



PHD IN COMPUTER SCIENCE

FUNCTIONAL BRAIN OSCILLATIONS:

HOW OSCILLATIONS FACILITATE INFORMATION

REPRESENTATION AND CODE MEMORIES

BY GEORGE PARISH

SUPERVISED BY PROFESSOR HOWARD BOWMAN

51,925 words over 187 pages

TABLE OF CONTENTS

1. Introduction	1
2. Computational Neuroscience	8
2.1 Neurophysiology	9
2.1.1 Membrane Current.....	10
2.1.2 Hodgkin & Huxley Model	18
2.1.3 Integrate & fire model	19
2.1.4 Rate-Coded Modelling	21
2.2 Networks of Neurons	22
2.2.1 Post-synaptic-potential	22
2.2.2 Background Noise	24
2.2.3 Spike-Time-Dependent-Plasticity (STDP)	26
2.2.4 Information Transmission	32
2.3 Summary	36
3. Human Memory	37
3.1 Complimentary Learning Systems	37
3.1.1 Neurological Counterparts.....	38
3.1.2 Memory Interactions	39
3.2 Episodic Memory	41
3.2.1 On What & Where	42

3.2.2 On Time	46
3.4 Summary	53
4. Functional Oscillations in the Brain	54
4.1 Oscillation Functionality.....	55
4.1.1 Excitation-Inhibition Balance	55
4.1.2 Amplitude & Frequency	57
4.1.3 Phase	59
4.2 Oscillations & Memory.....	61
4.2.1 Autoassociative Memory in Networks with Theta/Gamma Oscillations	61
4.2.2 How Inhibitory Oscillations Can Train Neural Networks.....	65
4.2.3 Theta Coordinated Error-Driven Learning	67
4.3 Oscillations & Information Processing	71
4.3.1 Information vis Desynchronisation	72
4.3.2 The Inhibition-Timing-Hypothesis.....	76
4.4 Summary	79
5. The Sync/deSync Conundrum	82
5.1 Sync, deSync and Memory Formation	83
5.1.1 Hippocampal Theta	84
5.1.2 Neo-Cortical Alpha	85
5.1.3 A Note on Other Frequencies	85

5.2 Materials and methods	86
5.2.1 Modelling Principles & Procedure	86
5.2.2 Neuron physiology	89
5.2.3 Neo-Cortical system	91
5.2.4 Hippocampal system	93
5.2.5 Simulation procedure	98
5.3 Summary	99
6. The Sync/deSync Model - Results	100
6.1 Results	100
6.1.1 Hippocampal weight change	101
6.1.2 Hippocampal activity	103
6.1.3 Theta phase	106
6.1.4 Alpha De-Synchronisation	108
6.1.5 Theta Synchronisation	110
6.1.6 Varying Stimulus Strength	111
6.1.7 Sync/De-Synch Predicts Learning	117
6.2 Discussion	119
6.2.1 Relevance to Other Models	120
6.2.2 Relevance to Literature & Predictions	121
6.2.3 Future Works	123

7. Time-Keeping in the Brain.....	125
7.1 Introduction	126
7.2 Theoretical Framework.....	129
7.3 Model Architecture.....	133
7.3.1 Neurophysiology	134
7.3.2 Neo-Cortical Alpha	136
7.3.3 The Binding Pool	138
7.3.4 Hierarchical Synfire Chains.....	141
7.3.5 Neuro-Modulators	148
7.4 Experimental Paradigm.....	149
7.5 Summary	151
8. Hierarchical Synfire Chain Model - Results	153
8.1 Results.....	153
8.1.1 Cortical Pattern Representation	153
8.1.2 Binding Discrete Events.....	155
8.1.3 Recalling Episodic Sequences.....	162
8.1.4 Content Specific Reinstatement of Temporal Patterns	167
8.2 Discussion.....	171
9. Discussion.....	180
References	188
Appendix A – Further Oscillations Literature.....	197

Frequency & Local-Field-Potential	197
Phase Synchrony	204
Phase Synchrony and Memory	209

The overall aim of the modelling works within this thesis is to lend theoretical evidence to empirical findings from the brain oscillations literature. Oscillations occur when groups of neurons (nerve cells that form the base computational unit within the brain) discharge within a precise time window, repeating at a particular frequency (Fell & Axmacher, 2011). We therefore hope to solidify and expand the notion that precise spike timing through oscillatory mechanisms facilitates communication (Fell & Axmacher, 2011), learning (Backus, et al., 2016; Burke, et al., 2014; Heusser, et al., 2016), information processing (Hanslmayr, et al., 2012; Klimesch, et al., 2007) and information representation (Michelmann, et al., 2016) within the brain. Whilst other models in this field typically use connectionist methods in a top-down manner, i.e. remain relatively distant from neurophysiology (Ketz, et al., 2013; Norman, et al., 2006; O'Reilly, et al., 2011), here we attempt to model oscillatory phenomena in a bottom-up fashion through the simulation of single cells and the neurophysiological interactions between them. In order to fulfil this overall criterion, this thesis begins by describing the field of computational neuroscience in Chapter 2, where the equations used to model individual neurons (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000) and the underpinning physiology (Shepherd, 1994) are described.

The primary hypothesis of this thesis is that it can be shown computationally that neural desynchronisations can allow information content to emerge, as has been argued elsewhere (Hanslmayr, et al., 2012; Michelmann, et al., 2016). We do this using two neural network models, the first of which shows how differential rates of neuronal firing can indicate when a single item is being actively represented (Parish, et al., 2018). The second model expands this notion by creating a complimentary timing mechanism, thus enabling the emergence of qualitative temporal information when a pattern of items is being actively represented, as has been shown in a recent study (Michelmann, et al., 2016).

The secondary hypothesis of this thesis is that it can be also be shown computationally that oscillations might play a functional role in learning. We do this by firstly expanding the use of the complimentary learning systems (CLS) framework (O'Reilly, et al., 2011). In this proposed framework, distinct brain regions are thought to interact in the learning and storage of information. Specifically, the older archi-cortical regions (the hippocampus and medial-temporal-lobe) are thought to quickly bind together our experiences in an on-line fashion (Squire, et al., 2004), whereas the newer neo-cortical regions are thought to consolidate a version of these experiences in long-term memory (Winocur, et al., 2010), possibly during sleep (Walker, 2018). Both of the models presented within this thesis propose a sparsely coded and fast learning hippocampal region that engages in the binding of novel episodic information (Squire, 1992; Squire, et al., 2004), which are then integrated within a distributed store of contextualised memories (Rumelhart, et al., 1986) located in neo-cortical outer layers of the brain (McClelland, et al., 1995). In this respect, the first model of this thesis, demonstrates how active cortical representations enable learning to occur in their hippocampal counterparts via a phase-dependent learning rule (Parish, et al., 2018). The second model expands this notion, creating hierarchical temporal sequences to encode the relative temporal position of a cortical representation. Therefore, we provide computational mechanisms that demonstrate how the activation of simulated hippocampal units relative to ongoing oscillatory phase can be crucial for synaptic learning, as has been shown in several studies (Huerta & Lisman, 1995; Pavlides, et al., 1988) and computational models (Hasselmo, 2005; Jensen, et al., 1996; Ketz, et al., 2013; Norman, et al., 2006).

We demonstrate in both of these models, how cortical brain oscillations might provide a gating function to the representation of information, whilst complimentary hippocampal oscillations might provide distinct phases of learning in a CLS framework. Chapters 3 & 4 therefore detail the experimental evidence behind such a hypothesis. This begins with an overview of human episodic memory in Chapter 3, where we discuss the rationale behind and evidence for the CLS framework (O'Reilly, et al., 2011), as well as an introduction to the notion of time-keeping in the brain through the evaluation of several neural network models (Bowman & Wyble, 2007; Goldman, 2009; Itskov,

et al., 2011; Shankar & Howard, 2012). Chapter 4 then details how oscillations have been found to play crucial roles in many brain functions (Basar, et al., 2001; Fell & Axmacher, 2011), in particular within attention (Klimesch, et al., 2007; Jensen & Mazaheri, 2010) and memory (Buzsaki, 2002; Hanslmayr, et al., 2012; Khader, et al., 2010) processes.

An oscillation is primarily observed through non-invasive methods, such as electro-encephalography (EEG) and magneto-encephalography (MEG), whereby the electrical currents that flow through neurons can be traced with high temporal - but low spatial – resolution (Neidermeyer & da Silva, 2004). Through this means, one can detect the mass activation of a large group of neurons that typically follows some internal processing in response to an external stimulus. One can capture higher spatial resolution of the same phenomenon by using more invasive methods, such as intra-cranial electro-encephalography (iEEG), which involves the placement of electrodes within the skull, as well as single cell recording, where a microelectrode inserted within the brain can record from a single or small group of nerve cells. From this, we can detect when a neuron sends a spike event to its neighbours, whereby it discharges an electrical signal in an all-or-nothing response (Dayan & Abbot, 2001; O'Reilly & Munakata, 2000).

All of these methods suggest that brain processes are efficiently organised through a series of interacting oscillations (Basar, et al., 2001; Fell & Axmacher, 2011), where several frequencies can co-exist within the same brain region. One can observe oscillatory changes through adjustments in the amplitude of specific frequencies (Neidermeyer & da Silva, 2004), which is thought to be dictated by the rhythmicity of spike events that occur within the up-state of that frequency. As an oscillation increases in amplitude, it is said that there is a synchronisation in that specific frequency, as more neurons in that region are entrained to activate more regularly within oscillatory up-states. A de-synchronisation is said to occur when neurons fire asynchronously with no specific organisation. One of the ways that this is thought to occur is if neuronal activation becomes so high that it overcomes inhibitory down-states to fire more often (Klimesch, et al., 2007), one of the central hypotheses of this thesis. This might occur during stimulus-related activation (Haegens, et

al., 2011), whereby neo-cortical de-synchronisations have been observed over brain regions that activate for specific stimuli or functions (Berger, 1929; Adrian & Matthews, 1934). From a mathematical viewpoint, information gain is thought to be minimal during periods of high synchronisation, as all assemblies oscillate in unison (Klimesch, et al., 2007). However, it is thought that desynchronisations enable information flow, as specific assemblies can be active when all others are silent (Hanslmayr, et al., 2012).

Oscillatory phase has also been found to play a crucial role in memory formation (Buzsaki, 2002; Fell & Axmacher, 2011). The phase describes when activation of a single cell or cell assembly takes place in relation to a reference oscillation. For example, if a local spike event occurs during the up-state of the reference oscillation it is said to occur at the peak of that oscillatory frequency, whereas if it occurs during the down-state it is said to occur at the trough (Fell & Axmacher, 2011). Single cell recordings within the hippocampus suggest that stimulation of neurons during alternating phases of the ongoing oscillation produce different effects (Huerta & Lisman, 1995). Here, the strength of the connections between neurons was found to increase when neurons were stimulated to spike in the opposite phase to the local reference oscillation. This showed that local learning might be dictated by global oscillatory phase (Hasselmo, 2005), whereby synaptic change might occur between any active assemblies that shift in phase together to a time when other assemblies were silent (Hanslmayr, et al., 2016; Hasselmo, et al., 2002).

The findings described thus far form the basis of the first computational model within this thesis (Parish, et al., 2018), described in Chapters 5 & 6. Throughout the neural oscillations literature, the inconsistency of findings has sewn discord regarding the role of oscillations in human episodic memory formation (Hanslmayr, et al., 2016). For the most part, a mid-frequency neocortical desynchronisation and a low-frequency hippocampal synchronisation have both been associated with successful memory formation (Backus, et al., 2016; Burke, et al., 2014; Hanslmayr, et al., 2012; Khader, et al., 2010), although a smaller number of experiments have presented different findings with respect to the latter (Crespo-Garcia, et al., 2016; Fellner, et al., 2016; Greenberg, et al., 2015).

The purpose of this model is to explore an alternative theory that might reconcile the argument as to whether synchronisation is a good predictor for memory (Hanslmayr, et al., 2016).

In this first modelling work, a CLS framework is realised through the creation of a dual neocortical and hippocampal network model (O'Reilly, et al., 2011). The model describes how large amplitude cortical activation desynchronises the ongoing oscillation, thus enabling information flow (Hanslmayr, et al., 2012). This then triggers stimulus-specific hippocampal assemblies to increase in activity, as observed experimentally (Ison, et al., 2015; Quiroga, 2012). Here, the majority of spikes occur phase-locked to the reference oscillation, a key predictor of memory formation (Rutishauser, et al., 2010), whilst some spikes overcome inhibition to fire in the down-state of the reference oscillation. The increase in rhythmic activation synchronises the hippocampal frequency, whilst activation in the trough of the reference oscillation enables selective learning via a phase dependent learning mechanism, similar to other models (Hasselmo, et al., 2002; Ketz, et al., 2013; Norman, et al., 2006). Thus, the model shows how both oscillatory mechanisms might be necessary for memory formation in a CLS framework (Parish, et al., 2018). We also show that for varying degrees of stimulus amplitude, one can capture the evolution of frequency synchronisation to desynchronisation, as posited by other theories (Klimesch, et al., 2007). Within the model, this process occurs at different rates in the cortex and hippocampus due to variation in neuronal firing rates. Therefore, through parameter evaluation, we are able to describe many disparate findings that have been found within the oscillations and memory field (Hanslmayr, et al., 2016), providing a theoretical platform from which to conduct further experiments.

Evidence has also been provided suggesting that qualitative temporal information can be decoded from phase-reset patterns that occur during cortical desynchronisations (Michelmann, et al., 2016). A phase-reset occurs when some event triggers the ongoing oscillation to complete its cycle early, restarting it with a relative phase shift (Canavier, 2015). This mechanism allows the brain to synchronise phase between regions, thus enabling efficient communication over long distances through the creation of shared windows of activation (Fell & Axmacher, 2011). A recent study has

identified unique phase reset patterns when subjects viewed video clips (Michelmann, et al., 2016), where a reset occurs upon changes in visual scenes as the relevant cortical assemblies activate. An examiner could then decipher when a subject later thought about a specific clip through the detection of these unique temporal signatures over cortical areas. Whilst the first modelling work of this thesis explored how a period of desynchronisation can signal information flow, the second modelling work of this thesis will go further by using this signal to actually decipher information content.

In order to achieve this, our second model (described in Chapters 7 & 8) similarly realises a CLS framework with dual neocortical and hippocampal systems (O'Reilly, et al., 2011), though with additional complexity to account for further behavioural requirements. As such, we note the importance of developing a more dynamic cortical oscillation through interacting populations (Brunel, 2000; Hansel & Mato, 2003), rather than feeding a stationary alternating signal into populations as in the first model (Parish, et al., 2018). This enables events to reset the ongoing phase as well as desynchronise the ongoing frequency, such that a sequential pattern of events can then be presented and a temporal signature for each unique phase-reset pattern can be acquired. Importantly, our second model must also account for the encoding of temporal sequences. Therefore, our hippocampal region takes inspiration from time cells (MacDonald, et al., 2011) and synfire chains (Diesmann, et al., 1999), where we design a novel time keeping mechanism that stores the temporal location of each event. Time cells have been identified in the hippocampus that aptly encode for our perception of time, where cells fire sequentially to cover a whole temporal period (Eichenbaum, 2014). A synfire chain is a structure of sequentially connected neuronal groups, where activation spreads in sequence (Diesmann, et al., 1999). Previous models have used the latter structure to describe time cell behaviour (Goldman, 2009; Itskov, et al., 2011), though criticisms have emerged regarding whether such a solution is scalable (Shankar & Howard, 2012). This thesis presents an alternative model that might reduce the computational cost of using synfire chains to encode for time. This model stores a temporal sequence through the use of an event triggered binding pool, similar to other models (Bowman & Wyble, 2007), which can subsequently

be replayed, thus enabling the distinction between multiple simulated cortical phase-reset patterns as observed experimentally (Michelmann, et al., 2016).

The second model thus builds on the first hypotheses of this thesis by demonstrating computationally how information content might be deciphered from oscillatory desynchronisations. We also theorise that our mechanism for encoding hierarchical temporal sequences might provide a functional role for oscillations in the segregation of temporal episodes, furthering the second hypothesis of our thesis that learning is dependent on spike timing relative to the phase of on-going oscillatory activity. This thesis concludes with an exploration of the effectiveness of both of these models in satisfying these hypotheses, presented in Chapter 9.

Computational neuroscience is a branch of neuroscience that aims to understand brain functionality through mathematical models and statistical analysis (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). This can be done with varying levels of abstraction, dependent on the type of modelling work being pursued. Some models focus on describing biophysical neural components, such as Hodgkin & Huxley's original description of the transparent squid axon in 1952. By such means, one aims to capture emergent cognitive operations from the simulation of small interactions in a bottom-up approach. In this way, one can be more confident in generating spatio-temporal signals that can be directly compared to those recorded from the brain during experiments (Dayan & Abbot, 2001), thus framing hypothesis for future experimental work.

Alternatively, one can abstract from fine-grain neural interactions, representing the activity of a single or a group of cells with nodes (Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). As the brain is a vastly distributed information store (Rumelhart, et al., 1986) with many interacting regions (McClelland, et al., 1995), it is difficult to isolate and simulate areas of interest. Therefore, connectionist models take a step back to produce theoretical schematics to understand large-scale neural interactions, providing powerful and innovative solutions to cognitive operations (O'Reilly & Munakata, 2000). For every observable neural phenomenon, there might be several sets of equations that vary in abstraction depending on the scope of the modelling work. As it can be computationally expensive to simulate an amalgamation of the most biophysically plausible equations (Dayan & Abbot, 2001), many models tend to operate on an Ockham's Razor principle, where the simplest solution is often the most preferable (O'Reilly & Munakata, 2000).

As this thesis explores neural network models of varying abstraction, the following Chapter will present the required level of neurophysiology necessary to understand this body of modelling work. Whilst multiple abstractions of various cognitive operations will be examined, we do not capture the whole biophysical complexity that exists in the field, as it lies outside the scope of this thesis.

2.1 NEUROPHYSIOLOGY

A neuron can be thought of as the base computational unit in the brain (Abeles, 1991). The defining feature of this nerve cell is its ability to transmit information to surrounding cells (Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). This means they are prevalent in the brain and nervous system (Shepherd, 1994), the purposes of which are to receive and transmit signals across the body. Like other living cells, a neuron consists of a cytoplasm surrounded by a cell membrane (Figure 1A; see Shepherd, 1994, for overview). The cytoplasm is made up of different types of ions with a positive or negative charge. Ion channels embedded in the cell membrane (Figure 1C) act as tunnels through the impermeable membrane that allow ions to flow in and out, altering the overall charge of the neuron. As there are different types of ions, various ion channels exist that will only allow passage to a certain type of ion molecule, as originally observed and modelled by Hodgkin & Huxley in 1952.

It is also important for the cell to maintain differences in the concentration of ions inside and outside of the cell (see Dayan & Abbot, 2001, for overview). This is performed by ion pumps within the membrane that expend energy to maintain equilibrium (Figure 1D). Once the overall charge within the cell body of the neuron is raised sufficiently, an action potential is generated. This is effectively a wave of electrical transmission through the cell body and axonal length (Figure 1A; nerve impulse), as channels and pumps push ions through the cell to retain equilibrium. Charged ions at pre-synaptic axonal terminals can then influence the electrical charge of neighbouring post-synaptic cells through any axonal-dendritic connections (Figure 1B), enabling signal transmission between connected neurons (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000).

The overall charge of a neuron is determined by the membrane potential, the voltage across the cell membrane (Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). Realistically, the charge varies across the membrane (Shepherd, 1994), but typically mathematical models assume an integrative value for the potential, termed a point-neuron equation (Abeles, 1991). Similarly, the neuronal body and axonal length are made up of a varying number of compartments (Shepherd, 1994),

dependent on the neuron type. There are many multi-compartment mathematical representations of axonal lengths (Dayan & Abbot, 2001), but this thesis focuses on single-compartment models, which assume a holistic integration of internal ionic changes via a single term, the membrane potential (V_m) (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000).

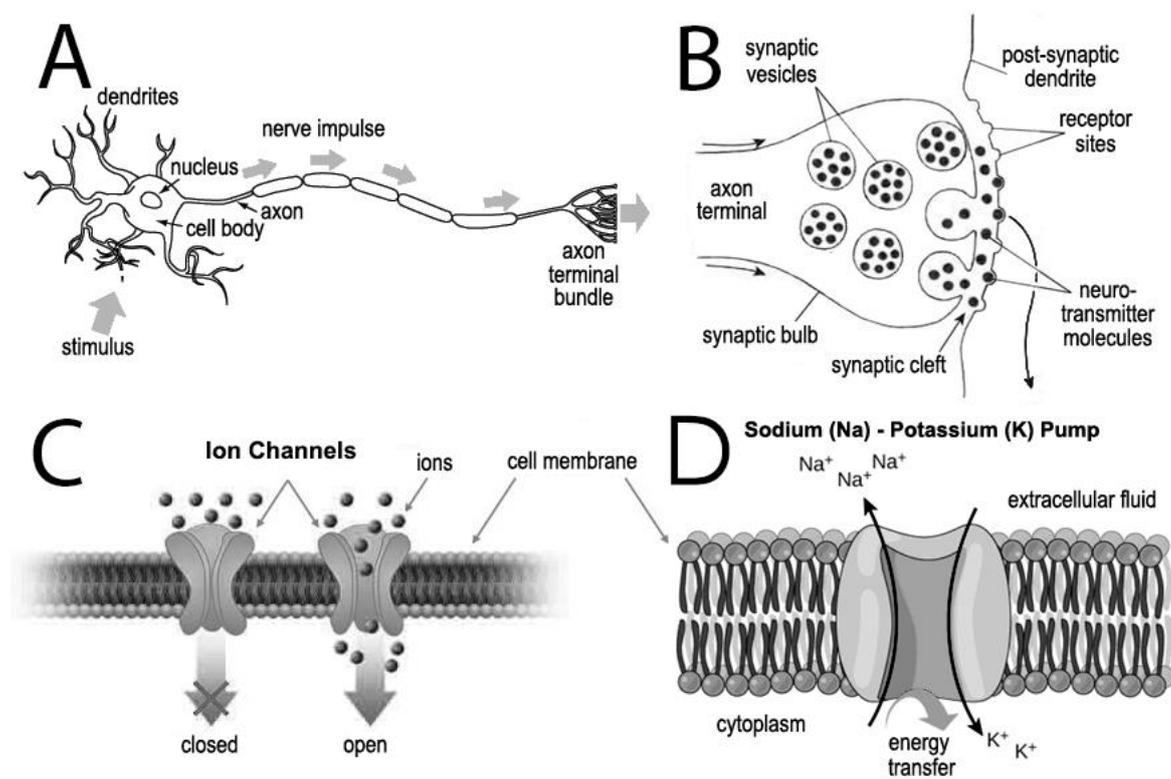


Figure 1 – **A** Structure of a neuron and direction of signal transmission. **B** Axonal-dendritic synaptic cleft that allows transmission of charged ions from one cell to another. **C** Channels in the cell membrane allow ions to traverse and influence the overall charge of the neuronal cell body. **D** Sodium-potassium pump enables diffusion by expending energy to maintain equilibrium. (Jarosz, 2009; OpenStax College, 2013).

2.1.1 MEMBANE CURRENT

The membrane current is simply the total flow of ions across the cell membrane, where a neuron has thousands of channels of various ion types operating to alter the electrical charge of the neuron in a hyperpolarising (negative) or depolarising (positive) way (Shepherd, 1994). These ion channels

enable molecules to move in line with concentration gradients to maintain an equilibrium within the cell. Ion pumps (Figure 1D) also contribute to the movement of molecules, though in contrast to ion channels, they work to actively transport ions against concentration gradients via an energy transfer. This process transports 3 sodium ions (Na^+) out of the cell and 2 potassium ions (K^+) into the cell, causing an unbalanced charge transfer that lowers the overall charge of the neuron (Dayan & Abbot, 2001). We consider the functional roles that these contributions to the membrane current provide in the following section. Throughout the following sections, we present equations in a unit-less fashion. We therefore introduce the units used throughout this thesis in Table 1, below.

Symbol	Magnitude	Unit	Unit symbol
V_x	Voltage potential	millivolts	<i>mv</i>
E_x	Reversal/resting potential	millivolts	<i>mv</i>
\bar{g}_x	Conductance	siemens	<i>S</i>
I_x	Current	ampere	<i>A</i>
m, n, h	Gating variables	dimensionless	0 – 1
C_x	Capacitance	farad	<i>F</i>
R_x	Resistance	ohm	Ω
τ_x	Time constant	milliseconds	<i>ms</i>

Table 1 – Introduction of units used for subsequent equations in this thesis.

2.1.1.1 LEAK CURRENT

Time-independent contributions to the membrane current, such as ion pumps, are typically lumped together in a leak current (Equation 1; see Dayan & Abbot, 2001, for overview). As these mechanisms operate to maintain equilibrium in the neuron, they occur at a rate relative to the reversal potential of the membrane (E_L). This is typically an approximation to match the membrane potential of the observed neuron at rest, though more detailed neurophysiological models allow the value of E_L to fluctuate dependent on the internal state of the neuron. Figure 2 shows how the relative difference between the membrane potential and the reversal potential drives the net flow of ions over a current, where an equilibrium exists when both potentials are equal.

$$I_L = \bar{g}_L(V_m - E_L)$$

Equation 1 – The leak current.

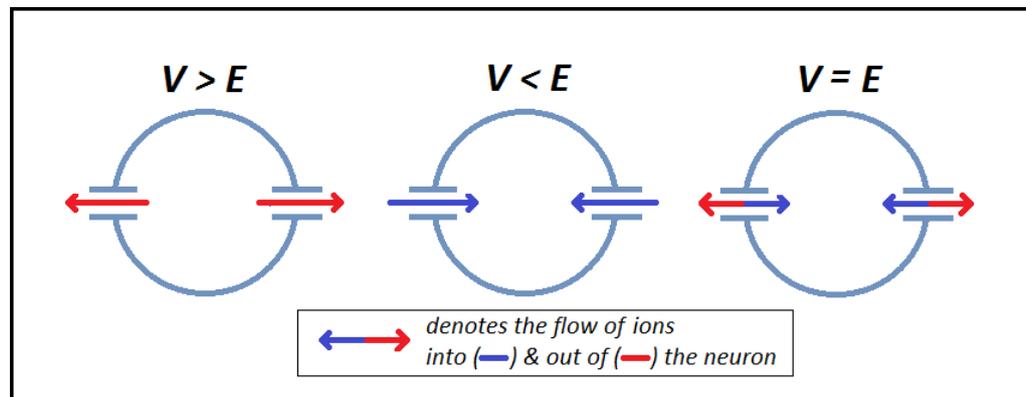


Figure 2 - The net flow of ions over a specific current depends on the relative difference between the membrane potential (V) and the reversal potential (E) of the current.

In the leak current, this driving force is derived from the difference between the membrane potential (V_m) and the leak reversal potential (E_L). A constant \bar{g}_L is used here to model the conductance of the leak current, acting as a multiplier to the magnitude of the driving force. This is typically an approximation to match the conductance of the membrane at rest, assuming that contributions work at a near constant rate and do not fluctuate much (Dayan & Abbot, 2001). Like

the leak reversal potential, more complex neurophysiological models would enable \bar{g}_L to fluctuate dependent on the state of the membrane. The leak current (Equation 1) thus drives the membrane potential to fall to the resting potential of the neuron, pushing the model neuron to dynamically maintain equilibrium due to the continual flow of ions traversing the cell membrane.

2.1.1.2 SODIUM CURRENT

We can also ascertain reversal potential values for channel currents of different ion types, where the value of E_i lies within different ranges dependent on the type of ion, where i denotes the channel type. Potassium (K^+) reversal potentials typically lie between -70 and -90mV, whilst sodium (Na^+) and calcium (Ca^{2+}) reversal potentials are above 50 and 150mV, respectively (Dayan & Abbot, 2001). This shows how ion channels can be highly selective, as when the membrane potential (V_m) changes value to greater than or less than E_i , the flow of ions in that channel reverses as it seeks equilibrium. Therefore, selective ion channels have a hyperpolarising or depolarising effect dependent on the voltage of their resting potential (Hodgkin & Huxley, 1952). As with the leak current, here we approximate over many thousands of channels to calculate a single term that estimates channel behaviour for each ion type (Dayan & Abbot, 2001).

Sodium channels provide a fast reacting, depolarizing change to a neuron's overall charge due to their high conductance ($\bar{g}_{Na} = \sim 100$) and reversal potential ($E_{Na} = \sim 50$) (Hodgkin & Huxley, 1952; Shepherd, 1994). In Equation 2, the voltage-dependent gating variable, m , ties the total conductance to the membrane potential in a positive feedback loop, such that conductance increases with membrane depolarization and vice versa. This initiates a large response to slight increases in the membrane potential, a key component of action potential generation (Dayan & Abbot, 2001). When the membrane potential reaches a threshold, the sodium current switches to an inactivation state, whereby the voltage-dependent gating variable, h , quickly diminishes. This dual behaviour can be described by the equation for the total sodium current below, that were derived from the famous experiments on a squid axon by Hodgkin and Huxley in 1952. Here, the

driving force is given by the difference between the membrane voltage (V_m) and the resting potential of sodium channels (E_{Na}), where a constant \bar{g}_{Na} approximates the conductance of open sodium channels and total sodium conductance fluctuates with voltage-dependent gating variables m and h .

$$I_{Na} = \bar{g}_{Na} m^3 h (V_m - E_{Na})$$

Equation 2 – The sodium (Na) current.

2.1.1.3 POTASSIUM CURRENT

Potassium channels work in tandem with the sodium current in producing membrane action potentials (Hodgkin & Huxley, 1952), slowly opening due to a lower conductance ($\bar{g}_K = \sim 80$) and the sharp sodium induced membrane depolarisation. Their function is to restore the membrane potential to its resting level and maintain equilibrium in the nerve cell (Shepherd, 1994), otherwise the nerve cell would be prone to persistent and uncontrollable activity. Due to the much lower reversal potential ($E_K = \sim -100$), this current increases in magnitude when the membrane potential becomes more depolarised. Equation 3 captures this behaviour, where the voltage-dependent gating variable n controls the total conductance.

$$I_K = \bar{g}_K n^4 (V_m - E_K)$$

Equation 3 – The potassium (K) current.

2.1.1.4 CHANNEL CONDUCTANCE

Voltage-dependent gates dictate the openness of each channel, whereby change in the gates m , n and h over time influence the total conductance of sodium and potassium currents (Hodgkin & Huxley, 1952). The equations below show how these are modelled, where each gate opens at a rate of $\alpha_x(V)$ and closes at a rate of $\beta_x(V)$, where x indicates the current value of the gates m , n or h .

Equation 4 shows how the total change over time in the gate is calculated as the sum of both of these processes. In Equation 5, $x_{\infty}(V)$ indicates the steady-state activation curve of the gate, which is the value at which the gate will get asymptotically close to if the voltage is held at a constant value. In Equation 6, $\tau_x(V)$ is the time constant of the gate that varies with the voltage and dictates how fast the gate will open or close.

$$\frac{dx}{dt} = \alpha_x(V)(1 - x) - \beta_x(V)x$$

$$x_{\infty}(V) = \alpha_x(V)/(\alpha_x(V) + \beta_x(V))$$

$$\tau_x(V) = 1/(\alpha_x(V) + \beta_x(V))$$

Equations 4, 5 & 6 – Voltage-dependent (V) gate update, steady-state & time constant equations, respectively, where x denotes gate type.

$$\alpha_m(V) = 0.32(V + 54) / (1 - e^{-(V+54)/4})$$

$$\alpha_n(V) = 0.032(V + 52) / (1 - e^{-(V+52)/5})$$

$$\alpha_h(V) = 0.128e^{-(V+50)/18}$$

Equations 7, 8 & 9 – Voltage dependent opening of gates m , n & h , respectively.

The opening and closing dynamics of each gate type are shown in the next set of equations (Hodgkin & Huxley, 1952), where fixed constants determine the rate and reversal potentials. More recent parameter estimation has found that there are various functional parameter sets (Willms, et al., 1999), one of which is shown here. Equations 7-9 show the rate of opening in gates m , n & h , respectively, whereas Equations 10-12 show the rate of closing in gates m , n & h , respectively.

$$\beta_m(V) = 0.28(V + 27) / e^{(V+27)/5}$$

$$\beta_n(V) = 0.5e^{-(V+57)/40}$$

$$\beta_h(V) = 4 / (1 + e^{-(V+27)/5})$$

Equations 10, 11 & 12 – Voltage dependent closing of gates m , n & h , respectively.

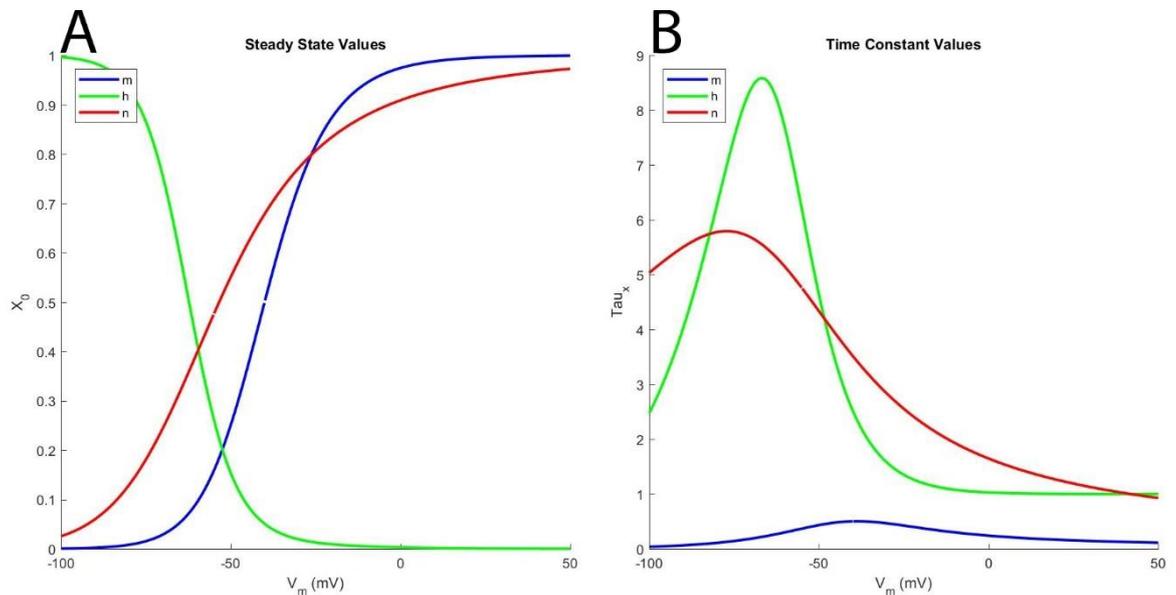


Figure 3 – Voltage clamping simulations show sodium activation (m , blue) and inactivation (h , green) curves and potassium activation (n , red) curve, where the steady-state values (**A**, Equation 5) and time-constant values (**B**, Equation 6) of each gate varies dependent on the voltage.

Figure 3 shows the steady-state and time constant activation curves in a voltage-clamping simulation for each of the gates m (blue), n (red) and h (green) from the Hodgkin & Huxley model, 1952. In terms of the sodium response, one can see that as the voltage increases, the m gate opens as it moves towards a higher steady-state value, initiating a positive feedback loop with the membrane potential. The rate at which a gate opens can also be solved as $dx/dt = (x_\infty(V) - x)/\tau_x(V)$. This shows us how the sodium activation current (m) occurs very quickly after -50mV to initiate the burst like response typical of action potentials (Dayan & Abbot, 2001), opening fully and at a fast rate as m_∞ and τ_m approach 1, respectively. The sodium inactivation current (h) then takes hold to diminish total sodium conductance, where h_∞ falls to 0 at an increasingly fast rate (τ_h). Total sodium conductance is therefore maximal at ~ -50 mV, in between these processes. The potassium activation current (n) then comes in to stabilise the action potential process, which

would be prone to falling into an oscillation due to the sharp swings in sodium activation and inactivation currents. The n gate opens more slowly during depolarisation, even remaining slightly open when the voltage becomes repolarised ($V < -50\text{mV}$) to soften the blow of rapid voltage changes. This helps to enable a more stable transition during fast, sodium induced depolarisations that occur during action potentials (Hodgkin & Huxley, 1952; Shepherd, 1994).

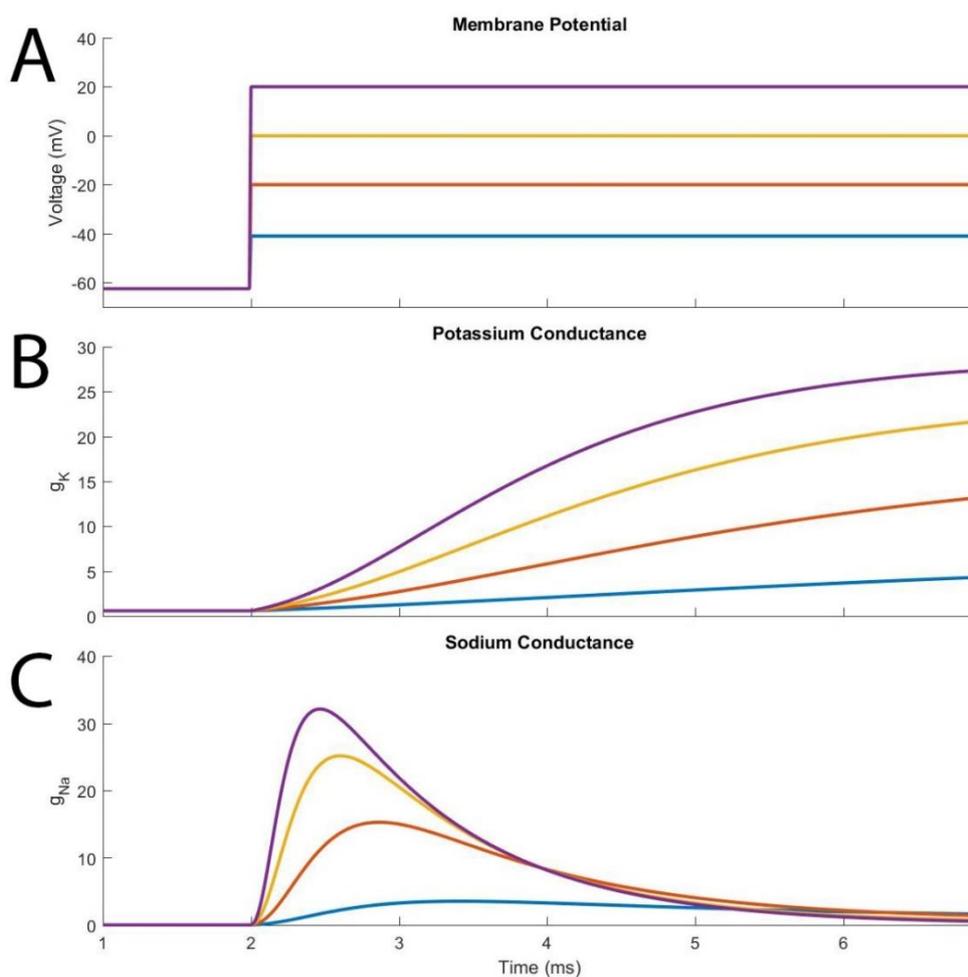


Figure 4 – Voltage clamping simulations showing changes in potassium (B) and sodium (C) conductance as the membrane potential is clamped at increasing values (A).

One can see the variability in total channel conductance in voltage clamping simulations, where the membrane potential is clamped to a range of values (Hodgkin & Huxley, 1952). Figure 4A indicates increasing values of clamped voltage, whilst Figures 4B & 4C shows the colour co-ordinated potassium and sodium conductance changes, respectively. Here, it is much easier to see the short

spike in sodium current as the activation and inactivation currents work against one another (Figure 4C), whereas the stabilising potassium current gradually increases over time (Figure 4B) to re-assert equilibrium after the spike event has occurred (Shepherd, 1994).

2.1.2 HODGKIN & HUXLEY MODEL

The previously described channel properties were first modelled by Hodgkin and Huxley in their famous squid axon experiments in 1952, seen in Equation 13. Using empirical evidence, they utilised channel and conductance equations to model the change in voltage of the membrane potential of a neuron. Here, the change in voltage (dV/dt) is equal to the negative of the summation of the aforementioned sodium, potassium and leak currents, as well as input from the dendritic connections of other neurons (I_{syn}) and any externally applied direct current (I_{DC}), divided by the specific capacitance of the neuron (c_m).

$$c_m \frac{dV}{dt} = -(I_{Na} + I_K + I_L + I_{syn} + I_{DC})$$

Equation 13 – The Hodgkin & Huxley model, 1952.

Constants for the conductance of each current dictate the rate of change in those channels ($\bar{g}_{Na} = 100, \bar{g}_K = 80, \bar{g}_L = 0.1$), whilst the resting potential drives voltage-dependent change in each current ($E_{Na} = 50, E_K = -100, E_L = -67$). The specific membrane capacitance is the capacitance of the model neuron per unit area (Dayan & Abbot, 2001), which must be overcome by any external charge ($c_m = 1$). Action potentials are self-generated by the sodium (I_{Na}) and potassium (I_K) currents, whilst the leak current (I_L) drives the membrane potential to seek equilibrium at the neuron's resting potential (Hodgkin & Huxley, 1952). Due to the biophysical complexity of this equation, a time-step update of less than 0.1ms must be assured to mitigate wild fluctuations in gate and channel update equations. A spike-detecting algorithm must also be implemented due to the dynamic spike-generation properties of current equations.

2.1.3 INTEGRATE & FIRE MODEL

The integrate-and-fire model is a commonly used, simplified model of a biological neuron (Abeles, 1991; Dayan & Abbot, 2001). Here, a single-compartment model is used that describes electrotonically compact neurons, i.e. uni-compartmental cells where a single term can describe the whole membrane potential (V_m). Realistically, biological pyramidal cells contain several interacting compartments that dictate internal current flow (Shepherd, 1994). Unlike the Hodgkin & Huxley model, this model only uses the leak current to alter the internal state of the neuron. As the sodium and potassium currents dynamically model an action potential, the integrate-and-fire model thus requires additional rules to detect and generate spike events.

$$\frac{dV_m}{dt} = \frac{E_L - V_m}{\tau_m} + \frac{I_{syn} + I_{DC}}{C_m}$$

Equation 14 – The integrate-and-fire model.

Equation 14 shows the integrate-and-fire model, where as a model neuron can be thought of as an electrical circuit (see Dayan & Abbot, 2001, for overview), one can simplify from the Hodgkin & Huxley model using equations from the resistor-capacitor (RC) circuit. As such, when considering the leak current in Equation 1, the membrane conductance (\bar{g}) is equal to $1/r_m$, the specific membrane resistance or resistance per unit area. Like the specific membrane capacitance (c_m), the specific membrane resistance (r_m) can be divided by the total neuronal area (A , typically set to a value of 1 in single compartment models) to obtain the total membrane resistance (R_m). The membrane time constant (τ_m) is then given as the product of R_m & C_m . This is summed up in Equation 14, where voltage change through time is driven by the difference between the leak reversal potential (E_L) and the membrane potential (V_m), divided by the membrane time constant (τ_m), which essentially dictates the rate at which the model neuron will decay back to the resting potential. Any current injected into the model neuron (I_{DC}) and current from dendritic connections (I_{syn}) are modulated by the total membrane capacitance of the neuron (C_m), whilst the membrane

potential exponentially decays to its resting potential in the absence of input due to the leak current (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000).

There are also several assumptions made within the framework of this model. In the absence of an explicit description of action potential generating currents, when the membrane potential reaches a threshold (-55mV) an action potential is assumed to occur (Dayan & Abbot, 2001). Once this happens, the membrane potential is reset to the membrane resting potential (-70mV), which assumes the immediate voltage decrease that is otherwise due to the inactivation of the sodium currents. Where previously the model neuron would dynamically repolarise after an action potential and inhibit incoming input in the process (Hodgkin & Huxley, 1952), now we must enforce a refractory period onto our model to capture a similar process. This is typically done in one of two ways (Dayan & Abbot, 2001), where the membrane potential is absolutely or relatively clamped to the resting potential for a specific period (typically $\sim 2\text{ms}$). In the former, no inputs may affect the model neuron for the determined period, whilst in the latter an exponential function incrementally allows input back into the neuron, more closely following the dynamic action-potential behaviour of a biological neuron (Shepherd, 1994).

The refractory period also ensures that the neuron cannot immediately spike after just having done so, mimicking the length of time it takes for sodium and potassium currents to come back online in biological neurons (Hodgkin & Huxley, 1952). Due to the simplicity of the equations, one can simulate the integrate-and-fire model with 1ms time steps, vastly increasing computational capacity. Due to direct involvement in the generation of action potentials, one also has all spike-detection data readily available, further reducing on-line computational cost.

2.1.4 RATE-CODED MODELLING

One can abstract even further when considering the simulation of networks of neurons. In rate-coded modelling, nodes can represent anything from an individual cell to hundreds of neurons, where the node outputs the summation of activation through time (see O'Reilly & Munakata, 2000, for overview). The aim here then is to capture emergent perceptual phenomena without incorporating all the underlying biophysical details, as the many layered model in Figure 5 shows (Aisa, et al., 2008), where several interacting neuronal layers forward activation through the network to perform a cognitive function. Such models, often termed connectionist, are useful for offering other interpretations of observable data or theorising on the underlying functional connectivity that might give rise to cognitive processes. There is also a massive reduction in computational complexity, leading to reduced simulation times and the ability to build larger and more interconnected neuronal architectures. However, as they make many biophysical assumptions, they cannot be relied on to accurately foresee in-depth neural processes, for example modelling the effect of discrete neural spiking in relation to the activity of the population (Dayan & Abbot, 2001; O'Reilly & Munakata, 2000).

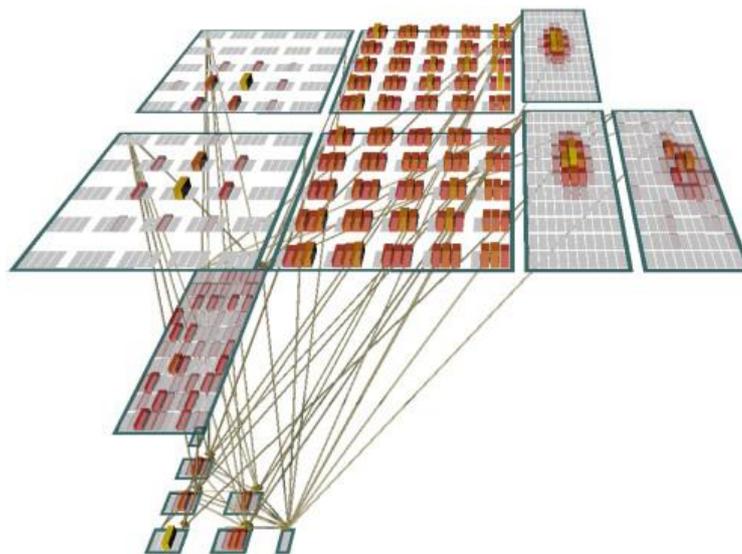


Figure 5 – Rate coded model of many interacting layers. Adapted from Aisa, et al., 2008.

2.2 NETWORKS OF NEURONS

The previous section focused on the neurophysiology of single neurons (Shepherd, 1994) and the various mathematical abstractions that are commonly used to model them (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). Next, we will consider how neurons connect and interact with one another, an important facet of computational neural network modelling.

2.2.1 POST-SYNAPTIC-POTENTIAL

As discussed previously, neurons are connected through axonal-dendritic connections (Figure 1B), where neurotransmitter molecules traverse the synaptic cleft that separates the pre-synaptic axon terminal and the post-synaptic dendrite (see Shepherd, 1994, for overview). Therefore, a pre-synaptic action potential effects a change in the membrane potential of other connected neurons over time. Such a process is typically termed a post-synaptic potential (PSP), where a positive change is termed an excitatory post-synaptic potential (EPSP) and a negative change an inhibitory post-synaptic potential (IPSP), depending on the charge of transmitted molecules. Whilst there are many neurophysiological models of both the electrical gap junctions and chemical synapses that facilitate PSPs (Dayan & Abbot, 2001; Wang, 2010), this work only considers simple mathematical solutions to chemical synapses.

There are several commonly used methods for modelling how a post-synaptic neuron reacts to spike-events from pre-synaptic neurons (Abeles, 1991; Dayan & Abbot, 2001). The first and simplest method is to add the strength of the weighted spike immediately onto the membrane potential of the post-synaptic neuron, where a spike equates to a binary event multiplied by a weight. A weight here refers to calcium dynamics within the neuron, acting as a modulator that facilitates the transmission of electrical signals and forms the basis of learning between neurons (Hebb, 1949). Another method attempts to model the behaviour of chemical molecular transmission at the synaptic cleft during an action potential (Dayan & Abbot, 2001), as seen in Equation 15.

The Alpha function of Equation 15 creates a dual exponential curve that rises sharply until t reaches the synaptic time constant (τ_s), before exponentially falling back to zero. This is termed the post-synaptic current (PSC), and is added to the incoming current of post-synaptic neurons to cause a PSP. Each spike event is multiplied by the weight of the connection between the two neurons (w). The value of τ_s determines how long the post-synaptic dendrites will allow neurotransmitters to influence the post-synaptic membrane potential (Dayan & Abbot, 2001). Therefore, the area under the PSP curve, exemplified in Figure 6A, is the total voltage change added to the post-synaptic input current over time (I_{syn}). Figure 6B also plots the membrane potential of an integrate-and-fire neuron ($C_m = 10, \tau_m = 50$) influenced by a pre-synaptic action potential, using several values for τ_s . Here, one can see that a larger τ_s leads to larger voltage change due to a more sustained input. The leak current of the model neuron is also fighting against any input as it pulls the membrane potential back to the resting potential, successfully doing so when the sustained PSP subsides.

$$w \cdot \left(e \cdot \frac{t}{\tau_s} \right) \cdot e^{-\frac{t}{\tau_s}}$$

Equation 15 – Exponential Alpha function for a post-synaptic potential (PSP).

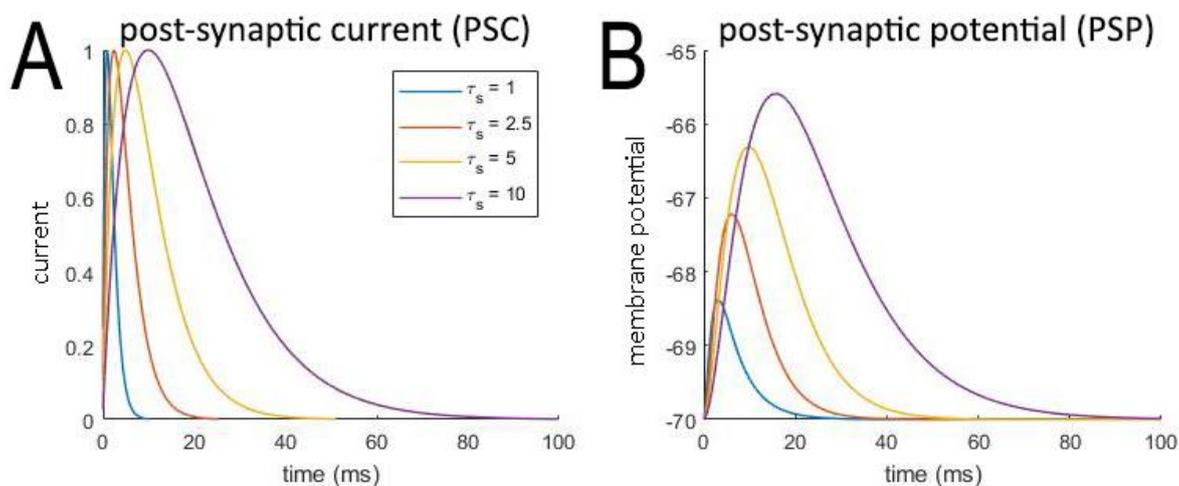


Figure 6 – **A** Post-synaptic current (PSC) for various synaptic time constants (τ_s). **B** The post-synaptic potential (PSP) caused by pre-synaptic action potentials for various τ_s .

2.2.2 BACKGROUND NOISE

Typically, research in computational neuroscience is focussed on isolating and recreating specific functional networks in the brain (O'Reilly & Munakata, 2000). Though it is a long-term goal, there is currently limited capacity to simulate the brain as a whole. As such, the interconnectedness of the brain must be appreciated in neural network models, where each of the 100 billion neurons in the human brain has on average 7000 synaptic connections (Shepherd, 1994). Therefore, there is need to account for any number of discrete spike events originating from external sources. This is typically modelled by using a Poisson distribution, where a known number of independent events occur within a known time frame (Abeles, 1991; Dayan & Abbot, 2001). This mechanism is commonly used to introduce background noise, often to keep model neurons in the ready-to-fire state that is typically observed in biological neurons (Dayan & Abbot, 2001). Functionally, this allows the network to quickly react to incoming inputs, as the membrane potential is kept just below the threshold for a neuron to fire.

Throughout the modelling work of this thesis, noise will be generated using an approximation of a Poisson distribution of spike events. Figure 7 shows the mathematical implementation, where a string of uniformly distributed random numbers is first initialised, beginning and ending in a pre-defined period where every millisecond has a corresponding random number between 0 and 1. All milliseconds where the random number is above the rate of firing required in kHz are set to 0, whilst those below are set to 1. Thus, a sequence of binary spike events is expressed, where 1 indicates a spike event. To enable multiple incoming spike-events per millisecond, multiple trains of sequences are generated and subsequently summed together. This latter step is important when generating distributions with a high firing rate, as well as to reflect the multitude of independent neuronal sources for spike events. For a true Poisson distribution, a similar probabilistic approach is calculated for each time-point independently. The approximation used here, enabled by the uniformly distributed set of random numbers, is a reduction in complexity that enables simulations to be instantiated and simulated at a faster pace.

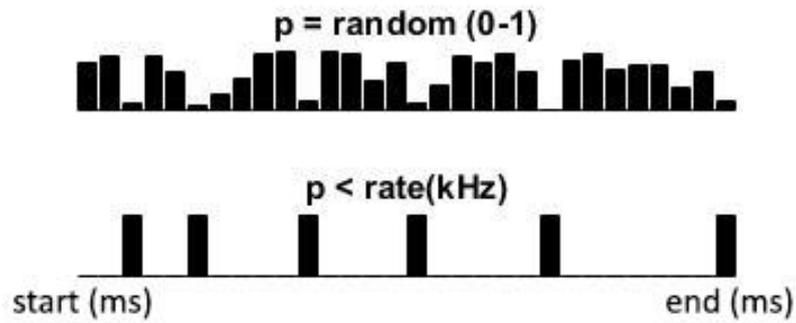


Figure 7 - Creating an approximation of a Poisson distribution by initialising a string (start to end in milliseconds) of uniform random numbers between 0-1. Any number below or above the desired rate (kHz) is then set to 1 or 0, respectively, generating a sequence of independent spike events.

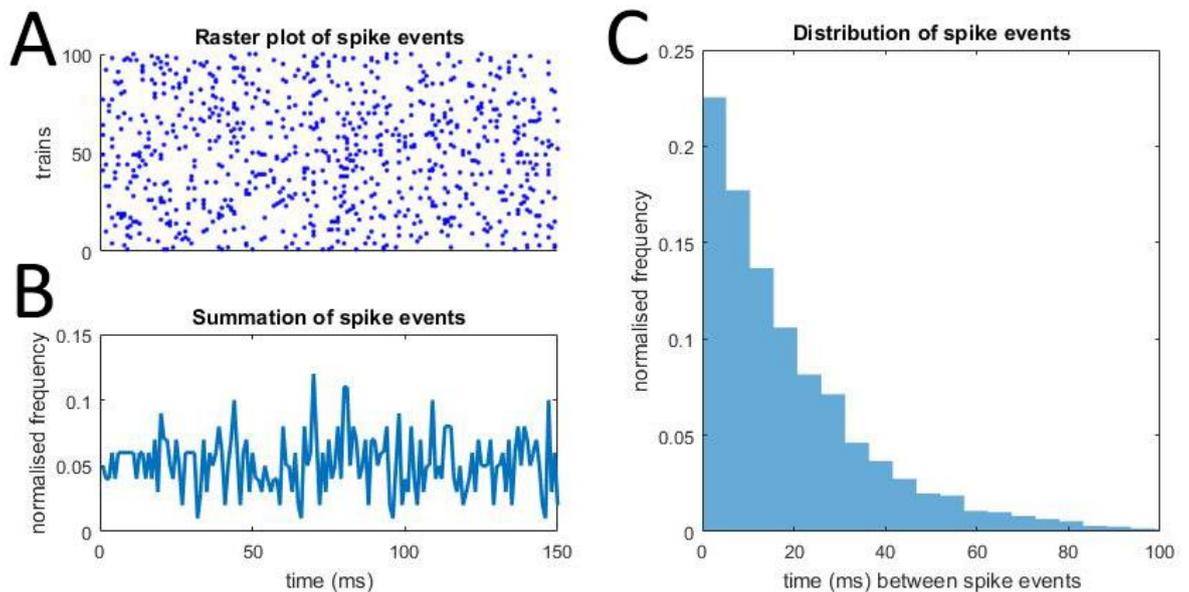


Figure 8 – Several trains of Poisson distributed spike events that are typically used to generate background noise. **A** Raster plot of many generated trains, **B** signal generation for summation of all spike events and **C** a histogram of the time between spike events.

EEG & MEG studies reveal that neurons fire precisely together in groups (Fell & Axmacher, 2011), therefore it is reasonable to assume that the background noise a neuron receives is bursty in nature (Dayan & Abbot, 2001). From Figure 8, one can see how a Poisson distribution is ideal for this consideration. Here, several bursts are apparent in the raster plot of many generated trains of spike events (Figure 8A), which are also visualised as erratic bursts in subsequent signal generation

(Figure 8B). However, a histogram of the time between each consecutive spike-event exemplifies this bursty behaviour best (Figure 8C), where it is probable that an incoming spike-event will be generated soon after one has already been received. This is done independently for each event in a probabilistic manner, as a Poisson distribution is memoryless (Dayan & Abbot, 2001).

2.2.3 SPIKE-TIME-DEPENDENT-PLASTICITY (STDP)

Learning is enabled in the brain through changes in the synaptic conductance between neurons (Shepherd, 1994). This was classically found in Hebb's work (1949) looking at the effect of activity on neuronal circuits, finding that:

"When an axon of cell A excites cell B and repeatedly or persistently takes part in firing with it, some growth process or metabolic change takes place in one or both cells so that A's efficiency as one of the cells firing B is increased."

This phrase neatly sums up Hebb's rule (Hebb, 1949), simply put that if neurons fire together, they also wire together. This is the essence of Hebbian learning, a subsequent modelling method where the connections between neurons or rate-coded nodes increase if they are concurrently activated at the same moment in time. Whilst this method echoes the principles of neuronal learning and is very useful for rate-coded models (O'Reilly & Munakata, 2000), it simplifies the biological processes that are thought to enable competitive learning in the brain (Graupner & Brunel, 2012; Song, et al., 2000). The following section details two other methods that more precisely model how neurons compete for control of shared post-synaptic nodes.

2.2.3.1 COMPETITIVE HEBBIAN LEARNING

One of the typical aspects of biological networks is that neurons have been observed to contest for control of shared post-synaptic neurons (Bi & Poo, 1998; Markram & Sakmann, 1995; Markram, et al., 1997). This enables the strongest pathways to activate faster upon successive iterations (Hebb, 1949), see Figure 9, successfully associating incoming stimuli from the environment with a library

of competing assemblies. To enable this competitive environment, a spike-time-dependent-plasticity (STDP) concept was proposed in a theoretical framework that modelled the changes in synaptic efficacies due to activation of pre and post-synaptic neurons (Song, et al., 2000), as Hebb had observed (Hebb, 1949). The STDP model we examine here, initially theorised elsewhere (Gerstner, et al., 1993; Gerstner, et al., 1996), rewards synapses that must have contributed to a post-synaptic spike whilst punishing synapses that are deemed to be in competition to the post-synaptic spike. A reward comes in the form of long-term-potential (LTP), an increase in the synaptic efficacy between neurons that can last for months (Shepherd, 1994), whilst punishment entails long-term-depression (LTD) of the synapse, a decrease in the neuronal weight. Efficacy changes are likely due to changes in calcium amplitudes within the neuron, which facilitate information transmission between neurons in response to a pre-synaptic action potential (Shepherd, 1994).

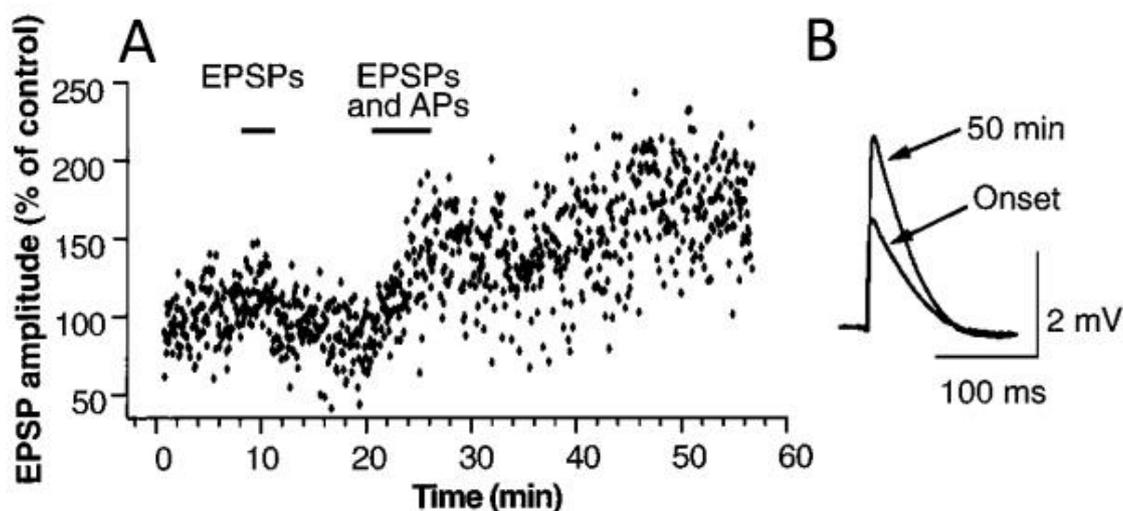


Figure 9 – Observations of the amplitude of excitatory-post-synaptic-potentials (EPSPs) of observed neurons over time (Markram, et al., 1997). Amplitude increases after pre-synaptic action potentials (APs) are induced (>20 mins), indicating synaptic modification through STDP.

This is neatly summed up in Equation 16 (Song, et al., 2000), where the difference in time (Δt) is the timing of the first spike minus the timing of the second spike. LTP is induced if $\Delta t < 0$ (i.e. if

the pre-synaptic neuron spiked before the post-synaptic neuron) and is calculated as the amount A_+ multiplied by the exponentially decayed time differential. Similarly, LTD is induced if $\Delta t > 0$ with an amount A_- , where $A_-/A_+ = 1.05$ to promote competition and maintain network stability (Song, et al., 2000). This dual behaviour is summarised in Figure 10, where $F(\%)$ indicates the fraction of change in weights due to Equation 16. Weights are additionally bound between 0 and a maximum amount, in order to minimise the impact of rampant weight change.

$$F(\Delta t) = \begin{cases} A_+ \exp(\Delta t/\tau_+) & \text{if } \Delta t < 0 \\ -A_- \exp(-\Delta t/\tau_-) & \text{if } \Delta t > 0 \end{cases}$$

Equation 16 – Function for competitive STDP (Song, et al., 2000).

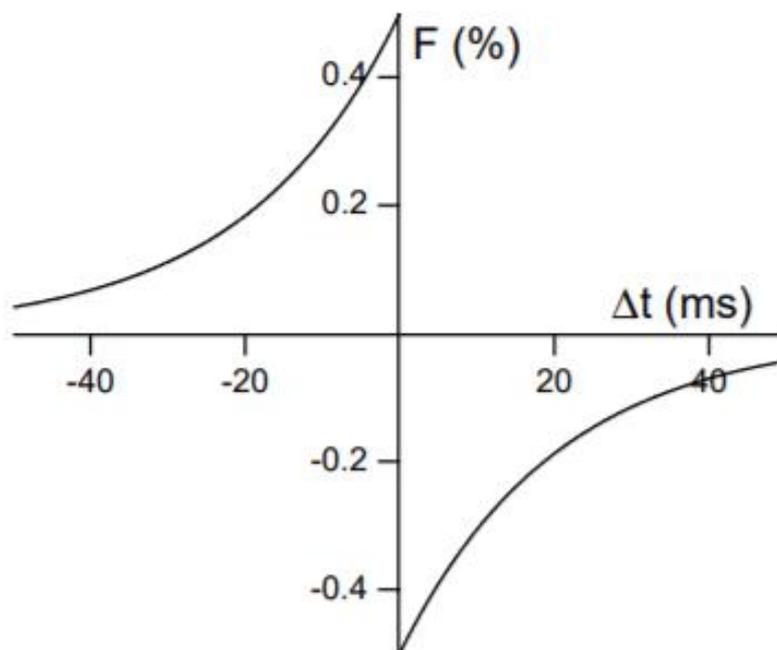


Figure 10 –STDP modification function, where the difference in spike pair timing determines the fraction of change in LTP or LTD of the synapse. Using this learning rule, synapses changed by up to 40% after a single activation of a spike-pair. Figure re-represented from Song, et al., 2000.

2.2.3.2 CALCIUM BASED PLASTICITY MODEL (GRAUPNER AND BRUNEL, 2012)

Whereas the previously defined STDP function captures the behaviour of a synapse in relation to pre and post-synaptic activation (Song, et al., 2000), there have been many different types of STDP curves observed experimentally in different neuron types and brain regions (Graupner & Brunel, 2012). A recent modelling study has adapted this STDP method by introducing additional functions that model calcium accumulation at the post-synaptic buton due to pre and post-synaptic spike events (Graupner & Brunel, 2012). This functionality allows the model to capture multiple STDP curves and predicts that differences in plasticity outcomes in different studies can be explained by a variation of the parameter space that define calcium dynamics.

Equation 17 summarises synaptic change in the model, where synapses are represented by the variable $\rho(t)$ which exists in a state between 0-1. Synapses gravitate towards being active ($\rho = 1$) or inactive ($\rho = 0$) once they pass a threshold (ρ_*). Change in synaptic efficacy is dependent on whether the amount of calcium $c(t)$ is over specific thresholds for LTP (θ_p) and LTD (θ_d), whereby LTP increases and LTD decreases synaptic efficacy at a rate γ_p or γ_d , respectively (H denotes the Heaviside function, which returns 0 or 1 if the function within [] is below or above 0, respectively).

$$\tau \frac{d\rho}{dt} = -\rho(1 - \rho)(\rho_* - \rho) + \gamma_p(1 - \rho) \cdot H[c(t) - \theta_p] - \gamma_d \rho \cdot H[c(t) - \theta_d]$$

Equation 17 – Synaptic efficacy dynamics in calcium-based STDP (Graupner & Brunel, 2012).

$$\frac{dc}{dt} = -\frac{c}{\tau_{Ca}} + C_{pre} \sum_i \delta(t - t_i - D) + C_{post} \sum_j \delta(t - t_j)$$

Equation 18 – Calcium dynamics in calcium-based STDP (Graupner & Brunel, 2012).

Put simply, the first term describes the synaptic decay that forces synapses to gravitate to fully closed or fully open dependent on the threshold (typically set to 0.5). The second term describes LTP induced changes, whereby the Heaviside function detects whether the calcium concentration

surpasses a threshold and subsequently enables LTP to occur at a predefined rate. Similarly, the final term describes LTD with its own threshold and rate (Graupner & Brunel, 2012).

In Equation 18, the amount of calcium at a synapse can be calculated by the summation of all spike events from pre (i) and post (j) synaptic neurons, where the Dirac delta function $\delta(t - t_i)$ captures the elapsed time since each spike event. Calcium is then linearly summed (after a delay, D , of $\sim 13\text{ms}$ for pre-synaptic spikes to account for the physical distance from the pre to the post-synaptic neuron), multiplied by a constant C_{pre} for pre-synaptic spikes and a constant C_{post} for post-synaptic spikes. Calcium decays exponentially over time ($\tau_{Ca} = 20\text{ms}$), to capture the spike-time-dependent aspect of this plasticity rule (Graupner & Brunel, 2012).

