards the empirical distribution for a massed $\mathrm{TC}_{\mathrm{massed}}$, a spaced $\mathrm{TC}_{\mathrm{spaced}}$ and a neutral protocol $TC_{neutral}$ involving very long repetition times. Using two different statistical goodness-of-fit measures we found that convergence times for small filter $\Theta = 4$ required at least $n > 60$ samples while larger filters $\Theta = 13$ increased convergence to $n > 1000$ samples. These results are discouraging because that means that single synapses require at least n applications of a particular memory repetition protocol before they can detect a spaced protocol being induced with one standard deviation safety margin. However, the interpretation of the experimental protocols into induction stimuli for filters perhaps has been naive.

We then considered that the HFS experimental protocols that induce l-LTP are perceived by synapses not as single induction stimuli on every stimulus repetition but as multiple

induction steps on every encoding repetition. Such a reinterpretation would intensify the memory encoding protocols we have been attempting to detect and this may accelerate the convergence time of the temporal averaged TC distributions. Our results revealed that indeed such protocols reduce the convergence time by effectively reducing a filter's size through multi-step encoding at each repetition. However, the convergence still required a few tens of samples for a filter of $\Theta = 7$ even under four-step encoding protocol with $n_r = 4$ repetitions. If we increased the number of steps further there would still be an overlap between $\mathrm{TC}_{\mathrm{massed}}$ and $\mathrm{TC}_{\mathrm{spaced}}$ and therefore a synapse would require tens of samples before it can identify the shape of the sampled distribution.

Finally, we examined if we could replace the complex accumulation of TC distributions held by every synapse by a simpler mechanism which could probabilistically identify if a spaced protocol is being induced. By measuring the total occupancy in TC distributions above some $c_\Theta = c$ for $\mathrm{TC}_{\mathrm{massed}}$ and $\mathrm{TC}_{\mathrm{spaced}}$ we found that with increasing n_r massed protocols display relatively higher occupancies than spaced protocols when summing above low values c_Θ. The contrast between the two distributions increased with n_r while $\mathrm{TC}_{\mathrm{spaced}}$ distribution favoured larger threshold cycles c. However, the differences in the ratios of the fraction of synapses above some c_Θ under massed and spaced protocols of few repetitions ($n_r \leq 8$) were small and the relative gain of $\mathrm{TC}_{\mathrm{spaced}}$ over high c attenuated as the filter size increased. Therefore, a weak preference towards spaced protocols could be based on detecting when synapses exceed a particular number of threshold cycles c_Θ. Although measuring the distance between distributions improves our criterion as the number of trials is increased, we do not intend to apply repetition protocols multiple times and thus the probabilistic interpretation of a simple TC threshold may suffice.

In conclusion, contrary to experimental evidence showing that even a single application of a spaced stimulation protocol is sufficient to induce stable forms of plasticity, single filter synapses under a few applications of such a protocol do not obtain sufficient statistics to detect if a spaced protocol is being induced. Therefore, we are compelled to improve the statistical sample to which single synapses have access to in order to make a memory system selective to the stimulation protocol being used. In the next chapter we draw inspiration from the biological mechanisms of l-LTP we reviewed in Chapter 2 to build models that increase the information provided to single synapses on the protocol being induced by sampling the signal at each memory repetition.

Chapter 9

Global mechanisms of memory allocation

In the previous chapter we explored the effects that various stimulation protocols have on the mean memory signal and defined a *strong* stimulation protocol. This consists of a few regularly spaced memory repetitions under which synapses should respond by making a transition to stable forms of plasticity such as late-phase long-term plasticity. There, the question was raised whether an individual synaptic filter alone can detect when such a memory encoding protocol is being applied so as to initiate the processes that lead to late-phase plasticity. We found that from the point of view of isolated synaptic filters at single synapses the small number of regularities in the plasticity induction signals are not significant to reliably detect a single application of this memory encoding protocol.

In this chapter we propose that the distinct form of filter signals under different memory repetition protocols can be used to detect strong regularly spaced repetition patterns. We begin with a short discussion of our approach and then proceed to examine how spaced memory repetition intervals can be discriminated from massed. Subsequently, we postulate that a molecular signalling pathway driven by the filter's mean signal dynamics can implement a spatio-temporal filter on the profile of the memory signal that detects optimal storage conditions. A model of this mechanism is constructed in two steps and its ability to discriminate between massed and spaced protocols is examined in relation to selectively initiating memory allocation under strong spaced protocols.

9.1 Introduction

We anticipate that an efficient memory system should prevent the long-lasting storage of poorly relevant information. It is known from behavioural and cellular studies, which were reviewed in Chapter 8, that memory presentation protocols of a few repetitions

are sufficient to induce LTM and late-phase plasticity. If single filters are inadequate to quickly detect temporal patterns in memory encoding stimuli, how do neurons detect the conditions that lead to stable forms of plasticity in a synapse specific manner?

The axiom that cellular processes underlie the spacing effect pervades a reductionist view of memory. This has been confirmed in behavioural long-term learning of the simple Aplysia reflex pathway (Montarolo et al., 1986, Mauelshagen et al., 1998). In general, spaced patterns of stimulation could be relevant for long-term behaviour of an animal because they relate to some persistent feature of an environment. Therefore, taking notice of this temporal pattern may be important for more complex forms of learning and memory as well but it may be argued that the underlying mechanisms could be different. Taking a high level view of behavioural experiments on more complex organisms exhibiting the spacing effect we may postulate that perhaps a system level mechanism is in place to selectively respond when information is relevant for the long-term. Thus, one neural system could be responsible for detecting stimulation conditions and upon detection it then signals a second memory system to store a memory for long-term. Indeed, such organizational patterns in memory systems where discussed in Chapter 1 on the role of memory modulating systems and then these can be linked the essential role of neuromodulators in forming LTM and long-term plasticity we discussed in Chapter 2.

This study is based on the assumption premises that it is within the roles of the target memory system to integrate all signalling sources before it makes the final decision to allocate resources for the long-term storage of particular memories. Consequently, we view the actions of external systems through neuromodulators as necessary but not sufficient to instruct memory allocation. Therefore, the idea that the neural substrates of memory systems can independently detect the stimulation pattern is worth researching further and we proceed to explore possible mechanisms that would enable filter synapses to detect the induction protocol. Here, it is proposed that filter synapses use global neuron-wide information provided by the memory signal. The memory signal is reflected in the magnitude of neural depolarization. A mechanism that reads the memory signal may be tuned to detect its distinct profile under strong regularly spaced repetition patterns against the one obtained under massed protocols. We look for signal profile criteria that could discriminate between massed and spaced protocols and suggest a mechanism inspired by the cAMP pathway that could implement a filter tuned to spaced repetition patterns.

9.2 Memory signal as a readout of the encoding protocol

In this section we aim to examine if certain criteria can be set to detect the differences in mean signal $\mu(t)$ profiles exhibited between massed and spaced protocols. After locating

these criteria, a mechanism can be constructed to act as a filter that selectively initiates the transition to long-term memory under strong regularly spaced stimulation protocols.

Before we proceed, we need to discriminate between two events that follow any memory encoding event. The first event is the neural depolarisation in response to a memory $\boldsymbol{\lambda}$ being presented for synaptic encoding. We define a signal $\mu_\lambda(t)$ to be the memory signal exactly prior to encoding $\boldsymbol{\lambda}$ at time t_r, which can be written as $\mu_\lambda(t_r) = \mu_\lambda(t_r - \epsilon)$. We take the signal to reflects the neural depolarisation in response to input pattern $\boldsymbol{\lambda}$ being presented. Therefore we expect that upon presentation of a new unrecognised memory the postsynaptic response will be zero on average, while presenting a previously encoded memory will produce a postsynaptic response proportional to the signal stored synaptically and therefore at initial encoding $\mu_\lambda(t_0) = 0$. We are interested in studying how memory repetition at times $t_r \in 0, t_1, \cdots, t_{n_r}$ (with $0 \geq r \geq n_r$) of some pattern $\boldsymbol{\lambda}$ affects the memory signal $\mu_\lambda(t_r)$. The $\mu_\lambda(t_r)$ gives a sample of the signal as $\mu(t_r - \epsilon)$ just before any changes in filter states take place due to the repetition of pattern λ at time t_r.

Every memory repetition event would re-encode a memory. If the signal due to its prior encoding at t_{r-1} has not decayed by the time of repetition t_r the re-encoding should cause the signal to augment. Thus, the repetition protocol of a memory λ influences the sampled signal $\mu_\lambda(t_r)$ but also the future profile of the signal. Due to the influence of repetition at t_{r-1} on $\mu(t_r)$, the postsynaptic response at time of re-encoding t_r can be used as a reporter of past memory repetitions.

Naively, one would expect that against all other fixed interval repetition protocols the one with repetition intervals timed to the filter's peak signal t_p would obtain maximum signal samples $\mu_\lambda(t_r)$, where $t_r \in \{0, t_p, 2t_p, \cdots, n_r t_p\}$. Clearly, this true for a protocol with one ($n_r = 1$) repetition as the sample is obtained at the time of the filter's unique signal peak. However, as it has already been discussed, this is not the case with an increasing number of repetitions due to the relationship between the timing of memory re-encoding and the filter dynamics.

In the previous chapter we examined the effect of different memory repetition protocols on the mean signal. Typical filter signal responses are reproduced here on Figure 9.1 showing that massed repetition $T_r = 1$ increases the signal peak to a maximum when compared with other spacing intervals for the same number ($n_r = 8$) repetitions. Also, a regularly spaced protocol timed to the peak of a $\Theta = 7$ filter at $T_r = 18$ shows a precession in the timings of the apparent signal peaks as well as the signal exhibiting an asymptotic maximum value.

The size of the signal during the last repetition reveals the ability of an encoding protocol to augment the filter's signal. Figure 9.2(a) shows the sampled signal $\mu_\lambda(t_r)$ obtained against a fixed repetition interval (T_r) for protocols of various repetition lengths n_r on a $\Theta = 6$ filter. We find that the sampled signal $\mu_\lambda(t_r)$ can be enhanced by increasing

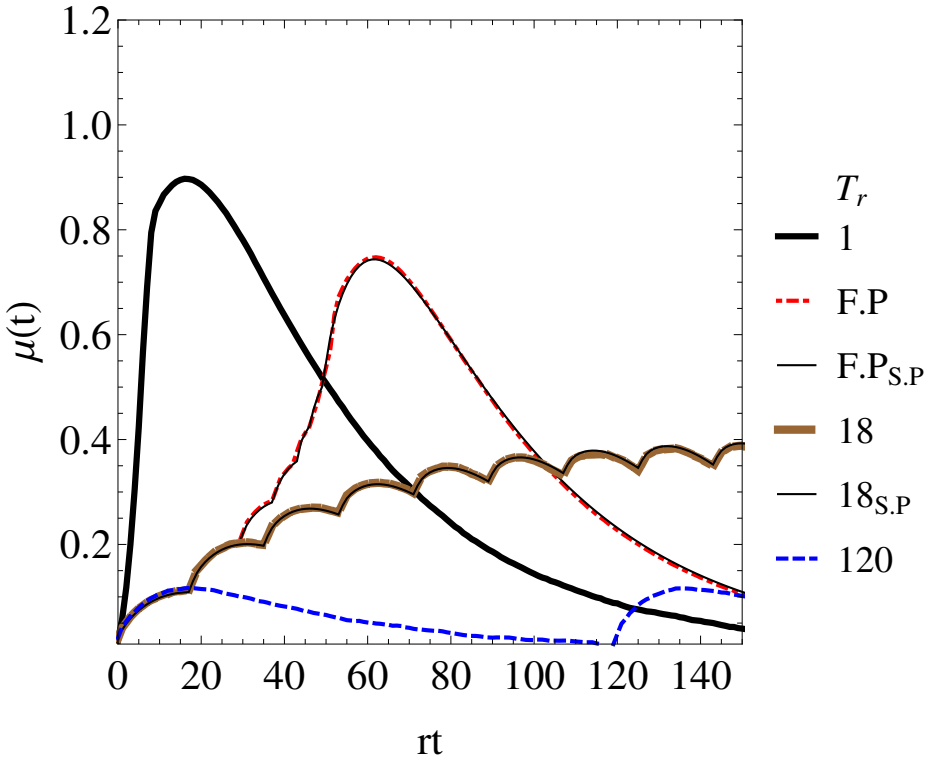
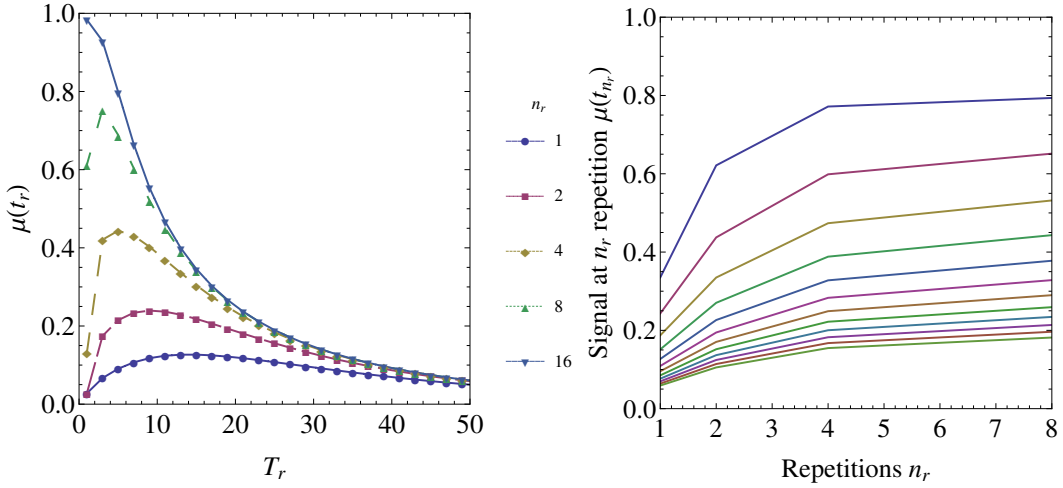


FIGURE 9.1: Filter mean memory signal under different repetition protocols. The effect T_r on the mean signal of a $\Theta = 7$ synaptic filter under $n_r = 8$. $T_r = 1$ shows the highest peak signal overall. Fixing the repetition interval to $T_r = 18$ shows that regular spaced repetition makes the maximum signal asymptote after approximately $n_r > 4$. The peak signal following each repetition moves towards earlier intervals under this protocol. Adding a filter signal made from superimposing the signal of a single filter on top of mean signal at every repetition ($T_r = 18$ S.P) reveals that this effect of the moving peaks is not due to a change in the filter internal dynamics but a results of signal superposition repetition. The $T_r = F.P$ signals result from a protocol timed to repeat on the apparent mean signal peak, which progressively get smaller with a ratio 3/2. Finally, changing the repetition interval to relatively long values $T_r = 120$ shows that the maximum peak signal does not rise and thus repetitions do not interact.

the number of repetitions n_r and by reducing the repetition interval T_r. Consequently, small T_r intervals (massed) become more effective at increasing the sampled signal as n_r increases. Conversely, increasing the number of repetitions n_r shortens the optimal repetition interval T_r that maximizes the sampled signal.

These results are simply a reflection of the fact that the signal is augmented under massed protocols by each memory repetition (see Figure 9.1). In contrast, under regular spaced repetition protocols with repetition intervals set to a filter's signal-peak time ($T_r = t_p = 0.375\Theta$) we find that signals asymptote. Plotting the size of the last sampled signal $\mu_\lambda(t_r)$ as the number of repetitions increase reveals this asymptotic behaviour. Figure 9.2(b) plots the last sample of the signal $\mu_\lambda(t_r)$ for filter sizes up to $\Theta = 15$ showing that all of these asymptote after a few repetitions. Small filters give higher signal samples and asymptote faster in n_r than larger filter sizes. Therefore, one characteristic

of regularly spaced on-peak protocols is that the signal sample obtained after a few repetitions becomes almost constant.



(a) $\Theta : 6$, $n_r : 1 - 16$ (b) $\Theta \in \{2, \cdots, 13\}$, $n_r : 1 - 8$

FIGURE 9.2: Comparing maximum signal obtained under different regularly spaced repetition protocols. **a**. Increasing the number of repetitions n_r shifts the optimal repetition interval towards shorter intervals. The time at which the signal sampling is maximized is shifted from the filter peak signal time to earlier repetition times. Here, the sampled signal from a filter of $\Theta = 6$ is shown across repetition intervals ranging from $n_r = 1$ (lower trace) where the maximum repetition time is at signal peak $t_r \approx 13$, up to $n_r = 16$ repetitions (top trace) where the sample is maximized if repetitions are delivered massed at $T_r = 2$. **b**. Mean signal value on last repetition $\mu(t_r)$ obtained under a $T_r^\star$ protocol, with n_r repetitions. Each curve from top to bottom represents a filter size from $\Theta \in \{2, \ldots, 13\}$ respectively. The signal obtained due to spaced repetition asymptotes as the number of repetitions increase while small filters give higher signal values.

However, here we have only examined protocols of regular fixed intervals and not F.P protocols with advancing interval timings aimed at the apparent signal peaks (see Figure 9.1). Figure 9.3(a) shows the variation of the sampled signal $\mu(t_r)$ as the length of the protocol n_r increases for a set of typical repetition intervals ranging from massed $T_r = 1$ to $T_r = 2t_p$ with the addition of the F.P protocol. Indeed, the F.P protocol appears to sample a higher signal at $t = t_r$ up to few repetitions. Beyond these, massed protocols exceed F.P as they result in significantly higher signal. This can also be seen by comparing the F.P and the massed protocol on Figure 9.1 on which the maximum peak signal results under a massed protocol for $n_r = 8$ and not an F.P protocol but also on Figure 9.2(a) where the optimal repetition interval switches to massed protocols as n_r is increased. Massed protocols of enough $n_r \geq (2\Theta - 1)$ result in deterministically pushing all synapses through the correct threshold recruiting all synapses to the encoding of the repeated memory $\boldsymbol{\lambda}$ vector. Therefore it is expected that these would provide the highest signal samples with sufficient n_r although for few n_r the samples are initially low

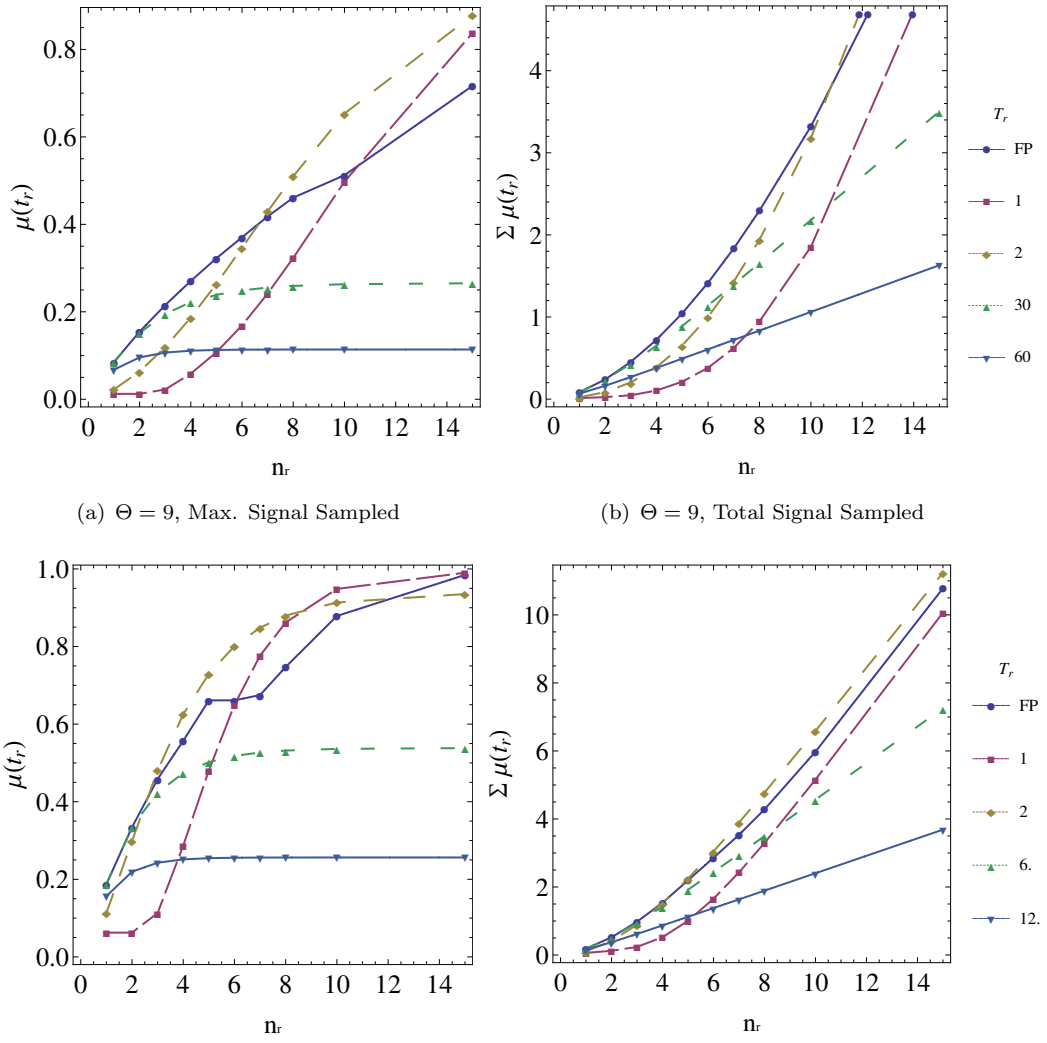
compared to spaced protocols as the former are obtained during the signal's rise phase while the later at the signals peak.

On Figures 9.3(b) and 9.3(d) we also observe that the F.P protocol integrated signal is close to optimal across n_r. This rather more complicated protocol may optimize the sampling but as the length of the protocols increases in n_r the intervals T_r become so small as to be considered massed repetitions. This would make the interpretation of such protocols rather difficult if we are taking the spacing effect to mean anything for memory systems at all. Thus, for simplicity we will be considering protocols of fixed interval spacings.

Overall, if we focus on the fixed interval protocols only and ignore the F.P protocol, which has variable intervals, we observe that the longer the repetition intervals T_r the quicker the signal asymptotes with increasing n_r. Decreasing the filter size from $\Theta = 9$ to $\Theta = 4$ shows no qualitative differences, see Figure 9.3(c). These results raise the possibility that strong regularly spaced on-peak protocols can be detected by specifically looking at the asymptotic signal behaviour seen under protocols with long repetition intervals.

Further, we may examine how the accumulated signal samples behave between protocols. Here, instead of looking at the value of the signal at the time of final repetition t_r, we examine the aggregate signal sampled over all t_i for $i \in \{1, \cdots, n_r\}$ as $S = \sum_{r=1}^{r=n_r} \mu(t_r)$. Figure 9.3(b) shows that for a $\Theta = 9$ filter the total signal S is greater under protocols with spaced intervals ($T_r = t_p$) than massed ($T_r = 1$) up to a few repetitions $n_r \leq 10$. This changes when $n_r > 10$ where massed repetition protocols give larger aggregate signal than spaced protocols. This behaviour is also found with smaller filters but the crossing point occurs sooner around $n_r = 8$, see Figure 9.3(d). The crossing point also moves towards less n_r if the massed protocol intervals are slightly increased to $T_r = 2$, but this is to be expected as intervals become more spaced. Nevertheless, for protocols of a few repetitions in length we may discriminate spaced on-peak protocols simply by integrating signal samples. Imposing a minimum signal constraint would make the signal integrator selective to spaced protocols as massed protocols require a few repetitions before the signal rises to the asymptotic level of spaced protocol. This is a promising result, but extending to protocols of more repetitions would require to filter temporal aspects of repetition intervals so to exclude short intervals from contributing to the integrated quantity. Thus, we would need to build spatial constraints in the signal range that ensure a minimum signal size is integrated combined with temporal constraints that enforce a minimum repetition interval between signal samples.

The general idea, is that the form a repetition protocols in terms of number of repetitions and the repetition intervals produce specific signal patterns, as shown previously on Figure 9.1. It would be possible to identify a particular pattern out of ones shown on the figure by setting certain criteria on signal size (spatial constraints) at particular times (temporal constraints). For example, identifying the $T_r = 18$ protocol on Figure



(a) $\Theta = 9$, Max. Signal Sampled

(b) $\Theta = 9$, Total Signal Sampled

(c) $\Theta = 4$, Max. Signal Sampled

(d) $\Theta = 4$, Total Signal Sampled

FIGURE 9.3: Plot of maximum sampled and total sampled signal per repetition count n_r for five sets of T_r values. The FP protocol has progressively smaller T_r designed to hit a the apparent peak of the filter signal, the other protocols consist of equal intervals T_r. Samples are the taken as the value of signal at the time t_r of signal repetition before any synaptic state changes occur due to re-encoding. The FP protocol appears near optimal across n_r for maximum and total signal. We see that filters naturally do not favour massed repetition at very short intervals $T_r = 1$ for low repetition counts $n_r < 4$, however increasing the number of repetitions shows that massed repetition provides an overall stronger input. For the $\Theta = 9$ the total signal under $T_r = 1$ meets on-peak repetition at $n_r = 11$ while for the smaller filter $\Theta = 4$ this point requires less repetitions ($n_r = 8$).

9.1, would require to set a constraint that the signal sampled at each repetition after $n_r > 5$ sits within range of 0.3 to 0.5. In this manner we would have constructed a spatio-temporal filter which is specifically looking for an asymptotic signal between a range of values, or in simply above some minimum value. Such spatio-temporal filtering of signal samples may actually be easily implemented biophysically and we explore this potential over the next sections.

9.3 Reading signals of neural excitation

A series of studies have shown that increased excitability of neurons can predict the subset of neurons that will be allocated in the encoding of long-lasting fear memories (see Benito and Barco, 2010, Zhou et al., 2009, Josselyn, 2010). Such evidence hint that postsynaptic depolarisation could be used as an indicator of suitability for neurons to allocate a particular memory. Neural excitability has been shown to be regulated by the activation levels of CREB that affects the balance of Na^+ and K^+ ion channels. In the amygdala neurons in particular, Zhou et al. (2009) suggest that fear memory allocation is based on competitive mechanism operating through neural excitability. Their hypothesis is that the allocation mechanism may be bidirectional and a CREB repressor may be expressed to drive away further memory allocation from the subset of neurons that have recently allocated a memory.

Therefore, these studies suggest that depolarisation triggers the processes that lead to long-term plasticity and memory allocation. In this chapter we will use neural depolarisation as a signal that provides a spatial average of synaptic strength. This signal will then be used as part of a molecular mechanism to provide an output variable indicating the level of regularities in the memory encoding stimuli.

At a first level, the magnitude of the depolarisation maybe used to indicate that a particular memory has been previously encoded. However, our examination of the signal readout $\mu(t_r)$ in the previous Section showed that it is possible to develop a mechanism that optimally responds under strong spaced protocols. The biophysical nature of this postsynaptic depolarisation signal could be in the form of a change in the concentration of an ion occurring within a dendritic compartment or the whole neuron. For example, during memory repetition the synaptic currents can cause an increase in an intracellular ion-flux concentration that is proportional to the postsynaptic response $\mu(t_r)$. It is known that neural depolarisation biophysically relates to transient increases in the concentration of calcium Ca^{2+} (see Ghosh and Greenberg, 1995) while these may be initiated neuron-wide via back propagating action potentials (BAPs) (Paulsen and Sejnowski, 2000). The entry of Ca^{2+} acts also as second-messenger to initiate downstream processes through its interaction with kinases (see Chin and Means, 2000) that relate to synaptic plasticity (see Xia and Storm, 2005).

We may model a stereotypical activation of such a kinase in response to increases in the concentration of Ca^{2+} by initially defining the transient concentration as a brief signal denoted by $c(t)$ that is proportional to a memory signal sample $\mu(t_r)$:

$$c(t) = \mu(t_i)\delta(t - t_i) \tag{9.1}$$

obtained using the Dirac delta function $\delta(t)$. Under some particular repetition protocol we let $c(t)$ be a piecewise function defined at each repetition time $t_r \in \{t_1, \cdots, t_r\}$ of n_r memory repetitions and with $c(t) = 0$ if $t < t_1$.

Standard functions can used to model concentration dependent activation of downstream kinases. A typical function that models the activation of kinases via cooperative binding is given by the Hill equation (Hill, 1913) :

$$h(t) = \frac{c(t)^n}{c(t)^n + K^n} \tag{9.2}$$

For a review of how this expression is derived and a discussion on a biophysical implementation of this threshold mechanism see Appendix 3. In essence, this a threshold function giving a half maximal response when the signal sample $c(t)$ reaches K. The level of K can be used to set a lower bound on the size of signal samples that will be processed downstream of $h(t)$. This function essentially implements a minimum repetition requirement before the signal reaches K. The parameter n sets the sharpness of this threshold function, Figure 9.4 shows responses of the Hill threshold function for a range of n values from smooth $n = 1$ to sharp $n = 9$. In summary, the magnitude of the

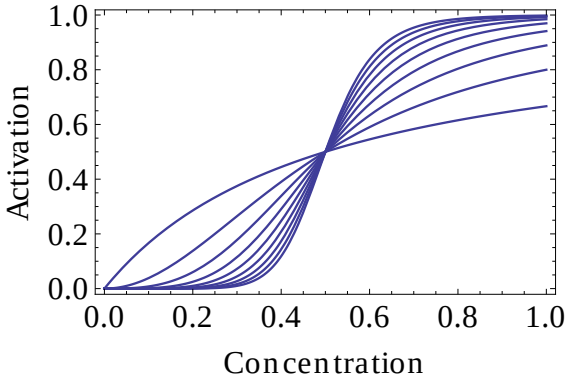


FIGURE 9.4: A Hill (Hill, 1913) threshold activation function for orders $n = 1$ (smooth transition) to $n = 9$ (sharp rise), in response to a hypothetical input of some compound concentration. The threshold has been set half way $K = 0.5$ and we find that for all orders n a half maximal response occurs when the input is equal to threshold. However, the slope of the curves around that half-maximum point increase with the order n. For $n = 9$, there is a sharp transition on the output from almost 0 to maximum as the input moves from 0.35 to 0.7.

$h(t)$ response is determined by $c(t)$ which gives a sample of $\mu(t)$ only at repetition times $t_r \in \{1, \cdots n_r\}$. The usual assumption holds that during memory repetition at t_r the sampled $\mu(t_r)$ is the value obtained before any filter state is modified and therefore any plasticity event takes place. Consequently, to obtain a non-zero $h(t)$, a memory needs to have been previously encoded at least once and thus at initial encoding $h(t_0) = 0$ while under a single memory repetition at t_1 the maximum mean signal sampling would occur if the repetition time t_1 is set at the time of a filter's peak signal $t_p = 0.375 \times \Theta^2$.

Following our suggestion that signal dynamics can be used to determine massed from spaced protocols here we have developed a model of how neural depolarization signals could be read. It remains however to develop the downstream mechanisms that will process these such that spaced protocols selectively initiate late-phase plasticity.

9.4 Towards a biophysical mechanism of signal profile detection

A separate filter may, in principle, be built to detect the characteristic features of the memory signal samples arising under strong spaced stimulation protocols. According to the results of Section 9.2 such protocols quickly reach an asymptotic signal value while the total signal sampled exceeds other protocols within a few repetitions. Yet to our knowledge a systematic search of the repetition number versus repetition interval trade-off (n_r/T_r) for the magnitude of plasticity is not available in the experimental literature. We will come back to this issue in the Discussion chapter, but here this behaviour is taken as a prediction of the filter-synapse model.

The working hypothesis is that the depolarization signals arising at the level of larger compartments stimulate molecular signalling pathways that implement a filter on the signal that is tuned to optimally respond to strong repetition protocols. These compartments could be a section of a dendrite or a whole dendritic branch containing hundreds or thousands of synapses. For simplicity, we will work on an abstract single large compartment that contains all synapses of a neuron or in any case is able to read the depolarization of the whole neuron. The spatio-temporal filters we propose are built using abstract biophysically plausible mechanisms that provide a compartment-wide signal indicating that a particular induction protocol is being used. According to our model, single-synapses combine this compartment-wide signal with synapse-specific filter statistics to decide on a transition to stable late-phase plasticity.

To communicate these ideas we shall take a systems' biology approach to *qualitatively* model how compartments could detect strong stimulation protocols. As it has already been discussed in Chapter 2, signal transmission in biological systems occurs mostly through enzymatic reactions of phosphorylation and protein-protein interactions, but also through controlled production of intracellular molecules that act as downstream messengers in a concentration depended manner.

Various mechanisms could be devised that specifically look for the asymptotic signal behaviour that is characteristic of strong spaced protocols. For example, we could imagine that measuring the rate of change in the sampled signals could offer a mechanism that broadly detects spaced protocols by checking if this rate gets close to zero within a few repetitions. Thus, a filter would have to be built to detect that the sampled depolarizations are above some minimum value and that this value becomes almost constant after

few repetitions. Alternatively, a filter could be built that only integrates signal samples within a range w above and below the asymptotic value for a given filter size Θ. In this manner, the temporal aspects of repetition protocols would be filtered by spatial constraints imposed on the signal because short intervals would raise the sampled signals above the allowed region while longer intervals would make the signal samples fall below. Strictly speaking, such a mechanism would not enforce a spacing between all signal samples obtained under repetition but it would require pauses between some stimuli so the signal drops within the acceptable range before any further repetitions occur. With these sampling constraints in place, a minimum repetition criterion for late-phase plasticity can be simply added by setting a threshold P on the recently integrated signal samples R.