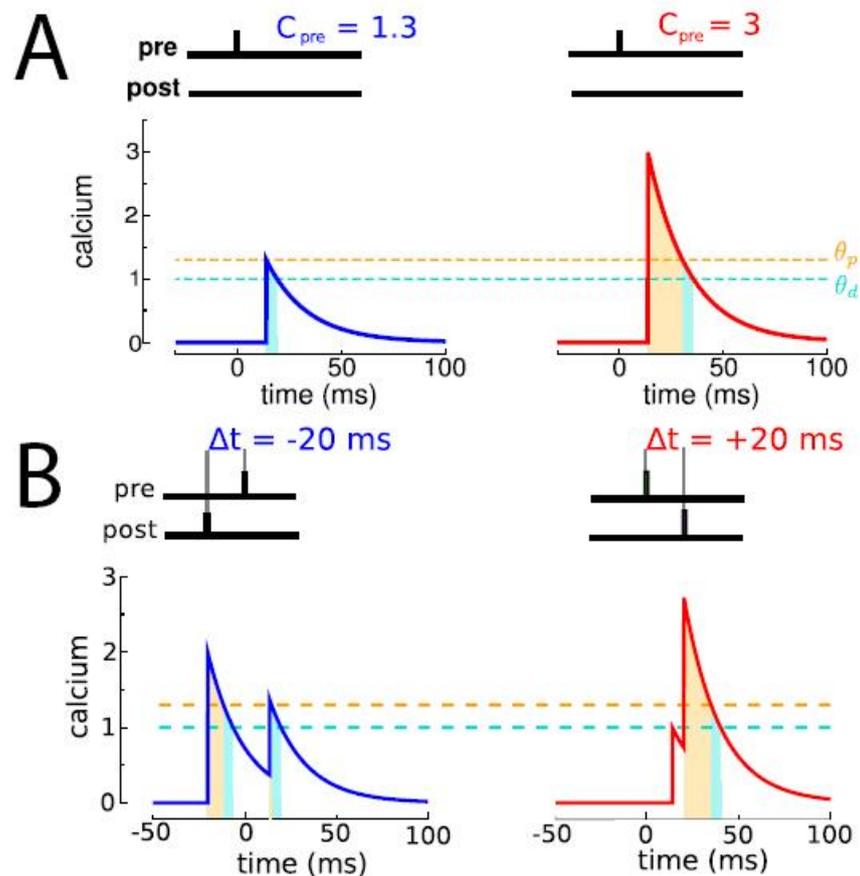


Figure 11 – Calcium based STDP rule. **A** Variation in pre-synaptic calcium increases, where C_{pre} is set to a value of 1.3 or 3. **B** Spike pair timing dictates whether LTP (red shaded) or LTD (blue shaded) occurs. Figure adapted from Graupner & Brunel, 2012.

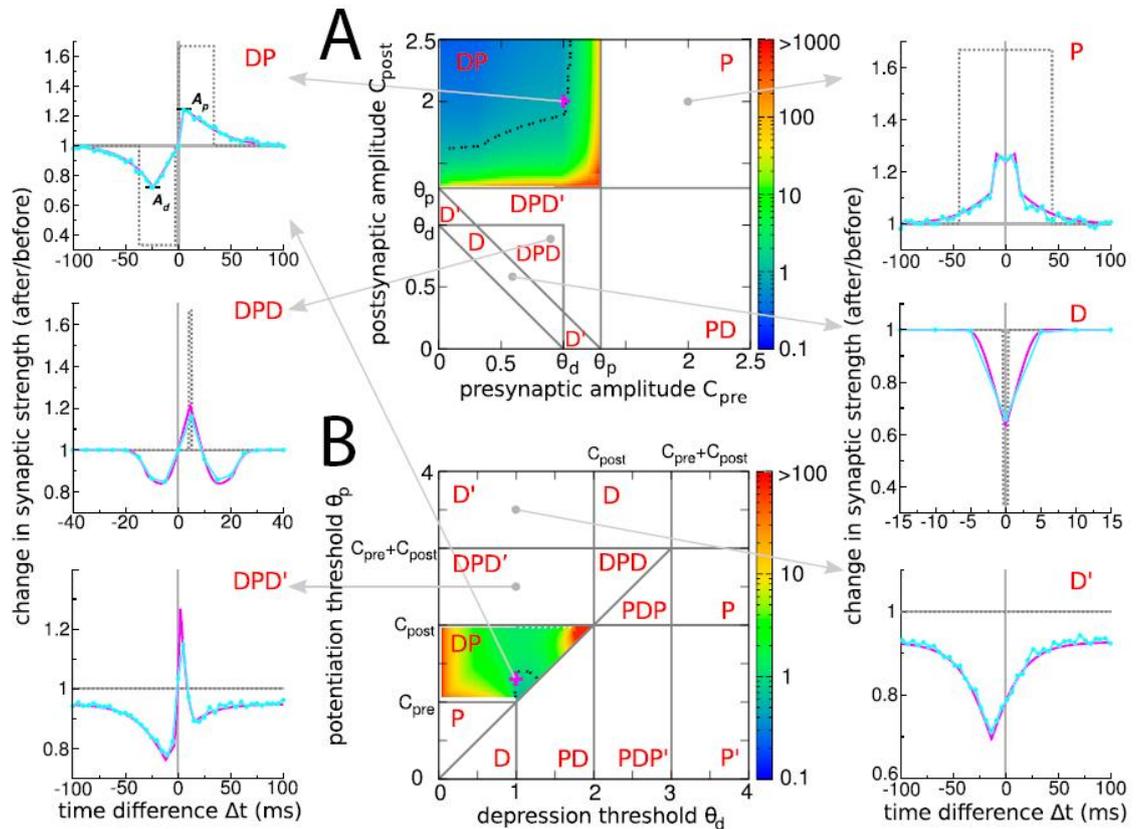


Figure 12 – Multiple STDP curves by varying pre and post-synaptic calcium amplitudes (A) and thresholds for LTP/LTD (B). Grey arrows indicate where each STDP curve lies in each respective parameter space. STDP curves typically have some combination of depression (D) and potentiation (P) of synaptic spike-pairs. Z-axis of parameter space shows total change in synaptic efficacy for a simple spike-pair paradigm. Graph adapted from Graupner & Brunel, 2012. A prime indicates the case where potentiation and depression are unbalanced for large Δt .

Figure 11 shows the calcium function of Equation 18 in action, where in 11A multiple values for the constant C_{pre} are used to show the relation between calcium and LTP/LTD. If the amount of calcium increases above the LTD threshold (θ_d) but not the LTP threshold (θ_p), then LTD will occur at the synapse (blue shaded), whereas if calcium levels surpass both thresholds then LTP is induced (red shaded). As the calcium transient decays, it typically spends some time in between thresholds to cause a short dip in synaptic efficacy that is typical of many experimentally observed STDP curves (Graupner & Brunel, 2012). Figure 11B shows how the timing between pre/post-synaptic spike pairs

dictates whether the synapse will undergo LTP or LTD, where C_{post} is larger than C_{pre} . If the pre-synaptic spike occurs before the post-synaptic spike then LTP is induced, otherwise LTD is induced.

The adaptability of the calcium-based STDP model can be seen in Figure 12, where variation of the parameters for pre and post-synaptic calcium amplitude (C_{pre} & C_{post}) in 13A, as well as the variation of the thresholds for LTP and LTD (θ_p & θ_d) in 13B, generate many types of STDP curve, as seen experimentally (Graupner & Brunel, 2012; Shepherd, 1994). Total change in synaptic strength for a simple spike-pair paradigm is shown on the z-axis, from low (blue) to high (red). When drawing boundaries on the parameter space in Figure 12A, the constant calcium thresholds are drawn on, whereas in Figure 12B, constant pre and post-synaptic calcium amplitudes are drawn on. This enables the parameter space to be split into several STDP regions, some of which are depicted in the peripheral STDP curves. For example, the most commonly observed curve in pyramidal cells in the depression/potential (DP) curve in the top left, where spike pair timing dictates STDP in a similar curve to the Song et al. (2000) model. Alternate curves are shown that match other experimental observations, or that might enable interesting computational learning functions (Graupner & Brunel, 2012). By increasing the complexity of parameter space, Graupner & Brunel have shown that one can also create a more adaptive and powerful model of STDP.

2.2.4 INFORMATION TRANSMISSION

It used to be thought that the brain was too stochastic a system to be able to transmit information through precise spike timing, where it was thought that noisy fluctuations in synaptic input prevented neurons from producing action potentials with high temporal accuracy (Diesmann, et al., 1999). However, once evidence began to show that precisely timed action potentials were key to behavioural events (Abeles, et al., 1993), a model was proposed that investigated the conditions that might give rise to reliable and well-timed information transmission in cortical assemblies (see Diesmann, et al., 1999 for model). Here, groups of consecutively connected cell assemblies forwarded a pulse of activity, initiated at the first layer (Figure 13a, for schematics).

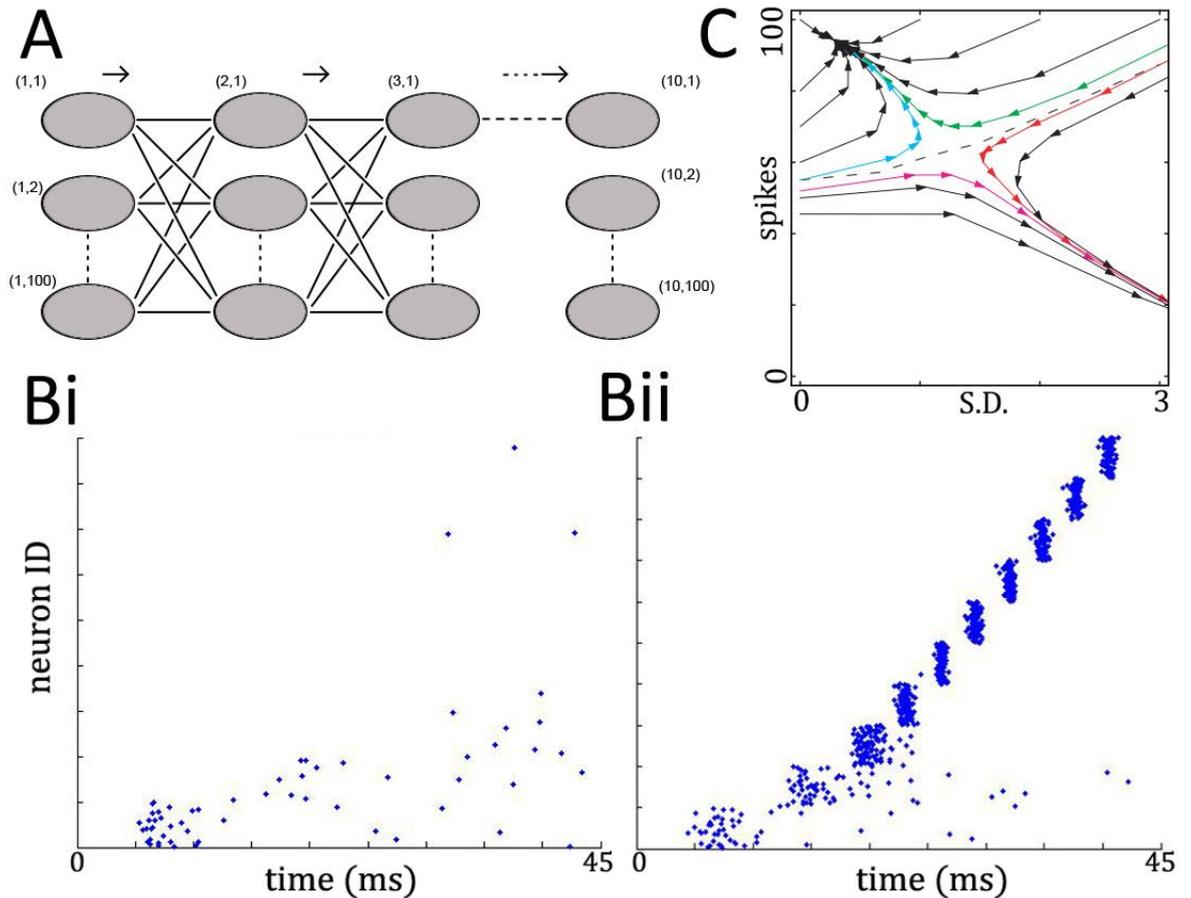


Figure 13 – Stable propagation of synchronous spiking in cortical neural networks. **A** Schematic of consecutively connected neural populations, where 10 groups of 100 neurons are connected with feedforward synaptic connections. When a burst of activation was fed into the first population, activity either desynchronised and did not propagate (**Bi**) or synchronised and propagated (**Bii**), dependent on the number of spike events and synchronisation (i.e. standard deviation, S.D.) of successive bursts (**C**), indicated by arrowheads connected by a directional line. Here, signal propagation either fell into a stable attractor (spikes = ~ 95 , S.D. = ~ 0.5) or dissipated if the trajectory fell below the dashed line. Figure shows recreated findings from Diesmann, et al., 1999.

This model found that the stable propagation of activity depended on the number and variance of spike events within sequentially active groups (Figure 13C, directional arrowheads), where activity either desynchronised (Figure 13Bi) or synchronised (Figure 13Bii). The parameters of the initial volley were significant in determining whether the volley would eventually stabilise into an attractor state (Figure 13C, cluster of arrowheads), where each consecutive wave fully fired with a

minimal standard deviation of spike-timing variation, or dissipated if the signal trajectory began to falter (Figure 13C, below the dotted line). By providing each neuron with a noisy train of Poisson generated spike events, where neurons were kept in a ready-to-fire state (Abeles, 1991; Dayan & Abbot, 2001), it was shown that signal transmission through precise spike timing was possible in a stochastic environment (Diesmann, et al., 1999).

These consecutively connected neuronal groups, called synfire chains, might operate as the highways through the brain that enable disparate brain regions to communicate over long distances, where signals either dissipate or transmit dependent on the power of activity at the initiating end (Diesmann, et al., 1999). To further demonstrate the plausibility of synfire chains to convey information in a noisy environment, a later model (see Kumar, et al., 2008 for model) realised a more biophysically realistic circuitry. Here, a body of excitatory and inhibitory neurons were created with interacting synaptic connections, where the former initiated an EPSP and the latter an IPSP to downstream neurons. Each node formed connections on a nearest neighbour basis to form a locally connected random network, where activity was naturally rhythmic. A feed-forward network, similar to Diesmann, et al., 1999, was then embedded within this framework such that the network and environment interacted with one another. This allowed the authors to consider the properties of the environment that determine whether activity would dissipate, propagate into stable waves or develop into an epileptic spreading. To this effect, it was found that signal propagation was favoured by environments with asynchronous and low firing rates.

Several feed-forward networks were also embedded to ascertain whether several chains could co-exist, or even overlap, within the environment. Figure 14 shows how several chains could coexist in an environment with asynchronous network activity, where the amount of overlap in a neuronal link determines whether the chains will interact or not, i.e. trigger one another when the signal traverses the overlapping node. By further investigating the possibility of synfire chains to convey information through precise spike timing in a noisy environment, Kumar, et al., 2008 determined

that the environment itself must be in an initial state that is favourable to the transmission of multiple overlapping signals.

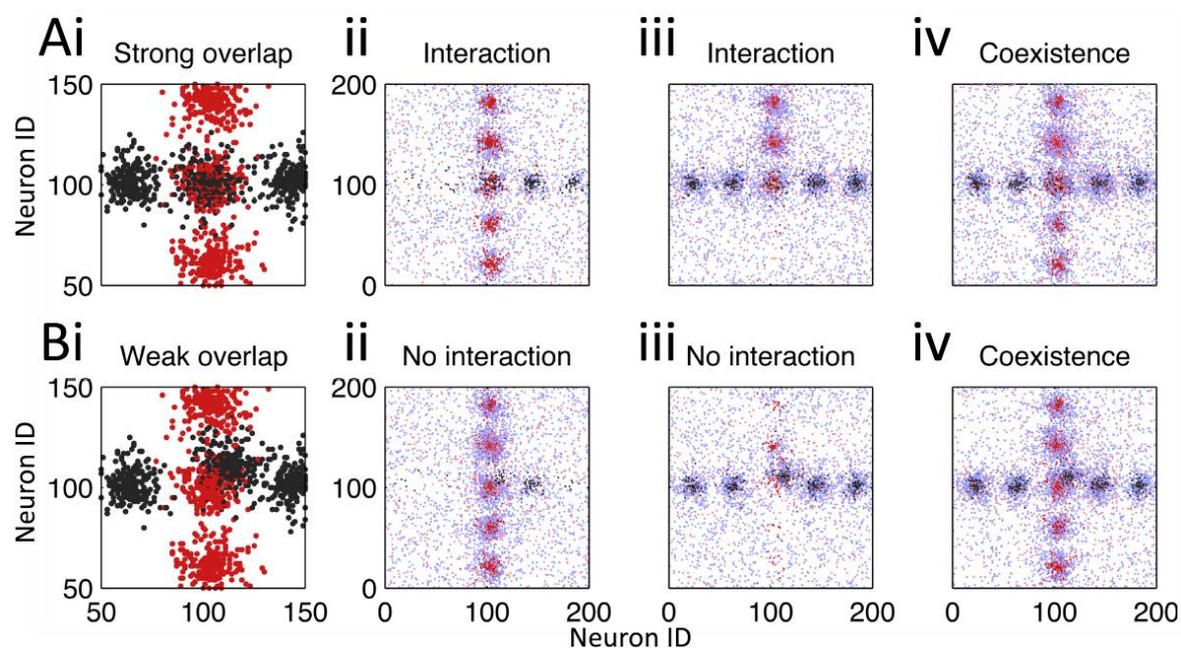


Figure 14 – Conditions for propagating synchronous spiking. Multiple synfire chains were embedded in a recurrent random network with strong (A) or weak (B) overlap in one neuronal group. **Ai/Bi** shows a weight matrix with positive weights of two synfire chains (red/black). **Aii-iv/Bi-iv** show mean firing rate of the entire environment through time, where dense red/blue regions indicate the activation of distinct synfire chains. Chains did or did not interact when initiating either one chain or the other (**ii-iii**) depending on the amount of shared overlap. Both chains could also coexist (**iv**) if both are simultaneously initiated, and do not lead to an epileptic spreading of activity. Figure adapted from Kumar, et al., 2008.

2.3 SUMMARY

This methods chapter has explained some of the neurophysiological processes of neuronal interactions in the brain (Shepherd, 1994), as well as presented some models that attempt to capture these findings with varying degrees of abstraction (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000). This includes how the biophysical structure of a neuron enables dynamic action potential generation (Hodgkin & Huxley, 1952), which can be further abstracted to an integrative point neuron with a single decay term (Abeles, 1991; Dayan & Abbot, 2001), or further still to a node that represents activation from a group of neurons (O'Reilly & Munakata, 2000). A common formula (Equation 15) for the effect of a pre-synaptic action potential on a post-synaptic neuron has been described (Dayan & Abbot, 2001), as well as how independent spike events from various unknown sources keep neurons in a ready-to-fire state (Abeles, 1991; Dayan & Abbot, 2001). We have explored how calcium and neurotransmitter dynamics facilitate communication and learning (Hebb, 1949; Shepherd, 1994), where theories of competitive (Song, et al., 2000) and calcium-based (Graupner & Brunel, 2012) spike-time-dependent-plasticity have accurately described these processes. A theory of signal transmission was then explored, where chains of neurons facilitated waves of activity to propagate through a noisy environment (Diesmann, et al., 1999; Kumar, et al., 2008).

An overarching theme of this thesis is that it is important to choose the simplest abstraction of neurophysiological processes for the desired functional requirements, in an Ockham's Razor approach. The theories presented here contain all the mathematical principles that will be used throughout the literature review and modelling work of this thesis. This allows us to discuss with confidence an overview of neural anatomy (O'Reilly, et al., 2011) and some key principles of learning (Tulving & Donaldson, 1972), as well as the principles of oscillatory dynamics in the brain (Fell & Axmacher, 2011) and what effect these might have on cognitive functions (Buzsaki, 2002; Klimesch, et al., 2007; Hanslmayr, et al., 2016). Further Chapters will refer to the methods presented in this Chapter when designing novel neural network models.

3. HUMAN MEMORY

Now that we have described some of the neurophysiological mechanisms (Shepherd, 1994) that can be used to model neural operations (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000), we here overview some anatomical regions that are key to some theories of human memory. To begin with, we review the anatomy of cortical and medial-temporal-lobe (MTL) regions, such that we can present a theory on the general schematics of information flow and episodic learning. In essence, it is theorised that the MTL is responsible for online learning of the environment, whilst the neocortical outer layers of the brain slowly integrate these newly formed memories to store them in stable long-term memories (O'Reilly, et al., 2011). In this Chapter, we present evidence for this argument alongside theories of ordinal (Bowman & Wyble, 2007) and temporal (Goldman, 2009; Itskov, et al., 2011; Shankar & Howard, 2012) sequence encoding. This Chapter describes the functional brain regions that form the basis of the neural network models presented in this thesis.

3.1 COMPLIMENTARY LEARNING SYSTEMS

The Complimentary Learning Systems framework (CLS) was one of the first connectionist models to provide a theoretical understanding of the functional organisation of the brain (McClelland, et al., 1995), by combining two distinct and interacting systems that could engage in different types of learning. One of the fundamental problems in connectionist neural network models at the time was that a highly overlapping & distributed representation led to the problem of catastrophic interference (McCloskey & Cohen, 1989), where learning new information would overwrite what had already been learnt. Whilst some argued that this indicated a major flaw in the ability of distributed representation approaches to explain human memory, such systems were highly effective at performing generalisation and inference tasks and were appropriate for learning new information in an on-line environment (Rumelhart, et al., 1986). At the other end of the spectrum, a non-overlapping sparse representation is much more robust to interference, able to store memories effectively but unable to make quick comparisons between old and new information

(McClelland, et al., 1995). The CLS framework thus proposed that any learning system that requires both on-line learning and the storage of stable memories, as the human brain does, requires the versatile pairing of both an overlapping distributed and a non-overlapping sparse representation (O'Reilly, et al., 2011). The general neural network model that realised these interacting memory systems was then able to account for a wide range of neurophysiological and behavioural data.

3.1.1 NEUROLOGICAL COUNTERPARTS

It is thought that the role of the cortex is to act as a distributed and stable store for functions, behaviour and long-term memory (O'Reilly, et al., 2011). Cortical representations are argued to be highly distributed (O'Reilly, et al., 2011), meaning that a memory might cover many functional regions. This requires incredibly precise communication over the entire brain (Fell & Axmacher, 2011), as thoughts seek to concurrently activate disparate regions to internally represent the world to us. This type of information store is finely balanced, not wishing to override old information (McCloskey & Cohen, 1989) nor to refuse entry to vital new information (Rumelhart, et al., 1986), often described as the stability-plasticity dilemma (Carpenter & Grossberg, 1988). This prompts consideration of the medial-temporal-lobe (MTL), see Figure 15, a densely connected region in the centre of the brain with an incredibly fast learning rate (Squire, 1992; Squire, et al., 2004). Here, new memories can be formed quickly as we go about our daily lives, ready to be copied to the stable storage of the neocortex (O'Reilly, et al., 2011; Winocur, et al., 2010), a process thought to occur during sleep (Diekelmann & Born, 2010; Walker, 2018).

The MTL is key for memory encoding (Squire, et al., 2004) and recall (Eichenbaum, et al., 2007), interacting with cortical regions through the gateway of the entorhinal cortex (Figure 15; O'Reilly, et al., 2011). Upon entering the MTL, information is integrated within the hippocampus, as depicted in the illustration of Figure 15. The recurrent architecture of hippocampal pathways is thought to enable learning, recognition and familiarity-based memory (Eichenbaum, et al., 2007). Integrated information is then passed back to the entorhinal cortex, ready to be internally represented by

further cortical regions (O'Reilly, et al., 2011). In this way, the MTL is thought to be a functionally distinct information store (Squire, et al., 2004) that navigates our on-line existence, tying together events in space (O'Keefe, 1979) and time (MacDonald, et al., 2011) and contextualising them with experiences from long-term memory (Eichenbaum, et al., 2007).

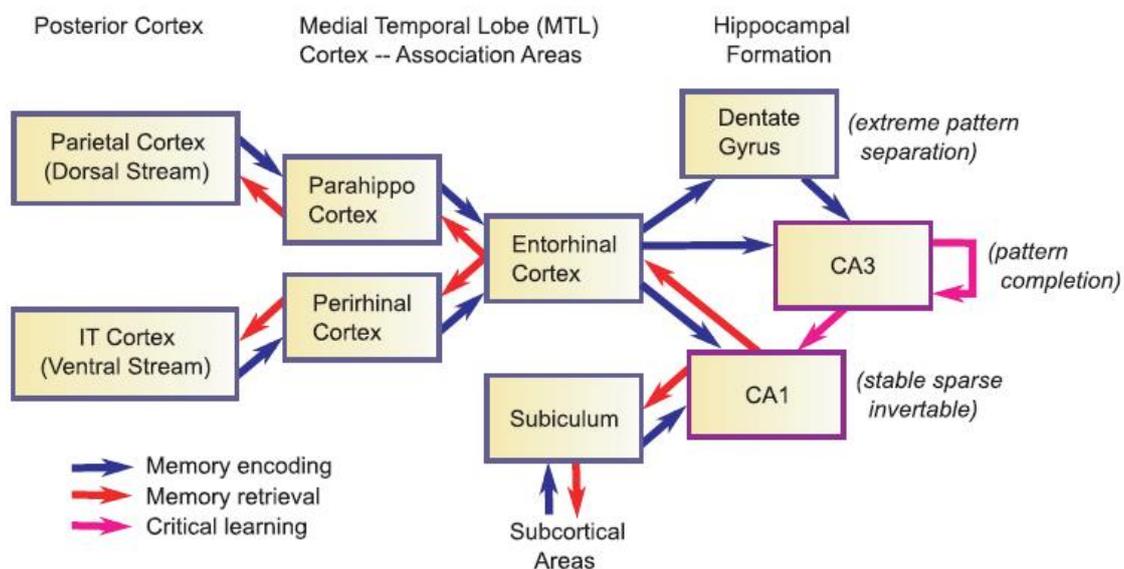


Figure 15 – Theorised informational processing pathways from cortical to medial temporal lobe (MTL) regions. Figure re-represented from O'Reilly et al., 2011.

3.1.2 MEMORY INTERACTIONS

The hippocampus is believed to be essential for declarative memory, a form of long-term memory (Tulving & Donaldson, 1972), required to temporarily bind together distributed sites in the neocortex that together represent a whole memory (Squire, 1992). Long-term memory is made up of several components, specifically declarative (i.e. explicit) and procedural (i.e. implicit) memory (Tulving & Donaldson, 1972). The latter is skill-based memory that we do not need to consciously recall, thus it does not require the hippocampus to function (Squire, 1992). The former is split further into episodic (i.e. contextualised events) and semantic (i.e. fact based) memory, both of which depend on the hippocampus and MTL (Squire, 1992). As defined by Tulving (1972), semantic memories are predominantly factually based with an absence of temporal or spatial details,

therefore lacking any particular context. It is thought that episodic memory evolved from semantic memory, as many animals are assumed to be able to encode factual information about the world but lack the ability to recollect detailed episodes of the past (Tulving & Donaldson, 1972).

One of the key elements of the CLS framework is the distinction between recollection and familiarity-based recognition memory (Eichenbaum, et al., 2007; O'Reilly, et al., 2011). Due to the pattern separation mechanism and conjunctive representations within the hippocampus, this has been theorised to be the main driver for recollection memory – where specific episodes can be matched and re-instantiated quickly (Eichenbaum, et al., 2007). The cortical system on the other hand contains much more overlapping and distributed representations, suggesting that it may naturally compute a long-lasting, global-match familiarity signal (O'Reilly, et al., 2011). The former can quickly recall specific episodic details, whilst the latter produces something more like a continuous signal of general recognition.

This key difference in neural architecture is also believed to play a role in the substance of the memories stored. Memories in the hippocampus are argued to be contextualised episodes with fine details (Squire, 1992), whereas the neocortex may extract a highly semanticised and generalised version as it integrates the original episodic memory with all lived experience (O'Reilly, et al., 2011). Hippocampal memory might also continue to exist in its original form if it remains in active use, meaning that memory consolidation is rather two complimentary learning systems encoding their own version of events (Winocur, et al., 2010) than a transfer of memory from short to long-term. This interweaving of learning systems has been theorised to eliminate the catastrophic interference dilemma (Carpenter & Grossberg, 1988; McCloskey & Cohen, 1989), and predictions that older memories should remain intact with damage to the on-line learning system have been upheld by observations of patients with hippocampal damage, where patients could recall long-term memories but could not store any newly formed memories (O'Reilly, et al., 2011).

Novel information is thought to be encoded through working memory, which acts as a short-term memory storage (Tulving & Donaldson, 1972) that is dependent on interactions with the attentional control system of the prefrontal cortex (Klimesch, et al., 2007). Neural representations must be sustained over a period of several seconds, even without continuous input, whilst attentional control performs operations on them (Fell & Axmacher, 2011). Throughout the day, cortical sites project onto the MTL so that the region can encode new episodic memories from the contextualised sensory information that is maintained in working memory (O'Reilly, et al., 2011). The direction is reversed, however, during sleep. During this time, memories are repeatedly presented by the hippocampus to the neocortex (Diekelmann & Born, 2010), where temporal sequences are played at a faster rate than that at which they were encoded, effectively squeezing a days' worth of memories into the few hours of slow-wave-sleep (SWS) that we get each night (Walker, 2018).

3.2 EPISODIC MEMORY

Episodic memory is the memory of autobiographical events, contextualised to integrate who, what, when, where and why information (O'Reilly, et al., 2011). It has been proposed to be encoded in the MTL, a structure consisting of the hippocampus, entorhinal cortex, parahippocampal cortex, perirhinal cortex and amygdala, as shown in Figures 15. As the hippocampus is theorised to be placed atop of a hierarchy of informational processing pathways (Figure 15; O'Reilly, et al., 2011), it can integrate information from all over the cortex so that it can form an on-line conjunctive representation of events (Squire, 1992; Squire, et al., 2004). Information is funnelled into the hippocampus through the entorhinal cortex, which in turn integrates information from the parahippocampal cortex (feeding spatial information from parietal areas) and the perirhinal cortex (feeding categorical information from inferior-temporal areas).

The hippocampus is subdivided into several functional areas to dissect arriving information, culminating in the activation of CA3 and CA1 areas. The CLS theory suggests that the highly distributed CA3 area enables pattern completion due to its dense recurrent synapses, which in turn

drives the CA1 region, illustrated in Figure 15. Successful memory encoding entails the strengthening of associations between active neurons in CA1/CA3. These can later be recalled by a triggered completion of CA3 patterns driving CA1 neurons, which themselves have feedback connections to the entorhinal cortex and are thus able to re-instantiate original activity patterns across the cortex through a series of feedback cortical routes. Within this commonly accepted theory of hippocampal architecture (Eichenbaum, et al., 2007; O'Reilly, et al., 2011; Squire, 1992; Squire, et al., 2004), the dentate gyrus is thought to play an important role in avoiding interference, whereby sparse levels of activity might perform pattern separation, thus enabling the representation of non-overlapping patterns in the downstream CA3 region despite its densely interconnected architecture.

This description of hippocampal architecture leaves the CA1 region with no immediate functional role, except that when it is removed from this conceptual framework the problem of catastrophic interference (McCloskey & Cohen, 1989) remains due to the integration of incoming information and pattern completion within the entorhinal cortex (O'Reilly, et al., 2011). Thus, the CA1 region must play a role in creating an integrative sparse representation of an event occurring in time and space, acting as a middle-man relay between pattern completion and the encoding and retrieval of selective events. Indeed, various cells have been identified in this area that specifically encode for spatial (O'Keefe, 1979), temporal (MacDonald, et al., 2011) and conceptual (Quiroga, 2012) information. Hence, the integrative hippocampal CA1 region will be dissected in the following subsections on the what, where and when of episodic memory.

3.2.1 ON WHAT & WHERE

In order for us to create memories rooted in the environment around us, we must be able to internally represent our place within it relative to the external world. This is supported by two types of cells, place cells, that fire when the subject is at a specific location, and grid cells, that map the current environment in a grid-like approach to aid relative navigation (Figure 16; Moser, et al.,

2008). Place cells are located in the CA areas of the hippocampus and together, the entire environment is represented by a local cell population where each node codes for a particular spot in physical space (O'Keefe, 1979). A recording from a single place cell in Figure 16A shows how the neuron only spikes at specific locations (Moser, et al., 2008). Grid cells were also located in the medial entorhinal cortex (MEC) that showed multiple sharply tuned spatial firing fields (Figure 16B), where it is thought that they create a path integration-based neural map. Through recurrent entorhinal-hippocampal projections (see Figure 15), this may allow us to engage in route-planning, as we integrate our current position relative to the environment (Moser, et al., 2008), a conjunctive representation that is thought to be processed in the MEC (McNaughton, et al., 2006).

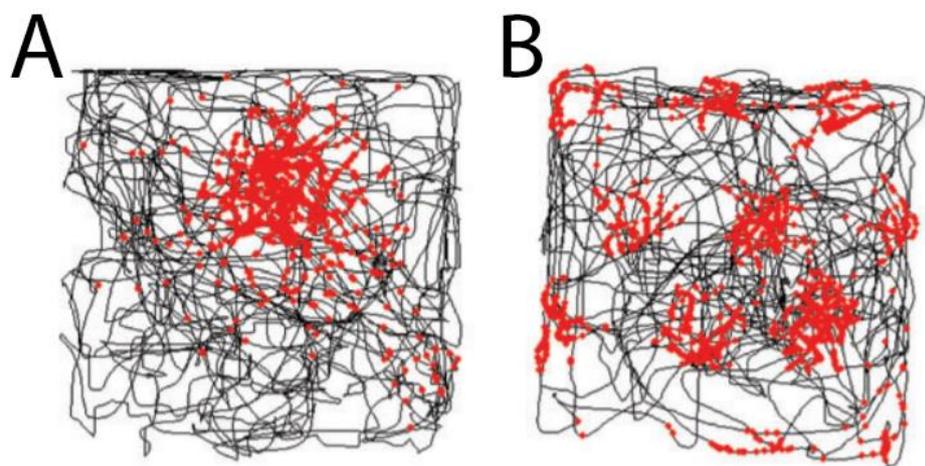


Figure 16 – (A) Firing of a hippocampal place cell and (B) firing of a medial entorhinal cortex grid cell. Spike locations (red) are superimposed on the trajectory of the animal as it moves around an enclosure. Figure adapted from Moser, et al., 2008.

In addition to spatial and temporal information, cells have been identified in the MTL that act as unique identifiers to specific concepts (Quiroga, 2012). By recording from single cells within various parts of the MTL, these concept cells seem to be mostly located in entorhinal cortex and hippocampal sites, shown by high levels of selectivity and visual invariance (i.e. sparse firing for unique concept) in the table of Figure 17A. Here, response times increase as one traverses the informational processing hierarchy (as seen in Figure 15; O'Reilly, et al., 2011) and responses are

passed forward from one area to the next. The number of responsive units (neurons that fire for at least 1 image) also decreases in this direction, indicating that each area encodes for increasingly sparse representations. Units in the entorhinal cortex and hippocampus are also more likely to respond to text and sound versions of the image, indicating that these cells integrate over many sensory sources when representing conceptual information (Quiroga, 2012). Figure 17B also shows a theoretical construction of the sparse representation of concepts in the MTL. In this instance, the sparsely coded cell assemblies that encode for the concept of 3 distinct characters from the film Star Wars might share some nodes, whereby the activation of Luke Skywalker might lead to the activation of Yoda and Darth Vader due to overlapping representations (Quiroga, 2012).

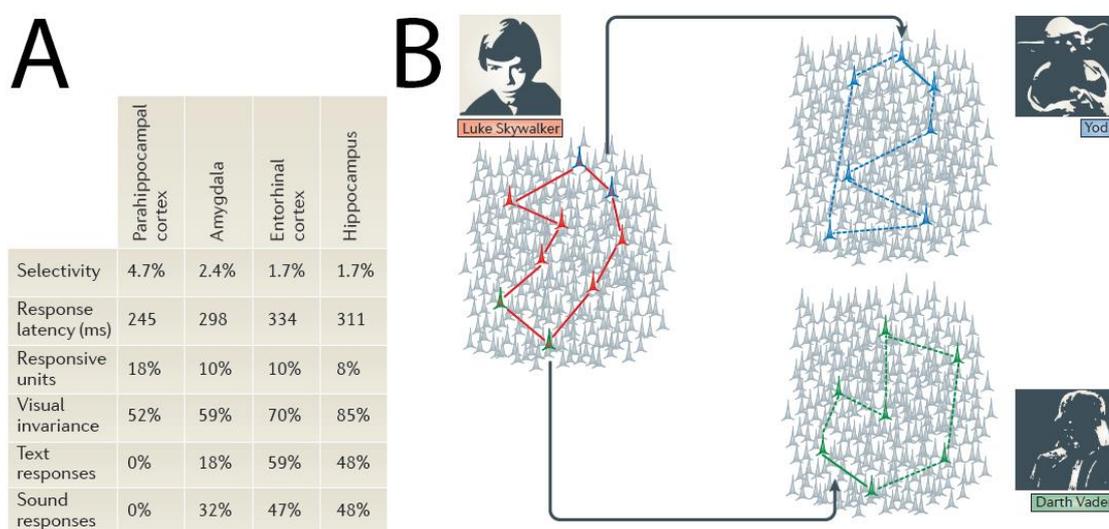


Figure 17 – **A** Table showing selectivity, response latency, responsive units (number of cells responding to any image), visual invariance, text and sound responses of cells in the medial temporal lobe in response to pictures of people and places. **B** Hypothesised overlapping representations of several cell assemblies that encode for specific concepts. Figure adapted from Quiroga, 2012.

A further study has shown that identified concept cells can associate with one-another after a single presentation (Ison, et al., 2015). In Figure 18, an image was found that the isolated MTL cell responded to, named the preferred (P) stimulus, and an image was found that the cell did not

respond to, named the non-preferred stimulus (Figure 18A), where the activity of the cell was recorded over several repeated trials (Figure 18B). Then, both images were layered over one another and presented as a composite (C) stimulus during a learning phase, whereby associations were built up between the relevant cell assemblies encoding for the P and NP stimuli. The effect of the increased synaptic efficacies could be seen after just a single presentation, as when the P and NP images were screened after the learning phase, the isolated cell responded strongly to the NP image as well as to the P image (Ison, et al., 2015).

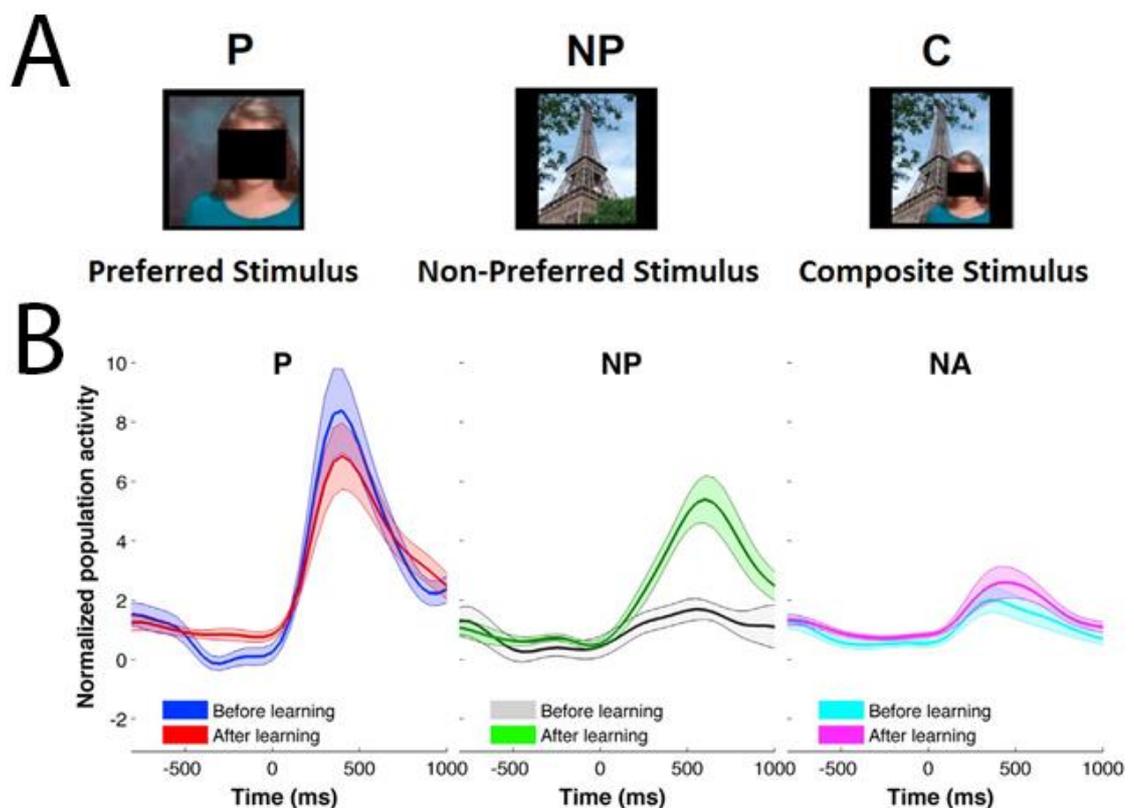


Figure 18 – **A** Images of people and places, where the preferred (P) stimulus is the image the cell responds to, the non-preferred (NP) stimulus is an image the cell does not respond to, and the composite (C) stimulus is a combination of the two. **B** Activation of the cell as it responds before and after learning, where the P & NP images are presented individually prior to the composite stimulus, then again afterwards. A non-associated (NA) unit was also found, acting as a control case for P & NP learning. Figure adapted from Ison et al., 2015.

This indicates that the cell assembly encoding for the P image now had some overlapping nodes with the cell assembly encoding for the NP image, where the presentation of one encourages the pattern completion of the other, as predicted by Quiroga (2012) in Figure 17B. This notion is prominent for the first modelling work of this thesis (Parish, et al., 2018), which models the strong synaptic plasticity that enables hippocampal episodic memory (O'Reilly, et al., 2011; Squire, 1992; Squire, et al., 2004).

3.2.2 ON TIME

There appears to be multiple time scales that humans are capable of perceiving (see Mauk & Buonomano, 2004, for review). One such time scale occurs during the encoding and retrieval of episodic information (see Eichenbaum, 2014, for review), where conscious time estimation occurs in the realm of seconds. In this way, the cortex has been described as a distributed time-keeper (Mauk & Buonomano, 2004), where long-term episodic memory is thought to be stored in a CLS framework (O'Reilly, et al., 2011). Whilst selectivity to temporal stimuli appears to be localised to right parietal, dorsolateral pre-frontal and auditory cortical areas, it is thought that cortical neural networks are generally capable of intrinsically processing temporal sequences (Mauk & Buonomano, 2004). To this effect, a recent model has shown how temporal sequences might even be inherited from higher-order chains (Itskov, et al., 2011) in a CLS framework, whereby cortical episodic sequences could be inherited from newly encoded episodic memories in the hippocampus.

As discussed previously, episodic memory is thought to be mediated by the hippocampus (O'Reilly, et al., 2011; Squire, 1992), which has been found to play an integral role during encoding and retrieving of both ordinal representation, the order of events, and temporal representation, the delays between events (see Eichenbaum, 2014, for review). On a cellular level in the hippocampus, whilst the aptly named place cells were found that activate for specific spatial fields, there are also cells of the same type that have been found to uniquely activate for specific temporal periods (MacDonald, et al., 2011). Together, time & place cells can perform spatio-temporal integration

that enables us to encode episode patterns. Temporal information can be extrapolated from this, where cells fired at the same time regardless of distance covered (Pastalkova, et al., 2008) and cells fired in sequence to cover timed durations independent of spatial or behavioural information (Eichenbaum, 2014; MacDonald, et al., 2011).

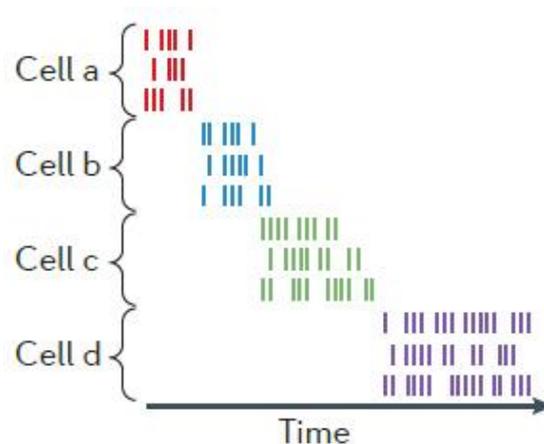


Figure 19 – Time cells in the hippocampus reliably activate in sequence to cover a defined period, observed over multiple trials for each identified cell. Figure adapted from Eichenbaum, 2014.

Research has shown that the MTL can encode temporal information that persists over long periods (see Howard & Eichenbaum, 2013 & 2015, for reviews). This is thought to occur through a gradually changing state of temporal context that serves as a cue for episodic recall, where it is possible for a repeated item to enable a ‘jump back in time’ and recover a previous state of temporal context, as shown with MTL epilepsy patients (Howard, et al., 2012). Similarly, rat studies have found that a position on a previous temporal sequence can be recovered, where distinct movements prior to a wheel running task in rats triggered separate chains of time cells, and rats could also distinguish between odours that had occurred in past positions on competing lists (see Eichenbaum, 2014, for review). This indicates that the hippocampus holds many temporal sequences that it can successfully associate, disambiguate and navigate between, as has been hypothesised to occur in spatial working memory (McNaughton, et al., 2006), where the lateral-entorhinal-cortex (LEC) is thought to perform this integrative operation of time, based on feedback interactions with the hippocampus (Tsao, et al., 2018).

Further evidence comes from timing experiments, which indicate that there is a degraded precision in memory for longer intervals, and un-cued memory tasks, where forgetting takes place over multiple time-scales with similar decay characteristics (see Howard & Eichenbaum, 2013, for review). In support of these findings, cells that came later in time cell sequences were found to be active for increasingly longer durations (as seen in Figure 19; Eichenbaum, 2014), suggesting that these cells enable a scalar coding of time (see Wearden & Lejeune, 2007, for review). This is exemplified by Figure 20B (Itskov, et al., 2011), where a reliability score, $R(t)$, is used to assess the extent to which population vectors from single trials resemble trial-averaged population vectors at time t , as opposed to trial-averaged population vectors at other times, where a population vector represents the pattern of activation through the network. Here, a positive score indicates that sequential time cell behaviour is reliably generated upon successive trials. Declining scores across time indicate scalar properties, where population vectors are more variable at later times, as indicated in the cellular activity of a single trial in Figure 20Aii between 10-15 seconds.

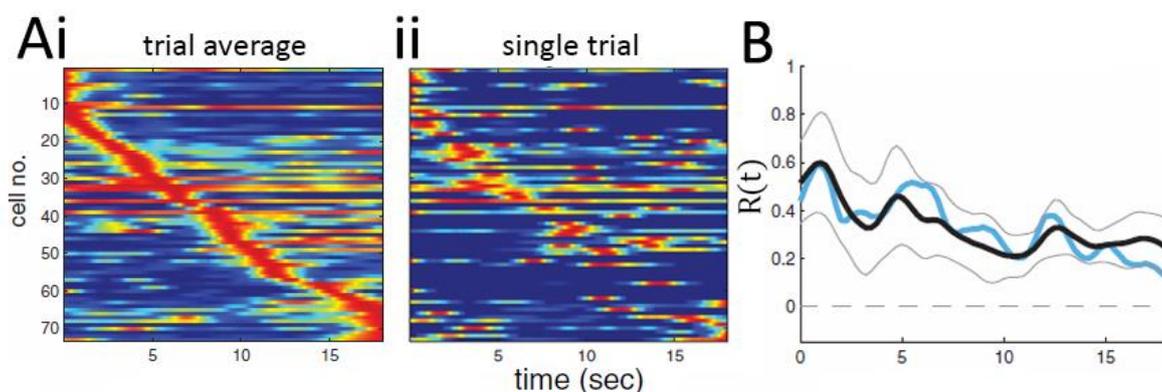


Figure 20 – **A** Time cell behaviour, where activity in rat hippocampal cells is averaged over many trials (**i**) and shown for a single trial (**ii**). **B** The reliability score ($-1 < R(t) < 1$) compared population activation vectors from individual trials to the mean through time. Black line with confidence intervals indicates average reliability across trials, blue line indicates the reliability for the single trial shown in **Aii**. Figure adapted from Itskov, et al., 2011.

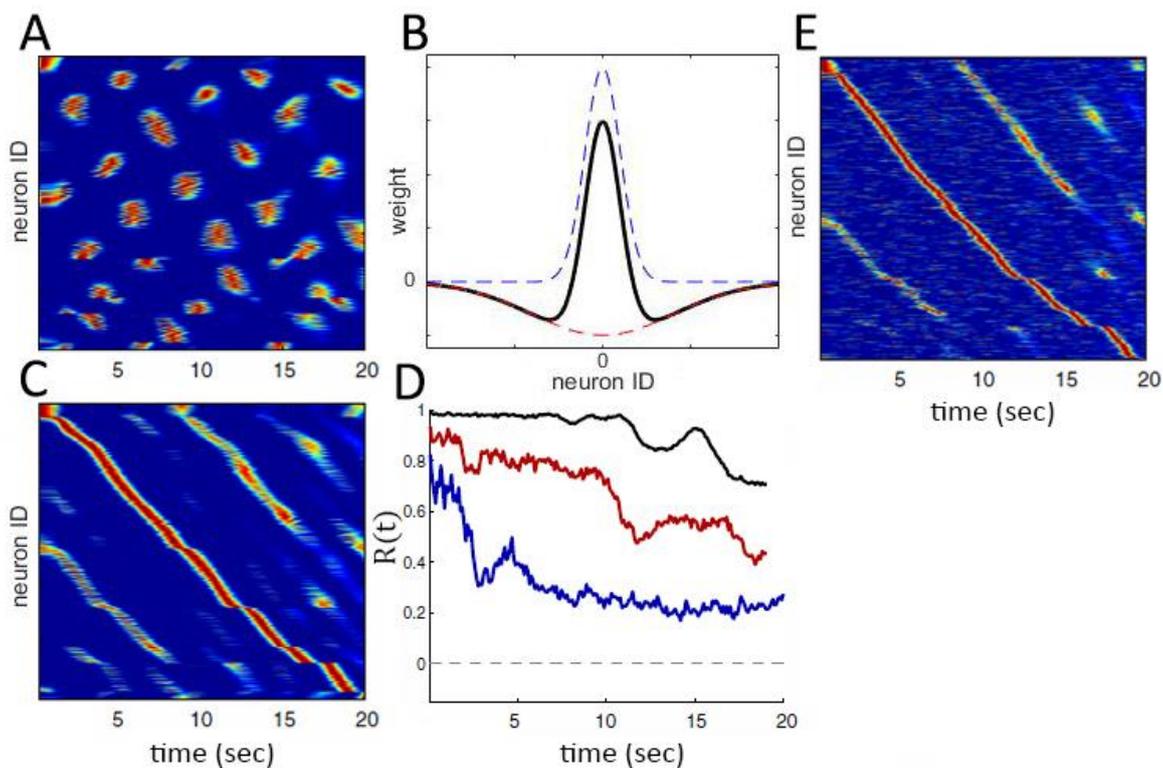


Figure 21 – ‘Mexican-hat’ topology (**B**, black line) with short-range excitatory (blue dashed) and long-range inhibitory (red dashed) connections encouraging clustered activation (**A**), where activation shifts between dominant clusters by increasing the threshold of activation for spiking neurons. When ordering neuron ID by time of activation (**C**) it can be seen that this architecture promotes time-cell behaviour comparable to that of Figure 20A, where cells transiently activate one after another in sequence to cover a duration. A reliability score of population vectors, $R(t)$, indicates scalar properties (**D**) that increasingly resemble those of hippocampal time-cells when increasing levels of noise are introduced into the network (black-red-blue). It was even shown that a separate network without a Mexican-hat topology can reliably inherit a sequence (**E**) if exposed to repeated presentations in the presence of STDP. Figure adapted from Itskov, et al., 2011.

One attractive method of modelling time cell behaviour with scalar properties is through evolving transiently active cell-assemblies, i.e. naturally occurring synfire chains (Diesmann, et al., 1999). In comparing simulated activity to that of time cells, one such model predicted that temporal sequences can be internally generated, which are reliable from trial to trial, context dependent and long lasting (see Itskov, et al., 2011, for model). Assuming the existence of a competitive

architecture within the hippocampus (McNaughton, et al., 2006), this model created a locally connected random network with short-range excitatory and long-range inhibitory connections (Figure 21B), similar to a previously described synfire chain model (Kumar, et al., 2008). This 'Mexican-hat' topology encourages a winner-take-all environment, where activation converges to a bump attractor point within the network, emulating the sustained activation of time cells (Eichenbaum, 2014). In order to shift activation out of the attractor point a threshold-adaptation mechanism was employed, whereby the membrane threshold varied relative to activation, a mechanism that has been observed in rodent hippocampi (Henze & Buzsaki, 2001). Thus, activation spread through the neuronal group (Figure 21A) dependent on the initial conditions of the network. This behaviour is similar to hippocampal time cells when neuron ID is ordered by time of activation (Figure 21C), where an accompanying reliability score indicates inherent scalar properties (Figure 21D), especially when there is a high level of noise within the network (black line is score for parameter set of Figure 21C, red-blue lines indicate increasing noise). The model also showed that a distinct network with no recurrent architecture, thus one that is unable to generate its own temporal coding, can nonetheless inherit a sequence through random sparse connectivity and an accompanying STDP mechanism (Figure 21E).

Another model further explores several recurrent and feedforward architectures (see Goldman, 2009, for model), showing that it is the feedforward nature between nodes that is most important for sequence generation, rather than the recurrent architecture within nodes that is itself important to sustain activation for longer durations. However, the purely feed-forward architecture of synfire chains demonstrated here (Goldman, 2009; Itskov, et al., 2011) has been criticised for lacking scalability (Shankar & Howard, 2012), as the desired duration of timed events linearly scales with physical length of the synfire chain. This complexity issue is compounded by the finding that the hippocampus encodes for temporal durations of many 10s of seconds (Howard & Eichenbaum, 2013) with the apparent existence of many competing temporal sequences that enable working memory, navigation and path integration (Eichenbaum, 2014; McNaughton, et al., 2006). This

criticism is further explored by the second model within this thesis, which employs a hierarchical synfire chain to tackle the issue of scalability in feedforward networks.

Whilst we have thus far discussed the hypothetical importance of a distributed neural time-keeper for episodic memory, another model argues that an additional mechanism is required to bind discrete events to a unique ordinal index when encoding sequences. This notion is an important building block for the 2nd modelling work of this thesis, discussed in later Chapters. Thus, the simultaneous type, serial token (ST²) model (see Bowman & Wyble, 2007, for model) instantiates a competitive architecture that separates events into distinct episodes by selecting a unique token (i.e. ordinal position) and binding it to concurrently active types (i.e. item representations). Here, it is theorised that the use of a binding pool facilitates the repetition of items, where each presentation is recast as a discrete event with a unique identifying tag. Similarly, it is also theorised that these events must be bound in sequence to enable ordinal memory. As such, Figure 22A shows how a winner-take-all environment is facilitated by recurrent connectivity between the binding pool and token layer, such that a unique identifying binding unit emerges alongside a unique ordinal position. These are bound to any concurrently active types, thus allowing the encoding of a discrete event (possibly containing multiple items) to be bound to a serial position in time. These units self-inhibit after this binding process, ensuring a sparse representation of ordinal events.

Within the model, item representation is sustained by activation from an additional 'blaster circuit', which is itself inhibited during the binding operation. This arrangement fuels the central hypothesis of the model; that our ineffectiveness at correctly recalling target items that are very close together in time (Figure 22C, attentional blink) is due to the functional process of binding (Bowman & Wyble, 2007; Swan & Wyble, 2014), where it is argued that in order for us to be able to encode distinct, non-overlapping episodes this binding process must not be interrupted, a notion that we will return to during the 2nd modelling work of this thesis. It is worth noting that this dip in perception performance is contained to within fractions of a second, perhaps indicating that this hypothesised trade-off between binding processes and attention has co-evolved to be as minimal as possible.

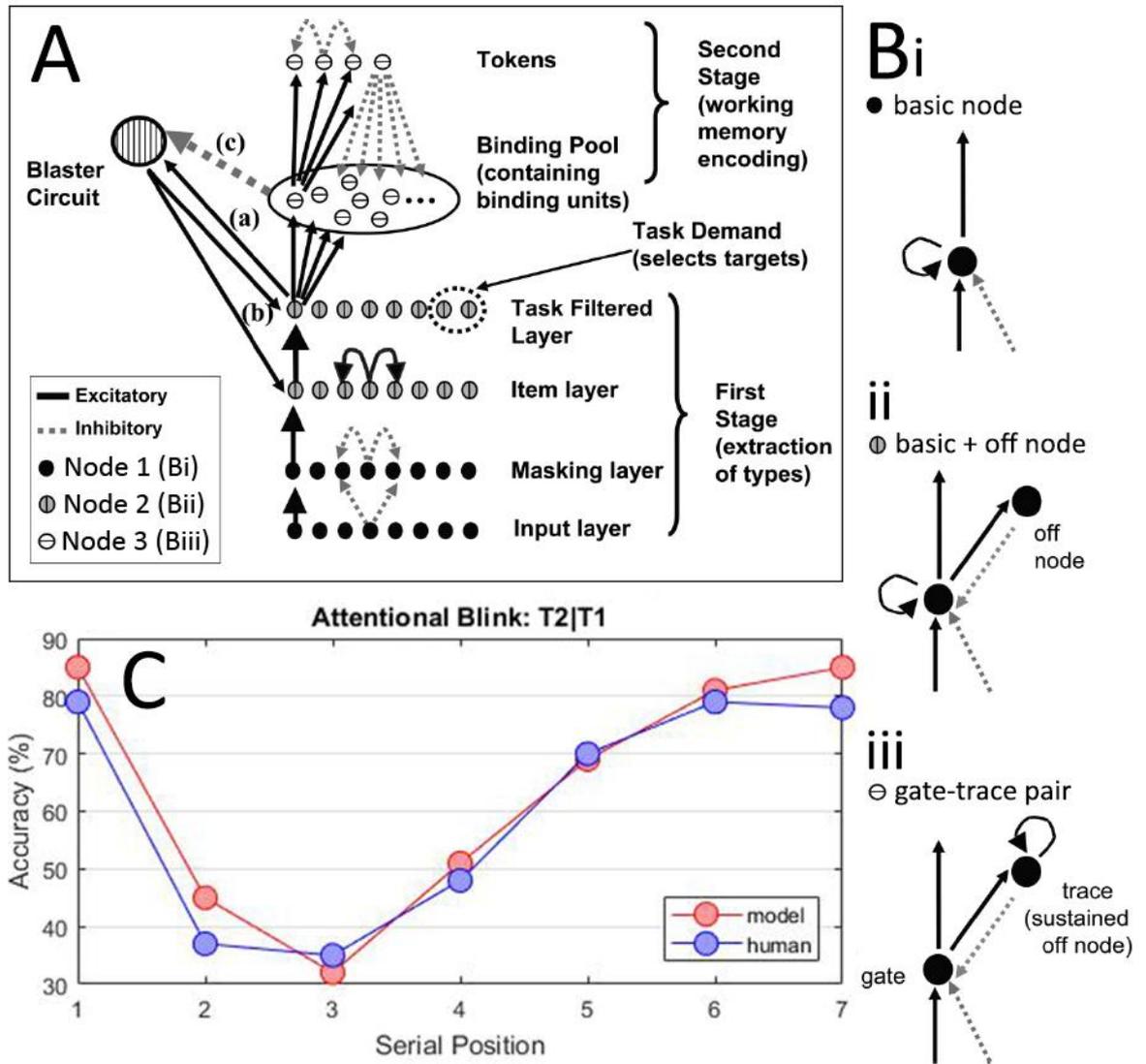


Figure 22 – Simultaneous type, serial token (ST^2) model. **A** Schematics, where the first stage, extraction of types, enables simultaneous event representation. Active types then stimulate the binding pool in second stage, working memory, which itself stimulates the token layer. Recurrent connectivity between binding units and tokens facilitates winner-take-all behaviour, assigning a unique binding unit and a unique token for a discrete event potentially containing several types. Excitatory (black lines) and inhibitory (grey dashed) connections enable distinct nodes to propagate sustained activation (**Bi**) or a single burst (**Bii**) or switch off altogether through sustained inhibition (**Biii**). The addition of an attentional blaster circuit enables the model to replicate the experimentally observed attentional blink (**C**), where performance is lower when two target items (T1 & T2) are presented closely together in time. Performance here is shown as the success rate of recalling T2 given that T1 was correctly reported. Figure adapted from Bowman & Wyble, 2007.

3.4 SUMMARY

This Chapter explored theories of human memory (Tulving & Donaldson, 1972), specifically the division of labour in a complimentary learning systems framework (O'Reilly, et al., 2011). In order for a learning system to avoid catastrophic interference (McCloskey & Cohen, 1989), it is thought that there exists a stable store of retrievable memories that is able to assimilate novel information from a separate learning mechanism with high plasticity (McClelland, et al., 1995). In the brain, this division is achieved by interactions between a distributed cortex and a sparsely coded hippocampus (O'Reilly, et al., 2011) that work in tandem to encode novel experiences during wakefulness (Squire, et al., 2004) and replicate them into long-term memory during sleep (Diekelmann & Born, 2010). This theory is fundamental to the first modelling work of this thesis (Parish, et al., 2018).

Episodic memory is key to human memory (Tulving & Donaldson, 1972), allowing the integration of unique concepts (Quiroga, 2012) in time (MacDonald, et al., 2011) and space (O'Keefe, 1979). Feedforward chains might enable us to generate temporal sequences (Goldman, 2009; Itskov, et al., 2011), though questions of scalability have emerged (Shankar & Howard, 2012). A binding pool might also be necessary to encode discrete events with repetition, where one model generates a unique tag to distinguish events (Bowman & Wyble, 2007). This binding operation has been theorised to account for lapses in our short-term working memory (Swan & Wyble, 2014). In later Chapters, the second modelling work of this thesis will take these ideas further in a unifying framework on the use of feedforward chains for temporal sequence encoding and retrieval.

Meanwhile, we next explore how the brain conjunctively represents information through precise spike timing, a transfer of information thought to be made possible by synchronised activity between disparate brain regions (Fell & Axmacher, 2011). The degree of the synchronisation between regions has been found to have a major impact on cognitive processes (Basar, et al., 2001); such as attention (Klimesch, et al., 2007), memory (Buzsaki, 2002; Fell & Axmacher, 2011) and information flow in the brain (Hanslmayr, et al., 2012), the subject of the following Chapter.

4. FUNCTIONAL OSCILLATIONS IN THE BRAIN

As discussed in the previous Chapter, the brain is a highly distributed storage of information (O'Reilly, et al., 2011). Therefore, precise communication is key between disparate brain regions (Fries, 2005). It has been argued that this is mostly achieved through neural oscillations (Fell & Axmacher, 2011). An oscillation occurs when assemblies of neurons synchronously spike during precise timing windows due to some modulating inhibitory effect (Klimesch, et al., 2007; Hanslmayr, et al., 2012). The first section of this Chapter explores the hypothesised functionality that oscillations can bring to neural circuitry in general (Fell & Axmacher, 2011), where synchronisation between disparate brain regions is thought to facilitate communication and learning. Some computational methods to simulate and analyse oscillations are also discussed here.

The second section explores the evidence for the secondary hypothesis of this thesis, namely how oscillatory functions might enable human working and long-term memory (Buzsaki, 2002; Fell & Axmacher, 2011). This section explores frequencies that originate from the medial-temporal-lobe (MTL), which have been found to modulate phases of learning (Hasselmo, 2005; Huerta & Lisman, 1995). Several models in the field are described which explore possibilities for the functionality of oscillations in MTL memory processes (Jensen, et al., 1996; Ketz, et al., 2013; Norman, et al., 2006).

The final section of this Chapter explores the evidence for the primary hypothesis of this thesis, namely how oscillatory functions might enable attentional processes (Klimesch, et al., 2007) and information flow (Hanslmayr, et al., 2012). The sources of these mid-frequency oscillations tend to be more distributed in nature, where cortical columns are thought to self-generate an oscillation at their own local phase (Jones, et al., 2000). Separate cell assemblies are then thought to be synchronised through a centralised top-down control process that helps to organise the flow of neural information (Klimesch, et al., 2007), forming the basis for attention and neural information flow (Hanslmayr, et al., 2012). This phenomenon is only more recently understood, meaning that

there is a noticeable lack of modelling work studying this phenomenon (Hanslmayr, et al., 2016), paving the way for the first modelling work of this thesis (Parish, et al., 2018).

4.1 OSCILLATION FUNCTIONALITY

This section describes some key concepts of neural oscillations and the possible functions that they may bring to communication and learning (Buzsaki, 2002; Fell & Axmacher, 2011; Klimesch, et al., 2007). An in-depth description of oscillatory processes, terms and hypothetical functions can be found in Appendix A, recommended for the reader wishing to gain a full insight into the functional role oscillations may play in the organisation of information flow within the brain – which is the overarching theme of this thesis. This section, however, focuses on the most relevant findings and models of interest that are important for the modelling works of this thesis.

4.1.1 EXCITATION-INHIBITION BALANCE

Neurons in the brain exist in a balance between excitation and inhibition, with a typical ratio of 80% excitatory pyramidal cells and 20% inhibitory interneurons in cortical cell assemblies (Markram, et al., 2004). A neuron is excitatory if its internal currents excite post-synaptic neurons at the time of an action potential, or inhibitory if they suppress activity (Dayan & Abbot, 2001; Shepherd, 1994). There is some variation in the internal dynamics between many classes of observed interneurons (Markram, et al., 2004), let alone between pyramidal cells and interneurons (Shepherd, 1994). In order to reduce complexity, a common modelling tactic is to use the same equations for both excitatory and inhibitory neurons, changing the sign of a post-synaptic-potential (PSP) dependent on the class of the simulated cell (Brunel, 2000; Hansel & Mato, 2003). Whilst there are many modelling studies that more precisely model the cellular generation of oscillations (Jones, et al., 2000; Wang, 2010), this thesis only considers simpler frameworks.

The excitation-inhibition balance enables assemblies to create their own local oscillations (Jones, et al., 2000). This balance has been explored in many computational models, with the emergence

of various synchronous and asynchronous states (Brunel, 2000; Hansel & Mato, 2003; Kitano & Fukai, 2007). One such cell assembly can be seen in the simulation of Figure 23, where an excitatory and inhibitory neuron share bidirectional connections. A Poisson distribution of spike events is fed into the excitatory neuron (Figure 23A, black synaptic current) to cause periodic spike events. These excitatory spike events are fed into the inhibitory neuron (Figure 23B, turquoise synaptic current) which in turn raises the membrane potential and generates several spike events. These inhibitory spike events are then summated in the excitatory neuron's synaptic currents (Figure 23A, cyan synaptic current) and decrease the membrane potential – effectively inhibiting the neuron from spiking again for a defined period. The frequency here is determined by the synaptic time constant of the inhibitory post-synaptic-potential (IPSP) coming into the excitatory neuron, where the higher the value then the longer the inhibition, thus the slower the frequency (Brunel, 2000).