Given that there are various possible mechanism we could chose to develop we have adopted an approach that takes into account known cellular signalling processes. The mechanism we propose is inspired by the cAMP-pathway that can be jointly stimulated by the synergistic action of converging neurotransmitter and neuromodulator signals on the AC enzyme (Tomchik and Davis, 2009, Lin et al., 2010, Abrams et al., 1991). Particular forms of this enzyme have been shown to be critical for long-term memory (Wong et al., 1999, Livingstone et al., 1984, Shan et al., 2008, Zhang et al., 2008, Wu et al., 1995); these initiate cAMP production in response to joint stimulation from neuromodulators, representing reinforcement signals, and sufficient Ca^{2+} concentration. As discussed in the previous Section, the Ca^{2+} concentration reflects the level of postsynaptic depolarization which in the cAMP system is sensed by co-bound calmodulin kinase (CaM) on ACs (see Figure 9.5). Therefore, the size of the postsynaptic depolarization during stimulation would influence cAMP production by these AC enzymes and consequently also influence the activity of downstream processes. A well known downstream process is the activation of the PKA family of kinases. These are known for their involvement in activating CREB transcriptional processes necessary for the initiation and maintenance of late-phase long-term memory (see Chapter 2) and are selectively recruited in response to spaced tetanization (see Woo et al., 2003). The level of PKA activation is therefore directly affected by the overall cAMP concentration and it would appear that PKA integrates the cAMP signals.

PKA is believed to initiate the protein synthesis and transcription required for long-term plasticity. Its concentration level, denoted by $R^\star$, will be taken as the biophysical representation of our mechanism's output variable. We will assume that some synaptic growth process is unlocked when the output variable $R^\star$ exceeds a critical level P. Reaching a sufficient level would allow to allocate a memory through transitions to late-phase synaptic plasticity specifically at the synapses that encode this memory through the utilization of synapse specific information.

Previously it was shown that the total sampled signal is maximized under on-peak spaced protocols if these are confined to a few repetitions. In the next Section we describe

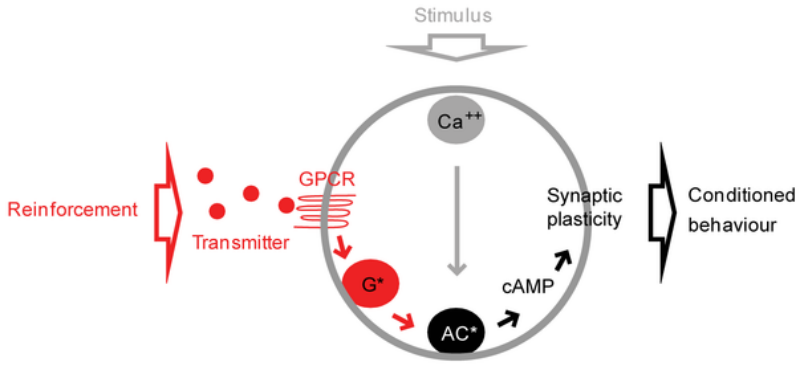


FIGURE 9.5: Schematic representation of excitatory stimuli converging with reward neuromodulatory stimuli on adenylate cyclase (AC). The stimulus results in raising the intracellular Ca^{+2} concentration which interacts with AC via calmodulin (CAM not shown). The reward signal carried by a neuromodulator binds to a G-protein coupled receptor (GPCR) on the same neuron and activates G-protein which in-turn also activates AC. This signal convergence system allows coincidence detection between excitation stimuli and reward signalling and can be found in various memory systems. Appropriately timed delivery of the two signals can synergistically activate cAMP production of AC leading to synaptic plasticity. Source:(Yarali et al., 2012)

a biophysically plausible model that relates to the cAMP-pathway in order to show how signal samples could be integrated. According to our above results, for a limited number of repetitions, cAMP production should be optimal if repetition intervals are timed accordingly and therefore spaced protocols should also maximize the $R^{\star}$ output.

9.5 An allocation model of direct kinase activation

9.5.1 Integrating signal samples

This model will rely on the fact that the total signal sampled under spaced protocols exceeds massed protocols within a limited number of repetitions, as previously shown on Figures 9.3(b) and 9.3(d). We aim to construct a global signal process that initiates late-phase plasticity under spaced protocols by using the fact that spaced protocols optimize the signal sampled within some n_r limits.

The signal samples are provided by the non-linear kinase activation $h(t)$ of equation (9.2). Our earlier results showed that the first few signal samples under massed repetition protocols are inferior to spaced protocols. The small samples of massed protocols can be suppressed by adding a threshold function to provide a suitable window above which samples are integrated. This window simultaneously defines a minimum repetition requirement before the signal exceeds the threshold for integration. The level of the K threshold sets the point above which the AC→cAMP system is engaged. A high enough K would impose a requirement for minimum number of repetitions to occur before depolarisation is sufficient to adequately initiate cAMP production. This point cannot be

arbitrary selected however. K should be set in the region below the asymptotic signal value obtained under spaced protocols for a filter Θ, otherwise the signal samples of spaced protocols would be suppressed.

The threshold response to depolarization provided by $h(t)$ could represent the stimulation of effector enzymes such as AC by CaM kinases. Accordingly, these ACs respond to the Ca^{2+} concentration under the simultaneous stimulation by G-proteins activated by neuromodulators. The neuromodulating signals are here assumed to occur simultaneously with Ca^{2+} and will not be explicitly modelled. Stimulation of these effector enzymes results in the production of second messenger molecules, such as cAMP, that are in turn responsible for propagating the stimuli downstream.

The output of this system is the activation level of a kinase, which in our case may represent PKA and we denote it by R. We assume that reaching a minimum level of activated R is sufficient to enable the processes of late-phase plasticity and thus R will act as surrogate for all molecular processes required downstream so as to initiate structural plasticity within a compartment. Activated states are denoted with a star superscript, and thus here the output variable representing the level of the activated PKA is denoted by $R^\star$. Forming long-term memory in our model requires a compartment-wide global signal that is activated when the output kinase $R \rightarrow R^\star$ has been activated to a sufficient level P. It can be envisaged that the ease or strength of the repetition protocol required to trigger memory allocation can be modulated by changing the minimum level P of $R^\star$ required. In the extreme case when the threshold is set to baseline (assume zero) then the synapses that become tagged even after a single memory encoding episode would make a transition to late-phase plasticity.

The input $c(t)$ providing signal samples to $h(t)$ is a piecewise function defined at repetition times t_r and therefore $h(t)$ is a also a piecewise function. In this initial model, $h(t)$ directly stimulates the production of the cAMP signalling molecule u through the activation of AC enzymes which synthesize it:

$$\frac{du(t)}{dt} = h(t) - F_u u(t). \tag{9.3}$$

Thus, cAMP is produced by AC and its concentration u naturally decays with rate F_u. However, as mentioned in the previous section (see Figure 9.5), the production of cAMP by particular types of AC appears to be synergistically modulated by a neuromodulator $D(t)$ and a calcium signal $c(t)$, something that our model above does not describe. As a first approach to our model we assume the the $D(t)$ signal is always delivered with repetition of the particular memory we are examining. These dynamics effectively define a low-pass filter of the input $h(t)$ with a timescale set by $1/F_u$:

$$u(t) = \frac{1}{F_u} \int_0^t \exp(-(t - t')/F_u) h(t') dt'. \tag{9.4}$$

The signal carried by the concentration $u(t)$ is transferred by stimulating the activation of the effector kinase $R^\star$. Assuming the decay of $R^\star$ is too slow for the timescales of the protocols used here we may define $R^\star$ as a non-decaying linear integrator :

$$R^\star(t) = \int_0^t u(t)dt. \tag{9.5}$$

Therefore, the dynamics of $u(t)$ are governed by the signal threshold K embedded in $h(t)$ but also by the natural rate of decay term F_u. Essentially F_u defines the filter's memory that allows us to set the time window for interaction between repetition stimuli. Slow decay dynamics would allow distant stimuli to augment u before it has decayed to zero, while a fast decay (ex. $F_u = 1/2$) would quickly forget any recent contribution to u by the last repetition which stimulated $h(t)$.

Next, we examine if conditions for the induction of long-term memory can be found that are specifically triggered under *strong* protocols consisting of at least four repetitions with regular pause intervals. To detect this protocol we set the threshold K of equation (9.2) to the value of $\mu(t_r)$ obtained after four spaced repetitions $n_r = 4$ and make the Hill threshold function $h(t)$ moderately sharp by setting $n = 4$ (see Figure 9.4). Furthermore, we initially choose F_u so the decay is slow enough to augment $u(t)$ when repeating a memory at any time during which the signal is above zero assuming a $\Theta = 6$ filter. For the typical filter sizes examined here their mean signal lifetimes would range from $100 < rt_{max} < 400$.

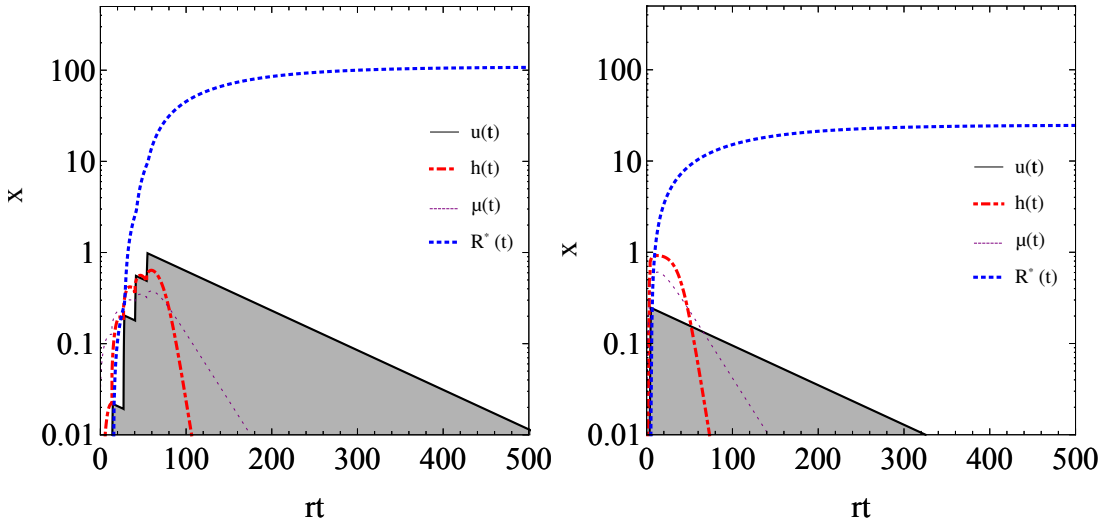
Relating the memory-signal evolution to real-time will require to fix some mean rate of memory encoding events r, since the evolution of the memory signal is dependent only on further encoding events in the palimpsest memories framework we are working in. We wish to choose a decay rate slow enough that would ensure that any repetitions occurring while the $\mu(t)$ signal is above zero would interact on the concentration of u. If the biophysical nature of u reflects a molecular concentration (ex. cAMP) then there may be limitations in the range of values we can set for F_u.

For the purposes of examining the function of this model, we assume $r = 1$ and then choose a decay rate of $F_u = 1/100$ by taking into account that after three time constants have passed an exponentially decaying component will drop by more than 95%. This choice of relatively slow decay and fast memory encoding rate would ensure that any repetitions occurring while the $\mu(t)$ signal is above zero would interact on the concentration of u if they occur within $rt < 300$. As we will see later, the decay parameter in this model only scales the output but we retain this parameter choice so we as to compare against later models.

The mean dynamics of each of the components in the above system are shown on Figure 9.6 for a $\Theta = 6$ filter under massed and spaced repetition protocols using $n_r = 4$ memory

repetitions. The output $R^\star$ from each protocol is also shown and it may be compared between the example massed and spaced protocol.

The spaced $n_r = 4$ protocol here achieves higher kinase $R^\star$ activation thus exhibiting a spacing effect; the massed protocol $T_r = 1$ gives $R^\star \approx 25$ while spaced on-peak protocol $T_r = t_p \approx 13$ gives $R^\star \approx 100$. In contrast however, comparing these to results shown on Figure 9.7 we find that the advantage of spaced intervals goes away when the number of repetitions is increased to $n_r = 8$ for this filter size ($\Theta = 6$). These outcomes are not surprising and are in accordance with earlier results on the total signal sampled shown on Figure 9.3(b). On the same figure we may observe that the total signal sampled is maximized under spaced protocols confined to a few repetitions while massed protocols became as effective beyond $n_r = 7$ repetitions.

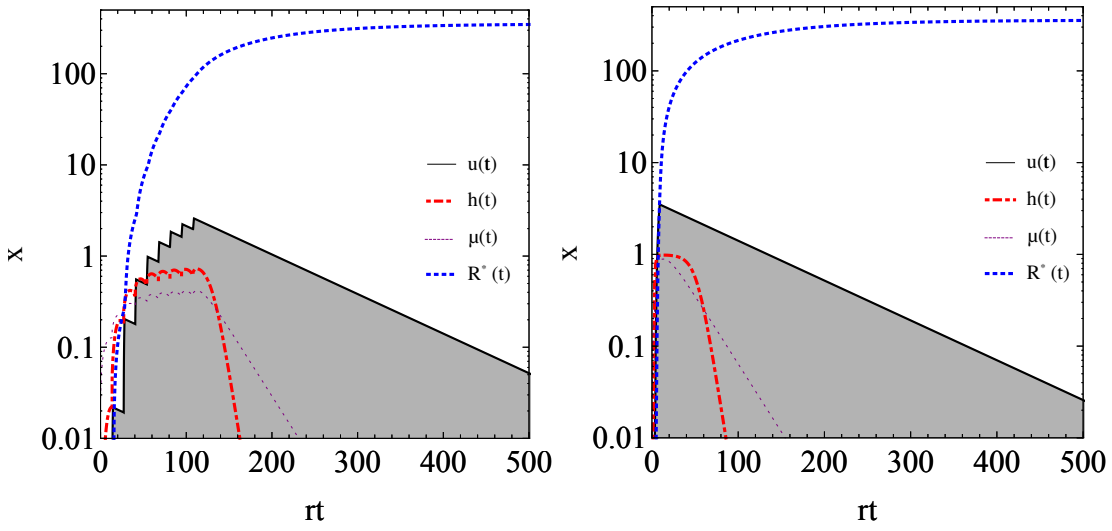


(a) $n_r = 4$, $T_r = 13.5$ Spaced (b) $n_r : 4$, $T_r = 1$ Massed

FIGURE 9.6: Dynamics of signal in a model of direct kinase activation by stimulating input produced second messenger u. The thick dotted line represents the overall activation of output kinase $R^\star(t)$ which is the integral of the shaded area representing $u(t)$. $h(t)$ shown here as a red thick dot-dashed line is driving the production of second messenger $u(t)$. The signal $h(t)$ implements a threshold function on $\mu(t)$, and thus it roughly follows the profile of the thin dotted line of $\mu(t)$. For this low number of repetitions, a spaced protocol produces higher $R^\star$ than a massed protocol.

A more detailed examination of the relationship between the output $R^\star$ against the interval T_r for protocols of $n_r = \{2, 4, 8\}$ is shown on the left column of Figure 9.8.

The intervals T_r are relative to the on-peak timing t_p for each filter size, thus $T_r = 0$ indicates a spacing interval timed to filter's peak signal. The $R^\star$ level is normalized across Θ because the activation threshold K in $h(t)$ is set at the asymptotic signal obtain for each Θ accordingly. We find that short protocols of $n_r = 2$ and $n_r = 4$ the activation of $R^\star$ against intervals exhibits a curve tuned to the on-peak interval protocols. We shall denote on-peak interval protocols as $T_r^\star$ from now on. The shape of the curve is



(a) $n_r : 8$, $T_r = 13.5$ Spaced (b) $n_r = 8$, $T_r = 1$ Massed

FIGURE 9.7: Dynamics of signal in a model of direct kinase activation by stimulating input produced second messenger u. The thick dotted line represents the overall activation of output kinase $R^\star(t)$ which is the integral of the shaded area representing $u(t)$. $h(t)$ shown here as a red thick dot-dashed line is driving the production of second messenger $u(t)$. The signal $h(t)$ implements a threshold function on $\mu(t)$, and thus it roughly follows the profile of the thin dotted line of $\mu(t)$. Here the total levels of $R^\star$ are almost equal between spaced and massed protocol.

however distorted as the number of repetitions increases. On Figure 9.8(e) we find that the tuning of $R^\star$ to $T_r^\star$ is nearly lost under $n_r = 8$ repetitions. Increasing the decay to $F_u = 1/2$ does not change the results qualitatively, it only scales $R^\star$ downwards since the integrated area under u is now reduced, see Figure 9.9. Nevertheless, it becomes evident that for $n_r \leq 4$ it is possible to identify spaced protocols by via the setting of a high enough threshold on $R^\star$ such that it is reached by spaced protocols only.

9.5.2 Triggering late-phase plasticity for allocation

Here we construct a simple allocation switch at the level of a compartment. This compartmental allocation switch is turned on once the $R^\star \geq P$, where the threshold P is specifically positioned so it is reached under $n_r = 4$ spaced interval protocols. The threshold level under slow decay dynamics $F_u = 1/100$ can be obtained by examining Figure 9.8(c) on which we find that maximum $R^\star \approx 100$ and thus we let the allocation threshold be $P = 100$.

Once the allocation switch turns on, a global signal is received by all synapses that a particular stimulus conditions have been met. Each synapse would need to integrate the global signal with synapse specific information so as to model memory allocation via synapse-specific transitions to long-term plasticity. Thus, we need to describe which synapses become "locked" once the allocation switch is on.

In the previous chapter we examined the theoretical ability of single synapses to identify the protocol being induced from an accumulated distribution of threshold-cycles (TC). The process of counting the size of a TC is assumed to initiate after the initial encoding of memory under the instruction of neuromodulatory signals. Assuming that one TC sample is obtained each time a repetition protocol is applied showed that single synapses could not safely determine which protocol is being induced within a few TC samples. The sampled distribution distance to TC distributions of spaced and massed protocols was not well separated unless tens or hundreds of samples were obtained. Further, given the overlap between the $\mathrm{TC}_{\mathrm{massed}}$ and $\mathrm{TC}_{\mathrm{spaced}}$ empirical distributions it is not possible to find a single point on the distribution that separate the two cases. We did find however, that TC values above some c are relatively more frequent under spaced rather than massed protocols. The converse is also true, massed protocols show a relatively higher occupancy if summing over small c values. For example, under $n_r = 4$ repetitions a $\Theta = 7$ synapse that has conducted $c \geq 1$ TC is almost 20% more likely experiencing a massed protocol rather than spaced but for $c \geq 5$ it is 10% more likely its experiencing a spaced protocol. Arguably, these criteria are deficient to operate as reliable synapse-local information on the protocol being induced. Alternatives could be to look for variability in threshold crossing behaviour, as $\mathrm{TC}_{\mathrm{massed}}$ distributions show reduced negative values in comparison to $\mathrm{TC}_{\mathrm{spaced}}$. Spaced protocols display a small fraction of synapses performing negative TCs. This because synapses may perform opposite threshold crossings under stimuli that arrive during repetition intervals. Such criteria however would only work for long $n_r \geq \Theta$ protocols relative to the filter size (see Section 8.2.2) or otherwise under the multi-step protocols we examined in Section 8.3.2. The $\mathrm{TC}_{\mathrm{spaced}}$ distribution from these protocols showed that a fraction of synapses had negative TC cycles thus showing variability in threshold crossing behaviour ($< 20\%$ for $\Theta = 4$ $n_r = \{4, 8\}$ 2-step and 4-step protocols). We have been unable to find any alternatives that could offer some simple means of obtaining more reliable local-signals. Overall, we could say that one important information contained in TCs is the direction of threshold crossings. Positive TCs indicate that synapses have crossed over the "correct" threshold for the memory being strongly encoded and consequently such synapses take part in the encoding of this memory.

Therefore, with TC as our local criteria, we create a system that begins counting correct/wrong same-threshold-crossings in response to a neuromodulatory signal given at the time of initial memory encoding. This may represent some attentional motivation which occurs every time the tracked memory is being repeated and is integrated by the global mechanism described above. When the global $R^\star \geq P$ allocation signal signal is set, the system stops counting same-threshold crossings and obtains the size of TCs. At this point the TC information of each synapse is evaluated and may result in converting to late-phase plasticity depending on the local criteria we have adopted. Initially the simplest criteria we can think of is a threshold c ensuring a synapse has conducted a positive TC. This method may ignore local information on the protocol being induced

and somewhat favour massed repetitions (see Section 8.4), but we can use it here to solely evaluate the potential of the global mechanism to induce late-phase plasticity and memory allocation in response to strong spaced protocols.

Whatever this local criteria may be, they actually operate to "tag" the synapses for late-phase plasticity. Using the positive TC as a local criterion means we are simply tagging synapses that encode the desired memory. Once the allocation switch is on, all synapses that have been tagged in the last memory encoding episode have their states locked. This mechanism is analogous to the capture process of the STC hypothesis. However, in contrast to STC the allocation is confined to the tagged synapses at the time of memory encoding alone, there is no late-associativity. We could allow late-associativity if we assigned a lifetime to the global allocation signal during which we would allow future encoding events to add tagged synapses. However for simplicity, we choose to only allocate tagged synapses based on the state of synapses at the time when global allocation is triggered.

Using this memory allocation mechanism we examine the variation in the allocated signal as a function of the number of repetitions n_r and the interval T_r. Figure 9.8 shows the total $R^\star$ and the allocated SNR for filters $\Theta \in \{7, 9, 11, 13\}$ under a range of repetition intervals $T_r \in \{1 \cdots 200\}$ centred around filter peak time by plotting $T_r - t_p$ on the x axis. The allocation threshold P was set near the peak of the $R^\star$ curve under n_r spaced repetitions, $P = 100$.

We find that for low repetition numbers $n_r = 2$ there is insufficient R^* to cross P and thus there is no allocated signal. Under four repetitions we find that $R^\star > P$ and the allocated SNR displays a repetition interval preference as required. The maximum allocated signal is not exactly centred around the zero point however and there is a precession even under four repetitions which can also been seen in the $R^\star$ curves. Nevertheless, there is a clear spacing requirement for all Θ examined. Notwithstanding an increase in the number of repetitions n_r, this spacing effect is abolished when $n_r = 8$. Under $n_r = 8$ massed protocols we get $R^\star > P$ and thus allocation is switched on. We observe that higher allocated signals are obtained under massed protocols because by increasing n_r more synapses are driven through the correct threshold for the encoded memory which is also reflected by the higher peak $\mu(t)$ (Figure 9.2(a)) and thus there are more tagged synapses.

For comparison we also check the effect of changing F_u to faster decay on Figure 9.9. These plots confirm the earlier findings that a spacing effect in R^star is only retained under low repetition counts as $n_r \leq 4$. The allocation threshold we have chosen for $F_u = 1/2$ is equal to the mean peak $R^\star$ activation level, just above $P = 2$. In this case also, massed $n_r = 8$ exceeds P while the allocated signal SNR is also stronger than spaced protocols. Although the $R^\star$ activation mechanism controls the conditions for allocation, under $n_r = 8$ the $R^\star$ threshold is exceeded even under massed protocols. In

this case, due to our choice for the local-allocation criteria ($c = 1$), massed protocols result in a higher number of tagged synapses and thus higher allocated signals as shown on Figures 9.8(f) and 9.9(f).

As discussed previously, we may use the limited local information offered by a TC criterion to partially occlude massed protocols from tagging synapses. Increasing the local TC threshold criterion from $c = 1$ to $c = 4$ would remove the advantage of massed protocols over spaced (see Section 8.4). Figure 9.10 compares the allocated SNR between a $c = 1$ and $c = 4$ under four and eight repetitions. The results shown demonstrate that an increase in the local TC threshold rule to $c = 4$ does indeed suppress the advantage of massed protocols when these are strong enough to activate the global $R^\star$ signal. Therefore. improving this local rule to further suppress massed protocols would result in a consistent spacing effect even under longer n_r where massed protocols are strong enough to activate $R^\star$.

Overall, the above results demonstrate that our simple signal integrating mechanism can provide a global memory allocation signal which is selectively initiated under strong spaced encoding protocols. However, the well known phenomenon of a spacing requirement for triggering memory allocation is not honoured across n_r. In our model, memory allocation is compartmentally initiated when $R^\star$ exceeds a threshold P. At the same time, setting P to a level that can be exceeded under a spaced protocol does not occlude the possibility that a stronger massed protocol with more repetitions n_r to exceed it too.

It may well be that the spacing effect is only retained under a limited number of repetitions ($n_r = 4$) beyond which massed protocols can also initiate late-phase plasticity and allocation. To our knowledge there are no experimental data that could provide a definite answer. A major part of parameter space on differences between massed and spaced protocols under various numbers of repetitions n_r remains unexplored (Kornmeier and Sosic-Vasic, 2012). Further, behavioural experiments on memory repetition spacing phenomena may be confounded due to the requirement for the activation of attentional and motivational centres. These may habituate after a few repetitions of a stimulus and stop providing the required neuromodulatory signals. Thus protocols of long sequences of repetitions, perhaps even beyond $n_r = 4$, may be irrelevant to a memory system because other brain systems or phenomena occlude the stimulation patterns from ever reaching the memory system in question.

On the other hand, long massed protocols may indeed be able to initiate late-phase plasticity but some mechanism may be in place to ensure that spaced protocols are always more efficient in activating the global allocation signal. We discuss a possible simple extension of this model to account for such phenomena next.

9.6 A saturating model of kinase activation

In this section we extend the previous model of allocation to add a limit on the maximum $u(t)$. By adding a saturation limit on $u(t)$ we may impose temporal constraints on the integrated signal samples $h(t)$. The extension proposed here adds a mechanism operating to limit the maximum activation concentration of the second messenger u. This limit effectively makes the u a saturating function. Such functions can seen as operating in two characteristic modes, a linear and a sub-linear addition mode. When the function is away from saturation it can increase in an approximately linear fashion, while when close to a saturation it operates in a sub-linear fashion. Thus, in terms of the biophysical interpretation of $u(t)$ as cAMP, saturation would result in sub-linear integration of samples obtained over short repetition intervals because u will be operating near the saturation limit. The timescale is defined by the decay timeconstant of cAMP, F_u. If sufficiently long intervals are used, such that cAMP decays between repetitions, new signal samples would be linearly integrated by cAMP. The result of a saturation on

$u(t)$ would therefore be to limit the overall $R^\star$ (PKA) when cAMP stimulation occurs before the time defined by the decay rate F_u.

In terms of biological mechanism the saturation limit could be imposed by a negative feedback suppressing the synthesis of u, or by a mechanism directly degrading u messenger molecules. Saturation may also be implemented passively by assuming that some component in the signalling pathway is depleted and we have to wait for its recovery. We do not explicitly model any particular mode of saturation but simply assume that u_{max} is the maximum amount of $u(t)$ within a compartment. The previous model may be extended by rewriting equation (9.3) to include an upper bound as :

$$\frac{du}{dt} = (u_{max} - u(t))h(t) - F_u u(t) \tag{9.6}$$

and retain the activation of output kinase as :

$$R^\star(t) = \int_0^t u(t)dt. \tag{9.7}$$

We choose $u_{max} = 1$ so $u(t)$ saturates within a few repetitions. The dynamics of $u(t)$ and the form input is shown on Figure 9.11 for massed and spaced protocols. Comparing Figure 9.11 to the unsaturated model on Figure 9.6 we may observe the constraint imposed by $u_{max} = 1$.

The effects of the saturation mechanism can also be seen by comparing the output kinase $R^\star$ activation curves per repetition interval of Figure 9.13 with the $R^\star$ curves of the previous unsaturated model under $F_u = 1/100$.

These results demonstrate that saturation does indeed relatively suppress the ability of massed repetition intervals to activate R in comparison to the unsaturated activation model (see Figure 9.8). In the previous model under $n_r = 8$ or higher the optimal repetition times for R activation shift towards earlier repetition intervals T_r. Here, we find that the addition of saturation retains the curvature of $R^\star$ plots although with a lesser relative gain between massed and spaced protocols; from approximately tenfold in $n_r = 4$ to just twofold under $n_r = 8$.

However, as in the previous model the saturation model also suffers from unconstrained growth of R activation under increasing n_r. For example, for an allocation threshold at $P = 80$ we need can initiate allocation under $n_r = 4$ spaced repetitions at filter peak times only, see Figure 9.13(c). Increasing to $n_r = 8$ however, allows a massed protocol to reach this allocation threshold too, see Figure 9.13(e). Consequently, any allocation threshold set on $R^\star$ based on the values reached under a spaced protocol for some value of n_r can be exceeded by a massed protocol of more n_r repetitions.

The effects of saturation only become relevant under high n_r where the signal exceeds K. For low n_r the spacing effect is enforced by the the simpler signal integrator model which

relies on differences in the filter signal dynamics between massed and spaced protocols. We find that the saturation mechanism offers limited effectiveness in suppressing massed repetitions from exceeding the allocation threshold P under arbitrary long repetition protocols but it does ensure that spaced protocols optimize R kinase activation. As previously discussed, perhaps strong massed protocols are capable of inducing late-phase plasticity and their difference to spaced protocols of the same n_r is that the latter are simply more efficient at stimulating the relevant signalling pathways. Our saturation model on the global signal therefore offers a simple model of a mechanism that could explain such phenomena. Once the global signal is initiated when $R^\star \geq P$, it is down to the local synaptic rules to enforce some form of spacing effect in the allocated signal by choosing which synapses are tagged for allocation. As with the non-saturating model our initial results have used a simple local-rule of tagging synapses with threshold-cycles of $c \geq 1$. Figure 9.14 compares the allocated SNR between $c_\Theta = 1$ and $c_\Theta = 4$, showing that the latter suppresses the massed protocol SNR under $n_r = 8$ where P threshold is reached by both massed and spaced protocols.

In conclusion, saturation has extended the spacing effect in memory allocation, which is naturally exhibited by integrating signal samples of synaptic filters, to longer repetitions protocols but its ability to enforce a spacing constraint is still limited by n_r. The saturation mechanism does not actually rely on anything specific about synaptic filter signal dynamics, it only implements part of a spatio-temporal filter on the signal samples that stimulate the signalling pathways responsible for long-term plasticity.

9.7 Summary

In this chapter we have focused on answering how synapses can detect strong memory encoding protocols that induce stable forms of plasticity. Following the findings of the previous chapter, that single synapses cannot reliably detect the pattern of repetitive memory encoding, we have been led to assert that repetitive encoding is detected at the level of a population of synapses either neuron-wide or at neural compartments that may contain hundreds of synapses. Based on this assertion we investigated how the synaptic filter's mean memory signal $\mu(t)$ behaves under strong encoding protocols. In particular, we looked at how $\mu(t)$ is augmented under an increasing number of memory repetitions n_r and how the repetition time-interval T_r changes the maximum mean signal obtained. Our conclusion is that a threshold on $\mu(t)$ could be used to detect a minimum number of repetitions at some T_r, but this threshold detection would not be specific to this particular protocol. Increasing the number of repetitions n_r allows crossing the threshold if the interval T_r is decreased simultaneously.

A simple model of how a neural depolarisation signals could be biophysically processed by early components of the cAMP pathway was constructed. According to our initial

model, the CaM kinase bound on the AC enzyme implements a threshold on $\mu(t)$ by requiring the cooperative binding of Ca^{2+} ions whose concentration depends on neural depolarisation. The activated kinase stimulates the AC enzyme to accelerate the production of the cAMP second messenger molecule. The concentration of cAMP provides a signal $u(t)$ that in turn activates the PKA kinase R, taken here to be a surrogate of all downstream processes for late-phase plasticity. In our models the processes that lead to stable forms of memory within a neural compartment are initiated once a sufficient level P of the activated kinase $R^\star$ has been reached. Once $R^\star$ exceeds a threshold P then we allocate a memory by freezing the state of the "tagged" synapses. These are synapses that have conducted a minimum number of same-threshold-crossings c_Θ along the "correct" strength state according to the strongly encoded memory. The correct states could be indicated to synapses by combining neuromodulation with information about the direction of plasticity.

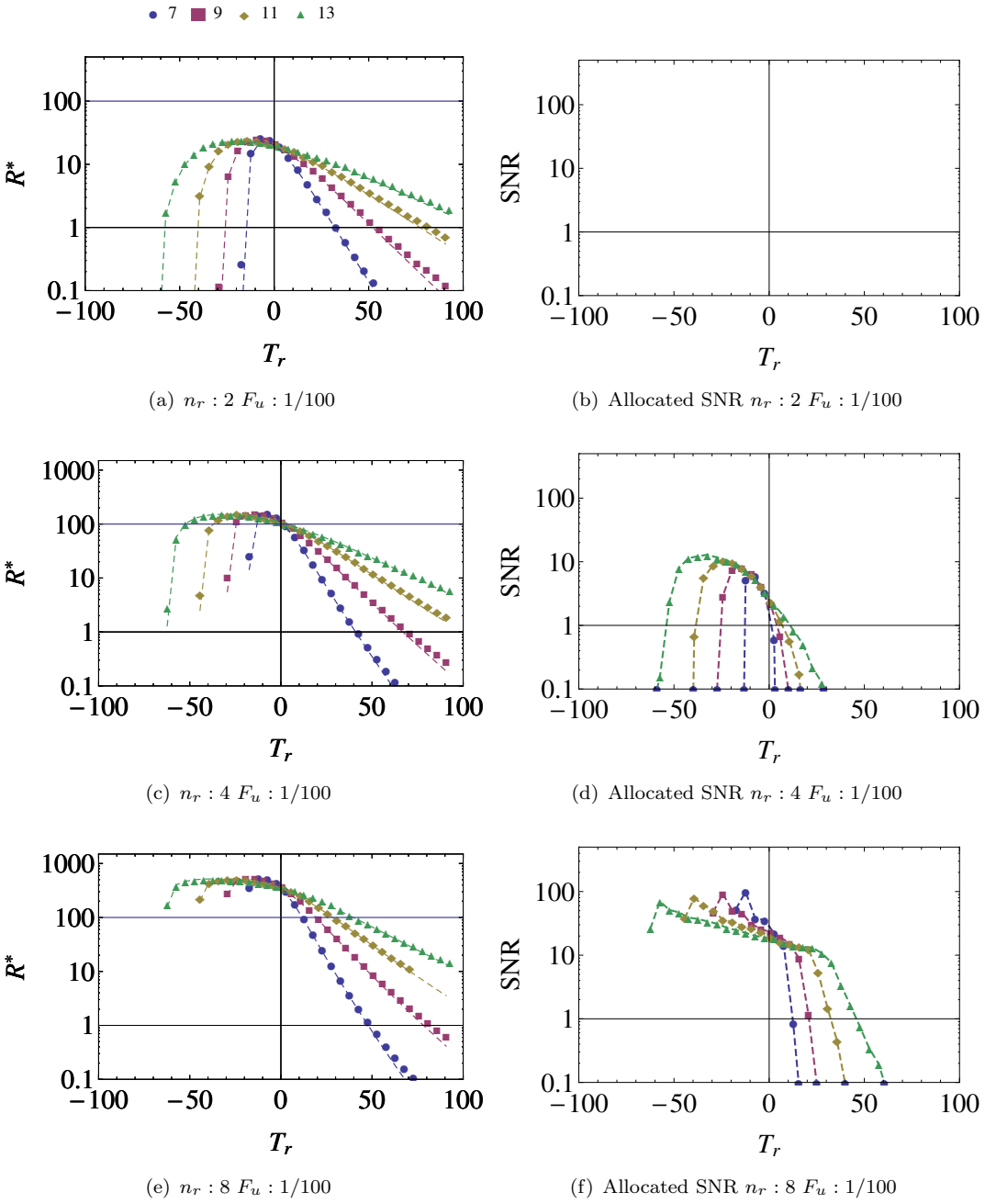
This model confirmed that relying on neural depolarisation at the level of a neuron or a dendritic compartment makes it possible to obtain mechanisms that detects a minimum level of spaced memory repetition in order to trigger or initiate memory allocation. However, our simple model, which is only driven by neural depolarisation, cannot be used to explain the phenomenon of spacing in the formation of long-term memory with increasing number of repetitions n_r. This comes as a result of the filter signal dynamics which is optimally augmented under massed protocols as the number of repetitions n_r increases. Therefore we considered taking issues of compartmental signalling control into account that may be able to explain the spacing effect.

We examined a model which assumes that signal pathways within a compartments saturate under very strong massed protocols. Saturation could be due to active negative feedback control which protects cells from toxic over-stimulation or due to natural constraints in the maximum level of enzymes activation for the production of the $u(t)$ signal per compartment. Our simple model was extended to include saturation at u_{max} and demonstrated that a spacing effect in the activation of R is imposed under long n_r protocols. The spacing here works by sub-linearly integrating signal samples that occur before $u(t)$ has decayed away from the saturation level u_{max}. Such temporal constraints on signal sampling allowed to extend the spacing effect for allocation to longer repetitions protocols. Although with saturation in place spaced protocols optimally initiate the global allocation signal $R^\star$, it is still possible for long massed repetition protocols to sufficiently activate R and initiate memory allocation.

Further, our results imply that a spacing effect needs to be maintained by specifically controlling the protocols that initiate a global signal for allocation. Once the global allocation processes are initiated the limited ability of local-information to identify spaced to massed protocols is exposed. However, here we have not explored all possible local tagging rules that could be derived from TC statistics. We have used a basic TC limit rule to identify the synapses that should make a transition to late-phase plasticity in

response to the global allocation switch. The global and local information in effect here cooperate by having the spacing effect imposed by the global signal pathway under low n_r while under high n_r, where global processes buckle, local information to identify the protocol are utilized to impose a spacing effect. This cooperation may possibly be improved if global mechanism further delay massed protocols from initiating allocation, as we have seen in Chapter 8 that local synapse TC statistics improve as protocols become longer in n_r.

The saturation model can impose sufficiently long temporal constraints to the repetition intervals but requires arbitrarily slow degradation rates of u within a compartment and thus low values of F_u. The requirement that $u(t)$ decay needs to be very slow raises questions on how the cAMP pathway could actually retain such low rates. F_u needs to be such that $u(t)$ does not decay to baseline between spaced repetition intervals which could be 15 to 30 minutes according experimental settings discussed earlier. Even if the rate of degradation of cAMP is slow the molecular signal would nevertheless diffuse over long distances in the absence of physical boundaries between compartments. Therefore, a model relying on arbitrary slow kinetics of u maybe unrealistic. Although the specifics of cAMP signalling between AC to PKA in this model may be over simplified we believe it still contributes a framework to consider signal sampling and the role of saturation in the spacing effect. The key components of CaM stimulation of AC via depolarization and the integration of cAMP by PKA are well known and thus our models offer to the interpretation of their role for late-phase plasticity. However, specifics of how saturation may actually be biophysically implemented are not implied by our model and remain to be elaborated by other models dealing with how compartments may saturate the relevant signalling.



(a) $n_r : 2$ $F_u : 1/100$

(b) Allocated SNR $n_r : 2$ $F_u : 1/100$

(c) $n_r : 4$ $F_u : 1/100$

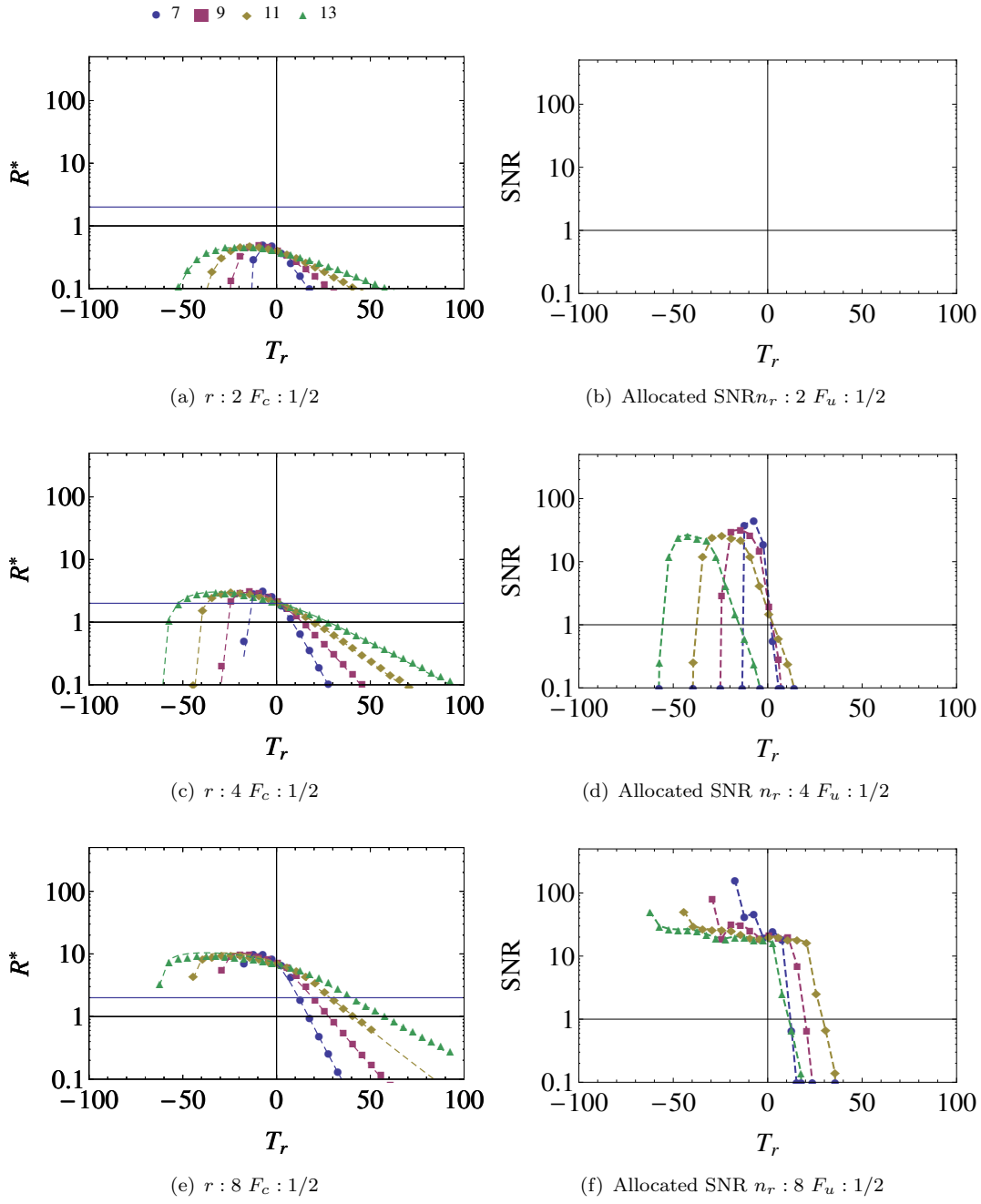
(d) Allocated SNR $n_r : 4$ $F_u : 1/100$

(e) $n_r : 8$ $F_u : 1/100$

(f) Allocated SNR $n_r : 8$ $F_u : 1/100$

FIGURE 9.8: Figures on the left column show the direct kinase activation model's overall output kinase $R^\star$ produced per repetition interval T_r with $F_u = 1/100$ for four filter sizes $\Theta \in \{7, 9, 11, 13\}$. All curves are shifted to the left so $T_r = 0$ is the on filter peak signal time t_p and therefore a massed protocol is represented by the first point on each curve. The $R^\star$ curves shown where produced using simulation of the stochastic model (each marker $N = 10^4$ synapses over $T = 10^3$ trials) with the overlayed dashed lines showing numerical solutions of the mean signal dynamic system confirming the agreement with the simulation.

The right column shows the respective allocated signal for each repetition protocol-filter combination evaluated via simulation only with the $R^\star$ threshold set to $P = 80$. The allocated signal is given by measuring the $\mu(t)/\sigma(t)$ at the time when the global allocation signal ($R^\star = P$) occurs only over the population of synapses that have been "locked". below it no signal is allocated under $n_r = 2$ while at $n_r = 8$ massed repetitions manage to exceed P .



(a) $r : 2$ $F_c : 1/2$

(b) Allocated SNR$n_r : 2$ $F_u : 1/2$

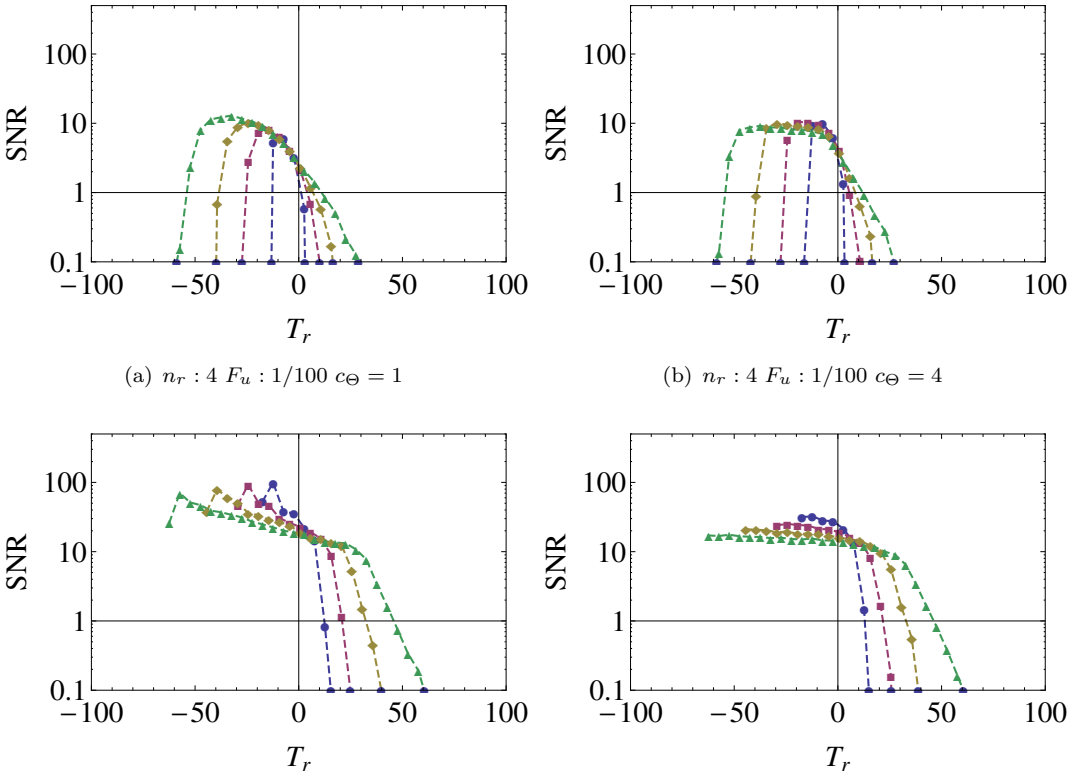
(c) $r : 4$ $F_c : 1/2$

(d) Allocated SNR $n_r : 4$ $F_u : 1/2$

(e) $r : 8$ $F_c : 1/2$

(f) Allocated SNR $n_r : 8$ $F_u : 1/2$

FIGURE 9.9: Direct kinase activation output per repetition interval and allocated SNR. Left column shows the activation of output kinase per repetition interval T_r for filter sizes $\Theta \in \{7, 9, 11, 13\}$. The dashed lines were obtained numerically from results using mean signal $\mu(t)$ and the overlayed points markers represent simulation results of the stochastic signal.($N = 10^4$ synapses and $T = 10^3$ trials for each point). Allocated SNR curve uses $c_\Theta = 1$ local was obtained in simulation only. Using a $c_\Theta = 1$ elucidates the role of the global mechanism in producing the spacing effect without contriving local synaptic rules. With $P = 2$ and low n_r no memory allocation occurs, while at high $n_r = 8$ the $R^\star$ is exceeded for all massed T_r.



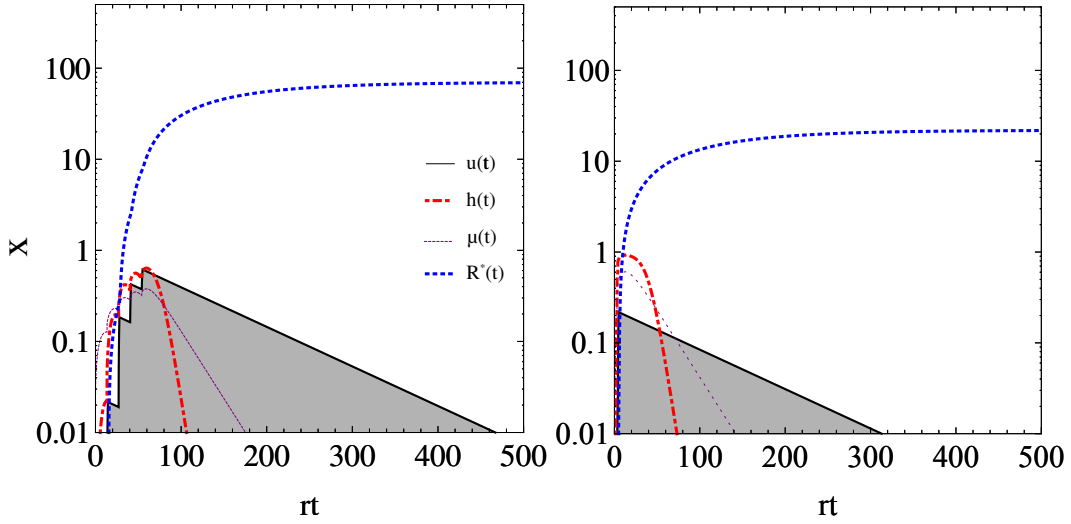
(a) $n_r : 4$ $F_u : 1/100$ $c_\Theta = 1$

(b) $n_r : 4$ $F_u : 1/100$ $c_\Theta = 4$

(c) $n_r : 8$ $F_u : 1/100$ $c_\Theta = 1$

(d) $n_r : 8$ $F_u : 1/100$ $c_\Theta = 4$

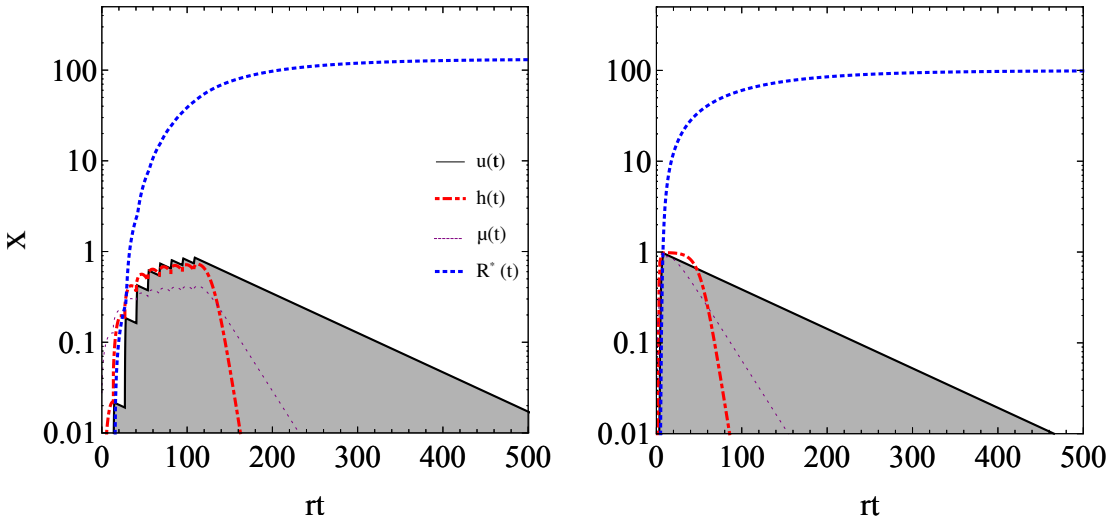
FIGURE 9.10: Direct kinase activation model, comparing allocated signal between $c \geq 4$ and $c \geq 1$ as allocation thresholds. Allocation curves are simulation only, with threshold set $P = 100$ and threshold cycle limit $c = 4$. Comparing to the $c = 1$ curves on Figure 9.8 we find that differences only arise for $n_r = 8$ case, where the massed SNR which used to be around 100 has now dropped to the levels of spaced at approximatelly 20. Although now massed repetitions do not exceed spaced for large Θ, still the two protocols do not differentiate. .



(a) $n_r : 4$ $F_u : 1/100$ Spaced

(b) $n_r : 4$ $F_u : 1/100$ Massed

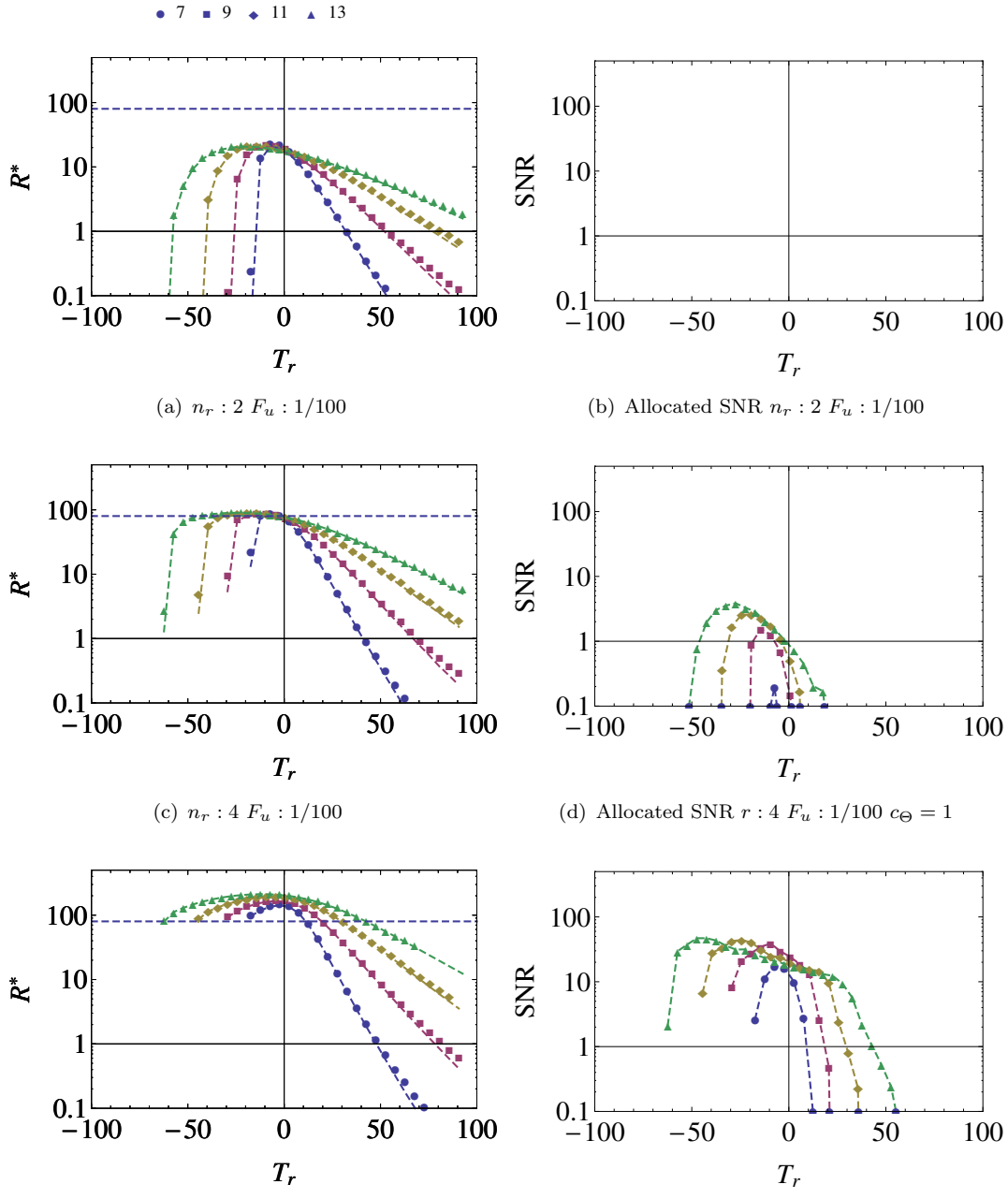
FIGURE 9.11: Dynamics of a saturating model of kinase activation. The thick dotted line reaching around hundred represents the overall activation of output kinase $R(t)$. The $R^{\star}(t)$ output is calculated by the integral of $u(t)$, shown as the shaded area. $h(t)$,shown here as thick dot-dashed line, drives the production of second messenger $u(t)$. The signal $h(t)$ implements a threshold function on $\mu(t)$, and thus it roughly follows the profile of the thin dotted line showing $\mu(t)$. As it can be seen on the plots on the right column, due to the saturation mechanism, $u(t)$ cannot exceed $u_max = 1$ as in the non-saturating model under $n_r = 8$. We observe that four spaced repetitions are better than four massed. although the system saturates $u(t)$ the overall $R(t)$ kinase activated under eight massed repetitions is slightly higher than four spaced repetitions.



(a) $n_r : 8$ $F_u : 1/100$ Spaced

(b) $n_r : 8$ $F_u : 1/100$ Massed

FIGURE 9.12: Dynamics of a saturating model of kinase activation. The thick dotted line reaching around hundred represents the overall activation of output kinase $R(t)$. The $R^{\star}(t)$ output is calculated by the integral of $u(t)$, shown as the shaded area. $h(t)$,shown here as thick dot-dashed line, drives the production of second messenger $u(t)$. The signal $h(t)$ implements a threshold function on $\mu(t)$, and thus it roughly follows the profile of the thin dotted line showing $\mu(t)$. As it can be seen on the plots on the right column, due to the saturation mechanism, $u(t)$ cannot exceed $u_max = 1$ as in the non-saturating model under $n_r = 8$. We observe that four spaced repetitions are better than four massed. although the system saturates $u(t)$ the overall $R(t)$ kinase activated under eight massed repetitions is slightly higher than four spaced repetitions.



(a) $n_r : 2$ $F_u : 1/100$

(b) Allocated SNR $n_r : 2$ $F_u : 1/100$

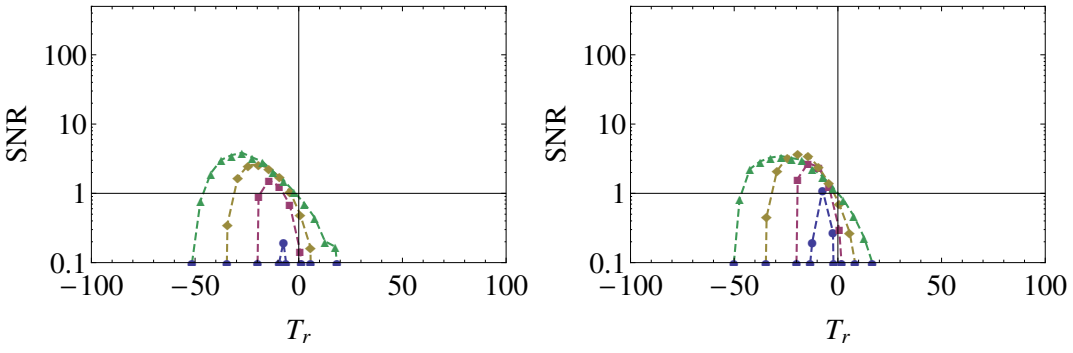
(c) $n_r : 4$ $F_u : 1/100$

(d) Allocated SNR $r : 4$ $F_u : 1/100$ $c_\Theta = 1$

(e) $n_r : 8$ $F_u : 1/100$

(f) Allocated SNR $n_r : 8$ $F_u : 1/100$ $c_\Theta = 1$

FIGURE 9.13: Saturating activation of target protein per repetition interval for four filter sizes $\Theta \in \{7, 9, 11, 13\}$ and $F_u = 1/100$ with low local allocation threshold $c_\Theta = 1$. The dashed curved lines are numerical solutions of the dynamic system and the overlayed points markers represent simulation results. The horizontal dashed line shows the R* threshold point $P = 80$. We have set P just under the peak values of R* for $n_r = 4$ and thus allocation occurs only around peak R*. Using a $c_\Theta = 1$ elucidates the role of the global mechanism in producing the spacing effect without contriving local synaptic rules. As before, for $n_r : 2$ R* activation falls below P and no memory allocation occurs. With high $n_r = 8$ we see that $R^\star$ curve is deformed to exceed the $P = 80$ threshold even at low T_r but due to saturation the R* curvature of remains unlike the non saturating model.



(a) $n_r : 4$ $F_u : 1/100$ $c_\Theta = 1$

(b) $n_r : 4$ $F_u : 1/100$ $c_\Theta = 4$

(c) $n_r : 8$ $F_u : 1/100$ $c_\Theta = 1$

(d) $n_r : 8$ $F_u : 1/100$ $c_\Theta = 4$

FIGURE 9.14: Comparing the allocated SNR of saturating kinase activation under a $c_\Theta = 1$ against a $c_\Theta = 4$ allocation threshold. Increasing c_Θ under $n_r = 8$ is shown to reduce the allocated signal for short $T_r < T_r^*$ interval protocols .

Chapter 10

Discussion

Memory is a ubiquitous property of neural systems and in complex organisms we find multiple memory systems each one adapted to a particular mental faculty. We believe that multiple memory systems evolved to allow organisms to deal with the complexity of their environment but in the absence of an evolutionary fossil record for brains we do not know the course of this evolution. If multiple memory systems evolved out of evolutionary tinkering of network parameters, then perhaps evolution would simultaneously evolve neurons and synapses in a manner that allows this specialization. The level of specialisation could be such that each memory system is constrained to areas where other memory systems would mostly fail (Sherry and Schacter, 1987). On the one end of the spectrum we find very specialized systems like the memory supporting song learning in birds, which express a critical period similar to human speech learning (see Doupe and Kuhl, 1999). However, we also have the example of the hippocampal memory system that targets broader content by supporting flexible representations used for navigation and episodic memory in general (Eichenbaum and Cohen, 2004). Human episodic memory acquires a stream of every-day facts and events while it is expected to serve healthy individuals to form new memories over their lifetime. The hypothesis emerged that episodic memory may be stored entirely synaptically in the hippocampus following evidence of persistent synaptic efficacy modifications after neural pathway stimulation (see Chapters 1 and 2).

If the hippocampus is a specialized online memory system for processing a stream of flexible representations, then it is reasonable to expect that its synaptic plasticity and network topology is specifically adapted to this function. The structure of the CA1 and CA3 hippocampal areas suggest that memory is stored in a recurrent network topology. However, recurrent network models with unbounded synapses of infinite resolution have limits in the number of patterns they are able to sustain. Beyond these limits they exhibit catastrophic forgetting, an inability to recall all stored patterns (French, 1999, McCloskey and Cohen, 1989, Ratcliff, 1990). To circumvent this problem we may either assume that synapses also unlearn stored patterns or that learning stops before a critical

limit and new neural resources are made available for further learning. Based on initial indications that neural resources in the brain remain fixed, theoretical investigations have attempted to explain how a number of patterns can overlappingly be synaptically stored within a single neural network. The underlying assumption has been that these resources can be exhausted by online learning of everyday facts and events and thus a level of forgetting is necessary.

According to this reasoning the saturation limit of a memory system becomes a relevant parameter and it assumes that biology works at the limits of its available resources. However, as we discussed in Section 1.3 forgetting could be under the control of an executive mechanism that attempts to keep only relevant information. An executive control mechanism could remove memories by reversing changes induced on synaptic efficacy. Accordingly, a memory system could operate under its saturation level by continual re-evaluation of memories in light of new memories. Thus, system level processes could have evolved to solve particular problems in a manner that would make evolutionary adaptations on the neural substrate unnecessary or irrelevant. For example, declarative memory encoding occurs between sleep cycles so its capacity and rate of learning could be such that well satisfy its online function for the animal. Current research on the relationship between sleep and memory consolidation reveals convincing evidence that offline memory reprocessing during sleep is an important component that shapes which memories are consolidated (see Stickgold, 2005). It has also been shown that sleep enhances relational memory performance on previously learned material (Ellenbogen et al., 2007).

Nonetheless, it could be argued that this re-organization of memory would still require resources to consolidate at least some of the memories occurring during the daytime and thus the usual network limits of fixed resources still apply. This last assumption perhaps is not that strong in the light of evidence of neurogeneration in the subgranular zone (SGZ) in the DG of the hippocampus, but significant questions remain to be addressed regarding contributions of new neurons to normal brain functions (see Ming and Song, 2011, Leuner et al., 2006). Nevertheless, neurogenesis could, in principle, incrementally increase the capacity of the memory system by adding more resources, but research in this direction has not yet produced conclusive results due to technical difficulties in achieving a disruption of neurogenesis without affecting normal circuit function (Leuner et al., 2006). Additionally, a system level re-organization phase could be limiting the rate of memory consolidation to such degree so the network limits with unbounded synapses leading to catastrophic forgetting are never reached, while neurogenesis may continuously add to the available neural resources. Therefore, overcoming catastrophic forgetting could be just a physist's focus on a saturation limit of memory networks that is not relevant to biology. However, beyond memory palimpsest, experimental evidence suggests that CA1-CA3 synapses have access to a only limited number of strength states,

which could be binary (Petersen et al., 1998, O'Connor et al., 2005) or ternary (Montgomery and Madison, 2004). Such evidence re-motivates the examination of overlapping memory encoding with "realistic" synapses that have discrete states of strength. As we have discussed extensively in the foregoing, automatic learning and unlearning can be produced by imposing appropriate bounds to synaptic strength introducing a palimpsest (Nadal et al., 1986) property that protects from catastrophic forgetting. As a result, significant theoretical efforts have concentrated on answering how memory systems assimilate new memories in a manner that minimizes the disruptive effects on previously stored memories. The presumption in these theoretical studies is that the single synapse is the computational unit of the memory system and thus theorists should address the notorious synaptic stability versus plasticity dilemma (Abraham and Robins, 2005) at this level.

In this thesis we have extended a synapse model previously applied to the control of fluctuations in neural development (Elliott, 2008, Elliott and Lagogiannis, 2009) to address the stability versus plasticity dilemma in synaptically stored memory. We then consider the question of how memory systems could consolidate memories selectively under protocols that have been experimentally shown to optimize storage conditions. The discussion of our findings in this study are organized accordingly. The results on the use of synaptic integration for resolving the stability versus plasticity dilemma are followed by a discussion of the results from our study in memory consolidation and the spacing effect. Finally, we propose future experimental and theoretical work that could elucidate the relation of our theoretical investigations to the biology of memory but also extend our current theoretical framework to consider further phenomena.