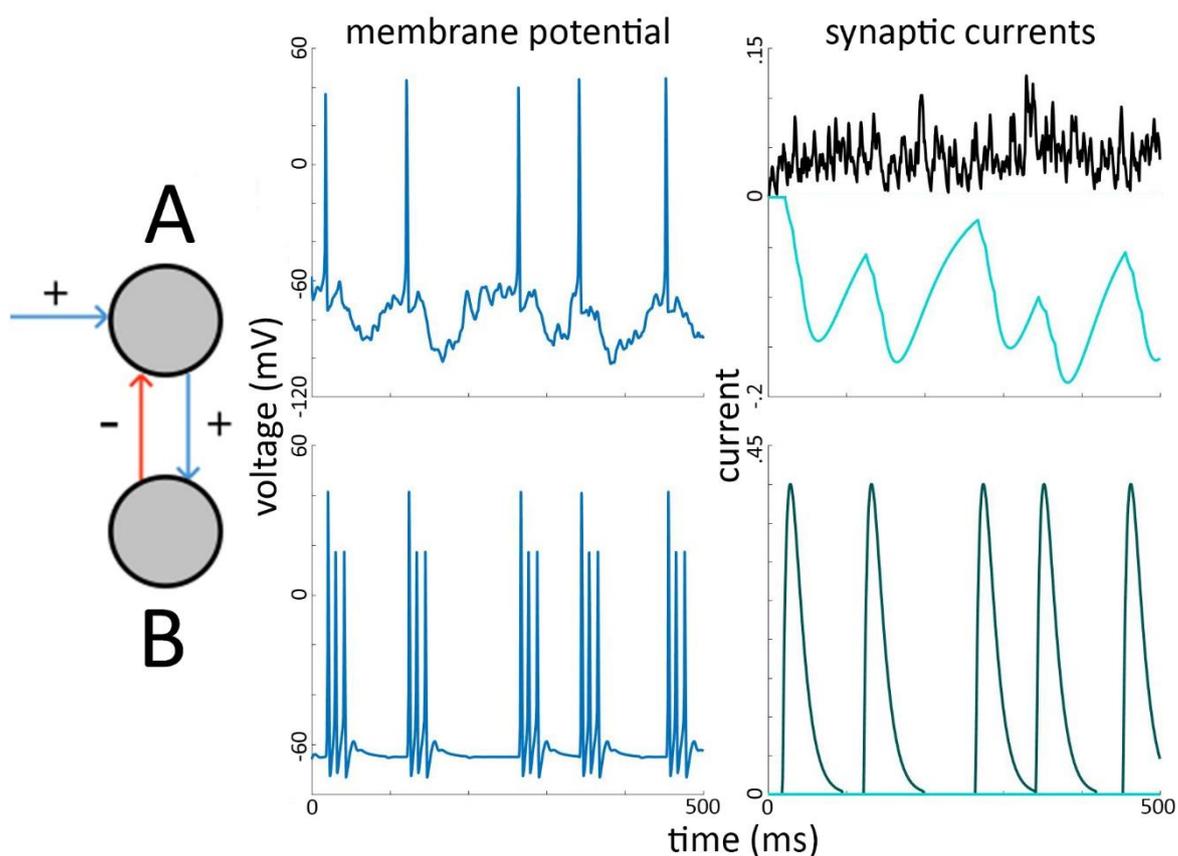


Figure 23 – **A** Membrane potential recording and synaptic currents of a model excitatory neuron, using Hodgkin & Huxley equations for action potentials. **B** The same for an inhibitory interneuron.

4.1.2 AMPLITUDE & FREQUENCY

As shown in Figure 23, populations of neurons are rhythmically activated through recurrent excitatory and inhibitory synapses, where excitatory phases are followed by an inhibitory phase, during which time neural assemblies are prevented from firing (Fell & Axmacher, 2011). This is especially true for cortical cells (Markram, et al., 2004), which dynamically oscillate in their local cell assemblies (Jones, et al., 2000). A pyramidal cell fires tonically if its excitation level is high enough to override the influence of this inhibition (Klimesch, et al., 2007), or rhythmically if either its excitation is low, or if the amplitude of the entraining oscillation is large (Figure 24). The important distinction then is that an increase in the amplitude of the oscillation, and therefore inhibition, may lead to more rhythmic activity rather than the suppression of activity due to the precise timing window for activation provided by each 'up' state (Klimesch, et al., 2007). Amplitude here refers to the strength, or power, of the observed oscillation. From this explanation, we can see that cell 1 of Figure 24A fires tonically as it overcomes the low-level inhibition of the oscillation, whereas cell 3 of Figure 24A fires much more rhythmically as spikes occur in precise time windows. Figure 24B shows how an increase in the amplitude can constrain all cells to fire within the more strictly defined up-states of the oscillation.

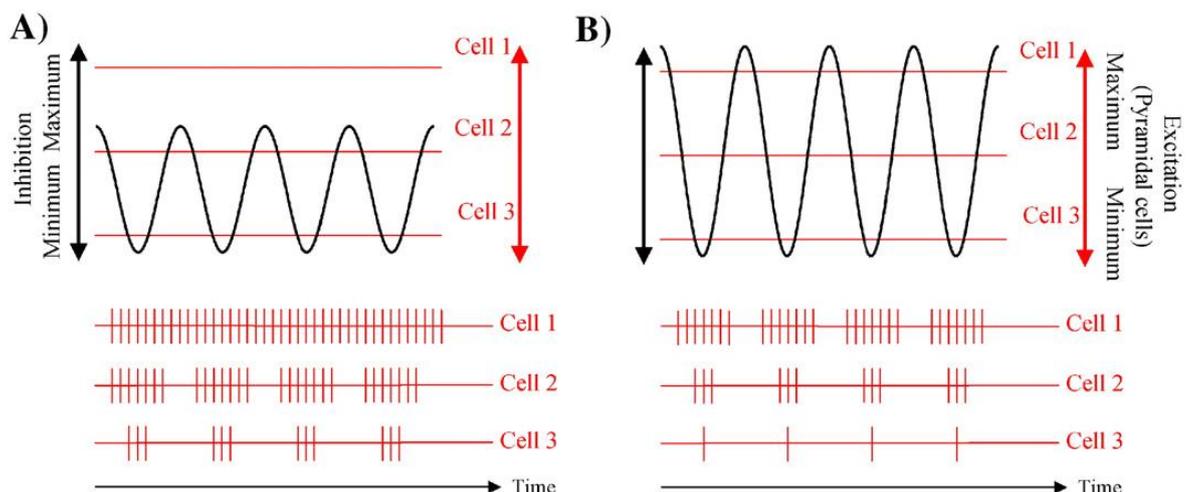


Figure 24 – Amplitude of oscillation impacts inhibitory, and therefore rhythmic, influence on target

neurons. **A** Low amplitude inhibition easily overcome for continuous firing. **B** High amplitude oscillation entrains neurons to fire precisely within time window. Figure from Klimesch, et al., 2007.

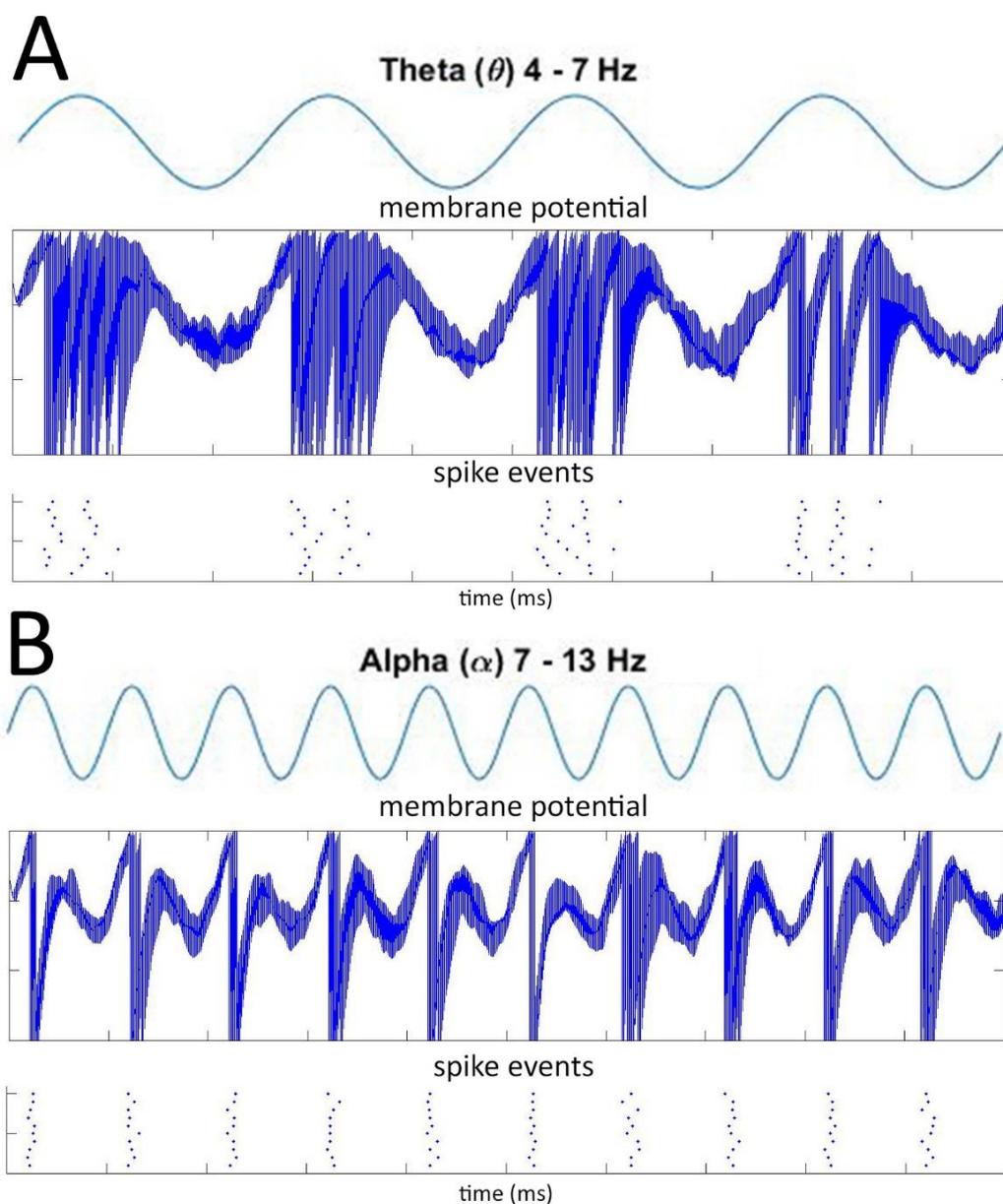


Figure 25 – Populations of integrate-and-fire neurons oscillating at a 4Hz (**A**) or 10Hz (**B**) frequency.

Some frequencies are better suited than others at mediating precise neural communication (Fell & Axmacher, 2011), as shown by the simulations of Figure 25. Here, integrate-and-fire neurons were fed a sine wave alternating-current (AC) generator in a couple of frequencies (see Appendix A for description of neural frequencies and their hypothesised neurological origins), shown in the top

panels. This effectively models a population of neurons under the influence of an external neuromodulator, thought to be more common in lower frequencies originating in archi-cortical matter (Buzsaki, 2002). Importantly, this means that the pyramidal cells cannot influence the source of the inhibition, limiting their ability to generate phase resets and drive phase synchrony.

In this set of simulations (Figure 25), neurons are fed with a Poisson distribution of spike events to simulate background noise, where the variability can be seen in the voltage trace plots in the middle panels. The bottom panels contain a raster plot, where spike events are shown as blue dots. These simulations show that lower frequencies provide a broader timing window for activation, thus allowing neurons to recover and spike again during the up-state (Figure 25A). The smaller window of activation in Figure 25B only allows cells to activate once during this time. This means that both the amplitude (Klimesch, et al., 2007) and the frequency (Fell & Axmacher, 2011) of the entraining oscillation determines the rhythmicity of activation.

4.1.3 PHASE

There is growing evidence that neural oscillations play a key role in the organisation of information flow within the brain (Basar, et al., 2001; Buzsaki, 2002; Fell & Axmacher, 2011; Klimesch, et al., 2007; Hanslmayr, et al., 2012), see Appendix A for further explanation of how phase synchrony enables information representation and memory processes. In short, phase synchronisation refers to the aligning of phase between the oscillations within distinct brain regions, a process that has been found to support neural communication, plasticity and many cognitive processes (Fell & Axmacher, 2011). Phase synchronised regions are thus pre-prepared for optimal communication, as both regions are entrained to the same up-states with no inhibition to overcome. Otherwise they are blocked from communicating, as information sent from one region cannot overcome the inhibitory down-state of the other. The phase of an on-going oscillatory cycle serves as a frame of reference for both internal and external events, where phase-resetting is an important mechanism

for phase synchrony and frequency coupling, allowing mutually coupled oscillators to coordinate their frequencies and phases to maximise communication (Canavier, 2015; Fell & Axmacher, 2011).

Due to the various theta phases existing between distinct hippocampal regions (Buzsaki, 2002; see Appendix A for more detail on theta rhythm and current generators), the timing of action potentials in relation to phase is also likely to play a role in information representation and learning. If recording the LFP from the CA1 pyramidal layer as the reference oscillation, the majority of hippocampal pyramidal cells fire on the negative phase of the theta cycle (Hasselmo, 2005), where fluctuations in phase depend on behavioural variables. For example, stronger dendritic excitation might cause a neuron to fire at increasingly earlier phases relative to the mean local field. This has most famously been observed in place cells, where activation not only increased but advanced in phase the closer a rat approached the place field of the recorded unit (O'Keefe & Reece, 1993). This indicates that both an increase in activation and the timing of spikes relative to a reference oscillation can signal 'what' is occurring and 'where,' respectively (Buzsaki, 2002; Fell & Axmacher, 2011).

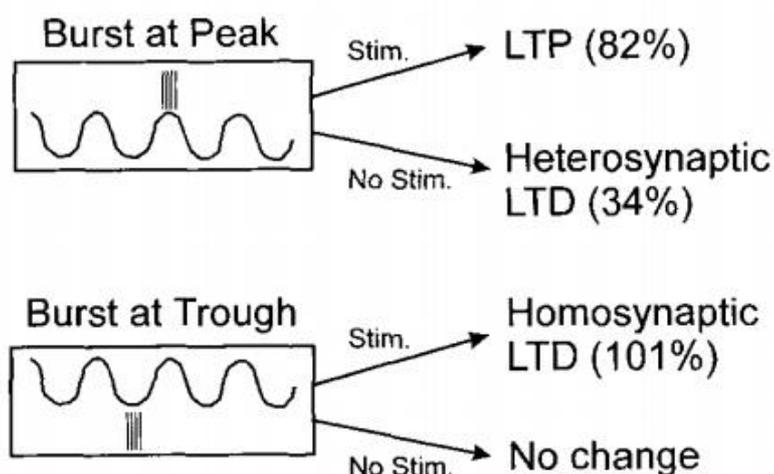


Figure 26 – Theta phase determines plasticity. LTP or LTD is induced when a burst of spikes is introduced during opposite cycles of an induced theta oscillation. Homosynaptic refers to synaptic change within the observed assembly, heterosynaptic refers to change in other outgoing/incoming connections from the cell assembly. Figure adapted from Huerta & Lisman, 1995.

Neural plasticity has also been found to be modulated by the phase of hippocampal theta oscillations. Initially, it was found that after a priming burst of spikes, bursts following at 200ms intervals optimally induced LTP. However, when theta oscillations were cholinergically induced, it could be further demonstrated that a burst applied to CA1 neurons at the peak of theta induced LTP, whilst a burst at the trough reversed the effect (Figure 26; Huerta & Lisman, 1995).

4.2 OSCILLATIONS & MEMORY

Since the early days of psychology and neuroscience, human memory has been one of the most observable and interesting fields of study (McClelland, et al., 1995; O'Reilly, et al., 2011; Rumelhart, et al., 1986; Tulving & Donaldson, 1972). This is equally true for the relatively nascent oscillations literature, where there have been a number of neural network models that theorise on the functional effects oscillations may have on memory (Hasselmo, et al., 2002; Jensen, et al., 1996; Ketz, et al., 2013; Norman, et al., 2006). This section will overview some of these models, taking inspiration from some of the ideas and methods presented here for the first modelling work of this thesis (Parish, et al., 2018), described in the next Chapter. These include the popular theta/gamma phase coupling model (Jensen, et al., 1996), where theta mediates gamma in a slot-like working memory, as well as theta phase learning models that enable phase dependent learning (Ketz, et al., 2013) and a winner-take-all mechanism within the hippocampus (Norman, et al., 2006).

4.2.1 AUTOASSOCIATIVE MEMORY IN NETWORKS WITH THETA/GAMMA OSCILLATIONS

A popular model captures the theta/gamma phase-amplitude coupling that occurs during working memory formation (Jensen, et al., 1996), theorising that a gamma frequency of working memory slots is maintained in sequential order every theta cycle. See Appendix A for more information that supports this notion, in particular the observed findings of gamma phase precession in place cells (O'Keefe & Reece, 1993) and theta phase-dependent learning (Huerta & Lisman, 1995).

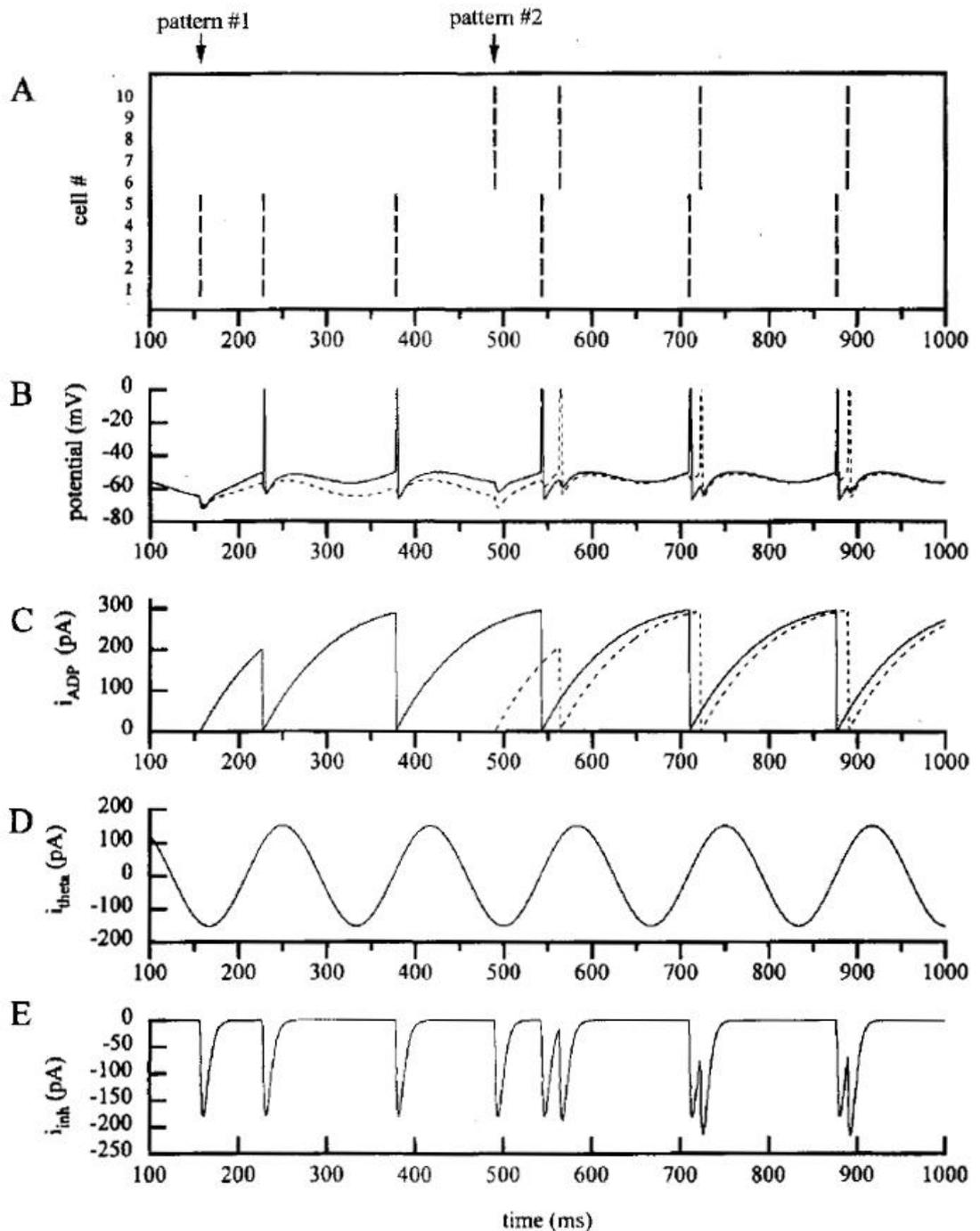


Figure 27 – **A** pattern representations; **B** Membrane potential of simulated neurons (solid – pattern #1, dotted – pattern #2); **C** After-De-Polarisation (ADP) function; **D** Theta AC current generator; **E** Inhibitory spikes from local interneuron. Figure re-represented from Jensen, et al., 1996.

In this model, a unique pattern made up of 5 simulated cells (Figure 27A) is maintained by a specific gamma slot within each theta up-state. This is achieved using an after-de-polarisation (ADP) function (Figure 27C), which is an exponentially increasing constant input that reaches a maximum

value after the length of a theta cycle ($\sim 250\text{ms}$). This input resets to zero after the target neuron spikes, ensuring that the maximal input entering the neuron will not occur until exactly one theta cycle later. There is evidence for a biophysical ADP function in hippocampal and cortical neurons (Andrade, 1991; Caesar, et al., 1993; Libri, et al., 1994).

Every excitatory cell is also linked to a single inhibitory interneuron, maintaining proportional excitation-inhibition balance (Markram, et al., 2004), such that the whole population is inhibited for a single gamma cycle after each spike event (Figure 27E). This ensures that multiple patterns cannot activate in the same gamma cycle and maintains the gamma frequency separation between pattern activations. To maintain theta cycles, an AC generator of 4Hz is fed into every excitatory cell (Figure 27D; see Figure 25 for example method).

This model is also interested in the presence of two types of *N*-methyl-D-aspartate (NMDA) synapse channels, categorised as fast NMDA and slow NMDA channels. They argue that different types of channels are more conducive to different types of learning. If NMDA channels have a fast de-activation time (Figure 28B1) that fits into the span of one gamma cycle, then LTP will only occur between the neurons coding for a single item (Figure 28B2) – forming the auto-associative memory that forms and maintains individual patterns. If NMDA channels have a slow de-activation that spans a theta cycle (Figure 28C1), then LTP would occur between neurons coding for many patterns (Figure 28C2) – forming the hetero-associative memory that is necessary for encoding the serial order of events. The hippocampus is mainly made up of slow NMDA receptors, therefore it is more inclined towards forming serial memory, whereas the cortex is made up of both types of receptors, indicating it is more conducive to forming a distributed memory store (Jensen, et al., 1996).

This very elegant model theorises a functional use for experimentally observed theta/gamma phase coupling in memory processes (Fell & Axmacher, 2011). Importantly, it showcases a proof of concept and not an attempt to create a dynamic functional memory store. Of interest to the later modelling work of this thesis is the use of an ADP function to control the rate of hippocampal

activity, as well as the idea of theta phase-dependent learning, highlighted in the slow NMDA synaptic modification in Figure 28C1-C2 (Jensen, et al., 1996). The latter is further explored in the next sections, which explore the role that hippocampal theta-phase might play in learning and memory (Hasselmo, 2005; Huerta & Lisman, 1995; Pavlides, et al., 1988) through the use of computational neural networks (Ketz, et al., 2013; Norman, et al., 2006).

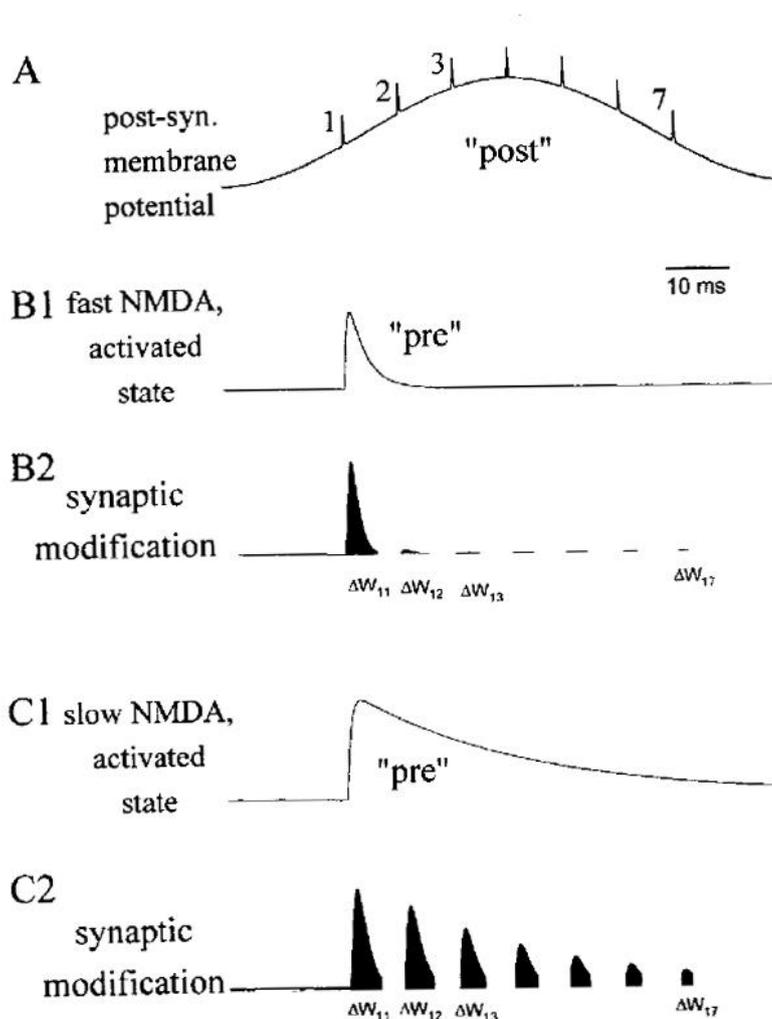


Figure 28 – **A** Post-synaptic membrane potentials of sequentially ordered patterns; **B1** NMDA channel that decays quickly after pre-synaptic spike event; **B2** Synaptic modification within cells encoding for pattern #1 (ΔW_{11}) due to fast NMDA channels; **C1** NMDA channel that decays slowly after pre-synaptic spike event; **C2** Synaptic modification from pattern #1 to other sequentially activated patterns (ΔW_{1-}) due to slow NMDA channels. Figure re-represented from Jensen, et al., 1996.

4.2.2 HOW INHIBITORY OSCILLATIONS CAN TRAIN NEURAL NETWORKS

Considering the role of theta oscillations in human memory, one neural network model represented a proxy for a human memory store by instantiating many overlapping patterns in a rate-coded, Hebbian learning model (Norman, et al., 2006). The aim of this model was to show that in an environment of competing, overlapping representations, it is necessary to strengthen target memories whilst also weakening competing ones when attempting to recall specific memories. In this model, the finding that the hippocampus employs theta oscillations to segregate plasticity (Huerta & Lisman, 1995) inspired an oscillatory learning mechanism (Figure 29). Here, high inhibition was used to identify the most active target memories, which are strengthened through LTP, whilst low inhibition was used to identify the least active target memories, which are further weakened through LTD. In this way, the oscillating algorithm was used for pattern separation, strengthening the most discernible features of each representation whilst weakening overlapping nodes. This notion is important for the first modelling work of this thesis, where we employ a similar phase-dependent learning rule that uses phase to strengthen concurrently active memories. Whilst we do not explicitly model it, we imagine that the pattern separation notion proposed here could work in tandem with the learning rule we define in the next Chapter.

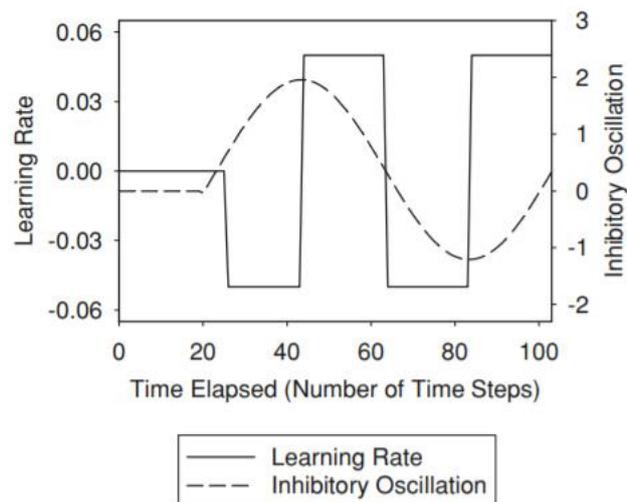


Figure 29 – Theta learning rule for punishing competitors. Learning rate increases when theta approaches its midpoint. Figure adapted from Norman, et al., 2006.

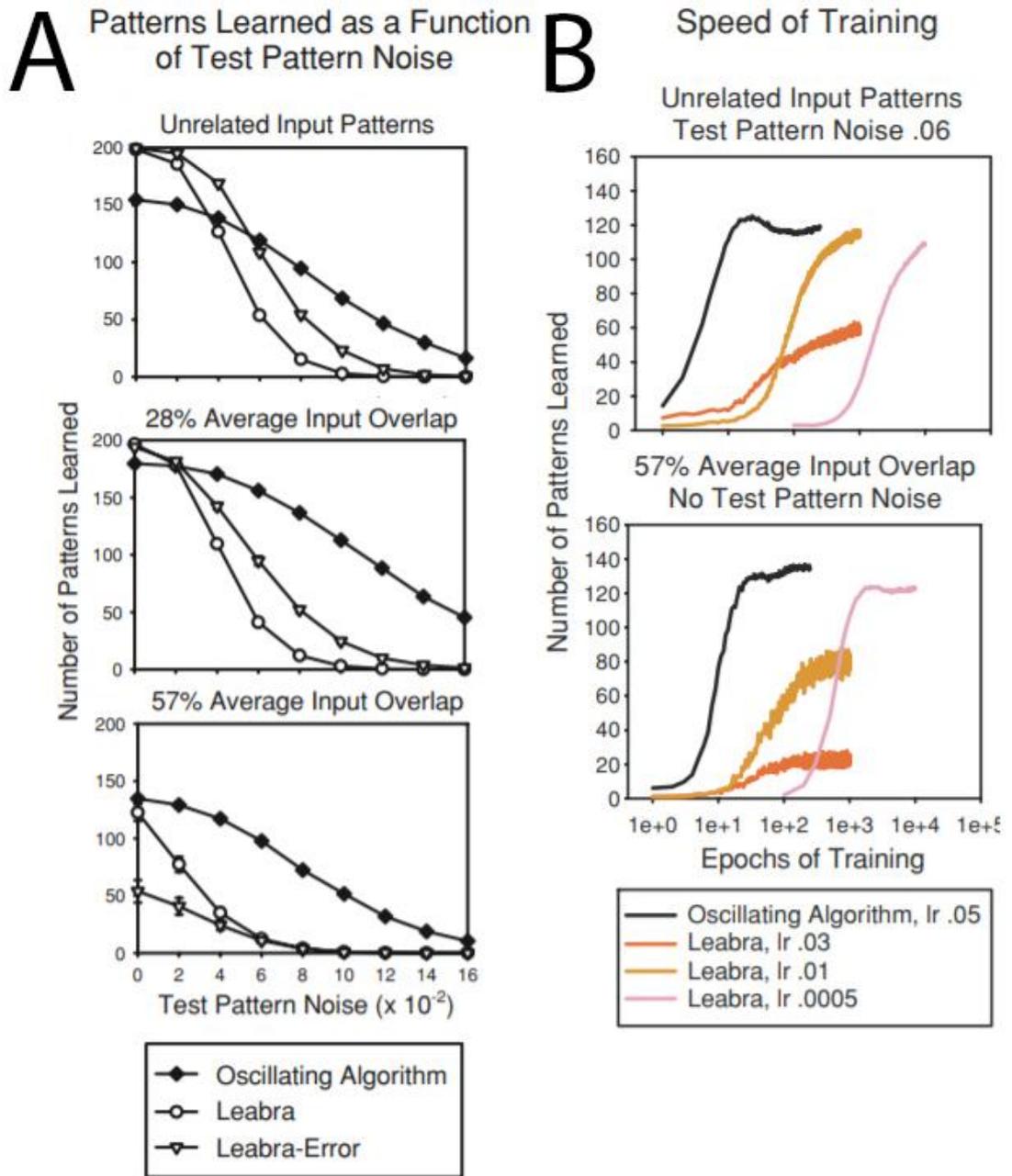


Figure 30 – Comparisons of theta-dependent learning algorithm with other Hebbian learning rules. Oscillating algorithm is better at pattern separation in noisier conditions (A) and can do so at a much faster rate than other algorithms (B). Figure adapted from Norman, et al., 2006.

Conceptually, it was shown that oscillations can be especially useful in identifying and separating patterns in an overlapping environment (Norman, et al., 2006). Figure 30A shows how the algorithm (described in Figure 29) compared to other standard and error-driven Hebbian learning rules (Hebb, 1949), where the former is superior in environments of increasing overlapping representations, as

is thought to exist in the hippocampus (O'Reilly, et al., 2011). The algorithm also performs better than others in noisy environments, indicating the role oscillations may play in facilitating communication and learning (Fell & Axmacher, 2011) in a vast and distributed information store. The oscillation algorithm is also much faster than other standard learning rules, managing to encode all pattern separated memories in half the time of other error-driven models (Figure 30B).

The interesting take-away from this model is the ability of an oscillatory learning rule to identify and separate overlapping patterns by segregating learning (LTP) and forgetting (LTD) into opposite phases, a hypothesised functional role of theta in the hippocampus (Hasselmo, 2005). Unlike previous models (Jensen, et al., 1996), this dynamic mechanism is more than a proof of concept and can be readily applied to large training sets. The ideas described here will be taken forward into the first modelling work of this thesis (Parish, et al., 2018), where theta similarly regulates learning and target concepts must be pushed forward in phase to be learnt.

4.2.3 THETA COORDINATED ERROR-DRIVEN LEARNING

There are several coexisting theta fields within the hippocampus that exist at different phases (Buzsaki, 2002, see Appendix A for a review on the origin of the theta frequency). As in previous models (Norman, et al., 2006), where an oscillatory algorithm with distinct functional phases was used to support pattern separation, a subsequent model has shown that differential phase relationships can further support a form of error-driven learning (Ketz, et al., 2013). This is an important notion for the 1st modelling work this thesis (Parish, et al., 2018), which similarly uses phase to modulate learning. The issues addressed in this model directly relate theta-phase modulated learning to regions and pathways in the hippocampus, where it is theorised that learning is in fact dependent on the existence of several coexisting theta fields. Whilst the 1st model of this thesis employs a much simpler mechanism, we imagine it might be scaled up to account for the varying theta phases described here, and the functional significance they might have for pattern separation processes – an operation not considered for the modelling works of this thesis.

In the model, theta oscillations switch between stages of encoding and recall through preferential modulation of the entorhinal-CA3 tri-synaptic (TSP) pathway and the entorhinal-CA1 mono-synaptic (MSP) pathway, respectively (Figure 31A). As discussed in the previous Chapter, representations in the entorhinal cortex (EC) project to their respective sparse version in hippocampal CA1 region via MSP, whilst also projecting to the dentate gyrus (DG) and the CA3 region via TSP (O'Reilly, et al., 2011; Squire, et al., 2004). Recurrent connections allow the CA3 region to pattern complete from incomplete or multiple incoming patterns (O'Reilly, et al., 2011), with the help of the DG to contribute sparse coding and avoid confusion from overlapping representations. Highly plastic connections in the Schaffer collaterals (from CA3 to CA1 region) then perform the main function of episodic memory, creating a sparse, conjunctive representation of EC inputs, where active representations are highlighted in orange in Figure 31B (Ketz, et al., 2013).

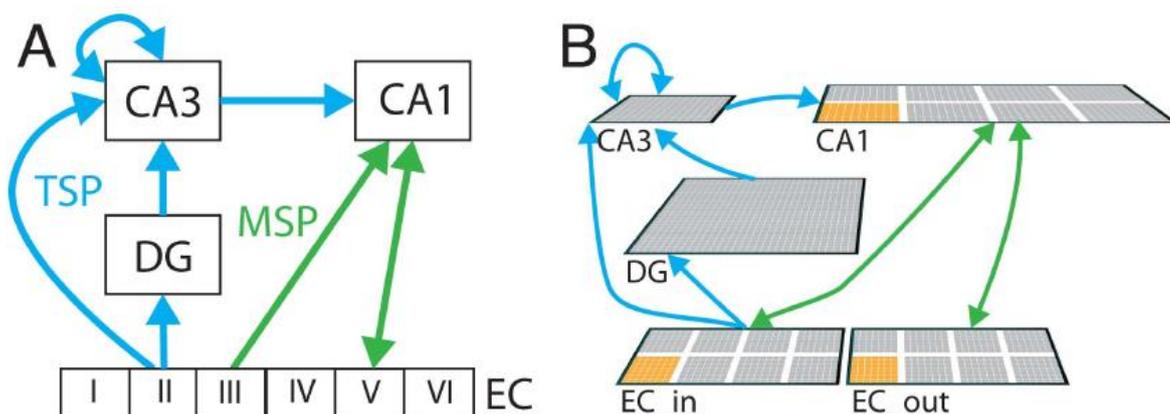


Figure 31 – **A** Schematic representation of hippocampal and entorhinal cortex connectivity. **B** Rate coded model implementation, where tri-synaptic pathways are highlighted in blue and mono-synaptic pathways are highlighted in green. Figure re-represented from Ketz, et al., 2013.

It is commonly accepted that the CA1 region projects back to outgoing layers of EC (Figure 31A; CA1 to EC-V via MSP), where the EC is thought to propagate the signal to neo-cortical representations and re-instantiate newly associated memories during recall (O'Reilly, et al., 2011; Squire, et al., 2004). Here, additional connections from CA1 to incoming layers of EC support back-propagated signal generation (Figure 31B; CA1 to EC-III via MSP), that forms the basis of error correction in the

model (Ketz, et al., 2013). It is hypothesised here that the difference between original EC activity and CA1 sparse representations generates further DG/CA3 activity via TSP, which will then continue to perform pattern completion until the error is reduced. This model further suggests a functional role for theta oscillations in this process (Ketz, et al., 2013), inspired by the finding that distinct hippocampal regions are active at various theta phases (Buzsaki, 2002). In the model, theta phase is used to modulate between TSP pattern completion, i.e. encoding, and MSP signal representation and comparison, i.e. recall. Here, the reference theta source is recorded from the hippocampal fissure, as depicted by the red theta oscillation in Figure 32A (between DG/CA1 regions, where the intra-hippocampal current generation is strongest; Buzsaki, 2002, see Appendix A for more information on theta rhythm and current generation). The CA3 region is thought to be the main rhythm contributor to this source (Buzsaki, 2002), and so it oscillates in phase as can be seen by the blue histogram of CA3 neural activity in Figure 32A. The EC region is thought to oscillate at an alternate phase to this reference theta source (Buzsaki, 2002), as can be seen in the green histogram of EC activity in Figure 32A. Phase synchrony then enables communication through shared windows of activation (Fell & Axmacher, 2011, see Appendix A for more on phase synchrony and neural communication), meaning that the CA1 region switches between communication with the EC and CA3 regions in alternating phases of theta (Ketz, et al., 2013).

Figure 32B shows active neural populations are associated together through an entire theta phase (Ketz, et al., 2013). The trough of fissure recorded theta (Figure 32B, left panel) enables MSP to perform retrieval processes, where representations are forwarded through the hippocampus via EC-III to CA1 and back via CA1 to EC-V (Figure 31A; MSP), re-instantiating representations that are distributed to further cortical sites (O'Reilly, et al., 2011). Plasticity here strengthens active EC representations with their respective sparse CA1 versions (Ketz, et al., 2013). During the peak of fissure recorded theta, activity in the CA3 and DG regions allows TSP to perform pattern completion from concurrently active EC representations (Figure 32B, middle panel). Phase synchrony then enables plasticity in the Schaffer collateral (CA3 to CA1 synapses), acting to create a novel

conjunctive representation in the CA1 region, as theorised during the previous Chapter on episodic memory (Ison, et al., 2015; Quiroga, 2012).

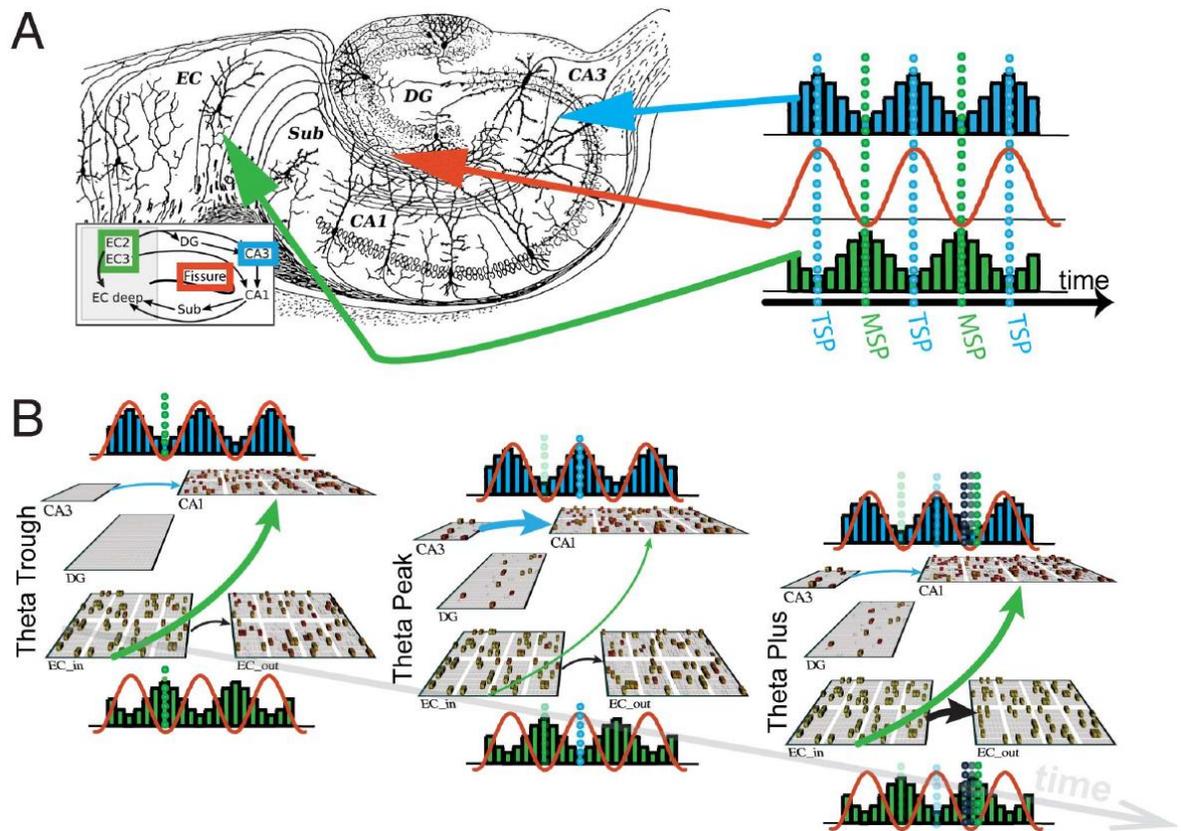


Figure 32 – **A** Hippocampal anatomy and the source of theta rhythm generators from CA3 (blue histogram) and EC (green histogram), as well as the hippocampal fissure source of recorded theta (red). **B** Activity of EC/CA1/CA3 through fissure recorded theta, where plasticity between regions depends on phase synchrony of regions. Figure re-represented from Ketz, et al., 2013.

As theta moves back towards the fissure recorded trough in a theta ‘plus’ stage (Figure 32B, right panel), the CA1 region settles into a conjunctive representation between target encoding (CA3 to CA1) and attempted recall (EC to CA1), generating an error-signal of differential representations that stimulates further encoding through CA1 to EC-III and further CA3/DG processing via TSP. The model thus builds on the previous notion of the functional use of oscillations in learning (Norman, et al., 2006), indicating that communication through phase synchrony between distinct regions can be facilitated by phasic differences (Ketz, et al., 2013).

4.3 OSCILLATIONS & INFORMATION PROCESSING

As discussed previously in our section on oscillation functionality, tonic and de-synchronised activity reflects states of high excitation, whilst synchronised activity has been argued to reflect states of inhibition (Fell & Axmacher, 2011). Thus, inhibitory states enable certain cognitive top-down control processes that limit activation through rhythmic entrainment (Buzsaki, 2002; Klimesch, et al., 2007). In this respect, oscillations act as an inhibitory filter that achieves a high signal-to-noise ratio by selectively allowing subsets of cells to process information whilst simultaneously silencing the majority of non-relevant cells (Klimesch, et al., 2007; Hanslmayr, et al., 2012). The previous section described this process operating in the theta and gamma frequencies (Buzsaki, 2002; Fell & Axmacher, 2011), acting to gate learning processes (see Appendix A for more on the hypothesised relationship of theta/gamma with learning). This phenomenon has also been observed many times in the alpha frequency ranges during attentional paradigms (Adrian & Matthews, 1934; Berger, 1929; Haegens, et al., 2011; Hanslmayr, et al., 2011b; Jensen & Mazaheri, 2010; Klimesch, et al., 2007), which will be described in the following section. See Appendix A for an overview on the origins of the mid-range frequencies that are typically involved in attentional processes, specifically alpha and beta, which have been observed since the early days of EEG research (Berger, 1929).

This section begins by describing the information via desynchronisation hypothesis (Hanslmayr, et al., 2012), where alpha oscillations are related to information theory and desynchronisations of alpha can be used to predict successful memory recall. An overview of the inhibition-timing hypothesis will then be given (Klimesch, et al., 2007), that describes how attention might be gated through the top-down inhibitory control mechanism that alpha provides. There is also a qualitative element to the desynchronisations of alpha that signify information flow. Typically, the phase of alpha is reset during each desynchronising period (Canavier, 2015), as neurons are triggered to fire irregularly by external or internal events. This section will thus conclude by summarising a recent paper (Michelmann, et al., 2016), where a stimulus can be successfully identified based on the unique temporal signature in this phase-reset pattern.

Whilst there has been substantial effort to simulate the alpha dynamics of a single cortical column and the way in which phase synchrony might occur through interactions with an intermediary assembly (Jones, et al., 2000; Vicente, et al., 2008; Wang, 2010), there has not been much focus on the theoretical relation between alpha and attention and memory processes (Hanslmayr, et al., 2016). This will be one of the founding principles of this thesis, whereby alpha desynchronisations will signify qualitative information flow in both models described in this thesis (Parish, et al., 2018).

4.3.1 INFORMATION VIS DESYNCHRONISATION

Event-related-desynchronisation (ERD) has been noted since the early days of EEG research, whereby opening one's eyes decreases observed alpha amplitudes (Berger, 1929). Once thought to be a bottom-up process of light stimulation, this effect has also been observed when participants are in darkness (Adrian & Matthews, 1934), suggesting that alpha is suppressed through a top-down mechanism. Alpha suppression over frontal brain areas also occurs during various other tasks such as semantic processing, as can be seen in the ERD during a retention task below (Figure 33b; blue line, >0.25 seconds). Here, one can also observe the increase in theta and gamma power that is typical in memory tasks (Burke, et al., 2014; Fell & Axmacher, 2011). ERD, calculated as the percentage of band power change during a task compared to some reference point, suggests synchrony is lost due to large populations of neurons beginning to fire asynchronously (Fell & Axmacher, 2011; Klimesch, et al., 2007; Hanslmayr, et al., 2012).

Low frequency (7-10Hz) alpha ERD is topographically widespread and is probably due to attentional demands (Hanslmayr, et al., 2011b; Klimesch, et al., 2007), whilst upper alpha (10-13.5Hz) and beta (13.5-40Hz) ERDs are restricted to local brain regions during semantic processing (Hanslmayr, et al., 2009; Klimesch, et al., 2005; Meeuwissen, et al., 2011). The time-course of stimulus locked ERD depends on the type of task, frequency band and stimulation, but typically shows a late onset post-stimulus of ~200ms, a de-synch peak around 350-650ms and a resynch peak around 900-2000ms (Klimesch, et al., 2007). A small ERD peak, typically resulting

from poor task performance, has a short latency and is followed by an early re-synchronisation, whereas a large ERD can last for longer and is followed by an even later re-synchronisation (Klimesch, et al., 2007). These features are summed up in Figure 33 (Hanslmayr, et al., 2011b), where a stimulus presented at 0 seconds causes an alpha ERD in the time-frequency-analysis (Figure 33A), also observable in the diminished LFP signal at 10Hz (Figure 33B). During this time, a theta (4Hz) and gamma (40Hz) synchronisation is also observable, indicating communication with other leading brain regions is underway (Fell & Axmacher, 2011).

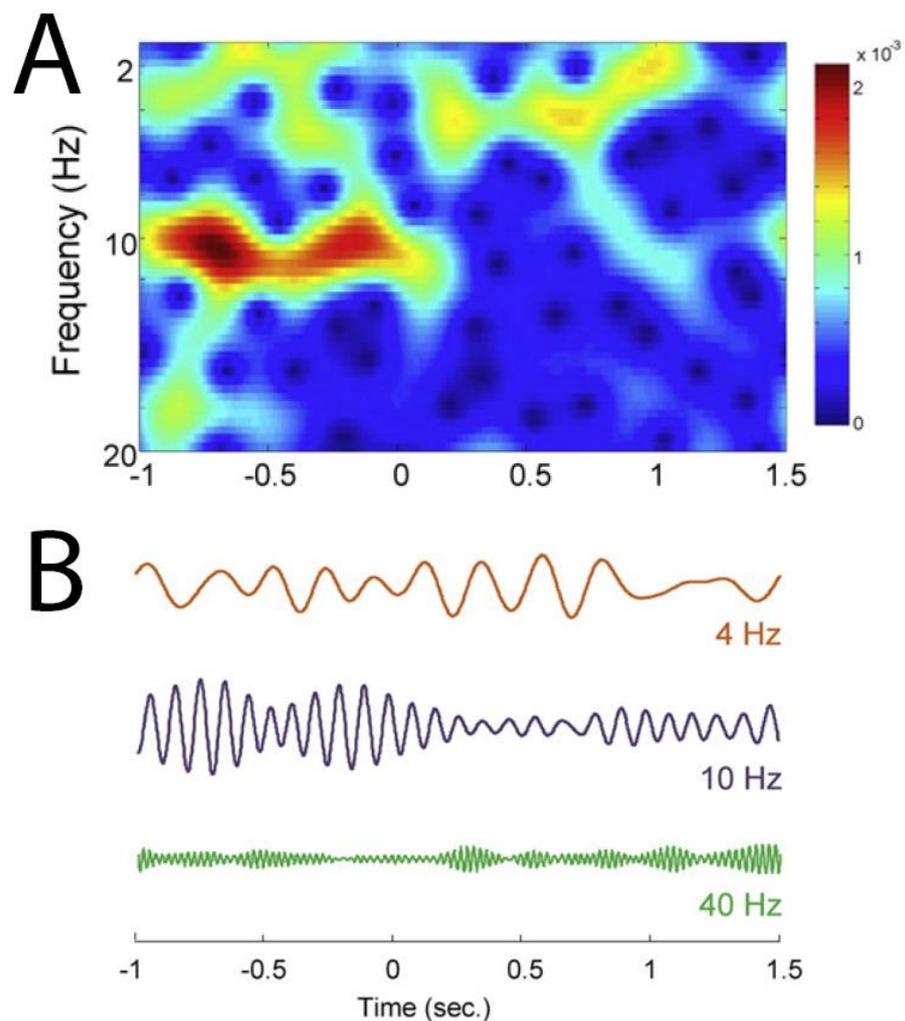


Figure 33 – Signal from the parietal EEG electrode during the presentation of a stimulus at $t = 0$ seconds. **A** Time frequency analysis of the signal. **B** Local field potential of various chosen frequencies. Figure adapted from Hanslmayr, et al., 2011b.

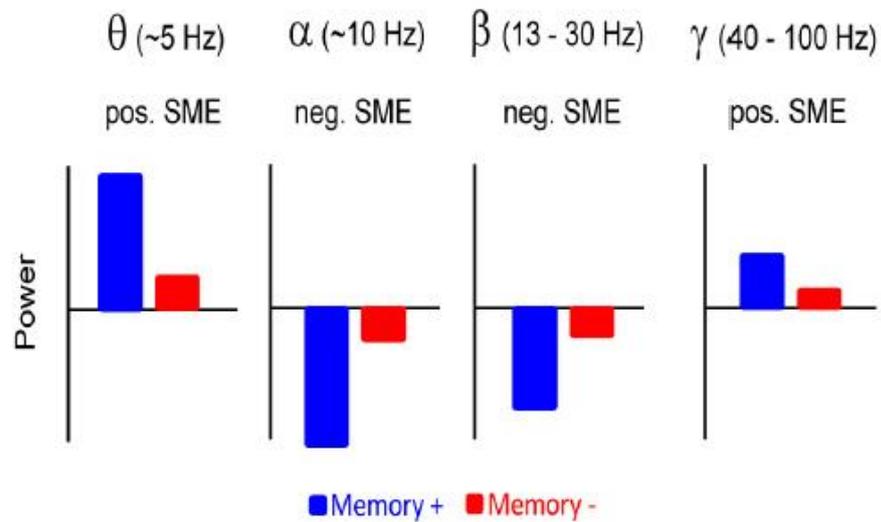


Figure 34 – Change in frequency power relative to stimulus being remembered, where blue indicates a successfully recalled memory and red a failure. Adapted from Hanslmayr, et al., 2012.

An ERD of cortical regions, specifically in the alpha and beta ranges, is also strongly correlated with long term memory formation and retrieval (see Hanslmayr, et al., 2012, for review). This was in part found by studies analysing successful memory encoding, whereby subsequent memory (SM) and difference in memory (DM) effects are used to compare between remembered and forgotten items. These differences are then associated with the concurrently occurring oscillatory dynamics, referred to as subsequent memory effects (SMEs), which can be positive or negative dependent on the direction of power change at encoding that successfully predicts memory retrieval. For example, positive SMEs are typically found in the theta and gamma frequencies (Fell & Axmacher, 2011), whereby an increase in power at encoding predicts successful retrieval, whereas negative SMEs are found in the alpha and beta frequencies (Hanslmayr, et al., 2012). These dual findings can be seen in Figure 34, where a blue bar indicates memories that are successfully recalled and a red those that are not. Furthermore, a simultaneous EEG/fMRI study has also demonstrated negative beta SME effects (Hanslmayr, et al., 2011a), this time using haemodynamics to indicate that high energy consumption can be correlated to the de-synchronisation of the on-going oscillation, corroborating previous studies that found low-energy consumption states are correlated with high alpha and beta power (Hanslmayr, et al., 2012).

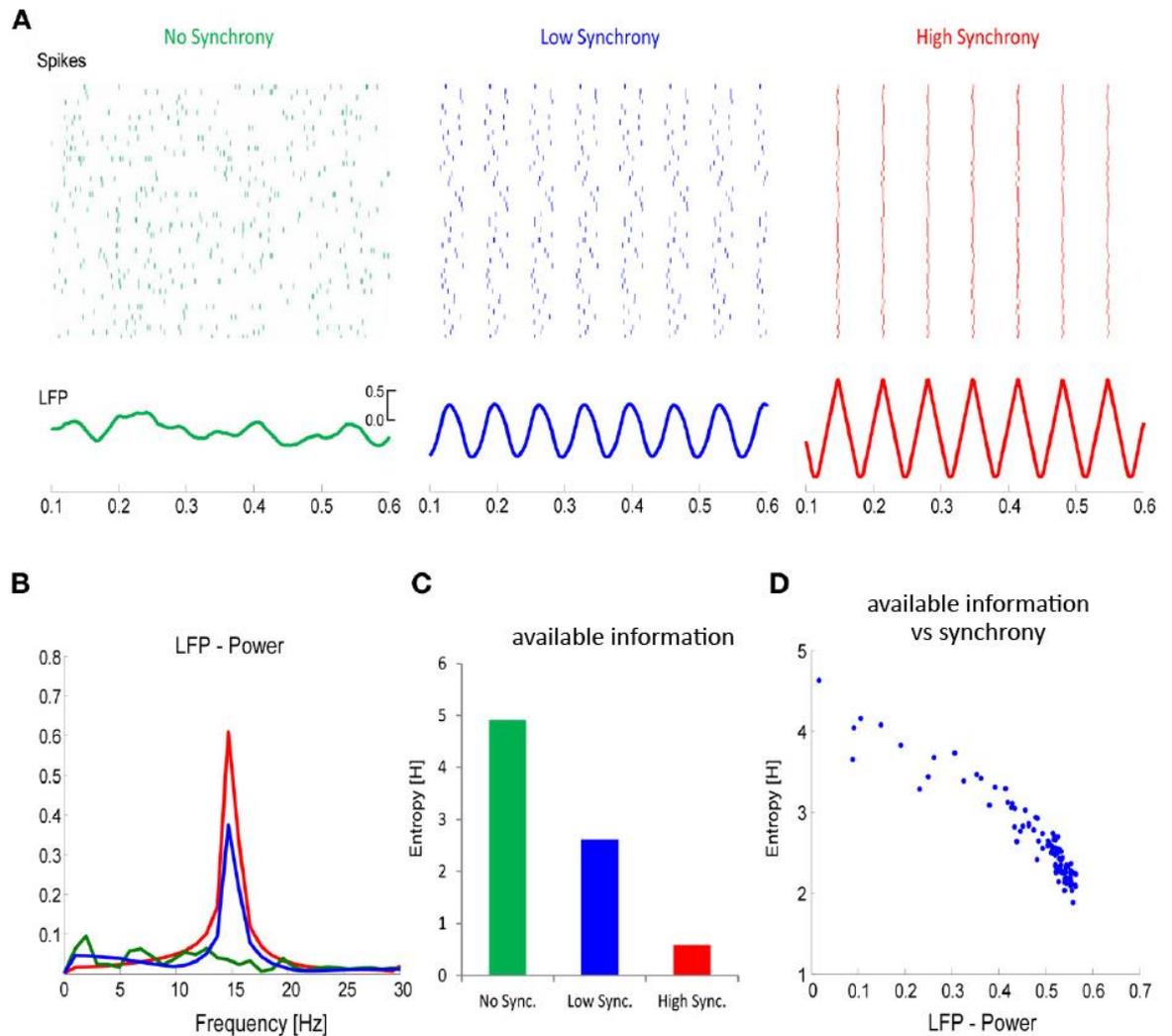


Figure 35 – **A** Raster plots of spike events and LFP signals for increasingly synchronous neural activity. **B** LFP power of increasingly synchronous activity. **C** Entropy for increasingly synchronous activity. **D** Entropy as a function of LFP power. Figure copied from Hanslmayr, et al., 2012.

The quality of information gained from the occurrence of events can be analysed with respect to entropy, which estimates the likelihood that an event can be predicted (Shannon & Weaver, 1949). When applying entropy to the activity of populations of neurons, one can surmise that oscillatory dynamics can be correlated with states of low entropy, whilst asynchronous neuronal firing can be correlated with states of high entropy, where entropy here predicts how much information can be gained from the occurrence of a spike event. If we consider various levels of synchronisation in neural populations (Figure 35A), we can plot LFP power (Figure 35B) as a function of entropy (Figure 35C) and see that they negatively correlate (Figure 35D; Hanslmayr, et al., 2012). From this analysis,

we can hypothesise that the desynchronisation of on-going oscillatory activity allows the once systemised neural subregion to increase its information gain (Hanslmayr, et al., 2012), as subsets of neurons separate themselves from baseline oscillatory behaviour to represent unique active features at a time when the majority are silent, thus providing more information to downstream target neurons (Fell & Axmacher, 2011). Alpha power is also negatively correlated with neural activity (Haegens, et al., 2011) and energy consumption (Hanslmayr, et al., 2011a), indicating that neurons increase in activity to overcome stable states of inhibition (Klimesch, et al., 2007), as hypothesised earlier (Figure 24). This is supported by the fact that cortical gamma increases also predict successful memory encoding (Sederberg, et al., 2007), indicating that cell assemblies shift into a higher-paced frequency with more activation when actively representing information.

4.3.2 THE INHIBITION-TIMING-HYPOTHESIS

Conversely, an event-related-synchronisation (ERS) of alpha will reflect a state of low-information processing, typically observed in idling or resting states (Adrian & Matthews, 1934). Additionally, upper alpha (10-13.5 Hz) ERS is thought to be involved when attentional resources must be managed, such as when a learned response must be withheld and over brain regions that are non-task relevant (Klimesch, et al., 2007). With respect to the former, alpha ERS has been observed when participants are instructed to remember a list of items and withhold responses until a probe is presented, where ERS increases relative to the number of items to be remembered (Sauseng, et al., 2005a; Sauseng, et al., 2005b). The same effect has also been observed when participants withhold motor actions (Klimesch, et al., 2007; Muthukumaraswamy, et al., 2004), where whilst focal ERD was observed in task-relevant areas, surround ERS was observed in non-relevant areas, suggesting that focal ERD reflects activation but surround ERS reflects inhibition of neighbouring cortical areas (Suffczynski, et al., 2001).

Rather interestingly, the ERS/ERD process can even be observed in real-time during memory retrieval tasks, where surround ERS is high when participants are searching for a memory trace and

ERD occurs the moment that the correct information is retrieved (Sauseng, et al., 2005b). Similarly, when participants had to manipulate information during working memory paradigms (Figure 36; Sauseng, et al., 2005b), ERS occurred at frontal sites where competing items were being suppressed (red), whilst ERD occurred at posterior sites where retained items were being visualised (blue). This interaction can be seen in EEG as a travelling alpha wave where frontal sites lead posterior sites, whereas gamma travelling waves can also be seen in the reverse direction. This indicates that alpha is involved with the top-down semantic encoding of familiar stimuli (Sauseng, et al., 2005b), where novel stimuli can be compared to pre-existing memories, whereas gamma is perhaps involved in the bottom-up activation of sensory processes, enabling separate attributes of a stimulus to be concurrently represented (Fell & Axmacher, 2011; Klimesch, et al., 2007; Sauseng, et al., 2005b).

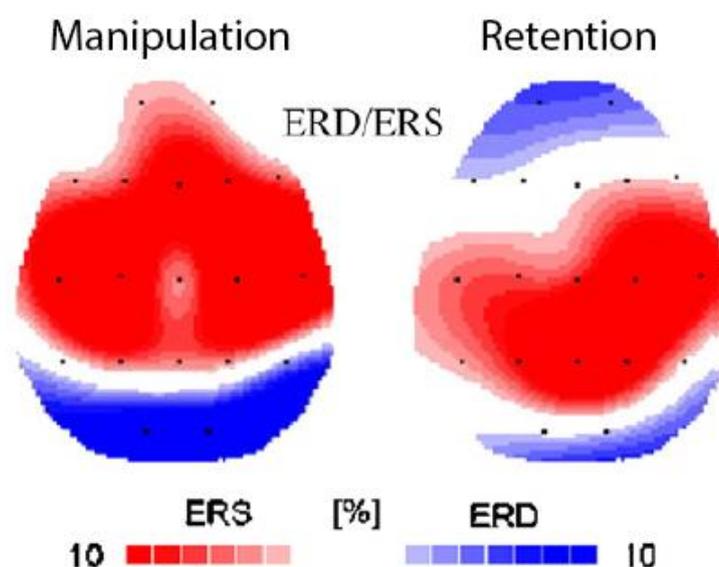


Figure 36 – Event related desynchronisation (ERD, blue) and synchronisation (ERS, red) during a mental manipulation and information retention task. Figure adapted from Sauseng, et al., 2005b.

It has even been suggested that intelligence is positively correlated with ERS over task-irrelevant areas (Doppelmayr, et al., 2005), in particular the ability to switch between long-term and working memory dependent on the task at hand. This suggests that intelligence, specifically the ability to sift through memories and cognitive tasks, depends on focused activation of task-specific regions through high ERD, whilst concurrently suppressing irrelevant pathways through high ERS

(Doppelmayr, et al., 2005). This idea is exemplified by further findings that ERD is exhibited over sensory-motor (observation) and ERS is exhibited over supplementary-motor (action) regions, when actions are internally visualised via mirror neurons but resulting movements are suppressed, during observations of complex motor tasks (Muthukumaraswamy, et al., 2004). Here, one might relate intelligence to the ability to control one's thoughts and actions, facilitated by alpha waves.

An important note is that pre-stimulus and surrounding oscillatory power may also be a key predictor in memory and attentional paradigms, providing a possible explanation for some of the variation between studies when measuring differences between baseline and post-stimulus activity (Hanslmayr, et al., 2012). For example, a large resting or reference alpha amplitude has been found to have a positive correlation with perception performance (Knight & Marica, 1995). This implies that perceptual performance is increased if the cortex is already activated, indicating that the brain can better match patterns if there is pre-existing rhythmic activation, where inputs require little extra activation to pattern complete and cause a subsequent ERD (Klimesch, et al., 2007). This is similar to our description of common methods for simulating background noise, described in Chapter 2, where neurons were kept just below threshold in a ready-to-fire state (Abeles, 1991; Dayan & Abbot, 2001). On the other hand, memory performance is increased if pre-stimulus alpha power is low (Hanslmayr, et al., 2012). This implies that de-activating the cortex beforehand enables the competitive process of re-instantiating a memory trace, where high levels of alpha might create too much inhibition for a self-completing pattern to overcome (Klimesch, et al., 2007).

In a similar fashion, the phase of on-going alpha also impacts behavioural measures (Klimesch, et al., 2007). For example, reaction times are reduced if stimuli are presented at alpha peaks, indicating that alpha phase mediates cortical excitability in an oscillatory manner, thus aiding pattern completion if the cortex is at its maximally active phase (Klimesch, et al., 2007). An event may also drive an alpha phase reset (Canavier, 2015; Fell & Axmacher, 2011), as alpha is reset to the maxima and the brain is put into a state of maximal awareness. The resulting P1 component, a signature of sensory processing in EEG recordings (Spehlmann, 1965), might be the earliest

manifestation of top-down processing as alpha oscillations re-set in phase to align with the timing of the stimuli, thus enabling all consecutive operations to function with typical millisecond precision (Hanslmayr, et al., 2011b). This reset in turn causes event-related alpha synchronisation between communicating brain regions, as the rest of the brain re-aligns with the now reset attentional mechanisms, whereby the lag to re-align in phase can even be correlated to distance from the location of the initial reset (Klimesch, et al., 2007). There is also a qualitative component hidden in the phase-reset patterns of alpha ERDs, where one can detect when a participant is internally visualising a dynamic stimulus (Michelmann, et al., 2016). This indicates that oscillations carry predictable, qualitative information as disparate neuronal assemblies concurrently represent elements of a stimulus (Canavier, 2015; Fell & Axmacher, 2011). We will return to this last finding in a later Chapter when introducing the second modelling work of this thesis.

4.4 SUMMARY

Whilst in general, oscillations seem to act to enable neural synchronisation and control information flow (Fell & Axmacher, 2011), this is most true for the neo-cortex. Here, it is thought that alpha frequencies represent a top-down control mechanism that allows specifically activated cell assemblies to maintain clarity in your consciousness by inhibiting all non-relevant areas (Klimesch, et al., 2007). A desynchronisation in alpha and beta frequencies is strongly correlated to activation (Haegens, et al., 2011; Hanslmayr, et al., 2011a) and information flow (Hanslmayr, et al., 2012), and along with a synchronisation in theta & gamma frequencies (Burke, et al., 2014; Khader, et al., 2010), acts as a good predictor for the successful encoding of memories and subsequent memory recall (Hanslmayr, et al., 2012). Conversely, a synchronisation in alpha frequencies is strongly correlated to the suppression of irrelevant brain regions (Sauseng, et al., 2005a; Sauseng, et al., 2005b), where the more intelligent one is, the larger the amplitude of alpha frequencies during working memory paradigms (Doppelmayr, et al., 2005). Alpha event-related-desynchronisations (ERDs) also convey qualitative information through unique phase-reset patterns (Canavier, 2015). During the presentation of dynamic stimuli, one can compare the phase angle time-series at

encoding and retrieval to detect the temporal signature of specific stimuli (Michelmann, et al., 2016). These findings help us to formulate the primary hypothesis of this thesis, which aims to provide computational support for the information-via-desynchronisation and the inhibition-timing hypothesis. We do this by creating neural network models that include a slow learning neo-cortical system, where information flow is regulated by an alpha rhythm.

Whilst there are several theoretical models on the creation of dynamic alpha oscillatory behaviour through recurrent connectivity or modulated activity (Jones, et al., 2000; Vicente, et al., 2008; Wang, 2010), there is a distinct lack of computational models on the role alpha oscillations play in cognitive processes such as memory and attention (Hanslmayr, et al., 2016). This leads to the first modelling work of this thesis (Parish, et al., 2018), which creates a neo-cortical model of alpha oscillations in order to predict memory performance, described in the next Chapter.

Theta oscillations are prevalent in memory processes, essential for the encoding and retrieval of episodic information (Backus, et al., 2016; Burke, et al., 2014; Buzsaki, 2002; Fell & Axmacher, 2011; Lega, et al., 2012), which originates from medial-temporal-lobe (MTL) structures (Buzsaki, 2002; Squire, et al., 2004). Gamma frequencies have also been found to successfully predict memory formation (Sederberg, et al., 2007), possibly by using a theta-gamma phase coupling to segregate information (Heusser, et al., 2016; Jensen, et al., 1996). Whilst this paints a rather complex picture of the interacting oscillations that enable learning and information storage in human memory (Fell & Axmacher, 2011), the modelling work of this thesis focuses predominantly on the much-studied theta frequency. This is because this frequency has been shown to act as the main driver for human learning (Buzsaki, 2002), acting as a top-down control mechanism that entrains areas of the brain into distinct phases of plasticity (Hasselmo, 2005; Huerta & Lisman, 1995).