10.1 Synaptic integration as a response to the stability Vs plasticity dilemma

10.1.1 Resistance to plasticity is futile

Formal models of synapses attempting to address stability versus plasticity dilemma have relied on dilating the time it takes for new memories to degrade previously stored ones. A concept initially put forward by Tsodyks (1990) and subsequently followed up by Amit and Fusi (1994) assumes that each synapse randomly expresses the plasticity signal induced by each new memory. Consequently, previous memory traces are partially overwritten in a stochastic manner to make room for new ones.

The output of a single formal neuron to a learned pattern can be used to measure the memory signal and capacity (see Barrett and Van Rossum, 2008), because under asynchronous recurrent network dynamics sufficient error rate in the recall of single neurons creates an avalanche effect disrupting recall over the whole network (see Chapter

4. Since the engram is stored synaptically obtaining memory capacity as signal-to-noise ratios in this setting gives one way to measure how new memory encoding disrupts previously stored memories. Such measures became standard in these models as they are divorced from network dynamics and take the view point of an ideal observer having access to the strength states of all synapses to measure fidelity of stored memories under the influence of new memory storage. With binary synapses that stochastically update their strength states to incoming stimuli the signal simply decays exponentially with the number of patterns stored. However, slowing the decay rate directly affects the initial signal of newly encoded memories, which needs to be significantly higher than the noise in the signal for successful recall in recurrent networks (see Huang and Amit, 2011, van Rossum et al., 2012).

Fusi et al. (2005) and Fusi and Abbott (2007) proposed a model with hidden states attached to each synaptic strength with progressively lesser degrees of plasticity. This idea was discussed in detail in Chapter 5, it achieves an extension of the signal lifetimes due to a combination of multiple exponential decay rates stemming from the fact that a memory trace is stored in synapses that exist in synaptic states of different transition probabilities. The most labile states give high initial SNR while the less labile states aim to prolong the decay of the memory signal. A theoretical treatment by the original authors went up to an approximation using the mean dynamics. In this thesis we have contributed an exact master equation approach which contains all higher moments but up to Laplace transform that has to be evaluated numerically. At its core, the cascade model does not really revise the solution to the stability-plasticity dilemma, it still relies on time dilation, but due to the superposition of multiple exponentials the signal decay gives power-law forgetting dynamics (see Chapter 5).

Such forgetting dynamics appear consistent with results from psychology research measuring forgetting dynamics (Ebbinghaus, 1885, Wixted and Ebbesen, 1991, Wixted and Carpenter, 2007, Rubin et al., 1999). But these protocols assume memory oblivescence only and do not measure how the recalled content changes at various time intervals. For example, the influential work conducted by Ebbinghaus used a savings index that measures the time required to relearn a material to perfect recall, yet it does not examine how memory contents evolve and there is no guarantee these tests are engaging a single memory system. On the other hand, protocols measuring recall relative to previous attempts allow a comparison of the contents between recalls and have shown that memory dynamics contain both an oblivescence and a reminiscence component (Ballard, 1913, Erdelyi, 2010, Brown, 1923, Tulving, 1967, Campbell et al., 2011). Nevertheless, it is risky to take the memory dynamics from psychology literature as a guide of the forgetting dynamics of a single neural network because test-subjects probably engage various memory faculties to the task and each faculty may have its own memory dynamics. In Chapter 1 we argued that it is this complexity that may underlie the power-law

forgetting dynamics reported. It is known that power-law like functions are ubiquitous in nature because they emerge when various functions of different time-scales are combined (Anderson, 2001) and thus an attempt to validate a model against power-law forgetting dynamics derived from an unknown number of memory faculties invalidates the approach.

10.1.2 Plasticity processing by separating induction from expression

All previous models of synaptically stored memory requiring stability in the face of ongoing plasticity have relied on state models that link the strength state of a synapse with a state representing the readiness of a synapse for plasticity. Interestingly, synapses are faced with a similar dilemma within the context of neural development. For example, during the developing phase of the visual system afferent neurons compete for control of a target neuron giving rise to segregated patterns of the innervation that result in the ocular dominance columns of the visual system (Purves, 1984). However, fluctuations in the plasticity inducing stimuli driving the development of the afferent neurons can destabilize the slowly developing patterns of target innervation (Elliott and Lagogiannis, 2009). Here, before they express plasticity synapses need to identify if the last stimulus is part of a persistent trend in the input statistics or just destabilizing noise. A simple solution to this problem can be given if synapses integrate stimuli to only express when they detect a particular trend. Indeed, the integration mechanism may underlie the requirement to deliver numerous plasticity stimuli before synaptic plasticity is expressed in typical experiments and this delay in expression would allow synapses to integrate plasticity stimuli before deciding on expressing any changes in strength. Theoretically, integration of the induction signals has been shown to robustly control developmental fluctuations (Elliott and Lagogiannis, 2009, Elliott, 2011b).

In this thesis we have adopted synaptic integrative dynamics to replace the stochastic updating mechanisms in models of synaptically stored memory. Our proposed model separates the processes of plasticity induction and expression and assumes that the hidden states of the model processing induction stimuli are independent of the strength state of a synapse. Beyond this revision, we follow a standard theoretical framework to compare our plasticity model against others. We isolate a single neuron from Hopfield associative memory (Hopfield, 1982) to obtain a perceptron on which to test the number of random uncorrelated binary patterns it recognizes under supervised association of all patterns to one class (see Barrett and Van Rossum, 2008). In this manner, the random uncorrelated patterns can be translated to the POT and DEP plasticity stimuli that feed the synaptic filters and, modulo the fact that low inputs (-1) via weak synapses (-1) enhance the output, the $h(t)$ perceptron output can be interpreted as neural depolarization. Nevertheless, the formal device providing plasticity stimuli will not influence the filter dynamics as long as it provides a random sequence of equiprobable potentiation

and depression stimuli and thus in general we anticipate that our results generalize to more realistic neuron models so long as the plasticity stimuli preserve the dynamics of our synaptic filters.

Further, we have assumed that synaptic strength is binary, but the filter framework is not restricted to a binary view of synapses and it can easily be extended to consider further discrete states of strength. This is contrasted to cascade models where the hidden metaplastic states are tightly bound with synaptic strength. With the induction process separated however, the filters could still indicate that stimulus conditions have been met for plasticity expression but if synaptic strength is saturated then this framework allows for not expressing any plasticity although the filter perform a threshold transition and reset to zero state. On the other hand we could modify filters to convert the threshold towards the saturated strength into a reflecting barrier. In this case the saturated strength states would become more stable and increase capacity, but such modifications would compromise the independence between the induction and the plasticity expression processes. We assume this separation makes sense for the biophysical substrate as well, the induction stimuli are specifically detected by one mechanism which could then signal multiple downstream process that perceive this signal in a state dependent manner.

To an external observer, who has no access to the internal states, an integrative synapse subjected to random uncorrelated stimuli would appear to randomly express induction stimuli. Critically however, the expression time of low-pass (LP) filter integrating synapses is a non-linear function of the drift in the sequence of induction signals (Elliott, 2011a). Under equipotential random uncorrelated stimuli expression time is maximally suppressed and sufficiently large fluctuations are required to drive to synapses to expression threshold. As we showed in Chapter 6, in radical contrast to previous models, which aim to extend memory lifetimes by dilating time and contain an oblivescence component only, these synapses give very different memory signal dynamics that exhibit both an oblivescence and a reminiscence component (Elliott and Lagogiannis, 2012). With synaptic filters after the initial encoding of a tracked pattern further memory encoding will enhance the output of a perceptron showing a signal rising to a peak. The rise is caused by a bias introduced in the distribution of filter states towards a particular threshold dictated by the tracked memory. Beyond the peak the bias is lost and the signal begins to decay asymptotically towards zero.

These remarkable dynamics also weaken previous constraints faced by plasticity models where a trade-off between initial signal and signal decay provided a hard lower bound to the learning rate. Filter synapses, may exhibit an initial signal that is below a threshold for recall but it may subsequently rise and allow memory recall. The initial and peak signal of LP filters are determined by Θ, the signal rises Θ-fold from Θ^{-2} to $\beta\Theta^{-1}$ where $\beta \approx 0.766$. Therefore, although there is a square dependence of initial signal to Θ the peak signal drops linearly with Θ relaxing the constraints on tuning learning rates to network size N. For example, taking noise as $\sqrt{N}$ a population N of SU synapses

requires $N > q^{-2}$ for the initial SNR to exceed unity while LP filter synapses with an equivalent $q = \Theta^{-2}$ would require $N > (\Theta/\beta)^2 \equiv q^{-1}/\beta^2$ for the signal peak to exceed the noise. Hence, there is a considerable weaker dependence on q.

Along with our LP filter synapse we have produced another version of an integrative synapse that we named the *dual filter* as it contains two independent integrators, one for each induction stimulus. These synapses effectively make induction stimuli compete for expression in a race towards collecting a minimum number of Θ induction stimuli. The plasticity expressed is determined by the stimulus that wins the race and subsequently both integrators are reset to restart the process. We showed that this synapse also exhibits a rising signal as initial encoding gives an advantage to one of the two integrators therefore biasing plasticity expression.

10.1.3 Superior capacity with low-pass filtering

Filter synapses have superior memory lifetimes to cascades of an equal number of states $n = \Theta$ over a biologically relevant range of network size N. Initially, we measured lifetimes as the time until the mean signal $\mu(t)$ drops below some lower threshold and found that LP filters outperform a cascade over a wide range of biologically relevant number of synapses N and across all ranges of recall thresholds tested. Large LP filters $\Theta \geq 15$ in the region of $N < 10^4$ do not reach sufficient initial signal and therefore fall behind the cascade yet they perform better than the cascade over larger N. The low $N < 10^4$ region is suited for smaller filters $\Theta < 8$ which robustly outperform cascades for the same state count $\Theta = n$. These, performance gains can be significant, for example between $10^3 < N < 10^4$ a $\Theta = 15$ filter outperforms a cascade's capacity of an equivalent number of states ($n = \Theta$) by a factor ranging from 30 to 8, with performance dropping as N increases. Comparing against even larger filters ($\Theta = 20$) shows further relative gains in capacity. However, these results also reveal that although capacity grows approximately proportionally with the $log(N)$, cascades grow faster but manage to exceed filters in a region closer to $N > 10^5$. Nevertheless, this region in N is not biologically relevant as synaptic count measurements in hippocampal pyramidal neurons have been shown to be around 10^4 to 3×10^4 synapses (see Megıas et al., 2001).

Critically, the performance gains reported ignore any low initial signal and only focus on the point in time when the mean signal drops below an SNR of unity. Such methods are standard but they are based on view of a decaying signal only. With the rising filter signals however these methods may be misleading because the point where memory signal drops does cannot be used as a measure of the length of time the signal has been above the noise. The filter signal may rise for a short time period around the peak before it drops below the noise again. These dynamics make the interpretation of filter signals as memory recall dynamics more complicated but as we discussed earlier these same dynamics also relax any hard lower constraints on the initial signal.

We confirmed that cascade sizes n need to be tuned to network size N to optimize capacity (Fusi et al., 2005). However, this relationship between N and n was abolished by changing they way memory lifetimes are measured from the time that the mean signal reaches a lower bound to measuring the mean first passage time (MFPT) for the memory signal to reach a lower bound instead. Under MFPT the cascade's capacity increases with n to a plateau instead of decreasing after the optimal n as originally reported. Besides these differences, MFPT measures verify our earlier results showing LP filter synapses exhibiting superior memory retention to cascade. However this was not the case for dual-filters, they appeared to scale capacity very poorly with N, as capacity would quickly rise to a plateau under both MFPT and for the mean memory lifetime measures. Compared to the cascade its capacity only appears comparable for low $N < 10^4$ and only outperforms the cascade for large state counts $n > 10$ and small networks $N < 10^3$.

Following next we explored the idea of creating a filter-cascade which will be equivalent to the original model by replacing the stochastic updaters at each cascade state by filters of matched mean expression time. Here, we considered an extension of LP filter synapses with stochastic decay, such a model having previously been examined in a development context (Elliott, 2011a). The decay component allows synapses to passively return towards the zero state in manner similar to radioactive decay. Such synapses not only allow matching the expression times of each cascade state exactly but also account for the fact that memories encoded distant in time do not interact. Thus, now we have a more generic form of a LP filter synapse which has the added parameter of passive decay η. Clearly for $\eta = 0$ we obtain the previous simpler LP synapse.

To create a filter-cascade equivalent using generic LP synaptic filters we need to choose an appropriate Θ and decay η to match each cascade's state escape time to the mean expression time of an LP filter. The choice of Θ and decay η to match a particular expression time is not unique. In our version of a filter-cascade equivalent relatively small filters where used with $\Theta < 5$, which limited the low-pass filtering ability of our synapses to very small fluctuations. Nevertheless, the filter-cascade robustly outperformed the original cascade up to the $N < 10^6$ examined. The choice of filter parameters to match expression times is not unique for some cascade states. Increasing the decay and making the filters smaller would result in synapses that approach the stochastic response limit and thus we return to the behaviour of the original cascade.

The filter-cascades under mean memory lifetimes it also exhibited the tuning of capacity to the number of cascade states n, which again disappears when under MFPT capacity measures. The filter-cascade equivalent however is composed of more hidden states than the original cascade as matching is only obtained in terms of expression times at each cascade state. Theoretically, this gives the filter-cascade the ability to store more information in the hidden states. But our results comparing single filters against

cascades on matched number of states show that it is the particular signal dynamics of filter synapses that increase the capacity.

We did not check however if arbitrarily tuned filter-cascades always outperform cascades of an equivalent size. Since single filters mostly outperform cascades of equivalent states counts we expect that a such filters arranged in a cascade would also outperform the original cascade even though the would not retain the progressive stabilization structure of the cascade.

In conclusion our results demonstrate that LP filter synapses exhibit superior memory lifetimes as measured by the output of a single formal neuron but also within recurrent network dynamics. Our filter synapses can operate over a wider parameter space by allowing future recall even in the absence of an sufficiently initial signal.

10.1.4 Why do low-pass filters have superior capacity?

It is seems rather contradictory that we have proposed LP filter synapses as relevant to plasticity due to their signal processing ability but at the same time we are testing capacity with random uncorrelated stimuli where apparently there is no "signal". However, we have quantitative evidence that LP filter synapses can exhibit superior memory lifetimes, at least within a convenient region of N. These can be further extended if one ignores low initial signals to consider the rise and fall of the filter signal. Naively, by matching filter expression times to SU synapses one would expect that memory lifetimes would remain the same between models and all we would have obtained is a re-interpretation of expression times. Off course there are obvious differences between synaptic behaviour of the two models. For example, filter synapses exhibit a refraction in the expression time because there is a minimum number of induction stimuli to reach threshold after a re-injection to the zero state occurs. However, by modifying filters to re-inject into a random state according to the equilibrium PDF of filter states we find that refraction is not the cause of this large expansion of memory lifetimes.

At the same time, the integration phase of our synapses alone is also insufficient to explain the superior capacity performance. Taking a high level view, the dual-filter synapse appears to share characteristic with the LP filter synapse. Both have a refractory period between plasticity events equal to Θ, they both exhibit a rising signal and both integrate plasticity stimuli before expression. In contrast to LP filters, the dual filter has a maximum waiting time of plasticity expression equal to 2Θ. In the event of alternating induction stimuli the integrators will fill up after 2Θ induction stimuli have arrived and therefore plasticity will be expressed. This suggests that the ability of the LP filter to suppress expression indefinitely until a large enough fluctuation drives them to threshold may be key for increasing memory lifetimes. The fact that dual-filters consistently show

inferior capacity to the low-pass filter synapse and that their capacity does not grow above some N indicates that the particular form of synaptic integration is important.

It is difficult to explain how LP filter synapses re-organize synaptic strength to accommodate more patterns in an ongoing memory system beyond the explanations we have provided through the analysis of the filter signal dynamics which exhibit both a bias and a time dilation through a reduced rate in expression. The bias towards threshold provides a rising component in the signal that effectively increases memory lifetimes. Our intuition is that low-pass filter synapses provide a meaningful re-organization of the weight vector for the memory patterns that have been encoded by allowing each synapse to utilize a history of stimuli to express a strength state which relates to the majority of recently seen patterns. The ability to indefinitely delay expression allows them to account for relationships between stored patterns stored far in the memory encoding history.

We anticipate that the capacity gains of LP filter synapses generalize to any two layer feed-forward network where pattern retrieval involves presenting a stored cue pattern that causes sufficient activity in the input layer to evoke activity in the target layer. On the other hand in recurrent network models of memory, the retrieval of a pattern relies on producing self-sustained activity due to a stable state of network dynamics (see Chapter 4). The stable states are still stored synaptically but the dynamics of retrieval complicate matters. Consequently, it may be argued that our SNR results do not directly transfer through as setting criteria for memory recall based on mean memory signals does not guarantee retrieval (Huang and Amit, 2011).

For this reason we compared memory capacities within a Hopfield network (Hopfield, 1982) by measuring the mean number of patterns stored before retrieval probability dropped to $1/2$. Checking our earlier results over a network of $N = 10^3$ neurons by comparing the original cascade model against filter-cascades or single LP and a dual-filter showed that filters models retain superiority, at least for this network size. For single filters however the superior memory lifetimes went up to a maximum model size $n = \Theta$, beyond which the initial signal was too low to allow immediate recall. These recurrent network capacity results were in agreement with earlier SNR and MFPT capacity results where the LP filter synapse exhibited higher capacity to other models but also to a matched dual-filter cascade. The key difference underlined is the collapse of capacity of filters beyond some n while the original cascade's capacity degrades slowly beyond the optimal size $n = 5$ under $N = 10^3$. As discussed previously, the initial signal of single filters changes as Θ^{-2} but cascades retain their most labile states and all that changes with an increase in cascade size n is the uniform probability $2/n$ that a synapse will occupy the two most labile states. Thus, cascades appear more robust against model size to network size effects and indeed filter-cascades do not show this capacity collapse either. However, we showed that taking the filter's rising signal into account by testing pattern retrieval up to at least the mean time of a filters peak signal we can significantly relax the

state count to network size constraints while we may obtain further significant increases in capacity. Thus, with synaptic filters under recurrent network dynamics a stored pattern that cannot be initially retrieved can have its retrieval probability increased by further memory storage. The memory dynamics therefore with integrative synaptic filtering mechanisms are radically different from those in which synapses stochastically respond to induction signals.

Critically the same LP filter synapse previously employed to suppress destabilizing fluctuations in neural development (Elliott and Lagogiannis, 2009) has been shown to successfully generalize over to synaptically stored ongoing memory. It is possible that the same synaptic plasticity processes serves neural systems across phases from development to ongoing memory processes. Consequently, synapses would have to somehow tune transition probabilities to the particular phase the neural system is undergoing. This could occur over a wide range, from very low transition probabilities that slow down development and reduce fluctuations to high encoding rates that allow one-shot memory encoding so the signal stands well above the noise level. A filter synapse provides a promising mechanism for modifying the learning rate because its transition probabilities naturally emerge from the model under the statistics of the input stimuli. It is not hard to imagine that expression times could be under dynamic control through the adjustment of filter thresholds $\Theta_{\pm}$. Further, we may envisage the generic form of LP filter synapses as being under dynamic control of the decay rate η. Increasing η would make a memory system enter a more stable state which would resemble the end of a critical development period. Such parametrization would also be useful for the evolutionary development of multiple memory systems. Instead of evolving a generic memory system, evolutionary pressure could have favoured the development of generic components that can be adapted to particular memory faculties, like for example synaptic filters with adjustable thresholds and decay rates that can adjust the learning rate qualitative and quantitatively. In contrast, the learning rate in stochastically updating synapses is explicitly modelled probabilistically and does not emerge from the structure of the model. Therefore, we find it difficult to imagine how very slow probabilistic learning rates could be actually implemented. Further, cascade synapses and synapses that stochastically express induction stimuli in general would be unsuitable to serve this dual role from development to memory as they are not designed to suppress fluctuations. Such synapses may only reduce fluctuations by assuming very low plasticity probabilities with the added side effect of unrealistically long periods for neural development (Elliott and Lagogiannis, 2009).

10.1.5 Biophysical locus of synaptic filters

Our abstract LP filter synapse or any theoretical device for that matter does not aim to describe the function of any particular molecular mechanism as this can be distributed in

the interaction of numerous molecular processes. It is known that potentiation depends on kinase activity (Malenka et al., 1989, Malinow et al., 1989) and that depression on phosphatase activity (Mulkey et al., 1993). In earlier work on neural development, it has been proposed that the calcium-calmodulin-dependent kinase II (CaMKII) (see Section 2.5) could provide for the integration of induction stimuli and also exhibit threshold type behaviour (Elliott, 2011a). This enzyme is highly expressed in the nervous system and can serve as a link between stimuli and synaptic plasticity (Lisman et al., 2002, Hanson and Schulman, 1992) while its phoshorylation state may represent filter states. Its phosphorylation and dephosporylation is mediated by induction stimuli through the actions of kinases and phosphatases while it also has the ability to sustain its phoshorylation state in the absence of plasticity stimuli. Phosphorylation of its subunits occurs in response to sufficient levels of Ca^{2+} concentration because calmodulin acts as Ca^{2+} sensor (for a review see Chin and Means, 2000). On the other hand, subunits can be dephosphorylated by the PP1 type phospatases (see Pi and Lisman, 2008, for an example of coupling phosphatase to kinase activity). Once a sufficient phosphorylation level is reached the kinase can autophosphorylate which allows it to become partially autonomous to act on its substrates for prolonged intervals in the absence of a sustained Ca^{2+} signal before its phosphorylation state presumably resets. The dodekameric ring structure of the CaMKII subunits would imply that realistic filter-state counts for the filter could be assumed to be in that range, if a direct mapping of our filter states is made to individual CaMKII subunits, however there is experimental data to suggest that the number of functionally coupled subunits is six (Rosenberg et al., 2005). Nevertheless, the collective properties of a population of these enzymes may be more relevant and thus number of states encoded may not directly map to the subunits (see Elliott, 2011a, for a discussion on the link between filter and CaMKII).

It is easy to assume that the decay rates introduced in the generalized LP filter model could represent the constitutive phosphatase PP1 activity levels in synapses (Elliott, 2011a). At the same time, we can hypothesize that the threshold for e-LTP could be represented by the autonomous self-sustained form of CaMKII. Supporting evidence indicates that CaMKII plays a crucial role in e-LTP of the CA1 area (see Fink and Meyer, 2002, Lisman et al., 2002, for a review) but also that it mediates the setting of LTP specific "tags" under the synaptic-tagging and capture hypothesis (Sajikumar et al., 2007). Further, the rapid reset mechanism could be implemented in the bistability of the CaMKII phosphorylation (see Graupner and Brunel, 2007) coupled with the ability of upregulating the PP1 decay once the filter enters its terminal state. Once the terminal state is reached an upregulation of decay could move the CaMKII operating point to the unstable steady-states that would result in a rapid return to the inactive state. Yet a direct mapping to state molecule of competition introduced between induction stimuli for the expression of either e-LTD or e-LTP cannot be clearly mapped onto separate states of a state-holding molecule in light of evidence that calcium signals are similar between STDP protocols that trigger either LTP or LTD in the somatosensory cortex (Nevian

and Sakmann, 2006). It appears that there may be two separate calcium sensors but nevertheless a competition for expression may still underlie the plasticity mechanism.

10.1.6 Filters as a new framework for synaptically stored memories

Overall, we find that synaptic filters can be applied to development and to ongoing memory processes. In retrospect this result is perhaps not surprising as low-pass filtering is a fundamental signal processing operation in communications systems. Therefore nervous systems, where information is encoded in temporal patterns of spiking activity, should employ these techniques so neurons can detect behaviourally relevant information in spike trains (George et al., 2011).

Here we have examined the encoding of a stream of equally important memories and we have assumed that this processes is based on early phases of plasticity. However, some memories have higher behavioural significance and it is therefore worth to examine if synaptic filters can be extended further to consider memory consolidation. Within a single memory system we can have particular memories that resist the effects of further memory encoding through the transition of encoding synapses to late-phase plasticity. Showing that synaptic filters can be extended to cover the transition to late-phase plasticity would further demonstrate that they offer a general and powerful framework to study diverse plasticity related phenomena. In the second part of this thesis we therefore asked how filter synapses could initiate memory consolidation by detecting significant memories from the pattern of stimulation.

10.2 Memory consolidation within SPM

10.2.1 Relevance of spacing effect to memory function

An efficient memory system should select relevant information for consolidation while allowing irrelevant ones to decay (McGaugh, 2000). There are probably multiple selection criteria to assist an animal in adapting its behaviour to the natural environment for long-term. One such criterion for long-lasting memories is the requirement for repetitive learning which is reproduced in behavioural training protocols of humans and animals. A general principle is that properly spaced inter-trial repetition intervals preferentially consolidate memories (Carew et al., 1972, Pinsker et al., 1973, Sutton et al., 2002, Beck et al., 2000, Tully et al., 1994, Melton, 1970, Hintzman, 1974). At the cellular level the spacing requirement for LTM finds parallels to the induction of late-phase synaptic plasticity (see Kornmeier and Sosic-Vasic, 2012, Litman and Davachi, 2008). Although the consolidation of memories appears as a general function of neural systems the particular form of the spacing effect is probably adapted to the constraints of the natural

environment an animal inhabits. For example, honey-bees need to form memories of foraging sites they discover but of all the sites they explore they should retain those that have been highly rewarding. Interestingly, the inter-trial intervals of the spacing requirement for LTM in the honey bee correlate with the flight times between patches of flowers when a rich foraging site is being harvested (Menzel, 1999). This particular tuning would allow honey bees to selectively store in LTM the particular properties of a rewarding flower site so it can be retraced in the future. In general, a requirement for a spaced pattern of stimulation for memory consolidation can be one of the mechanism that allow animals to behave with reference to persistent and salient information of the environment assimilated over long periods of time.

The STC hypothesis we reviewed in Chapter 2 is also a variant of the memory selection idea, it predicts that one memory may determine the relevance of another. At the cellular level electrophysiology experiments on the rat hippocampus have shown that the transient early form of LTP, which is induced by a weak stimulus, can be converted to the stable late LTP form; the time constraints of the interaction show that the weak stimulus should be applied within a time window of 1–2 hours before or less than 2.5–3 hours after a strong stimulation of separate pathways affecting different sets synapses of a neuron (Frey and Morris, 1997, 1998). The experimental results have been interpreted as the mechanism behind *distal reward* according to which one memory assigns importance to another that occurred close in time and consequently both become consolidated (see Frey, 2001, Päpper et al., 2011). Thus, at the level of memory function the theory predicts that a significant event causing strong stimulation on one pathway can be associated via consolidation to other events that occurred within a few hours and only weakly stimulated a separate pathway.

As in the case of plasticity and memory in general, addressing the physiological relevance of these cellular processes is difficult. Behavioural tagging experiments in rodents have been devised to show that STC is relevant to memory (Moncada and Viola, 2007, Ballarini et al., 2009, Wang et al., 2010). In these experiments a weak memory encoding episode, which induces short but not long term memory, can be consolidated to long-term if a strongly encoded memory which uses novelty exploration is encoded close in time with the weak memory. These experiments support that STC is relevant to memory but have also verified the crucial role of dopamine in establishing the strongly encoded memory which triggers consolidation (Wang et al., 2010). The usual caveat still stands however on whether complex behaviour can be attributed to individual memory systems and isolated neural circuits.

Although STC is a rather more complicated process of memory selection involving the interaction of two memory events, it still requires strong stimulation of one pathway by spaced repeated stimulation conjointly with dopaminergic activation (Huang and Kandel, 1995, Sajikumar and Frey, 2004a) in order to trigger l-LTP. Consequently, one

question arising within the SPM framework is whether individual synapses can independently detect the patterns of stimulation so as to trigger the transition to late-phase plasticity. Showing that it is theoretically possible for single synapses to detect the stimulation pattern raises the possibility that individual synapses are the computational unit handling memory consolidation as well as initial early forms of memory encoding. However, as we discussed in Chapter 4, models of l-LTP have directed efforts towards theoretical descriptions of the STC hypothesis mostly by explicitly modelling the experimentalist's view of the results while overlooking the question of stimulation patterns raised above.

In our study we initially focused on how synapses could respond to the stability versus plasticity dilemma in the absence of any neuromodulation within the early protein-synthesis independent phase of plasticity to encode memories of equal importance in one-shot. However, we contrasted our model against the cascade model of synaptically stored memory and expressed objections on how it attempts an overinterpretation of memory dynamics but also on whether the abstract hidden cascade states can represent the different molecular pathways associated with different phases of plasticity (Fusi et al., 2005, Kandel, 2009). We argued that modelling the interference of new memory encoding on previously stored memory should consider that biologically there appear to be mainly two forms of persistent plasticity. Hence, in a cascade model only the last few states could relate to persistent plasticity but there is no explicit consideration for the conditions that lead to the most stable plasticity form.

A proposal to extend the cascade model to account for spacing effect by introducing refractory times in the metaplastic transitions down the cascade states (Fusi et al., 2005) would appear to circumvent problems of generalizing this framework to l-LTP. According to this revision spacing is imposed by a purposefully built-in mechanism that requires a pause between repetition if a synapse is to be driven to lower cascade states. Such a modification however obliges further changes in order for the cascade to retain the mean expression times of each cascade state. The transition probabilities need to be recalculated to account for a period of refraction. In essence the decision of a synapse to change state in response to each induction stimulus in the cascade is a Bernoulli trial and a refractory period fixes the outcome of the first k of these trials. We could extend this to completely remove the stochastic component and have synapses that are totally refractory with deterministic expression times. Our results indicated that introducing the refractory revision may exhibit a reduction in capacity around 10%. Arguably this is not a significant reduction considering that this refraction was introduced to the extreme. Nevertheless, it shows that since the model was not built with signal detection in mind attempting such revisions has other consequences for the learning dynamics expressed that influence the behaviour of the model. Additionally, exploring the potential of a fully refractory cascade showed that a synapse with access to multiple refractory periods that completely remove the stochastic transition component

could still operate with comparable performance to the stochastic cascade therefore raising the question whether stochastic transitions in the cascade are justified given the relative small performance gains they offer.

10.2.2 Extending the filter framework to l-LTP

We postulated that filter synapses could be naturally suited to detect optimal storage conditions as these could emerge due to the interesting signal dynamics without imposing a refractory period. We chose to examine memory consolidation using the LP filter synapses since they displayed superior memory capacities to dual filter synapses. If repeating a memory at the natural timescale of signal peak significantly alters the statistics of filter threshold crossings compared to other stimulation protocols then individual synapses could detect the pattern of repetition and trigger late-phase plasticity in response. In Chapter 8 we examined how different repetition protocols influence the number of consecutive same threshold crossings (threshold-cycles) a synapse experiences after the initial encoding of the repeated pattern. The repetition protocols examined differed in the number of repetitions n_r and in repetition intervals T_r. In line with the earlier filter framework, where memories only decay due to further memory storage, time is discrete and thus repetition interval T_r represent the number of memory patterns stored in between repetitions. Our results showed that on average if a synapse experiences the same repetition protocol a large number of times the distributions of threshold-cycles between protocols differ but nevertheless overlap. Examining the convergence of the accumulated threshold-cycle distribution towards the empirical distributions of either a massed or spaced protocol showed that the number of samples required for detection ranged from tens for small filters to hundreds of samples for larger filters.

Taking the electrophysiology protocols literally would mean that a filter synapse experiencing a spaced repetition protocol of $n_r = 4$ repetitions would probably require hundreds of stimulation trials before it could detect a spaced protocol being induced. A more careful evaluation of the electrophysiology experiments however would suggest that the 100 pulses delivered at each repetition by the 1 second tetanus probably induces multiple filter state transitions and not just a single-step transition. We found that a protocol which induced multiple transition per repetition effectively reduces the filter size and therefore the convergence time. However, even under multiple-step encoding the distributions of massed and spaced protocols overlap and therefore tens of samples are required before the accumulated threshold-cycle distribution can be identified.

One key difference found between the distributions of the two protocols is that a spaced protocol allows synapses to conduct opposite threshold crossings in between repetitions. This is clearly seen under multi-step protocols where spaced protocol distributions have a distinctive negative threshold cycle occupancy. In contrast, massed protocols do not allow any variability as the repetitions being delivered close in time create a long train

of stimuli that deterministically drive synapses through the desired threshold only. Exploiting this fact as a signature of variability in the stimulation protocol for detection purposes is however difficult, as variability and therefore negative threshold cycles occur very frequently under no repetition too for example. Thus, a synapse would require to first detect or be told that a strong multiple-step protocol is being induced and then check for variability to safely determine the protocol being induced. In any case, even if we were to use the negative cycle for protocols with many repetitions, it is not clear how strict the requirement for a spaced protocol should be for initiating stable plasticity and thus enforcing such a constraints would be restricting a model to a confined set of data that has not fully explored the trade-off between number of repetitions and repetition intervals.

Our results demonstrate that it is theoretically possible for single filter synapses to identify the protocol being induced since different protocols change the statistics of threshold-crossing behaviour. However, this would require synapses to accumulate statistics from at least tens of protocol delivery trials. Thus, the integrative dynamics of filter synapses are not sufficient to interpret results drawn from electrophysiology experiments showing that even a single trial of an $n_r = 4$ spaced repetitions is enough to elicit l-LTP or LTM.

10.2.3 Filter signal dynamics to the rescue

Some experiments have indicated that the level of postsynaptic depolarization is crucial for l-LTP induction (Sjöström et al., 2001, Dudek and Fields, 2002) or that increased neuronal excitability preferentially allocates memories (Zhou et al., 2009). The electrical activity of a neurons can be directly read as a chemical signal involving transient changes in Ca^{2+} concentration. These occur via various mechanisms either involving voltage gated calcium channels (VGCC), entry via ligand-gated ion channels like NMDA channels or release from intracellular stores (see West et al., 2001, Chetkovich et al., 1991). Within the context of memory the magnitude of postsynaptic depolarization caused by presenting a previously encoded memory is representative of the preserved signal trace.

Building on the idea that synapses could utilize this information to detect storage conditions we extended our model synapses to consider a parallel signalling pathway that integrates signal samples taken at the time of memory repetition. Our l-LTP model combines a filter with an abstract molecular pathway model in a way that the dynamics of the pathway are ultimately influenced by the dynamics of the synapse model's signal. The molecular pathway mechanism we proposed contains a leaky integrator u of signal samples at times t_r. The output of the pathway is taken to be the total activity of u measured as its integral in continuous time by a quantity $R^\star$. The model assumes that the pathway is activated under sufficient postsynaptic depolarization and therefore the signal samples are passed from a threshold before reaching the input of a leaky integrator. This threshold is implemented by a non-linear sigmoidal response function which

largely suppresses low depolarizations. To implement a minimum repetition requirement our threshold levels are set to the asymptotic value of the signal obtained after regularly spaced on-peak memory repetitions. Due to the dynamics of the filters, these levels are reached after almost $n_r = 4$ spaced repetitions thus providing an interesting explanation as to why usually such few repetitions are required to induce l-LTP in experiments.

Arguably, we could have enforced a minimum spacing requirement by adding a slow-filter between $c(t)$ and $\mu(t)$ instead of the direct relationship of Equation (9.1). However, such an approach would have enforced a spacing requirement in an attempt to fit our model to a relatively small amount of data that has not fully explored the relationship between number of repetitions and repetition intervals for LTM. In contrast, our simple mechanism utilizes properties of the filter signal so that the output $R^\star$ is maximized under spaced repetition protocols by employing components of a well-known biochemical pathway involved in late-phase plasticity. For instance, on average a memory encoded for the first time will result in the sampling of a zero signal while a memory repeated at the time of signal peak for the first repetition will sample a signal size of approximately β/Θ^2 (the average peak signal size). Thus, if we sampled the filter signal at the peak at every repetition this should maximize the overall signal integrated. Setting a simple threshold on the total output $R^\star$ showed that it was sufficient to preferentially detect minimum number of spaced rather than massed protocols, at least below a limited number of memory repetitions $n_r < 5$. Moreover, results obtaining the size of the allocated signal under $n_r = 4$ repetitions for various repetition intervals showed that the local filter statistics counting the number of same threshold crossings were not required to obtain a spacing effect. A single correct threshold crossing that tagged synapses for allocation was sufficient to produce an allocated signal standing above the noise ($SNR \approx 10$). Therefore, our results show that local synapse-specific filter information indicating the protocol being induced through statistics of threshold crossings can be ignored. The form of the protocol can be derived from spatio-temporal filtering of the global information provided by the postsynaptic depolarization signal alone.

Here we have therefore shown that a simple system that integrates signal samples could result in a spacing effect due to the dynamics of filter signal. We note, however, there are limitations with this approach due to how the signal behaves under memory repetition. On the one hand, as the number of repetitions increase maximizing the total signal sampled requires protocols with smaller regular T_r repetition intervals. The consequence of this fact is that beyond a few repetitions $n_r > 4$, massed protocols give higher total sampled signal than spaced protocols. Therefore, relying simply on the filter dynamics to give the spacing effect only works for protocols of few repetitions, beyond which massed rather than spaced protocols become more effective. There is also the option of inducing a protocol to hit the apparent signal peak we discussed in Chapter 9. Although this strict on-peak protocol would be near optimal across n_r we would have to progressively reduce the repetition intervals to end up with a massed protocol as n_r increases and

thus spacing interval condition is broken. Under increasing numbers of repetition we find that the total signal sampled with the progressive interval protocol (F.P) is the same between massed and spaced protocols.

10.2.4 Well-known signalling processes could underlie the filter's extension

Biophysically the basic components of the well known cAMP-pathway could be well suited to perform the above function. The cAMP-pathway has been shown to be critically important for protein synthesis dependent LTM (Abel et al., 1997, Abel and Nguyen, 2008, Frey et al., 1993, Huang and Kandel, 1994) and it appears that PKA is selectively recruited in response to spaced tetanization (see Woo et al., 2003). Initiation of the cAMP-pathway could result from the membrane bound enzyme adenylyl cyclase (AC) which synthesizes cAMP in response to sufficient Ca^{+2} concentration. There are two forms of AC which are calcium-stimulated and have been shown to play a critical role in long-term memory function of mammals and invertebrates (Wong et al., 1999, Livingstone et al., 1984, Shan et al., 2008, Zhang et al., 2008, Wu et al., 1995). AC synthesizes cAMP and the Ca^{+2} regulation of this activity is achieved through bound calmodulin (CaM) to these AC enzymes. In our models, the synergistic action of neuromodulators via G-protein is presumed to occur concurrently under a single biochemical step for the activation of AC occurring in response to each synaptic stimulus. The cooperative binding of Ca^{+2} to CaM can be modelled by a sigmoidal function (see Appendix 3) and therefore cAMP production is stimulated every time Ca^{+2} exceeds a threshold. Further, the cAMP synthesis and degradation dynamics can make the concentration of cAMP represent the leaky integrator which activates PKA. Since the concentration of cAMP activates PKA, the dynamics of PKA would appear to be integrating cAMP, thus giving the output quantity $R^\star$ of our model. Finally, reaching a sufficient level of activated PKA could define the point where local protein synthesis is initiated which selectively affects tagged synapses in a manner similar to the STC hypothesis.

10.2.5 Is the spacing requirement strictly enforced?

The question arising next is how does a memory system respond under an increasing number of repetitions. Should it allocate memories that are repeated a large number of times in small time-frame or does a spaced pattern of repetition signify something very special about the environment and thus memory systems should be specifically tuned to this pattern? To our knowledge there are no experimental data that could provide a definite answer. A major part of parameter space on differences between massed and spaced protocols under various numbers of repetitions n_r remains unexplored (Kornmeier and Sosic-Vasic, 2012). On the other hand, we may draw upon indications from sparse

experimental results. For example, Yin et al. (1995) report that 48 massed training sessions in Drosophila aversive olfactory learning where insufficient to induce LTM while a spaced training protocol of the same total duration consisting of 10 session with a 15 minute rest interval where effective at inducing seven day LTM. Maximum LTM was achieved at $n_r = 10$ with a progressive improvement as the number of repetitions increased while further repetitions ($n_r = 15$ or $n_r = 20$) did not improve memory retention further. Furthermore, the locus of the Drosophila aversive olfactory LTM is distinct from earlier memory phases and it appears to specifically respond to spaced repetitive protocols (Isabel et al., 2004). Therefore, at least in the Drosophila learning model we obtain an example of a cAMP-pathway mediated LTM which is particularly tuned to a spaced encoding protocol.

Answering questions around the relationship of spaced and massed training protocols is admittedly difficult at this point and probably there is no general principle. LTP can be expressed under massed or spaced protocols but the cAMP-pathway and thus PKA mediated allocation seems to preferentially respond to spaced patterns at least in some animal models of learning. Therefore, it may well be that multiple pathways for plasticity exist each processing particular forms of stimulation while some have to integrate multi-modal input combining information arriving as neuromodulation (see Shohamy and Adcock, 2010). Under this hypothesis the particular processors online in neurons probably depend on the function of the particular memory system they belong to and it may well be that the cAMP-pathway can be adapted to fit particular stimulus patterns where such patterns are functionally meaningful. In this context we could then postulate that some pathway is tuned to specifically process spaced patterns of stimulation and thus under this hypothesis we may proceed to further investigate its form. However, although experiments use regular fixed interval spacing this may not be optimal in our model since a strict on-peak protocol under regressing intervals appears to be optimum across n_r. Under sufficient repetitions the intervals in such a protocol would end up being similar to a massed protocol thus making interpretations difficult. Hence, we adhered to a fixed interval spacing because this pattern is simpler and the results obtained are easier to interpret through a qualitative comparison against the experimental literature.

10.2.6 Saturation can enforce a spacing effect

We extended our simple model by introducing saturation to the leaky-integrator to remedy the problems arising from increasing the number of repetitions. Biophysically, saturation could represent active regulatory mechanisms or other rate limiting steps in signal transduction pathways. Saturation converts the previous mechanism of signal sample integration to a low-pass filter of stimulus repetition. The cut-off frequency is set by the rate of decay of the leaky-integrator. Massed repetitions that are close to

the saturation level add sub-linearly while with an appropriate rate of decay spaced repetitions allow each signal sample to linearly increase the output $R^\star$ by some fixed amount.

Our results showed that a saturation mechanism is sufficient to express a spacing effect. Under high n_r the $R^\star$ output of regularly spaced protocols can double compared to massed protocols. Although spaced protocols are more efficient the saturation mechanism does not occlude the possibility that massed protocols initiate l-LTP. For some fixed allocation threshold on $R^\star$, increasing the number of repetitions can produce sufficient output to exceed this threshold even under massed protocols. Nevertheless, regardless of the number of repetitions spaced protocols will remain more efficient due to the action of saturation. This model predicts that with enough repetitions (n_r) there would be sufficient PKA activated within a neuron, or in one of its compartments, to initiate protein synthesis. Experimental evidence to support the possibility of PKA mediated LTM under massed training have been reported.

Contrary to Drosophila learning where massed training was insufficient to elicit LTM and previous work on the Aplysia gill-withdrawal reflex LTF, more recent experimental work has provided evidence that a single massed training session could activate PKA and protein synthesis but in different learning mode examining *aversive feeding behaviour* (Michel et al., 2012). Interpreting these phenomena in our model would require to either lower the position of the $R^\star$ threshold for allocation or raising the presumed number of n_r massed repetitions induced during a single training session.

10.2.7 A synaptic integration framework for the initial and late-phases of plasticity

We have demonstrated that biophysically plausible integrating mechanisms can constitute a combination of low-pass filters operating first at the level of single induction stimuli arriving at synaptic terminals and then at the level of memory patterns received by neurons. The combined two simple models reproduce the spacing effect of memory allocation but also provide insights on the role of each integration unit in early and late forms of plasticity. Our results show that appropriately filtering global signal dynamics is sufficient to reproduce a spacing effect while incorporating local statistics has limited effects. Therefore, a spacing effect in memory consolidation was present even if synapse specific information was limited to a simple "tag" indicating a synapse has expressed plasticity over the threshold dictated by a given memory. There is evidence to support that single synapses can induce l-LTP under massed repetitive encoding if the cAMP pathway has previously been pharmacologically stimulated. Govindarajan et al. (2011) used a glutamate uncanging protocol to stimulate single spines of proximal apical dendritic branches at CA1 pyramidal neurons with a train of thirty pulses delivered very close in time (0.5Hz). These experiments showed that l-LTP could be expressed at single

spines if the cAMP-pathway had been previously stimulated via forskolin or via a DA receptor agonist.

Although, the synapse specific information may be limited to a "tag" we have assumed that this is set when a "correct" threshold is crossed. The question raised is how the synapses in our model determine which threshold is the "correct" one for the memory being allocated. In the absence of any salient extra signals dictating which threshold is correct this does not seem possible. If setting the "tag" required neuromodulation we could assume that the desired strength state is set during initial encoding of a memory being strongly encoded. However, our earlier discussions on plasticity have indicated that the expression of e-LTP and the setting of "tag" does not require neuromodulator activity. Thus, our requirement that "tags" are set over the "correct" threshold may be excessive but nevertheless our choice of how a "tag" is set is not critical. Freezing the state of all synapses that expressed plasticity in response to the last encoding when $R^\star$ exceeds the critical threshold for allocation would suffice to consolidate the desired pattern. Thus "tags" in our model could well be set at the time of a filter-threshold-crossing and expire within some time-frame in accordance with other models of STC.

Despite these concerns, the results reported by Govindarajan et al. (2011) give further support to our model as they indicate a role of sufficient depolarization for l-LTP. The researchers showed that single spine stimulation could only induce l-LTP if the Mg^{+2} block from NMDA receptors was absent. In physiological conditions removal of the Mg^{+2} block requires sufficient postsynaptic depolarization. Eliciting l-LTP in the presence of Mg^{+2} required stimulation of proximal multiple spines along with the usual forskolin requirement for stimulating cAMP production. These findings imply that there is a threshold of synapse activation below which L-LTP induction does not occur, a result which agrees with our model's predictions.

10.2.8 The nature of the saturation mechanism

Besides some parallels to biology, a particular biophysical implementation of saturation is not suggested by our model. Although the simple integration of signal samples can be easily attributed to the initial components of the cAMP pathway, the saturation mechanism is a completely theoretical construct. Nevertheless, our simple formal description of saturation gives rise to new hypothesis on functional and mechanistic aspects of the signalling cascade for l-LTP.

It could be embodied in a single or multiple feedback mechanisms regulating the signalling pathway from the initial cAMP synthesis at ACs to the maximum PKA activation rate allowed at any time. Alternatively, depletion of a renewable necessary component involved in the signalling pathway could also reflect a form of saturation as it would

require a period of inactivity before its availability is re-established. Thus, in the absence of saturation the signal sample integrating mechanism could well relate to the well known cAMP-pathway, but there may be problems if saturation is taken literally. For example, the model relies on the slow dynamics of the first quantity in the signalling cascade u to set the minimum spacing interval. For the optimal repetition interval to be around the filter peak then the decay rate should be slow enough so to ensure sub-linear addition of shorter intervals. Consequently, if the biophysical implementation of our model relates to the cAMP-pathway initiated by stimulating AC then arbitrary slow dynamics of u may be unrealistic and there exists experimental evidence to suggest so. Recent advances in fluorescent cAMP probe technology have provided the opportunity to assess real-time changes in cAMP with unprecedented temporal sensitivity (Nikolaev et al., 2004). Essays on the response of calmodulin-stimulated AC (AC8) to produce cAMP in response to induced Ca^{2+} oscillations showed that cAMP transients had a peak frequency of three per minute, while the cAMP signalling system acted as a low-pass filter under higher frequency Ca^{2+} oscillations (Willoughby and Cooper, 2006). Therefore, AC is quite sensitive to Ca^{2+} rises and it is difficult to assume that cAMP decay can be prolonged to the extend required by memory spacing effects (i.e 15-40 minutes in Drosophila). The fast transients of cAMP have been shown to be mostly due to an active feedback mechanism arising through the hydrolysis of cAMP from phosphodiesterases (PDE4) activated by PKA (Willoughby and Cooper, 2006).

Consequently, the mechanism we proposed perhaps offers a simple conceptual framework showing how spacing phenomena may arise under a simple series of biochemical processes but the actual biophysical system appears to be operating somewhat differently perhaps due to other constraints involved with containing cAMP signals. Nevertheless, we showed that a signalling cascade expressing saturation produces sub-linear augmentation of responses for short-interval repetitions thus favouring longer repetition intervals. Addressing the biophysical nature of a saturation mechanism within the cAMP-pathway would require more elaborate models. For example, in the biophysical model by Smolen et al. (2006) saturation arises by assuming that a precursor protein required for l-LTP is depleted under successive stimulation.

If our proposed signalling cascade relates to the cAMP-pathway then perhaps we could look for a negative feedback mechanism in the phosphodiesterases that hydrolyse cAMP. These limit the diffusion of cAMP gradients so as to organize them into micro-domains (Zaccolo and Pozzan, 2002).

10.3 Model implications for experimental tests

In this thesis we have proposed that synaptic filters provide a meaningful mechanism for synapses to solve the synaptic plasticity versus stability dilemma and have demonstrated that filters also increase the theoretical storage capacity compared to cascade models or stochastic updating synapses within some biologically relevant limits on the number of synapses $N < 10^5$. The prospect that synapses bestowed with the task of encoding a stream of memories in one-shot utilize a synaptic plasticity filter therefore becomes appealing. If filter synapses handle early memory encoding the question that was subsequently raised is whether these filters have the ability to detect particular storage patterns known to lead to memory consolidation. Our examination in Chapter 8 suggested that under a few repetitions the ability of synaptic filters to detect stimulation patterns through threshold-crossing statistics is limited. We then considered the possibility that a parallel signalling mechanism which read the postsynaptic depolarization (PD) could be used to detect optimal storage conditions.

This very fact introduced a divide as the contribution of local synaptic filter information was minor compared to the parallel mechanism taking advantage of PD. In essence we could do away with local filter statistics that could indicate when a spaced repetition pattern is being administered and still get a spacing effect relying on the mechanism using global information from the PD. Nevertheless, for a limited number of memory repetitions ($n_r < 5$) the proposed mechanism for memory allocation exploits the filter signal dynamics to detect spaced patterns up to. For a larger number of repetitions the synaptic filter dynamics are not effective at reproducing the optimal storage conditions of spaced repetition patterns and these have to be enforced by the parallel signalling mechanism through an embedded saturation in response. Thus, here the divide has been introduced and essentially two mechanisms, one for solving the stability versus plasticity dilemma and the other for detecting optimal storage conditions, can be searched for independently.

10.3.1 An experimental signature of synaptic filters

First and foremost we require experiments to discern if hippocampal synapses integrate induction stimuli before expressing plasticity as the filter framework suggests. Here, we reproduce proposals for experimental protocols published in Elliott and Lagogiannis (2012) and outlined for the cascade in Elliott and Lagogiannis (2011a).

Most of the experimental data we have presented involve whole pathway stimulation using HFS or LFS stimuli to induce LTP (Bliss and Lømo, 1973) or LTD (Lynch et al., 1977). The number of induction stimuli delivered to synapses under such protocols cannot be controlled and given their frequency and duration these may infact be saturating synapses (Elliott, 2010a). Furthermore, to gauge the amount of expressed plasticity

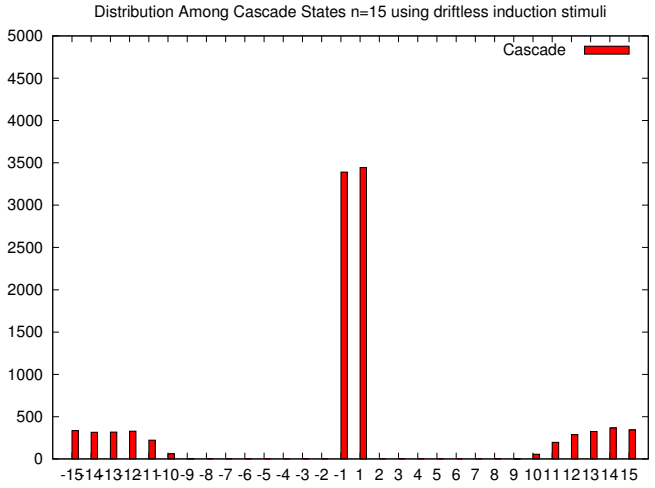
these electrophysiology protocols measure fEPSP which measures the activation of the postsynaptic targets under the average behaviour of a very large population of synapses and not that of individual synapses. Avoiding these issues leads to consider rather more delicate experimental procedures at the level of single synapses by using spike-pairing protocols to induce STDP (Markram et al., 1997, Bi and Poo, 1998). Under such protocols a single spike-pair can provide a single plasticity inducing stimulus and by separating spike-pairs in time to avoid any higher-order spike interactions (Froemke and Dan, 2002) we may produce controlled trains of plasticity induction stimuli. We propose experiments operating at this level to tightly control the statistics of the train of induction stimuli perceived by a synapse and establish qualitative criteria for detecting integrative synapses against stochastically updating ones.

We use the same approach as in Chapter 6, to examine the change in expression time against changes in the statistics of the induction stimuli. Assuming we can obtain an appropriate random number generating source we let the statistic of the stimuli train be characterized by the probability that the next stimulus (POT or DEP) is identical to the previous one $\text{Prob}[I_2 = \pm I_1] = (1 \pm C_2)/2$ and let parameter C_2 move between two extremes, with $C_2 = -1$ the train of stimuli will toggle deterministically between POT and DEP and with $C_2 = 0$ return to random uncorrelated stimuli.

We have already shown that an LP filter mean expression time is modified in response to changes in the statistics of the stimulus train (see Section 6.3.2). In summary, a homogeneous train of stimuli will drive towards quicker expression than a mixture of induction stimuli because all single induction step will be directed towards the same threshold. In contrast, with driftless induction stimuli there will be on average no change in synaptic strength because the synapse will alternate between two internal filter states and thus the mean time to reach threshold will be infinite.

Previously we showed that a stochastic synapse has a constant expression time regardless of the statistics in the stimulus train. Here, we propose a slightly different experimental protocol to detect if synapses operate as cascade, which still at its core, it contains SU (Elliott, 2011a, Elliott and Lagogiannis, 2012). A cascade model is still based on Bernoulli trials but the expression time of plasticity can be delayed further if a synapse experiencing a train of induction stimuli is moved deeper in the cascade towards less labile states. In the first instance, extending expression times in response to a particular succession of induction stimuli may appear similar to a filter but in reality it is very different. A filter delays a decision for plasticity expression when the induction train shows no clear trend towards either plasticity direction. The LP filter for example will not express plasticity at all under completely driftless ($C_2 = -1$) induction stimuli when POT and DEP signals are continuously alternated. In contrast, driftless stimuli make all cascade synapses move to the most labile state (see Figure 10.1(a)). This will occur regardless of the state a synapse is initially found in because alternating stimuli will make each SU of any cascade state eventually express plasticity. Once plasticity is expressed

the cascade model moves synapses to the most labile states where the mean number of induction stimuli for plasticity expression is minimum.

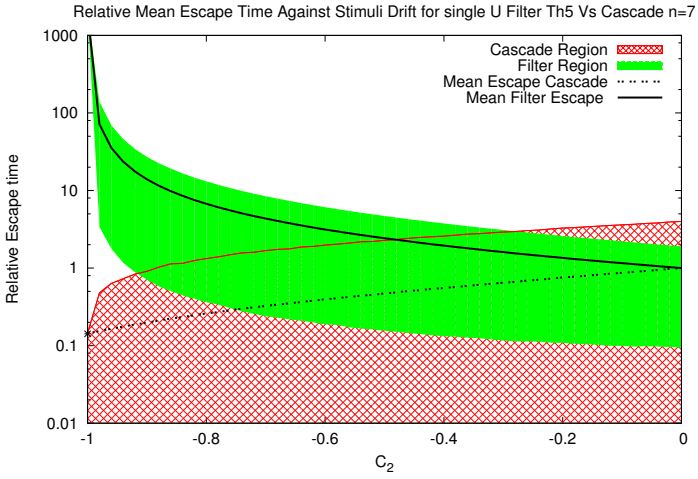


(a) Distribution among cascade states using driftless stimuli

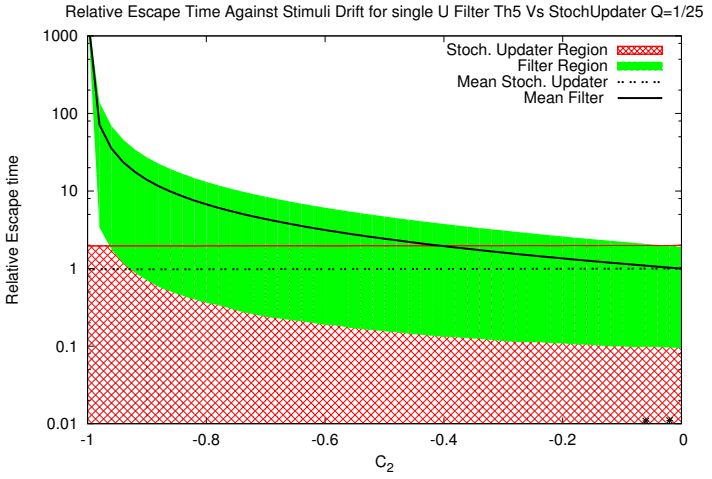
FIGURE 10.1: **a**. Using driftless stimuli that continuously toggle between POT DEP breaks the uniform distribution. This distribution is for shows a snapshot of the distribution after 1500 iterations continuously encoding a memory and its reverse. This effectively toggles between POT/DEP stimuli at every synapse eventually moving all synapses to the most labile state.

Thus, unlike LP filters where plasticity expression would be suppressed, alternating POT and DEP stimuli for a sufficient amount of time would make cascade synapses toggle between strengths and thus minimize the expression times relative to the one obtained under random uncorrelated stimuli. Figure 10.2(a) shows results from a simulation comparing the relative expression time of a cascade and LP filter synapse confirming that under $C_2 = -1$ where stimuli alternate between POT/DEP the cascade synapse's mean expression time is minimized while that of an LP filter synapse goes to infinity.

Although theoretically these experiments appear simple, in practice *single* synapse experiments are not actually easily implemented with the accuracy required here. On the one side standard electrophysiology methods that measure synaptic strength via current probes that keep postsynaptic voltage fixed could be regularly used to asses when synaptic strength changes occur. On the other hand, using these methods to generate plasticity stimuli via spike pairing could be hindered by synaptic transmission losses that could alter the assumed statistics of the spike train thus affecting results. Therefore, standard electrophysiology methods previously used at single CA3-CA1 synapses (see Petersen et al., 1998) will not suffice. Glutamate uncaging protocols avoid transmission failures of chemical synapses caused by the stochastic nature of vesicle fusion by directly releasing glutamate at the postsynaptic density. It appears then that these methods can



(a) Cascade Vs Filter synapse on relative escape times



(b) Stochastic updater Vs Filter synapse on relative escape times

FIGURE 10.2: The escape time of a $\Theta = 5$ filter is compared against a cascade of $n = 7$ under a various stimulus drift regimes. C_2 denotes the probability that the next stimulus (POT or DEP) is identical to the previous one $Prob[I_2 = \pm I_1] = (1 \pm C_2)/2$. When $C_2 = -1$ the train of stimuli toggles deterministically between POT and DEP. The plotted escape times are relative to the $C_2 = 0$ case where POT or DEP occur with equal probability regardless of the previous stimulus. A clear difference between filter and cascade model is seen as the induction stimuli used approaches the driftless regime $C_2 = -1$. The cascade mean expression time drops while the filter's increases to infinity. The same filter is also compared against a matched stochastic updater which has a constant escape time regardless the induction stimulus. This result is trivial, the stochastic updater expresses 1/25 stimuli randomly without any memory of previous events.

safely deliver presynaptic stimuli that can be coupled with direct postsynaptic depolarization to provide the desired spike-pair statistics (Bagal et al., 2005).

10.3.2 Mechanisms of late-phase plasticity and consolidation

We have extended the filter based framework to consider how the optimal storage conditions for memory consolidation arise and found that the dynamics of the signals could support a spacing requirement for a limited number of stimulus repetitions. Indeed, most experimental results that exhibit a spacing requirement for late-phase plasticity compare the minimum number of required spaced repetitions, with usually four repetitions being sufficient, against the same number of repetitions delivered with very short repetition intervals.

We have consider how our filter framework could enforce spacing effect beyond a few memory repetitions based on the hypothesis that spacing stimuli is a strong requirement for some forms of memory consolidation which is also reflected in the requirements for the induction of late-phase plasticity. However, there are limited experimental data that could reveal the relationship between the number of repetitions (n_r) and the repetition intervals (T_r) in the induction of l-LTP. A strong spacing requirement is seen for example in the Drosophila behavioural aversive olfactory memory experiments we mentioned earlier where 48 massed training sessions were unable to induce LTM (Yin et al., 1995) but linking these behavioural results directly to synaptic plasticity is rather risky as system wide functions may be at play. We therefore suggest that experiments should be designed to systematically search the parameter space of optimal storage conditions for various repetition protocols in eliciting l-LTP. These should not be constrained to fixed regular repetition intervals and should explore the potential of advancing or precessing repetition protocols as well. We have already discussed the complexities involved with interpreting results from classic electrophysiology protocols of LTP and LTD. Here, the requirements for extended stimulation by massed protocols may introduce even further constraints due to transmitter depletion in synapses and therefore again glutamate uncaging protocols appear indispensable. Govindarajan et al. (2006) have already provided evidence that in the absence of direct cAMP production stimulation via forskolin single spine stimulation via glutamate uncanging is insufficient to elicit l-LTP. If cAMP-pathway plays a crucial role in the spacing requirement, as our model suggests, then the experiments need to be conducted in the absence of forskolin. However, we have discussed the action of neuromodulators in gating l-LTP and how glutaminergic and neumodulatory signals converge to synergistically activate AC that produce cAMP (see Chapter 2). In our model of l-LTP we have assumed synapses are simultaneously stimulated by neuromodulator and neurotransmitter to stimulate signalling pathways necessary for l-LTP. Therefore, such experiments should be conducted by delivering pulses of DA and glutamate simultaneously to stimulate the production of cAMP in waves.

Here, the authors used imaging to detect spine enlargement as an indicator of l-LTP and it is unclear if that influenced results. If multiple spine stimulation is required for l-LTP, as Govindarajan et al. (2006) demonstrated, then the use of electrophysiology, we

suggested in the previous section, would require specifically tracing and stimulating the afferents stimulated. Hence, electrophysiology methods may be impractical in measuring synapse specific stable late-phase strength changes. We would therefore have to reside to the same spine-head enlargement imaging methods.

If a spacing requirement is revealed even under a larger number of repetitions then we could begin searching for the mechanisms involved by examining the relationships of the administered protocols with the activation of PKA. In our model we have suggested that the leaky-integrator u could map to the early c-AMP pathway components but we have also introduced a saturation component to enforce spacing for protocols of larger repetition counts. As we have previously discussed, saturation is computational tool that could be implemented by multiple molecular machines. A direct mapping of our simple saturation l-LTP model to the cAMP-pathway would suggest that the maximum concentration of cAMP is tightly controlled and therefore experiments should specifically focus on probing cAMP concentrations. Such experiments however would probably fail to address the general hypothesis if saturation enforces spacing as saturation could be located in components downstream of cAMP.

As shown in our results, simple saturation can only ensure that spaced protocols are more efficient than massed, but as the number of repetitions increase massed protocols could also sufficiently activate a signalling pathway to initiate l-LTP. Thus, having established the minimum number of repetitions required by an optimally spaced protocol, the effects of a reduction in the repetition interval (T_r) should be rescued by an increase in the number of repetitions. If this relationship is established then this would indicate that some form of saturation is taking place where a reduction caused by decreased output when operating close to saturation can be overcome by increasing the number of stimuli to increase the overall output.

10.4 Future work

Synaptic filters provide a new framework on which to base further research relating to plasticity and memory phenomena. We have already described extensions in this thesis, first by introducing a passive decay η which stochastically returns filter states to zero, and second via an extension to address the transition to late-phase plasticity and the spacing effect.

Despite these extensions there are still theoretical points concerning synaptic filter and memory capacity that could be further pursued. First, in this thesis we focused against a comparison to the cascade model but in terms of capacity the linear multi-state model we saw in Section 4.4.3.1 outperforms the cascade (Barrett and Van Rossum, 2008, Leibold and Kempter, 2008). Therefore, it would be interesting to extend our capacity analysis to

a comparison against such linear multi-state models that have similar transition matrices to filters but exclude a reset at the terminal state.

Further, beyond measuring the response of a single neuron to consider recurrent network dynamics where retrieval is more difficult. Here, for retrieval a pattern needs to sustain itself in the neuronal dynamics, after the initial cue stimulus is removed. We have already obtained indicative Hopfield (Hopfield, 1982) recurrent network results that have confirmed the superior capacity of synaptic filters to stochastic synapse models. However, Huang and Amit (2011) use a recurrent network with binary patterns $\xi \in \{0, 1\}$ (instead of our $\xi \in \{-1, 1\}$) and define capacity as the expected number of retrieved patterns over all ages (Amit and Huang, 2010). A coding level sets the number of neurons participating in each pattern being encoded while a threshold needs to be adjusted so to minimize the number of non-encoding neurons from firing. The authors report that multi-state models provide only small increases in capacity when compared against a basic two-state learning model, mainly due to a requirement for high initial signal which is compromised in multi-state models by slow learning. Given that our filter synapses relax this requirement it would be interesting to further compare our model in this setting.

Our l-LTP models have focused on reproducing stimulus pattern selectivity by maximizing the global signal for l-LTP under spaced repetition patterns and demonstrated that a simple saturation mechanism can give rise to the spacing requirement in stimulation patterns Despite the fact that these are not models specifically addressing STC phenomena they nevertheless hold the basic components that could provide for synaptic capture from a weakly stimulated synapse. Stimulated synapses can be tagged and a global signal is set that signal's the conversion of tagged synapses to stable long-term forms of plasticity. Hence, sufficiently strong stimulation of one synapse can initiate the global signal that is subsequently captured by weakly stimulated synapses. To describe heterosynaptic interactions of weakly and strongly activated synapses first we should assume that the $R^\star$ is only increased via strong stimulation (coupled with DA) and then we need to introduce a decay in $R^\star$ activation so that we create a time-window for interaction. In contrast to other models where the lifetime of the "tags" is explicitly modelled via passive decay, in our model these are associated with the lifetimes of threshold-cycles; an opposite threshold transition would invalidate a tag. Therefore, the extension of our model to l-LTP opens the prospect of considering STC phenomena albeit by very simple modifications.