For this reason, theta frequencies have been the main enabler for several popular theoretical models in the field. One such model theorises how a number of gamma slots organises sequential information in a theta-phase modulated plasticity mechanism (Jensen, et al., 1996). Theta-phase

modulation was also shown to be superior to alternate learning rules when encoding patterns in a noisy environment, whereby target memories can be strengthened and competing memories weakened in opposite phases of plasticity (Norman, et al., 2006). By allowing for the various theta-phases that exist in neighbouring hippocampal assemblies (Buzsaki, 2002), one can introduce error-driven learning into such a model (Ketz, et al., 2013), potentially increasing the computational power to segregate memories in an overlapping and noisy environment. These models each describe neural network frameworks that theorise on the functional use of such experimentally observed frequency interactions in human memory processes. The first modelling work of this thesis (described in the next Chapter) takes inspiration from some of the methods described in this Chapter in furthering this aspiration. Here, a model (Parish, et al., 2018) describes the functional importance of theta phase-dependent learning and theta phase-synchrony in much the same way, thus hoping to provide yet more computational evidence for this theory in line with the secondary hypothesis of this thesis.

As we have discussed in Chapter 4, brain oscillations, via their ability to synchronize and desynchronize neuronal populations (Fell & Axmacher, 2011), play a crucial role in the formation and retrieval of episodic memories. However, little is known about how oscillations implement the necessary mechanisms for encoding and retrieval of such memories. Despite several models that theorise oscillatory mechanisms (Jones, et al., 2000; Wang, 2010) and functionality (Jensen, et al., 1996; Ketz, et al., 2013; Norman, et al., 2006), this knowledge gap is partly due to a lack of computational models that simulate oscillatory behaviours as observed in human EEG/MEG recordings during memory tasks (Hanslmayr, et al., 2016). The link between oscillations and memory is further complicated by empirical data, which has fuelled a conundrum as to how oscillations relate to memory. As we have shown in the preceding Chapter, hippocampal theta (~3-8 Hz) and gamma (~40-80 Hz) synchronisation (Fell & Axmacher, 2011) and the de-synchronisation of alpha and beta (8-30 Hz) in cortical regions (Hanslmayr, et al., 2012) have both been reported as important for memory encoding and retrieval. Indeed, even some researchers have found that a desynchronisation in hippocampal theta is key (Crespo-Garcia, et al., 2016; Fellner, et al., 2016). This brings us to the first modelling work of this thesis, which describes a theoretical framework that encapsulates these disparate findings (Parish, et al., 2018). We build on previous theories of oscillations (Klimesch, et al., 2007) to show that with increasing levels of activation, oscillatory power increases until tonic activation overcomes inhibitory down states and power decreases. We show that for distinct neural structures, this journey of sync to desync is affected by internal parameter settings and stimulus strength, thus showing how various findings can be described by a single theoretical framework on oscillations and memory formation.

This Chapter begins to explore the synchronisation conundrum through the creation of a spiking neural network model, with the intention of showing that both alpha de-synchronisations and theta synchronisations may be necessary for memory encoding and retrieval. The mathematical methodology of the model and design decisions will be discussed in this Chapter.

5.1 SYNC, DESYNC AND MEMORY FORMATION

As discussed in Chapter 3, classic computational models theorise that hippocampal and neo-cortical regions offer functionally distinct mechanisms to form episodic memory (O'Reilly, et al., 2011; McClelland, et al., 1995), where a sparse coded hippocampus learns new information quickly and a distributed neo-cortex incorporates this information slowly. This is a hypothetical solution to the catastrophic interference problem (McCloskey & Cohen, 1989), whereby newly acquired information erases established memories.

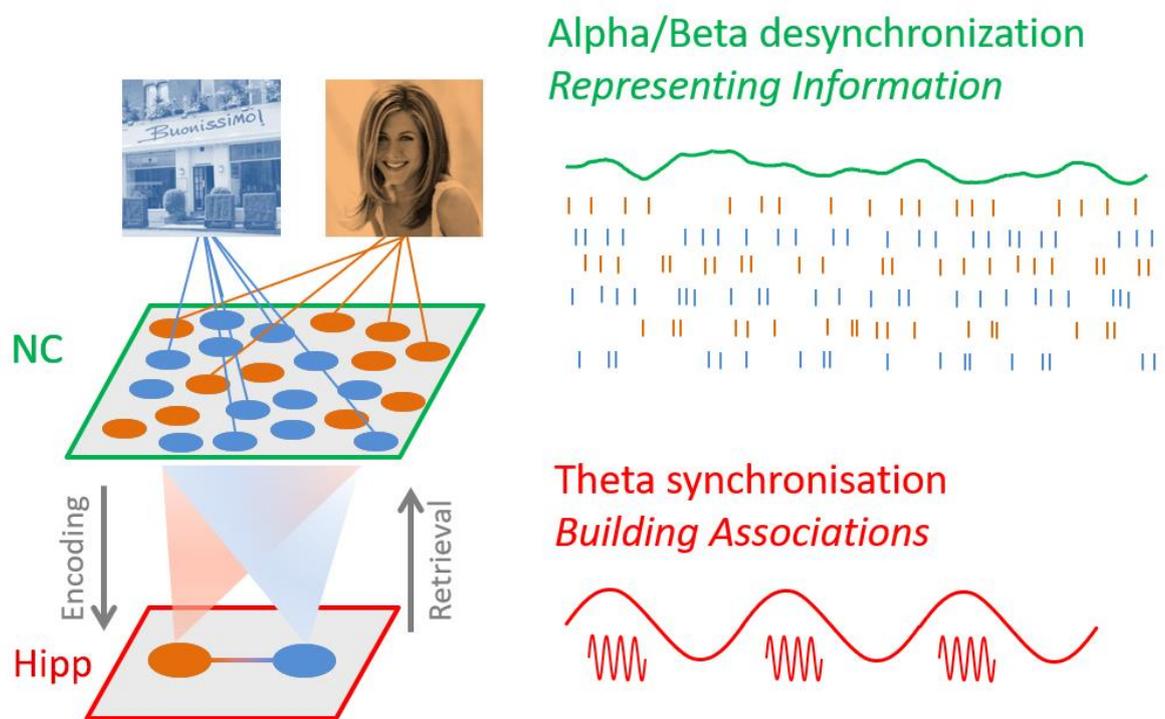


Figure 37 – The neo-cortex (NC) has been shown to desynchronise in alpha frequencies when stimuli are presented, representing regional activation. At the same time, the hippocampus (Hipp) has been shown to synchronise in theta frequencies during encoding and retrieval of memories. It is therefore theorised that alpha desynchronises to allow information flow in cortical regions. This in turn influences the hippocampus to synchronise in theta, which has been found to modulate phases of learning in the brain. Adapted from Hanslmayr, et al., 2016.

In Chapter 3, we described how projections, normally leading from neocortical to hippocampal sites in wakefulness, reverse during NREM sleep (Diekelmann & Born, 2010; Walker, 2018) such that the long-term memory store can incorporate a contextualised version of recent events from a shorter time-frame memory. In Chapter 4, we discussed the oscillations that are prevalent during these stages of memory consolidation (Buzsaki, 2002; Fell & Axmacher, 2011, see Appendix A for a further description). Here, delta waves appear to facilitate this stage of slow-wave-sleep, whilst theta and gamma synchrony were related to memory consolidation and recall. In contrast, alpha and beta waves appear to organise information flow, where synchronisations indicate the inhibition of task-irrelevant regions and desynchronisations indicate information flow. Building on these complementary learning systems, a recently presented theory to the synchronisation/de-synchronisation conundrum (Hanslmayr, et al., 2016) suggested that hippocampal theta synchronisation ($\sim 4\text{Hz}$) mediates the building of associations (Hasselmo, 2005; Huerta & Lisman, 1995), while neocortical alpha de-synchronisation ($\sim 10\text{Hz}$) is due to the representations of these concepts becoming active (Hanslmayr, et al., 2012; Klimesch, et al., 2007), as shown in Figure 37.

5.1.1 HIPPOCAMPAL THETA

Theta oscillations in medial temporal lobe (MTL) are assumed to play a key role in the formation of memories (Buzsaki, 2002), where learning is dependent on the power of observed theta oscillations and the timing of action potentials in relation to the ongoing theta cycle (Backus, et al., 2016; Heusser, et al., 2016; Rutishauser, et al., 2010; Staudigl & Hanslmayr, 2013). Studies in rodents have provided a mechanism by which theta oscillations exert their influence on memory in showing that long-term-potential (LTP) and long-term-depression (LTD) occur in specific phases of the theta cycle (Huerta & Lisman, 1995; Pavlides, et al., 1988). Building on theories of synaptic plasticity (Hasselmo, 2005), it has been postulated that LTD occurs whilst most neurons in region CA1/CA3 are active in the excitatory phase of theta (as recorded from CA1/CA3 hippocampal regions), whereas LTP occurs in the inhibitory phase of theta when most neurons are silent. The model we describe here shows that stimulated hippocampal cells demonstrate a phase shift forward in theta,

enabling LTP to occur in the inhibitory phase of theta where other non-stimulated cells are silent (Parish, et al., 2018).

5.1.2 NEO-CORTICAL ALPHA

It can be assumed that there is a negative relationship between alpha power and neural activity (Haegens, et al., 2011; Hanslmayr, et al., 2011a), leading to the notion that alpha provides functional inhibition (Klimesch, et al., 2007; Jensen & Mazaheri, 2010). Supporting this notion, alpha power decreases are often localized in cortical regions relevant for a given task, whereas alpha power increases occur in competing areas that are being inhibited (Jokisch & Jensen, 2007; Sauseng, et al., 2005b; Waldhauser, et al., 2012). These findings suggest that the de-synchronisation of alpha represents the flow of information to a targeted group of neurons, as is shown in Figure 37. Consistent with this general gating function of alpha, power decreases are strongly evident in episodic memory tasks where cortical alpha power decreases predict successful encoding (Hanslmayr, et al., 2012) and retrieval (Khader, et al., 2010; Waldhauser, et al., 2016). In addition to the hippocampal theta dynamics, our model also simulates such memory dependent alpha power decreases in the neocortex during the encoding and retrieval of episodic memories.

5.1.3 A NOTE ON OTHER FREQUENCIES

Beta frequencies appear to occur exclusively in semantic memory paradigms (Hanslmayr, et al., 2009; Meeuwissen, et al., 2011). As it behaves similarly to other cortically generated frequencies (Hanslmayr, et al., 2016), it is treated as functionally synonymous with alpha for this modelling work, where any reference to alpha might be extended to beta frequencies. As this model will realise a simple 'one-shot' learning paradigm as a proof of concept theoretical framework, long-term memory consolidation and delta frequencies will not be considered here. Gamma oscillations were once considered an epiphenomenon that simply represented high frequency activation (Basar, et al., 2001). This seems to be untrue, as gamma frequencies might provide fast-paced and accurate transmission of information that enables targeted communication and spike-time-

dependent-plasticity (STDP) to occur in the brain (Fell & Axmacher, 2011). However, this is not the case for the forthcoming model, where a simple alternating-current modulator was used to generate theta and alpha frequencies, yet no such mechanism was realised for gamma.

5.2 MATERIALS AND METHODS

We here present a first computational network model which implements the aforementioned mechanisms and simulates the opposing synchronising and desynchronising behaviours in the hippocampus and neocortex during a typical episodic memory task (Parish, et al., 2018). Our model, while being very simple, successfully simulates a number of empirical findings ranging from human single neuron recordings, intracranial EEG recordings, to non-invasive EEG/MEG recordings. Therefore, it represents a useful theoretical link between different levels of human electrophysiological recordings.

This section details the mathematical principles underlying this spiking neural network model (Parish, et al., 2018). First, we explain the principles guiding the computational model and experimental paradigm to be simulated (O'Reilly, et al., 2011). Then, we describe the neurophysiological mechanisms that this model seeks to emulate, revisiting and adapting some methods from Chapter 2 (Abeles, 1991; Dayan & Abbot, 2001; O'Reilly & Munakata, 2000; Song, et al., 2000). Finally, we end with a description of the separately created cortical and hippocampal assemblies, along with their respective neural dynamics and oscillatory mechanisms. This part also evaluates some decisions that were taken through the design process of this model.

5.2.1 MODELLING PRINCIPLES & PROCEDURE

Here we describe a simple computational neural network model, which takes inspiration from the complementary learning systems framework (CLS; O'Reilly, et al., 2011), and lends credence to the previously theorised notion that opposing oscillatory behaviour in cortical and hippocampal regions both contribute to episodic memory formation (Hanslmayr, et al., 2016). We do not fully detail the

different steps of how information enters and exits the hippocampus through different subregions, e.g. via the perforant pathway from entorhinal cortex.

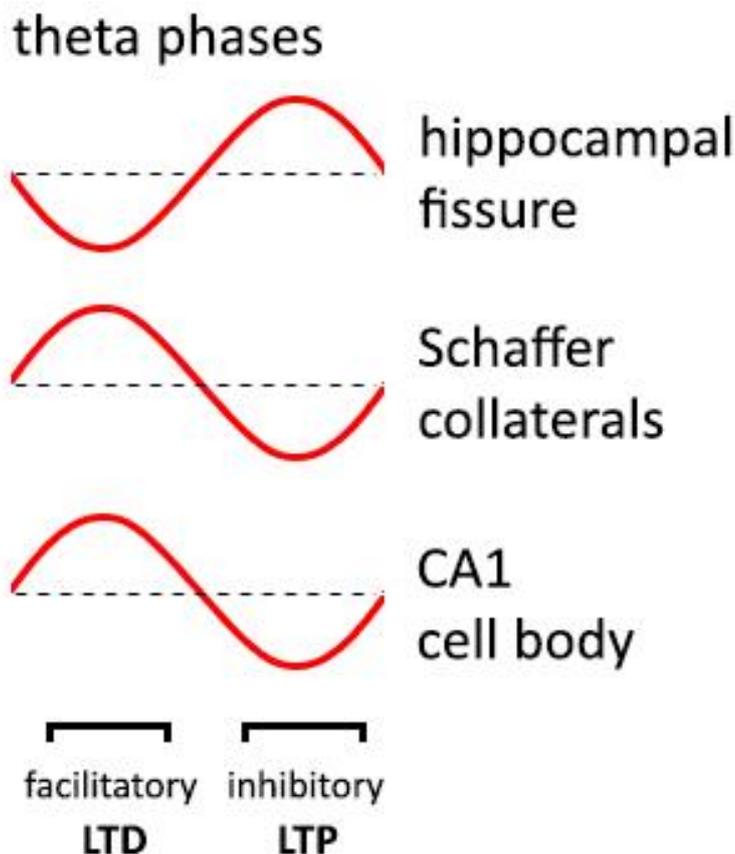


Figure 38 – Coexisting theta phases in the hippocampus that are assumed for the purpose of our model. Here we define a facilitatory phase as functional facilitation of CA1 neurons, where local theta is positive and CA1 neurons are active. An inhibitory phase indicates functional suppression of CA1 neurons, where local theta is negative and CA1 neurons are typically silent. Long-term-depression (LTD) and long-term-potential (LTP) have been found to occur in opposing theta phases (Huerta & Lisman, 1995), thus for the purpose of this model LTD and LTP occur during functional facilitation and functional suppression of CA1 neurons, respectively.

Importantly, theta oscillations show a phase reversal between the two pathways from entorhinal cortex to CA1 (the monosynaptic perforant pathway and the tri-synaptic pathway, via the Schaffer collaterals), which is the focus of previous models describing the computational utility of theta in

providing discrete time windows for encoding and retrieval (Hasselmo, 2005) or error-driven learning (Ketz, et al., 2013). Our model draws inspiration from these works, but focusses particularly on the dynamics in region CA1 (see Figure 38). The key functional property we have constructed our model upon is that theta sets up an inhibitory phase at the soma of pyramidal cells, at which LTP occurs, and a facilitatory phase at the soma of such cells, at which LTD occurs. Neurophysiologically, this could arise from the coincidence of a trough of fissure theta (which is known to coincide with LTP; Huerta & Lisman, 1995); a peak at stratum radiatum (input from Schaffer collaterals to CA1); and a trough at stratum pyramidale (i.e. functional inhibition at the cell body). This pattern is justified in Hasselmo et al. (2005, section “Induction of LTP”), and is consistent with Hanslmayr, et al., 2016, which refers to the peak in stratum radiatum. To simplify presentation, through the following Chapters we use functional descriptors, i.e. we talk in terms of the inhibitory phase of theta, meaning functional suppression at the pyramidal cell body, and the facilitatory phase of theta, meaning functional facilitation at the pyramidal cell body. In these terms, we will model a simple mechanism to simulate a typical episodic memory paradigm where an association between stimuli has to be learnt in one trial, as described in the section on episodic memory in Chapter 3 (Ison, et al., 2015). As echoed throughout Chapter 2, a principle of this modelling endeavour will be to identify the simplest neural instantiation of our theory under an Ockham’s razor principle.

We chose to compare our model to an experiment that recorded from medial-temporal-lobe (MTL) neurons within epilepsy patients (Ison, et al., 2015), see Chapter 3.2. As shown in Figure 18, and depicted here in Figure 39, the experimenters screened many images of people to each participant to find one that the neuron under observation responded to, denoted from here on as the preferred (P) image. A separate image of a location was chosen that the neuron did not respond to, denoted as the non-preferred (NP) image. The P image of the person was then digitally superimposed onto the NP image of the location (denoted here as the composite (C) image), before being presented to the participant in what is termed here as the learning phase. The experimenters then conducted the screening process again, presenting both the NP & P images, to assess the impact of learning

on the activity of the hippocampal neuron. Figure 39 shows how we simulated this paradigm, where there is a screening phase before and after the presentation of the composite stimulus.

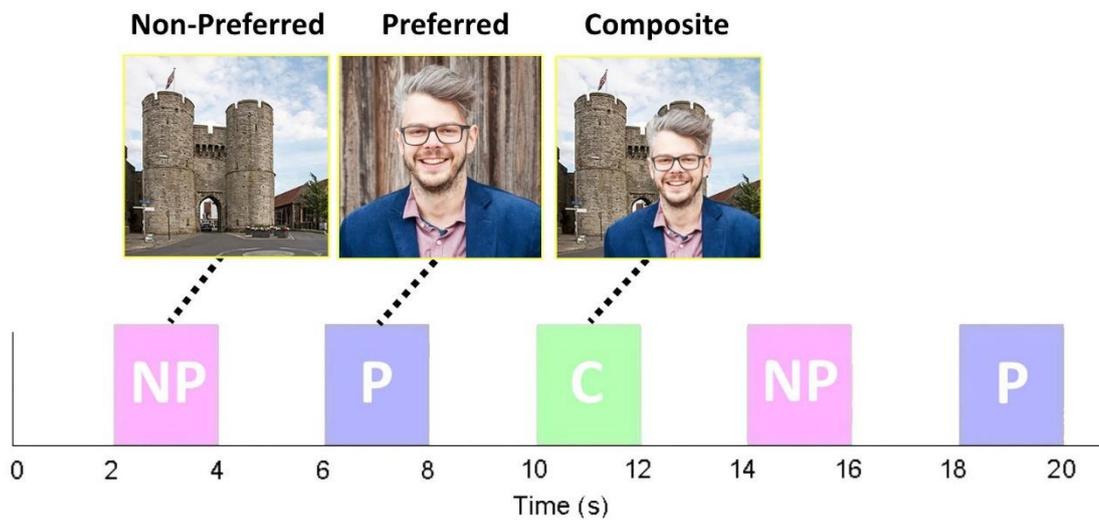


Figure 39 – Experimental paradigm. A non-preferred (NP) and preferred (P) image are found that the neuron does not and does respond to. These are then combined and presented in a composite (C) stimulus. Both P and NP images are presented again after this learning phase.

5.2.2 NEURON PHYSIOLOGY

Our model comprises two groups of neurons representing the neo-cortex (NC) and the hippocampus (Figure 40), split again into two subgroups coding for the P and NP images (where the number of neurons in each group was ($N_{NC} = 20, N_{hip} = 10$)). All neurons are simulated using an integrate-and-fire equation (Equation 19; $V_{th} = -55, E = -70, C_m = 240, D_{ref} = 2, \tau_m = 20$), adapted from that which was described in Chapter 2 (Dayan & Abbot, 2001).

$$\Delta V_m(t) = \frac{E - V_m(t-1)}{\tau_m} + \frac{I_{tonic}(t) + I_{syn}(t) + I_{ADP}(t)}{C_m}$$

Equation 19 – The integrate-and-fire equation for the Sync/deSync model

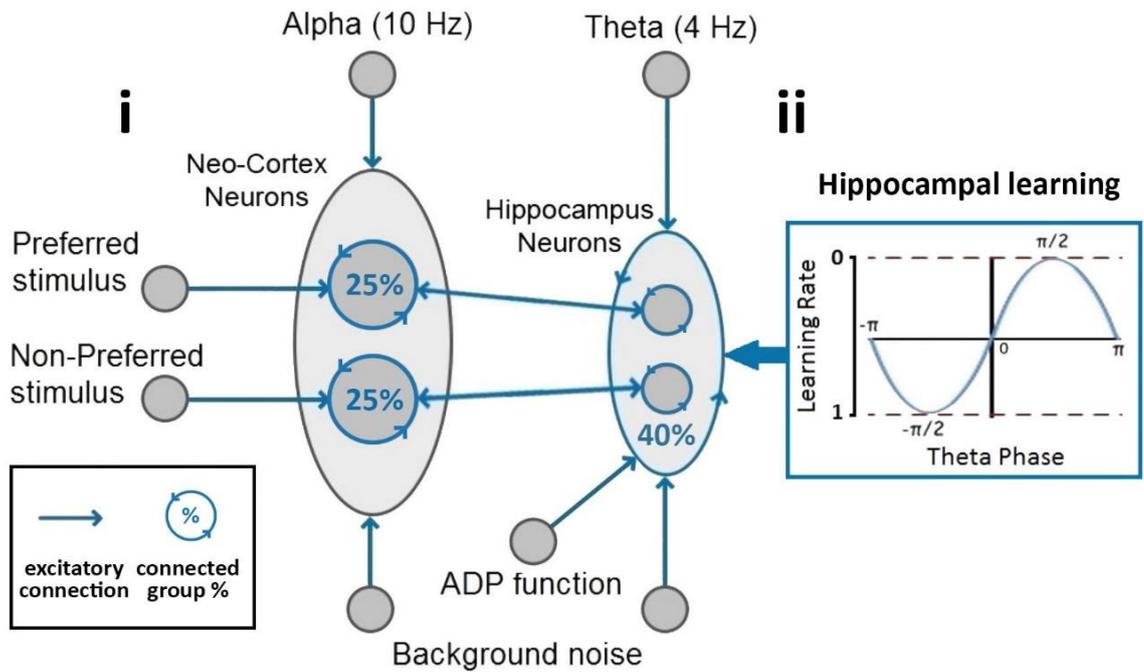


Figure 40 – The architecture of the network (i) shows how a group of neo-cortical (NC) neurons and a group of hippocampal neurons receive input from a 10Hz and 4Hz tonic wave, respectively, and both groups receive background noise from Poisson distributed spikes. Two subgroups of NC neurons receive input from higher level areas that represent the P and NP image. Each subgroup of NC and Hip neurons have reciprocal connectivity between themselves, 25% for NC and 40% for Hip. Hippocampal neurons also receive an after-de-polarisation (ADP) function. Hippocampal neurons are interconnected (i.e. not just within subgroups), again with 40% connectivity, and spike-time-dependent-plasticity (STDP) is enabled with a theta phase dependent learning rate (ii).

The neural dynamics here are as described in Chapter 2. In summary, a spike event is sent to other downstream connected neurons if the membrane potential ($V_m(t)$) of a neuron surpasses the threshold for firing (V_{th}). After a spike, the neuron enters a refractory period, where the membrane potential is clamped to the resting potential (E) for a set period (D_{ref}). With this equation, the membrane potential of a neuron is constantly decaying to its resting potential (E) at a rate dictated by the membrane time constant (τ_m). The sum of all inputs at t is divided by the capacitance (C_m) of the membrane potential. Inputs originate from constant alternating currents (I_{tonic}), the sum of

excitatory-post-synaptic-potentials (EPSPs) from spikes at each input synapse (I_{syn}) and an after-de-polarisation function (I_{ADP}), which will be described in more detail later.

$$EPSP(t) = \left(e \cdot \frac{\Delta t}{\tau_s} \right) \cdot \exp\left(-\frac{\Delta t}{\tau_s}\right), \quad \Delta t = t - t_{fire}$$

Equation 20 – Excitatory-post-synaptic-potential (EPSP) for the Sync/deSync model

As described in Chapter 2.2.1, an Alpha function was used to model EPSPs for spike events (Equation 20), where Δt is equal to the current time (t) minus the time of the eliciting spike (t_{fire}). The higher the synaptic time constant τ_s , the larger the integral through time of the EPSP, ensuring that a spike has a more sustained effect on the receiving neuron's membrane potential. All synapses within the NC integrated with a τ_s of 1.5ms, whilst synapses within the hippocampus integrated with a slightly larger synaptic time constant ($\tau_s = 5ms$) to allow them to more easily interact with one another. Spikes originating from external noise generators had a synaptic time constant of 1.5ms.

5.2.3 NEO-CORTICAL SYSTEM

Based on CLS (O'Reilly, et al., 2011), the NC system learns slowly from repeated presentations, where the transfer of short-long term memory is thought to take place during sleep (Diekelmann & Born, 2010; Walker, 2018). As our model emphasises the effect of oscillations on a single learning event (Ison, et al., 2015), we assumed the existence of two pre-established NC populations, one representing the P and the other the NP concept, where neurons within each population had a 25% chance of being connected and synaptic modification was not implemented due to an assumed slow cortical learning rate (Figure 40i). Each NC neuron received background noise, representing "chatter" from other brain regions that kept neurons in a ready-to-fire state (Abeles, 1991; Dayan & Abbot, 2001), as described in Chapter 2.2.2, in the form of Poisson distributed spike-events (~42k spikes/s).

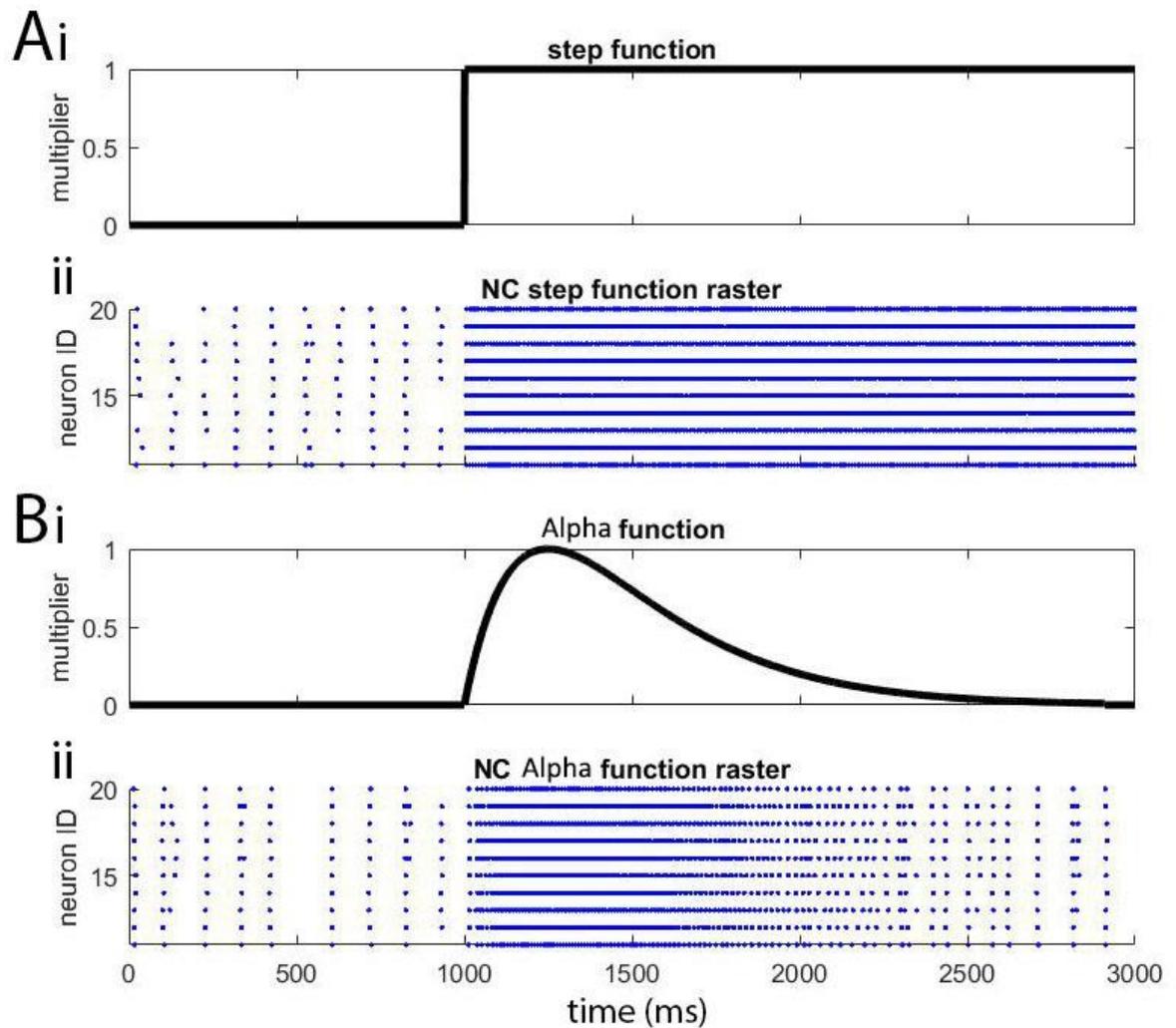


Figure 41 – Neo-cortical input from stimuli as a step function (**Ai**) or as a more realistic Alpha function (**Bi**), with accompanying raster plots of NC activity (**Aii/Bii**).

We do not explicitly model a neural mechanism for oscillations, thus a sine wave of frequency 10Hz (amplitude = 21pA) was fed into NC neurons via I_{tonic} to model ongoing Alpha. This approximates the dominance of Alpha oscillatory activity in the cortex, which arise via pacemaker regions like the thalamus (Hughes, et al., 2004) or emerge via cortico-cortical top-down interactions (van Kerkoerle, et al., 2014), where an excitation-inhibition balance is thought to generate oscillations in cortical cell assemblies (Markram, et al., 2004).

Two separately generated Poisson distributed spike-trains ($\sim 80k$ spikes/s) were then paired with each NC subgroup upon stimulus presentation, modelling the activation of the P and/or NP images from higher cortical and visual areas. Initially, activation in the model was caused by an on/off step

function (Figure 41Ai), where the stimulus provided a wall of instantaneous excitation. Upon reflection, stimulus related spike-trains were then multiplied by an EPSP function ($\tau_s = 250\text{ms}$; Figure 41Bi) to more realistically model the typical shape of event-related-desynchronisations occurring at stimulus onset (Klimesch, et al., 2007). Raster plots of NC spike events influenced by a step (Figure 41Aii) and an EPSP (Figure 41Bii) function are shown in each respective bottom panel.

5.2.4 HIPPOCAMPAL SYSTEM

Hippocampal neurons were similarly organised into two subgroups (Figure 40i), where each neuron received background noise ($\sim 4\text{k spikes/s}$) and a sine wave of 4Hz (amplitude = 28pA) to model ongoing theta. This ongoing theta oscillation approximates input into the hippocampus from pacemaker regions like the septum (Petsche, et al., 1962), or interactions between different types of interneurons acting as local theta generators (Rotstein, et al., 2005). Neurons here received additional input from an after-de-polarisation (ADP) function to control activation ($A_{ADP} = 100\text{pA}, \tau_{ADP} = 250\text{ms}$), described in Chapter 4.2 (Jensen, et al., 1996) and shown by Equation 21. This provided ramping input, reset after each spike-event (t_{fire}), which eventually falls exponentially back to zero like an Alpha function.

$$I_{ADP}(t) = \frac{A_{ADP} \cdot \Delta t}{\tau_{ADP}} \cdot \exp\left(1 - \frac{\Delta t}{\tau_{ADP}}\right), \quad \Delta t = t - t_{fire}$$

Equation 21 – After-de-polarisation function, from Jensen, et al., 1996.

Evidence for an ADP function in hippocampal neurons has been found experimentally during cholinergic (Andrade, 1991; Caesar, et al., 1993; Libri, et al., 1994) and serotonergic (Araneda & Andrade, 1991) modulation. Here, this has the effect here of modelling an inhibitory input for each hippocampal neuron that occurs after each spike event, possibly from a connected inhibitory interneuron. The activity of a single hippocampal integrate-and-fire neuron is shown in Figure 42. Here, one can see the influence of the AC generator and the ADP function as they coordinate

hippocampal activity to occur in the peak of theta phase, where some variation at each cycle is due to the Poisson distribution of spike events that are fed into hippocampal neurons.

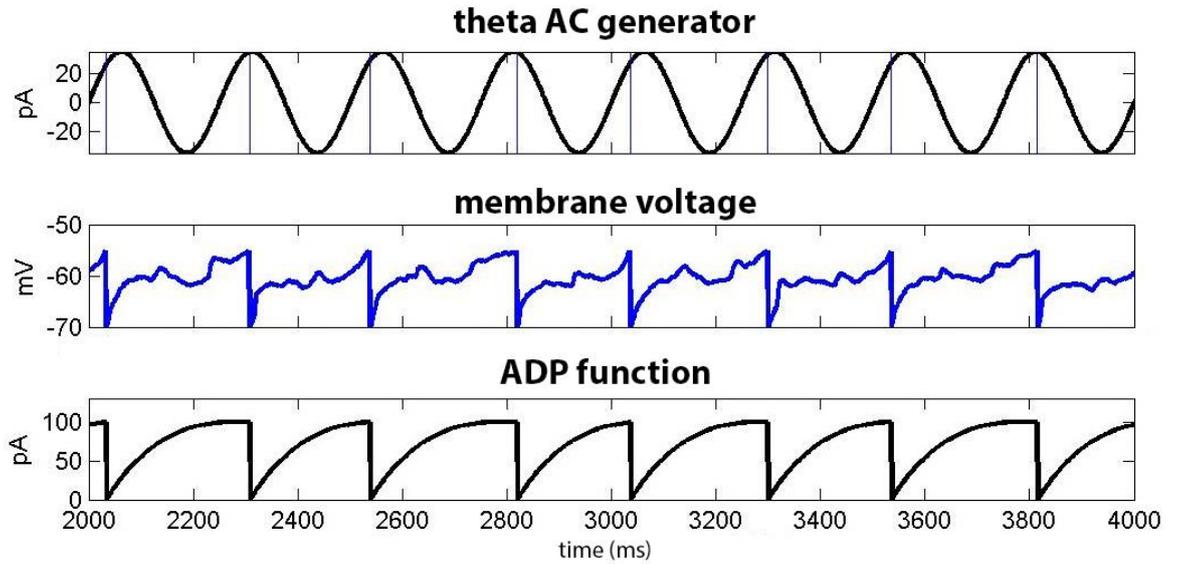


Figure 42 – Activity of a single hippocampal neuron, receiving input from a theta frequency AC generator, an after-depolarisation (ADP) function and some noise. Spike events are indicated as vertical blue lines in the top panel.

Based on CLS, the hippocampal system learns quickly from a single presentation (O'Reilly, et al., 2011). Therefore, hippocampal synaptic modification was enabled via an adapted spike-time-dependent-plasticity (STDP) learning rule, as described in Chapter 2 (Song, et al., 2000).

$$A_+ = \varepsilon \cdot \theta(t) \cdot W_{max}, \quad A_- = 1.1 \cdot A_+$$

Equation 22 – Synaptic LTP and LTD changes in an adapted STDP mechanism (Song, et al., 2000).

We adjusted this rule to relate to empirical evidence that hippocampal learning is theta phase dependent (Huerta & Lisman, 1995), with LTP occurring in the functionally inhibitory phase and LTD in the functionally excitatory phase of theta (Hasselmo, 2005). To this end, synaptic LTP was implemented by multiplying STDP weight modifications by the phase of the theta sine wave, with a value between 0 and 1, with 0 on the excitatory “up” phase and 1 on the inhibitory “down” phase

(Figure 40ii). Equation 22 shows that upon spiking, an increase (A_+) for contributing synapses was calculated as the product of a constant learning rate ($\varepsilon \in \mathbb{R}. 0 \leq \varepsilon \leq 1$), theta at time t ($\theta \in \mathbb{R}. 0 \leq \theta \leq 1$) and the maximum weight (W_{max}), whilst punishments for competing synapses (A_-) were given a slight priority over rewards to maintain network stability (Song, et al., 2000).

The function in Equation 23 was then used to calculate the STDP acting on the post-synaptic synapse between spiking neurons, where an exponential weighting of A_+ was applied if the pre-synaptic spike occurred before the post-synaptic spike and of A_- if the post-synaptic spike occurred first.

$$F(\Delta t) = \begin{cases} A_+ \cdot \exp(\Delta t/\tau_s), & \text{if } \Delta t < 0 \\ -A_- \cdot \exp(-\Delta t/\tau_s), & \text{if } \Delta t \geq 0 \end{cases}$$

Equation 23 – STDP function from Song, et al., 2000.

Whenever a spike event occurs at unit i or j , an accumulated STDP update $v_{ij}(t)$ for synapse i to j is calculated from its history of previous spiking, in Equation 24. Here, STDP synaptic modification at time t for a network with node labels $\mathfrak{N} = \{1, \dots, n\}$. $C(i, j)$ is true if and only if i and j are connected. $S(t)_i$ indicates a spike event at the i th neuron at time t . $T(k, t)$ returns the set of all times before time t , at which there was a spike at neuron k . This is used to provide spike events paired across synapse ij , with the spike at time t . In addition, we use auxiliary weight variables v_{ij} and V_{ij} to enable application of a piecewise linear bounding function.

$\forall i, j \in \mathfrak{N} \text{ s.t. } C(i, j)$.

$$v_{ij}(t) = \begin{cases} \sum_{t' \in T(i, t)} F(t' - t), & \text{if } S(t)_j \\ \sum_{t' \in T(j, t)} F(t - t'), & \text{if } S(t)_i \\ 0, & \text{otherwise} \end{cases}$$

$$T(k, t) = \{ d \in \mathbb{R}^{0,+} \mid S(d)_k \wedge t \geq d \}$$

Equation 24 – Equation for accumulated STDP updates for synapse ij .

In the weight update of Equation 25, the final term is an additional exponential passive decay multiplied by the complement of the phase of theta ($1 - \theta(t)$). The presence of this decay is consistent with the non-specific LTD that might occur during oscillatory spiking in the facilitatory phase of theta (Hasselmo, 2005). This decay was larger for smaller weights, establishing a transition point whereby weakly interacting synapses were pruned ($\tau_w = 20$). A piecewise linear bounding function, see Equation 26, was also used to protect against sign reversal and run-away weights ($W_{max} = 120$; $W_{min} = 0$).

$$\forall i, j \in \mathfrak{N} \text{ s.t. } C(i, j) . V_{ij}(t) = W_{ij}(t-1) + v_{ij}(t) - \frac{(1 - \theta(t)) \cdot \exp\left(-\frac{W_{ij}(t-1)}{\tau_w}\right)}{\tau_w}$$

Equation 25 – Weight update equation.

$$W_{ij}(t) = \begin{cases} W_{min}, & \text{if } V_{ij}(t) < W_{min} \\ W_{max}, & \text{if } V_{ij}(t) > W_{max} \\ V_{ij}(t), & \text{otherwise} \end{cases}$$

Equation 26 – Piecewise linear bounding function.

This learning rule ensures that STDP (both LTP & LTD) occurs maximally in the trough of theta, though due to the piecewise linear bounding function weights tend to undergo LTP during periods of activation. Conversely, a form of LTD is implemented in the peak of theta through the passive decay, simulating the assumed effect of accumulated spiking during this period. As such, we do not explicitly model LTP in the trough and LTD in the peak of theta, but mechanisms that operate to such an effect. Therefore, our learning rule might better be regarded as theta-modulated STDP.

Hippocampal neurons were interconnected with a probability of 40% to form a connection (Figure 40i). This was decided, since hippocampal weights tend to increase uni-directionally when approaching 100% intra-hippocampal connectivity (Figure 43). This is likely due to one group winning completely via the competitive STDP rule (Song, et al., 2000), where one concept leads the

other through direct pathways as shown in Figure 43A. In Figure 43C, the blue line indicates that P->NP weights increase at the expense of the NP->P weights in red (dashed line), though the direction is reversed for another simulation with the same parameter set in Figure 43A.

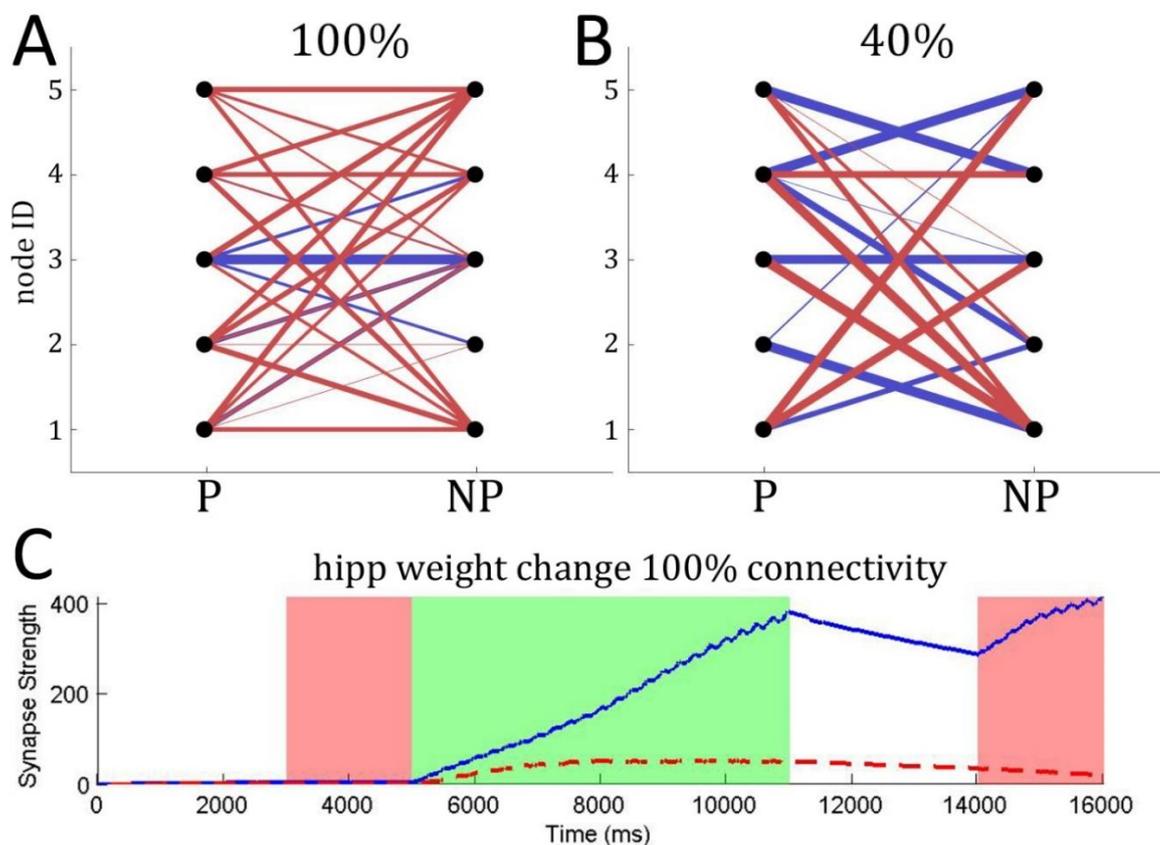


Figure 43 – STDP driven connectivity patterns between hippocampal P & NP subgroups when intra-hippocampal connectivity is 100% (A) or 40% (B). Weights here begin at zero and increase during the simulated hippocampal learning paradigm. When looking at hippocampal weight change for 100% connectivity (C), weights tend to go up in only one direction. For all graphs, lines go in the P->NP (blue) direction and the NP->P (red) direction. For A-B, line width indicates synaptic strength. For C, shaded phases indicate the learning phase (green shaded, P & NP stimuli shown together) and screening phase (red shaded, P or NP stimuli shown).

The direction of transmission tends to rely on which neuronal group begins firing first, enabling one group to control the post-synaptic potential of neurons in the other population. This in turn typically relies on which neuronal group has the strongest initial connections, which in this case is the most

recently presented image due to the dual synaptic plasticity and decay mechanisms. Lower levels of connectivity (Figure 43B), however, enable multiple competing pathways to emerge between sparsely connected neural assemblies, where bi-directional loops emerge to intertwine populations encoding for the P & NP stimulus. As it was assumed that both images were previously known to the participants but not associated, a random 50% of synapses within each hippocampal subgroup had initial synaptic weights of W_{max} whilst all others were set to 0. This ensured the random assignment of pre-established sets of winning and losing pathways within the subgroups coding for the P & NP images.

5.2.5 SIMULATION PROCEDURE

We simulated our model based on a learning paradigm used in an MTL single cell recording experiment (Ison, et al., 2015), described in Chapter 3. During the initial screening phase, both the P & NP images were presented individually. This was simulated by independently creating two Poisson distributed spike trains ($\sim 80\text{k/s}$ for 2 seconds) that fed into each respective P & NP subgroup of NC neurons (Figure 39; P = blue, NP = magenta). An inter-stimulus interval of 2 seconds was used. Afterwards, we presented both images in a composite stimulus (green), where both subgroups of NC neurons concurrently received spike-trains. Following this learning phase, we repeated the screening phase to assess the capability of the network to associate these stimuli together. The whole process was simulated 1000 times to assess the variability of the network, where for each simulation the alpha and theta sine waves each began at a different random phase (choosing a random 30° angle between $0-360^\circ$, i.e. $N \times 30^\circ$ where $N \in \mathbb{N}$ s. t. $0 \leq N \leq 12$), new noisy spike trains were generated, and new initial patterns of connectivity were established. Thus, there was no carry-over of weight values between runs. Resulting data take an average over all simulations, where each simulation is treated as an individual trial.

5.3 SUMMARY

Here we described the Sync/deSync conundrum, whereby both neo-cortical alpha desynchronisation and hippocampal theta synchronisation are strong predictors of learning (Hanslmayr, et al., 2016). To explain how such a combination might be possible, we have created a neural network model of interacting neo-cortical and hippocampal regions in a complimentary-learning-systems (CLS) framework (O'Reilly, et al., 2011). Here, a neo-cortical system was created by feeding a sine alternating-current (AC) of alpha frequency into a group of integrate-and-fire neurons. Similarly, the hippocampal region was created with an AC current of theta frequency, with the addition of an after-depolarisation (ADP) function to control the rate of activation. Spike-time-dependent-plasticity (STDP) was initialised in the hippocampal population to enable learning, where synaptic changes were modulated by theta to model the experimentally observed theta-phase dependent learning in the hippocampus (Hasselmo, 2005; Huerta & Lisman, 1995). In addition, it was identified that neuronal populations developed asymmetric learning when intra-hippocampal connectivity was too high, fitting in with theories of competitive STDP whereby nodes compete for control of other post-synaptic nodes (Song, et al., 2000). An experimental paradigm was identified (Ison, et al., 2015), that would enable us to directly compare simulated hippocampal spiking data with that from single-cell recording experiments. We hope that this model provides evidence for both of the hypotheses of this thesis, namely that cortical oscillatory desynchronisations enable the active representation of concepts in relation to irrelevant items, whilst spike-timing relative to the phase of hippocampal frequencies maximises synaptic learning.

The following chapter analyses and discusses the data simulated through this model, where we can compare results to both hippocampal single cell data (Ison, et al., 2015; Rutishauser, et al., 2010) and the previously described EEG/MEG studies (Backus, et al., 2016; Fell & Axmacher, 2011; Fellner, et al., 2016; Hanslmayr, et al., 2012; Khader, et al., 2010; Staudigl & Hanslmayr, 2013). The latter will be done by calculating the local-field-potential (LFP) and time-frequency-analysis (TFA) from simulated spike data through the method described in Appendix A.

6. THE SYNC/DESYNC MODEL - RESULTS

The methods and literature review in previous Chapters has led us to the first modelling work of this thesis (Parish, et al., 2018). In Chapter 5, a neural network model was described to propose a plausible solution for the Sync/deSync (SDS) conundrum, outlined in the previous Chapter, where a synchronised hippocampus and a desynchronised neo-cortex seem to be required for successful memory encoding (Hanslmayr, et al., 2016). This Chapter will go step-by-step through the data generated by the model, analysing results and comparing them to experimental findings. Afterwards, we conclude our findings, along with any limitations of the model, and discuss the impact of the model and possible future steps.

6.1 RESULTS

This section analyses the data produced by the Sync/DeSync framework. This begins by looking at hippocampal weight change, which is realised via our theta-phase dependent learning rule described in the preceding Chapter. The next section then discusses how weight change encourages increased activation in hippocampal populations in response to future stimuli presentations. The phase of this hippocampal activity is then looked at, providing evidence that corroborates some experimental findings (Ison, et al., 2015; Rutishauser, et al., 2010). From there, we overview the crux of the SDS conundrum, namely neo-cortical alpha desynchronisation and hippocampal theta synchronisation within the model.

The benefit of creating a computational model of biophysical phenomena is that one can simulate a range of paradigms that predict what one might see in the brain. This Chapter follows this philosophy, modifying stimulus strength to analyse the SDS conundrum through parameter space, providing theoretical evidence for a possible mechanism behind oscillatory power changes. We also correlate oscillatory power with learning, making some key predictions regarding the degree to which a sync/desync is required in order to predict successful learning, whilst simultaneously giving evidence for some of the alternate findings within the oscillations and memory field.

It has been proposed that maximal synaptic modification occurs between hippocampal neurons that are stimulated to shift forward in phase and fire in the inhibitory cycle of an ongoing theta oscillation (Hasselmo, 2005). Due to this, synaptic learning only occurs during the screening and learning phases of the simulation (Figure 44; NP stimulus-magenta; P stimulus-blue; C stimulus-green) and not during the inter-stimulus intervals. Weight change after stimulus onset follows the Alpha function shape of the activation fed into these neurons. Due to the maximisation of a random 50% of synapses within each P & NP subgroup, the average weights of these groups begin at $W_{max}/2$ (Figure 44A). Throughout the entire simulation, there is weight change within each subgroup (P-blue line; NP-magenta dash) when the respective image they are coding for is presented. With the competitive STDP rule, winning and losing weights are pushed towards W_{max} or W_{min} respectively, causing a capping effect where a weight in one direction can still change whilst the reverse direction is capped. In the screening process before learning (NP/P BL), this means that the average weight of each subgroup rises a small amount to stabilise just above $W_{max}/2$ every time the respective image is presented.

When the composite stimulus is presented (green), there is only marked synaptic change between both subgroups (Figure 44B; P->NP - blue line; NP->P - magenta dash). Here, weights go up bi-directionally as both subgroups of neurons are concurrently stimulated to become active during the inhibitory phase of theta. In this phase, there are short term increases and decreases in weights, as paths are found between subgroups. As indicated by Figure 44B, DL period, sustained changes are positive. When the screening phase is repeated after the learning phase, weights fluctuate and eventually settle with an increase in the direction from the active population to the non-active population. Before learning, concepts are only strengthened when the relevant image is presented. After learning, both concepts are reinforced upon the presentation of either image, indicating how previously associated but non-present concepts can remain strong over time. Weights passively decay very slowly according to an exponential pattern to model the effect of a large population of

neurons spiking during the facilitatory phase of theta, where LTD has been found to occur (Hasselmo, 2005). As LTP occurs over a spectrum of 1 to 0, small weight increases occur as neurons spike on either side of the point at which theta maximally inhibits. The passive decay implemented here is stronger for smaller weights, to mitigate these gradual weight increases and prune irrelevant synapses. This can be seen most prominently in Figure 44B during the initial screening phase (2-4 & 6-8 seconds), where small weight increases to stimulated neurons decay quickly. LTD weight decay is also prominent in the inter-stimulus periods, where all weights slowly reduce over time.

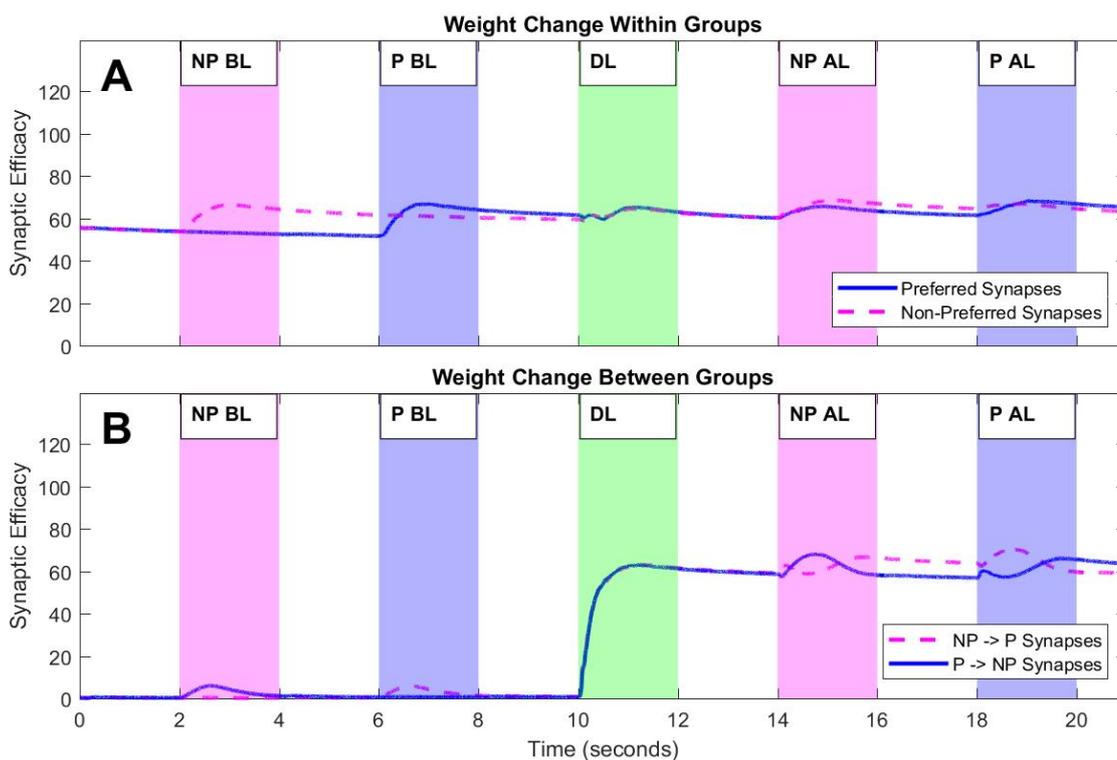


Figure 44 – Hippocampal weight change throughout the simulation both within (A) and between subgroups (B) that code for the P and NP stimulus. Weights within each subgroup increase when the relevant image is presented (A), where the magenta and blue periods indicate the presentation of the NP and P images, respectively, and the green period indicates the presentation of both images combined into a composite image. During this learning period (DL), weights from the NP to the P subgroup (magenta dashed) and vice-versa (blue solid) increase (B). Outgoing weights then increase upon the presentation of the relevant stimulus after learning (AL). Incoming weights also increase a small amount before learning (BL), then decay back to zero.

6.1.2 HIPPOCAMPAL ACTIVITY

Activity is measured as the sum of spikes within bins of a 20ms width throughout the length of a simulation, taking an average of 1000 simulations with varying random phases for alpha and theta oscillations, where the mean firing rate is shown with bootstrapped confidence intervals (Figure 45A). As we have access to data from both preferred (P) and non-preferred (NP) neurons, we can capture the network's capability of recognition, where P & NP units respond to their own stimulus, and cued recall, where P & NP units respond to the opposite stimulus. During the initial screening phase before learning (BL), we see that neurons respond to their relevant images (Figure 45A), where activation at stimulus onset seems to cause a phase reset. This generates a high-frequency damped oscillation that is phase consistent across replications, and rides on top of a much lower frequency evoked transient, which plays out over a second or more.

When the C image is presented during learning (Figure 45Ci), activity increases dramatically. Figure 45Cii shows the cause of this increase by breaking down the average input coming into neurons during learning, where the sum of all input sources follows the grey area (I). Here, we see an external force (I_{ext}) drive the hippocampus at stimulus onset, which then causes the ADP current (I_{ADP}) to reset before it can reach maximum conductance (A_{ADP}), thus reducing its effect. The relative increase in activation is due to substantial weight change, and resulting additional input, between subgroups ($I_{\text{H}<>\text{H}}$). Activation then feeds back into each subgroup dependent on how weights develop. When the screening phase is repeated after learning, the network successfully performs cued recall (Figure 45Bii) due to the aforementioned weight change, showing that our model efficiently learns associations between two arbitrary stimuli in one short presentation, a crucial requirement for a model of episodic memory. Similarly, random reciprocal feedback of activity between subgroups causes a relative increase in activation (Figure 45Bi). Raster plots show the activation of a single random P and NP neuron, as they respond to presentations of the P stimulus through a randomly chosen trial, where each line corresponds to a spike event (Figure 45Aiii, Biii & Ciii). These are colour co-ordinated with the relevant activation plots seen above.

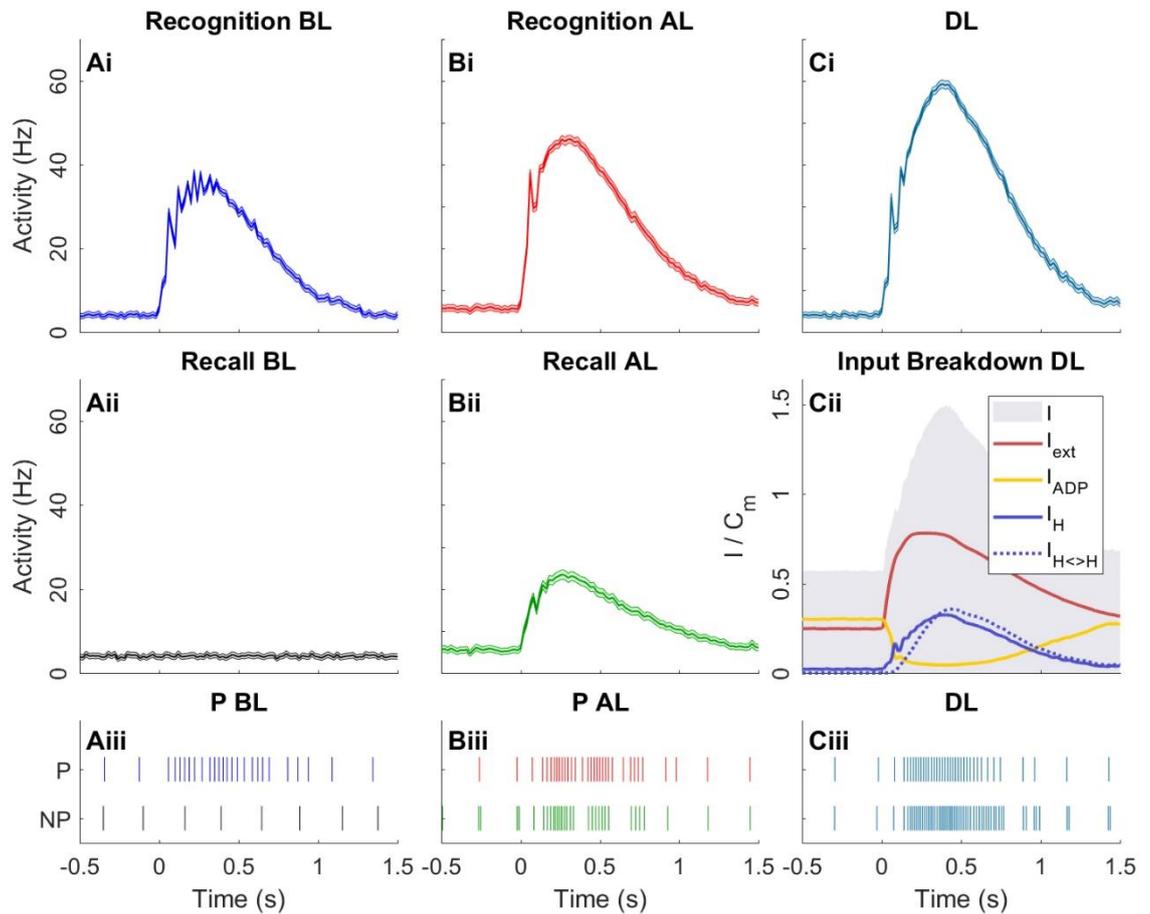
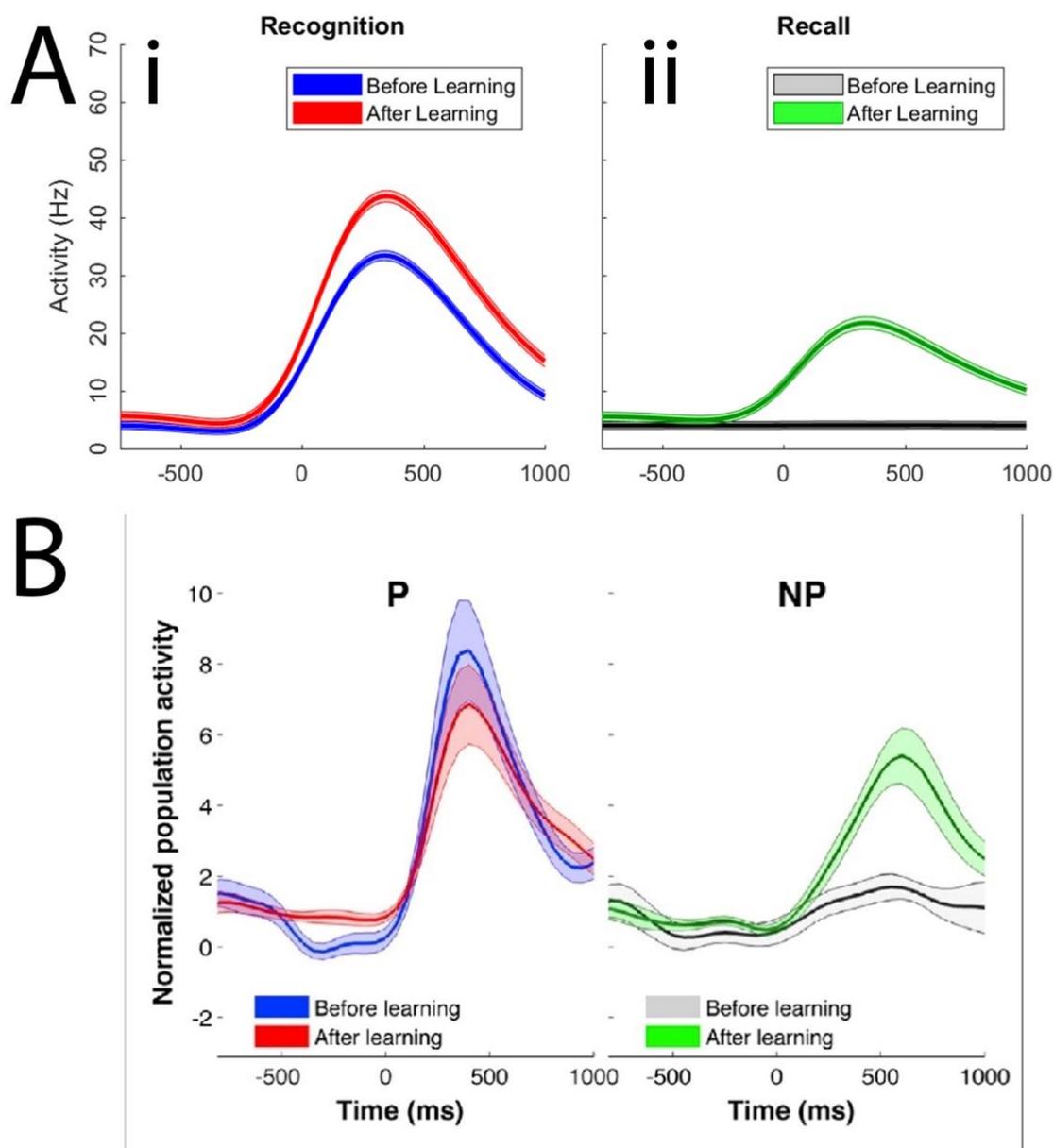


Figure 45 – Activity of Hippocampal neurons. Recognition reflects neurons responding to their own stimulus, e.g. P units activating for the P stimulus. Cued recall reflects neurons responding to the opposite stimulus, e.g. P units activating for the NP stimulus. Here, activation from before learning (BL) **(A)**, after learning (AL) **(B)** and during learning (DL) **(C)** is shown. Raster plots show the activity of a single P and NP neuron during presentations of the P stimulus BL **(Aiii)**, AL **(Biii)** and DL **(Ciii)**. The average input into both P and NP neurons across all trials is shown in **Cii**, where coincidental external drive (I_{ext}) during stimulus onset counteracts the effect of the ADP function (I_{ADP}). Additional activation causes an increase in input from other neurons within the group (I_H) and also from the opposite group ($I_{H<>H}$) as weights increase during learning.

We compare the results of our simulation to those from experimental evidence from a recent human single unit learning paradigm (Ison, et al., 2015). Figure 46Ai-ii shows smoothed curves (smoothing spline; $p = 1e^{-7}$) following simulated recognition and cued recall performance before

and after learning, compared to experimental evidence of the same data in Figure 46B. Despite some overlap of confidence intervals, Figures 46Ai-ii suggest that there is an increase in pre-stimulus activation after learning for recognition and recall in both sets of data. Raster plots show that this could be caused by occasional double spike events during the excitatory phase of theta, due to increased weights between neurons (Figure 45Biii; -500-0ms).



Rapid Encoding of New Memories by Individual Neurons in the Human Brain
Ison et al. 2015

Figure 46 – Smoothed activation data at recognition (**Ai**) and recall (**Aii**) is then compared to data reported in an MTL neuron study, adapted from Ison, et al., 2015 (**B**).

Both the model and experimental data indicate successful cued recall after learning (Figure 46Aii/B; green), however, recognition after learning shows differences (Figure 46Ai/B; red). The experimental finding is that encoding neurons become less active with successive presentations of the same stimulus (Ison, et al., 2015), perhaps due to a repetition suppression effect (Pedreira, et al., 2010). In our model, an increase in recognition activation after learning is caused by the overall increase in synaptic efficacies both between and within subgroups. This could be countered by implementing an habituation mechanism that lies outside of the scope of this model. Such a mechanism could involve the re-balancing of weights or the storing of short-term-memory in a higher brain structure.

6.1.3 THETA PHASE

Figure 47 shows theta phase for the cued recall condition during the 3 stages of the simulation. The red and green halves of the polar distribution represent the excitatory and inhibitory phases of the 4 Hz sine wave used to model theta, where $\pi/2$ is maximum excitation and $-\pi/2$ is maximum inhibition. The total number of spikes occurring within each phase quadrant of theta was recorded (Figure 47Ai, Bi & Ci), as well as the first spike of each neuron after maximum inhibition ($>-\pi/2$) (Figure 47Aii, Bii & Cii). The latter analysis was performed to show how hippocampal neurons shift forward in theta phase once stimulated. Spike numbers were normalised over 1000 simulations.

Before learning, neurons are un-responsive to the image they do not encode for and oscillate at theta, where all spikes occur during the excitatory phase (Figure 47Ai 0 to $\pi/2$ to π), with the first spikes generally occurring just before maximum excitation (Figure 47Aii; 0 to $\pi/2$). When the C image is presented during the learning phase, both subgroups become active across all phases of theta (Figure 47Bi). Importantly, in order for activation to overcome inhibition, more activity will occur during the inhibitory phase of theta. Neurons also exclusively spiked first immediately after the inhibitory maximum (Figure 47Bii; $-\pi/2$ to 0), indicating that all neurons in the P subgroup successfully phase-shifted forward once stimulated during learning.