Another interesting extension would be to explicitly consider the coupling of neuromodulation and synaptic activity for the induction of late-phase plasticity. It is well established that l-LTP is dependent on dopamine receptor activation (Frey et al., 1990, Swanson-Park et al., 1999, Navakkode et al., 2007, Sajikumar and Frey, 2004a). We have already discussed a crucial molecular mechanisms for the synergistic action of neuromodulator and neurotransmitter operating at the level of AC enzyme activation to

initiate cAMP signalling and late-phase plasticity. This pattern is found in all of the three animal models of learning we discussed, but we did not explicitly model how the converging signals interact.

The behavioural experiments reviewed in Section 2.5.2 revealed that the time window for interaction appears to be quite broad, enabling neuromodulatory signals to indicate the significance of a memory either protectively or retroactively as a late reward signal. Moreover, neuromodulatory signals have also been associated with gating or modulating the threshold for STDP (see Pawlak et al., 2010) and thus the phenomenon we are focusing on is not specific to LTP/LTD but can be traced to more tightly defined spike-to-spike interaction plasticity protocols.

Phenomenological models of STDP have attempted to capture this late interaction of spikes with neuromodulation by either assuming that sustained neural activity of a few seconds is able to link a CS and US in conditioning experiments (Drew and Abbott, 2006), or by postulating the existence of slow synaptic processes that acts as decaying "tags" to identify eligible synapses (Izhikevich, 2007). According to the later model, plasticity is only induced in response to an interaction of a "tag" synapse with a neuromodulatory signal. Other models have focused on describing the biochemical reactions for the synergistic activation of AC by transmitter and neuromodulator in order to model the dynamics of cAMP production which is assumed to lead to synaptic plasticity Yarali et al. (2012). Specifically, Yarali et al. (2012) attempt to capture the dynamics of Ca^{2+}/calmodulin-sensitive AC regulation by letting Ca^{2+} accelerate reaction rates of either activation or deactivation of AC in order to bimodally regulate cAMP production depending on the relative timing of the CS and US. This chemical reaction model explores the idea that AC can account for the late interaction of CS and the US in associative learning such as the Drosophila odour conditioned response learning we reviewed in Chapter 2.

Given the importance of this cellular mechanism for reward learning it would be interesting to extend our model of l-LTP presented in Chapter 9 to explicitly address the conjoint action of neuromodulatory signals. The extension could focus on providing a further level of integration to account for the synergistic action of neuromodulators in controlling the activity of the enzyme that produces the $u(t)$ signal. Adding state dependent integration of neuromodulation and neurotransmitter would allow the two signals to interact in time. Thus, we may then investigate how reward signals and memory encoding could interact to activate $R^{\star}$ and allocate particular memories to provide a simple model of reward learning.

10.5 Closing summary

In this thesis we focused on synaptic plasticity for memory systems that are continuously faced with new memory encoding and are required to selectively consolidate some memories upon detecting optimal storage conditions. Memory is studied at multiple levels from cognitive psychology to molecular biology but we have yet to obtain seamless understanding of memory across levels. The separate scientific fields studying memory are introspective and it appears that a theorist's ultimate goal should be to cross boundaries.

To maintain perspective in our study, we began with a short history of how modern theories on the organization of memory evolved and how these classified multiple memory systems and learning forms in the brain. We then organized a range of experimental results on the cellular basis of learning so that a common theme becomes evident across animal models of memory. Distinct synaptic plasticity phases have been identified on the basis of molecular and temporal dissociations. The late-phase of plasticity is believed to be the basis of memory consolidation within a memory system and it appears to be selectively induced under specific stimulation conditions. We proceeded to discuss theoretical models linking synaptic plasticity to memory processes that need to retain previous memories in the face of new memory encoding. This requirement manifests as the notorious stability versus plasticity dilemma faced by synapses embedded in such memory systems (Abraham and Robins, 2005). We noticed that previous models addressing this dilemma have all resided to solutions that attempt to delay the ultimate demise of an initial memory trace via slowing down the learning rate. The majority of models assume synapses stochastically respond to plasticity induction stimuli and thus the learning rate is set by the set of transition probabilities between synaptic strengths. However, the learning rate is ultimately linked with the size of the initial signal and therefore these models are also faced with a direct compromise between initial encoding strength and memory lifetime. One prominent example is the cascade model (Fusi et al., 2005), it extends the stochastic transition idea by allowing synapses to access multiple transition probabilities in a state dependent manner.

To resolve the plasticity-stability dilemma we proposed synapses that separate plasticity induction from expression in order to filter synaptic stimuli. These filter synapses process the recent history of plasticity stimuli by acting as a low-pass filter to identify a trend before expressing plasticity. With this simple mechanism in place we obtained unique signal dynamics exhibiting both a rise and fall in the memory signal dynamics while a markable capacity increase was measured within the standard framework used by earlier models. The tight compromise between initial signal and learning rate was now broken because with filters a low initial signal may now rise above the signal noise as new memories are being processed. These dynamics allow filter synapses to exceed cascade synapses in memory capacity specifically within a biologically relevant region

of network size without requiring the cascade's multiple levels of plasticity stability. Additionally, synaptic filters allow the setting of a learning rate by changing the size of thresholds without residing to arbitrary transition probabilities. Thus, they provide a new framework to consider continuous memory encoding processes that may be extended to consider homoeostatic changes that either modify the thresholds or modify the rate of an additional passive decay of filter integration to prolong or decrease expression times.

Further in our study, we suggested that the filter framework could be extended to consider transitions to late-phase plasticity. The filter's signal peak offered a natural timescale that could relate to the optimal time window for repetition protocols to induce late-phase plasticity. Indeed, adding a mechanism that integrates the signal samples obtained at each memory repetition provided a simple model for stimulus induced late-phase plasticity. Triggering late-phase plasticity depended on the post-synaptic depolarization and due to the filter's signal dynamics, under stimulation protocols limited to a few repetitions, achieving maximum depolarization requires a spaced repetition pattern. To retain the spacing effect for protocols consisting of an arbitrary number of repetitions required imposing a simple saturation constraint in the mechanism that integrates signal samples. The mechanisms proposed have been chosen so they relate to known cellular signalling processes. The biophysical nature of the proposed signal integration and saturation mechanism for late-phase plasticity can be paralleled to component interactions of the well known cAMP signalling pathway, which has been strongly implicated with the processes that lead to l-LTP formation and the spacing effect.

Our study has aimed at developing abstract models that balance between biological reality and susceptibility to analysis to propose that synaptic integration operating at the level of single synapses and neurons could be used to describe key mechanisms underlying memory formation. We obtained a biophysically plausible synapse model of early and late-phase plasticity for ongoing memory processes that operates by concurrently integrating plasticity induction stimuli and postsynaptic depolarizations while it exhibits superior memory retention relative to previous models. We believe that obtaining a successfully balanced description between theory and biology can have concomitant ramifications on the understanding of the intracellular molecular signalling during memory formation but also on how memories are formed and organized at the level of memory systems.

Calmodulin cooperative binding

In models of memory allocation we have used a threshold function $h(t)$ on the postsynaptic depolarization which we take is reflected by the mean signal $\mu(t)$. Here we aim to establish the biological plausibility of this threshold mechanism. In Chapter 2 we reviewed the synergistic action of neurotransmitter and neuromodulator. The neurotransmitter action was mediated by a Calcium-modulated protein (CaM). This protein is used as a Ca^{2+} sensor in multiple cellular mechanisms. A key action of CaM for learning is that it allows Ca^{2+} signals to regulate the concentration of intracellular cAMP by regulating the activity of cAMP producing adenylate cyclases (AC) (see also Wang and Storm, 2003, for a review).

We know that postsynaptic depolarizations cause a rapid increase in Ca^{2+} concentration and thus we have to examine whether CaM could function as the threshold mechanism which once activated could in turn act on its AC enzyme substrate. CaM requires the cooperative binding of four Ca^{2+} ions to switch to a conformation that stimulates AC.

There exist a few simple models to describe the Ca^{2+} binding to CaM (see Valeyev et al., 2008). Since we have used a Hill equation for the signal threshold, here we show how the Hill equation (Hill, 1913) can be used to model the CaM cooperative binding.

We let a function $H(t)$ denote the activation of enzyme process $U(t)$ due CaM under a Ca^{2+} ion concentration I. $H(t)$ acts as a sensor of Ca^{2+} ions by binding Ca^{2+} at a fixed number of binding sites C_T due to collision of Ca^{2+} ions with concentration I with unbound ion sites C_R. The Ca^{2+} ion sensor CaM is activated when nI ions are bound to it to form the $[nIC_R]$ complex. The number of binding sites remains fixed and thus

$$C_T = C_R + [nIC_R].$$

Collisions rates are proportional to concentration of ions and free binding sites with a rate $K_{on}I^nC_R$. The ion concentration is raised to the power n because it represents the probability that n independent binding events occur when n ions are found at the same binding site C_R. We let Ca^{2+} ions dissociate with rate $K_{off}[nIC_R]$ giving :

$$\frac{[nIC_R]}{dt} = K_{on}I^nC_R - K_{off}[nIC_R] \tag{1}$$

At steady state we have :

$$\frac{K_{on}}{K_{off}} C_R I^n = [nIC_R] \tag{2}$$

$$\frac{K_{on}}{K_{off}} (C_T - [nIC_R]) I^n = [nIC_R] \tag{3}$$

$$\frac{K_{on}}{K_{off}} C_T I^n = [nIC_R](1 + \frac{K_{on}}{K_{off}} I^n) \tag{4}$$

We then simplify the expression by writing :

$$K^n = \frac{K_{off}}{K_{on}}$$

to obtain a Hill equation (Hill, 1913) :

$$\frac{[nIC_R]}{C_T} = \frac{I^n}{K^n + I^n} \tag{5}$$

where the ratio of bound $[nIC_R]$ over the total number of binding sites C_T can be interpreted as the probability that the ion sensor CaM has been activated in response to a ion concentration I at time t. We write this probability as a function:

$$H(t) = \frac{I(t)^n}{I(t)^n + K^n} \tag{6}$$

where the K and I are the microscopic equilibrium dissociation constant and the ligand concentration, respectively.

By taking into account that the Ca^{2+} ion concentration $I(t)$ is proportional to the mean signal $\mu(t)$, we see that CaM could operate as a biophysical mechanism of the threshold function $h(t)$ we wrote in equation (9.2), and thus stimulate its substrate in response to sufficient postsynaptic depolarization.

As we discussed in Section 2.7, the sensitivity of AC to Ca^{2+} concentration dependents on the concentration of CaM and Mg^{2+}. Increasing the concentration of Mg^{2+} raised the threshold required for the concentration of Ca^{2+} to activate AC (Yovell et al., 1992).

Filter-cascade parameters

Intermediate State				Terminal States		
i⇒i+1	τ	η	$\Theta_{p,q}$	i⇒1	η	Θ
1⇒(2,1)	1	0.000	1,1	1⇒1	-	-,-
2⇒(3,1)	2.5	0.000	2,2	2⇒1	0.000	1
3⇒(4,1)	4	0.636	2,2	3⇒1	0.000	2
4⇒(5,1)	8	0.355	3,3	4⇒1	0.693	2
5⇒(6,1)	16	0.602	3,3	5⇒1	0.385	3
6⇒(7,1)	32	0.402	4,4	6⇒1	0.608	4
7⇒(8,1)	64	0.506	4,4	7⇒1	0.406	4
8⇒(9,1)	128	0.607	4,4	8⇒1	0.507	4
9⇒(10,1)	256	0.707	4,4	9⇒1	0.607	4
10⇒(11,1)	512	0.807	4,4	10⇒1	0.707	4
11⇒(12,1)	1024	0.909	4,4	11⇒1	0.808	4
12⇒(13,1)	2046	1.013	4,4	12⇒1	0.909	4
13⇒(14,1)	4092	1.117	4,4	13⇒1	0.013	4
14⇒(15,1)	8184	1.223	4,4	14⇒1	1.117	4
15⇒(16,1)	16368	1.330	4,4	15⇒1	1.223	4

TABLE 1: Parameters of the dual filter with decay to match the transition rates of the cascade model for induction rate $r = 1.0$. The left column contains the sets for thresholds and decay rates to reproduce the required escape time τ. The right column displays the values required for the terminal states. As each state consists of two independent filters with a single absorbing boundary we just ignore the $p - filter$ and assign the parameters of the previous cascade index to the remaining $q - filter$. For this reason only a single threshold value is given as there is no $p - filter$ on the terminal state.

Intermediate State				Terminal States		
i⇒i+1 or 1	τ	η	Θ	i⇒1	η	$\Theta_{p,q}$
1⇒(2,1)	1	0.000	1,1	1⇒1	-	-,-
2⇒(3,1)	2	0.000	1,2	2⇒1	0.000	1,1
3⇒(4,1)	4	0.000	2,2	3⇒1	0.000	2,1
4⇒(5,1)	8	1.098	2,2	4⇒1	0.000	4,1
5⇒(6,1)	16	0.000	4,4	5⇒1	1.022	3,2
6⇒(7,1)	32	0.167	4,4	6⇒1	1.934	3,2
7⇒(8,1)	64	0.312	4,4	7⇒1	2.705	3,2
8⇒(9,1)	128	0.444	4,4	8⇒1	0.860	4,3
9⇒(10,1)	256	0.569	4,4	9⇒1	1.117	4,3
10⇒(11,1)	512	0.690	4,4	10⇒1	1.363	4,3
11⇒(12,1)	1024	0.809	4,4	11⇒1	1.603	4,3
12⇒(13,1)	2046	0.926	4,4	12⇒1	1.839	4,3
13⇒(14,1)	4092	1.042	4,4	13⇒1	2.074	4,3
14⇒(15,1)	8184	1.158	4,4	14⇒1	2.307	4,3
15⇒(16,1)	16368	1.273	4,4	15⇒1	2.539	4,3

TABLE 2: Parameters of the unified filter with decay to match the transition rates of the cascade model for induction rate $r = 1.0$. The left column contains the sets for thresholds and decay rates to reproduce a the required escape time τ. The set of thresholds for each synapse are randomly assigned as either upper or lower with equal probability. A synapse that moves to cascade index 2 maybe assigned the threshold for p, q as 1,2 or vice-versa. The right column displays the values required for the terminal states at n where a holding boundary is used. These filters are only used at the upper cascade state and the thresholds are not inter-changeable. The first value is the holding boundary and the second is the absorbing boundary (through which q transitions will be expressed).

Intermediate State				Terminal States		
i⇒i+1 or 1	τ	η	Θ	i⇒1	η	$\Theta_{p,q}$
1⇒(2,1)	1	0.000	1,1	1⇒1	-	-,-
2⇒(3,1)	2	0.000	1,2	2⇒1	0.000	1,1
3⇒(4,1)	4	0.000	2,2	3⇒1	0.000	1,1
4⇒(5,1)	8	0.000	2,4	4⇒1	0.000	1,2
5⇒(6,1)	16	0.000	4,4	5⇒1	0.000	1,4
6⇒(7,1)	32	0.000	4,8	6⇒1	0.000	1,16
7⇒(8,1)	64	0.000	8,8	7⇒1	0.000	1,32
8⇒(9,1)	128	0.000	8,16	8⇒1	0.000	1,64
9⇒(10,1)	256	0.000	16,16	9⇒1	0.000	1,128
10⇒(11,1)	512	0.000	16,32	10⇒1	0.000	1,256
11⇒(12,1)	1024	0.000	32,32	11⇒1	0.000	1,512
12⇒(13,1)	2048	0.000	32,64	12⇒1	0.000	1,1024
13⇒(14,1)	4096	0.000	64,64	13⇒1	0.000	1,2048
14⇒(15,1)	8192	0.000	64,128	14⇒1	0.000	1,4096
15⇒(16,1)	16384	0.000	128,128	15⇒1	0.000	1,8192

TABLE 3: Table of threshold value progression of the unified filter to match the transition rates of the cascade model for induction rate. The left column contains the sets for thresholds and decay rates to reproduce the required escape time τ. The set of thresholds for each synapse are randomly assigned as either upper or lower with equal probability. A synapse that moves to cascade index 2 maybe assigned the threshold for p, q as 1,2 or vice-versa. The right column displays the values required for the terminal states at n where a holding boundary is used. These filters are only used at the upper cascade state and the thresholds are not inter-changeable. The first value is the holding boundary and the second is the absorbing boundary (through which q transitions will be expressed).

Simulation Algorithms

Below we include pseudo-code of the simulation algorithms used in Chapter 6 and Chapter 5 to simulate memory lifetimes of synaptic filter synapses with decay and cascade models in continuous time. The simpler results of the filter model with no decay process can be obtained by setting decay rates to zero and the discrete time simulations can be obtained by changing the time-step to a fixed constant.

Algorithm 1 Simulation of single LP filter with decay in continuous time

Input: $ThetaL < 0, ThetaP > 0, stimRate > 0, decayRate \geq 0, 0 > ts \geq 1$
{*thetaL ← Filter Lower Boundary Threshold value*}
{ *thetaP ← Filter Upper Boundary Threshold value*}
{*stimRate ← Rate of memory enconding events*}
{*decayRate ← Rate of Filter-state decay back to zero-state*}
{ *ts ← Simulation timestep - set 0.001 or 1 for discrete time*}
{*RecallPeriod ← A regular time interval at which we wish to obtain signal samples*}

for $t = 0$ to *Trials* **do**
  $t \leftarrow 0$
  $LastRecallt \leftarrow 0$
  {Initialize Synapses}
  for $i = 0$ to *Number of Synapses* **do**
    $strength[i] \leftarrow$ *Random either -1 or 1*
    $filterState[i] \leftarrow$ *Random state according to Filter PDF*
  end for
  {Do simulation Loop}
  for $p = 0$ to *Number of Patterns to store* **do**
    for $i = 0$ to *Number of Synapses* **do**
      $probDecay \leftarrow (1.0 - \exp(-decayRate * timeStep))$
      $decaySteps \leftarrow$ *Random from Binomial Bin(probDecay,filterState[i])*
      $filterState[i] \leftarrow$ *decrement towards zero by decaySteps*
      $stim \leftarrow$ *Random between DEP(-1) and POT(+1)*
      {Store Tracked pattern vector}
      if $p = 0$ **then**
        $X[i] \leftarrow stim$
      end if
      $filterState[i] \leftarrow filterState[i] + stim$
      {Check for threshold event}
      if $filterState[i] = ThetaL$ **then**
        $filterState[i] \leftarrow 0$
        $strength[i] \leftarrow -1$
      else if $filterState[i] = ThetaP$ **then**
        $filterState[i] \leftarrow 0$
        $strength[i] \leftarrow +1$
      end if
    end for
    {Advance timestep to the time of next pattern encoding}
    $timeStep \leftarrow$ *Random exponentially distributed with rate stimRate*
    $t \leftarrow t + timeStep$
    {If the next timePoint exceeds our regular measurement interval, then obtain signal sample}
    if $t > (LastRecallt + RecallPeriod)$ **then**
      $LastRecallt \leftarrow (LastRecallt + RecallPeriod)$
      {Here measure and store signal sample μ and μ^2 using $\boldsymbol{strength} \cdot \boldsymbol{X}$ for this time $t_i = LastRecallt$ }
    end if
  end for
end for

Output: Calculate mean signal and variance from the collected samples of μ and μ^2 across t_i and save to output file.

Algorithm 2 Simulation of cascade signal in continuous time

Input: $CSize, q[1] \cdots q[CSize], p[1] \cdots p[CSize], stimRate, 0 > ts \geq 1$
{$CSize \leftarrow$ *Cascade size n*}
{ $q[] \leftarrow$ *Array of transition prob. q*}
{ $p[] \leftarrow$ *Array of transition prob. q*}
{$stimRate \leftarrow$ *Rate of memory enconding events*}
{ $ts \leftarrow$ *Simulation timestep - set 0.001 or 1 for discrete time*}
{$RecallPeriod \leftarrow$ *A regular time interval at which we wish to obtain signal samples*}

for $t = 0$ to *Trials* **do**
  $t \leftarrow 0$
  $LastRecallt \leftarrow 0$
  {Initialize synapses}
  for $i = 0$*toNumber of synapses* **do**
    $Strength[i] \leftarrow$ *Random either -1 or 1*
    $CState[i] \leftarrow$ *Random state uniformly drawn from* $1 \cdots CSize$
  end for{Do simulation Loop}
  for $p = 0$ to *Number of patterns to store* **do**
    for $i = 0$ to *Number of synapses* **do**
      $stim \leftarrow$ *Random between DEP(-1) and POT(+1)*
      {Store tracked pattern vector}
      if $p = 0$ **then**
        $X[i] \leftarrow stim$
      end if{Make cascade state transition according to stimulus and state}
      if $stim = strength[i]$ **then**
        if $CState[i] < CSize$ **then**
          if $p[CState[i]] > r$ **then**
            $CState[i] \leftarrow CState[i] + 1$ {Metaplastic transition}
          end if
        end if
      else if $stim \neq strength[i]$ **then**
        if $q[CState[i]] > r$ **then**
          $CState[i] \leftarrow 0$ {Plastic transition}
          $Strength[i] \leftarrow -1 * Strength[i]$
        end if
      end if
    end for
    {Advance timestep to the time of next pattern encoding}
    $timeStep \leftarrow$ *Random exponentially distributed with rate stimRate*
    $t \leftarrow t + timeStep$
    {If the next timePoint exceeds our regular measurement interval, then obtain signal sample}
    if $t > (LastRecallt + RecallPeriod)$ **then**
      $LastRecallt \leftarrow (LastRecallt + RecallPeriod)$
      {Here measure and store signal sample μ and μ^2 using $\boldsymbol{strength} \cdot \boldsymbol{X}$ for this time $t_i = LastRecallt$ }
    end if
  end for
end for

Output: Calculate mean signal and variance from the collected samples of μ and μ^2 across t_i and save to output file.

Bibliography

T. Abel, P.V. Nguyen, M. Barad, T.A.S. Deuel, E.R. Kandel, and R. Bourtchouladze. Genetic demonstration of a role for pka in the late phase of ltp and in hippocampus-based long-term memory. *Cell*, 88(5):615–626, 1997.

Ted Abel and Peter V Nguyen. Regulation of hippocampus-dependent memory by cyclic amp-dependent protein kinase. *Progress in brain research*, 169:97–115, 2008.

W.C. Abraham. How long will long-term potentiation last? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1432):735, 2003.

W.C. Abraham and M.F. Bear. Metaplasticity: the plasticity of synaptic plasticity. *Trends in neurosciences*, 19(4):126–130, 1996.

W.C. Abraham, B. Logan, J.M. Greenwood, and M. Dragunow. Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *Journal of Neuroscience*, 22(21):9626, 2002.

W.C. Abraham and S. Otani. Macromolecules and the maintenance of long-term potentiation. *Kindling and Synaptic Plasticity. The Legacy of Graham Goddard*, pages 92–109, 1991.

W.C. Abraham and A. Robins. Memory retention-the synaptic stability versus plasticity dilemma. *TRENDS in Neurosciences*, 28(2):73–78, 2005. ISSN 0166-2236.

W.C. Abraham and W.P. Tate. Metaplasticity: a new vista across the field of synaptic plasticity. *Progress in neurobiology*, 52(4):303–323, 1997.

T.W. Abrams, K.A. Karl, and ER Kandel. Biochemical studies of stimulus convergence during classical conditioning in aplysia: dual regulation of adenylate cyclase by ca2+/calmodulin and transmitter. *The Journal of neuroscience*, 11(9):2655, 1991.

R. Adolphs, L. Cahill, R. Schul, and R. Babinsky. Impaired declarative memory for emotional material following bilateral amygdala damage in humans. *Learning & Memory*, 4(3):291, 1997.

Sriram M Ajay and Upinder S Bhalla. A role for erkii in synaptic pattern selectivity on the time-scale of minutes. *European Journal of Neuroscience*, 20(10):2671–2680, 2004.

I. Akirav and G. Richter-Levin. Biphasic modulation of hippocampal plasticity by behavioral stress and basolateral amygdala stimulation in the rat. *The Journal of neuroscience*, 19(23):10530, 1999.

I. Akirav and G. Richter-Levin. Mechanisms of amygdala modulation of hippocampal plasticity. *The Journal of neuroscience*, 22(22):9912, 2002.

D.J. Amit and S. Fusi. Constraints on learning in dynamic synapses. *Network: Computation in Neural Systems*, 3(4):443–464, 1992.

D.J. Amit and S. Fusi. Learning in neural networks with material synapses. *Neural Computation*, 6(5):957–982, 1994.

D.J. Amit, H. Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. *Physical Review A*, 32(2):1007–1018, 1985a.

D.J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530–1533, 1985b. ISSN 1079-7114.

D.J. Amit, H. Gutfreund, and H. Sompolinsky. Statistical mechanics of neural networks near saturation. *Annals of Physics*, 173(1):30–67, 1987.

Yali Amit and Yibi Huang. Precise capacity analysis in binary networks with multiple coding level inputs. *Neural computation*, 22(3):660–688, 2010.

P. Andersen. *The hippocampus book*. Oxford University Press, USA, 2007.

M.C. Anderson. Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, 49(4):415–445, 2003.

R.B. Anderson. The power law as an emergent property. *Memory & cognition*, 29(7): 1061, 2001.

S. Arakawa, J.D. Gocayne, W.R. McCombie, D.A. Urquhart, L.M. Hall, C.M. Fraser, and J.C. Venter. Cloning, localization, and permanent expression of a drosophila octopamine receptor. *Neuron*, 4(3):343–354, 1990.

B.A. Armitage and S.A. Siegelbaum. Presynaptic induction and expression of homosynaptic depression at aplysia sensorimotor neuron synapses. *The Journal of neuroscience*, 18(21):8770, 1998.

A.A. Bagal, J.P.Y. Kao, C.M. Tang, and S.M. Thompson. Long-term potentiation of exogenous glutamate responses at single dendritic spines. *Proceedings of the National Academy of Sciences*, 102(40):14434, 2005.

C.H. Bailey and M. Chen. Morphological basis of long-term habituation and sensitization in aplysia. *Science*, 220(4592):91, 1983.

C.H. Bailey and M. Chen. Long-term memory in aplysia modulates the total number of varicosities of single identified sensory neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 85(7):2373, 1988a.

C.H. Bailey and M. Chen. Morphological basis of short-term habituation in aplysia. *The Journal of neuroscience*, 8(7):2452, 1988b.

P.B. Ballard. Obliviscence and reminiscence. *British Journal of Psychology*, 1913.

Fabricio Ballarini, Diego Moncada, Maria Cecilia Martinez, Nadia Alen, and Haydée Viola. Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences*, 106(34):14599–14604, 2009.

A. Barco, J.M. Alarcon, and E.R. Kandel. Expression of constitutively active CREB protein facilitates the late phase of long-term potentiation by enhancing synaptic capture. *Cell*, 108(5):689–703, 2002.

A. Barco, M. Lopez de Armentia, and J.M. Alarcon. Synapse-specific stabilization of plasticity processes: The synaptic tagging and capture hypothesis revisited 10 years later. *Neuroscience & Behavioural Reviews*, 32(4):831–851, 2008.

A.B. Barrett, G.O. Billings, R.G.M. Morris, and MC Van Rossum. State based model of long-term potentiation and synaptic tagging and capture. *PLoS Comput. Biol*, 5, 2009.

A.B. Barrett and MCW Van Rossum. Optimal learning rules for discrete synapses. *PLoS Computational Biology*, 4(11), 2008.

Andres Barria, Dominique Muller, Victor Derkach, Leslie C Griffith, and Thomas R Soderling. Regulatory phosphorylation of ampa-type glutamate receptors by cam-kii during long-term potentiation. *Science*, 276(5321):2042–2045, 1997.

D. Bartsch, M. Ghirardi, P.A. Skehel, K.A. Karl, S.P. Herder, M. Chen, C.H. Bailey, and E.R. Kandel. Aplysia creb2 represses long-term facilitation: relief of repression converts transient facilitation into long-term functional and structural change. *Cell*, 83(6):979–992, 1995.

M.F. Bear. Bidirectional synaptic plasticity: from theory to reality. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1432):649, 2003.

MF Bear, A. Kleinschmidt, QA Gu, and W. Singer. Disruption of experience-dependent synaptic modifications in striate cortex by infusion of an nmda receptor antagonist. *The Journal of Neuroscience*, 10(3):909, 1990.

CDO Beck, Bradley Schroeder, and Ronald L Davis. Learning performance of normal and mutantdrosophila after repeated conditioning trials with discrete stimuli. *The Journal of Neuroscience*, 20(8):2944–2953, 2000.

Eva Benito and Angel Barco. Creb's control of intrinsic and synaptic plasticity: implications for creb-dependent memory models. *Trends in neurosciences*, 33(5):230–240, 2010.

I. Bethus, D. Tse, and R.G.M. Morris. Dopamine and memory: Modulation of the persistence of memory for novel hippocampal nmda receptor-dependent paired associates. *The Journal of Neuroscience*, 30(5):1610, 2010.

A.V. Beylin, C.C. Gandhi, G.E. Wood, A.C. Talk, L.D. Matzel, and T.J. Shors. The Role of the Hippocampus in Trace Conditioning: Temporal Discontinuity or Task Difficulty?* 1. *Neurobiology of Learning and Memory*, 76(3):447–461, 2001.

G. Bi and M. Poo. Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. *Journal of Neuroscience*, 18(24):10464–10472, 1998.

G.Q. Bi and J. Rubin. Timing in synaptic plasticity: from detection to integration. *TRENDS in Neurosciences*, 28(5):222–228, 2005.

R.A. Bjork. Retrieval inhibition as an adaptive mechanism in human memory. *Varieties of memory and consciousness: Essays in honour of Endel Tulving*, pages 309–330, 1989.

TV Bliss, GV Goddard, and M. Riives. Reduction of long-term potentiation in the dentate gyrus of the rat following selective depletion of monoamines. *The Journal of Physiology*, 334(1):475, 1983.

T.V.P. Bliss and G.L. Collingridge. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, 361(6407):31–39, 1993.

T.V.P. Bliss and T. Lømo. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of physiology*, 232(2):331, 1973.

J.M. Brader, W. Senn, and S. Fusi. Learning real-world stimuli in a neural network with spike-driven synaptic dynamics. *Neural computation*, 19(11):2881–2912, 2007. ISSN 0899-7667.

W. Brown. To what extent is memory measured by a single recall?. *Journal of Experimental Psychology*, 6(5):377, 1923.

M. Brunelli, V. Castellucci, and ER Kandel. Synaptic facilitation and behavioral sensitization in aplysia: possible role of serotonin and cyclic amp. *Science*, 194(4270): 1178, 1976.

D. Byers, R.L. Davis, and J.A. Kiger. Defect in cyclic amp phosphodiesterase due to the dunce mutation of learning in drosophila melanogaster. *Nature*, 289:79 – 81, 1981.

J. Byrne, V. Castellucci, and E.R. Kandel. Receptive fields and response properties of mechanoreceptor neurons innervating siphon skin and mantle shelf in aplysia. *Journal of Neurophysiology*, 37(5):1041, 1974.

J.H. Byrne, V.F. Castellucci, T.J. Carew, and E.R. Kandel. Stimulus-response relations and stability of mechanoreceptor and motor neurons mediating defensive gill-withdrawal reflex in aplysia. *Journal of neurophysiology*, 41(2):402, 1978.

J.H. Byrne and E.R. Kandel. Presynaptic facilitation revisited: state and time dependence. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 16(2):425, 1996.

Donald Peter Cain, Deborah Saucier, Jeff Hall, Eric L Hargreaves, and Francis Boon. Detailed behavioral analysis of water maze acquisition under apv or cnqx: contribution of sensorimotor disturbances to drug-induced acquisition deficits. *Behavioral neuroscience*, 110(1):86, 1996.

J. Campbell, L. Nadel, D. Duke, and L. Ryan. Remembering all that and then some: Recollection of autobiographical memories after a 1-year delay. *Memory*, 19(4):406–415, 2011.

Thomas J Carew, Harold M Pinsker, and Eric R Kandel. Long-term habituation of a defensive withdrawal reflex in aplysia. *Science*, 1972.

T.J. Carew, E.T. Walters, and E.R. Kandel. Classical conditioning in a simple withdrawal reflex in aplysia californica. *The Journal of Neuroscience*, 1(12):1426, 1981.

C.M. Carroll, S.J. Martin, J. Sandin, B. Frenguelli, and R.G.M. Morris. Dopaminergic modulation of the persistence of one-trial hippocampus-dependent memory. *Learning & Memory*, 13(6):760, 2006.

A. Casadio, K.C. Martin, M. Giustetto, H. Zhu, M. Chen, D. Bartsch, C.H. Bailey, and E.R. Kandel. A transient, neuron-wide form of creb-mediated long-term facilitation can be stabilized at specific synapses by local protein synthesis. *Cell*, 99(2):221–237, 1999.

Stijn Cassenaer and Gilles Laurent. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, 482(7383):47–52, 2012.

V. Castellucci, H. Pinsker, I. Kupfermann, and E.R. Kandel. Neuronal mechanisms of habituation and dishabituation of the gill-withdrawal reflex in aplysia. *Science*, 167 (3926):1745, 1970.

VF Castellucci, H. Blumenfeld, P. Goelet, and ER Kandel. Inhibitor of protein synthesis blocks longterm behavioral sensitization in the isolated gill-withdrawal reflex of aplysia. *Journal of neurobiology*, 20(1):1–9, 1989.

V.F. Castellucci and E.R. Kandel. A quantal analysis of the synaptic depression underlying habituation of the gill-withdrawal reflex in aplysia. *Proceedings of the National Academy of Sciences*, 71(12):5004, 1974.

CP Cavafy. Collected poems, translated by edmund keeley and philip sherrard, 1975.

C.B. Cave and L.R. Squire. Intact and long-lasting repetition priming in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(3):509–520, 1992.

F.L.F. Chang and W.T. Greenough. Transient and enduring morphological correlates of synaptic activity and efficacy change in the rat hippocampal slice. *Brain Research*, 309(1):35–46, 1984.

Z. Chen, S. Fujii, K.I. Ito, H. Kato, K. Kaneko, and H. Miyakawa. Activation of dopamine d1 receptors enhances long-term depression of synaptic transmission induced by low frequency stimulation in rat hippocampal ca1 neurons* 1. *Neuroscience letters*, 188(3):195–198, 1995.

DM Chetkovich, R Gray, D Johnston, and JD Sweatt. N-methyl-d-aspartate receptor activation increases camp levels and voltage-gated ca2+ channel activity in area ca1 of hippocampus. *Proceedings of the National Academy of Sciences*, 88(15):6467–6471, 1991.

D. Chin and A.R. Means. Calmodulin: a prototypical calcium sensor. *Trends in Cell Biology*, 10(8):322–328, 2000.

K. Cho, JP Aggleton, MW Brown, and ZI Bashir. An experimental test of the role of postsynaptic calcium levels in determining synaptic strength using perirhinal cortex of rat. *The Journal of physiology*, 532(2):459–466, 2001.

B.R. Christie and W.C. Abraham. Priming of associative long-term depression in the dentate gyrus by [theta] frequency synaptic activity. *Neuron*, 9(1):79–84, 1992.

A. Claridge-Chang, R.D. Roorda, E. Vrontou, L. Sjulson, H. Li, J. Hirsh, and G. Miesenböck. Writing memories with light-addressable reinforcement circuitry. *Cell*, 139(2):405–415, 2009.

C. Clopath, L. Ziegler, E. Vasilaki, L. Büsing, and W. Gerstner. Tag-trigger-consolidation: a model of early and late long-term-potentiation and depression. *PLoS computational biology*, 4(12):e1000248, 2008. ISSN 1553-7358.

M.A. Cohen and S. Grossberg. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, & Cybernetics*, 1983.

N.J. Cohen. Preserved learning capacity in amnesia: Evidence for multiple memory systems. *Neuropsychology of memory*, 1, 1984.

N.J. Cohen and L.R. Squire. Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210(4466):207, 1980.

Graham L Collingridge, Stephane Peineau, John G Howland, and Yu Tian Wang. Long-term depression in the cns. *Nature Reviews Neuroscience*, 11(7):459–473, 2010.

L.N. Cooper. Stdp: Spiking, timing, rates and beyond. *Frontiers in Synaptic Neuroscience*, 2, 2010.

RJ Cormier, AC Greenwood, and JA Connor. Bidirectional synaptic plasticity correlated with the magnitude of dendritic calcium transients above a threshold. *Journal of Neurophysiology*, 85(1):399, 2001.

M. Costa-Mattioli, D. Gobert, H. Harding, B. Herdy, M. Azzi, M. Bruno, M. Bidinosti, C.B. Mamou, E. Marcinkiewicz, M. Yoshida, et al. Translational control of hippocampal synaptic plasticity and memory by the eif2α kinase, gcn2. *Nature*, 436(7054):1166, 2005.

D.R. Cox and H.D. Miller. *The theory of stochastic processes*, volume 134. Chapman & Hall/CRC, 1977.

P.K. Dash, B. Hochner, and E.R. Kandel. Injection of the camp-responsive element into the nucleus of aplysia sensory neurons blocks long-term facilitation. *Nature*, 345 (6277):718–721, 1990.

SN Davies and GL Collingridge. Role of excitatory amino acid receptors in synaptic transmission in area CA1 of rat hippocampus. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 236(1285):373–384, 1989.

R.L. Davis. Olfactory memory formation in drosophila: from molecular to systems neuroscience. *Annu. Rev. Neuroscience.*, 28:275–302, 2005.

R.L. Davis and J.A. Kiger. Dunce mutants of drosophila melanogaster: mutants defective in the cyclic amp phosphodiesterase enzyme system. *The Journal of cell biology*, 90 (1):101, 1981.

S. Davis, SP Butcher, and RG Morris. The nmda receptor antagonist d-2-amino-5-phosphonopentanoate (d-ap5) impairs spatial learning and ltp in vivo at intracerebral concentrations comparable to those that block ltp in vitro. *The Journal of neuroscience*, 12(1):21, 1992.

D. Debanne, B.H. Gähwiler, and S.M. Thompson. Asynchronous pre-and postsynaptic activity induces associative long-term depression in area CA1 of the rat hippocampus in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 91(3):1148, 1994.

D. Debanne, B.H. Gähwiler, and S.M. Thompson. Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. *The Journal of Physiology*, 507(1):237, 1998.

K. Deisseroth, H. Bito, and R.W. Tsien. Signaling from synapse to nucleus: postsynaptic creb phosphorylation during multiple forms of hippocampal synaptic plasticity. *Neuron*, 16(1):89–101, 1996.

D.M. Diamond, C.R. Park, M.J. Puls, and G.M. Rose. Differential effects of stress on hippocampal and amygdaloid ltp. *Neuronal mechanisms of memory formation: concepts of long-term potentiation and beyond*, page 379, 2001.

Allison J Doupe and Patricia K Kuhl. Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, 22(1):567–631, 1999.

CA Doyle, WK Cullen, MJ Rowan, and R Anwyl. Low-frequency stimulation induces homosynaptic depotentiation but not long-term depression of synaptic transmission in the adult anaesthetized and awake rat hippocampus¡ i¿ in vivo¡/i¿. *Neuroscience*, 77(1):75–85, 1997.

Patrick J Drew and LF Abbott. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences*, 103 (23):8876–8881, 2006.

Y. Dudai, G. Corfas, and S. Hazvi. What is the possible contribution of ca 2+-stimulated adenylate cyclase to acquisition, consolidation and retention of an associative olfactory memory indrosophila. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 162(1):101–109, 1988.

Y. Dudai, Y.N. Jan, D. Byers, W.G. Quinn, and S. Benzer. dunce, a mutant of drosophila deficient in learning. *Proceedings of the National Academy of Sciences*, 73(5):1684, 1976.

Serena M Dudek and R Douglas Fields. Somatic action potentials are sufficient for late-phase ltp-related cell signaling. *Proceedings of the National Academy of Sciences*, 99 (6):3962–3967, 2002.

S.M. Dudek and M.F. Bear. Homosynaptic long-term depression in area ca1 of hippocampus and effects of n-methyl-d-aspartate receptor blockade. *Proceedings of the National Academy of Sciences of the United States of America*, 89(10):4363, 1992.

S.M. Dudek and M.F. Bear. Bidirectional long-term modification of synaptic effectiveness in the adult and immature hippocampus. *Journal of Neuroscience*, 13(7):2910, 1993.

T. Dunwiddie and G. Lynch. Long-term potentiation and depression of synaptic responses in the rat hippocampus: localization and frequency dependency. *The Journal of Physiology*, 276(1):353, 1978.

H.E. Ebbinghaus. *Memory: A contribution to experimental psychology*. New York:Dover, 1885.

H. Eichenbaum and N.J. Cohen. *From conditioning to conscious recollection: Memory systems of the brain*. Oxford University Press, USA, 2004.

H. Eichenbaum, P. Mathews, and N.J. Cohen. Further studies of hippocampal representation during odor discrimination learning. *Behav Neurosci*, 103(6):1207–1216, 1989.

J.M. Ellenbogen, P.T. Hu, J.D. Payne, D. Titone, and M.P. Walker. Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences*, 104(18):7723, 2007.

T. Elliott and K. Lagogiannis. Taming fluctuations in a stochastic model of spike-timing-dependent plasticity. *Neural Computation*, 21(12):3363–3407, 2009.

T. Elliott and K. Lagogiannis. The rise and fall of memory in a model of synaptic integration. *Neural Computation*, 24(10):2604–2654, 2012.

Terry Elliott. Temporal dynamics of rate-based synaptic plasticity rules in a stochastic model of spike-timing-dependent plasticity. *Neural computation*, 20(9):2253–2307, 2008.

Terry Elliott. Discrete states of synaptic strength in a stochastic model of spike-timing-dependent plasticity. *Neural computation*, 22(1):244–272, 2010a.

Terry Elliott. A non-markovian random walk underlies a stochastic model of spike-timing-dependent plasticity. *Neural computation*, 22(5):1180–1230, 2010b.

Terry Elliott. The mean time to express synaptic plasticity in integrate-and-express, stochastic models of synaptic plasticity induction. *Neural Computation*, 23(1):124–159, 2011a.

Terry Elliott. Stability against fluctuations: scaling, bifurcations, and spontaneous symmetry breaking in stochastic models of synaptic plasticity. *Neural Computation*, 23 (3):674–734, 2011b.

F. Engert and T. Bonhoeffer. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *J. Physiol*, 260:C917–925, 1991.

M.H. Erdelyi. *The recovery of unconscious memories: Hypermnesia and reminiscence*. University of Chicago Press, 1996.

M.H. Erdelyi. The ups and downs of memory. *American Psychologist*, 65(7):623, 2010.

M.H. Erdelyi and J. Becker. Hypermnesia for pictures* 1:: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology*, 6(1):159–171, 1974.

F.J. Evans and W.A.F. Thorn. Two types of posthypnotic amnesia: Recall amnesia and source amnesia. *International Journal of Clinical and Experimental Hypnosis*, 14(2): 162–179, 1966.

G.D. Ferguson and D.R. Storm. Why calcium-stimulated adenylyl cyclases? *Physiology*, 19(5):271–276, 2004.

J.V. Filoteo, W.T. Maddox, et al. Characterizing rule-based category learning deficits in patients with parkinson's disease. *Neuropsychologia*, 45(2):305–320, 2007.

C.C. Fink and T. Meyer. Molecular mechanisms of CaMKII activation in neuronal plasticity. *Current opinion in neurobiology*, 12(3):293–299, 2002.

Rosalina Fonseca, U Valentin Nägerl, and Tobias Bonhoeffer. Neuronal activity determines the protein synthesis dependence of long-term potentiation. *Nature neuroscience*, 9(4):478–480, 2006.

Rosalina Fonseca, U Valentin Nägerl, Richard GM Morris, and Tobias Bonhoeffer. Competing for memory: hippocampal ltp under regimes of reduced protein synthesis. *Neuron*, 44(6):1011–1020, 2004.

N.J. Fortin, K.L. Agster, and H.B. Eichenbaum. Critical role of the hippocampus in memory for sequences of events. *Nature neuroscience*, 5(5):458–462, 2002.

Paul W Frankland and Sheena A Josselyn. Neuroscience: memory and the single molecule. *Nature*, 493(7432):312–313, 2013.

Robert M French. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4):128–135, 1999.

JU Frey. Long-lasting hippocampal plasticity: Cellular model for memory consolidation. *Cell Polarity and Subcellular RNA Localization*, pages 27–40, 2001.

S. Frey and JU Frey. 'Synaptic tagging'and'cross-tagging'and related associative reinforcement processes of functional plasticity as the cellular basis for memory formation. *Progress in brain research*, 169:117, 2008. ISSN 0079-6123.

U. Frey, YY Huang, and ER Kandel. Effects of cAMP simulate a late stage of LTP in hippocampal CA1 neurons. *Science*, 260(5114):1661, 1993.

U. Frey, M. Krug, K.G. Reymann, and H. Matthies. Anisomycin, an inhibitor of protein synthesis, blocks late phases of ltp phenomena in the hippocampal ca1 region in vitro. *Brain research*, 452(1-2):57–65, 1988.

U. Frey, H. Matthies, K.G. Reymann, and H. Matthies. The effect of dopaminergic D1 receptor blockade during tetanization on the expression of long-term potentiation in the rat CA1 region in vitro. *Neuroscience letters*, 129(1):111–114, 1991. ISSN 0304-3940.

U Frey and R. Morris. Synaptic tagging and long-term potentiation. *Nature*, 385:533 – 536, 1997.

U Frey and RGM Morris. Weak before strong: dissociating synaptic tagging and plasticity-factor accounts of late-ltp. *Neuropharmacology*, 37(4):545–552, 1998.

U. Frey, K. Schollmeier, KG Reymann, and T. Seidenbecher. Asymptotic hippocampal long-term potentiation in rats does not preclude additional potentiation at later phases. *Neuroscience*, 67(4):799–807, 1995.

Uwe Frey, Helmut Schroeder, and Hansjurgen Matthies. Dopaminergic antagonists prevent long-term maintenance of posttetanic ltp in the ca1 region of rat hippocampal slices. *Brain Research*, 522(1):69 – 75, 1990. ISSN 0006-8993.

R.C. Froemke and Y. Dan. Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*, 416(6879):433–438, 2002.

W.N. Frost, V.F. Castellucci, R.D. Hawkins, and E.R. Kandel. Monosynaptic connections made by the sensory neurons of the gill-and siphon-withdrawal reflex in aplysia participate in the storage of long-term memory for sensitization. *Proceedings of the National Academy of Sciences of the United States of America*, 82(23):8266, 1985.

S. Fusi. Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biological Cybernetics*, 87(5):459–470, 2002.

S. Fusi and LF Abbott. Limits on the memory storage capacity of bounded synapses. *Nature Neuroscience*, 10(4):485–493, 2007.

S. Fusi, M. Annunziato, D. Badoni, A. Salamon, and D.J. Amit. Spike-driven synaptic plasticity: theory, simulation, VLSI implementation. *Neural Computation*, 12(10): 2227–2258, 2000.

S. Fusi, P.J. Drew, and LF Abbott. Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611, 2005.

S. Fusi and W. Senn. Eluding oblivion with smart stochastic selection of synaptic updates. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16:026112, 2006.

D. Gabor. Microscopy by reconstructed wave-fronts. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 197(1051):454–487, 1949.

J.D.E. Gabrieli. Cognitive neuroscience of human memory. *Annual Review of Psychology*, 49(1):87–115, 1998.

J.D.E. Gabrieli, D.A. Fleischman, M.M. Keane, S.L. Reminger, and F. Morrell. Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science*, 6(2):76, 1995.

A. Gasbarri, A. Sulli, and M.G. Packard. The dopaminergic mesencephalic projections to the hippocampal formation in the rat. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 21(1):1–22, 1997.

D. Genoux, U. Haditsch, M. Knobloch, A. Michalon, D. Storm, and I.M. Mansuy. Protein phosphatase 1 is a molecular constraint on learning and memory. *Nature*, 418 (6901):970–975, 2002.

Andrew A George, Ariel M Lyons-Warren, Xiaofeng Ma, and Bruce A Carlson. A diversity of synaptic filters are created by temporal summation of excitation and inhibition. *The Journal of Neuroscience*, 31(41):14721–14734, 2011.

Bertram Gerber, Hiromu Tanimoto, and Martin Heisenberg. An engram found? evaluating the evidence from fruit flies. *Current opinion in neurobiology*, 14(6):737–744, 2004.

M. Ghirardi, O. Braha, B. Hochner, P. Giorgio Montarolo, E.R. Kandel, and N. Dale. Roles of pka and pkc in facilitation of evoked and spontaneous transmitter release at depressed and nondepressed synapses in aplysia sensory neurons. *Neuron*, 9(3): 479–489, 1992.

Anirvan Ghosh and Michael E Greenberg. Calcium signaling in neurons: molecular mechanisms and cellular consequences. *Science*, 268(5208):239–247, 1995.

D.T. Gillespie. *Markov processes: an introduction for physical scientists*. Academic Pr, 1992.

AR Gilliland. The rate of forgetting. *Journal of Educational Psychology*, 39(1):19, 1948.

David L Glanzman. Pkm and the maintenance of memory. *F1000 biology reports*, 5, 2013.

D.L. Glanzman. Habituation in aplysia: the cheshire cat of neurobiology. *Neurobiology of learning and memory*, 92(2):147–154, 2009.

D.L. Glanzman, S.L. Mackey, R.D. Hawkins, A.M. Dyke, P.E. Lloyd, and E.R. Kandel. Depletion of serotonin in the nervous system of aplysia reduces the behavioral enhancement of gill withdrawal as well as the heterosynaptic facilitation produced by tail shock. *The Journal of neuroscience*, 9(12):4200, 1989.

Mark Good and David Bannerman. Differential effects of ibotenic acid lesions of the hippocampus and blockade of¡ em¿ n¡/em¿-methyl-{d}-aspartate receptor-dependent long-term potentiation on contextual processing in rats. *Behavioral neuroscience*, 111 (6):1171, 1997.

Marco Gori and Franco Scarselli. Are multilayer perceptrons adequate for pattern recognition and verification? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11):1121–1132, 1998.

A. Govindarajan, I. Israely, S.Y. Huang, and S. Tonegawa. The dendritic branch is the preferred integrative unit for protein synthesis-dependent ltp. *Neuron*, 69(1):132–146, 2011.

Arvind Govindarajan, Raymond J Kelleher, and Susumu Tonegawa. A clustered plasticity model of long-term memory engrams. *Nature Reviews Neuroscience*, 7(7):575–583, 2006.

Peter Graf, Arthur P Shimamura, and Larry R Squire. Priming across modalities and priming across category levels: extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(2):386, 1985.

Michael Graupner and Nicolas Brunel. Stdp in a bistable synapse model based on camkii and associated signaling pathways. *PLoS computational biology*, 3(11):e221, 2007.

G. Grimmett and D. Stirzaker. *Probability and random processes*, volume 2. Clarendon press Oxford, 1992.

Stephen Grossberg. How does a brain build a cognitive code? *Psychological Review*, 87 (1):1, 1980.

B Gustafsson, F Asztely, E Hanse, and H Wigström. Onset characteristics of long-term potentiation in the guinea-pig hippocampal ca1 region in vitro. *European Journal of Neuroscience*, 1(4):382–394, 1989.

B. Gustafsson, H. Wigstrom, WC Abraham, and YY Huang. Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *Journal of Neuroscience*, 7(3):774–780, 1987.

J. Hanoune and N. Defer. Regulation and role of adenylyl cyclase isoforms. *Annual review of pharmacology and toxicology*, 41(1):145–174, 2001.

P.I. Hanson and H. Schulman. Neuronal Ca2+/calmodulin-dependent protein kinases. *Annual review of biochemistry*, 61(1):559–601, 1992.

R.D. Hawkins, TW Abrams, TJ Carew, and E.R. Kandel. A cellular mechanism of classical conditioning in aplysia: activity-dependent amplification of presynaptic facilitation. *Science*, 219(4583):400, 1983.

D.O. Hebb. The Organization of Behavior. A Neuropsychological Theory., 1949.

M. Heisenberg. Mushroom body memoir: from maps to models. *Nature Reviews Neuroscience*, 4(4):266–275, 2003.

M. Heisenberg, A. Borst, S. Wagner, and D. Byers. Drosophila mushroom body mutants are deficient in olfactory learning. *Journal of neurogenetics*, 2(1):1–30, 1985.

Archibald Vivian Hill. The combinations of haemoglobin with oxygen and with carbon monoxide. i. *Biochemical Journal*, 7(5):471, 1913.

Douglas L Hintzman. Theoretical implications of the spacing effect. *Theories of cognitive psychology: The Loyola Symposium*, 1974.

J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554, 1982.

J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81 (10):3088, 1984.

E.P. Huang. Synaptic plasticity: going through phases with ltp. *Current biology*, 8(10): R350–R352, 1998.

Q. Huang, D. Zhou, K. Chase, J.F. Gusella, N. Aronin, and M. DiFiglia. Immunohistochemical localization of the d1 dopamine receptor in rat brain reveals its axonal transport, pre-and postsynaptic localization, and prevalence in the basal ganglia, limbic system, and thalamic reticular nucleus. *Proceedings of the National Academy of Sciences*, 89(24):11988, 1992a.

Y. Huang and Y. Amit. Capacity analysis in multi-state synaptic models: a retrieval probability perspective. *Journal of computational neuroscience*, 30(3):699–720, 2011.

Y.Y. Huang, A. Colino, D.K. Selig, and R.C. Malenka. The influence of prior synaptic activity on the induction of long-term potentiation. *Science*, 255(5045):730, 1992b.

YY Huang and ER Kandel. Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of hippocampus requires repeated tetanization. *Learning & Memory*, 1(1):74, 1994.

Y.Y. Huang and E.R. Kandel. D1/d5 receptor agonists induce a protein synthesis-dependent late potentiation in the ca1 region of the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, 92(7):2446, 1995.

Guillaume Isabel, Alberto Pascual, and Thomas Preat. Exclusive consolidated memory phases in drosophila. *Science*, 304(5673):1024–1027, 2004.

Eugene M Izhikevich. Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral Cortex*, 17(10):2443–2452, 2007.

J.S. Janowsky, A.P. Shimamura, and L.R. Squire. Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology*, 17(1):3–11, 1989.

Sheena A Josselyn. Continuing the search for the engram: examining the mechanism of fear memories. *Journal of psychiatry & neuroscience: JPN*, 35(4):221, 2010.

Sheena A Josselyn, Chanjun Shi, William A Carlezon, Rachael L Neve, Eric J Nestler, and Michael Davis. Long-term memory is facilitated by camp response element-binding protein overexpression in the amygdala. *The Journal of Neuroscience*, 21 (7):2404–2412, 2001.

B.K. Kaang, E.R. Kandel, and S.G.N. Grant. Activation of camp-responsive genes by stimuli that produce long-term facilitation in aplysia sensory neurons. *Neuron*, 10(3): 427–435, 1993.

E.R. Kandel. The molecular biology of memory storage: a dialogue between genes and synapses. *Science*, 294(5544):1030, 2001.

E.R. Kandel. The biology of memory: a forty-year perspective. *Journal of Neuroscience*, 29(41):12748, 2009.

Hyejin Kang and Erin M Schuman. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science*, 273(5280):1402–1406, 1996.

J.W. Kebabian and D.B. Calne. Multiple receptors for dopamine. *Nature*, 277:93–96, 1979.

J.W. Kebabian and P. Greengard. Dopamine-sensitive adenyl cyclase: possible role in synaptic transmission. *Science*, 174(4016):1346, 1971.

A.C. Keene and S. Waddell. Drosophila olfactory memory: single genes to complex neural circuits. *Nature Reviews Neuroscience*, 8(5):341–354, 2007.

R.P. Kesner, B.L. Bolland, and M. Dakis. Memory for spatial locations, motor responses, and objects: Triple dissociation among the hippocampus, caudate nucleus, and extrastriate visual cortex. *Experimental Brain Research*, 93(3):462–470, 1993.

J.J. Kim and M.G. Baxter. Multiple brain-memory systems: the whole does not equal the sum of its parts. *Trends in neurosciences*, 24(6):324–330, 2001.

J.J. Kim, H.J. Lee, J.S. Han, and M.G. Packard. Amygdala is critical for stress-induced modulation of hippocampal long-term potentiation and learning. *The Journal of Neuroscience*, 21(14):5222, 2001.

MyungSook Kim, Ted Huang, Ted Abel, and Kim T Blackwell. Temporal sensitivity of protein kinase a activation in late-phase long term potentiation. *PLoS computational biology*, 6(2):e1000691, 2010.

B.J. Knowlton, J.A. Mangels, and L.R. Squire. A neostriatal habit learning system in humans. *Science*, 273(5280):1399, 1996.

Jürgen Kornmeier and Zrinka Sosic-Vasic. Parallels between spacing effects during behavioral and cellular learning. *Frontiers in human neuroscience*, 6, 2012.

M.J. Krashes, S. DasGupta, A. Vreede, B. White, J.D. Armstrong, and S. Waddell. A neural circuit mechanism integrating motivational state with memory expression in drosophila. *Cell*, 139(2):416–427, 2009.

M.J. Krashes and S. Waddell. Rapid consolidation to a radish and protein synthesis-dependent long-term memory after single-session appetitive olfactory conditioning in drosophila. *The Journal of Neuroscience*, 28(12):3103, 2008.

M. Krug, B. Lössner, and T. Ott. Anisomycin blocks the late phase of long-term potentiation in the dentate gyrus of freely moving rats. *Brain research bulletin*, 13(1): 39–42, 1984.

H. L. Roediger and L. A. Thorpe. The role of recall time in producing hypermnesia. *Memory & Cognition*, 6(3):296–305, 1978.

Raphael Lamprecht and Joseph LeDoux. Structural plasticity and memory. *Nature Reviews Neuroscience*, 5(1):45–54, 2004.

K.S. Lashley. In search of the engram. In *Symposia of the society for experimental biology*, volume 4, pages 454–482, 1950.

J.E. LeDoux. Emotion: Clues from the brain. *Annual review of psychology*, 46(1): 209–235, 1995.

Anna M Lee, Benjamin R Kanter, Dan Wang, Jana P Lim, Mimi E Zou, Chichen Qiu, Thomas McMahon, Jahan Dadgar, Sarah C Fischbach-Weiss, and Robert O Messing. Prkcz null mice show normal learning and memory. *Nature*, 493(7432):416–419, 2013.

F.J.S. Lee, S. Xue, L. Pei, B. Vukusic, N. Chéry, Y. Wang, Y.T. Wang, H.B. Niznik, X. Yu, and F. Liu. Dual regulation of nmda receptor functions by direct protein-protein interactions with the dopamine d1 receptor. *Cell*, 111(2):219–230, 2002.

Seok-Jin R Lee, Yasmin Escobedo-Lozoya, Erzsebet M Szatmari, and Ryohei Yasuda. Activation of camkii in single dendritic spines during long-term potentiation. *Nature*, 458(7236):299–304, 2009.

G. W. Leibniz. *New Essays Concerning Human Understanding (1764)*. The Open Court Publishing Co., 1916, 1916.

C. Leibold and R. Kempter. Sparseness constrains the prolongation of memory lifetime via synaptic metaplasticity. *Cerebral Cortex*, 18(1):67, 2008.

N. Lemon and D. Manahan-Vaughan. Dopamine d1/d5 receptors gate the acquisition of novel information through hippocampal long-term potentiation and long-term depression. *The Journal of neuroscience*, 26(29):7723, 2006.

Benedetta Leuner, Elizabeth Gould, and Tracey J Shors. Is there a link between adult neurogenesis and learning? *Hippocampus*, 16(3):216–224, 2006.

L.R. Levin, P.L. Han, P.M. Hwang, P.G. Feinstein, R.L. Davis, and R.R. Reed. The drosophila learning and memory gene rutabaga encodes a ca2+/calmodulin-responsive adenylyl cyclase. *Cell*, 68(3):479–489, 1992.

WB Levy and O. Steward. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, 8(4):791–797, 1983.

S. Li, W.K. Cullen, R. Anwyl, and M.J. Rowan. Dopamine-dependent facilitation of ltp induction in hippocampal ca1 by exposure to spatial novelty. *Nature neuroscience*, 6 (5):526–531, 2003.

XG Li, P. Somogyi, A. Ylinen, and G. Buzsaki. The hippocampal ca3 network: an in vivo intracellular labeling study. *Journal of Comparative Neurology*, 339(2):181–208, 1994.

Allison H Lin, Jonathan E Cohen, Qin Wan, Katelyn Niu, Pragya Shrestha, Steven L Bernstein, and Thomas W Abrams. Serotonin stimulation of camp-dependent plasticity in aplysia sensory neurons is mediated by calmodulin-sensitive adenylyl cyclase. *Proceedings of the National Academy of Sciences*, 107(35):15607–15612, 2010.

J. Lisman. A mechanism for the hebb and the anti-hebb processes underlying learning and memory. *Proceedings of the National Academy of Sciences*, 86(23):9574, 1989.

J. Lisman, H. Schulman, and H. Cline. The molecular basis of CaMKII function in synaptic and behavioural memory. *Nature Reviews Neuroscience*, 3(3):175–190, 2002.

J.E. Lisman and A.A. Grace. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46(5):703–713, 2005. ISSN 0896-6273.

John Lisman. The cam kinase ii hypothesis for the storage of synaptic memory. *Trends in neurosciences*, 17(10):406–412, 1994.

John Lisman, Robert C Malenka, Roger A Nicoll, and Roberto Malinow. Learning mechanisms: the case for cam-kii. *SCIENCE-NEW YORK THEN WASHINGTON-*, pages 2001–2001, 1997.

Leib Litman and Lila Davachi. Distributed learning enhances relational memory consolidation. *Learning & Memory*, 15(9):711–716, 2008.

Margaret S Livingstone, Patricia P Sziber, and William G Quinn. Loss of calcium/-calmodulin responsiveness in adenylate cyclase of¡ i¿ rutabaga¡/i¿, a drosophila learning mutant. *Cell*, 37(1):205–215, 1984.

D.M. Lovinger, K.L. Wong, K. Murakami, and A. Routtenberg. Protein kinase c inhibitors eliminate hippocampal long-term potentiation* 1. *Brain research*, 436(1): 177–183, 1987.

Gary S Lynch, Thomas Dunwiddie, and Valentin Gribkoff. Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. *Nature*, 1977.

J.C. Magee and D. Johnston. A synaptically controlled, associative signal for hebbian plasticity in hippocampal neurons. *Science*, 275(5297):209, 1997.

R.C. Malenka et al. Long-term potentiation–a decade of progress? *Science*, 285(5435): 1870, 1999.

R.C. Malenka, J.A. Kauer, R.S. Zucker, and R.A. Nicoll. Postsynaptic calcium is sufficient for potentiation of hippocampal synaptic transmission. *Science*, 242(4875):81, 1988.

Robert C Malenka, Julie A Kauer, David J Perkel, Michael D Mauk, Paul T Kelly, Roger A Nicoll, and M Neal Waxham. An essential role for postsynaptic calmodulin and protein kinase activity in long-term potentiation. *Nature*, 340(6234):554–557, 1989.

R. Malinow, H. Schulman, and R.W. Tsien. Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. *Science*, 245(4920):862, 1989.

Denise Manahan-Vaughan, Alexander Kulla, and J Uwe Frey. Requirement of translation but not transcription for the maintenance of long-term depression in the ca1 region of freely moving rats. *The Journal of Neuroscience*, 20(22):8572–8576, 2000.

F. Manseau, X. Fan, T. Hueftlein, W.S. Sossin, and V.F. Castellucci. Ca2+-independent protein kinase c apl ii mediates the serotonin-induced facilitation at depressed aplysiasensorimotor synapses. *The Journal of Neuroscience*, 21(4):1247, 2001.

H. Markram, J. Lubke, M. Frotscher, and B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297):213, 1997.

Vincenzo Marra, Michael O'Shea, Paul R Benjamin, and Ildikó Kemenes. Susceptibility of memory consolidation during lapses in recall. *Nature communications*, 4:1578, 2013.

K.C. Martin, A. Casadio, H. Zhu, et al. Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell*, 91(7):927–938, 1997a.

K.C. Martin and K.S. Kosik. Synaptic tagging who's it? *Nature Reviews Neuroscience*, 3(10):813–820, 2002.

K.C. Martin, D. Michael, J.C. Rose, M. Barad, A. Casadio, H. Zhu, and E.R. Kandel. Map kinase translocates into the nucleus of the presynaptic cell and is required for long-term facilitation in aplysia. *Neuron*, 18(6):899–912, 1997b.

SJ Martin, PD Grimwood, and RGM Morris. Synaptic plasticity and memory: an evaluation of the hypothesis. *Neuroscience*, 23(1):649, 2000.

M. Martone, N. Butters, M. Payne, J.T. Becker, and D.S. Sax. Dissociations between skill learning and verbal recognition in amnesia and dementia. *Archives of Neurology*, 41(9):965, 1984.

Juliane Mauelshagen, Carolyn M Sherff, and Thomas J Carew. Differential induction of long-term synaptic facilitation by spaced and massed applications of serotonin at sensory neuron synapses of aplysia californica. *Learning & Memory*, 5(3):246–256, 1998.

M. Mayford, T. Abel, and E.R. Kandel. Transgenic approaches to cognition. *Current Opinion in Neurobiology*, 5(2):141–148, 1995.

M. McCloskey and N.J. Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. *The psychology of learning and motivation: Advances in research and theory*, 24:109–165, 1989.

W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5(4):115–133, 1943.

R.J. McDonald and N.M. White. A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, 107(1):3–22, 1993.

R.J. McEliece, E.C. Posner, E.R. Rodemich, and S.S. Venkatesh. The capacity of the Hopfield associative memory. *IEEE Transactions on Information Theory*, 33(4):461–482, 1987. ISSN 0018-9448.

J.L. McGaugh. Memory–a century of consolidation. *Science*, 287(5451):248, 2000.

J.A. McGeoch. Forgetting and the law of disuse. *Psychological Review*, 39(4):352, 1932.

JA McGeoch. *The psychology of human learning.* Longmans, Green, 1942.

S.E. McGuire, P.T. Le, A.J. Osborn, K. Matsumoto, and R.L. Davis. Spatiotemporal rescue of memory dysfunction in drosophila. *Science*, 302(5651):1765, 2003.

M Megıas, ZS Emri, TF Freund, and AI Gulyas. Total number and distribution of inhibitory and excitatory synapses on hippocampal ca1 pyramidal cells. *Neuroscience*, 102(3):527–540, 2001.