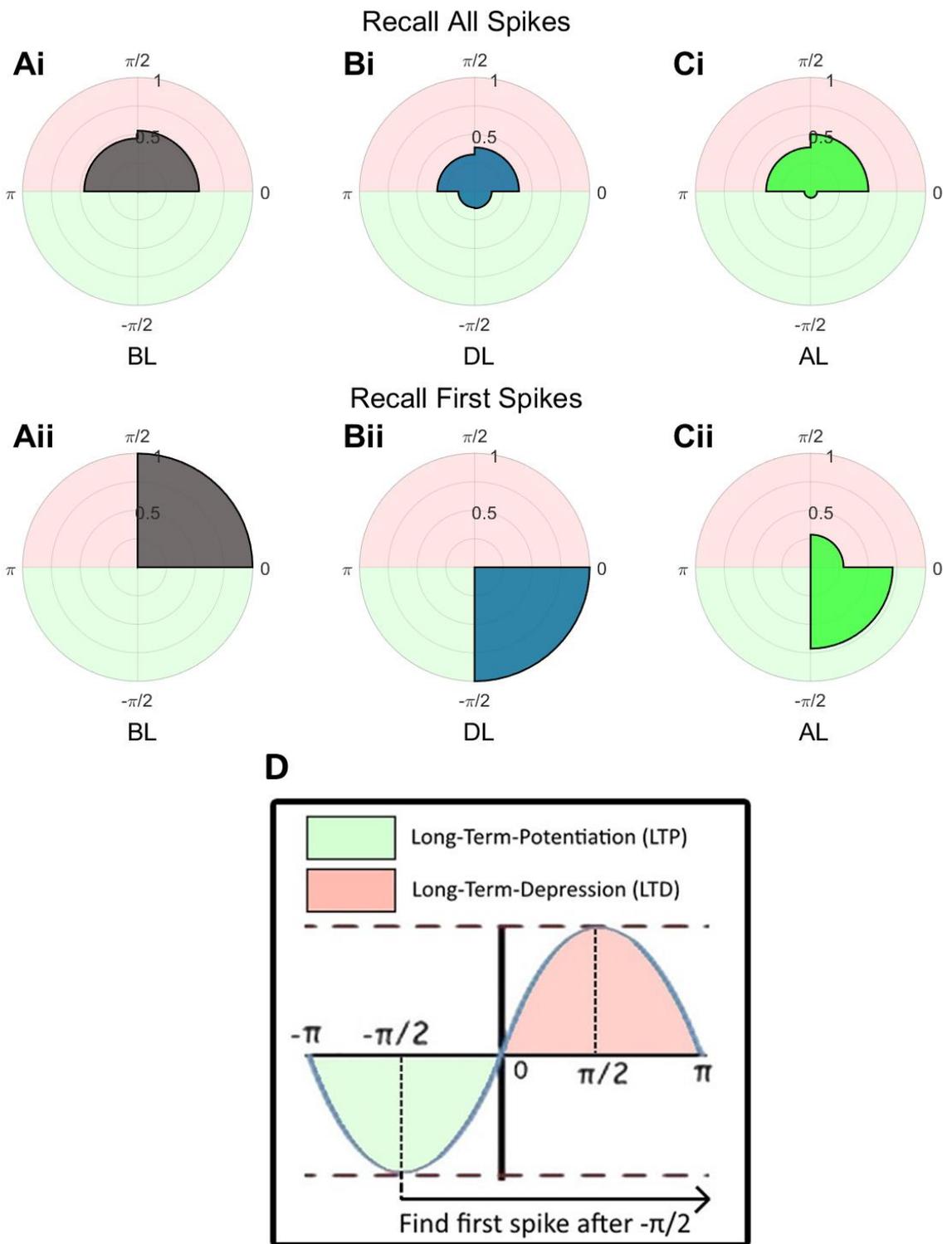


Figure 47 – Polar histograms for the recall condition of all spikes before (**Ai**), during (**Bi**) and after learning (**Ci**), and of first spikes after $-\pi/2$ before (**Aii**), during (**Bii**) and after learning (**Cii**). **D** shows the distinction between the excitatory (red) and inhibitory (green) phases of theta, where LTD and LTP are more prevalent, respectively. Data is binned in quadrants for ease of viewing.

When the screening phase occurs again after learning, neurons now respond to the opposite image. Spikes occur in most phase quadrants of theta (Figure 47Ci), but in the main during the excitatory phase. However, inhibition can now be overcome, allowing spikes to first occur during the negative phase of theta (Figure 47Cii) and demonstrating a phase shift forward in theta. This shift in phase is an index of successful learning and has been well documented in rodents for neurons encoding a particular place when the rodent approaches that place (Huxter, et al., 2003). Our model shows a similar behaviour and predicts that this shift in phase is responsible for associative memory formation. Importantly, this phase shift is most evident when analysing only the first spike within a theta cycle, starting at the theta trough (i.e. where inhibition is maximal). This prediction can be tested in studies recording single units and LFPs in human epilepsy patients (Ison, et al., 2015).

6.1.4 ALPHA DE-SYNCHRONISATION

Figure 48A shows time-frequency power spectra (8-12Hz) of the LFP of the NC neurons for the recall, recognition and learning phases. A thick band at 10Hz during the recall condition before learning shows non-stimulated neurons oscillating at alpha (Figure 48Ai), as they do not respond to an image at this time. When neurons are responsive to the image they encode for in recognition and learning conditions, a strong de-synchronisation of alpha is exhibited (Figure 48Aii/iii/v; 0 to 1s), simulating the well-documented effect of alpha suppression upon visual stimulation (Berger, 1929). A similar, but weaker effect can be seen in the cued recall condition after learning (Figure 48Aiv; 0 to 1s). This de-synchronisation is due to learning driven activation of hippocampal neurons caused by the association between the P and NP stimuli. This low-frequency drive (from hippocampus to neo-cortex) de-synchronises alpha by causing substantial activation in the inhibitory phase. The effect can be more clearly seen in Figure 48Bii, where a 20% relative decrease in alpha power from pre to post stimulus is exhibited (Figure 48Bii; 0 to 1s), consistent with the findings that memory retrieval can be predicted by this same alpha de-synchronisation (Hanslmayr, et al., 2012). Pre-stimulus alpha power is also slightly stronger (Figure 63Bi; -1 to 0s), indicating that pre-stimulus alpha/beta power can be used to predict memory formation (Salari & Rose, 2016).

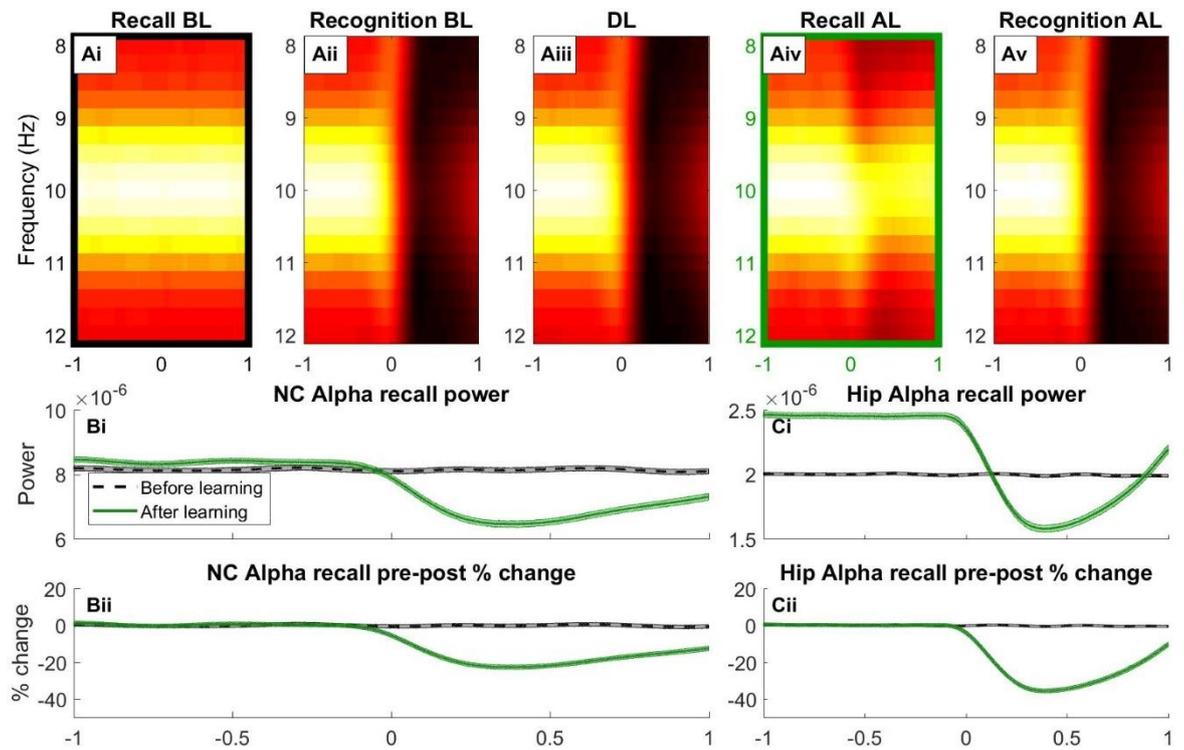


Figure 48 – Time-frequency-analysis (TFA) of Neo-Cortical alpha for the recall and recognition conditions before and after learning (**Ai-ii, Aiv-v**), as well as during learning (**Aiii**). A time-course of alpha power is shown for the colour-coded boxes around the recall condition before (**Ai**) and after (**Aiv**) learning, where pure power (**Bi**) and percent change in pre-post stimulus power (**Bii**) are shown. The same analysis can be seen for hippocampal alpha, where raw power (**Ci**) and relative power change (**Cii**) are shown. Recall and recognition here refer to all subgroups that do and do not originally respond to the given stimulus, respectively, before the learning phase takes place.

This is due to stronger weights within hippocampal subgroups causing knock-on activation during the excitatory phase of alpha. This activation feeds back into hippocampal units to cause an even more pronounced increase in pre-stimulus alpha after learning (Figure 48Ci), where after stimulus onset alpha also significantly decreases in these hippocampal units (Figure 48Cii), which is consistent with a previous study (Staresina, et al., 2016). This behaviour of our model mimics several findings in the literature showing memory dependent alpha power decreases during the reinstatement of episodic memories (Khader, et al., 2010; Michelmann, et al., 2016; Waldhauser,

et al., 2016). Here, the de-synchronisation of alpha represents the flow of information in the NC caused by activation of relevant stimuli (Klimesch, et al., 2007; Jensen & Mazaheri, 2010).

6.1.5 THETA SYNCHRONISATION

Figure 49Ai-v shows time-frequency power spectra (2-4Hz) of the LFP of hippocampal neurons for the recall, recognition and learning conditions. In the recall condition before learning, neurons do not respond to any image and oscillate at theta (Figure 49Ai). An increase in theta power accompanies increased activation, as neurons respond to the image they encode for before and during learning (Figure 49Aii-iii).

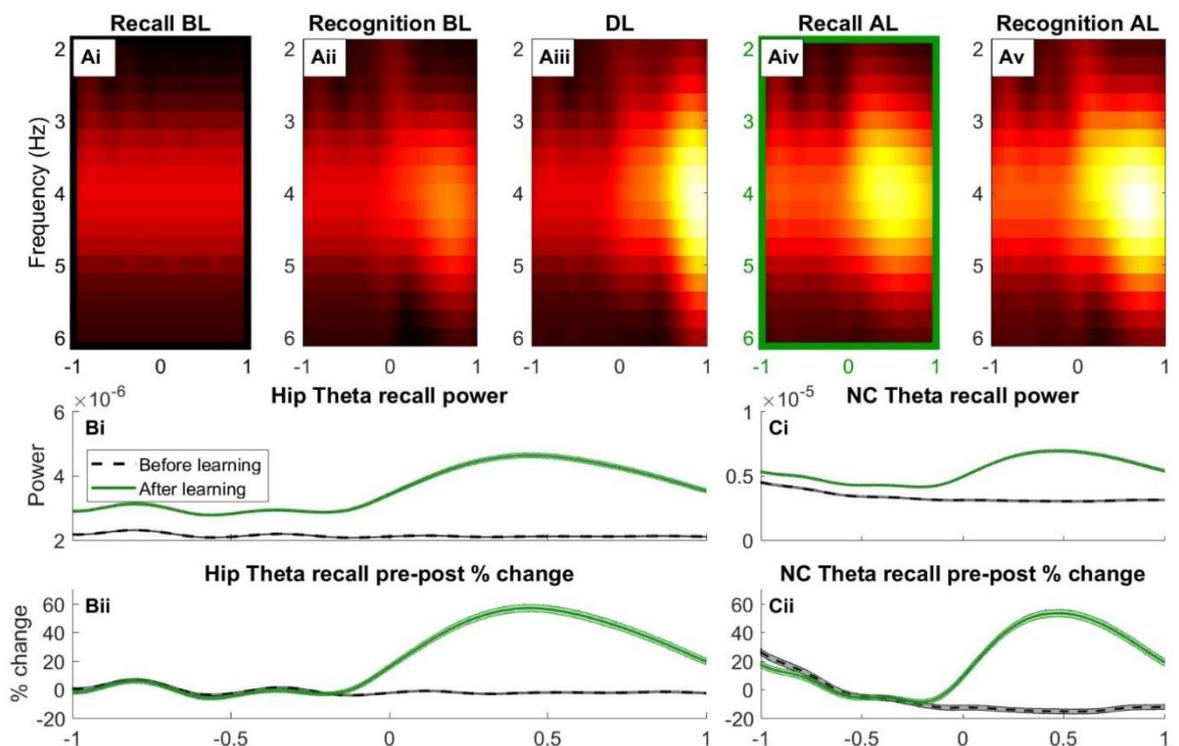


Figure 49 – Time-frequency-analysis (TFA) of hippocampal theta for the recall and recognition conditions before and after learning (**Ai-ii, Aiv-v**), as well as for during learning (**Aiii**). A time-course of theta power is shown (**B**) for the colour-coded highlighted boxes (**Ai, Aiv**), where raw power (**Bi**) and percent change in pre-post stimulus power (**Bii**) are shown. The same analysis is shown for neocortical theta power during the same time periods (**Ci-ii**).

Theta synchronisation is stronger during learning, consistent with experimental evidence (Backus, et al., 2016; Lega, et al., 2012; Staudigl & Hanslmayr, 2013). This is due to the rapid increase in synaptic weights during this period (Figure 49B; 10 to 12s) causing feedback activation, which, in turn, causes more neurons to fire above threshold, but according to the theta rhythm. After the learning phase, neurons are also responsive to the opposite image, where a synchronisation of theta occurs due to an increase in activity post stimulus (Figure 49Aiv). This can be seen more clearly in Figure 49Bii, where there is up to a 60% increase in theta power relative to the pre-stimulus period. Due to stronger weights between the P & NP cluster, there is increased feedback activity during the normal oscillatory rhythm. This activity is amplified by a higher synaptic time constant ($\tau_s = 5\text{ms}$) for hippocampal neurons, causing an increase in pre-stimulus theta power (Figure 49Bi; -1 to 0s). The same changes in theta power are passed through to the NC (Figure 49Ci-ii), which is consistent with experimental evidence of increases of theta in NC areas after learning paradigm experiments (Burke, et al., 2014; Klimesch, et al., 2005).

6.1.6 VARYING STIMULUS STRENGTH

We next varied how strongly our simulated participant perceived the P & NP images during the encoding and recall after learning conditions, allowing us to explore the sync/de-sync of hippocampal theta and NC alpha over time at different strengths. This is achieved by varying stimulus strength, i.e. the rate of spikes per second being fed into NC neurons at stimulus onset, and taking the average power during the post-stimulus period across frequencies (0-30Hz). This information is displayed as heatmaps of frequency vs stimulus strength (Figure 50Ai-ii & Di-ii), where stimulus strength is shown on a logarithmic scale from 10^0 to 10^6 . We can extract from this information to show the evolution of NC alpha (Figure 50B; Red, 8-12Hz) and hippocampal theta (Blue, 3-5Hz) as neurons are driven more. It can be shown that for weakly perceived stimuli, the NC actually synchronises in alpha within the model (see around 10^3 strength). This is due to input activity being too weak to overcome the trough of the 10Hz sine input, but strong enough to cause more spiking in the peak.

As stimulus strength increases, a de-synchronisation of alpha is obtained as neurons overcome inhibition to spike across all phases of alpha (see around 10^5 strength). In contrast, the hippocampus exhibits a strong synchronisation of 4Hz (Figure 50B) with increasing stimulus strength. This is due to the ADP function preventing neurons recovering quickly after a spike event.

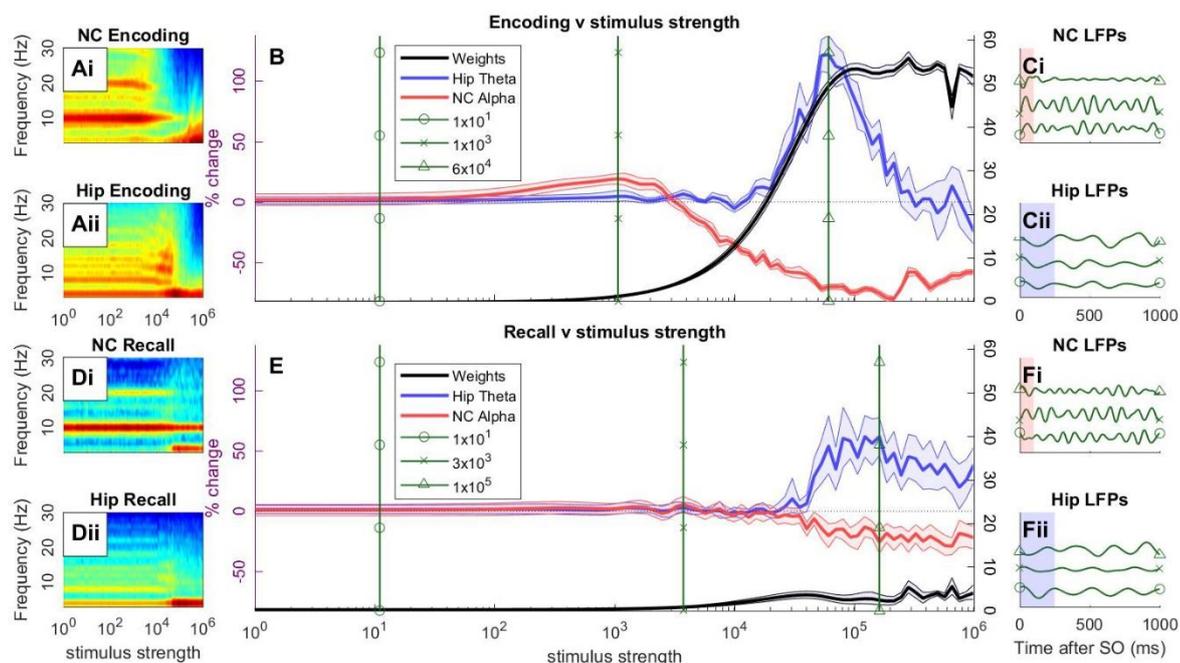


Figure 50 – Increasing stimulus strength (number of spikes being fed into NC neurons) during the encoding (DL) and recall after learning conditions, where stimulus strength is depicted on a logarithmic scale. During the encoding stage (A-C), frequency by strength heatmaps of raw power at NC (Ai) and hippocampus (Aii) are shown. From this data, relative changes in NC alpha (B; red, 8-12Hz) and hippocampal theta power (B; blue, 3-5Hz) are plotted, as well as weight change between P and NP hippocampal subgroups (B; black). From this plot, three different stimulus strength values are chosen: normal oscillatory activity ($\sim 10^1$ strength), small alpha power increases ($\sim 10^3$ strength) and maximal theta power increases ($\sim 10^5$ strength). At these points, local-field-potentials (LFPs) are calculated using specific 2-6 or 8-12Hz filters for hippocampal theta (Cii) & NC alpha (Ci), respectively, where blue and red highlighted regions indicate the possible stimulus onset area due to re-aligning phases across multiple trials. The same symbols indicate at which point an LFP represents. The same format is applied for the recall after learning condition (D-F).

This then is an important difference between the neo-cortical and hippocampal systems, which underlies why (apart from with very strong inputs) the hippocampus synchronises rather than desynchronises – essentially the ADP function prevents the hippocampus from desynchronising. Weight change between P & NP units also increases monotonically with stimulus strength, plateauing at the same level that theta and alpha maximally sync/de-sync, respectively. This indicates why alpha de-synchronisation and theta synchronisation are both markers of successful memory encoding (Backus, et al., 2016; Lega, et al., 2012; Hanslmayr, et al., 2012; Staudigl & Hanslmayr, 2013). Hippocampal theta synchronisation can also be seen to bleed into NC neurons as stimulus strength increases (Figure 50Ai; 10^4 to 10^6 strength), corroborating experimental evidence (Burke, et al., 2014; Klimesch, et al., 2005).

When we push the model past normal levels of activation (the model's default is $\sim 8 \times 10^4$), hippocampal theta eventually de-synchronises, indicating that although the ADP function essentially acts as a break on hippocampal units, it can eventually be overcome. Weight change remains high as units are spiking across all phases of theta. This gives a possible explanation for why some experimental evidence also finds a positive correlation with successful memory encoding and hippocampal theta de-synchronisation (Crespo-Garcia, et al., 2016; Greenberg, et al., 2015; Fellner, et al., 2016).

We also choose three important points from Figure 50B that best convey the model's sync/de-sync characteristics, indicated by vertical green lines during first normal oscillatory behaviour, second, alpha sync and third, maximal theta sync and alpha de-sync. The corresponding LFPs (indicated by the same symbol) are shown for these three points for NC (Figure 50Ci) and hippocampal units (Cii). NC alpha LFPs show how power can increase when more spikes during the excitatory phase cause larger amplitudes of activity (Ci; cross), and how power decreases when activation occurs throughout an oscillation (Ci; triangle). Similarly, hippocampal theta LFPs show how power can increase with increased activation in the peaks, despite the low-level activation in the trough (Cii; triangle) that is responsible for learning.

The same analysis has been performed for the recall condition after learning, with similar results. Importantly, the method of de-synchronisation is different in this condition. As Figure 50Di shows, in the NC an alpha de-sync at recall is accompanied by a theta sync, indicating that alpha is de-synced by theta as activation feeds into the hippocampus, which in turn feeds activation back to the NC. This ensures we do not see a small synchronisation of alpha with low levels of stimulus strength as we saw in the encoding condition. As theta and alpha phases are rarely aligned (as seen by comparing LFP plots in Figures 50Fi-ii), maximal theta excitability is just as likely to de-synchronise by occurring during an alpha inhibitory phase as it is to be facilitated by aligning with an alpha excitatory phase. As stimulus strength increases, one observes both hippocampal theta synchronisation and NC alpha desynchronisation accordingly, indicating that both are important for successful memory retrieval. Figure 50E shows that the model is able to exhibit re-instantiation of a memory's content. That is, neo-cortical alpha desynchronises during recall for the stimulus cued, but not presented. This represents a purely endogenous activation of rich content.

The variation of stimulus strength can also be alternatively visualised (Figures 51/52), whereby one can see the sync/desync effect across the whole 4-30Hz frequency range. Here, the behaviour of cells encoding for the P stimulus were observed during the screening phases of the simulation, where the P & NP stimuli are presented individually before and after learning. The default behaviour for hippocampal theta (Figure 66B) and NC alpha (Figure 67B) can be seen when the NP stimulus is presented before learning, as the P units do not respond to anything but the P stimulus at this phase of the simulation. Hippocampal theta and NC alpha can be seen to repeat weakly at frequencies in increasing orders of magnitude, i.e. the harmonics: 10/20Hz for alpha and 4/8/12/16Hz for theta.

When considering stimulus strength against frequency power in active hippocampal populations, there is a power increase around 4Hz (Figure 51A) with increasing stimulus strength, as we have shown in Figure 50. This is stronger for learnt items, where synaptic weights are stronger (Figures 51C/D). Interestingly, as stimulus strength increases, there is a diagonal band of increasing power

across frequency space. This suggests that the frequency of hippocampal activity increases linearly with stimulus strength, indicating that the up-state of the theta frequency sine wave is broad enough to enable multiple spike event patterns within it. For example, a spike doublet might come out as an 8Hz frequency, whilst spike triplets might be recorded as a 12Hz frequency. This might lend some theoretical evidence for the gamma/theta coupling observed during memory tasks (Burke, et al., 2014; Ison, et al., 2015), where high frequency activation during a theta peak is recorded as a gamma oscillation. In the case of this model, no such gamma mechanism is realised, so we see a gradual increase in the frequency of activity within theta peaks, when experimentally one might only observe activation to occur in gamma frequencies due to biophysical circuitry.

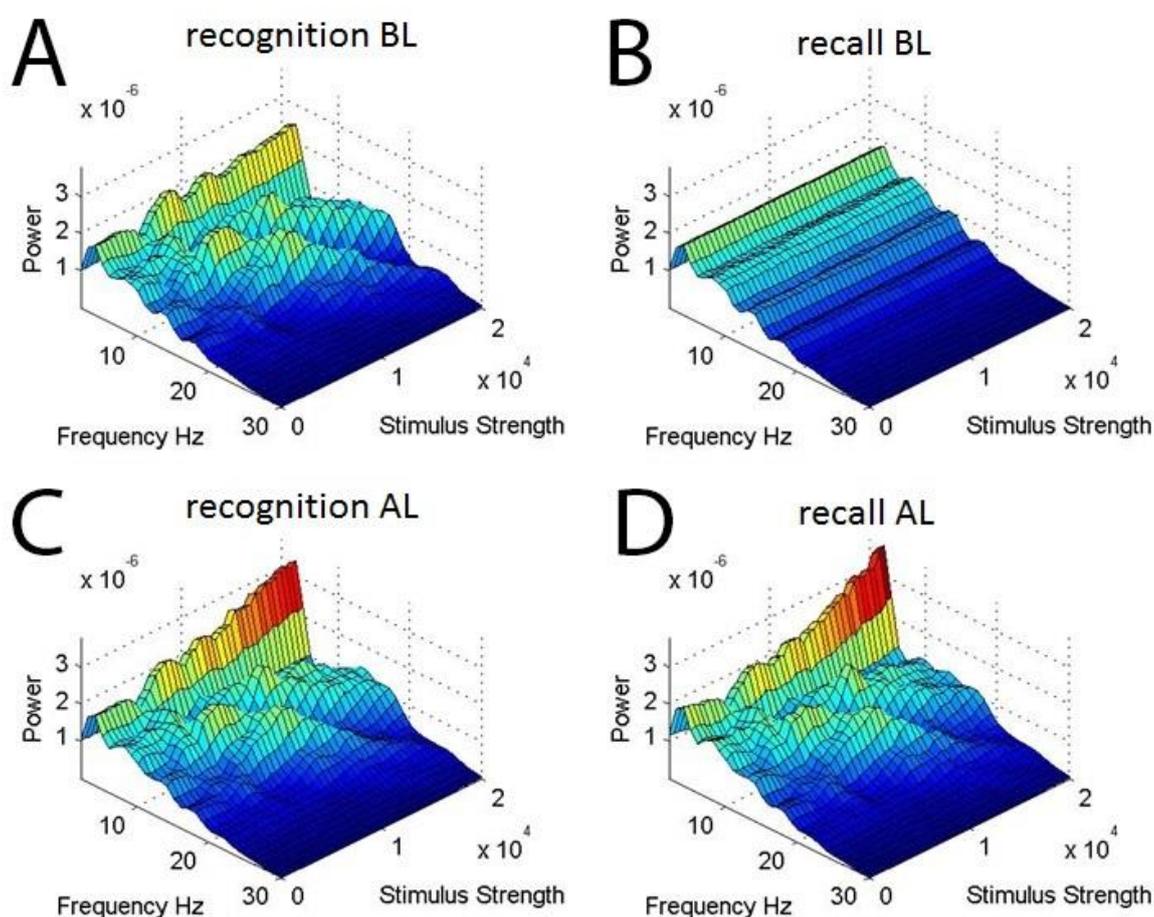


Figure 51 – 4-30Hz frequency power against stimulus strength for hippocampal units at different stages of the simulation, namely, the recognition condition before (A) and after (C) learning and the recall condition before (B) and after (D) learning.

In neo-cortical populations, the same behaviour is shown as is witnessed in Figure 50, where alpha power increases with weak stimulus strength (Figures 52A/C) before desynchronising with increasing stimulus firing rates. Interestingly, as stimulus strength increases and therefore, feedback interactions from the hippocampus increase, the same diagonal band of power increases across frequency space is visible. This pattern is most evident in the recall condition (Figure 52D) due to the fact that those NC units only receive input from hippocampal interactions and not from the source of the stimulus itself. This indicates that the hippocampus is successfully driving the neo-cortex, thus the same activity patterns emerge in neo-cortical sites. This influence also reduces alpha 10Hz power, increasing with larger amounts of stimulus strength.

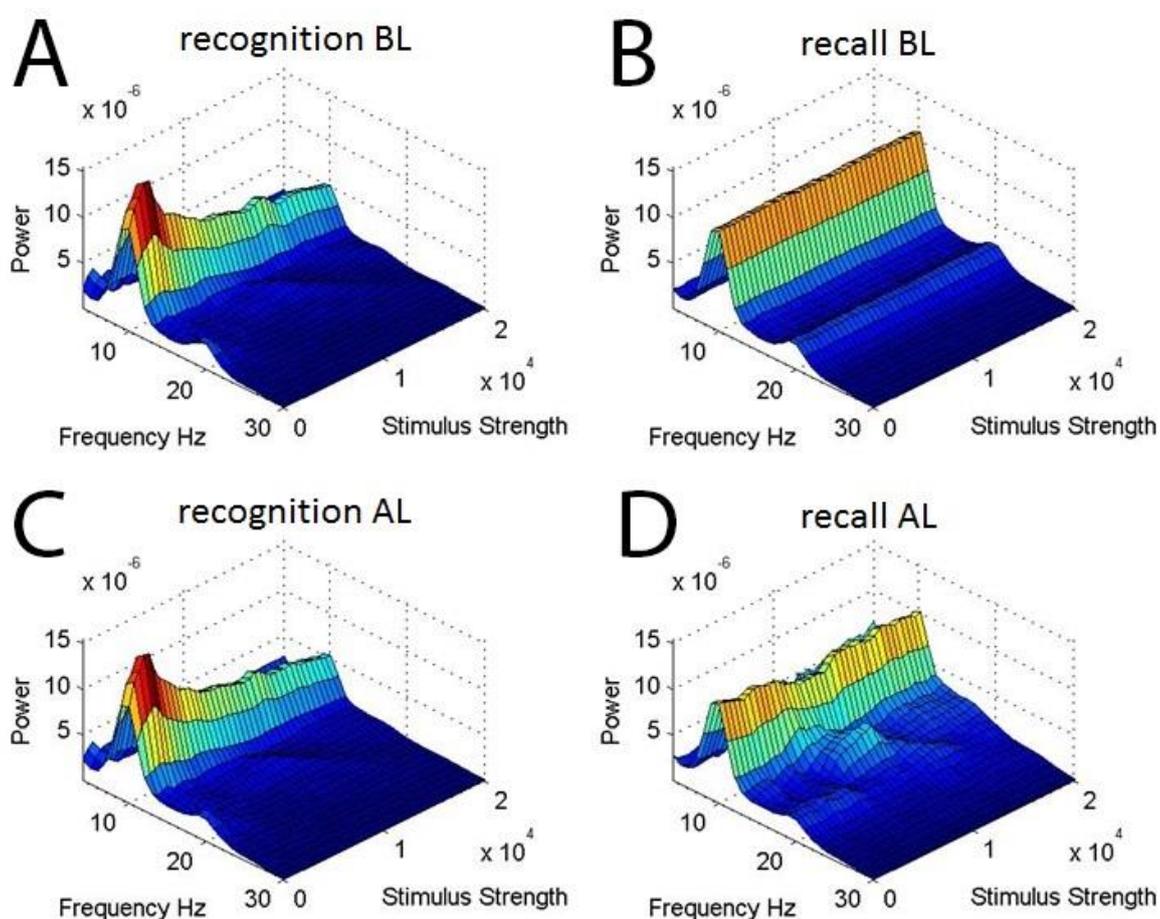


Figure 52 – 4-30Hz frequency power against stimulus strength for neo-cortical units at different stages of the simulation, namely, the recognition condition before (A) and after (C) learning and the recall condition before (B) and after (D) learning.

6.1.7 SYNCH/DE-SYNCH PREDICTS LEARNING

Having demonstrated that our model mimics the behaviour of alpha power decreases in the NC, and theta power increases and phase dynamics in the hippocampus, we now link these contrasting synchronisation behaviours with learning (Figure 53). By varying the learning rate of STDP weight change (ϵ) between 0-1, it was possible to assess how the model behaves with different learning outcomes. The average of all bi-directional hippocampal weights between subgroups P & NP increased with ϵ (Figure 53C), which is used here to assess learning, i.e. the stronger the weight change, the better the memory. We then calculate the effectiveness of recall (P response to NP + NP response to P) as a percent change in power at a particular frequency from before learning to after learning, effectively allowing us to isolate the effect of learning on power. A bootstrap procedure then provided the confidence intervals (shaded area) around a mean (solid line) of recall power for incremental values of ϵ for pre-stimulus (black) and post-stimulus (red) periods.

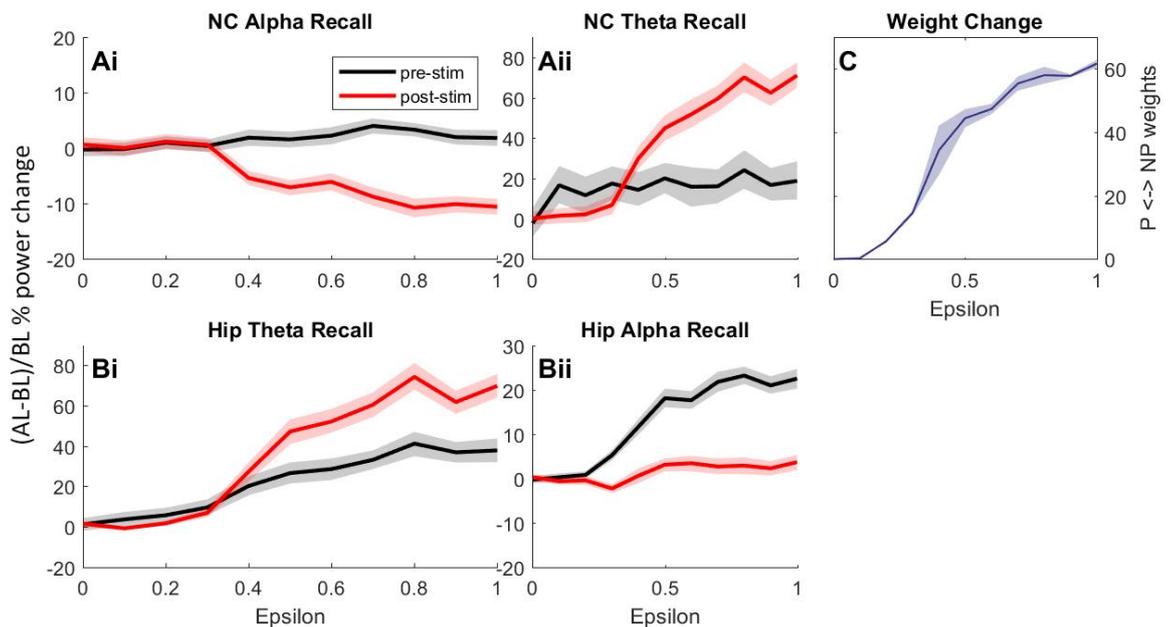


Figure 53 – The effect of increasing the learning rate (ϵ), and therefore synaptic efficacy between P and NP subgroups, on NC alpha power (**Ai**), hippocampal theta power (**Bi**), NC theta power (**Aii**) and hippocampal alpha power (**Bii**). **C** plots the mean and variance of P <-> NP weights from 1000 simulations, where the learning rate (ϵ) was incremented gradually from 0 to 1.

From this we can use power at a particular frequency to predict whether learning has successfully occurred in our model, and vice versa. In respect of the sync/de-sync theory (Hanslmayr, et al., 2016), the model indicates that both a de-synchronisation of alpha in NC areas (Figure 53Ai) and a synchronisation of theta in hippocampal areas (Figure 53Bi) during recall can predict successful memory retrieval.

Interestingly, one could also look at pre-stimulus theta and alpha power in the hippocampus to predict whether learning has occurred (Figure 53Bi-ii; black), where both increase by 30-40% due to stronger weights within the hippocampus and reciprocal connectivity between the hippocampus and NC. This is consistent with evidence that reports the importance of pre-stimulus theta for learning (Fell, et al., 2011; Gyderian, et al., 2009). The effect of feedback activity plays a smaller role in NC areas, where a small increase (<5%) in pre-stimulus alpha power (Figure 53Ai; black) and an increase (<20%) in pre-stimulus theta power (Figure 53Aii; black) can also predict learning (Salari & Rose, 2016). Importantly, there is a large synchronisation of theta (<70%) at recall (Figure 53Aii; red) in NC areas, consistent with experimental findings (Burke, et al., 2014; Klimesch, et al., 2005).

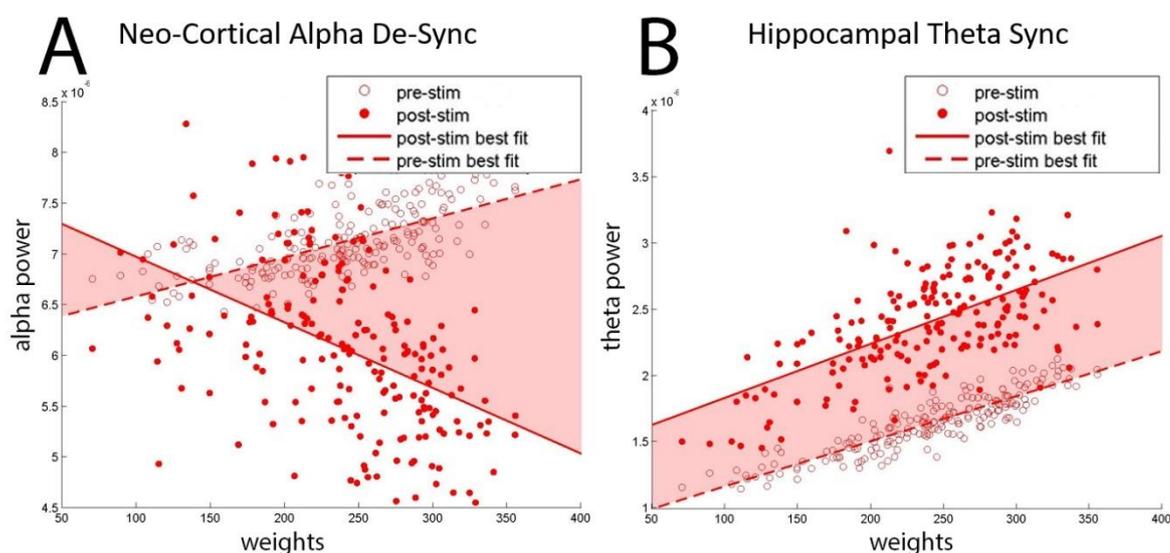


Figure 54 – The sync/desync model predicts learning. **A** Neo-cortical alpha de-sync and **B** hippocampal theta sync, both plotted as a function of raw weight strength between P->NP subgroups.

These findings are re-displayed in an alternate graph (Figure 54), which further visualises the concept behind the sync/desync model. Here, weights between hippocampal subgroups are plotted as a function of power at recall, i.e. the activation of P units in response to the NP stimulus after they have been associated together. NC pre-stim alpha power increases linearly with the strength of this association (Figure 54A), where a linear line of best fit indicates this positive correlation. Equally, the line of best fit for NC post-stim alpha power shows a negative correlation with learning. Together, these intersecting lines show that both an increase in pre-stim and a decrease in post-stim alpha power are good predictors for memory, potentially showing that increased pre-stimulus alpha represents a heightened ready-to-fire state that helps pattern completion, and therefore memory recollection (Klimesch, et al., 2007). Alternatively, hippocampal pre and post-stim theta increase linearly with learning (Figure 54B), indicating how increased theta power is a good indicator for successful memory encoding and retrieval (Fell, et al., 2011; Gyderian, et al., 2009; Hanslmayr, et al., 2016).

6.2 DISCUSSION

We have presented a relatively simple spiking neural network model, which captures the complex synchronising and desynchronising behaviours of hippocampus and neocortex during encoding and retrieval in a typical memory task. This model, which we term the Sync/deSync (SdS) model, simulates hippocampal theta synchronisation and neocortical alpha desynchronisation in the service of encoding and retrieving novel stimulus associations – a key requirement of episodic memory. Consistent with the notion that one-shot learning occurs in the hippocampus, but not in the neocortex (O'Reilly, et al., 2011), our model only implements synaptic modifications in the hippocampus. This hippocampal learning uses two well-described synaptic modification mechanisms. The first is spike-timing-dependent-plasticity (Song, et al., 2000), where synaptic modifications increase with decreasing time lag between the firing of pre and post-synaptic neurons. The second mechanism is theta phase-dependent plasticity, where synapses between neurons firing in the inhibitory phase of theta are strengthened, whereas synaptic connections

between neurons firing in the excitatory phase are weakened (Hasselmo, 2005). In the model neocortex, neurons fire phase-locked to an alpha oscillation when they receive no input (Jensen & Mazaheri, 2010; Klimesch, et al., 2007). When these neurons are driven by a stimulus, they increase their firing rate and gradually desynchronise from the ongoing alpha, especially when the input is strong enough to overcome maximum inhibition. Therefore, alpha power decrease is negatively related to the neural firing rate (apart from the small power increase at low stimulus intensities), thereby mimicking the well-known negative relationship between alpha and neural firing (Haegens, et al., 2011).

6.2.1 RELEVANCE TO OTHER MODELS

The Sync/deSync model draws inspiration from and resonates with a number of previous models that incorporate oscillations into the complementary learning systems framework. In particular, the concept of theta phase-dependent plasticity in the hippocampus has inspired aspects of a number of influential neural models (Hasselmo, et al., 2002; Ketz, et al., 2013; Norman, et al., 2005). An important component in two of these models (Hasselmo et al., 2005; Ketz et al., 2013) is a phase reversal between the two pathways from entorhinal cortex to CA1 (the monosynaptic performant pathway and the tri-synaptic pathway, via the Schaffer collaterals), which could provide a powerful mechanism in terms of separating encoding from retrieval cycles, as discussed in Chapter 4.2. We chose not to fully model this aspect in detail, but focused particularly on the dynamics in area CA1 in order to keep the model as simple as possible. Norman et al. (2005) present an important refinement of the basic complementary learning systems model, in which the strength of k Winner-Take-All (kWTA) inhibition is varied across theta phases. This modulation of inhibition provides a theta-phase dependent learning, with parallels to the Sync/deSync model. That is, in the Norman et al. (2005) model, the high inhibition phase of theta generates selective activation, restricting above-threshold activation to strongly responding units. LTP is then applied just to the active units, enabling selective weight update. This has similarities to the Sync/deSync idea that strongly active

units move their spiking forward in the phase of theta, enabling LTP (which only obtains in the inhibitory phase) to be selectively applied.

The match between the Norman et al and Sync/deSync models for the low inhibition phase of theta is a little weaker than for the high inhibition phase, but there are still parallels. Specifically, both models exhibit activation of a broader profile of units in the low inhibition phase. In the Norman et al model, this enables LTD to be applied to competitor units (that are not strongly tuned to the memory being encoded). Sync/deSync similarly applies LTD in this low inhibition phase, however, it is a non-specific, passive, decay.

Our use of an ADP function to reduce the capacity for units to spike multiple times in quick succession is inherited from the Jensen & Lisman (2005) model. Additionally, while advancing the phase of theta at which a unit spikes plays a key role in the Sync/deSync model, it is somewhat different to precession in the Jensen & Lisman model, where it encodes serial order.

6.2.2 RELEVANCE TO LITERATURE & PREDICTIONS

The Sync/deSync model is also able to capture a number of human electrophysiological findings. Human single neuron recordings revealed that hippocampal neurons can change their tuning, by showing an increase in firing rate to a non-preferred stimulus after this stimulus has been associated with a preferred stimulus (Ison, et al., 2015). Furthermore, Rutishauser et al. (2010) showed that a significant portion of neurons in the MTL are phase-locked to the ongoing theta rhythm during memory encoding, with an increase in theta phase-locking predicting later memory performance. Our model is consistent with these findings in showing an increase in activation for newly associated neurons, these responses being theta phase-locked, and increased theta synchronicity to be related to later memory performance. However, Sync/deSync also suggests that responsive neurons during learning are less locked to the ongoing theta phase (Figure 47A and B), which seems at odds with Rutishauser et al. (2010). This decrease in theta phase-locking is present for responsive neurons only, occurring since these units overcome maximum inhibition and thus

fire at the plastic phase of theta. Importantly, Rutishauser et al. (2010) did not separate neurons into stimulus responsive (i.e. showing an increase in firing rate) or not, therefore these findings cannot be directly linked to our model. However, an interesting prediction that arises from the model is that the preferred phase of firing differs between responsive and non-responsive neurons, and that this phase difference is related to later memory performance. Indeed, Rutishauser et al. (2010) found that different neurons were locked to different phases of ongoing theta. In our model, this difference is most prominent when only the first spike occurring after maximum inhibition is considered, a specific prediction that can be tested in future experiments.

Inherent to the SdS model is that the same neurons can be either synchronised or de-synchronised depending upon the strength of driving input. By gradually increasing stimulus strength, a population with more inhibition/slower integration can exhibit a synchronisation at stimulus strengths when faster spiking populations exhibit a de-synchronisation (Figure 50B; $\sim 10^5$ strength). This provides a neat explanation for the Sync/deSync conundrum, suggesting that it reflects the point where active neurons in different brain regions are on their trajectory towards a ceiling firing rate. We show in Figure 50B that the slower spiking hippocampal population synchronises with normal levels of input ($\sim 10^5$), but will eventually de-synchronise ($\sim 10^6$). In fact, non-invasive studies in humans have linked successful encoding of stimulus associations in the MTL with both theta power increases (Backus, et al., 2016; Kaplan, et al., 2012; Staudigl & Hanslmayr, 2013), and decreases (Crespo-Garcia, et al., 2016; Greenberg, et al., 2015; Fellner, et al., 2016). SdS indicates that both eventualities could yield successful memory encoding (Figure 50B; black line & blue line, which is trending negative at the top range of stimulus strengths).

With respect to alpha, many studies have shown that a decrease in alpha power coincides with successful encoding and retrieval of episodic memories (Hanslmayr, et al., 2012; Hanslmayr & Staudigl, 2014). In most previous studies, these effects extend also to beta. For this reason, and to ensure model simplicity, we have assumed only one cortical alpha rhythm, we, though, see no reason why the same principles would not also apply to beta. During successful encoding of episodic

memories, alpha/beta power decreases have been found in left frontal areas for verbal material (Hanslmayr, et al., 2009; Hanslmayr, et al., 2011a; Meeuwissen, et al., 2011) and occipital for visual material (Noh, et al., 2014). During retrieval, alpha/beta power decreases indicate the areas that are being reactivated, i.e. house the memory representation (Khader & Rosler, 2011; Michelmann, et al., 2016; Waldhauser, et al., 2016). This targeted alpha/beta power decrease is exactly what is modelled here, with only neural assemblies that actively process the stimulus during encoding or retrieval showing power decreases, and the degree of this power decrease predicting memory performance. A key element of formal modelling is the identification of predictions that give the opportunity for the model to be falsified. The key predictions that SdS makes are presented in Figure 50B, which shows that as driving stimulus strength increases, neo-cortical alpha goes through an initial phase, (strength around 10^3), of alpha power increase (i.e. synchronisation), followed by a much more marked alpha power decrease (i.e. desynchronisation), which is maximal just below a strength of 10^5 . This pattern could be argued to be inherent to the way synchronisation and desynchronization are modelled, i.e. a small increase in drive will generate more spikes at an oscillation's peak, and power will increase, while a large drive will cause spiking during the trough of the oscillation and power will go down. This pattern is our main prediction.

A further prediction is that the degree of alpha power decrease should correlate with the degree of hippocampal theta power increase, and the degree of phase precession of responsive neurons in the hippocampus. This prediction can be tested in intracranial EEG, which often records simultaneously from the neocortex and the hippocampus.

6.2.3 FUTURE WORKS

This model has explored the Sync/deSync conundrum using constant sine waves to model our oscillations. It is known that there are varied sources of rhythm and current generation in hippocampal (Buzsaki, 2002) and cortical (Jones, et al., 2000; Wang, 2010) populations, from excitatory-inhibitory interactions to membrane resonance within single cells. It would therefore be

prudent to consider these other oscillatory sources for both cortical alpha generation and hippocampal theta generation if we are to consider our findings robust. In this way, one could also consider the cooperative dynamics of event-related synchronisation and desynchronisation (ERS/ERD) in cortical populations, by creating many cortical columns with an intermediate thalamic pacemaker to encode phase synchrony with zero lag, as shown by other models (Vicente, et al., 2008). By using such dynamic oscillatory rhythm generators, one could further examine the phase synchrony relationships between un-related and associated objects at recall, perhaps making key predictions for future experiments.

This Chapter revisits previously described notions of time-keeping in the brain, specifically on the existence of multiple time-keepers that were alluded to in Chapter 3 (Mauk & Buonomano, 2004). We previously gave a brief overview of episodic memory and time cells in Chapter 3 (Eichenbaum, 2014), specifically on several feed-forward computational solutions (Goldman, 2009; Itskov, et al., 2011). Here, we expand these solutions in order to examine possible hierarchical relationships between independent forms of time-keeping, such as the centralised circadian rhythm (Hofstra & de Weerd, 2008) and the more distributed cortical episodic memory (Mauk & Buonomano, 2004), which might contextualise one-another through specific neuronal pathways (Aggleton, et al., 2010; Eichenbaum, et al., 2007). We here overview these distinct time-keepers such that we can offer a theoretical framework, whereby a neural substrate mechanism might enable the contextualisation of multiple and simultaneous hierarchical neural time-keepers.

In the model, a hierarchy of interacting feed-forward time-keepers is used to encode an episodic memory paradigm, where it was found that a temporal signature of memories can be identified through the phase-reset patterns of on-going cortical alpha oscillations (Michelmann, et al., 2016). To replicate these findings, it was important to instantiate a complimentary learning systems (CLS) framework (O'Reilly, et al., 2011), where discrete events were bound to a moment in time. We also produce a separate network that is dynamically entrained in an alpha frequency, as discussed in Chapter 4, where event-driven activation causes a phase-reset to occur. By binding each event as a discrete moment, similar to other models (Bowman & Wyble, 2007), then recalling them in temporal order, we show how one can detect the encoding and retrieval of unique temporal signatures within the brain. As such, we hope to provide computational evidence for the findings from Michelmann, et al. (2016) and for our primary hypothesis, that cortical desynchronisations can indicate information flow and can be used to decipher information content. We also hope to support our secondary hypothesis by creating complimentary binding and temporal processes, where learning is dependent on spike-timing relative to a hierarchical oscillator.

7.1 INTRODUCTION

As discussed in Chapter 3, there has been much work in recent years to understand how the brain represents ‘what’ and ‘when.’ With respect to the ‘when,’ competing notions of time-keeping at the systems level distinguish between the possibility that the brain generates a central timing mechanism and distributes it (Church, 1984), or that local subsystems might produce their own time-keeping (Mauk & Buonomano, 2004). A biophysical mechanism for the former is the circadian rhythm, the mammalian biological clock that consists of a hierarchy of oscillators, the master circadian pacemaker of which is formed of suprachiasmatic nuclei (SCN) within the anterior hypothalamus (see Hofstra & de Weerd, 2008, for review). This cycle moderates a host of bodily functions and also receives adjustments through entrainment by external cues. In this manner, seizures have been found to occur in specific temporal patterns mediated by the light/dark cycle (Gowers, 1885) and disorders of the circadian rhythm have been found to occur in the majority of epileptic patients (Hofstra & de Weerd, 2008). This leads us to the question as to whether there is a ‘termination’ operation occurring at slower temporal dimensions, where a fault in the distribution of time originating from a low frequency central timing mechanism leads to the spread of rampant higher frequency activity in the brain.

With respect to a more distributed form of time-keeping, episodic memory paradigms have demonstrated that humans can integrate experiences in time and space within the realm of milliseconds to 10s of seconds (Mauk & Buonomano, 2004). As discussed in Chapter 3, this is thought to be mediated by the hippocampus within the medial-temporal-lobe (MTL) (Howard & Eichenbaum, 2015; O'Reilly, et al., 2011; Squire, 1992; Squire, et al., 2004). Here, a gradually changing state of temporal context is thought to serve as a cue for episodic recall (Howard & Eichenbaum, 2013; Howard, et al., 2012), as theorised by other models of working memory (Burgess & Hitch, 1999; Shankar & Howard, 2012). The encoding of temporal sequences is thought to be enabled by the sequential activation of time cells within the hippocampus (see Eichenbaum, 2014, for review), which have been theorised to exist as a series of feed-forward chains (Goldman,

2009; Itskov, et al., 2011). Cortical areas might then act as a distributed store of long-term episodic memories (McClelland, et al., 1995; O'Reilly, et al., 2011), where sequences must be reliably encoded (Winocur, et al., 2010) and recalled (Eichenbaum, et al., 2007), possibly through the inheritance of hippocampal sequences through repeated exposure (Itskov, et al., 2011).

Feed-forward activation is thought to be the best enabler for fast communication with high temporal precision, thought necessary for the distributed time-keeping that occurs during sensory and motor events (Mauk & Buonomano, 2004), possibly by enabling target cells to conjunctively represent distinct elements of a stimulus (Fries, 2005). As we have shown in Chapter 2, models of such synfire chains have shown that this kind of signal transmission can operate within a noisy environment (Diesmann, et al., 1999), can co-exist within an embedding LCRN (Kumar, et al., 2008), and might even naturally emerge during activation within an LCRN with plastic connectivity (Fiete, et al., 2010). We further demonstrated in Chapter 3 that a feed-forward architecture can support time (Goldman, 2009; Itskov, et al., 2011), predicting that temporal sequences can be internally generated, reliable from trial to trial, context dependent and long lasting, in a manner similar to time cells (Eichenbaum, 2014). However, a common criticism of such a solution for time is that the length of the chain must increase linearly with the desired duration (Shankar & Howard, 2012). As discussed in Chapter 3, this complexity issue is compounded by the finding that there might be many competing chains within the hippocampus (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). A fundamental goal of this modelling project is to thus reduce the complexity of feed-forward synfire chains in the encoding of time.

It has been hypothesised that there is a competitive architecture underpinning spatial working memory, navigation and path integration in the hippocampus (see McNaughton, et al., 2006, for review). As discussed in Chapter 3, the same might be said for the timing aspect of episodic memory, where studies have indicated that the hippocampus holds many columns of sequences that it can successfully associate, disambiguate and navigate between (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). Competition is often modelled as short-range

excitatory and long-range inhibitory connections in a locally connected random network (LCRN). Together, this encourages winner-take-all behaviour, an architecture that has commonly been used in hierarchical models of vision, recognition and attention (Carpenter & Grossberg, 1987; Itti, et al., 1998; Reisenhuber & Poggio, 1999). A hypothesis of this modelling work is that any model of ordinal and temporal sequencing would enable competition through a winner-take-all architecture, whereby the emergence of a winning pathway maintains temporal and spatial context through the suppression of losing pathways.

With respect to the 'what' of episodic memory, we have discussed in previous Chapters how neural oscillations might play a key role in structuring our informational capacity (Fell & Axmacher, 2011); where it is thought that information flow depends on whether neural activation is restricted to the inhibitory phase of an entraining oscillation or not (Klimesch, et al., 2007; Hanslmayr, et al., 2012; Parish, et al., 2018). If this were the case, information representation would be measurable by oscillatory power changes in specific frequencies. We have previously discussed how alpha frequency de-synchronisations are thought to signify information processing in cortical regions, as de-regulated neural activation enables potential information gain (Hanslmayr, et al., 2012). We have also shown that successful episodic memory encoding and retrieval can be predicted by alpha power in the literature review of Chapter 4 (Fell, et al., 2011; Hanslmayr & Staudigl, 2014; Hanslmayr, et al., 2012; Khader, et al., 2010; Klimesch, et al., 2005; Waldhauser, et al., 2016), and in the modelling work of Chapters 5-6 (Parish, et al., 2018).

Whilst we have previously shown that alpha power decreases can signal high information states, we here relate the phase of on-going alpha to information content (Michelmann, et al., 2016; Ng, et al., 2013; Schyns, et al., 2011). To this effect, representational similarity analysis (RSA) has been used to detect the re-instatement of neural patterns in several studies (Ng, et al., 2013; Staresina, et al., 2012; Staudigl, et al., 2015; Wimber, et al., 2015). Whilst these previous studies focused on detection, a new study associated these temporal signatures with specific stimuli, enabling RSA to dynamically decode which temporal pattern is being actively reproduced by the brain, and when

(Michelmann, et al., 2016). It is the aim of this second modelling work to reproduce the findings of this episodic paradigm, both to showcase our method of temporal encoding and to provide theoretical evidence for the usage of RSA in the detection of unique dynamic stimuli.

7.2 THEORETICAL FRAMEWORK

Now that we have established how sequences might be encoded in the brain, we can consider a unifying theoretical framework (Figure 55). Here, we consider evidence from recognition-based memory studies when discussing possible interactions between the distinct thalamic, MTL and cortical time-keepers that have been described. Recognition memory is also facilitated by the MTL, which, it has been argued, can be split into the functionally distinct familiarity-based and recollection-based memory (Tulving & Donaldson, 1972).

The hippocampus is thought to be critical for recollection-based memory (Eichenbaum, et al., 2007), as it is important for contextualising events in time and space. Research has shown that the hippocampus contains many competing sequences that can be navigated and disambiguated (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000), possibly through a winner-take-all environment as is hypothesised to exist in spatial working memory (McNaughton, et al., 2006). Here it might be that temporal encoding is constrained to a serial order of events through the maximisation of a set of winning pathways, where irrelevant nodes at each time-step are inhibited, as indicated by simulations of LCRN models (Itskov, et al., 2011; Kumar, et al., 2008). As well as this, hippocampal-anterior thalamic pathways are also critical for recollection-based memory (Aggleton, et al., 2010), possibly helping to further contextualise events in regards to the centralised thalamic pace-maker of the brain (Hofstra & de Weerd, 2008). Assuming that neural substrates are partly responsible for these time-keepers, the termination of each thalamic period might trigger spreading activation to propagate multiplicatively down the many competing hippocampal chains. In a system with faulty inhibitory mechanisms, this could plausibly trigger epileptic seizures that occur in specific temporal patterns throughout the day (Gowers, 1885). The

same hypothesised neural substrates for thalamic time-keeping might also replay the day's events at a sped-up rate during REM sleep, as indicated by the increased rate of suprachiasmatic nuclei (SCN) activity at this time. This may, in turn, increase the risk of hippocampally induced seizures during sleep (see Hofstra & de Weerd, 2008, for review).

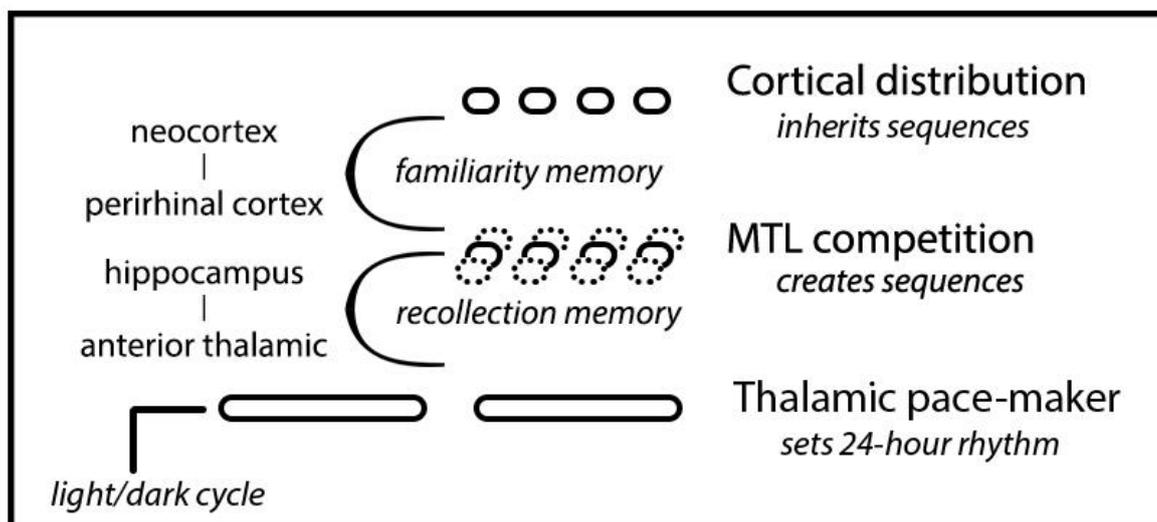


Figure 55 – Theoretical framework for temporal sequences. A thalamic pace-maker sets the circadian rhythm, initiated by the light/dark cycle and updated through external cues (Hofstra & de Weerd, 2008). Hippocampal-anterior thalamic connections have been found necessary for recollection-based memory (Aggleton, et al., 2010), where the hippocampus is thought to contextualise events in time & space (Eichenbaum, et al., 2007), possibly with updates from thalamic projections. The medial-temporal-lobe (MTL) contains many competing sequences that the brain can navigate in a competitive framework (Eichenbaum, 2014), enabling the creation and recognition of episodic sequences. Sequences might then be inherited through the perirhinal-cortical projections that are necessary for familiarity-based memory (Eichenbaum, et al., 2007), a process that might compare concurrently active sequences to generate a similarity measure.

Familiarity-based memory is thought to reflect a similarity measure generated through the comparison of concurrently active features (Eichenbaum, et al., 2007). This requires the perirhinal cortex (Eichenbaum, et al., 2007), which is thought to feed categorical information from the

inferior-temporal cortex (O'Reilly, et al., 2011). This pathway might allow active episodic sequences in the MTL to be compared with distributed sequences within the long-term memory store of the neocortex, producing a signal of general recognition (Eichenbaum, et al., 2007). This comparison might be achieved in a manner similar to that which we described earlier, where it was shown that a sparsely connected network could inherit temporal sequences through repeated exposure and an STDP mechanism (Itskov, et al., 2011).

We propose here a novel time-keeping neural network model with two main considerations. Firstly, we would like to address the argument that feed-forward synfire chains are unlikely to encode for temporal sequences, since their physical size has to scale linearly with the imposed time requirements (Shankar & Howard, 2012). We do this by instantiating a hierarchical synfire chain model, where sequential nodes exist as positive-feedback loops that maintain persistent activation in an LCRN framework. As such, our sequence of recurrent nodes behaves as a “feedforward network in disguise,” as described in Goldman (2009), and can reduce the physical requirements of encoding for longer temporal durations.

Secondly, we wish to propose a mechanism for a neural substrate through which multiple time-keepers with distinct temporal dimensions can interact with one another. To do this, we envisage a simple cell assembly of Hodgkin & Huxley neurons that can be scaled up to encode for multiple, simultaneous and interacting temporal hierarchies. This cellular assembly exists as a compact unit with specific feedforward and feedback connections, initiating and terminating persistent activation upon the completion of hierarchical sequences. Thus, we instantiate a hierarchy of synfire chains, where completion of a lower-order, faster temporal dimension initiates the transmission of persistent activity to the next node in a higher-order, slower temporal dimension. As such, lower-order sequences repeat at every node of the higher-order sequence, satisfying our primary consideration, raised by Shankar & Howard (2012), by increasing the capacity of feedforward chains. A moment in time is then marked as the concurrent activation of multiple temporal positions on simultaneous hierarchies, as indicated in Figure 56. In this respect, our

purpose is not to create a model of interacting brain regions of thalamic or hippocampal origins, but merely to instantiate a theoretical neural mechanism through which multiple time-keepers might interact and synchronise, possibly through the aforementioned neurological pathways.

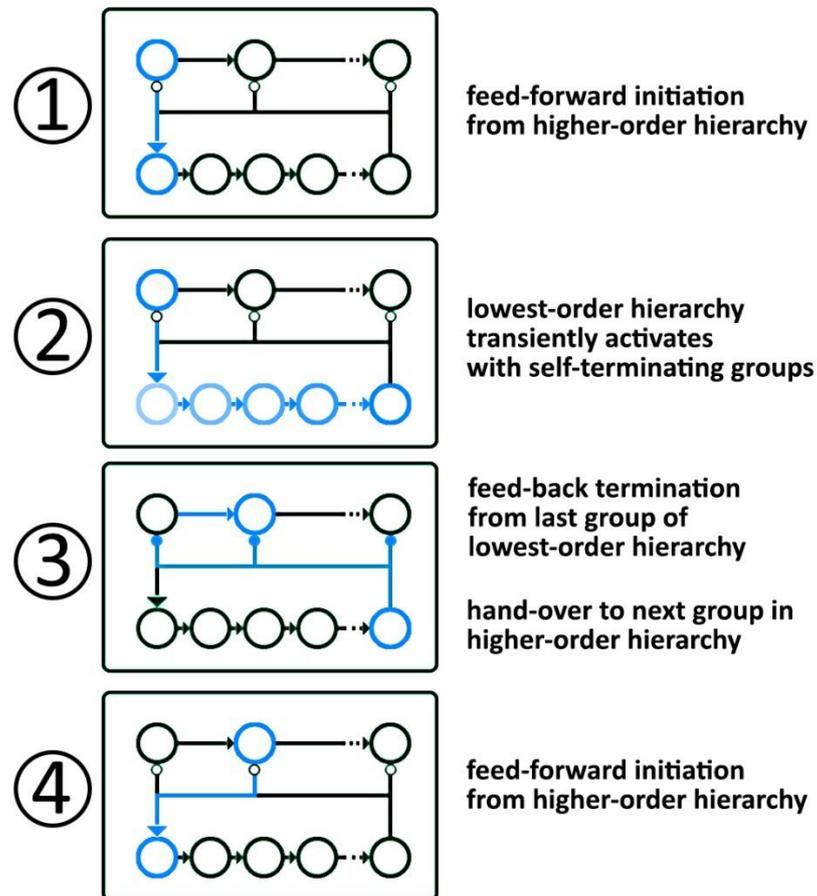


Figure 56 – Hierarchical updates between distinct time-keepers with varying temporal dimensions. Excitatory (lines with arrow-heads) and inhibitory (lines with circle-heads) connections dictate how neural assemblies (white circles) interact. Several temporal layers (vertical) simultaneously encode for the present moment. Transiently active groups (horizontal) pass activation through a chain. Upon completion of any lower-order chain (1st layer, bottom), a feedback signal is sent to any higher-order layers (2nd layer, top) with the effect of terminating persistent activation in the active group of that layer. This process is shown for the panels 1-4, where feedforward initiation from higher-to-lower order layers (1) and feedback termination from lower-to-higher order layers (3) organises feed-forward signal transmission. Active assemblies and connections are indicated by blue lines, just as inactive assemblies and connections are by green lines.

7.3 MODEL ARCHITECTURE

Our model comprises three distinct mechanisms, as shown in Figure 57, where we aim to reproduce an experiment that decodes temporal signatures from neocortical alpha phase patterns (Michelmann, et al., 2016). As such, information representation occurs in a neocortical area, where an alpha desynchronisation indicates information processing (Hanslmayr, et al., 2012), as described in the previous modelling work of this thesis (Parish, et al., 2018). In the current work, we create an intrinsic and dynamic alpha frequency (Figure 57; blue cortical region), such that sequences of stimuli can cause phase reset patterns to occur with some specific temporal pattern, thus conveying temporal information (Michelmann, et al., 2016; Ng, et al., 2013; Schyns, et al., 2011).

This information is then forwarded to a binding pool area (Figure 57; red region), where the occurrence of a stimulus is treated as a distinct event that activates a unique population of units in a winner-take-all environment. This mechanism has been thought to be necessary to allow for repetition, where any repeated stimulus can be re-represented by a unique tag (Bowman & Wyble, 2007), allowing us to differentiate between repeating items in a list. Binding pool units are associated to any active cortical units via a calcium dependent learning rule (Graupner & Brunel, 2012). As their local weights decrease during this process, which effectively models a form of hetero-synaptic long-term-depression (Volgushev, et al., 2016), active groups diminish the likelihood that they will be able to compete during successive events. Thus, active bindings are unique to the bound event. Concurrently active temporal units (Figure 57; green region) are also associated with binding pool units, thus sequencing the timing of events as they occur. Once the hierarchical time-keepers are re-started in a cued-recall paradigm (Figure 57; grey directional arrow), the relevant binding pool units become active at specific moments in time, causing a temporal pattern of events to be re-instantiated in the neo-cortex. We then use a representational-similarity-measure (RSA) to distinguish between multiple unique temporal sequences.

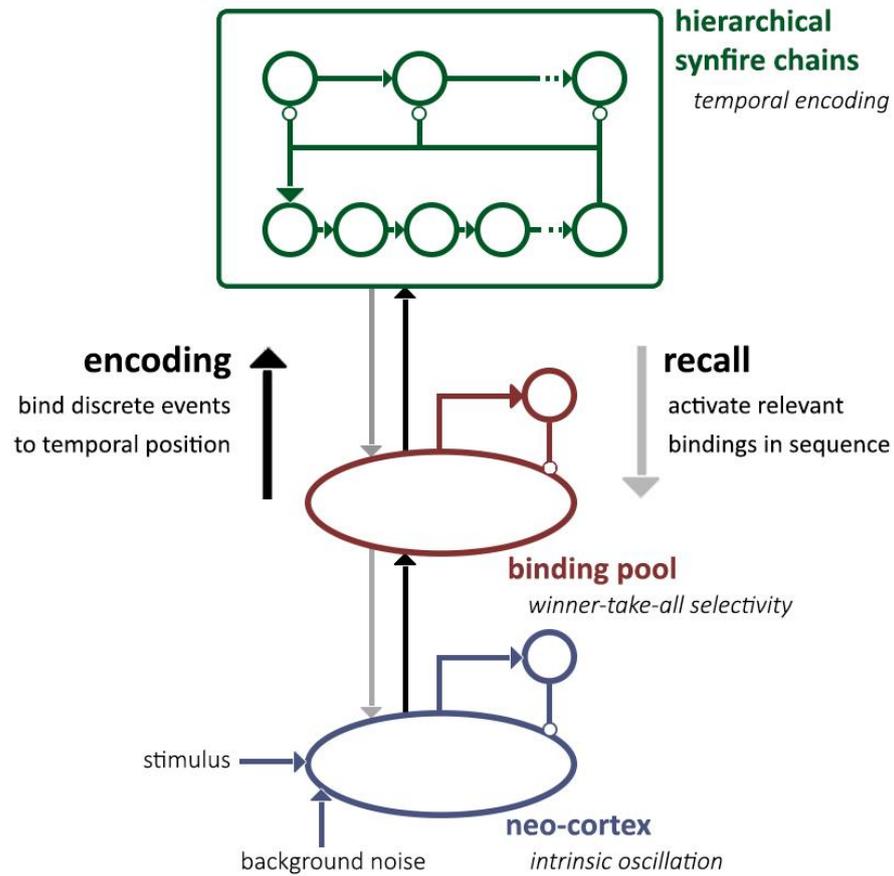


Figure 57 – Model architecture. A neo-cortical region (blue) intrinsically oscillates through excitatory-inhibitory interactions (lines with arrowheads, excitatory; circleheads, inhibitory). During encoding, incoming stimuli trigger activation in a binding pool (red), which encourages winner-take-all selectivity due to similar excitatory-inhibitory interactions. Active assemblies are then bound to the hierarchical synfire chains (green), which is a previously described (Figure 56) feed-forward structure used in the encoding of time. During recall, synfire chains are re-started and the relevant bindings are activated in sequence, causing observable phase-reset patterns to emerge in the intrinsically oscillating cortical region.

7.3.1 NEUROPHYSIOLOGY

Neurons in our model were simulated using Hodgkin-Huxley equations (as described in Chapter 2; Equations 4-13), whereby specific ion channels for sodium (Na), potassium (K) and the leak (L) were modelled and summed along with input from dendritic synapses (syn) and applied direct

current (DC). Constants for the conductance of each channel dictate the rate of change of that channel ($\bar{g}_{Na} = 100, \bar{g}_K = 80, \bar{g}_L = 0.1$), whilst the resting potential drives change in each channel dependent on the membrane potential ($E_{Na} = 50, E_K = -100, E_L = -67$). Voltage dependent gates dictate the degree to which each channel is open, whereby changes in the gates m, n and h over time were modelled with opening rates of $\alpha_x(V)$ and closing rates of $\beta_x(V)$ (Chapter 2; Equations 7-12). Spike events effect model neurons in the shape of an excitatory-post-synaptic-potential Alpha function (Chapter 2; Equation 20) multiplied by a weight (W), where the synaptic time constant (τ_s) dictates the amount of time it takes for the input to decay to 33% of its maximum value and t_{fire} is the time of the up-stream spike event. All spike events have a delay of 2ms to reach down-stream neurons.

Spike-time-dependent-plasticity (STDP) was enabled via an adapted calcium-based plasticity model (as described in Chapter 2; Equations 17-18 & adapted below in Equations 27-28; Graupner & Brunel, 2012). Synapses were represented by the variable $\rho(t)$ (Equation 27), which existed in a state between 0-1 and was multiplied by specific weight variables. Synapses gravitate towards being active ($\rho = 1$) or inactive ($\rho = 0$) once they pass an attractor state (ρ_*). Synapses changed dependent on whether the amount of calcium $c(t)$ at the post-synaptic buton was over specific thresholds for long-term-potiation (LTP) θ_p or long-term-depression (LTD) θ_d , whereby LTP occurred at a rate γ_p and LTD occurred at a rate γ_d (Equation 27; H denotes the Heaviside function, which returns 0 or 1 if the expression within $[]$ is below or above 0, respectively).

The amount of calcium at a synapse can be calculated by the summation of all spike events from pre (i) and post (j) synaptic neurons (Equation 28), where the Dirac delta function $\delta(t - t_i)$ takes the elapsed time since each spike event to capture its decayed amount. In this model, an additional constraint was added to reduce multiple-spike calcium increases (Equation 29), such that the function $f(t_i - t_{i-1})$ takes the time between spike events and returns an exponentially fitted value that punishes short periods. Calcium is then summed subject to a multiplicative constant C_{pre} for pre-synaptic spikes after a delay ($D = 13ms$), and a constant C_{post} for post-synaptic spikes.

Calcium also decayed exponentially over time ($\tau_{Ca} = 20ms$). See Chapter 2 for full overview of this calcium-based spike-time-dependent-plasticity rule.

$$\tau \frac{d\rho}{dt} = -\rho(1 - \rho)(\rho_* - \rho) + \gamma_p(1 - \rho) \cdot H[c(t) - \theta_p] - \gamma_d \rho \cdot H[c(t) - \theta_d]$$

Equation 27 – Synaptic efficacy dynamics in calcium-based STDP (Graupner & Brunel, 2012).

$$\frac{dc}{dt} = -\frac{c}{\tau_{Ca}} + C_{pre} \sum_i f(t_i - t_{i-1}) \cdot \delta(t - t_i - D) + C_{post} \sum_j f(t_j - t_{j-1}) \cdot \delta(t - t_j)$$

Equation 28 – Adapted calcium update equation.

$$f(x) = 0.1(1 - e^{-x})$$

Equation 29 – Adaptive calcium increase equation.

7.3.2 NEO-CORTICAL ALPHA

The neo-cortical region (NC) was modelled using a population of 35 neurons, 30% of which were inhibitory (I) and 70% excitatory (E). These neurons were connected using surround inhibition to enable a winner-take-all environment, as has commonly been used in hierarchical models of vision, recognition and attention (Carpenter & Grossberg, 1987; Itti, et al., 1998; Reisenhuber & Poggio, 1999). Excitatory neurons will oscillate in a steady state until some become more active, inhibiting their competitors in the process. To realise this, nearest-neighbour connections between two banks of neurons were implemented as a Gaussian distribution centred over each neuron, whereas farthest-neighbour connections were implemented as 1 minus a Gaussian centred over each neuron, where each distribution had a specific standard deviation (S.D.) width. STDP was not implemented in our neocortex, as in a complementary learning systems framework (O'Reilly, et al., 2011), it was assumed that learning occurs very slowly in this region. As our model is a proof of principle 'one-shot' learning paradigm, slow STDP was not implemented here.

In our neocortex (Figure 58), the excitatory bank of neurons connects to the inhibitory bank on a nearest neighbour basis ($S.D. = 1$, $\tau_s = 5$, $W = 1$), whilst the reverse inhibitory connections are made on a farthest neighbour basis ($S.D. = 1$, $\tau_s = 15$, $W = -1.25$). The excitatory bank self-connects to its nearest neighbours ($S.D. = 1$, $\tau_s = 1.5$, $W = 1$), whilst the inhibitory bank self-connects to its farthest neighbours ($S.D. = 1$, $\tau_s = 5$, $W = -1.25$). The lower τ_s for E \rightarrow E connections allows the activation to spread through the population, before the larger τ_s for E \rightarrow I connections activates the inhibitory bank, which in turn clamp down on all other E neurons. This causes oscillatory behaviour where the frequency is set by the τ_s for I \rightarrow E synapses (Brunel, 2000), thus here the frequency of idling activation is ~ 7 -8Hz, roughly approximating a resting alpha frequency.

Maintaining a steady state oscillation requires that E neurons are constantly driven to activate by low level noise. To do this, each excitatory neuron is fed by a different Poisson distribution of 1500 spike events per second ($\tau_s = 1.5$, $W = 0.05$), modelling activation from distant neurons and generating variation between neurons. As we described in our previous model of Chapters 5-6 (Parish, et al., 2018), the neo-cortical section of our model represents the flow of information, where the de-synchronisation and phase-angle time-series of alpha signals information processing (Hanslmayr, et al., 2012) and conveys information content (Canavier, 2015), respectively. Phase resets are thought to occur when there is a change in the scene during the presentation of dynamic stimuli (Michelmann, et al., 2016). In our model, scene changes are simulated as a strong pulse of Poisson distributed spike events (1000Hz, $\tau_s = 1.5$, $W = 3$), whose time course is multiplied by an Alpha function of $\tau_s = 60ms$ (Equation 15). This input is fed into a random NC unit, where an overlapping Gaussian in the spatial dimension effects nearby units ($S.D. = 1.5$). Upon activation of this select group, respective inhibitory neurons also activate due to the nearest neighbour connections. These inhibit all other excitatory and inhibitory units, thus promoting the unimpeded activation of the selected group until the event driven stimulus dissipates.

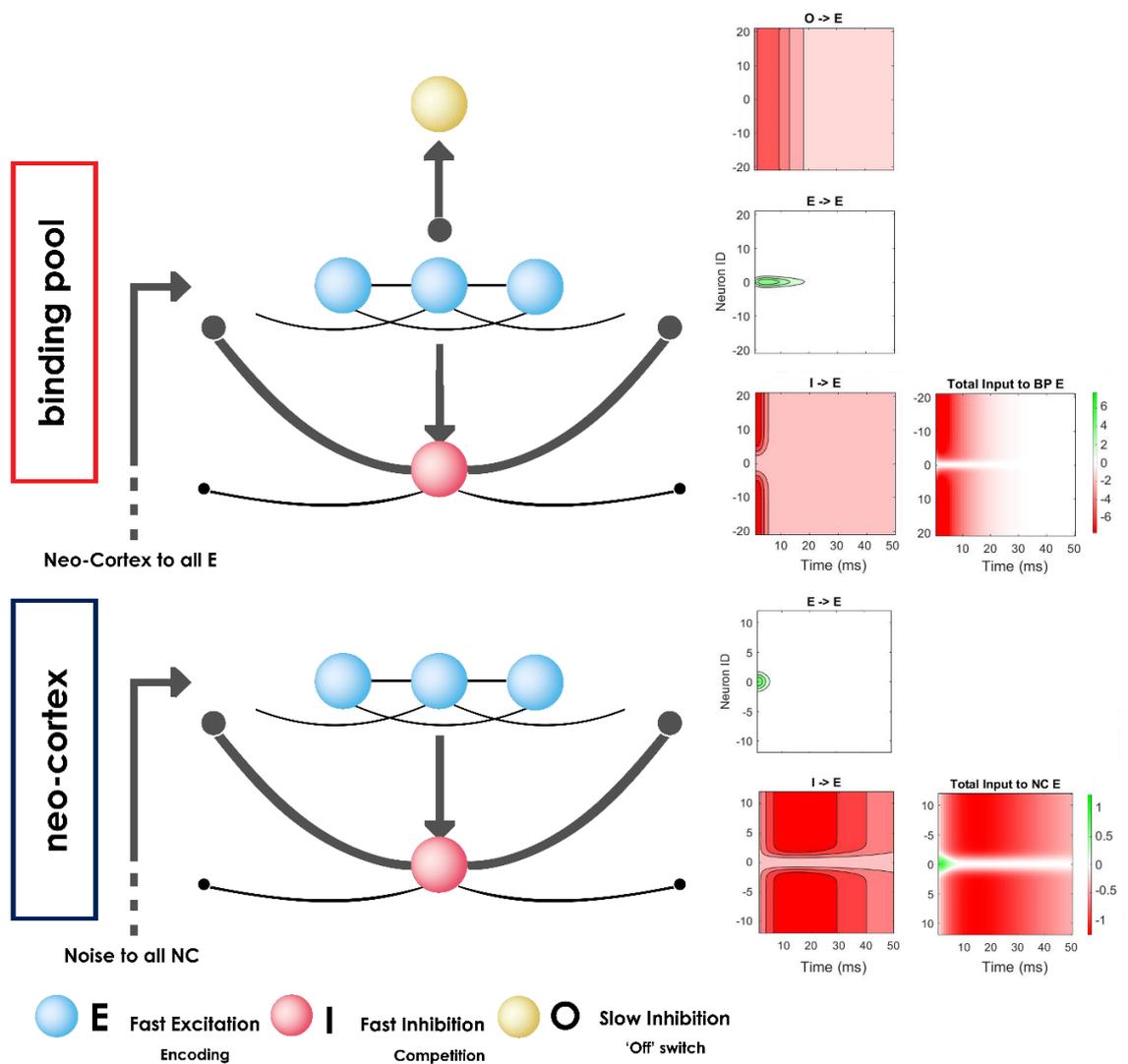


Figure 58 – Model architecture for the neo-cortex and binding pool areas. A surround inhibition architecture was used for both areas, which encouraged a winner-take-all behaviour via short-range excitatory and long-range inhibitory connections. Graphs in the right panels show the temporal and spatial response to an incoming spike event, which is the EPSP function (x-axis) multiplied by a weight (z-axis: green, positive; red, negative), from excitatory nodes (E; blue) and from fast (I; red) and slow (O; yellow) inhibitory nodes.

7.3.3 THE BINDING POOL

The binding pool (BP) was modelled using a population of 70 neurons, 30% of which are fast-inhibitory units (I), 10% slow-inhibitory units (O) and 60% excitatory (E). Like the neo-cortex,

surround inhibition was used here to promote a winner-take-all environment (Figure 58). Binding pool units are much more selective, however, remaining silent until an event occurs, upon which one group will become active for a potentially indefinite period. To do this, two additional mechanisms were implemented to promote selectivity and terminate activation. Firstly, STDP was added to synapses between binding pool units, such that weights go down between active units ($\theta_d = 1$, $\theta_p = 5$, $\gamma_p = 0.05$, $\gamma_d = 1$, $C_{pre} = 1$, $C_{post} = 1$). Here, calcium is prevented from crossing the high LTP threshold (θ_p), hovering between thresholds when successive spike events cause calcium to cross the LTD threshold (θ_d). When this occurs, weights go down at a relatively fast rate (γ_d). This weakens the active groups ability to participate in winner-take-all competition at successive events, creating a highly selective mechanism such that no one group of binding pool units will activate for any two events. Diminished internal weights also ensures that active neurons are less able to affect one another over time and activation within the group recedes. This mechanism effectively models heterosynaptic LTD, as observed experimentally and in other models (Volgushev, et al., 2016), which entails a balancing act where pre-existing synaptic pathways are diminished when new connections are strengthened through STDP. Alternatively, BP nodes could have been self-inhibited through sustained inhibition as in other models that describe a binding-pool (Bowman & Wyble, 2007). Despite the currently used mechanism, it is still possible for activation to spread through the network on occasion, disrupting the entire population. To this effect an 'off' switch (Figure 58; O units) was added, which is slow to activate but clamps down on the entire E population after a certain amount of activation has taken place.

A surround inhibition architecture was similarly realised in the BP (Figure 58), where the excitatory bank of neurons connects to the fast-inhibitory bank on a nearest neighbour basis ($S.D. = 0.5$, $\tau_s = 1.5$, $W = 8$), whilst the reverse fast-inhibitory connections are made on a farthest neighbour basis ($S.D. = 3$, $\tau_s = 1.5$, $W = -8$). The excitatory bank self-connects to its nearest neighbours ($S.D. = 1$, $\tau_s = 5$, $W = 5$), whilst the fast-inhibitory bank self-connects to its farthest neighbours ($S.D. = 1.5$, $\tau_s = 1.5$, $W = -8$). Here, the emphasis is placed on the low τ_s for E \rightarrow I and

recurrent I → E connections that allow competitive inhibition to spread quickly. A higher τ_s for E → E connections ensures that initial E activity can survive the subsequent inhibitory response, thus giving preference to the earliest spikers due to the speed at which they can inhibit their competitors. Weights are much higher in the binding pool than in the neo-cortex, as this was found to be necessary to allow swift winner-take-all competition to take place. As noted previously, slow inhibitory units were added to terminate spreading activation throughout the entire population. To this effect, every excitatory unit connects to every slow-inhibitory unit ($\tau_s = 25$, $W = 0.008$) and vice-versa ($\tau_s = 5$, $W = -5$). These parameters ensure that input for the O units builds up slowly over a relatively long period, but once activated they send an immediate inhibitory pulse, where the τ_s must be at least equal to that for E → E connections to mitigate any ongoing activation.

For events to be transmitted to the binding pool, the neo-cortical excitatory bank connects to the binding pool excitatory bank ($\tau_s = 15$, $W = 0.0035$). A large τ_s and low weight discourage normal cortical oscillatory activity from triggering the binding pool, while encouraging the build-up of input from quick successive spike-events that occur during cortical events. To maximise competition within the binding pool, randomness had to be implemented in NC → BP connections. To this effect, the banks connect via a random 15% of synapses, and weights vary around a normal distribution ($S.D. = 0.025$). In order for events to be bound, STDP was implemented on the reverse BP → NC connections ($\theta_d = 1$, $\theta_p = 1.3$, $\gamma_p = 1.5$, $\gamma_d = 0.05$, $C_{pre} = 0.8$, $C_{post} = 0.7$), where binding is encouraged with a fast LTP rate (γ_p) and slow LTD rate (γ_d). BP → NC connectivity was strong and long ($\tau_s = 60$, $W = 2$), to enable re-activation of NC events to occur on a similar time-scale as during encoding. Whilst this is an unrealistically large synaptic time-constant, this effectively models the accumulation of sustained activation through successive hippocampal-cortical pathways, e.g. presumed MTL pattern completion (O'Reilly, et al., 2011) and familiarity (Eichenbaum, et al., 2007) processes.

As shown in Figure 59B, hierarchical synfire chains (SC) are modelled as chains of sequentially connected groups (g) existing within hierarchical layers (l). Here, order denotes the position of the chain in the hierarchy, such that the slowest temporal dimension is termed the ‘highest-order’ chain and the fastest dimension the ‘lowest-order’ chain. Each group exists as a compact assembly of 8 units (Figure 59A), 12.5% of which are fast-inhibitory units (I), 12.5% of which are slow-inhibitory units (O), 12.5% are slow-excitatory units (P) and 62.5% are fast-excitatory units (E). We identify the need for two types of chains in Figure 59B; a lowest-order chain that must self-terminate due to it being the lowest-order hierarchy, denoted as $H1$; and higher-order chains that are terminated through feedback interactions with lower-order chains, denoted as Hn . As such, $H1$ chains have slightly different parameters to encourage the propagation of activation. Each group in an Hn chain receives a terminating signal from the combined culmination of every existing lower-order chain, upon which transient activity is propagated onto the next group.

Within each group, fast-excitatory units (E ; *blue units*) exist in a ring like structure, each unit connecting to its nearest neighbour (*blue to blue*; $S.D. = 0.5$, $\tau_s = 1.5$, $H1 W = 3$, $Hn W = 2.5$). These all connect to the fast-inhibitory unit (I ; *red units*) in the group (*blue to red*; $\tau_s = 1.5$, $W = 1$), which in turn connects to the E units (*red to blue*; $\tau_s = 1.5$, $W = -4$) and I units (*red to red*; $\tau_s = 1.5$, $W = -3$) of every other group. This surround inhibition ensures only one group may be active at any one time, a key concept of time cells (Eichenbaum, 2014). Slow-excitatory units (P ; *green units*) straddle between two groups, receiving activation from the preceding group of E units (*blue to green*; $\tau_s = 30$, $H1 W = 0.011$, $Hn W = 0.015$), and forwarding activation to the E units of both groups (*green to blue*; $S.D. = 2$, $\tau_s = 25$, $H1 W = 0.7$, $Hn W = 0.4$). All P units have feed-forward connections to the first E group in any existing lower-order chain ($\tau_s = 30$, $W = 0.1$), triggering the lower order chain to commence (Figure 59A; feed-forward initiation). P units also connect to the slow-inhibitory unit (O ; *yellow*) within their group (*green to yellow*; $\tau_s = 30$, $H1 W = 0.5$, $Hn W = 0.025$), where they either increase

input to just below the spiking threshold for Hn chains, or immediately activate O units in the $H1$ chain to enable self-termination. All O units in Hn chains then receive the summation of input from all of the P units in the final group of any existing lower order chain ($\tau_s = 15$, $W = 0.2$), which is enough to push the ready-to-fire O unit to spike (Figure 59A; feed-back termination). It is important that the weight here is modified dependent on the number of layers, such that $W = W/n$, where n denotes the number of feed-back connections from higher-order layers in relation to the current layer, l . Spiking O units then inhibit E units within their group ($\tau_s = 30$, $W = -3$), thus terminating the currently active group in the chain.

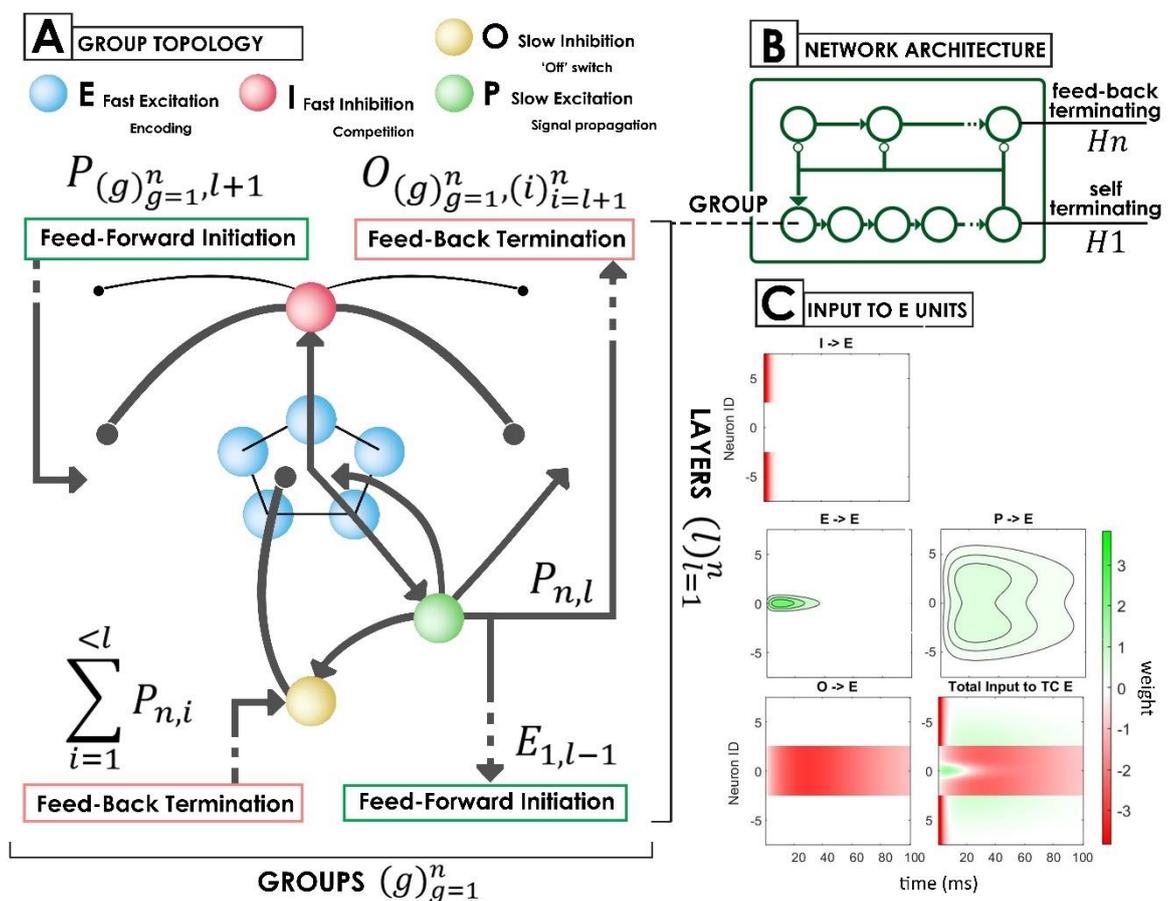


Figure 59 – Hierarchical synfire chains. **A** Illustration of the intrinsic connectivity of an assembly with accompanying notations for external connections. Here, hierarchical layers (l) of sequential groups (g) dictate information flow through feed-forward initiation and feed-back termination across hierarchies. **B** Network architecture of the hierarchy, where we identify higher-order chains (Hn) that terminate through feed-back interactions with a self-terminating lower-order chain

(*H1*). **C** Input to the fast-excitatory (E; blue) units from slow-inhibitory (O; yellow), slow excitatory (P; green) and fast inhibitory (I; red) units, where weights range from negative (red) to positive (green) and are sustained for a time determined by specific synaptic time constants.

This architecture enables P units to act as densely connected signal propagators, helping to maintain the input of both groups of E units that it connects to, thus facilitating the signal to move onto the next E group in the chain which it has been keeping in a ready-to-fire state. P units also perform feed-forward initiation and feed-back termination, kick-starting the 1st group of a lower-order chain via the former and terminating currently active groups in any higher-order chains via the latter. The *H1* chain has stronger weights within E->E & P->O connections, ensuring that activity propagates faster to enable self-termination, as no lower-order layer is available to do so. All *Hn* chains should be able to take the same parameter set, regardless of the desired temporal dimension. We thus partly satisfy a concern raised by Goldman (2009) that recurrent networks would not likely be able to encode for long durations, as infinitely small weights would be required to encode for increasingly longer temporal dimensions.

One can see how activation occurs in *Hn* chains in Figure 60, where the ring-like structure of E cells feeds activation through the assembly (Figure 60Aii; blue line) to cause rapid spike events in the voltage trace (Figure 60Ai; <5000ms). These in turn activate the local I cell (Figure 60Bii; blue line, <5000ms) which spikes (Figure 60Bi) to feed inhibitory input into neighbouring E assemblies (Figure 60Aiv; red line), thus preventing them from firing (Figure 60Aiii; <5000ms). A P cell connects these neighbouring E assemblies, receiving excitatory input from the first (Figure 60Cii; blue line) and sending excitatory input to both (Figures 60Aii & 60Aiv; green lines). An E assembly will only terminate once it receives a feed-back signal from a local O cell, which is kept in a ready-to-fire state due to excitatory input from the local P cell (Figure 60Dii; green line). Once a higher-order chain terminates, this O cell receives an additional boost (Figure 60Dii; ~4900ms), which causes it to spike (Figure 60Di) and send a strong inhibitory signal to the local E assembly (Figure 60Aii; yellow line), which terminates its activity (Figure 60Ai; ~4900ms). Ramping input then builds

up within the next consecutive E assembly (Figure 60Aiv; blue line) due to the propagating boost of the now terminated P cell (Figure 60Aiv; green line) and the diminishment of inhibition, thus starting the cycle over again in a feed-forward manner.

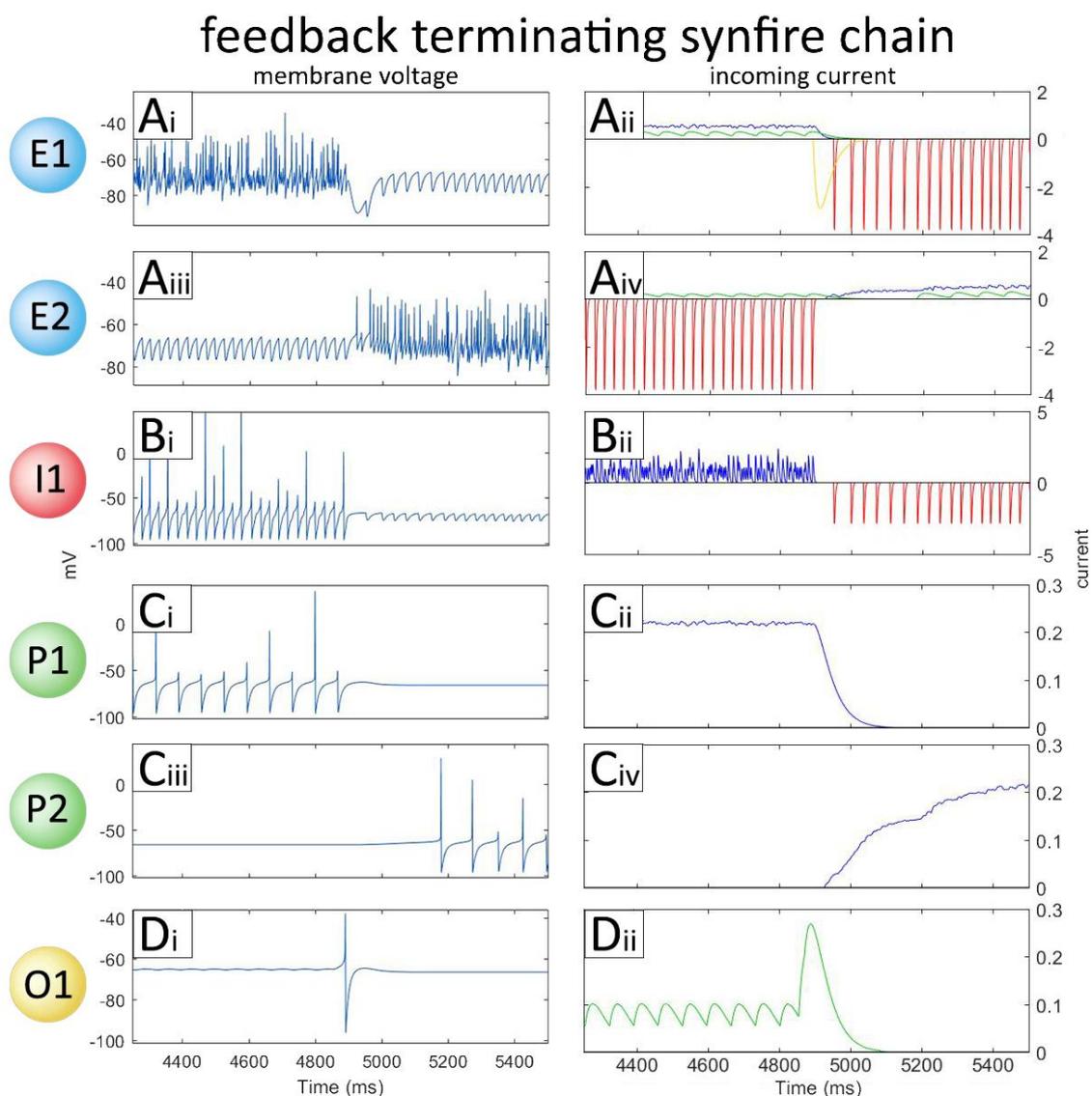


Figure 60 – Intra-cellular analysis of a higher-order chain (H_n), where incoming synaptic current (right panels) is shown for inputs with colour-coded lines that relate to the originating cell. Voltage traces are also shown (left panels). These chains are made up of rings of fast-excitatory cells (E; **A**, blue), which inhibit their neighbours through local fast-inhibitory cells (I; **B**, red). A slow-excitatory cell (P; **C**, green) ensures signal transmission between groups and layers. Activation is terminated by slow-inhibitory cells (O; **D**, yellow), which here receive a signal from the completion of a lower-order chain to terminate activation in the currently active E assembly.

The *H1* chain operates in much the same way as the previously described feedback-terminating *Hn* chain, but transmits the feed-forward signal at a faster rate due to increased synaptic strength within each E assembly and from P to E cells. Self-termination is fulfilled by the much larger synaptic strength from P to O cells, which ensures O cells will activate with no need for additional input from other layers. Figure 61 describes this process. As before, an E assembly maintains activation through recurrent circuitry (Figure 61Aii; blue line). This feeds into a local inhibitory cell (Figure 61Bii; blue line, 3600>t>3800ms), which in turn inhibits neighbouring E cells and I cells (Figures 61Aii & Bii; red lines) to ensure a winner-take-all environment.

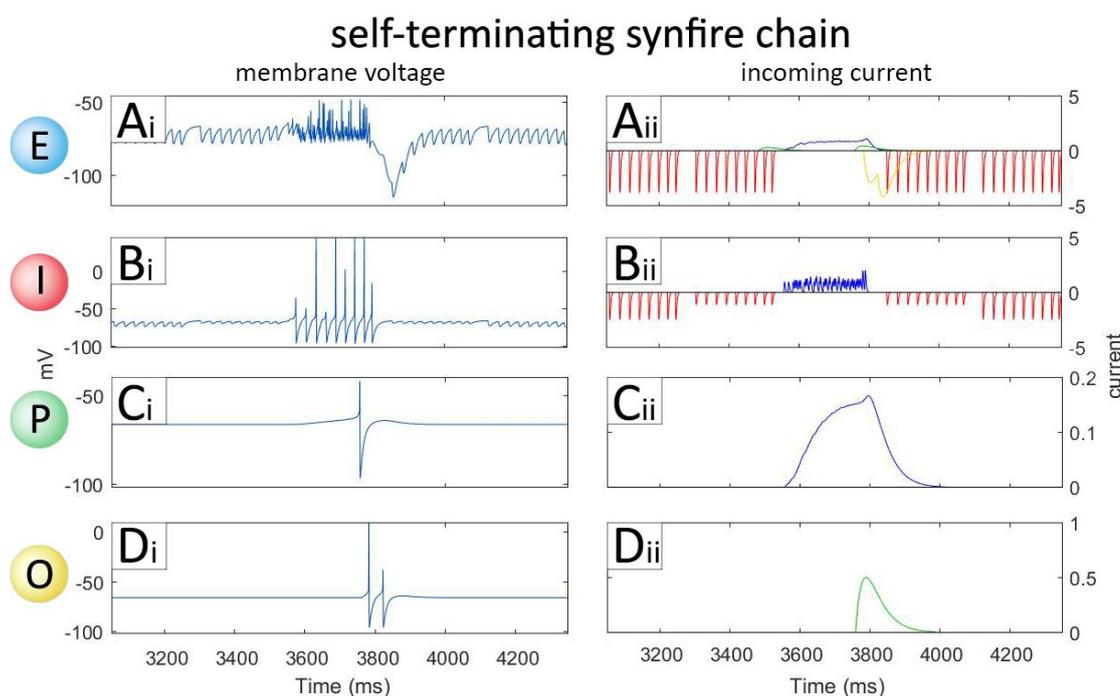


Figure 61 – Intra-cellular analysis of the lowest-order synfire chain (*H1*), which must self-terminate. Incoming synaptic current (right panels) is shown for inputs with colour-coded lines that relate to the originating cell. Voltage traces are also shown (left panels). Activation is passed through successive groups of fast-excitatory cells (E; A, blue), which inhibit one-another through the use of local fast-inhibitory cells (I; B, red). Communication is facilitated by slow-excitatory cells (P; C, green), which activate slow-inhibitory cells (O; D, yellow) that in turn terminate local E assemblies. P cells also pass activation on to the next E assembly to ensure signal transmission.

The local P cell in this chain receives weaker cumulative input from E cells than in *Hn* chains (Figure 61Cii; blue line), such that when it eventually spikes (Figure 61Ci; ~3800ms) it triggers the local O cell (Figure 61Dii; green line) to also spike (Figure 61Di; ~3800ms). This in turn has the effect of terminating ongoing activation in the local E assembly (Figure 61D; yellow line) so that the process can be repeated in the next consecutive neuronal group. In such a way, this chain self-terminates with no feedback activation from higher-order chains, thus keeping time for the lowest-order temporal dimension in the model. As all *Hn* chains terminate through feed-back interactions, the overall speed of the hierarchical synfire chain model is dictated by the parameters for the *H1* chain, which could be made to self-complete at a faster rate for a finer temporal dimension.

The interaction between feedback-terminating higher order chains and a self-terminating lowest-order chain can be seen in the raster plot of Figure 62, where each spike event is represented by a dot. Here, activation of each cellular type is colour coded according to Figures 60-61, with the faster-firing lowest-order chain represented in the top half of Figure 62. This simulation chose 3 neuronal groups per layer, though in theory any number may be chosen – as later simulations will demonstrate. The interesting finding in Figure 62 is that an alteration of P->E synaptic strength within and between hierarchical layers leads to separate scenarios. Using the previously described parameter set, one is left with a ‘hand-over’ period whilst consecutive E assemblies ramp up activation during signal transmission. During this time, incoming events might not be encoded (Figure 62A; red boxes), as it is presumed within the model that an event must be bound to every temporal dimension. Alternatively, one can minimise this hand-over period by increasing P->E synaptic strengths, such that consecutive E assemblies require no ramping up during signal transmission (Figure 62B). In this case, however, events may be bound to multiple layers within each temporal dimension (Figure 62B; red boxes). This in turn might have the effect of causing order-errors or conjunction errors in the encoding of episodic sequences (Botella, et al., 1992). This resonates with the simultaneous type/token theory of the attentional blink (Bowman & Wyble, 2007; Wyble, et al., 2009; described in Chapter 3), which argues that this blink in perception has the functional role of preventing spurious episode bindings.

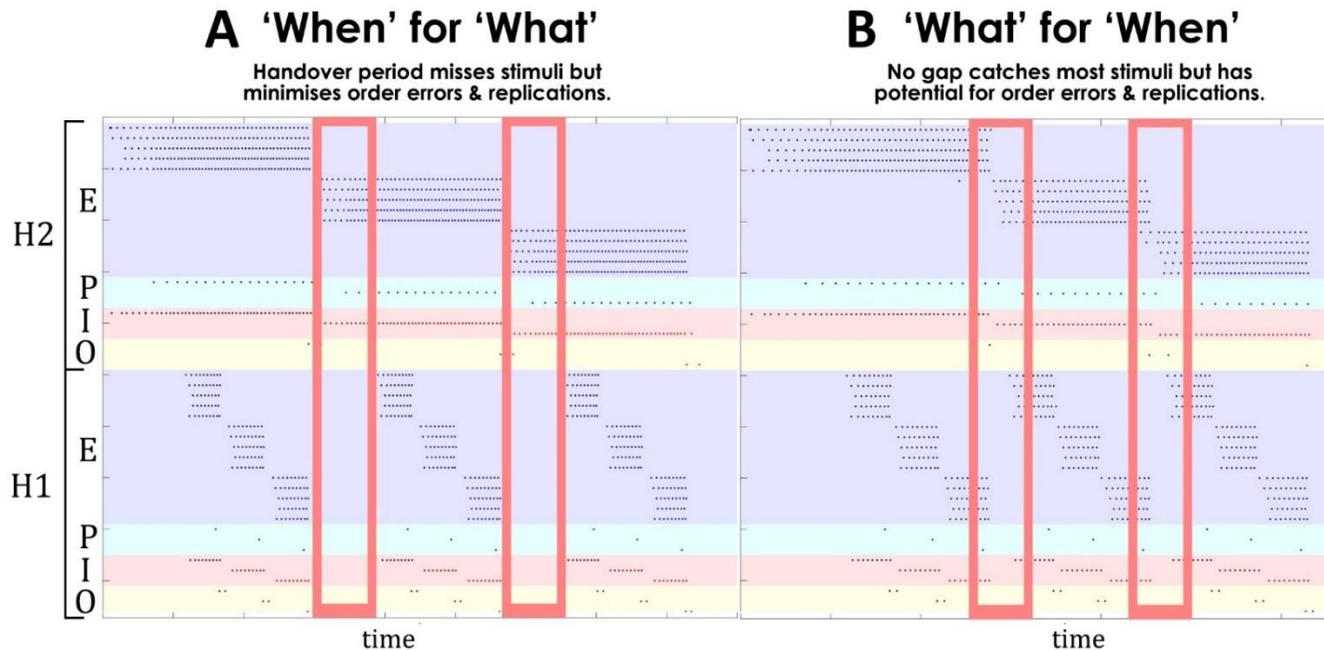


Figure 62 – Time-keeping using the hierarchical synfire chain model with varying parameter sets. The activation of fast-excitatory (E; blue), slow-excitatory (P; green), fast-inhibitory (I; red) and slow-inhibitory (O; yellow) cells keep time through multiple temporal dimensions. A ‘hand-over’ period in higher-order chains can be altered by varying P->E synaptic strength. A slight increase causes the scenario in **B**, which minimises the length of time that consecutive E assemblies must ramp up during signal transmission. In this case, incoming discrete events (indicated by red boxes) have a risk of being bound to multiple groups within each temporal dimension – possibly causing order and conjunction errors. Minimising this risk, however, entails including a blank ‘handover-period’ (**A**), which are temporal blind-spots where events might be missed altogether.

Within the context of this model, the highest-order temporal chain must be initiated by an external force, possibly attentional-related and fixed to the onset of a stimulus. This was implemented as an injection of a DC current to the first E unit in the first group lasting for 10ms ($W = 0.15$), which occurred 500ms before the onset of the first stimulus to allow the hierarchy to ramp up activation. For the subsequently described experimental paradigm, two hierarchical layers were used to encode a brief episodic sequence. These were organised as a lowest-order chain of 8 groups (H1) and a highest-order chain of 3 groups (H2), such that a stimulus of up to $\sim 6s$ could be recorded.

STDP enabled synfire chains to attach themselves to active binding pool units ($\theta_d = 1$, $\theta_p = 1.3$, $\gamma_p = 1.5$, $\gamma_d = 0.05$, $C_{pre} = 0.8$, $C_{post} = 0.8$), with the effect that active binding pool units associated in time to any active group from each hierarchical layer, thus indicating that the event occurred at a point on simultaneous temporal dimensions. It was necessary to balance SC->BP connections such that BP units only activated from the cumulative activation of every SC layer ($H1 \tau_s = 1.5$, $H2 \tau_s = 20$, $H1 W = 0.028$, $H2 W = 0.011$). As such, longer duration H2 input builds BP units up to sub-threshold levels, then shorter duration input from H1 triggers activation.

7.3.5 NEURO-MODULATORS

Simulations comprised of two distinct phases, an encoding phase and a recall phase. During the former, stimuli are presented to the NC and bound via the BP to the SC, where activation flows upwards through the network (see Figure 57; encoding phase, black directional arrow). During the latter, activation moves in reverse, whereby the SC is kick-started and activates related NC units via the BP in a downward direction (Figure 57; recall phase, grey directional arrow). To allow uni-directional flows of information and prevent feed-back activation to propagate through the hierarchy, we assume the presence of several neuro-modulators. During the encoding phase, downwards directional weights SC->BP & BP->NC are set to zero. Similarly, during the recall phase, the reverse upwards connectivity weights NC->BP are set to zero. It is important to note

that whilst weights are nullified, the synapse variable ρ is still able to fluctuate between an active (1) and inactive (0) state if STDP is enabled on those connections.

7.4 EXPERIMENTAL PARADIGM

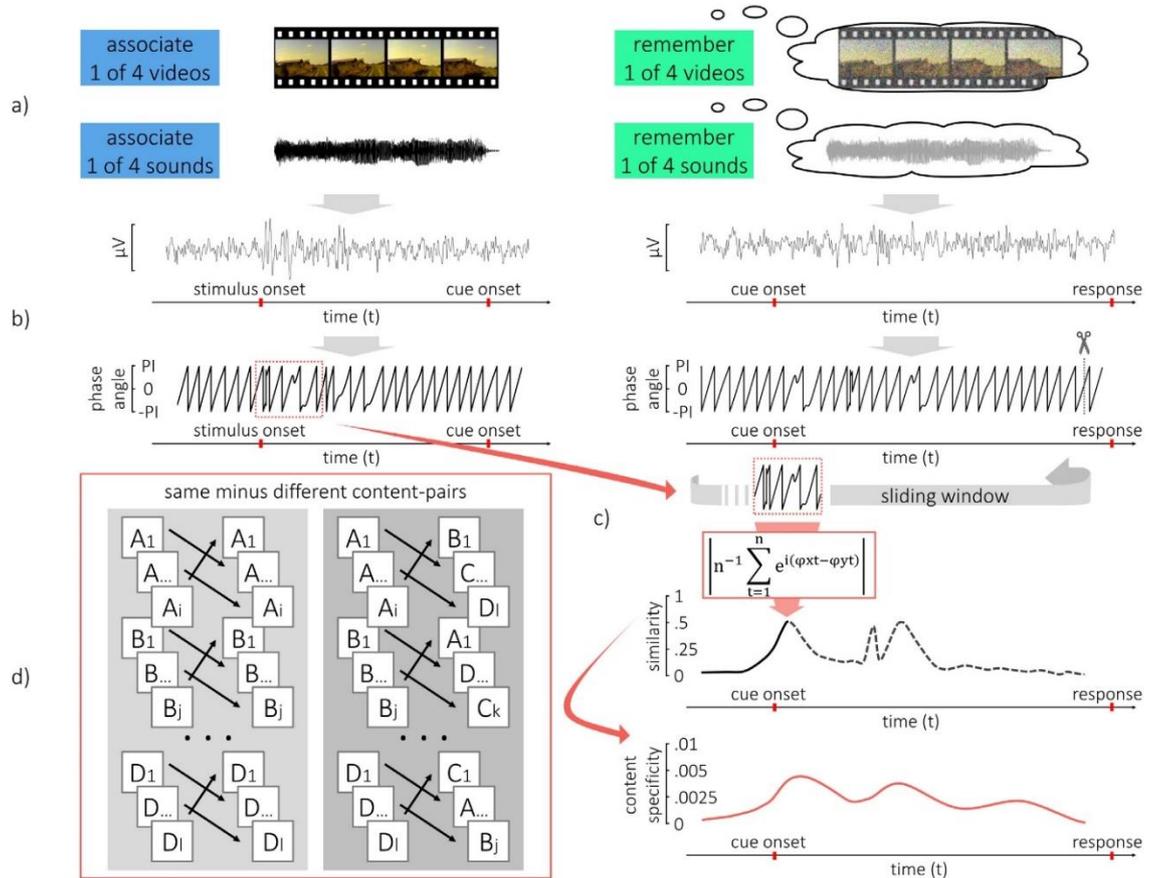


Figure 63 – Detection of content-specific reinstatement of temporal patterns. **(a)** During encoding (left panels), subjects associated one of four videos or sounds with a different cue word in every trial. During retrieval (right panels), subjects were asked to recall the dynamic stimulus after seeing the cue word on screen. **(b)** At every electrode, the oscillatory phase for a frequency of interest was extracted from the EEG activity from encoding (left) and retrieval (right). A time window from encoding was also selected (left; orange dotted box). **(c)** A sliding window was used to assess the similarity (top panel), based on the constancy of phase angle differences over time. **(d)** The difference in similarity between pairs of same and pairs of different content was interpreted as evidence for content specific reinstatement of temporal patterns, which is shown in the bottom panel of c. Figure adapted from Michelmann, et al., 2016.

The model was used to simulate an experimental paradigm to detect the content-specific reinstatement of temporal patterns in the neocortex (Figure 63; see Michelmann, et al., 2016, for experiment). In the experiment, several dynamic stimuli were presented to the participant, in this case a video or sound, which were each associated to a non-dynamic stimulus, in this case a word (Figure 63a; left panel). Alpha frequencies were interrogated during the presentation of each dynamic stimulus at encoding; a period of post-stimulus alpha desynchronisation was observed that is typical of active neuronal assemblies (Haegens, et al., 2011; Hanslmayr, et al., 2011a). During this period, the phase angle pattern was observed (Figure 63b; left panel), which has been shown to reliably convey information content (Ng, et al., 2013; Schyns, et al., 2011), where a window indicating a period of phase-resets was chosen for each pattern. Later, the associated word was presented to the participants, where they were asked to remember the associated dynamic stimulus (Figure 63a; right panel). Similarly, the phase angle pattern during a period of alpha desynchronisation was observed (Figure 63b; right panel). As the experimenters had no knowledge of the specific time that the participant internally recollected the dynamic stimulus, they had to dynamically find the period of time that the video was being mentally re-played.

This was done using a convolution approach, where the phase-reset pattern of the dynamic stimulus at encoding was compared to a window of equal length centred over every time point of the phase-angle time series at recall. This is summed up in Equation 30, where the sum of differences between two phase angles (φ_{xt} & φ_{yt}) at every time point was calculated. This approach gave a dynamic similarity signal through time (Figure 63c; top panel), where an increase in the similarity measure signalled high similarity between the phase-reset pattern at encoding and retrieval, thus indicating the re-instatement of neural patterns.

$$\left| n^{-1} \sum_{t=1}^n e^{i(\varphi_{xt} - \varphi_{yt})} \right|$$

Equation 30 – Representational similarity analysis (RSA)

This method was taken one step further with the use of several unique dynamic stimuli (Figure 63d), where each was presented several times. Here, the similarity time-series between encoding and retrieval was found between trials of the same stimulus and summated. From this, the similarity time-series between trials of different stimulus was taken. Thus, if the resulting content specificity signal was positive, then it indicated that the similarity of the same stimuli was bigger than for different stimuli. The bottom panel of Figure 63c shows such a time-series, indicating that each video initiated a unique alpha phase-reset pattern that it could later be identified by.

7.5 SUMMARY

In this Chapter, we have looked again at human time-keeping in the brain. Here, it was shown that there might be several hierarchies of independent time-keepers. This begins with the centralised circadian rhythm that moderates bodily processes (Hofstra & de Weerd, 2008) and possibly helps to update episodic information via hippocampal-anterior thalamic connections (Aggleton, et al., 2010), where it is thought that the hippocampus plays a substantial role in contextualising memories in space and time during recollection memory (Eichenbaum, et al., 2007). Here, a serial-order of events is thought to emerge from the competition of many sequences (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). It is thought that cortical matter stores a more distributed form of time-keeping (Mauk & Buonomano, 2004), where sequences that form episodic long-term memory might be inherited from medial-temporal-lobe (MTL) sequences (Itskov, et al., 2011) in a complimentary-learning-systems framework (O'Reilly, et al., 2011). These distributed sequences can then be simultaneously compared to concurrently active MTL sequences via interactions with the perirhinal cortex, generating a similarity signal that is important for familiarity-based memory (Eichenbaum, et al., 2007).

We then proposed a theoretical framework of a hierarchy of time-keepers, using consecutively connected recurrent groups as a 'feed-forward chain in disguise,' with the intention of theorising as to how multiple interacting and simultaneous time-keepers might interact with one-another.

Using a feedforward architecture to encode for time has been proposed before (Goldman, 2009; Itskov, et al., 2011), though issues of complexity and scalability have arisen (Shankar & Howard, 2012). Here, we present a computational neural network model based on a competitive hierarchy of synfire chains, where each moment in time is encoded by the simultaneous activity of multiple hierarchies, thus allowing for repetition within each hierarchy and reducing complexity. Using such an approach, we theorise how a slower time-keeper like the circadian rhythm might help contextualise or update the faster time-keepers in episodic memory (Aggleton, et al., 2010), where a chunking operation might explain temporally moderated hippocampal seizures (Gowers, 1885). Instead of persistent excitatory activation, these chains might similarly be organised through inhibitory heteroclinic orbits (Rabinovich, et al., 2006), or other such methods. Our main interest is in simulating some mechanism for the representation of a hierarchical encoding of time.

We identify an episodic memory paradigm to evaluate our time-keeping architecture, where a cortical alpha desynchronisation and subsequent phase-reset pattern convey information flow (Haegens, et al., 2011; Hanslmayr, et al., 2012) and information content (Ng, et al., 2013; Schyns, et al., 2011), respectively. Here, a representational similarity analysis (RSA) was performed to show that a dynamic stimulus could be identified by its unique phase-reset signature (Michelmann, et al., 2016). To simulate this experiment, it was important for the model to be able to create a unique tag for every event, as in other models (Bowman & Wyble, 2007), such that it might be bound in time. To do this, the model was created in three parts. Firstly, an intrinsically oscillating neo-cortical (NC) area was simulated using a competitive excitation-inhibition balance. Events, coded as a pulse of noise into NC units, then initiated activation in a binding pool area, upon which a unique set of neurons activated in a competitive winner-take-all environment. These units were then bound to an active hierarchy of synfire chains. Using this approach, one could re-start the synfire chains to recall the temporal order of events in the neo-cortex, where a subsequent RSA analysis would be performed to see if event patterns were unique enough to be identified by their respective phase-reset signatures. The following Chapter will analyse simulations from each of the interacting sections of this modelling project, as well as the final content specificity RSA.

8. HIERARCHICAL SYNFIRES CHAIN MODEL - RESULTS

This Chapter details the results of the hierarchical synfire chain model described in the preceding Chapter. These will analyse the synaptic currents and weight changes that enable the model to keep time and encode discrete events. From there, we discuss the findings in relation to the chosen experimental paradigm, where a representational-similarity-analysis (RSA) was performed to identify unique temporal signatures and thus dynamically detect when episodic memories were being replayed in the brain (Michelmann, et al., 2016).

8.1 RESULTS

This section will begin by analysing the manner in which alpha oscillations are generated, detailing the conditions in which a desynchronization and subsequent phase-reset might occur. We then look into intra-cellular dynamics to describe how our binding pool area enables a winner-take-all environment, where one cell assembly can come to dominate over the rest for a short period of time. A clustering algorithm is then identified, enabling us to detect discrete events in the binding pool such that we can look at the weight changes specific to active cell assemblies. Once we have described the full binding process within the model, it is important to analyse how sequences are replayed and dynamically detected using RSA, similar to our chosen experimental paradigm (Michelmann, et al., 2016).

8.1.1 CORTICAL PATTERN REPRESENTATION

Cortical assemblies were designed such that they intrinsically oscillated at a particular frequency, determined by parameters for excitation/inhibition interactions (Brunel, 2000). In addition, a winner-take-all environment was instantiated, as has been theorised in models of visual working memory (Itti, et al., 1998). Hypothetically, this ensures that only one locally connected neuronal group can be active at any one time, minimising the simultaneous activation of multiple representations in a distributed manner. As described in Figure 57, parameters were chosen such

that activation could spread linearly across the entire excitatory population before inhibitory interactions had time to clamp down, thus allowing intrinsic oscillations to emerge.

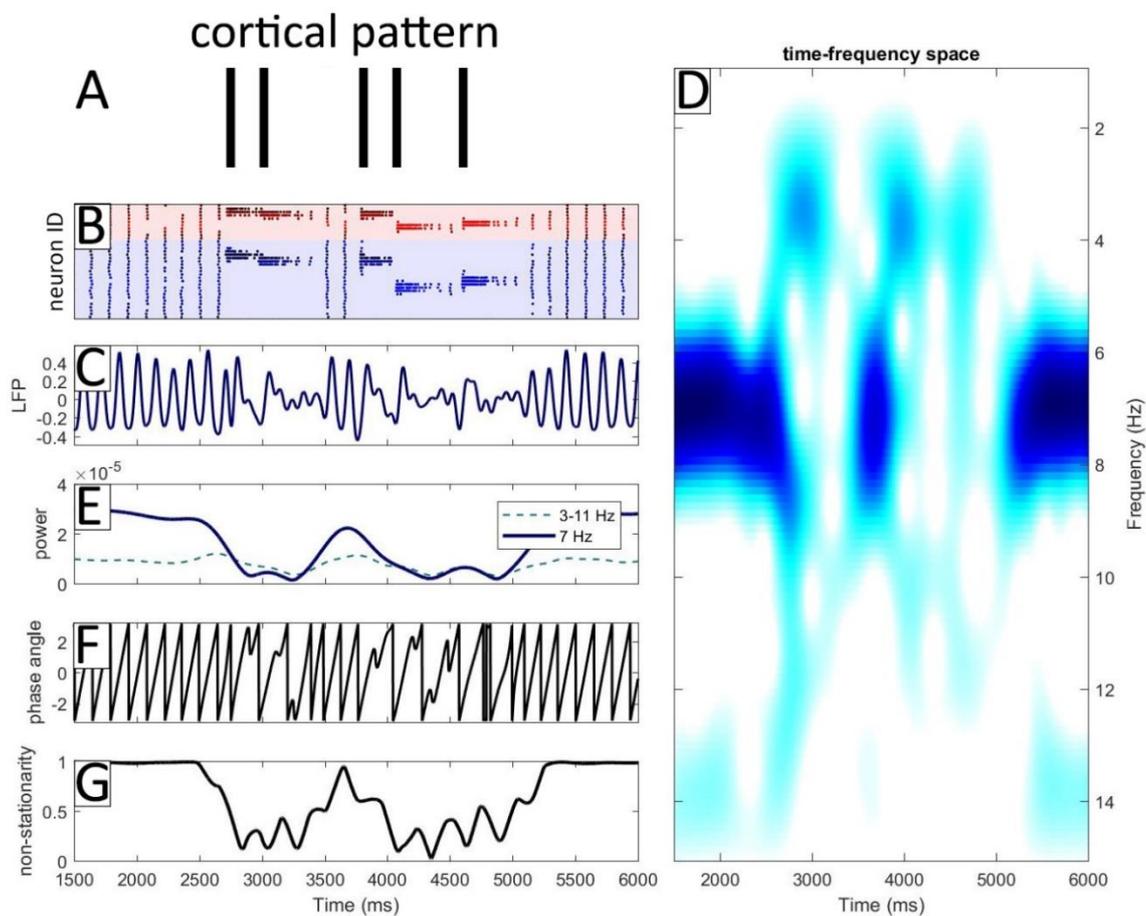


Figure 64 – Cortical representations and alpha phase-resets. The presentation of a temporal pattern (A) triggers activation in a cortical assembly that intrinsically oscillates at rest (B; excitatory and inhibitory spike events in blue and red, respectively). This causes irregularity in the local-field-potential (LFP; C), which in turn can be examined in time-frequency space (D). Pronounced desynchronisations in the ongoing oscillation can be observed during event-related activation, as seen by plotting power of the idling frequency through time (E; 7Hz). At these times, the phase angles of the idling frequency indicate the occurrence of multiple phase-resets (F), further indicated by analysing the non-stationarity of the phase-angle time-series (G; i.e. similarity to a phase-angle time-series from the pre-stimulus period).

In the case of sustained additional input, however, a selective excitatory population recurrently interacted with its local inhibitory cells to inhibit the rest of the cell assembly. This is best seen in Figure 64B, where pre-stimulus spike events (excitatory/inhibitory in blue/red, respectively) indicate an intrinsic oscillation (Figure 64D; 7Hz band in frequency domain), until event-related activation causes sustained activation of a subset of units. This subsequently causes a desynchronisation in the ongoing alpha oscillation (Figure 64E; blue line at 7Hz), as discussed in Chapter 6, this is a phenomenon shown in many studies (Haegens, et al., 2011; Hanslmayr, et al., 2011a; Klimesch, et al., 2007) and models (Parish, et al., 2018) to be related to neuronal activation.

During event-related desynchronisations in cortical alpha oscillations, evaluating the phase-angle time-series has been shown to render qualitative information about stimulus content (Ng, et al., 2013; Schyns, et al., 2011). Therefore, in the model, we show that the phase-angle time-series does indeed show evidence of phase-resets at this time (Figure 64F; 2500-5000ms), as the on-going intrinsic oscillation is externally affected through its representation of incoming stimuli. A further examination of non-stationarity in the phase-angle signal (Figure 64G) indicates that these phase-reset periods are linked to periods of event-related cortical activation, lending theoretical evidence to the argument that phase-angle time-series can be used to decode information content (Ng, et al., 2013; Schyns, et al., 2011). We later examine how this information content can be used to dynamically detect unique temporal signatures when replaying episodic memories within our model, as has been found in the literature (Michelmann, et al., 2016).

8.1.2 BINDING DISCRETE EVENTS

The binding pool (BP) within the model was developed with similar intentions as other models (Bowman & Wyble, 2007), where a unique node would activate to encode an incoming discrete event. It is argued that this method allows for repetitions, as each presentation is treated as an independent event. In this way, a BP node is only required to activate during an event-related cortical activation, as can be seen in Figure 65. Here, parameters were chosen such that intrinsic

oscillatory activity (Figure 65A; $<0s$) is not sufficient to cause activation in the BP layer. This is fulfilled by a large synaptic time constant from NC to BP layers ($\tau_s=15$), large enough to allow sustained input to eventually trigger BP activity but low enough that synaptic input decays by the next periodic oscillatory activity.

As discussed in the previous Chapter, the weights of sparse connections from NC to BP layers vary in a normal distribution ($W=0.0125$, $S.D.=0.0005$), such that when event-related sustained input triggers BP activity, it is likely that a single excitatory unit (E) will spike before others. This is the crux of winner-take-all behaviour in BP layers, such that a single unit can 'snowball' excitatory input to its locally connected neighbours (Figure 65B; blue line, 0.1-0.2s), whilst simultaneously activating local fast-inhibitory (I) cells that in turn inhibit excitatory neighbours (Figure 65C; red line 0.1-0.2s). Due to the nature of E \leftrightarrow I connectivity, there is visible inhibition when examining the average synaptic inputs for a cluster (Figure 65B; red line, 0.1-0.2s). This is largely accounted for by the neurons that form the edge of the cluster boundaries, and is irrelevant so long as E \leftrightarrow E input is larger. A global slow inhibitory cell (O) adds an additional safeguard mechanism to prevent runaway activation across the excitatory population, operating to inhibit the entire BP layer once sustained excitatory input reaches a threshold (Figure 65B & C; yellow $\sim 0.2s$).

A raster plot of spike events shows how all these forces interact (Figure 65D), where the top third shows NC intrinsic oscillatory activity interleaved with periodic event-related bursts of activity, as shown in Figure 64B. The bottom two thirds show the BP layer, where a subset of unique units activate in response to each event-related NC input. From this, a clustering algorithm was used to detect BP group activity in time and space (Figure 66; colour coded from blue-red). As BP groups dynamically activate in a winner-take-all environment, this procedure allowed the analysis of BP group dynamics, as seen in Figure 65B & 65C and in the following Figures that depict the synaptic weight changes that are fundamental to the overall binding process.

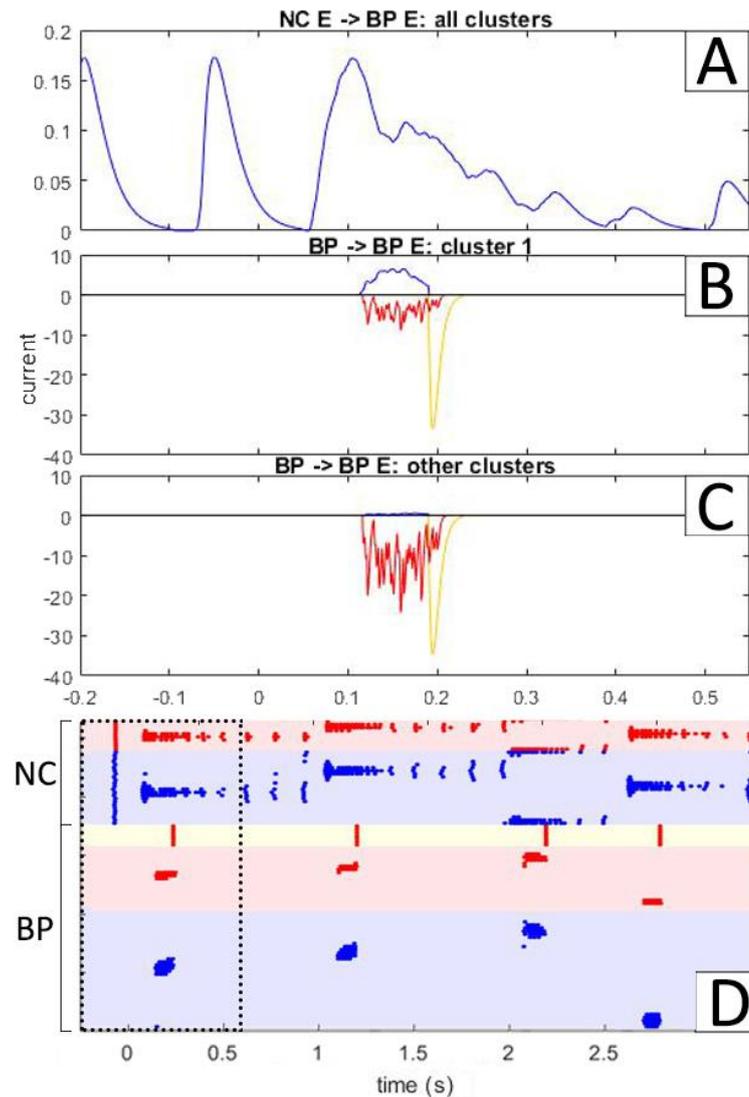


Figure 65 – Intra-cellular analysis of binding pool units, where incoming synaptic current (**A-C**) is shown for inputs with colour-coded lines that relate to the originating cell type. Synaptic input from neo-cortical (NC) areas to the whole binding pool (BP) layer (**A**; blue line) triggers winner-take-all behaviour within the BP layer. During event-related NC stimulation, recurrent BP excitatory input triggers activation in a subset of locally connected excitatory units (E) (**B**; blue line, 0.1-0.2s), which in turn activates local fast-inhibitory cells (I) that inhibit neighbouring excitatory units (**C**; red line, 0.1-0.2s). After a certain level of activation has been reached, a global slow-inhibitory unit (O) activates to shut down all BP excitatory activation (**B & C**; yellow, ~0.2s). Activation of NC and BP layers is depicted as a raster plot of spike events (**D**), where colour-coded dots (blue, excitatory; red, inhibitory) and shaded regions (blue, E; red, I; yellow, O) depict neuron type, and the black dotted region defines the temporal period shown in **A-C**.

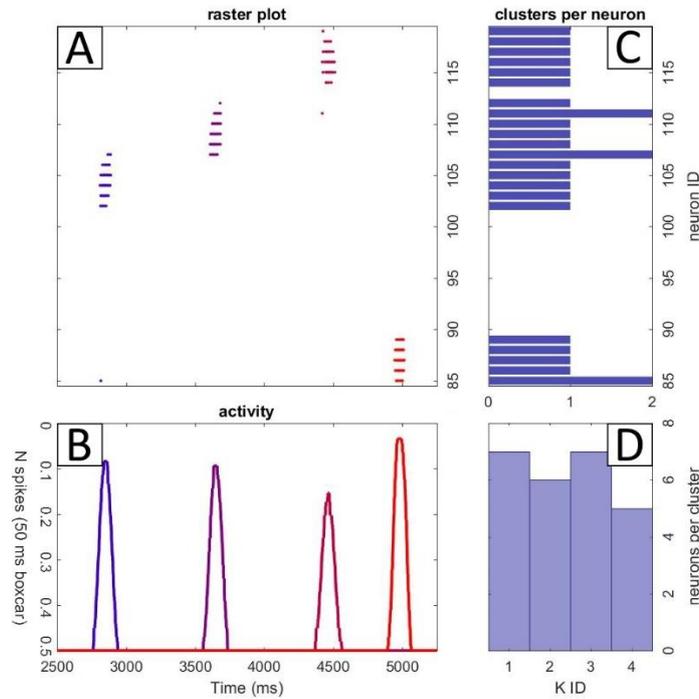


Figure 66 – Clustering algorithm used to identify active binding pool (BP) groups in time and space, indicated by a colour coded raster plot of spike events (A) and histograms of activity (B). The number of clusters per neuron (C) and the number of neurons per cluster (D) is also identified.

Synaptic modification was implemented in the model using a calcium-based STDP rule, as described in the previous Chapter. Here, separate parameter sets were chosen dependent on the desired behaviour of cellular populations. For example, intra-binding pool synapses underwent long-term-depression (LTD) during periods of activation (Figure 67A; black dotted boxes). Here, the threshold for long-term-potential (LTP) was too high for transient calcium amplitudes to reach (Figure 67A; colour coded shaded regions, left axis), meaning that BP synapses only underwent LTD (Figure 67A; black lines, right axis) once calcium passed the relatively low LTD threshold. This depression had a profound effect on active assemblies' ability to further participate in the winner-take-all environment prevalent in the BP layer, as once a group of units had been activated then they would not have sufficient strength to outcompete the now stronger untouched cell assemblies. Due to some overlap in clustered BP groups (Figure 66C; cases of 2 clusters per neuron), there is also depression within the synapses between these shared nodes (Figure 67A), where LTD occurs in a cluster despite that cluster not being currently active.