Arthur W Melton. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9(5):596–606, 1970.

R Menzel. Memory dynamics in the honeybee. *Journal of Comparative Physiology A*, 185(4):323–340, 1999.

Claude Meunier, Hiro-Fumi Yanai, and Shun-ichi Amari. Sparsely coded associative memories: capacity and dynamical properties. *Network: Computation in Neural Systems*, 2(4):469–487, 1991.

M. Mézard, JP Nadal, and G. Toulouse. Solvable models of working memories. *Journal de physique*, 47(9):1457–1462, 1986.

Maximilian Michel, Charity L Green, Jacob S Gardner, Chelsea L Organ, and Lisa C Lyons. Massed training-induced intermediate-term operant memory in aplysia requires protein synthesis and multiple persistent kinase cascades. *The Journal of Neuroscience*, 32(13):4581–4591, 2012.

GE Miiller and A. Pilzecker. Experimentelle beiträge zur lehre vom gedächtnis. *Zeitschrift für Psychologie*, 1:1–288, 1900.

Paul Miller, Anatol M Zhabotinsky, John E Lisman, and Xiao-Jing Wang. The stability of a stochastic camkii switch: dependence on the number of enzyme molecules and protein turnover. *PLoS biology*, 3(4):e107, 2005.

B. Milner. Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. *Physiologie de lhippocampe*, pages 257–72, 1962.

B. Milner. Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, 27(3):272, 1971.

B. Milner, L.R. Squire, and E.R. Kandel. Cognitive neuroscience and the study of memory. *Neuron*, 20(3):445–468, 1998.

Guo-li Ming and Hongjun Song. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron*, 70(4):687–702, 2011.

Diego Moncada and Haydée Viola. Induction of long-term memory by exposure to novelty requires protein synthesis: evidence for a behavioral tagging. *The Journal of neuroscience*, 27(28):7476–7481, 2007.

PG Montarolo, P. Goelet, VF Castellucci, J. Morgan, ER Kandel, and S. Schacher. A critical period for macromolecular synthesis in long-term heterosynaptic facilitation in aplysia. *Science*, 234(4781):1249, 1986.

Johanna M Montgomery and Daniel V Madison. Discrete synaptic states define a major mechanism of synapse plasticity. *Trends in neurosciences*, 27(12):744–750, 2004.

I. Morgenstern. Neural networks and chip design. *European Physical Journal B*, 67: 265–270, 1987.

RGM Morris, Elizabeth Anderson, GS a Lynch, and Michel Baudry. Selective impairment of learning and blockade of long-term potentiation by an n-methyl-d-aspartate receptor antagonist, ap5. 1986.

A. Morrison, M. Diesmann, and W. Gerstner. Phenomenological models of synaptic plasticity based on spike timing. *Biological cybernetics*, 98(6):459–478, 2008.

J.R. Moyer, R.A. Deyo, and J.F. Disterhoft. Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behavioral Neuroscience*, 104(2):243–252, 1990.

Rosel M Mulkey, Caroline E Herron, and Robert C Malenka. An essential role for protein phosphatases in hippocampal long-term depression. *Science*, 261(5124):1051–1055, 1993.

E.A. Murray and M. Mishkin. Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *Journal of Neuroscience*, 18 (16):6568, 1998.

Jean-P Nadal. Associative memory: on the (puzzling) sparse coding limit. *Journal of Physics A: Mathematical and General*, 24(5):1093, 1991.

Jean-Pierre Nadal and Gerard Toulouse. Information storage in sparsely coded memory nets. *Network: Computation in Neural Systems*, 1(1):61–74, 1990.

JP Nadal, G. Toulouse, JP Changeux, and S. Dehaene. Networks of formal neurons and memory palimpsests. *EPL (Europhysics Letters)*, 1:535, 1986.

L. Nadel, J. Campbell, and L. Ryan. Autobiographical memory retrieval and hippocampal activation as a function of repetition and the passage of time. *Neural plasticity*, 2:90472, 2007.

L. Nadel and M. Moscovitch. Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7(2):217–227, 1997.

S. Navakkode, S. Sajikumar, and J.U. Frey. Synergistic requirements for the induction of dopaminergic d1/d5-receptor-mediated ltp in hippocampal slices of rat ca1 in vitro. *Neuropharmacology*, 52(7):1547–1554, 2007.

Reviel Netz and William Noel. *The Archimedes Codex*. Hachette UK, 2011.

Thomas Nevian and Bert Sakmann. Spine ca2+ signaling in spike-timing-dependent plasticity. *The Journal of neuroscience*, 26(43):11001–11013, 2006.

P.V. Nguyen, T. Abel, and E.R. Kandel. Requirement of a critical period of transcription for induction of a late phase of ltp. *Science*, 265(5175):1104, 1994.

Viacheslav O Nikolaev, Moritz Bünemann, Lutz Hein, Annette Hannawacker, and Martin J Lohse. Novel single chain camp sensors for receptor-induced signal propagation. *Journal of Biological Chemistry*, 279(36):37215–37218, 2004.

V.O. Nikolaev and M.J. Lohse. Monitoring of camp synthesis and degradation in living cells. *Physiology*, 21(2):86–92, 2006.

L. Nowak, P. Bregestovski, and P. Ascher. Magnesium gates glutamate-activated. *Nature*, 307:2, 1984.

D.H. O'Connor, G.M. Wittenberg, and S.S.H. Wang. Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9679, 2005.

J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 34:171–175, 1971.

D.S. Olton, J.A. Walker, and F.H. Gage. Hippocampal connections and spatial discrimination. *Brain Research*, 139(2):295–308, 1978.

S. Otani and WC Abraham. Inhibition of protein synthesis in the dentate gyrus, but not the entorhinal cortex, blocks maintenance of long-term potentiation in rats. *Neuroscience letters*, 106(1-2):175–180, 1989.

S. Otani, CJ Marshall, WP Tate, GV Goddard, and WC Abraham. Maintenance of long-term potentiation in rat dentate gyrus requires protein synthesis but not messenger rna synthesis immediately post-tetanization. *Neuroscience*, 28(3):519–526, 1989.

N.A. Otmakhova and J.E. Lisman. D1/d5 dopamine receptor activation increases the magnitude of early long-term potentiation at ca1 hippocampal synapses. *The Journal of neuroscience*, 16(23):7478, 1996.

M.G. Packard and L. Cahill. Affective modulation of multiple memory systems. *Current Opinion in Neurobiology*, 11(6):752–756, 2001.

M.G. Packard and J.L. McGaugh. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65(1):65–72, 1996.

M.R. Pagani, K. Oishi, B.D. Gelb, and Y. Zhong. The phosphatase SHP2 regulates the spacing effect for long-term memory induction. *Cell*, 139(1):186–198, 2009. ISSN 0092-8674.

Günther Palm. On associative memory. *Biological Cybernetics*, 36(1):19–31, 1980.

M. Päpper, R. Kempter, and C. Leibold. Synaptic tagging, evaluation of memories, and the distal reward problem. *Learning & Memory*, 18(1):58, 2011. ISSN 1072-0502.

G. Parisi. A memory which forgets. *Journal of Physics A: Mathematical and General*, 19:L617, 1986.

O. Paulsen and T.J. Sejnowski. Natural patterns of activity and long-term synaptic plasticity. *Current opinion in neurobiology*, 10(2):172–180, 2000.

I.P. Pavlov and G.V. Anrep. *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Dover, 1960.

Verena Pawlak, Jeffery R Wickens, Alfredo Kirkwood, and Jason ND Kerr. Timing is not everything: neuromodulation opens the stdp gate. *Frontiers in Synaptic Neuroscience*, 2, 2010.

D.G. Payne. Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin*, 101(1):5, 1987.

D.G. Payne and H.L. Roediger. Hypermnesia occurs in recall but not in recognition. *The American journal of psychology*, 100(2):145–165, 1987.

Karl Pearson. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900.

Pierre Peretto. *An introduction to the modeling of neural networks*, volume 2. Cambridge University Press, 1992.

C.C.H. Petersen, R.C. Malenka, R.A. Nicoll, and J.J. Hopfield. All-or-none potentiation at CA3-CA1 synapses, 1998.

G.T. Philips and T.J. Carew. It's all about timing. *Cell*, 139(1):23–25, 2009.

G.T. Philips, E.I. Tzvetkova, and T.J. Carew. Transient mitogen-activated protein kinase activation is confined to a narrow temporal window required for the induction of two-trial long-term memory in aplysia. *The Journal of Neuroscience*, 27(50):13701–13705, 2007.

Hyun Jae Pi and John E Lisman. Coupled phosphatase and kinase switches produce the tristability required for long-term potentiation and long-term depression. *The Journal of Neuroscience*, 28(49):13132–13138, 2008.

H. Pinsker, I. Kupfermann, V. Castellucci, and E. Kandel. Habituation and dishabituation of the gm-withdrawal reflex in aplysia. *Science*, 167(3926):1740, 1970.

H.M. Pinsker, W.A. Hening, T.J. Carew, and E.R. Kandel. Long-term sensitization of a defensive withdrawal reflex in aplysia. *Science*, 182(4116):1039, 1973.

Pierre-Yves Plaçais, Séverine Trannoy, Guillaume Isabel, Yoshinori Aso, Igor Siwanowicz, Ghislain Belliart-Guérin, Philippe Vernier, Serge Birman, Hiromu Tanimoto, and Thomas Preat. Slow oscillations in two pairs of dopaminergic neurons gate long-term memory formation in drosophila. *Nature neuroscience*, 15(4):592–599, 2012.

K.H. Pribram. The neurophysiology of remembering. *Scientific American*, 1969.

L.H. Prisko. *Short-term memory in focal cerebral damage.* PhD thesis, McGill University., 1963.

Dale Purves. *Principles of Neural Development.* Sinauer Associates Inc.,U.S., 1984.

W.G. Quinn, W.A. Harris, and S. Benzer. Conditioned behavior in drosophila melanogaster. *Proceedings of the National Academy of Sciences*, 71(3):708, 1974.

RJ Racine and NW Milgram. Short-term potentiation phenomena in the rat limbic forebrain. *Brain Research*, 260(2):201–216, 1983.

R. Ratcliff. Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, 97(2):285, 1990.

S.G. Rayport and S. Schacher. Synaptic plasticity in vitro: cell culture of identified aplysia neurons mediating short-term habituation and sensitization. *The Journal of neuroscience*, 6(3):759, 1986.

Roger L Redondo and Richard GM Morris. Making memories last: the synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, 12(1):17–30, 2010.

K. Reymann, U. Frey, and H. Matthies. A multi-phase model of synaptic long-term potentiation in hippocampal CA1 neurones: protein kinase C activation and protein synthesis are required for the maintenance of the trace. *Synaptic Plasticity in the Hippocampus, Springer, Berlin*, pages 126–129, 1988.

K.G. Reymann and J.U. Frey. The late maintenance of hippocampal LTP: Requirements, phases, 'synaptic tagging', 'late-associativity'and implications. *Neuropharmacology*, 52 (1):24–40, 2007. ISSN 0028-3908.

J.P.A. Richard and C. Saunders. The relationships between temporal lobe and diencephalic structures implicated in anterograde amnesia. *Memory*, 5(1):49–72, 1997.

Thomas Riemensperger, Thomas Voller, Patrick Stock, Erich Buchner, and André Fiala. Punishment prediction by dopaminergic neurons in drosophila. *Current biology*, 15 (21):1953–1960, 2005.

S. Robb, TR Cheek, FL Hannan, LM Hall, JM Midgley, and PD Evans. Agonist-specific coupling of a cloned drosophila octopamine/tyramine receptor to multiple second messenger systems. *The EMBO journal*, 13(6):1325, 1994.

H.L. Roediger and D.G. Payne. Hypermnesia: The role of repeated testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(1):66, 1982.

E. Roggenhofer, P. Fidzinski, J. Bartsch, F. Kurz, O. Shor, and J. Behr. Activation of dopamine d1/d5 receptors facilitates the induction of presynaptic long-term potentiation at hippocampal output synapses. *European Journal of Neuroscience*, 32(4): 598–605, 2010.

R. Rojas. *Neural networks: a systematic introduction*. Springer, 1996. ISBN 3540605053.

B. Roozendaal and J.L. McGaugh. Amygdaloid nuclei lesions differentially affect glucocorticoid-induced memory enhancement in an inhibitory avoidance task. *Neurobiology of learning and memory*, 65(1):1–8, 1996.

Oren S Rosenberg, Sebastian Deindl, Rou-Jia Sung, Angus C Nairn, and John Kuriyan. Structure of the autoinhibited kinase domain of camkii and saxs analysis of the holoenzyme. *Cell*, 123(5):849–860, 2005.

F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386–408, 1958.

J.I. Rossato, L.R.M. Bevilaqua, I. Izquierdo, J.H. Medina, and M. Cammarota. Dopamine controls persistence of long-term memory storage. *Science*, 325(5943):1017, 2009. ISSN 0036-8075.

D.C. Rubin, S. Hinton, and A. Wenzel. The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5):1161–1176, 1999.

D.D.B.D. Rubin and S. Fusi. Long memory lifetimes require complex synapses and limited sparseness. *Frontiers in Computational Neuroscience*, 1, 2007.

RC Rubin and AE Wenzel. 100 years of forgetting: A quantitative description of Retention. *Psychological Review*, 103:734–760, 1996.

G. Ryle. The Concept of Mind. Hutchinson's University. *Library, London*, pages 286–292, 1949.

Todd C Sacktor. How does pkmζ maintain long-term memory? *Nature Reviews Neuroscience*, 12(1):9–15, 2010.

S. Sajikumar and J.U. Frey. Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiology of learning and memory*, 82(1):12–25, 2004a. ISSN 1074-7427.

S. Sajikumar and JU Frey. Resetting of 'synaptic tags' is time-and activity-dependent in rat hippocampal CA1 in vitro. *Neuroscience*, 129(2):503–507, 2004b.

Sreedharan Sajikumar, Sheeja Navakkode, and Julietta U Frey. Identification of compartment-and process-specific molecules required for "synaptic tagging" during long-term potentiation and long-term depression in hippocampal ca1. *The Journal of Neuroscience*, 27(19):5068–5080, 2007.

Deborah Saucier and Donald P Cain. Spatial learning without nmda receptor-dependent long-term potentiation. *Nature*, 378(6553):186–189, 1995.

B. Scatton, H. Simon, M. Le Moal, and S. Bischoff. Origin of dopaminergic innervation of the rat hippocampal formation. *Neuroscience Letters*, 18(2):125–131, 1980.

S. Schacher, F. Wu, J.D. Panyko, Z.Y. Sun, and D. Wang. Expression and branch-specific export of mrna are regulated by synapse formation and interaction with specific postsynaptic targets. *The Journal of neuroscience*, 19(15):6338–6347, 1999.

D. L Schacter. Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13:501–518, 1987.

C. Schroll, T. Riemensperger, D. Bucher, J. Ehmer, T. Völler, K. Erbguth, B. Gerber, T. Hendel, G. Nagel, E. Buchner, et al. Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in drosophila larvae. *Current biology*, 16(17):1741–1747, 2006.

M. Schwaerzel, M. Monastirioti, H. Scholz, F. Friggi-Grelin, S. Birman, and M. Heisenberg. Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in drosophila. *The Journal of neuroscience*, 23(33):10495, 2003.

J.H. Schwartz, V.F. Castellucci, and ER Kandel. Functioning of identified neurons and synapses in abdominal ganglion of aplysia in absence of protein synthesis. *Journal of Neurophysiology*, 34(6):939, 1971.

WB Scoville. The Limbic Lobe in Man*. *Journal of Neurosurgery: Pediatrics*, 11(1), 1954.

WB Scoville and B. Milner. Loss of recent memory after bilateral hippocampal lesions . *Journal of Neurology, Neurosurgery & Psychiatry.*, 20(1):11–21, 1957.

Julien Séjourné, Pierre-Yves Plaçais, Yoshinori Aso, Igor Siwanowicz, Séverine Trannoy, Vladimiros Thoma, Stevanus R Tedjakumala, Gerald M Rubin, Paul Tchénio, Kei Ito, et al. Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in drosophila. *Nature neuroscience*, 14(7):903–910, 2011.

W. Senn and S. Fusi. Convergence of stochastic learning in perceptrons with binary synapses. *Physical Review E*, 71(6):061907, 2005.

Geun Hee Seol, Jokubas Ziburkus, ShiYong Huang, Lihua Song, In Tae Kim, Kogo Takamiya, Richard L Huganir, Hey-Kyoung Lee, and Alfredo Kirkwood. Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron*, 55 (6):919–929, 2007.

Qiang Shan, Guy C-K Chan, and Daniel R Storm. Type 1 adenylyl cyclase is essential for maintenance of remote contextual fear memory. *The Journal of Neuroscience*, 28 (48):12864–12867, 2008.

S.K. Sharma, C.M. Sherff, J. Shobe, M.W. Bagnall, M.A. Sutton, and T.J. Carew. Differential role of mitogen-activated protein kinase in three distinct phases of memory for sensitization in aplysia. *The Journal of neuroscience*, 23(9):3899–3907, 2003.

Reut Shema, Sharon Haramati, Shiri Ron, Shoshi Hazvi, Alon Chen, Todd Charlton Sacktor, and Yadin Dudai. Enhancement of consolidated long-term memory by overexpression of protein kinase mζ in the neocortex. *Science*, 331(6021):1207–1210, 2011.

D.F. Sherry and D.L. Schacter. The evolution of multiple memory systems. *Psychological Review*, 94(4):439–454, 1987.

A.P. Shimamura and L.R. Squire. A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (3):464–473, 1987.

D. Shohamy and R.A. Adcock. Dopamine and adaptive memory. *Trends in cognitive sciences*, 2010.

K. Si, M. Giustetto, A. Etkin, R. Hsu, A.M. Janisiewicz, M.C. Miniaci, J.H. Kim, H. Zhu, and E.R. Kandel. A neuronal isoform of cpeb regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in aplysia. *Cell*, 115(7):893–904, 2003a.

K. Si, S. Lindquist, and E.R. Kandel. A neuronal isoform of the aplysia CPEB has prion-like properties. *Cell*, 115(7):879–891, 2003b.

M. Sidman, LT Stoddard, and JP Mohr. Some additional quantitative observations of immediate memory in a patient with bilateral hippocampal lesions. *Neuropsychologia*, 6(3):245–254, 1968.

S. Sikstrom. Power function forgetting curves as an emergent property of biologically plausible neural network models. *International Journal of Psychology*, 34(5):460–464, 1999.

A.J. Silva, J.H. Kogan, P.W. Frankland, and S. Kida. Creb and memory. *Annual Review of Neuroscience*, 21(1):127–148, 1998.

C.B. Sindreu, Z.S. Scheiner, and D.R. Storm. Ca¡ sup¿ 2+¡/sup¿-stimulated adenylyl cyclases regulate erk-dependent activation of msk1 during fear conditioning. *Neuron*, 53(1):79–89, 2007.

P.J. Sjöström, G.G. Turrigiano, and S.B. Nelson. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–1164, 2001.

Paul Smolen. A model of late long-term potentiation simulates aspects of memory maintenance. *PloS one*, 2(5):e445, 2007.

Paul Smolen, Douglas A Baxter, and John H Byrne. A model of the roles of essential kinases in the induction and expression of late long-term potentiation. *Biophysical journal*, 90(8):2760–2775, 2006.

H. Sompolinsky. Neural networks with nonlinear synapses and a static noise. *Physical Review A*, 34(3):2571, 1986.

H. Sompolinsky. The theory of neural networks: The Hebb rule and beyond. In *Lecture Notes in Physics, Berlin Springer Verlag*, volume 275, pages 485–527, 1987.

L. R. Squire. *Closing remarks In The biology of memory* . Stuggart, Germany: F. K. Schattauer Verlag., 1990.

L.R. Squire. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99(2):195–231, 1992.

L.R. Squire. Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory*, 82(3):171–177, 2004.

L.R. Squire and P. Alvarez. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Current opinion in neurobiology*, 5(2):169–177, 1995.

L.R. Squire, B. Knowlton, and G. Musen. The structure and organization of memory. *Annual Review of Psychology*, 44(1):453–495, 1993.

L.R. Squire and S. Zola-Morgan. The medial temporal lobe memory system. *Science*, 253(5026):1380, 1991.

L.R. Squire, S. Zola-Morgan, and K.S. Chen. Human amnesia and animal models of amnesia: Performance of amnesic patients on tests designed for the monkey. *Behavioral Neuroscience*, 102(2):210–221, 1988.

Ursula Staubli and Gary Lynch. Stable depression of potentiated synaptic responses in the hippocampus with 1–5 hz stimulation. *Brain research*, 513(1):113–118, 1990.

RJ Steele and RGM Morris. Delay-dependent impairment of a matching-to-place task with chronic and intrahippocampal infusion of the NMDA-antagonist D-AP5. *Hippocampus*, 9(2):118–136, 1999.

O. Steward and W.B. Levy. Preferential localization of polyribosomes under the base of dendritic spines in granule cells of the dentate gyrus. *The Journal of Neuroscience*, 2 (3):284, 1982.

R. Stickgold. Sleep-dependent memory consolidation. *Nature*, 7063:1272, 2005.

M.A. Sutton, J. Ide, S.E. Masters, and T.J. Carew. Interaction between amount and pattern of training in the induction of intermediate-and long-term memory for sensitization in aplysia. *Learning & Memory*, 9(1):29–40, 2002.

L.W. Swanson and G.D. Petrovich. What is the amygdala? *Trends in neurosciences*, 21(8):323–331, 1998.

JL Swanson-Park, CM Coussens, SE Mason-Parker, CR Raymond, EL Hargreaves, M. Dragunow, AS Cohen, and WC Abraham. A double dissociation within the hippocampus of dopamine d1/d5 receptor and [beta]-adrenergic receptor contributions to the persistence of long-term potentiation. *Neuroscience*, 92(2):485–497, 1999.

J.D. Sweatt. Mitogen-activated protein kinases in synaptic plasticity and memory. *Current opinion in neurobiology*, 14(3):311–317, 2004. ISSN 0959-4388.

R.F. Thompson. The neural basis of basic associative learning of discrete behavioral responses. *Trends in Neurosciences*, 11(4):152–155, 1988.

R.F. Thompson and D.J. Krupa. Organization of memory traces in the mammalian brain. *Annual Review of Neuroscience*, 17(1):519–549, 1994.

E.L. Thorndike. *Educational psychology, Vol 2: The psychology of learning.* Teachers College, 1913.

E.C. Tolman. There is more than one kind of learning. *Psychological Review*, 56(3): 144–155, 1949.

S.M. Tomchik and R.L. Davis. Dynamics of learning-related camp signaling and stimulus integration in the drosophila olfactory pathway. *Neuron*, 64(4):510–521, 2009.

G. Toulouse, S. Dehaene, and J.P. Changeux. Spin glass model of learning by selection. *Proceedings of the National Academy of Sciences*, 83(6):1695, 1986.

Thomas Trappenberg. *Fundamentals of computational neuroscience.* Oxford University Press, 2010.

D. Tse, R.F. Langston, M. Kakeyama, I. Bethus, P.A. Spooner, E.R. Wood, M.P. Witter, and R.G.M. Morris. Schemas and memory consolidation. *Science*, 316(5821):76, 2007. ISSN 0036-8075.

J.Z. Tsien, P.T. Huerta, S. Tonegawa, et al. The essential role of hippocampal ca1 nmda receptor-dependent synaptic plasticity in spatial memory. *Cell*, 87(7):1327–1338, 1996.

MV Tsodyks. Associative memory in neural networks with binary synapses. *MOD. PHYS. LETT. B.*, 4(11):713–716, 1990.

MV Tsodyks and MV Feigel'Man. The enhanced storage capacity in neural networks with low activity level. *EPL (Europhysics Letters)*, 6:101, 1988.

T. Tully, T. Preat, SC Boynton, and M. Del Vecchio. Genetic dissection of consolidated memory in drosophila. *Cell*, 79(1):35–47, 1994.

E. Tulving. The effects of presentation and recall of material in free-recall learning1. *Journal of Verbal Learning and Verbal Behavior*, 6(2):175–184, 1967.

E. Tulving. *Elements of episodic memory*. Clarendon Press/Oxford University Press, 1983.

E. Tulving. Organization of memory: Quo vadis. *The cognitive neurosciences*, pages 839–847, 1995.

E. Tulving. Episodic memory: From mind to brain. *Annual review of psychology*, 2002.

E. Tulving, D.L. Schacter, and H.A. Stark. Priming effects in word-fragment completion are independent of recognition memory. *Journal of experimental psychology: learning, memory, and cognition*, 8(4):336–342, 1982.

Najl V Valeyev, Declan G Bates, Pat Heslop-Harrison, Ian Postlethwaite, and Nikolay V Kotov. Elucidating the mechanisms of cooperative calcium-calmodulin interactions: a structural systems biology approach. *BMC systems biology*, 2(1):48, 2008.

L. Vallar and J. Meldolesi. Mechanisms of signal transduction at the dopamine d2 receptor. *Trends in pharmacological sciences*, 10(2):74–77, 1989.

J. L. van Hemmen. Nonlinear neural networks near saturation. *Phys. Rev. A*, 36(4): 1959–1962, Aug 1987.

Mark CW van Rossum, Maria Shippi, and Adam B Barrett. Soft-bound synaptic plasticity increases storage capacity. *PLoS computational biology*, 8(12):e1002836, 2012.

A. Volianskis and M.S. Jensen. Transient and sustained types of long-term potentiation in the CA1 area of the rat hippocampus. *The Journal of Physiology*, 550(2):459, 2003.

Lenora J Volk, Julia L Bachman, Richard Johnson, Yilin Yu, and Richard L Huganir. Pkm-[zgr] is not required for hippocampal synaptic plasticity, learning and memory. *Nature*, 493(7432):420–423, 2013.

G.V. Wallenstein, M.E. Hasselmo, and H. Eichenbaum. The hippocampus as an associator of discontiguous events. *Trends in Neurosciences*, 21(8):317–323, 1998.

R. Waltereit and M. Weller. Signaling from camp/pka to mapk and synaptic plasticity. *Molecular neurobiology*, 27(1):99–106, 2003a.

Robert Waltereit and Michael Weller. Signaling from camp/pka to mapk and synaptic plasticity. *Molecular Neurobiology*, 27:99–106, 2003b. ISSN 0893-7648. URL: http://dx.doi.org/10.1385/MN:27:1:99. 10.1385/MN:27:1:99.

Hongbing Wang and Daniel R Storm. Calmodulin-regulated adenylyl cyclases: cross-talk and plasticity in the central nervous system. *Molecular pharmacology*, 63(3):463–468, 2003.

H.X. Wang, R.C. Gerkin, D.W. Nauen, and G.Q. Bi. Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nature neuroscience*, 8(2):187–193, 2005.

S.H. Wang and R.G.M. Morris. Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annual review of psychology*, 61:49–79, 2010. ISSN 0066-4308.

Szu-Han Wang, Roger L Redondo, and Richard GM Morris. Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proceedings of the National Academy of Sciences*, 107(45):19537–19542, 2010.

L. Warrington, E.K. & Weiskrantz. A new method for testing long-term retention with special reference to amnesic patients. *Nature*, 217:972–974, 1968.

Anne E West, Wen G Chen, Matthew B Dalva, Ricardo E Dolmetsch, Jon M Kornhauser, Adam J Shaywitz, Mari A Takasu, Xu Tao, and Michael E Greenberg. Calcium regulation of neuronal gene expression. *Proceedings of the National Academy of Sciences*, 98(20):11024–11031, 2001.

E.M. Wexler and P.K. Stanton. Priming of homosynaptic long-term depression in hippocampus by previous synaptic activity. *Neuroreport*, 4(5):591, 1993.

N.M. White and R.J. McDonald. Multiple parallel memory systems in the brain of the rat. *Neurobiology of Learning and Memory*, 77(2):125–184, 2002.

W.A. Wickelgren. Single-trace fragility theory of memory dynamics. *Memory & Cognition*, 2(4):775–780, 1974.

N. Wiener. *Cybernetics*. J. Wiley, 1948.

H. Wigström and B. Gustafsson. On long-lasting potentiation in the hippocampus. a proposed mechanism for its dependence on coincident pre-and postsynaptic activity. *Acta physiologica scandinavica*, 123(4):519–522, 1985.

D.B. Willingham and K. Goedert. The role of taxonomies in the study of human memory. *Cognitive, affective, & behavioral neuroscience*, 1(3):250, 2001.

Debbie Willoughby and Dermot MF Cooper. Ca2+ stimulation of adenylyl cyclase generates dynamic oscillations in cyclic amp. *Journal of cell science*, 119(5):828–836, 2006.

D.J. Willshaw, O.P. Buneman, and H.C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222(7):960–962, 1969.

J.T. Wixted and S.K. Carpenter. The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science*, 18(2):133, 2007.

J.T. Wixted and E.B. Ebbesen. On the form of forgetting. *Psychological science*, 2(6): 409–415, 1991.

S.T. Wong, J. Athos, X.A. Figueroa, V.V. Pineda, M.L. Schaefer, C.C. Chavkin, L.J. Muglia, and D.R. Storm. Calcium-stimulated adenylyl cyclase activity is critical for hippocampus-dependent long-term memory and late phase ltp. *Neuron*, 23(4):787–798, 1999.

Newton H Woo, Steven N Duffy, Ted Abel, and Peter V Nguyen. Temporal spacing of synaptic stimulation critically modulates the dependence of ltp on cyclic amp-dependent protein kinase. *Hippocampus*, 13(2):293–300, 2003.

N.H. Woo and P.V. Nguyen. Silent Metaplasticity of the Late Phase of Long-Term Potentiation Requires Protein Phosphatases. *Learning & Memory*, 9(4):202, 2002.

G.Y. Wu, K. Deisseroth, and R.W. Tsien. Spaced stimuli stabilize mapk pathway activation and its effects on dendritic morphology. *Nat Neurosci*, 4(2):151–158, 2001.

Zhi-Liang Wu, Steven A Thomas, Enrique C Villacres, Zhengui Xia, Michele L Simmons, Charles Chavkin, Richard D Palmiter, and Daniel R Storm. Altered behavior and long-term potentiation in type i adenylyl cyclase mutant mice. *Proceedings of the National Academy of Sciences*, 92(1):220–224, 1995.

Zhengui Xia and Daniel R Storm. The role of calmodulin as a signal integrator for synaptic plasticity. *Nature Reviews Neuroscience*, 6(4):267–276, 2005.

X. Xie, G. Barrionuevo, and TW Berger. Differential expression of short-term potentiation by AMPA and NMDA receptors in dentate gyrus. *Learning & Memory*, 3(2-3): 115, 1996.

KK Yamamoto, GA Gonzalez, WH Biggs 3rd, and MR Montminy. Phosphorylation-induced binding and transcriptional efficacy of nuclear factor creb. *Nature*, 334(6182): 494, 1988.

Peter V Yan-You Huang, Ted Abel Nguyen, and Eric R Kandel. Long-lasting forms of synaptic potentiation in the mammalian hippocampus. *Learning & Memory*, 1996.

S.N. Yang, Y.G. Tang, and R.S. Zucker. Selective induction of ltp and ltd by postsynaptic [ca2+] i elevation. *Journal of neurophysiology*, 81(2):781, 1999.

Ayse Yarali, Johannes Nehrkorn, Hiromu Tanimoto, and Andreas VM Herz. Event timing in associative learning: From biochemical reaction dynamics to behavioural observations. *PloS one*, 7(3):e32885, 2012.

JCP Yin, M. Del Vecchio, H. Zhou, and T. Tully. Creb as a memory modulator: induced expression of a dcreb2 activator isoform enhances long-term memory in drosophila. *Cell*, 81(1):107–115, 1995.

Y. Yovell, E.R. Kandel, Y. Dudai, and T.W. Abrams. A quantitative study of the ca2+/calmodulin sensitivity of adenylyl cyclase in aplysia, drosophila, and rat. *Journal of neurochemistry*, 59(5):1736–1744, 1992.

Manuela Zaccolo and Tullio Pozzan. Discrete microdomains with high concentration of camp in stimulated rat neonatal cardiac myocytes. *Science*, 295(5560):1711–1715, 2002.

A.M. Zhabotinsky. Bistability in the Ca^{2+}/calmodulin-dependent protein kinase-phosphatase system. *Biophysical Journal*, 79(5):2211–2221, 2000.

Ming Zhang, Changjong Moon, Guy C-K Chan, Lan Yang, Fei Zheng, Alana C Conti, Lisa Muglia, Louis J Muglia, Daniel R Storm, and Hongbing Wang. Ca-stimulated type 8 adenylyl cyclase is required for rapid acquisition of novel spatial information and for working/episodic-like memory. *The Journal of Neuroscience*, 28(18):4736–4744, 2008.

Yili Zhang, Rong-Yu Liu, George A Heberton, Paul Smolen, Douglas A Baxter, Leonard J Cleary, and John H Byrne. Computational design of enhanced learning protocols. *Nature neuroscience*, 15(2):294–297, 2012.

Y. Zhou, J. Won, M.G. Karlsson, M. Zhou, T. Rogerson, J. Balaji, R. Neve, P. Poirazi, and A.J. Silva. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nature Neuroscience*, 12(11):1438–1443, 2009.

S.M. Zola-Morgan and L.R. Squire. The primate hippocampal formation: evidence for a time-limited role in memory storage. *Science*, 250(4978):288, 1990.