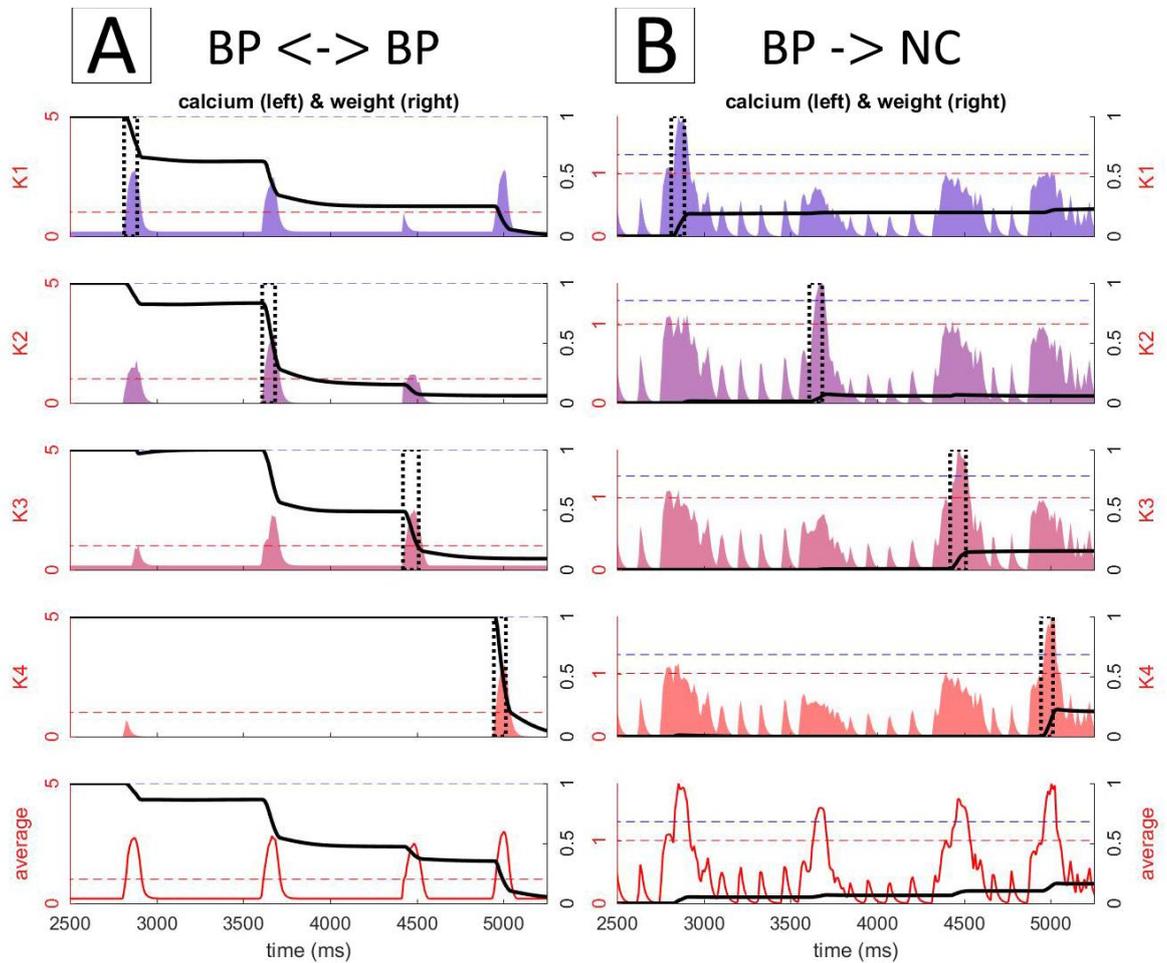


Figure 67 – Calcium induced synaptic changes within the binding pool layer (**A**; BP<->BP) and from binding pool groups to the neo-cortical layer (**B**; BP->NC). BP groups are colour coded using the clustering algorithm described in Figure 66, where clusters are labelled as ‘K-’ as identified in Figure 66D. Calcium increases at the synapse due to pre-& post-synaptic activation (left axis). Long-term-potential (LTP) is induced if calcium increases above an LTP threshold (blue dashed line) and long-term-depression (LTD) is induced if calcium is above an LTD threshold (red dashed line). Synapses (p ; black line, right axis) gravitate towards 0 or 1 dependent on a reversal state (ρ^*) of 0.5, where the average over all active synapses between selected groups is shown here. For visibility purposes, calcium amplitudes were shown as the mean plus the standard deviation of calcium dynamics in all active synapses between groups, as most of the synaptic change was due to a few synapses surpassing the threshold. The bottom panels depict the maximum at every time point of all the calcium dynamics shown in the panels above, as well as the average weight over all neuronal group pairings. Black dotted lines indicate periods of activation for pre-synaptic units.

In order for BP groups to encode for novel, discrete events, they had to be able to bind to active NC assemblies (Figure 67B). Here, BP->NC synaptic weights underwent LTP (Figure 67B; black lines, right axis) once transient calcium amplitudes (Figure 67B; colour coded shaded areas, left axis) surpassed the LTP threshold, which can be seen to occur when periods of pre-synaptic binding pool activation (Figure 67B; black dotted boxes) aligned with periods of event-related calcium increases from NC post-synaptic units. Here, it was important to find a balance in the parameters for BP->NC calcium dynamics such that LTP did not occur from the whole BP region or to the entire NC region, if pre-synaptic binding pool calcium or post-synaptic cortical calcium alone were enough to cross the LTP threshold, respectively. As such, it is required that the sum of both of these calcium amplitudes is necessary to achieve LTP. With the default calcium STDP mechanism (Graupner & Brunel, 2012; Equation 18, as described in Chapter 2), these calcium dynamics would be intrinsically tied to the firing rates of BP and NC units, and as such, would have to be adjusted for a range of levels of event-related cortical or binding pool activation. For this reason, an additional constraint was added to calcium dynamics (Equations 28-29, as described in Chapter 7) that punished additional calcium increases if a neuron had spiked in the recent past.

In order for discrete events to be bound to a moment in time, there had to be some STDP process between transiently active synfire chain groups and any active BP cluster. This is shown in Figures 68 & 69, where synapses from synfire chain to BP units increased if both were concurrently active at any moment in time. For the higher-order synfire chain (*H2*), this meant that the first 3 BP clusters bound to the 1st transiently active synfire group (Figure 68A; purple shaded), whilst the last BP cluster bound to the 2nd (Figure 68A; red shaded), where the colour coded clustered activity of synfire chain groups is shown in Figure 68B. As before, it was important to choose parameters for these calcium dynamics such that pre-synaptic synfire chain or post-synaptic BP calcium amplitudes were not sufficient on their own to cause LTP. Here, only the combination of both was enough to induce LTP (Figure 68A). The additional constraint of Equation 29 was also key for stabilising calcium increases for prolonged synfire chain activation.

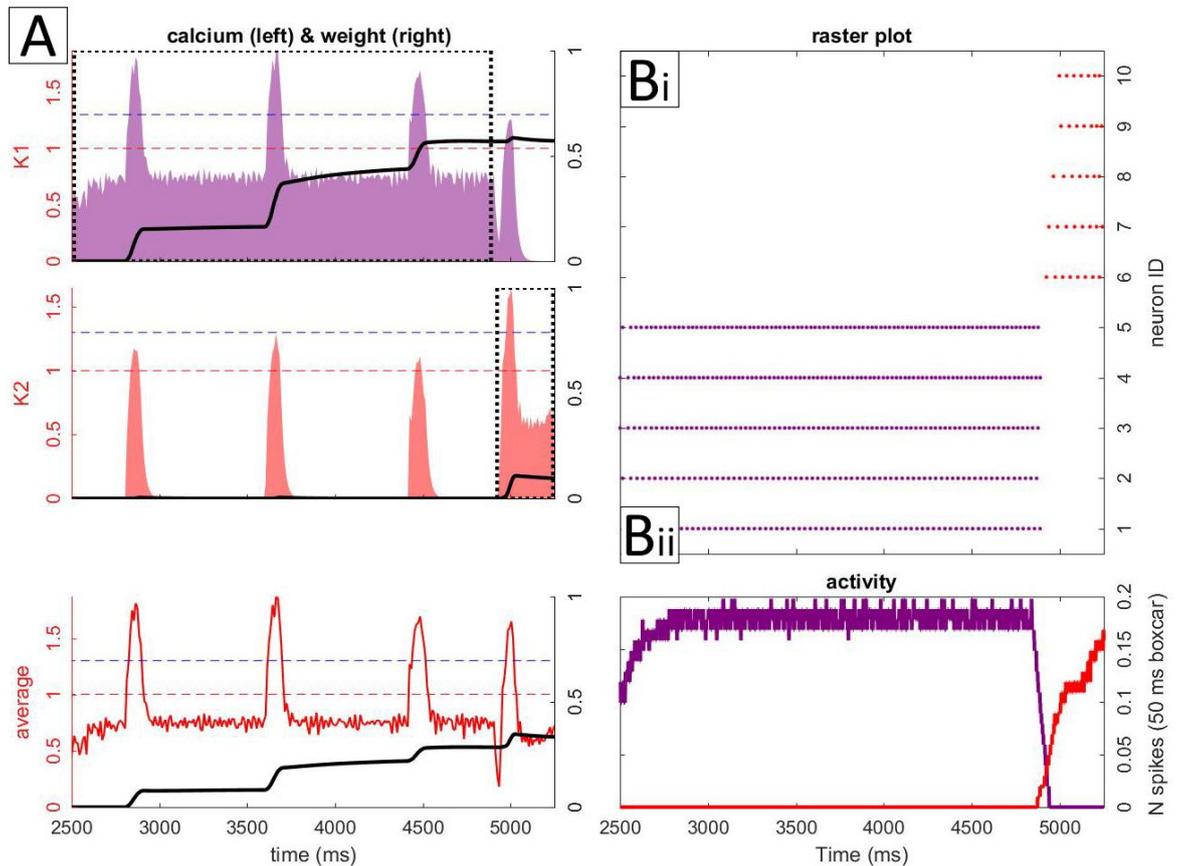


Figure 68 – Calcium induced synaptic changes from the higher-order synfire chain ($H2$) to binding pool units (A), where a colour coded clustering algorithm was used to detect spike events from separate synfire chain groups (Bi) and a histogram of activity per group is also shown (Bii). Calcium increases at the synapse due to pre- & post-synaptic activation (left axis) & the average synaptic change between paired neuronal groups (black line; right axis) is shown, as described in Figure 67. Black dotted lines indicate periods of activation for pre-synaptic synfire chain units.

Similarly, clustered BP groups bind to any active group within the lowest-order synfire chain, where activation of each synfire chain cluster is shown (Figure 69A; black dotted boxes). For the following simulations, a chain of 8 neuronal groups was chosen for the lowest-order ($H1$) synfire chain. Here, the last BP cluster does not bind to any of the active synfire chain groups, due to the ramping up process described in Figure 62.

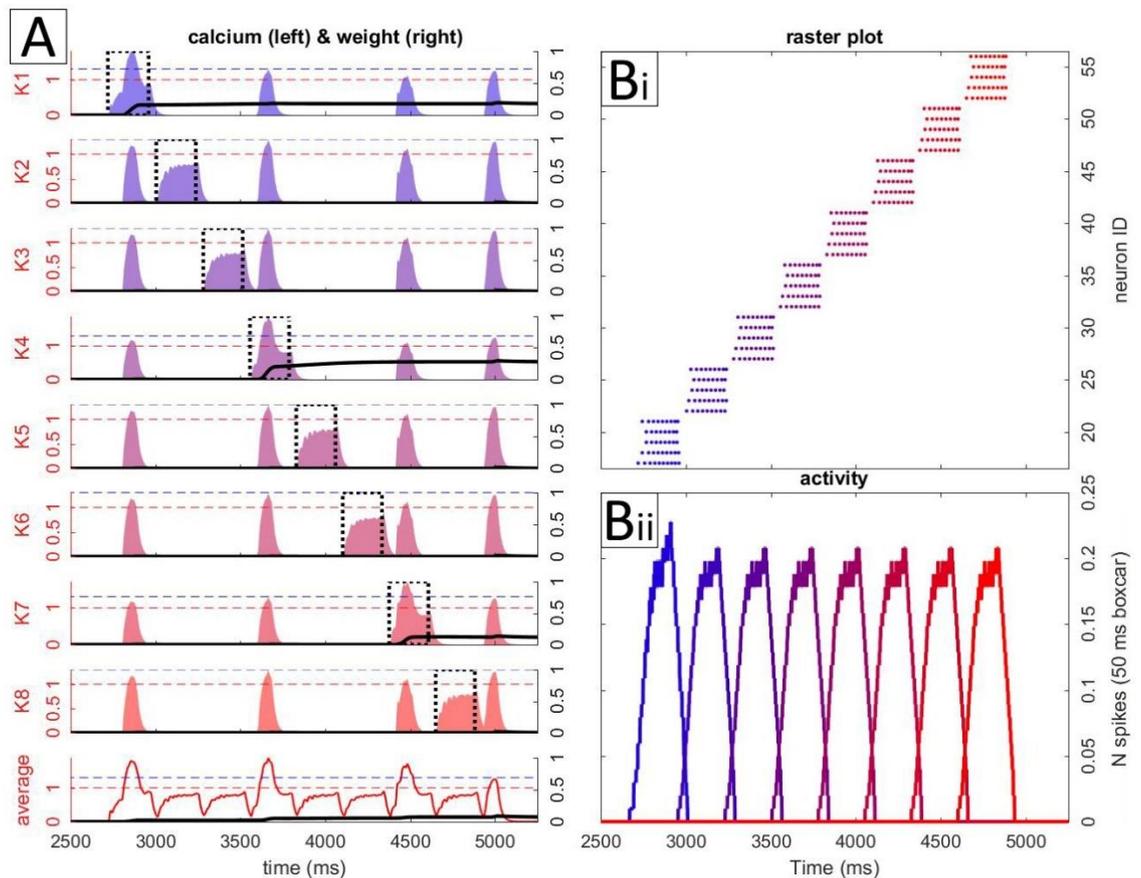


Figure 69 – Calcium induced synaptic changes from the lowest-order synfire chain ($H1$), that here contains 8 neuronal groups, to binding pool units (**A**), where a colour coded clustering algorithm was used to detect spike events from separate synfire chain groups (**Bi**) and a histogram of activity per group is also shown (**Bii**). Calcium increases at the synapse due to pre- & post-synaptic activation (left axis) & the average synaptic change between paired neuronal groups (black line; right axis) is shown, as described in Figure 67. Black dotted lines indicate periods of activation for pre-synaptic synfire chain units.

8.1.3 RECALLING EPISODIC SEQUENCES

Once we have encoded sequences to memory, it is important to be able to correctly replay relevant items in the original order. As described in the previous Chapter, this is achieved in the model through the use of neuromodulators, which restrict activation either in a top-down encoding or a bottom-up recall direction. During recall, downwards neo-cortical (NC) \rightarrow binding pool (BP) weights are set to zero, whilst the reverse synfire chain (SC) to BP and BP to NC weights are increased.

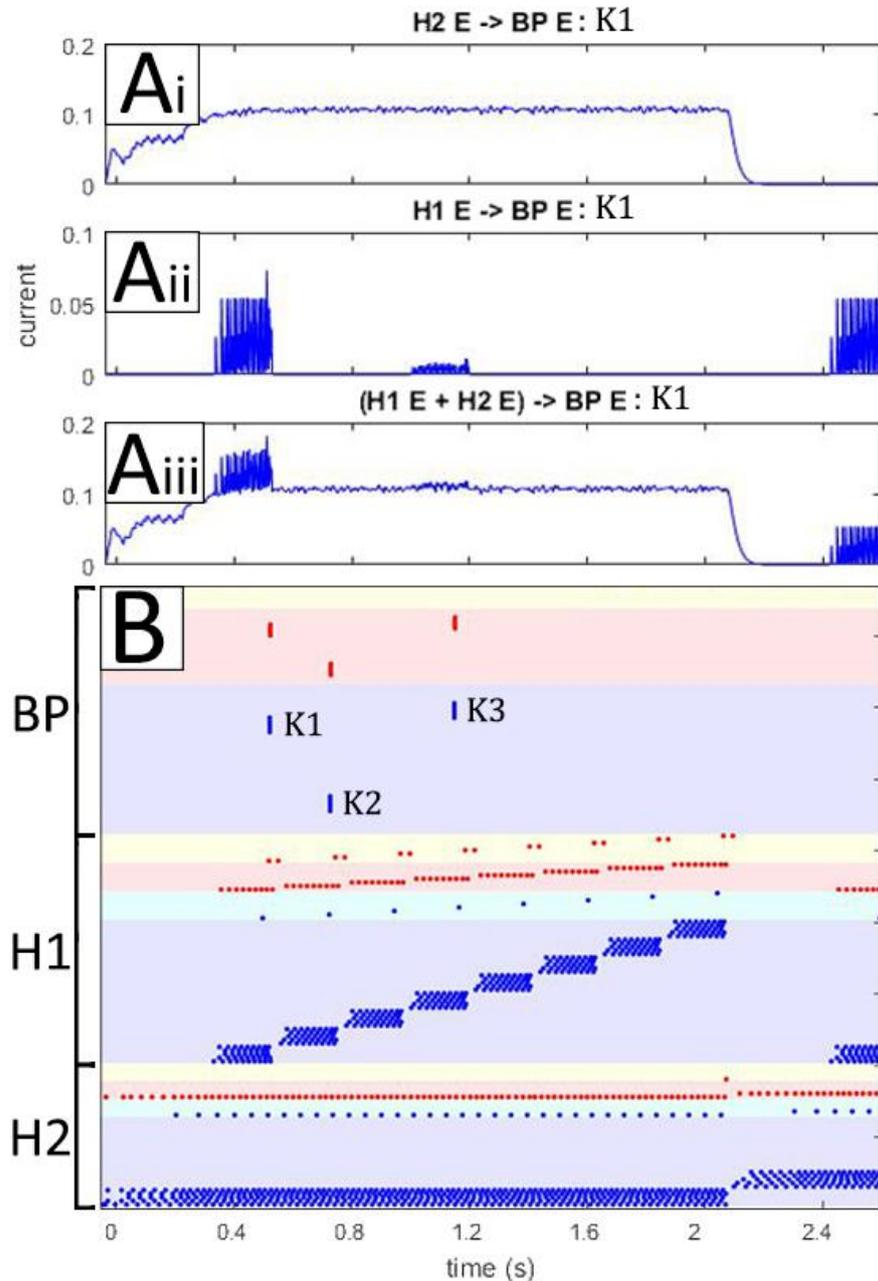


Figure 70 – Triggering the binding pool from synfire chain activation. During recall, the hierarchical synfire chain is restarted and a synaptic modulation amplifies synfire chain (SC) to binding pool (BP) and BP to neo-cortical (NC) weights, in preference to NC to BP weights. Here, input to BP cluster 1 (K1) is shown (A) from the slow H2 chain (i), fast H1 chain (ii) and the combination of both temporal layers (iii). Some additional input arises during alternate phases (~ 1.2 s) due to an overlap between clusters K1 & K3. The resulting raster plot of activation (B) shows how sequential SC activity encourages BP activity to occur at specific times, where colour-coded dots (blue, excitatory; red, inhibitory) and shaded regions (blue, E; green, P; red, I; yellow, O) depict neuron type.

This neuromodulation enables the directional flow of information and prevents recurrent activation between layers during encoding or recall. This is especially key to allow the binding pool to fulfil its dual purpose, as it is required to be highly sensitive to either cortical or temporal activation during encoding and retrieval, respectively. Where Figure 65 dealt with the encoding state, Figure 70 deals with the retrieval state. Here, it is key that binding pool units only activate when the combined synaptic currents from each hierarchical synfire chain surpasses a threshold. In Figure 70A, the synaptic input for the first BP cluster is shown, where 70Ai shows weighted synaptic input from the higher-order SC and 70Aii shows the same from the lowest-order SC. BP spiking occurs when summated synaptic currents are maximal, shown in 70Aiii. It is important to note that later activation of one SC layer but not the other does not result in further BP activity (Figure 70Aiii; >2.6s). This method allows for the repetition of each temporal dimension, decreasing the computational complexity of the use of feedforward chains to encode for time, as has been a major criticism elsewhere (Shankar & Howard, 2012).

We can now encapsulate the entire simulation procedure in Figure 71. Here, a pattern is presented to cortical sites during the encoding phase (Figure 71Ai), which causes subsequent NC & BP activation (Figure 71Bi) and bindings (Figure 71C). In the presence of the aforementioned alternating neuromodulators, the hierarchical synfire chain is then restarted, presumably following a top-down call to remember a specific sequence. The relevant NC & BP clusters are then successively activated dependent on the combined activation of synfire hierarchies (Figure 71Bii). We then examine our model in light of recent experimental findings (Michelmann, et al., 2016), where unique temporal signatures were detected for dynamic stimuli in the phase-angle time series. As such, we show the phase-angle time series of cortical regions at encoding (Figure 71Di, at 7Hz), where a phase-reset pattern coincides with the presentation of the pattern (black dotted box). Using representational-similarity-analysis (RSA, Equation 30, as shown in Chapter 7), we can then compare this phase-reset pattern at encoding (Figure 71Di) with the phase-angle time series at recall (Figure 71Dii), resulting in a similarity time-series (Figure 71E) that peaks at the time of the re-instatement of the encoded pattern in cortical regions.

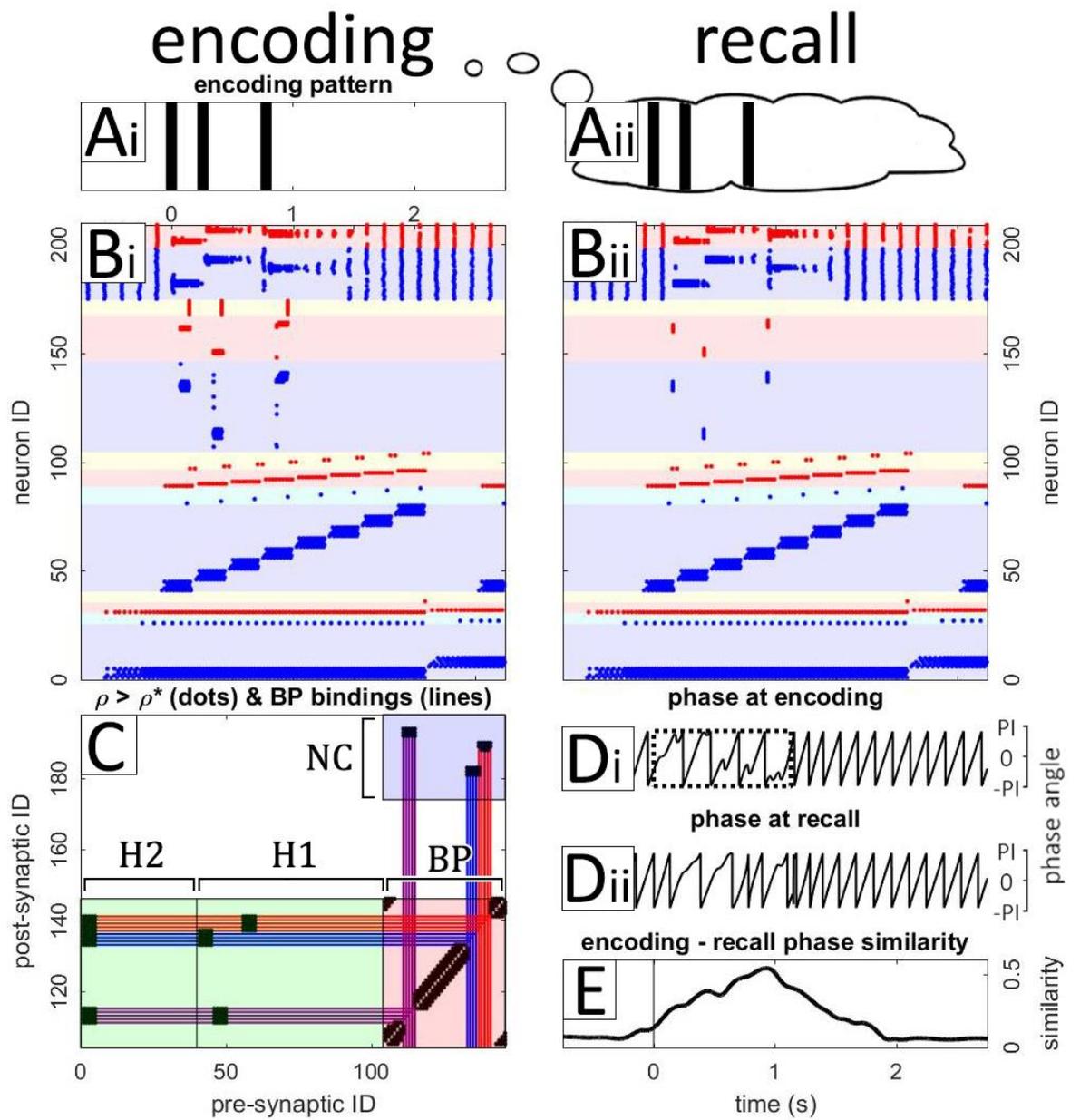


Figure 71 – Analysing a single trial consisting of both encoding and recall phases. A pattern is presented to the neo-cortex (NC) during encoding (**Ai**). A raster plot of spike events (**Bi**) shows subsequent NC & binding pool (BP) activity, where colour-coded dots (blue, excitatory; red, inhibitory) and shaded regions (blue, E; green, P; red, I; yellow, O) depict neuron type. Events are encoded through BP synaptic modification (**C**), where a Boolean (0 or 1; black areas) shows synapses (ρ) above the synaptic reversal state (ρ^*) for pre- & post-synaptic IDs that relate to the neuron ID from the y-axis of **B**. Shaded regions with accompanying notations indicate group connectivity (red, BP<->BP; blue, BP->NC; green, H1/H2->BP) and coloured lines depict BP cluster connections (red-purple-blue lines). The recall phase attempts to re-instantiate the encoded pattern (**Aii**), where

a raster plot shows the successful activation of the relevant BP & NC clusters (**Bii**). From there, the phase-angle time-series (at 7Hz) of the NC region at encoding is shown (**Di**), where a series of phase-resets coincides with the presentation of the pattern at encoding (black dotted box), and the same is shown for the recall phase (**Dii**). A phase similarity measure (**E**, 0-1) indicates when the phase-reset pattern at encoding (**Di**, black dotted box) best matches the phase-angle time series at recall (**Dii**), which was smoothed with a linear filter of 300ms.

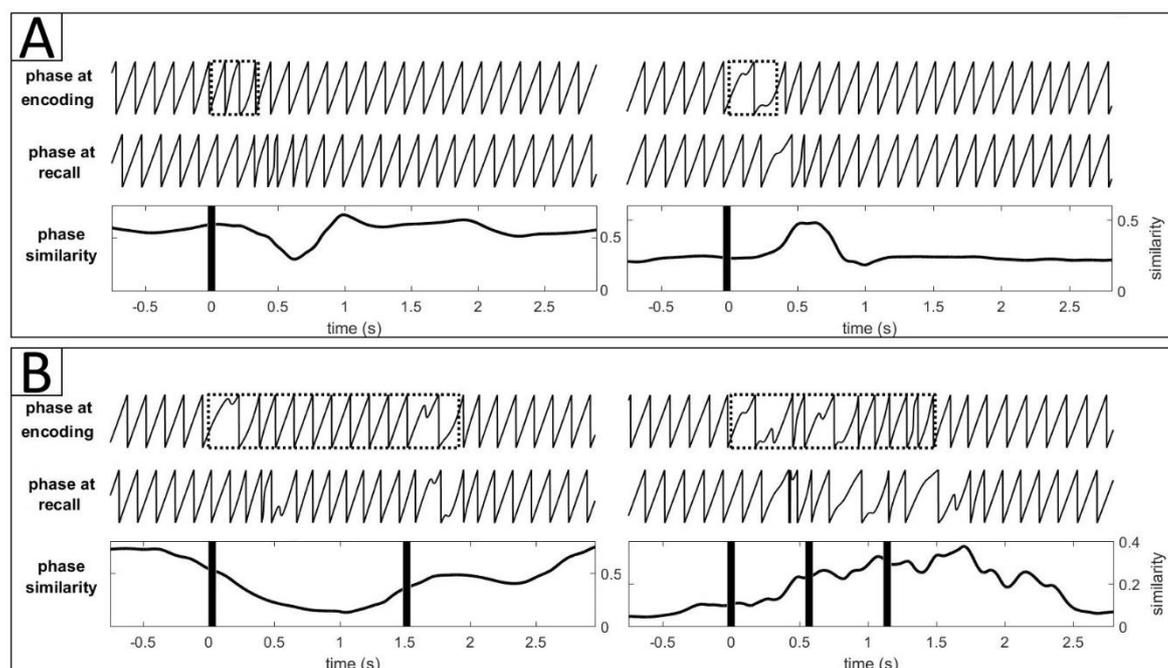


Figure 72 – Analysing the effectiveness of representational-similarity-analysis (RSA, Equation 30) for various circumstances within the model. A stimulus-related (vertical black lines) phase-reset can be slight (**A**; left panel, black dotted box) or major (**A**; right panel, black dotted box) dependent on the conditions of ongoing oscillatory activity. The amount of stationarity in this signal has an effect on the similarity analysis between encoding and retrieval (**A**; bottom panels). This is likewise the case if a phase-reset pattern is caused by events that are far apart (**B**; left panel) or close together (**B**; right panel) in time. Experimentally, one would discard the former scenario due to a large amount of stationarity in the signal. However, the model indicates to us that a pattern of some sort was presented during this time.

The model shows that a phase-reset pattern typically occurs in the presence of desynchronised neural activity (see Figure 64), however, the phase-reset can take various forms dependent on the timing of the event in relation to on-going oscillatory activity. Figure 72 shows how an event-related phasic signature at encoding can be similar to default oscillatory behaviour, and thus during RSA is more likely to produce a non-stationarity (Figure 72A; left panel) rather than a pattern similarity signal (Figure 72A; right panel). This can also occur if events in the pattern are too far apart (Figure 72B; left panel). Experimentally, one would discard these scenarios due to the large amount of stationarity that would fall within an analysis window. Therefore, within the model, presented patterns and the subsequent phase-reset signatures at encoding should contain a minimal amount of stationary signal given a window size if we are to closely follow analysis from the literature. It was chosen that for further simulations we would select patterns with several events that are not too far apart in time, which are known to contain less stationarity in their phasic signatures (as indicated in Figure 72B; right panel). We also simulate several trials of each pattern to ensure that, on average, a robust phase-reset pattern is generated for each event.

8.1.4 CONTENT SPECIFIC REINSTATEMENT OF TEMPORAL PATTERNS

As discussed in the previous Chapter, the experimental paradigm we chose to follow investigated whether one could detect the reinstatement of a temporal pattern from its content (Figure 63, Michelmann, et al., 2016). Therefore, we chose to present many patterns to our model with the aim of distinguishing each by its unique phasic signal. As such, all possible combinations of patterns consisting of 3 distinct events over a period of 1.5s were presented, with the first occurring at 0s and a minimum spacing of .235s between events, where each of these 10 patterns were repeated 10 times (Figure 73Ai; P1-P10). During encoding, we observed the familiar desynchronisation of the ongoing frequency (Figure 73Bi, 0-1.5s at 7Hz) and the resulting non-stationarity of the phase-angle time series (Figure 73Ci, 0-1.5s at 7Hz). The average similarity in this phasic signal between trials of the same pattern at encoding is also shown (Figure 73Di-Dii, 0-1.5s at 7Hz), where a positive number indicated high similarity and thus a robustness across phase-reset patterns.

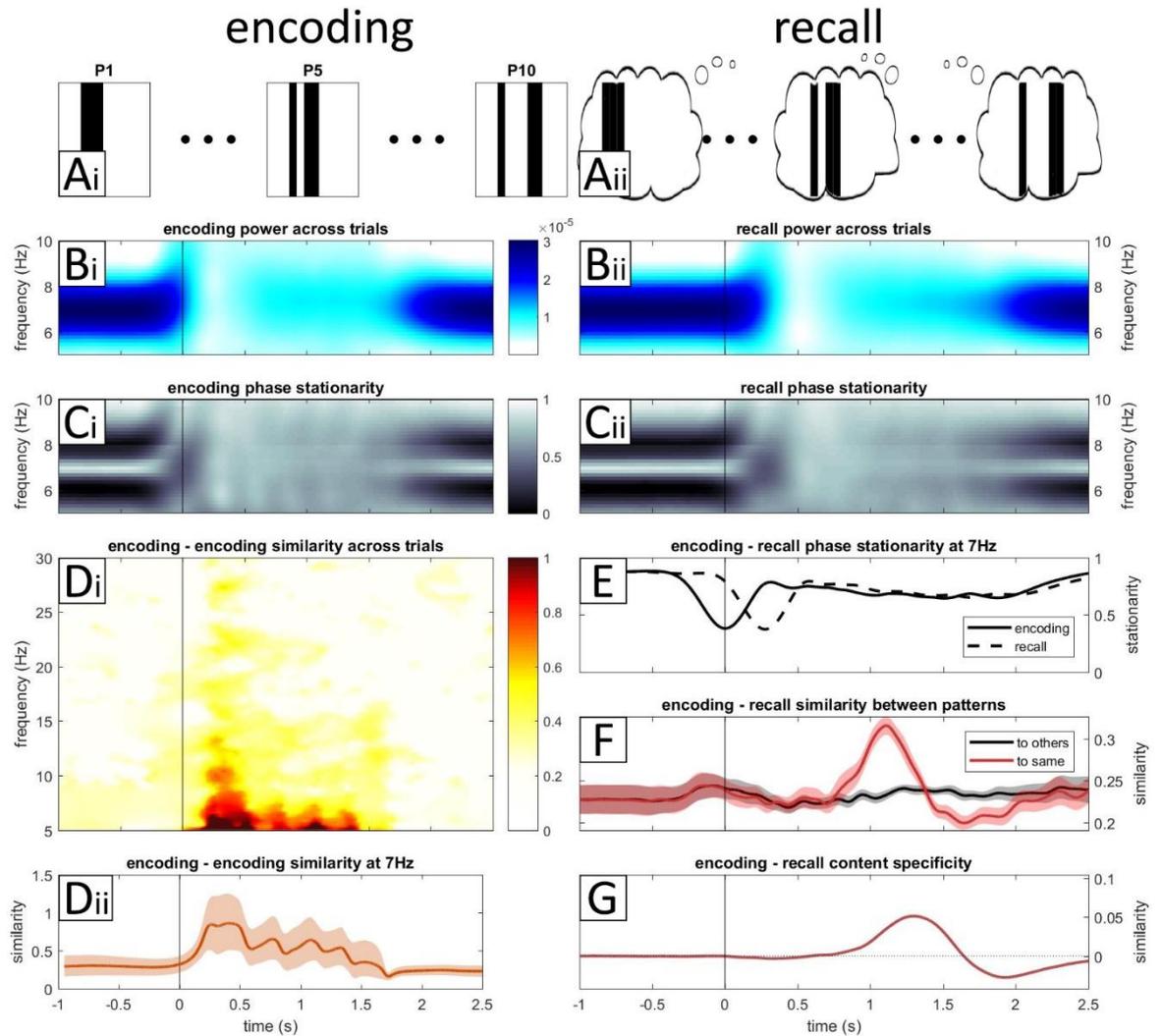


Figure 73 – Content specificity over several patterns comprising of 3 distinct events, where every possible combination of event timings is used over a period of 1.5s with gaps of .235s between events, with each pattern containing an event at 0s (Ai; P1-P10) and being repeated 10 times. During encoding, a time-frequency analysis shows a desynchronisation of the ongoing frequency across all patterns and trials (Bi; 0-1.5s at 7Hz, black vertical line indicates stimulus onset). Similarly, stationarity in the phase-angle series decreases at stimulus onset (Ci; 0-1.5s at 7Hz, 1 = stationary), indicating a period of phase-resets across patterns and trials. There is also a high degree of phase-angle similarity across time and frequency (Di; $\cos A \cdot \cos B + \sin A \cdot \sin B$) between the same patterns across trials at encoding (0-1.5s, >0), particularly at 7Hz (Dii; shaded region indicates confidence intervals across patterns and trials, 50ms linear smoothing filter applied). We then attempted to retrieve patterns during recall (Aii), where a similar desynchronisation was found (Bii) and non-stationarity detected (Cii). By comparing non-stationarity at 7Hz, it was found that there

was a lag in pattern representations at recall compared to encoding (**E**; encoding, black line; recall, dashed black line). Representation-similarity-analysis (RSA; Equation 29) was then performed between encoding and recall phase-angle time series (**F**), where the average of similarities between the same patterns is indicated as a red line and similarities to other patterns as a black line (shaded regions indicate confidence intervals across trials and patterns, 50ms linear smoothing filter applied). As in the paradigm presented in Michelmann et al. (2016), the difference between these similarities (**G**; same – others, 500ms linear smoothing filter applied) was interpreted as evidence for content specific reinstatement of temporal patterns.

As shown in Figure 73, the recall phase entailed restarting the hierarchical synfire chain, such that any relevant bindings are reactivated in sequence (Figure 73Aii). During this time, we again observed the familiar desynchronisation (Figure 73Bii) and non-stationarity of phasic signals (Figure 73Cii), though with a notable lag. This is best observed in Figure 73E, where the stationarity of the phase-angle series at 7Hz is shown for encoding (black line) and recall (black dash). Here, there is a difference of $\sim .3s$ in the non-stationarity between these phasic signals, indicating that the upwards direction of retrieval processes takes longer than the downwards direction of encoding processes, which is also indicated by experimental findings (Michelmann, et al., 2016). Neurophysiologically, there are likely to be many neuronal layers for detecting and processing stimuli that would subsequently increase this recollection lag. In the model, however, this is mostly due to the fact that BP units fire late in their respective temporal window (as indicated by Figure 70). This lag could be reduced within the model by increasing SC->BP weights such that BP units activated earlier in their window. However, this could have the effect of shifting activation forwards from encoding to recall, as events that are bound at a later point during the relatively broad window we have chosen for our *H1* chain ($\sim 200ms$ in length, $\sim 5Hz$), might be replayed at an earlier point. In response to this, one could reduce the error by choosing a finer temporal dimension for the fastest, self-completing synfire chain, such that bindings would be encoded and recalled with more temporal precision. Such a notion might go some way to addressing why high-frequency gamma oscillations ($>40Hz$) are prevalent during episodic memory formation (Sederberg, et al., 2007). This line of

reasoning has also been noted by Fell & Axmacher (2011) and other models of sequence encoding (Jensen, et al., 1996), who argue that gamma provides a fine-grain window of activation that might enable precise communication, learning and the segregation of sequences.

Previous research indicates that a cortical alpha desynchronisation and subsequent phase-reset pattern might signify information flow (Haegens, et al., 2011; Hanslmayr, et al., 2012) and convey information content (Ng, et al., 2013; Schyns, et al., 2011), respectively. Here, we have shown some theoretical evidence for these findings, where a reduction in power (Figure 73Bi-ii) and phasic non-stationarity (Figure 73Ci-ii & D) signifies the active representation of information. Further, experimental evidence has suggested that one can decode information content through analysis of the phase-angle time series, enabling the identification of dynamic stimuli through examination of their unique temporal signatures (Michelmann, et al., 2016). We next show that this position can also be supported theoretically through the use of computational simulations. By presenting many unique temporal patterns to our model (Figure 73Ai), we can compare phasic signatures from encoding to retrieval between trials of the same pattern as well as to other patterns (Figure 73F). Here, we use RSA, as in other research (Michelmann, et al., 2016), to compare these phase-reset patterns, where a red line indicates the average similarity between trials of the same patterns and a black line indicates the average similarity between trials of different patterns. The relatively flat line (<0.25) of the latter indicates that patterns were unique enough that on average, they did not resemble other patterns, whereas the peak in the former (>0.3 , $\sim 1-1.5s$) indicates that the phasic time-series of each pattern across trials was robust enough that on average, they resembled the same temporal signature. Taken together, the difference between these similarities (Figure 73G; to same minus to others) can indicate the content specific reinstatement of unique temporal patterns. The lag in the reinstatement of these patterns ($\sim 1s$) can partly be attributed to the lag discussed in the previous paragraph, and might partly be attributed to the nature of the patterns presented. The dip in performance after content specific recognition ($\sim 1.8s$), which is not present in the literature (Michelmann, et al., 2016), might also be due to the nature of the patterns presented and would require further investigating.

8.2 DISCUSSION

We have here presented a novel neural network with three distinct mechanisms. Together, these mechanisms enable us to bind discrete, observable events in time, though each mechanism allows us to explore a distinct set of hypotheses. We first instantiate a neo-cortical (NC) mechanism, where intrinsic oscillations are generated at a resting alpha frequency through recurrent excitatory-inhibitory interactions. Here, we have implemented a winner-take-all (WTA) locally connected random network (LCRN), as has also been described in hierarchical models of vision, recognition and attention (Carpenter & Grossberg, 1987; Itti, et al., 1998; Reisenhuber & Poggio, 1999). By introducing an external event to our NC region, a subset of winning units that coded for that specific stimulus remained active whilst others were silent. During this time the intrinsic frequency was desynchronised (Figure 64E), consistent with the experimental hypothesis that oscillatory desynchronisations are due to increased neural activation (Haegens, et al., 2011; Hanslmayr, et al., 2011a), which occurs due to relevant representations becoming active at the expense of others (Klimesch, et al., 2007). In this way, we lend evidence to our primary hypothesis that desynchronised neural states indicate information flow within the brain (Hanslmayr, et al., 2012), as in the first modelling work of this thesis shown in Chapter 5-6 (Parish, et al., 2018).

By adding a temporal dynamic to our model, we were also able to look at the phase-angle time series of the intrinsic frequency. In the literature, phasic patterns are thought to convey information content (Ng, et al., 2013; Schyns, et al., 2011), where one can even identify a stimulus by a unique temporal signature that can be used to later detect its re-instatement in memory (Michelmann, et al., 2016). The model lends theoretical evidence to both of these notions, showing that phase-reset patterns occur at a time when representations are active within our modelled NC region (Figure 64F). Taking this further, we show how the model can store and recall a sequence of events (Figure 71), where representational-similarity-analysis (RSA) was used to show that phase-reset patterns at encoding and recall were sufficiently similar. Thus, we provide computational support that phase is an important predictor in conveying information content (Ng, et al., 2013; Schyns, et al., 2011).

By presenting many sequences to our model, we were then able to show that one could on average detect a temporal stimulus by its phase-angle signature, so long as sequences were sufficiently unique (Figure 73F-G), lending theoretical evidence to the experimental technique of using RSA to identify the content specific reinstatement of unique dynamic stimuli (Michelmann, et al., 2016).

Whilst we have successfully generated identifiable phase-reset patterns in our model, the best method of conveying information through this means remains unclear. Events were generated using an Alpha function (Equation 15; $\tau_s = 60ms$) modulated series of excitatory spike events, though they could equally be generated in shorter or longer durations with varying degrees of strengths. It would be interesting to demonstrate if the timing of these events relative to the up-down states of the intrinsic frequency influence whether they would be remembered or not, as has been found in the literature (Klimesch, et al., 2007). Within the model, this would hypothetically rely on the strength of cortical I->E synapses, as well as the strength and duration of external events. Other models have also found that several regions, each with an independent intrinsic oscillation, can align in phase through the mediation of a coordinating pacemaker (Vicente, et al., 2008). As this is a hypothetical mechanism through which independent cortical columnal alpha oscillators are thought to align in phase, possibly through recurrent thalamo-cortical loops, it would be of further interest to instantiate a similar environment and assess whether phase can be robustly used to convey information content. One could then also explore whether a priming event might cause a general phase-reset in thalamic-cortical columns, as has been hypothesised to occur during the EEG P1 component of general recognition (Hanslmayr, et al., 2011b). This might ensure that cortical regions are pre-aligned in phase and thus optimally entrained to a given stimulus, such that any subsequent phase-reset patterns might be more consistent across trials and not as reliant on the intrinsic oscillatory conditions, as was found to be an issue in the model (Figure 72A).

Episodic memories have an inherent temporal dynamic, whereby studies have suggested that our perception is not continuous but is rhythmically sampled in discrete alpha-frequency time-steps (Hanslmayr, et al., 2013; Landau & Fries, 2012; VanRullen, et al., 2007). It might therefore be the

case that there is a more qualitative element to alpha desynchronisations. As neocortical and hippocampal gamma have both been found to be important when predicting successful memory encoding (Sederberg, et al., 2007), it might be the case that an alpha desynchronisation deregulates alpha-frequency inhibition in order for a gamma-frequency sequence to be allowed to activate. In this sense, our neocortex within the model might be arranged similarly to a single feed-forward chain, as in other models (Goldman, 2009), whereby the cortical sequence might inherit an equally spaced hippocampal sequence during encoding, as other models have shown (Itskov, et al., 2011). The deregulation of alpha inhibition might then enable the sequence to play in full, both decreasing alpha power and increasing gamma power.

The second mechanism within our model ensured that events were treated discretely such that they could be bound as independent components of a sequence. This method bears similarity to other models (Bowman & Wyble, 2007), where it was argued that the use of a binding-pool (BP) enables repetition. It would be interesting to analyse this effect further within our model by contrasting a simulation of the same episodic paradigm with and without the binding pool. It was found through the design of this model that if bindings were direct from synfire chain to cortical cells, then the cortical cells would activate upon every possible iteration of its bound hierarchical components. For example, if events were bound to the 3rd part of the 1st period (1,3) as well as the 1st part of 2nd period (2,1), then during recall the cortical assembly would activate at both the 1st and 3rd parts of both the 1st and 2nd period (1,1; 1,3; 2,1; 2,3). The best prevention method of this is if there were only one layer of synfire chains, as has been hypothesised in other models (Itskov, et al., 2011), however, then the timed duration of the synfire sequence must be long enough to encode for the entire episodic sequence, which is often a criticism levied against this type of model (Shankar & Howard, 2012).

An interesting finding of our model is the effect of the 'off switch' mechanism in the binding pool that operates to prevent spreading activation across the excitatory population. Here, it was found that a new event would not be encoded during this inhibitory pulse, as cortical impulses could not

overcome the increased global inhibition. This is reminiscent of the hypothesised function of binding in other models (Bowman & Wyble, 2007), where it is thought that attentional deficits for events very close together in time arise as a consequence of the need to limit the temporal extent of the binding process. In our model, it is sufficient that the synaptic time constant from the 'off' cells (O) to excitatory cells (E) merely match that for E to E cells such that it quenches the possibility that activation 'snowballs'. However, if we are to relate this hypothesis (see Bowman & Wyble, 2007, & Swan & Wyble, 2014, for overview of binding errors and the attentional blink) to our model, then this $O \rightarrow E \tau_s$ might be larger to act as a buffer between events, minimising binding errors as well as acting in its default role of preventing runaway activity. This might especially be necessary when encoding over very fine grain temporal dimensions, where a global inhibitory pulse could be useful in separating events close together in time, though it risks causing an attentional blink for those that follow before inhibition subsides. Importantly, our model would also allow for multiple events to be encoded within the same binding location, possibly leading to order errors when replaying these events later, as has also been found in the literature for the encoding of events close together in time (Botella, et al., 1992; Wyble, et al., 2009). It would be of interest to simulate more of these paradigms, as in other models (Bowman & Wyble, 2007; Swan & Wyble, 2014), to assert whether these findings can be robustly linked to the functional limitations when modelling an iteration of a binding pool.

Additionally, it would also be of interest to explore sequence completion issues in the model, possibly arising from the presence of multiple and overlapping sequences. Theoretically, the winner-take-all aspect of the binding pool in the model would enable a "winning pathway" to emerge from any overlapping nodes, possibly causing temporal sequences to "switch tracks" part way through sequence replay. This resonates with a modelling exploration of overlapping synfire chains in Kumar, et al., 2008, which showed how simultaneous signals can be kickstarted on multiple chains through overlapping nodes.

The third mechanism within our model builds on the use of feed-forward synfire chains in the encoding of time (Goldman, 2009; Itskov, et al., 2011). One of our principle purposes, however, is to address the criticism levied at such models (Shankar & Howard, 2012) that they are unlikely to be used for temporal encoding due to the fact that the physical dimensions of the chain have to increase linearly with temporal requirements. This complexity issue is compounded when one considers that episodic sequences can be in the many 10s of seconds (Howard & Eichenbaum, 2013) and that there are likely many competing temporal lists (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). We address this issue by proposing a biophysical neural network model of persistent activation, that uses recursive feedforward and feedback connections to instantiate simultaneous temporal hierarchies. In this way, each hierarchical layer repeats, decreasing the computational complexity issue raised by Shankar & Howard, 2012, and events are encoded by their association to a point on multiple time-scales that together encode for a unique moment in time. In doing so, we also address another criticism that has been levied against recurrent mechanisms that enable persistent activity, where it has been argued that their use is unlikely to encode for long temporal durations due to the ever decreasing weights that would be required to sustain activity (Goldman, 2009). This is overcome by the recurrent architecture of our simple and compact neuronal assembly (entailing E, I, P & O intra-connectivity, see Figure 59) that can hypothetically be scaled up indefinitely for any number of chains of any length. This assembly ensures that recurrent activation can be maintained indefinitely with no adjustments to the parameter set, where a feedback signal is required in order to terminate activation and move the signal to the next neuronal group. Despite these feedback projections, the overall signal is forward moving, hence our model can be described as a “feed-forward model in disguise,” as also noted by Goldman (2009).

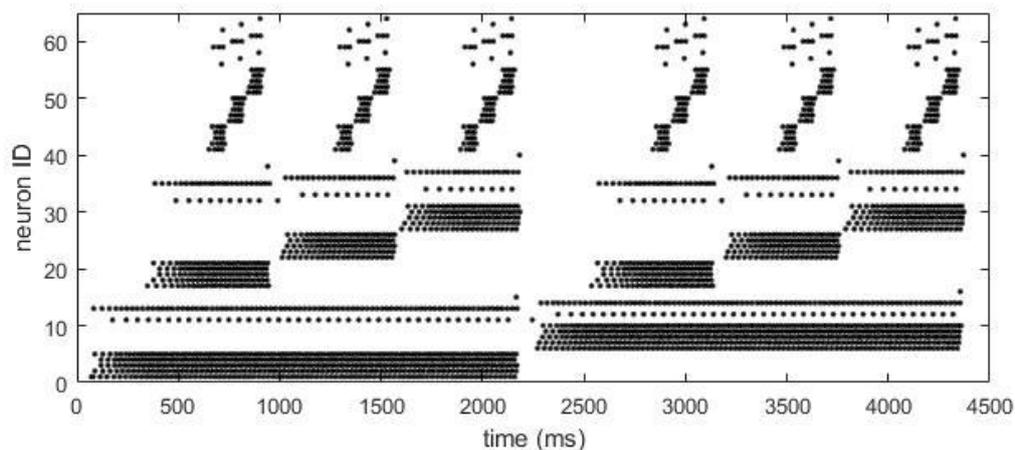


Figure 74 – Three distinct layers of synfire-chains, where each respective layer modulates phases of learning for each lower-order layer and items are segregated in oscillatory-like time-steps.

This leads us to the second consideration of our time-keeping mechanism: as to how the multiple co-existing and simultaneous time-keepers in the brain (Mauk & Buonomano, 2004) might interact, update and contextualise one another (Aggleton, et al., 2010; Eichenbaum, 2014). In proposing a neural substrate for such phenomena, we consider our feedback update method between temporal hierarchies, where a parameter-dependent ‘handover’ period might exist (see Figure 62A). This raises several interesting questions, the first of which is how the now well-established phenomenon of order-errors (Bowman & Wyble, 2007; Wyble, et al., 2009) and conjunction errors (Botella, et al., 1992) might arise in episodic memory. Here, one can choose whether to maximise ‘what’ or ‘when’, minimising missed stimuli or binding errors, respectively. Secondly, a period of unresponsiveness during this ‘handover’ period might provide a theoretical explanation for the origin of observable oscillatory dynamics, as gamma, theta and delta frequencies have all been found to be important for encoding and sequencing memories by modulating phases of learning. With respect to this idea, our model can be compared to a popular model of theta-gamma coupling (Jensen, et al., 1996), where items are segregated in time by gamma and theta time-steps. In our model, each respective synfire chain hierarchy might encode for a slower frequency, where items are bound in the respective up-states of gamma-theta coupling, and ‘handover’ periods at each hierarchy segregate memories into distinct episodes, as hypothetically shown in the simulation of Figure 74.

It might also be the case that a signal could traverse chains at varying speeds dependent on the strength of the initial burst or levels of background noise, as other feed-forward synfire models have explored (Diesmann, et al., 1999; Kumar, et al., 2008). In this way, one could replay events at a faster rate, as has been observed during sleep (Diekelmann & Born, 2010). It might also be the case that some attentional mechanism can skip through scene-sized chunks, as indicated by recent findings (Michelmann, et al., 2018), where selective additional excitatory input focussed on a specific hierarchy might compress memory replay until a point of interest is identified through an additional feedback mechanism.

As forgetting appears to take place over multiple time-scales with similar decay characteristics (see Howard & Eichenbaum, 2013, for review), one might imagine that previously active nodes within each hierarchy decay with similar time constants (Figure 75). Such a mechanism might allow for the smearing of previous events and precision in the moment, whilst allowing a previous temporal context to be re-instantiated upon the re-activation of a previous node (Howard, et al., 2012). In our model, this would require connections from the binding pool into the temporal region, such that the re-activation of an event in memory would re-start the temporal context from a new position.

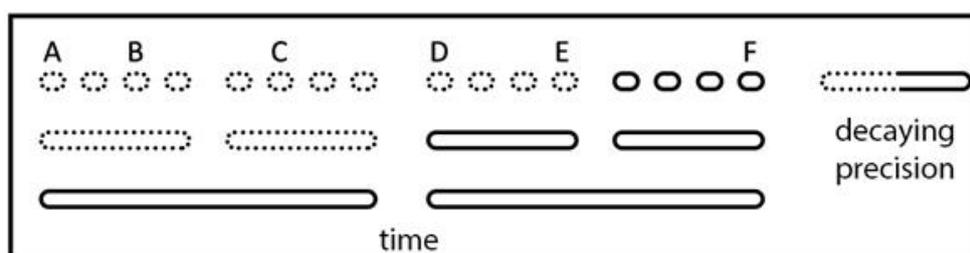


Figure 75 – A decay in activation might lead to decaying precision in temporal awareness, where one might perceive events A/B/C or D/E to have occurred together, whilst the recent event F can be remembered with higher temporal precision. However, the re-activation of event A would re-instate a previous temporal context.

An important component of our hierarchical synfire chains is the use of a WTA LCRN, where surround inhibition is essential for the maintenance of persistent activity at a singular moment in time. This architecture has also been hypothesised to be important for our navigational capacity (McNaughton, et al., 2006), as we must maximise the difference between our current and other irrelevant positions. Considering that place and time cells are thought to be the same type of cells that together integrate temporal and spatial information (Eichenbaum, 2014; Moser, et al., 2008), one might consider that a similar WTA environment frames our temporal perception, as hypothesised by other models (Itskov, et al., 2011). As noted in the previous Chapter, contention between positions on temporal lists has also been identified in the hippocampal literature (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). This could be further explored within the scope of our model, where one could add another dimension within our hierarchical synfire chains, such that assemblies at the group level (g) might spread in a horizontal dimension as well as the already defined group and layer dimensions (e.g. $X(p, g, l)$, where X denotes neuron type, p denotes position in the horizontal dimension, g denotes group and l denotes layer, as shown in Figure 59). In doing so, horizontal dimensions of fast-excitatory (E) and fast-inhibitory (I) cells might ensure that surround inhibition would select the most appropriate assembly at each moment in time. However, this would assume the presence of some feedback interactions that would be important in the WTA selection process for each position in time.

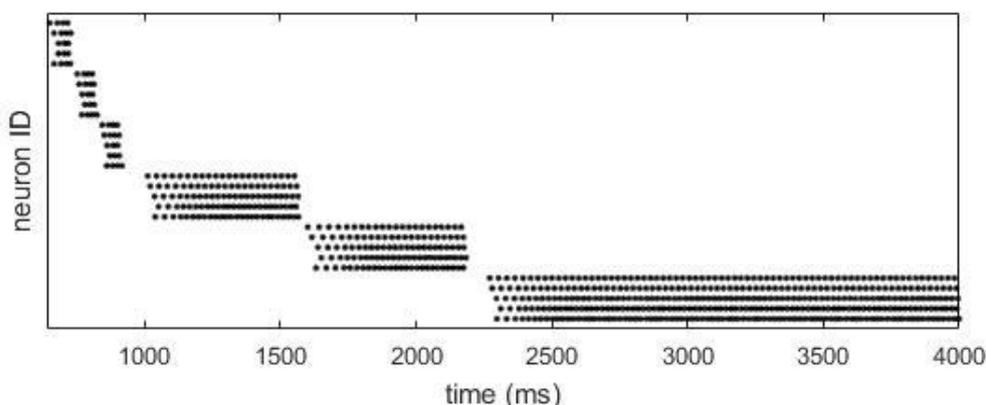


Figure 76 – Re-ordering synfire-chain activity by time of firing to more closely match time-cells.

Finally, our hierarchical synfire-chains resemble time cells in the sense that cells activate for a unit amount of time in sequence to cover a timed duration (Eichenbaum, 2014), where there is a scalar encoding of time in the sense that later cells in the sequence activate for longer durations (Howard & Eichenbaum, 2013). The difference in our model is that we have created multiple simultaneous hierarchies, such that there are multiple cells encoding for the same moment in time. To better match time cell findings, one could re-arrange neurons in our model by time of firing, as in Figure 76. One can immediately see that it might be possible that observed time cells in the brain are but a selection of the total number of cells that could be encoding our temporal dynamics. We therefore hope that this model has provided an alternate mechanism that might be describing observed neurological phenomenon, however, many more investigations of our model are required before such a hypothesis could be asserted.

Here we discuss the modelling work of this thesis in light of our two main hypotheses, defined in Chapter 1, namely, that it can be shown computationally that neural oscillations both gate information flow and learning in the brain. As such, we aimed to more generally lend theoretical evidence to empirical findings from the brain oscillations literature. To do so, we created neural network models in a bottom-up fashion, modelling neurophysiological processes with emergent behaviours. These behaviours can generally be typified as the interactions between complimentary learning systems (CLS) (O'Reilly, et al., 2011). Here, it was proposed that two distinct memory systems are required if one is to avoid the problem of catastrophic interference (McCloskey & Cohen, 1989), where the encoding of new information is likely to erase pre-existing information. However, such learning is required to allow the incorporation of new information (McClelland, et al., 1995), where long-term memory must be flexible enough to allow new information to be gradually incorporated. This problem is often also framed as the 'plasticity-stability' dilemma (Carpenter & Grossberg, 1988), which assumes there must be a balance between plastic learning in an online environment, and the storage of stable long-term memories. The CLS framework argues that the brain solves this dilemma through the interactions of a plastic hippocampus with a high-learning rate, which navigates our online existence through the recognition of patterns and encoding of novel information, and a stable neo-cortex with a slow learning rate, where distributed representations store our pre-existing knowledge (O'Reilly, et al., 2011). As such, novel memories learnt throughout the day are thought to be rapidly presented to neo-cortical regions during the altered brain state of sleep (Diekelmann & Born, 2010; Walker, 2018), where they are thought to be contextualised and incorporated into long-term memory (Winocur, et al., 2010).

In this thesis, we expand the use cases of the CLS framework through the simulation of distinct oscillating brain regions, as has been observed in the brain (Buzsaki, 2002; Basar, et al., 2001; Fell & Axmacher, 2011), thus providing an alternate mechanism through which these interacting learning systems might gate fundamental neurological processes. In our first model, described in

Chapters 5 & 6, we create a dual neo-cortical/hippocampal model, where each distinct brain region oscillates at an intrinsic frequency. In compliance with the CLS framework, this model instantiates a fast-learning hippocampus that is capable of 'one-shot' learning, where it has been observed that new concepts can be encoded after a single presentation (Ison, et al., 2015). In tandem with this system, a neo-cortical region was instantiated that stores a stable version of concepts, where learning was disabled as it was assumed that plasticity would occur at a much slower rate than the length of simulated episodic memory paradigms. By introducing the concept of oscillations, however, we allowed these co-operating brain regions a means with which to better communicate, as it is thought that oscillations provide precise windows of communication (Fell & Axmacher, 2011) that structure information flow (Klimesch, et al., 2007) and distribute phases of plasticity (Buzsaki, 2002). Therefore, in the Sync/deSync model (Parish, et al., 2018), information representation in the stable neo-cortex is facilitated by the desynchronisation of its local frequency, whereas hippocampal online learning is modulated by its own local frequency. In this respect, we present a method by which complimentary brain regions with co-operative but distinct mechanisms might gate the bi-directional flow of relevant information and facilitate learning.

The purpose of the second modelling work of this thesis was to further expand both of our main hypothesis, namely that it can be shown computationally that neural desynchronisations can both allow information content to emerge and play a functional role in learning. As such, we formulated a complimentary temporal system that enabled the encoding of unique episodic information, such that we could encode more information about the content of a stimulus from a period of cortical desynchronisation. We go on to show how sequences might be encoded in a plastic hippocampus, and hypothesise as to how these sequences might be stored in long-term cortical memory, where interactions between these memory stores are thought to occur during recognition and familiarity based memory (Eichenbaum, et al., 2007). At this time, the concept of memories being stored as sequences is only enabled in the hippocampal region, but future simulations might show how cortical sequences could be inherited through repeated exposure, as other models have shown (Itskov, et al., 2011). This would allow us to store these newly acquired sequences as stable long-

term memories and perform more complex learning paradigms between concurrently active concepts, as in the previous model (Parish, et al., 2018). Despite this, this second model again highlights the role that a fast-learning region might have in providing phases of learning. In the first Sync/deSync model, this functional role was achieved by using a phase-dependent learning rule. In the second model, hierarchical time-keepers provide windows during which an item will be bound in an oscillatory fashion (see Figure 62A, 'encoding when for what'). The second model thus extends the notion that memory depends on the interactions between a phase-dependent & fast-learning system, with a slow-learning system where the representation of qualitative information is regulated by oscillatory dynamics.

As well as providing support for the CLS framework, we also provide computational evidence for the mechanisms that oscillations appear to provide that enable communication, learning and information representation within the brain (Buzsaki, 2002; Basar, et al., 2001; Fell & Axmacher, 2011). Recent evidence-based hypotheses have argued that oscillations provide precise windows of communication during the relative up-states of specific frequencies, where distinct neurological regions have been found to intrinsically oscillate at different frequencies (Fell & Axmacher, 2011). The phase synchronisation (see Appendix A for more discussion) between regions has been used as further evidence as to which region is initiating communication, where the receiving region tends to synchronise in the unique frequency of the initiating region (Fell & Axmacher, 2011). It has been further shown that distinct frequencies perform different neurological functions and are active during different brain states (Basar, et al., 2001), lending evidence to the CLS framework hypothesis that brain regions have distinct mechanisms but co-operative purposes (O'Reilly, et al., 2011).

With respect to our primary hypothesis, cortical alpha frequency (~10Hz) oscillations are thought to gate information representation in the brain (Klimesch, et al., 2007), where a synchronisation implies higher levels of inhibition that prevents interference from irrelevant representations (Sauseng, et al., 2005a), whilst a desynchronisation implies the activation of relevant representations (Haegens, et al., 2011; Hanslmayr, et al., 2011a). A desynchronisation has also been

shown to have a qualitative aspect, where the analysis of phase shows that information content can be deciphered from unique temporal phase-reset signatures (Michelmann, et al., 2016). Both of the modelling explorations within this thesis address the first consideration, where a desynchronisation of alpha in both the Sync/deSync model (Parish, et al., 2018) and the hierarchical synfire-chain model indicates a signal of informational processing, as has been hypothesised in the literature (Hanslmayr, et al., 2012). We go on to show in the Sync/deSync model that a desynchronisation is correlated to neural activation (see Figure 50 in Chapter 6), which occurs due to the active representation of relevant concepts. The simplicity of the Sync/deSync model, however, does not enable the encoding of sequence-driven episodic memory, therefore it is difficult to extrapolate any more information about the content of active representations. This motivated us to design the hierarchical synfire-chain model, whose purpose is to encode discrete events in time, such that unique temporal patterns can be stored and replayed. In doing so, the second modelling work of this thesis provides computational support to the process of identifying patterns from alpha-phase, such that one can decipher information content (Ng, et al., 2013; Schyns, et al., 2011). In the model, it was further demonstrated that one could indeed detect the re-instatement of these unique temporal signatures (see Figure 73 in Chapter 8), as has been shown experimentally (Michelmann, et al., 2016), provided that encoded episodic content was sufficiently unique.

With respect to our second hypothesis, hippocampal theta frequency ($\sim 4\text{Hz}$) oscillations are thought to modulate phases of plasticity in the brain (Buzsaki, 2002), where a synchronisation has been found to predict successful memory encoding and retrieval (Backus, et al., 2016; Burke, et al., 2014) and the precise timing of spikes relevant to the prevalent theta oscillation determines synaptic plasticity (Huerta & Lisman, 1995). This gives rise to the conundrum that both a desynchronisation of neo-cortical alpha and a synchronisation of hippocampal theta have been argued to be predictors of memory formation (Hanslmayr, et al., 2016), which the first model in this thesis addresses. In the Sync/deSync model, we address the question of how a synchronisation can simultaneously be 'good' and 'bad' for memory by approaching the problem in a CLS framework. In such a framework, we argue that these two frequencies of distinct neurological

origins provide different mechanisms for complimentary purposes, where the desynchronisation of alpha signals the activation of representations whilst the synchronisation of theta signals the utilisation of learning processes. These mechanistic differences are achieved through the simple addition of an after-depolarisation function for hippocampal units, which has the effect of slowing down spike-rates to more closely resemble experimental observations (Ison, et al., 2015; Rutishauser, et al., 2010). We went on to show that in general, unit increases over stimulus strength in the model had the effect of first synchronising an ongoing frequency, as more units were active in the relative up-states, then desynchronising it, as more units overcame inhibition to fire in relative down-states (see Figure 50 in Chapter 6), a process hypothesised elsewhere (Klimesch, et al., 2007). In the case of our simulated complimentary learning systems, this had the effect of desynchronising cortical alpha whilst simultaneously synchronising hippocampal theta. Additionally, we were able to explain a range of divergent findings in the field where the desynchronisation of hippocampal theta was equally found to be a predictor of memory encoding (Crespo-Garcia, et al., 2016; Fellner, et al., 2016), by simply increasing stimulus strength past 'normal' levels. We also show with this model that active cortical representations could push their hippocampal counterparts forward in theta phase, to a phase where plasticity is maximal and the majority of cells are silent. This provides an oscillatory mechanism through which complimentary learning systems might dynamically interact with one-another to encode memories. We also provide evidence consistent with observations that alpha and theta dynamics propagate between these brain regions during pre- (Fell, et al., 2011; Gyderian, et al., 2009; Salari & Rose, 2016) & post- (Backus, et al., 2016; Hanslmayr, et al., 2016; Staresina, et al., 2012) stimulus periods (see Figure 53 in Chapter 6), further supporting the presence of communication between distinct neurological mechanisms that are facilitated by oscillatory mechanisms.

Further simulations using both of the models defined in this thesis could yield further evidence for the functional significance of oscillations in the encoding of complimentary memories. It has been hypothesised in other models that theta-gamma phase coupling might mediate complimentary phases of plasticity in working memory, where several items might be stored and reinforced in

sequential slots within a dominant theta cycle (Jensen, et al., 1996). Similarly, we have hypothesised that our hierarchical synfire-chain model might also mediate such behaviour, where a high-frequency sequence might be stored for every node of a low-frequency sequence, possibly giving rise to the observed theta-gamma dynamics. We thus hypothesise a functional role that oscillations might have in segregating our temporal experience into hierarchical episodes, that can then work together to encode and recall contextualised memories. In this theory, we point to findings that the synchronisation of simultaneous cortical and hippocampal gamma has equally been found to be integral to memory (Sederberg, et al., 2007), leading us to suggest that long-term cortical memories are stored as high-frequency sequences that are inherited through repeated exposure to their hippocampal counterparts, as other models have shown is possible (Itskov, et al., 2011). In the first edition of the Sync/deSync model within this thesis, we do not analyse gamma frequencies as we have not implemented any mechanistic source for such a frequency. However, one might imagine how observed desynchronisations of alpha might release cortical sequences from inhibition, enabling the reinstatement of unique cortical sequences at a gamma frequency. This implementation might go some way to address a common criticism of the Sync/deSync theory (Hanslmayr, et al., 2016), as it is thought that conjunctive representations of a relevant stimulus are best facilitated through neuronal synchronisation rather than desynchronisation (Fries, 2005).

Leading up to the design of our hierarchical synfire-chain model, which is concerned with the temporal aspect of episodic memory in the brain, we have examined a feed-forward mechanism that might be the means for a communications ‘highway’ in the brain (Diesmann, et al., 1999; Kumar, et al., 2008). Here, consecutively connected groups of neurons can synchronise a volley into precisely timed packets of information, dependent on the conditions of an initiating burst (Diesmann, et al., 1999) and an embedding network (Kumar, et al., 2008). It has been proposed that this structure can produce behaviour similar to time-cells (Itskov, et al., 2011), which are medial-temporal-lobe (MTL) cells that have been observed to transiently activate in sequence to cover a timed duration (MacDonald, et al., 2011) and have been related to many of the timing aspects of episodic memory (Eichenbaum, 2014). However, a criticism regarding the scalability of such a

solution has emerged, whereby the physical size of the chain must increase linearly with the required temporal duration (Shankar & Howard, 2012). The encoding of novel episodic memories might last for many 10s of seconds (Mauk & Buonomano, 2004) and there also appear to be many chains that compete at every temporal position (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000), such that we can navigate our online existence as is thought to occur during spatial navigation (McNaughton, et al., 2006). As these findings likely compound the initial criticism, Shankar & Howard's (2012) suggestion that feed-forward chains are 'unlikely' to encode for time seems plausible. Thus, a further consideration of this thesis is to propose a variation on the synfire chain approach that might address this legitimate criticism. The second modelling work of this thesis creates a hierarchical synfire-chain model, whereby each chain repeats for every transiently active node of a higher-order chain. As such, hierarchies initiate and terminate one another through a series of feed-forward and feed-back connections. This is essentially a 'feed-forward chain in disguise' (Goldman, 2009), as a continuous forward moving signal traverses multiple feed-forward chains simultaneously. Using such a structure, we substantially increase the capacity of feed-forward chains in the encoding of time, allowing a small set of cells to encode for very long durations. With the addition of another dimension, we also envisage an alteration that would allow competition to occur between several chains at every moment in time whilst maintaining temporal context, as has been observed experimentally (MacDonald, et al., 2013; Pastalkova, et al., 2008; Wood, et al., 2000). We also theorise that such a solution could be a possible neural substrate that would enable communication between multiple simultaneous time-keepers in the brain (Mauk & Buonomano, 2004), where it has been found that distinct neural projections appear to be critical for the contextualisation of familiarity- (Eichenbaum, et al., 2007) and recollection- (Aggleton, et al., 2010) based memory. The primary objective of our temporal structure is thus to provide a workable framework for a feed-forward chain in the encoding of time that responds to previous criticisms, as well as to stimulate debate on how simultaneous temporal hierarchies might contextualise one-another.

A secondary prediction of this model arose during the simulation of an episodic learning paradigm, where it was important to instantiate a 'go-between' layer that dealt with the binding of events to a moment in time, as in other models (Bowman & Wyble, 2007). Due to our hierarchical method for encoding time, conjunction errors of a temporal context would occur without this binding-pool layer, as any repeating event would get bound to multiple moments in time. Conjunction errors have been frequently observed experimentally (Botella, et al., 1992), and have been shown to be related to the functional pre-occupation of a binding resource in another model (Chennu, et al., 2011). Therefore, we argue that episodes are 'recast' as discrete events, where a unique identifying tag is associated with any contextualising elements of the episode. This tag is also bound to a moment in time (and hypothetically in space), thus enabling the differentiation of a repeating item in various positions on a temporal list (and hypothetically between spatial positions).

With our two main hypothesis in mind, we hope to have created neural network models that provide support for a CLS framework in the brain (O'Reilly, et al., 2011); expand the notion of the functional operations that brain oscillations might provide for communication (Fell & Axmacher, 2011), learning (Buzsaki, 2002) and information representation (Hanslmayr, et al., 2012; Klimesch, et al., 2007; Michelmann, et al., 2016); as well as to offer an alternative feed-forward mechanism that might enable the hierarchical contextualisation of simultaneous time-keepers in the brain (Mauk & Buonomano, 2004).

REFERENCES

- Abeles, M., 1991. *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.
- Abeles, M., Bergman, M., Margalit, E. & Vaadia, E., 1993. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, Volume 70, pp. 1629-1638.
- Adrian, E. & Matthews, B., 1934. The Berger rhythm: potential changes from the occipital lobes in man. *Brain*, Volume 57, pp. 355-385.
- Aggleton, J. P. et al., 2010. Hippocampal–anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *European Journal of Neuroscience*, Volume 31, pp. 2292-2307.
- Aisa, B., Mingus, B. & O'Reilly, R., 2008. The Emergent neural modeling system. *Neural Networks*, 21(8), pp. 1146-1152.
- Andersen, P. & Andersson, S. A., 1968. *Physiological Basis of the Alpha Rhythm*. New York: Appelton Century Crofts.
- Andrade, R., 1991. Cell excitation enhances muscarinic cholinergic responses in rat association cortex. *Brain Resources*, pp. 81-93.
- Araneda, R. & Andrade, R., 1991. 5-Hydroxytryptamine 2 and 5-hydroxytryptamine 1A receptors mediate opposing responses on membrane excitability in rat association cortex. *Neuroscience*, pp. 399-412.
- Backus, A. et al., 2016. Hippocampal-prefrontal theta oscillations support memory integration. *Current Biology*, Volume 26, pp. 450-457.
- Basar, E., Basar-Eroglu, C., Karakas, S. & Schurmann, M., 2001. Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, Volume 39, pp. 241-248.
- Berger, H., 1929. Über das Elektrenkephalogramm des Menschen. *H. Archiv f. Psychiatrie*, 87(1), pp. 527-570.
- Bi, G. Q. & Poo, M. M., 1998. Synaptic modifications in cultured Hippocampal. *Journal of Neuroscience*, 18(10), pp. 464-72.
- Botella, J., Garcia, M. L. & Barriopedro, M., 1992. Intrusion patterns in rapid serial visual presentation tasks with two response dimensions. *Perception & Psychophysics*, 52(5), pp. 547-552.
- Bowman, H. & Wyble, B., 2007. The Simultaneous Type, Serial Token Model of Temporal Attention and Working Memory. *Psychological Review*, 114(1), pp. 38-70.
- Brunel, N., 2000. Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience*, Volume 8, pp. 183-208.
- Burgess, N. & Hitch, G., 1999. Memory for Serial Order: A Network Model of the Phonological Loop and its Timing. *Psychological Review*, 106(3), pp. 551-581.
- Burke, J. F. et al., 2014. Theta and high-frequency activity mark spontaneous recall of episodic memories. *Journal of Neuroscience*, 34(34), pp. 11355-11365.

- Buzsaki, G., 2002. Theta Oscillations in the Hippocampus. *Neuron*, Volume 33, pp. 325-340.
- Caesar, M., Brown, D. A., Gahwiler, B. H. & Knopfel, T., 1993. Characterization of a calcium-dependent current generating a slow after depolarization of CA3 pyramidal cells in rat hippocampal slice cultures. *European Journal of Neuroscience*, pp. 560-569.
- Canavier, C. C., 2015. Phase-resetting as a tool of information transmission. *Current Opinion in Neurobiology*, Volume 31, pp. 206-213.
- Carpenter, G. A. & Grossberg, S., 1987. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, 37(1), pp. 54-115.
- Carpenter, G. A. & Grossberg, S., 1988. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21(3), pp. 77-88.
- Chennu, S., Bowman, H. & Wyble, B., 2011. Fortunate Conjunctions Revived: Feature Binding with the 2f-ST2 Model. *Annual Conference of the Cognitive Science Society: Cognitive Science Society*, pp. 2598-2603.
- Church, R. M., 1984. Properties of the Internal Clock. *Annals of the New York Academy of Sciences*, 423(1), pp. 566-582.
- Crespo-Garcia, M. et al., 2016. Slow-theta power decreases during item-place encoding predict spatial accuracy of subsequent context recall. *NeuroImage*, 142(15), pp. 533-543.
- Dayan, P. & Abbot, L. F., 2001. *Theoretical Neuroscience: Computational and Mathematical modeling of neural systems*. s.l.:Massachusetts Institute of Technology Press.
- Diekelmann, S. & Born, J., 2010. The memory function of sleep. *Nature Reviews*, 11(2), pp. 114-126.
- Diesmann, M., Gewaltig, M.-O. & Aertsen, A., 1999. Stable propagation of synchronous spiking in cortical neural networks. *Nature*, Volume 402, pp. 529-533.
- Doppelmayr, M. et al., 2005. Intelligence related upper alpha desynchronization in a semantic memory task. *Brain Res. Bull.*, Volume 66, pp. 171-177.
- Eichenbaum, H., 2014. Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews*, Volume 15, pp. 732-744.
- Eichenbaum, H., Yonelinas, A. R. & Ranganath, C., 2007. The Medial Temporal Lobe and Recognition Memory. *Annual Reviews Neuroscience*, Volume 30, pp. 123-152.
- Fell, J. & Axmacher, N., 2011. The role of phase synchronisation in memory processes. *Nature Review Neuroscience*, Volume 12, pp. 105-118.
- Fell, J. et al., 2011. Medial Temporal Theta/Alpha Power Enhancement Precedes Successful Memory Encoding: Evidence Based on Intracranial EEG. *Journal of Neuroscience*, 31(14), pp. 5392-5397.
- Fellner, M.-C. et al., 2016. Spatial Mnemonic Encoding: Theta Power Decreases and Medial Temporal Lobe BOLD Increases Co-Occur during the Usage of the Method of Loci. *eNeuro*, 3(6).
- Fellous, J., Rudolph, M., Destexhe, A. & Sejnowski, T., 2003. Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. *Neuroscience*, pp. 811-829.

- Fiete, I., Senn, W., Wang, C. & Hahnloser, R., 2010. Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*, 65(4), pp. 563-576.
- Fries, P., 2005. A mechanism for cognitive dynamics, neuronal communication through neuronal coherence. *Trends in Cognitive Science*, Volume 9, pp. 474-480.
- Gebber, G. L., Zhong, S., Lewis, C. & Barman, S., 1999. Human brain alpha rhythm: Nonlinear oscillation or filtered noise?. *Brain Research*, 818(2), pp. 556-560.
- Gerstner, W., Kempter, R., van Hemmen, J. L. & Wagner H, 1996. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, Volume 386, pp. 76-78.
- Gerstner, W., Ritz, R. & van Hemmen, J. L., 1993. Why Spikes? Hebbian learning and retrieval of time-resolved excitation patterns. *Biol. Cybern.*, Volume 69, pp. 503-515.
- Goldman, M. S., 2009. Memory without Feedback in a Neural Network. *Neuron*, Volume 61, pp. 621-634.
- Gowers, W., 1885. Course of epilepsy. In: G. W, ed. *Epilepsy and other chronic convulsive diseases: their causes, symptoms and treatment*. New York: William Wood, pp. 157-164.
- Graupner, M. & Brunel, N., 2012. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *PNAS*, 109(10), pp. 3991-3996.
- Greenberg, J. A. et al., 2015. Decreases in Theta and Increases in High Frequency Activity Underlie Associative Memory Encoding. *NeuroImage*, Volume 114, pp. 257-263.
- Gyderian, S., Schott, B., Richardson-Klavehn, A. & Duzel, E., 2009. Medial temporal theta state before an event predicts episodic encoding success in humans. *PNAS*, 106(13), pp. 5365-5370.
- Haegens, S. et al., 2011. α -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *PNAS*, 108(48), pp. 19377-19382.
- Hansel, D. & Mato, G., 2003. Asynchronous States and the Emergence of Synchrony in Large Networks of Interacting Excitatory and Inhibitory Neurons. *Neural Computation*, Issue 15, pp. 1-56.
- Hanslmayr, S., Gross, J., Wolfgang, K. & Shapiro, K., 2011b. The role of alpha oscillations in temporal attention. *Brain Research Reviews*, pp. 331-343.
- Hanslmayr, S., Spitzer, B. & Bauml, K.-H., 2009. Brain Oscillations Dissociate between Semantic and Nonsemantic Encoding of Episodic Memories. *Cerebral Cortex*, 19(7), pp. 1631-1640.
- Hanslmayr, S., Staresina, B. P. & Bowman, H., 2016. Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. *Trends in Neuroscience*, 39(1), pp. 16-25.
- Hanslmayr, S. & Staudigl, T., 2014. How brain oscillations form memories - a processing based perspective on oscillatory subsequent memory effects. *Neuroimage*, Volume 85, pp. 648-655.
- Hanslmayr, S., Staudigl, T. & Fellner, M.-C., 2012. Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in Human Neuroscience*.
- Hanslmayr, S. et al., 2013. Prestimulus Oscillatory Phase at 7 Hz Gates Cortical Information Flow and Visual Perception. *Curr Biol. Elsevier Ltd*, Volume 23, pp. 2273-2278.

- Hanslmayr, S., Volberg, G., Wimber, M. & Raabe, K.-H. T., 2011a. The relationship between brain oscillations and BOLD signal during memory formation: a combined EEG-fMRI study. *Journal of Neuroscience*, Volume 31, pp. 15674-15680.
- Harmony, T., 2013. The functional significance of delta oscillations in cognitive processing. *Frontiers in Integrative Neuroscience*, 7(83).
- Hasselmo, M. E., 2005. What is the function of hippocampal theta rhythm? - Linking behavioral data to phasic properties of field potential and unit recording data. *Hippocampus*, 15(7), pp. 936-949.
- Hasselmo, M. E., Bodelon, C. & Wyble, B. P., 2002. A Proposed Function for Hippocampal Theta Rhythm: Separate Phases of Encoding and Retrieval Enhance Reversal of Prior Learning. *Neural Computation*, 14(4), pp. 793-817.
- Hebb, D. O., 1949. *The Organization of Behavior*. New York: Wiley & Sons.
- Henze, D. & Buzsaki, G., 2001. Action potential threshold of hippocampal pyramidal cells in vivo is increased by recent spiking activity. *Neuroscience*, Volume 105, pp. 121-130.
- Herrmann, C. S., 2001. Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137(3-4), pp. 346-353.
- Heusser, A. C., Poeppel, D., Ezzyat, Y. & Davachi, L., 2016. Episodic sequence memory is supported by a theta-gamma phase code. *Nature Neuroscience*, 19(10), pp. 1374-1380.
- Hodgkin, A. L. & Huxley, A. F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4), pp. 500-544.
- Hofstra, W. A. & de Weerd, A. W., 2008. How to assess circadian rhythm in humans: A review of literature. *Epilepsy & Behavior*, Volume 13, pp. 438-444.
- Howard, M., Viskontas, W., Shankar, K. H. & Fried, I., 2012. Ensembles of human MTL neurons “jump back in time” in response to a repeated stimulus. *Hippocampus*, Volume 22, pp. 1833-1847.
- Howard, M. W. & Eichenbaum, H., 2015. Time and space in the hippocampus. *Brain Research*, Volume 1621, pp. 345-354.
- Howard, W. M. & Eichenbaum, H., 2013. The Hippocampus, Time, and Memory Across Scales. *Journal of Experimental Psychology*, 142(4), pp. 1211-1230.
- Huerta, P. T. & Lisman, J. E., 1995. Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 in vitro. *Neuron*, pp. 1053-1063.
- Hughes, S. W. et al., 2004. Synchronized oscillations at alpha and theta frequencies in the lateral geniculate nucleus. *Neuron*, 42(2), pp. 254-68.
- Huxter, J., Burgess, N. & O'Keefe, J., 2003. Independent rate and temporal coding in hippocampal pyramidal cells. *Letters to Nature*, pp. 828-832.
- Ison, M., Quiroga, R. & Fried, I., 2015. Rapid Encoding of New Memories by Individual Neurons in the Human Brain. *Neuron*, Issue 87, pp. 220-230.

- Itskov, V., Curto, C., Pastalkova, E. & Buzsaki, G., 2011. Cell Assembly Sequences Arising from Spike Threshold Adaptation Keep Track of Time in the Hippocampus. *Journal of Neuroscience*, 31(8), pp. 2828-2834.
- Itti, L., Koch, C. & Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), pp. 1254-1259.
- Jarosz, Q., 2009. *Neuron-no labels2.png*. [Art] (Wikipedia).
- Jensen, O., Idiart, M. & Lisman, J., 1996. Physiologically Realistic Formation of Autoassociative Memory in Networks with Theta/Gamma Oscillations: Role of Fast NMDA Channels. *Learning & Memory*, pp. 243-256.
- Jensen, O. & Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, 4(186).
- Jokisch, D. & Jensen, O., 2007. Modulation of Gamma and Alpha Activity during a Working Memory Task Engaging the Dorsal or Ventral Stream. *Journal of Neuroscience*, 27(12), pp. 3244-3251.
- Jones, S., Pinto, D., Kaper, T. & Koppell, N., 2000. Alpha-frequency rhythms desynchronize over long cortical distances: a modeling study. *Computational Neuroscience*, Volume 9, pp. 271-291.
- Kaplan, R. et al., 2012. Movement-Related Theta Rhythm in Humans: Coordinating Self-Directed Hippocampal Learning. *PLOS Biology*.
- Ketz, N., Morkonda, S. G. & O'Reilly, R. C., 2013. Theta coordinated error-driven learning in the hippocampus. *PLoS computational biology*, 9(6).
- Khader, P. H. & Rosler, F., 2011. EEG power changes reflect distinct mechanisms during long-term memory retrieval. *Psychophysiology*, Volume 48, pp. 362-369.
- Khader, P., Jost, K., Ranganath, C. & Rosler, F., 2010. Theta and Alpha oscillations during working-memory maintenance predict successful long-term memory encoding. *Neuroscience Letters*, 468(3), pp. 339-343.
- Klimesch, W., Barbel, S. & Sauseng, P., 2005. The Functional Significance of Theta and Upper Alpha Oscillations. *Experimental Psychology*, 52(2), pp. 99-108.
- Klimesch, W., Sauseng, P. & Hanslmayr, S., 2007. EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, 53(1), pp. 63-68.
- Knight, R. T. & Marica, G., 1995. Escape from linear time: Prefrontal cortex and conscious experience. In: M. S. Gazzaniga, ed. *The Cognitive Neurosciences*. Cambridge, MA, US.: The MIT Press, pp. 1357-1371.
- Kumar, A., Rotter, S. & Aertsen, A., 2008. Conditions for Propagating Synchronous Spiking and Asynchronous Firing Rates in a Cortical Network Model. *The Journal of Neuroscience*, 28(20), pp. 5268-5280.
- Landau, A. & Fries, P., 2012. Attention Samples Stimuli Rhythmically. *Curr Biol. Elsevier Ltd*, Volume 22, pp. 1000-1004.
- Lega, B. C., Jacobs, J. & Kahana, M., 2012. Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*, Volume 22, pp. 748-761.

- Libri, V., Constanti, A., Calaminici, M. & Nistico, G., 1994. A comparison of the muscarinic response and morphological properties of identified cells in the guinea-pig olfactory cortex in vitro. *Neuroscience*, pp. 331-347.
- MacDonald, C. J., Carrow, S., Place, R. & Eichenbaum, H., 2013. Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *Journal of Neuroscience*, Volume 33, pp. 14607-14616.
- MacDonald, C., Lepage, K., Eden, U. & Eichenbaum, H., 2011. Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. *Neuron*, 71(4), pp. 737-749.
- Markram, H., Lubke, J., Frotscher, M. & Sakmann, B., 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, Volume 275, pp. 213-215.
- Markram, H. & Sakmann, B., 1995. Action potentials propagating back into dendrites triggers changes in efficacy. *Soc. Neurosci. Abs.*, Volume 21.
- Markram, H. et al., 2004. Interneurons of the neocortical inhibitory system. *Nature Reviews*, Volume 5, pp. 793-807.
- Mauk, M. D. & Buonomano, D. V., 2004. The Neural Basis of Temporal Processing. *Annual Reviews Neuroscience*, Volume 27, pp. 307-340.
- McClelland, J., McNaughton, B. & O'Reilly, R., 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, Volume 102, pp. 419-457.
- McCloskey, M. & Cohen, N. J., 1989. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. *Elsevier*, Volume 24, pp. 109-165.
- McNaughton, B. L. et al., 2006. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, Volume 7, pp. 663-678.
- Meeuwissen, E. B., Takashima, A., Fernandez, G. & Jensen, O., 2011. Increase in posterior alpha activity during rehearsal predicts successful long-term memory formation of word sequences. *Human Brain Mapping*, pp. 2045-2053.
- Michelmann, S., Bowman, H. & Hanslmayr, S., 2016. The Temporal Signature of Memories: Identification of a General Mechanism for Dynamic Memory Replay in Humans. *PLoS Biology*, 14(8).
- Michelmann, S., Staresina, B., Bowman, H. & Hanslmayr, S., 2018. Speed of time-compressed forward replay flexibly changes in human episodic memory. *BioRxiv*.
- Moser, E., Kropff, E. & Moser, M.-B., 2008. Place Cells, Grid Cells, and the Brain's Spatial Representation System. *Annual Reviews Neuroscience*, Volume 31, pp. 68-89.
- Muthukumaraswamy, S. D., Johnson, B. W. & McNair, N. A., 2004. Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, Volume 19, pp. 195-201.
- Neidermeyer, E. & da Silva, F. L., 2004. *Electoencephalography: Basic principles, clinical applications, and related fields*. s.l.:Lippincot Williams & Wilkins.
- Ng, B., Logothetis, N. & Kayser, C., 2013. EEG Phase Patterns Reflect the Selectivity of Neural Firing. *Cereb Cortex*, Volume 23, pp. 389-398.

- Noh, E., Herzmann, G., Curran, T. & de Sa, V. R., 2014. Using Single-trial EEG to Predict and Analyze Subsequent Memory. *NeuroImage*, Volume 84, pp. 712-723.
- Norman, K., Newman, E., Detre, G. & Polyn, S., 2006. How Inhibitory Oscillations Can Train Neural Networks and Punish Competitors. *Neural Computation*, Volume 18, pp. 1577-1610.
- O'Keefe, J., 1979. A review of hippocampal place cells. *Progress in Neurobiology*, 13(4), pp. 419-439.
- O'Keefe, J. & Reece, M., 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3(3).
- OpenStax College, 2013. *OpenStax CNX*. [Online]
Available at: <http://cnx.org/content/m44427/latest...ol11448/latest>
[Accessed 29 11 2019].
- O'Reilly, R., Bhattacharyya, R., Howard, M. & Ketz, N., 2011. Complementary Learning Systems. *Cognitive Science*, pp. 1-20.
- O'Reilly, R. & Munakata, Y., 2000. *Computational Explorations in Cognitive Neuroscience*. Cambridge, Massachusetts; London, England: The MIT Press.
- Parish, G., Hanslmayr, S. & Bowman, H., 2018. The Sync/deSync Model: How a Synchronized Hippocampus and a Desynchronized Neocortex Code Memories. *Journal of Neuroscience*, 38(14), pp. 3428-3440.
- Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsaki, G., 2008. Internally generated cell assembly sequences in the rat hippocampus. *Science*, Volume 321, pp. 1322-1327.
- Pavlidis, C., Greenstein, Y., Grudman, M. & Winson, J., 1988. Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of θ -rhythm. *Brain Research*, 493(1-2), pp. 383-387.
- Pedreira, C. et al., 2010. Responses of Human Medial Temporal Lobe Neurons Are Modulated by Stimulus Repetition. *Journal of Neurophysiology*, pp. 97-107.
- Petsche, H., Stumpf, C. & Gogolak, G., 1962. The significance of the rabbit's septum as a relay station between the midbrain and the hippocampus. I. The control of hippocampus arousal activity by the septum cells. *Electroencephalography Clinical Neurophysiology*, Volume 14, pp. 202-11.
- Quiroga, R. Q., 2012. Concept cells: the building blocks of declarative memory functions. *Nature Reviews*, Volume 13, pp. 587-597.
- Rabinovich, M. I., Huerta, R. & Varona, P., 2006. Heteroclinic Synchronization: Ultrasubharmonic Locking. *Physical Review Letters*, Volume 96.
- Reisenhuber, M. & Poggio, T., 1999. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, Volume 2, pp. 1019-1025.
- Rotstein, H. G. et al., 2005. Slow and fast inhibition and an H-current interact to create a theta rhythm in a model of CA1 interneuron network. *J Neurophysiology*, 94(2), pp. 1509-18.
- Rumelhart, D. E., McClelland, J. L. & Asanuma, C., 1986. *Parallel Distributed Processing*. s.l.:MIT Press.

- Rutishauser, U., Ross, I., Mamelak, A. & Schuman, E., 2010. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*, pp. 903-907.
- Salari, N. & Rose, M., 2016. Dissociation of the functional relevance of different pre-stimulus oscillatory activity for memory formation. *Neuroimage*, Volume 125, pp. 1013-1021.
- Sauseng, P. et al., 2005a. EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping*, pp. 148-155.
- Sauseng, P., Klimesch, W., Schabas, M. & Doppelmayr, M., 2005b. Fronto-parietal coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, Volume 57, pp. 97-103.
- Schyns, P., Thut, G. & Gross, J., 2011. Cracking the Code of Oscillatory Activity. *PLoS Biology*, Volume 9.
- Sederberg, P. B. et al., 2007. Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex*, Volume 17, pp. 1190-1196.
- Shankar, K. & Howard, M., 2012. A Scale-Invariant Internal Representation of Time. *Neural Computation*, Volume 24, pp. 134-193.
- Shannon, C. E. & Weaver, W., 1949. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Shepherd, G. M., 1994. *Neurobiology*. 3rd ed. Oxford: Oxford University Press.
- Song, S., Miller, K. D. & Abbot, L. F., 2000. Competitive Hebbian learning through spike-timing-dependent-plasticity. *Nature Neuroscience*.
- Spehlmann, R., 1965. The average electrical responses to diffuse and to patterned light in the human. *Electroencephology and Clinical Neurophysiology*, Volume 19, pp. 560-569.
- Squire, L. R., 1992. Memory and the Hippocampus: A Synthesis From Findings With Rats, Monkeys, and Humans. *Psychological Review*, 99(2), pp. 195-231.
- Squire, L. R., Stark, C. E. L. & Clark, R. E., 2004. The Medial Temporal Lobe. *Annual Reviews Neuroscience*, Volume 27, pp. 279-306.
- Staresina, B., Henson, R. a., Kriegeskorte, N. & Alink, A., 2012. Episodic reinstatement in the medial temporal lobe. *Journal of Neuroscience*, Volume 32, pp. 18150-18156.
- Staresina, B. P. et al., 2016. Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *eLife*, 10(5).
- Staudigl, T. & Hanslmayr, S., 2013. Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. *Current Biology*, 23(12), pp. 1101-1106.
- Staudigl, T., Vollmar, C., Noachtar, S. & Hanslmayr, S., 2015. Temporal-Pattern Similarity Analysis Reveals the Beneficial and Detrimental Effects of Context Reinstatement on Human Memory. *Journal of Neuroscience*, Volume 35, pp. 5373-5384.
- Suffczynski, P., Kalitzin, S., Pfurtscheller, G. & Lopes da Silva, F., 2001. Computational model of thalamo-cortical networks: dynamic control of alpha rhythms in relation to focal attention. *International Journal of Psychophysiology*, Volume 43, pp. 25-40.

- Swan, G. & Wyble, B., 2014. The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention Perception & Psychophysics*, Volume 76, pp. 2136-2157.
- Tsao, A. et al., 2018. Integrating time from experience in the lateral entorhinal cortex. *Nature*, Volume 561, pp. 57-62.
- Tulving, E. & Donaldson, W., 1972. *Organisation of Memory*. Oxford: Academic Press Inc.
- van Kerkoerle, T. et al., 2014. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *PNAS*, 111(40), pp. 14332-41.
- VanRullen, R., Carlson, T. & Cavanagh, P., 2007. The blinking spotlight of attention. *Proc Natl Acad Sci*, Volume 104, pp. 19204-19209.
- Vicente, R. et al., 2008. Dynamical relaying can yield zero time lag neuronal synchrony despite long conduction delays. *Proc. Natl Acad. Sci. USA*, Volume 105, pp. 17157-17162.
- Volgushev, M. et al., 2016. Partial Breakdown of Input Specificity of STDP at Individual Synapses Promotes New Learning. *Journal of Neuroscience*, 36(34), pp. 8842-8855.
- Waldhauser, G. T., Braun, V. & Hanslmayr, S., 2016. Episodic memory retrieval functionally relies on very rapid reactivation of sensory information. *Journal of Neuroscience*, 36(1), pp. 251-260.
- Waldhauser, G. T., Johansson, M. & Hanslmayr, S., 2012. Alpha/Beta Oscillations Indicate Inhibition of Interfering Visual Memories. *Journal of Neuroscience*, 32(6), pp. 1953-1961.
- Walker, M., 2018. *Why We Sleep: The New Science of Sleep and Dreams*. s.l.:Penguin.
- Wang, X.-J., 2010. Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiol. Review*, Volume 90, pp. 1195-1268.
- Wearden, J. H. & Lejeune, H., 2007. Scalar properties in human timing: Conformity and violations. *The Quarterly Journal of Experimental Psychology*, pp. 1-19.
- Willms, A. R., Baro, D. J., Harris-Warrick, R. M. & Guckenheimer, J., 1999. An Improved Parameter Estimation Method for Hodgkin-Huxley Models. *Journal of Computational Neuroscience*, Volume 6, pp. 145-168.
- Wimber, M. et al., 2015. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience*, Volume 18, pp. 582-589.
- Winocur, G., Moscovitch, M. & Bontempi, B., 2010. Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal–neocortical interactions. *Neuro-psychologia*, Volume 48, pp. 2339-2356.
- Wood, E. R., Dudchenko, P., Robitsek, R. J. & Eichenbaum, H., 2000. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*, Volume 27, pp. 623-633.
- Wyble, B., Bowman, H. & Nieuwenstein, M., 2009. The attentional blink provides episodic distinctiveness; sparing at a cost. *J Exp Psychol Hum Percept Perform*, 35(3), pp. 787-807.

FREQUENCY & LOCAL-FIELD-POTENTIAL

Oscillations occur when groups of neurons are rhythmically activated by some modulatory source (Klimesch, et al., 2007). We can record the electrical impulses of these mass activations through electro-encephalogram (EEG) and magneto-encephalogram (MEG) recordings, which show us oscillatory power changes in electrical signals of particular frequencies (Fell & Axmacher, 2011). These types of recordings are termed local-field-potentials (LFPs), as we have to keep in mind that we are recording an electrical field from a local population and not specific neurons. Single-cell-recording does allow for the latter type of measurement, however, where one can determine how a single neuron behaves in relation to the local group and what effect that has on its ability to communicate and learn (Buzsaki, 2002; Fell & Axmacher, 2011). This section overviews several oscillatory functions that have been observed as brain regions seek to communicate and maximise learning (Fell & Axmacher, 2011), using some computational simulations to further explore theories of oscillations and communication (Dayan & Abbot, 2001; Wang, 2010).

To explain the neurophysiological underpinnings of oscillatory power changes, it is important to consider the model of EEG/MEG signal generation, whereby it is assumed that the local field potential that is being recorded reflects the summation of thousands of postsynaptic potentials, and thus can capture when observed populations of neurons become active. By concurrently observing the activity of single cells, we can see that neurons tend to activate phase-locked to the local field potential, thus showing that the highly synchronised activity of many neurons contributes to a larger power in a specific frequency band (Fell & Axmacher, 2011).

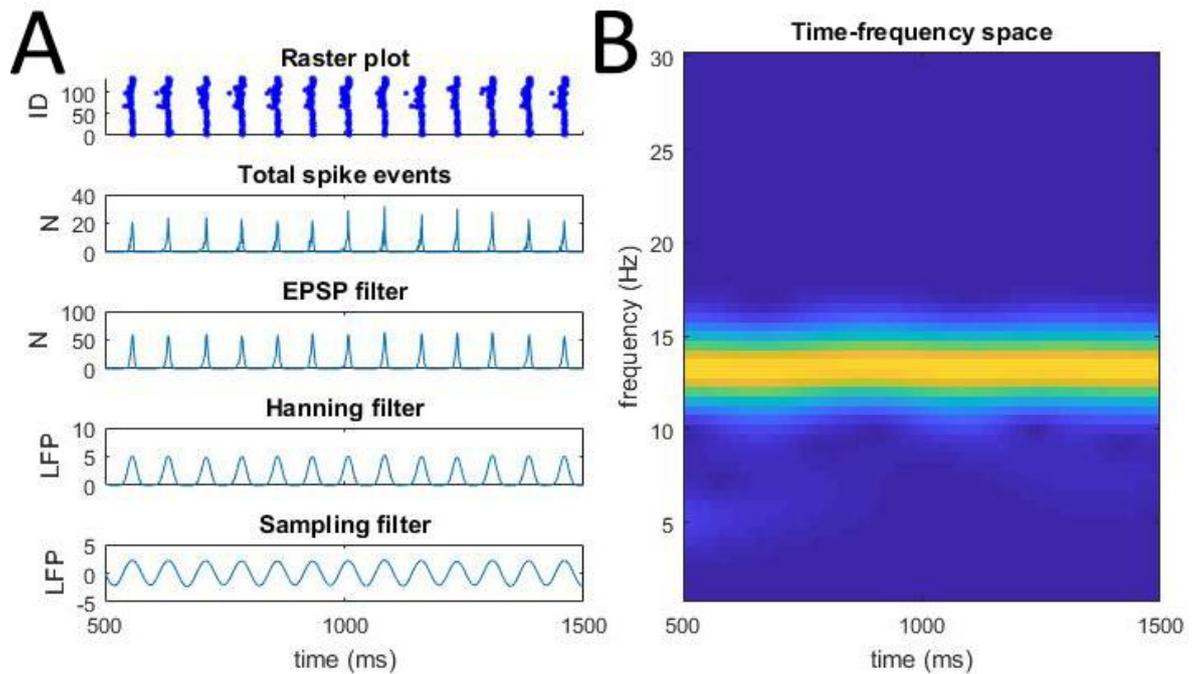


Figure 1 – **A** Conversion of spike data to local-field-potential (LFP). Signal generated by summing all spike events and convolving through an EPSP filter ($\tau_s = 1.5\text{ms}$), then filtering further with a Hanning window (window = 30ms) followed by a sampling filter with a frequency of 6-14Hz. **B** LFP signal further processed using a Gabor filter ($\gamma = 0.5$, lower bound = 1Hz, upper bound = 30Hz), where the absolute values are plotted in time-frequency space.

When calculating LFP from simulated neuronal data, the activity of a group of neurons can be recorded by first aggregating spikes through time. An EPSP filter can then be used over each spike event, followed by a Hanning filter ($\sim 30\text{ms}$ window) over the whole signal, to more accurately match the aggregated voltage data that is recorded through scalp EEG and MEG methods. To convolve the simulated signal and render it comparable to recorded EEG/MEG signals, it is important to further filter the signal with a sampling frequency in the desired range, for example a 2-6Hz filter for slow waves or an 8-12Hz filter for medium waves. The LFP signal can then be analysed in time-frequency space using a Gabor filter with an upper and lower bound dependent on the desirable frequency once again ($\gamma = 0.5$ for $< 30\text{Hz}$ or $\gamma = \pi/2$ for $> 30\text{Hz}$). The absolute values are then taken and plotted in time-frequency space.

Figure 1 shows these analytical methods applied to voltage traces from a group of simulated neurons like those from Figure 23, where neurons are modulated to fire at a specific frequency. In Figure 1A, the stages of each applied filter ends with a waveform effect that is similar to observed traces from actual neurological data. When filtering this signal further and plotting the absolute values in time-frequency space (Figure 1B), one can see a thick band in the z-axis at $\sim 13\text{Hz}$ frequency on the y-axis, indicating that neuronal spiking is occurring strongly in this frequency. The size of the z-values in time-frequency space indicates the amplitude of the oscillation, which varies with the synchronicity and relative number of spike events occurring in oscillatory up-states (Klimesch, et al., 2007; Fell & Axmacher, 2011; Hanslmayr, et al., 2012). Considering simulated neurological data in the modelling work throughout this thesis, this process of filters will be used to analyse spike data in time-frequency space.

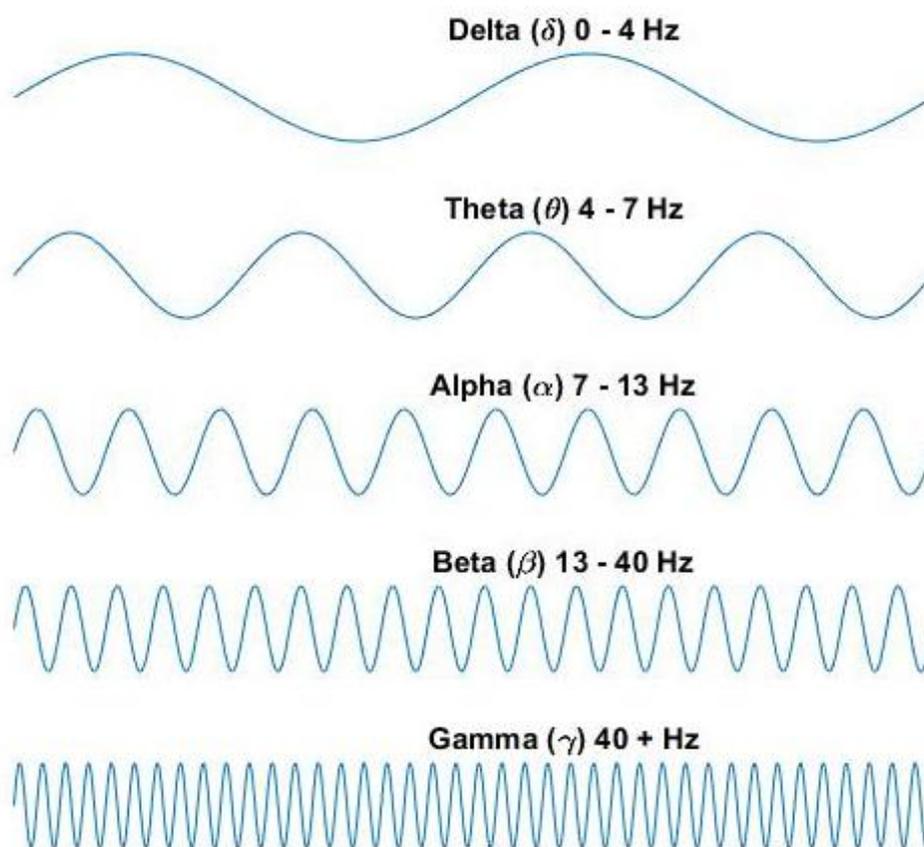
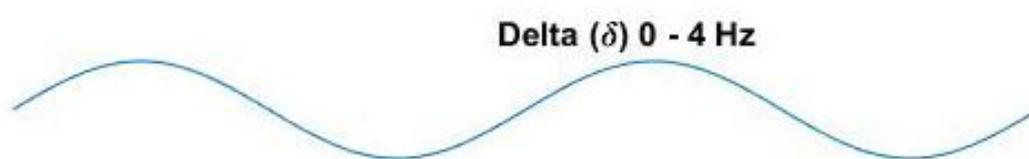
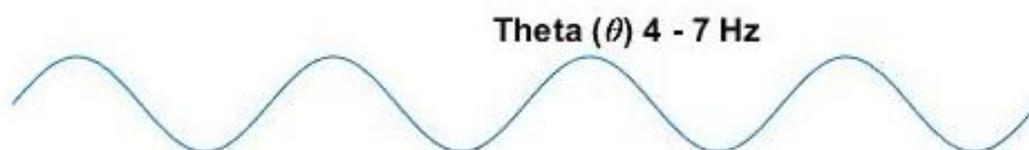


Figure 2 – LFP signals for several frequency bands of neural oscillations observed in the brain.

The frequency of an oscillation is determined by how many up-states occur during a second-long interval. Frequency ranges are found over 100Hz in the brain, occurring within a range from a few milliseconds to several seconds, where several frequencies can concurrently occur within the same brain region (Fell & Axmacher, 2011). Figure 2 shows the main frequencies that have been observed in the brain, which have been found to originate from distinct brain regions and are observed during specific cognitive operations (Basar, et al., 2001; Buzsaki, 2002; Fell & Axmacher, 2011; Klimesch, et al., 2007). This indicates which brain areas are key to observed functions as it is believed that they entrain other regions to their frequency to facilitate communication.



Delta oscillations are thought to have a thalamic or cortical origin and are present in frontal cortical areas during mental calculations, motor or semantic tasks when the subject is internally concentrating (Harmony, 2013). They are thought to play a role in the suppression of irrelevant tasks, essential for attentional processes to dictate the flow of information in working memory maintenance. They are also present during non-rapid eye movement (NREM) stage 3-4 deep sleep, also known as slow-wave-sleep (SWS) (Diekelmann & Born, 2010; Walker, 2018). Here, the brain is said to be in a state similar to that of deep thought, as opposed to the more hallucinogenic rapid eye movement (REM) sleep or the preparatory stages of NREM, where theta and higher frequencies are more prevalent (Diekelmann & Born, 2010).



Theta is one of the most studied frequencies in mammalian brains due to its perceived role in many fundamental neural operations (Backus, et al., 2016; Burke, et al., 2014; Buzsaki, 2002; Fell &

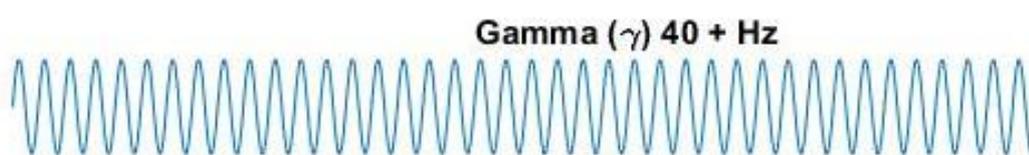
Axmacher, 2011; Lega, et al., 2012). Much research has attempted to understand how theta segregates neuronal assemblies and assigns specific computational tasks to them (Hasselmo, 2005), as well as the relationship between local-field-potential theta and the spiking of individual neurons (Buzsaki, 2002; Huerta & Lisman, 1995; Rutishauser, et al., 2010). Theta waves are present in the hippocampus during REM sleep and specific types of movement such as voluntary, preparatory, orienting or exploratory, yet absent in immobile animals (Fell & Axmacher, 2011).

The rhythm generation of the theta frequency is thought to be modulated by subcortical neurotransmitters (Buzsaki, 2002), either through rhythmic activation of recurrently connected neurons or a permissive pacemaker function. The current generation that determines theta amplitude is thought to be amplified by feedback from several cortical and subcortical regions (Buzsaki, 2002). Many of such regions exhibit phase-locked firing to hippocampal theta oscillations, indicating that the hippocampus is the main driver of this frequency.

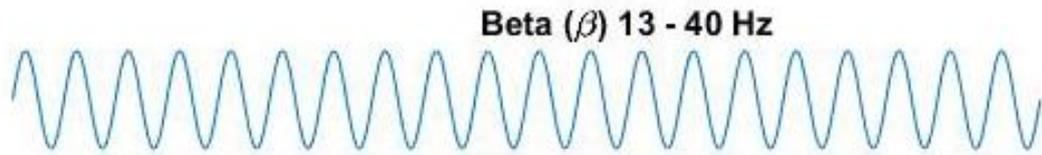
Hippocampal theta has many influences from several sources (Buzsaki, 2002), though GABA_A and NMDA receptors (inhibitory and excitatory neurotransmitters, respectively) appear to be critical in maintaining both rhythm and current generation, as shown by many drug blocker studies. Different hippocampal regions generate their own theta fields, whereby a travelling wave can be observed through regions firing at phases up to 180° apart. Recurrent synapses within the CA3 region have also been found to act as rhythm generators, which drive the CA1 region to fire in phase through the Schaffer collaterals, and it can even act as the sole intrahippocampal oscillator when other sources are nullified (Buzsaki, 2002). Due to this, intrahippocampal theta can dynamically adjust in phase and frequency with respect to extrahippocampal theta sources, thus allowing phase to have a profound effect on the timing of action potentials (Hasselmo, et al., 2002; Huerta & Lisman, 1995; Rutishauser, et al., 2010).

Inhibitory interneurons have also been found to contribute to hippocampal theta (see Figure 28; Buzsaki, 2002), where separate classes of interneuron are thought to contribute to the creation of

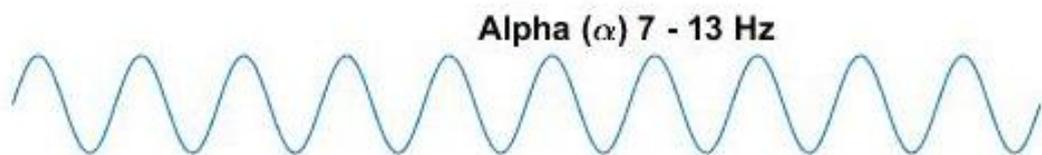
several gamma frequencies as well as helping to maintain theta phase. Interneurons also enable a competitive winner-take-all architecture (Hasselmo, 2005); whereby weakly activated neurons cannot overcome inhibition such that a clear line of activation can be maintained, exemplified in one of the following modelling studies (Norman, et al., 2006). Finally, the internal mechanisms of neurons themselves can contribute to the local-field-potential of recorded hippocampal theta, whereby a persistent Na^+ current can cause voltage-dependent oscillations within the membrane potential of several classes of neurons (Buzsaki, 2002). Interestingly, this membrane resonance might enable hippocampal neurons to be actively selective to input rather than act solely as passive signal integrators, the subject of several modelling studies (Wang, 2010).



Gamma frequencies are typically associated with periods of activity in most regions of the brain (Jensen & Mazaheri, 2010). It is now known that some classes of inhibitory interneurons can generate gamma rhythms in local cell assemblies (Buzsaki, 2002), though it was not initially clear if gamma in the brain represents anything more functionally relevant than simply tonic activation of cell assemblies (Basar, et al., 2001). As stated earlier, high-frequency ranges give a very precise window for activation, making gamma very important in the integration of communication on specific target cells (Fell & Axmacher, 2011). Indeed, one can observe phase synchrony in gamma oscillations between disparate brain regions as they communicate to represent various elements of a stimuli (Basar, et al., 2001). Gamma frequencies have also been found as an important predictor for successful memory formation in humans (Sederberg, et al., 2007), where a theta/gamma phase coupling is thought to be important for working (Jensen, et al., 1996; O'Keefe & Reece, 1993) and episodic memory (Burke, et al., 2014; Heusser, et al., 2016).



Beta oscillations occur exclusively during semantic memory encoding in the left inferior pre-frontal cortex (Hanslmayr, et al., 2009; Meeuwissen, et al., 2011). Alpha, however, is more widely observed throughout the neo-cortex and is thought to function as a top-down attentional control mechanism (Klimesch, et al., 2007). The way in which both frequencies behave is very similar (Hanslmayr, et al., 2016; Hanslmayr & Staudigl, 2014), perhaps due to their shared cortical architecture. Therefore, in the context of the limited scope of this thesis we consider them as analogous signals, where references to alpha throughout this work may also be applied to beta frequencies.



Alpha waves were the first observable frequencies in the brain upon the invention of the electroencephalogram, due to their dominance over the outmost cortical layers of the brain (Berger, 1929). Since the beginning, an increase in observable alpha was associated with relaxation and attentional control (Adrian & Matthews, 1934; Berger, 1929), indicating that alpha represents an internal equilibrium state. Though some studies have since suggested alpha is an epiphenomenon (Herrmann, 2001; Gebber, et al., 1999), there is overwhelming evidence of the causation between alpha and attentional (Haegens, et al., 2011; Hanslmayr, et al., 2011b; Klimesch, et al., 2007) and memory (Doppelmayr, et al., 2005; Fell, et al., 2011; Hanslmayr, et al., 2009; Khader, et al., 2010; Sauseng, et al., 2005) processes, whereby any transient changes in alpha activity inevitably leads to a predictable change in cognitive performance. These findings suggest alpha mediates a top-down processing (Klimesch, et al., 2007), where synchronisation between regions allows neurons in distributed networks to coherently influence common target cells (Fell & Axmacher, 2011) and

inhibition allows for the selective activation of specific subsets and pathways (Hanslmayr, et al., 2012).

Cortical alpha is modulated by thalamo-cortico-thalamic recurrent loops, where bursts of thalamic cells have been found to induce larger EPSPs in cortical units and thus a larger potential for attentional focus (Klimesch, et al., 2007), and dysfunctions in thalamic-cortical circuits is associated with a lack to inhibit irrelevant information (Knight & Marica, 1995). Whilst previous work indicated that alpha frequencies were primarily driven by a thalamic pacemaker (Andersen & Andersson, 1968), it has since been found that this is not the case (Klimesch, et al., 2007). It may also be the case that there are many distributed alpha systems in the brain, where cortical columns generate their own alpha oscillations through local cortical circuits (Jones, et al., 2000) and phase synchrony in cortical alpha is enabled through thalamic modulation (Vicente, et al., 2008).

PHASE SYNCHRONY

There is growing evidence that neural oscillations play a key role in the organisation of information flow within the brain (Basar, et al., 2001; Buzsaki, 2002; Fell & Axmacher, 2011; Klimesch, et al., 2007; Hanslmayr, et al., 2012). This section details the basic functionality of oscillations and how synchronisations in phase between disparate regions can encourage communication, inhibit irrelevant information or even enable associative learning between cell assemblies (Fell & Axmacher, 2011). Neurons discharge within defined periods of excitability that are set by the oscillatory phase, though not every neuron discharges at every oscillatory cycle. Oscillatory phase can be described as the current phase in comparison to a reference oscillation.

Phase synchronisation refers to the aligning of phase between the oscillations within distinct brain regions (Figure 3), a process that has been found to support neural communication, plasticity and many cognitive processes (Fell & Axmacher, 2011). Phase synchrony enables separate regions to concurrently represent specific attributes of any given stimulus, that together form a coherent representation (Fries, 2005). Indeed, targeted neurons may only be triggered by the integration of

multiple inputs (Figure 4, coincidence detection), which may only be feasible through fast integration in the very precise time windows of high-frequency ranges (Fell & Axmacher, 2011). Phase synchronised regions are thus pre-prepared for optimal communication, as both regions are entrained to the same up-states with no inhibition to overcome (Figure 4, neural communication). Otherwise they are blocked from communicating, as information sent from one region cannot overcome the inhibitory down-state of the other.

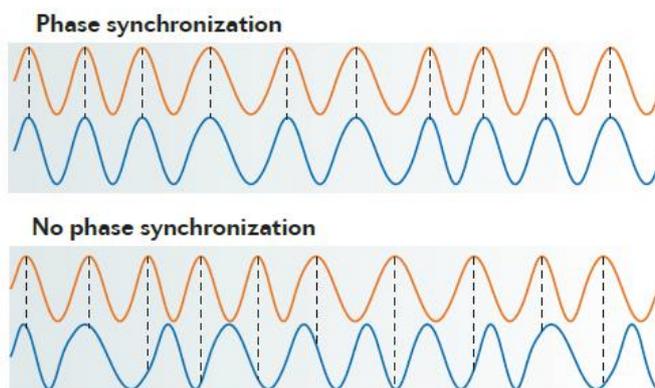


Figure 3 – Phase synchrony between oscillations, where synchrony does or does not occur due to an alignment in phase between disparate brain regions. Adapted from Fell & Axmacher, 2011.

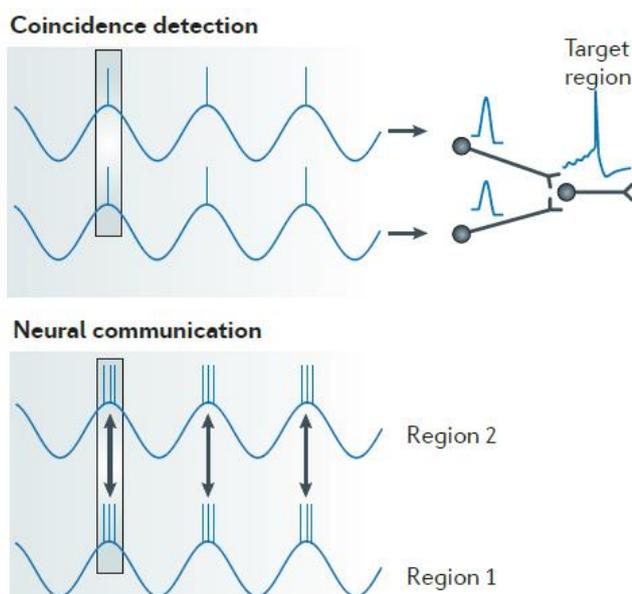


Figure 4 – Synchrony enables neural communication by allowing target neurons to integrate concurrent inputs from multiple regions. Figure adapted from Fell & Axmacher, 2011.

If neural regions synchronise in phase one would expect there to be a lag in phase, as communication from one brain region to another will take some amount of time (Fell & Axmacher, 2011). This would also tell us which region was leading, helping us to identify the source of several oscillatory frequencies. This is indeed the case in some types of neural synchrony, for example theta phase synchrony between the hippocampus and medial prefrontal cortex, where the hippocampus leads (Buzsaki, 2002). However, there are many other studies that show zero phase lag between regions, despite there being a relatively large neural distance between them (Fell & Axmacher, 2011). This may be due to delayed excitatory-inhibitory interactions in local populations, or if two regions receive a common input (Vicente, et al., 2008).

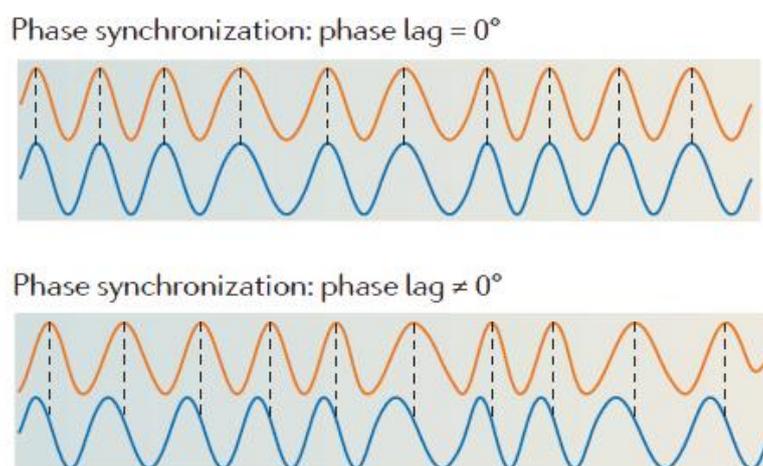


Figure 5 – Phase synchronisation with different lags, i.e. difference in time between peaks. Figure adapted from Fell & Axmacher, 2011.

The simulation in Figure 6 shows how several individual assemblies with no inter-connectivity oscillate intrinsically at various phases (Figure 6A). However, when each individual assembly connects bi-directionally to an intermediate assembly (Figure 6B), there is a zero-lag alignment in phase between the individual assemblies. This might be how the thalamus acts as a relay for cortico-cortico phase synchronisation via thalamo-cortical loops (Vicente, et al., 2008), a popular theory for phase synchrony between independently generated cortical columnar alpha (Jones, et al., 2000).

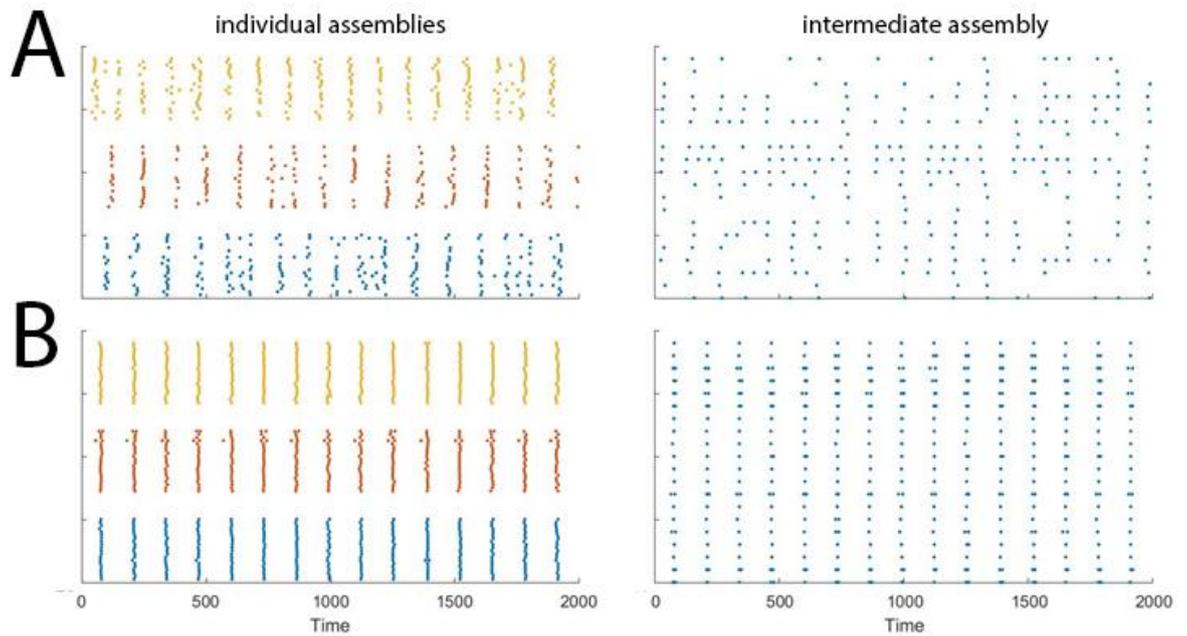


Figure 6 – **A** Several excitatory cell assemblies with underlying inhibitory nodes, distinguished by colour (left), and an intermediate assembly (right), with no connections between the assemblies. **B** The same individual assemblies except with bi-directional excitatory connections between each individual assembly (left) and the intermediate assembly (right). No inter-connections between any of the individual assemblies on the left in either case. Simulation made using Hodgkin & Huxley equations of excitatory and inhibitory populations, recreating findings from Vicente, et al., 2008.

When brain regions communicate with one another through phase synchrony, oscillations at each respective region can become coupled (Fell & Axmacher, 2011). Phase-amplitude coupling (Figure 7; phase-amplitude coupling) occurs when the phase of a low-frequency oscillation mediates the amplitude of a high-frequency oscillation, observed in the rat hippocampus, monkey auditory cortex, human EEG and iEEG (Buzsaki, 2002). Episodic memory has also been found to rely on theta/gamma phase-amplitude coupling (Sederberg, et al., 2007; Burke, et al., 2014).

Phase-phase coupling can also occur (Figure 7; phase-phase coupling), whereby the phase of a higher frequency oscillation is locked to the phase of a lower frequency oscillation (Fell & Axmacher, 2011), such that a ratio of $n:m$ cycles occurs (i.e. 1:8 for 5:40Hz frequencies). This type of coupling

has been found to occur in all aspects of human working memory tasks; between alpha and theta in working memory processing, between theta and gamma during working memory maintenance, and between alpha and gamma during manipulation of sustained information (Fell & Axmacher, 2011; Sauseng, et al., 2005b). This type of coupling may be useful for encoding non-interfering representations of multiple items in working memory due to the higher temporal precision of phase-phase coupling as opposed to phase-amplitude coupling.

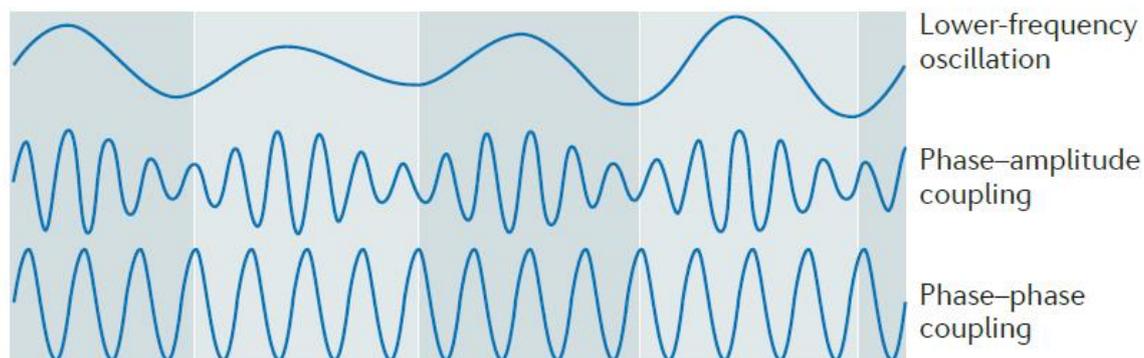


Figure 7 – Phase-amplitude coupling and phase-phase coupling of a higher frequency oscillation to a lower frequency oscillation. Figure adapted from Fell & Axmacher, 2011.

The phase of an on-going oscillatory cycle serves as a frame of reference for both internal and external events. In cell assemblies where oscillations are driven by an excitation-inhibition balance, events can directly influence the phase by resetting it, thereby supplying a new reference point specific to the event circumstances and allowing further events to be decoded more consistently (Canavier, 2015). In this fashion, periodic events may control the frequency and phase of a neural oscillator, maximise encoding and decoding by controlling timing windows for activation. This is referred to as phase-locking, and has even been found to predict successful memory encoding in humans (Rutishauser, et al., 2010).

Phase-resetting is also an important mechanism for frequency coupling, allowing mutually coupled oscillators to coordinate their frequencies and phases (Canavier, 2015; Fell & Axmacher, 2011). Phase-resetting is not always immediate but can often be due to the accumulation of mini resets

that advance or delay the phase of the oscillator, as seen in Figure X8. Here, an event (dotted vertical line) causes the first signal (pink) to delay its phase and peak later than it usually would, whilst causing the second signal (green) to advance its phase as neurons are driven to fire earlier than usual. Phase-resetting can be measured with respect either to a single oscillating neuron, or for a whole population with a network oscillator (Canavier, 2015). The phase synchrony occurring in the simulation of Figure 6 shows how a whole neuronal group can reset in phase relative to an intermediary population (Vicente, et al., 2008). Phase precession might occur if a single unit advances in phase in relation to the network oscillator, but the network oscillator remains unchanged, a phenomenon observed in theta/gamma phase-amplitude coupling during working memory processes (Jensen, et al., 1996).

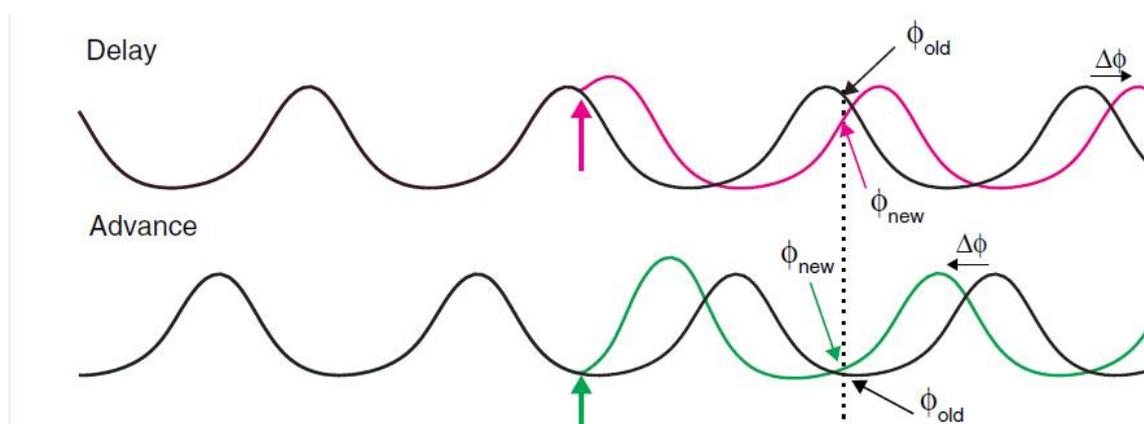


Figure 8 – Phase-resetting of coupled oscillators relative to an event (dotted line), where phase is referred to as ϕ and change in phase as $\Delta\phi$. Figure adapted from Canavier, 2015.

PHASE SYNCHRONY AND MEMORY

Neural communication and plasticity reinforce one another (Figure 9), whereby communicating regions share activation windows through the phase synchronisation of their respective oscillations (Fell & Axmacher, 2011; Fries, 2005). This enables STDP, as neurons are conditioned to spike close in time (Hebb, 1949). Phase synchrony occurs in the theta and gamma ranges within the hippocampus and neocortex of rodents, humans, cats and monkeys, and promotes the induction of LTP, the strengthening of synapses, between and within these regions, where many studies have

found that successfully remembered items are predicted by high phase synchrony in these frequencies (Burke, et al., 2014; Heusser, et al., 2016; Fell & Axmacher, 2011).

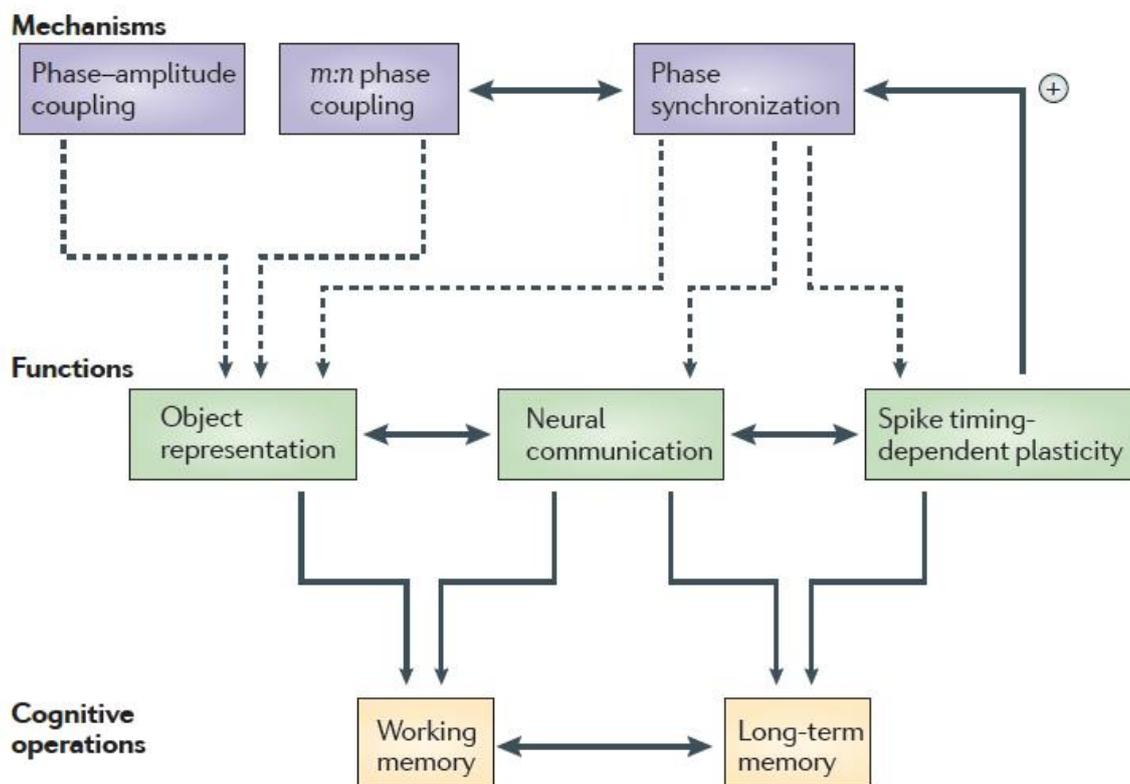


Figure 9 – How functional oscillations facilitate cognitive operations. Phase coupling promotes object representation, necessary for working memory. Phase synchrony induces neural communication, essential for working, long-term & working/long-term memory interactions, as well as inducing STDP. Dotted lines - mechanical support for functions; thin black lines - functional contribution to memory; thick bi-directional lines - interactions between mechanisms, functions, or operations. Figure adapted from Fell & Axmacher, 2011.

Functionally, gamma is optimally suited for enabling STDP due to its short and precise activation window that works on a similar time-scale (Song, et al., 2000). STDP then promotes further phase synchrony due to stronger connectivity between regions (Figure 9). Theta phase synchrony has a broader window of excitation, meaning it is unlikely to promote the precise relative timing of spikes that is necessary for STDP. Instead, theta mediates STDP through its ‘up/down’ states (Hasselmo, et al., 2002; Huerta & Lisman, 1995; Pavlidis et al., 1988), such that only inputs that are active at

the appropriate theta phase are encoded into long-term memory. Theta synchrony also reaches further across the brain than gamma synchrony, suggesting it represents a top-down control function that mediates neural plasticity (Fell & Axmacher, 2011).

During working memory paradigms, the complimentary mechanisms of theta/gamma phase-amplitude and m:n phase coupling are vital to represent non-interfering representations of overlapping patterns (Figure 9; Fell & Axmacher, 2011), necessary for maintaining coherent objects in working memory. During this time, theta phase synchrony is also present between the pre-frontal cortex and the temporal lobe during the encoding and retrieval of items, which persists during maintenance intervals (i.e. sustained representations with no input); between parietal and pre-frontal cortices during complex manipulations; and between frontal and temporal-parietal regions with increasing memory load (i.e. number of items to be stored). All of this indicates that theta phase synchrony organises long-range communication to recruit executive control functions and to sustain active representations in working memory paradigms (Fell & Axmacher, 2011).

Long-term and working memory interact during operations that depend on previous knowledge of the environment, typically facilitated by the hippocampus (Buzsaki, 2002; Fell & Axmacher, 2011). Rats making just such decisions in a T-maze have shown theta phase synchrony between pre-frontal cortex and hippocampal regions, and gamma phase synchrony within hippocampal CA1/CA2 regions. This suggests that theta and gamma phase synchrony within the hippocampus supports the encoding of information from short to long term memory, as well as the retrieval of information from long to short term memory (Fell & Axmacher, 2011; Sederberg, et al., 2007).