
THE ROLE OF SYNAPTIC TAGGING AND CAPTURE FOR MEMORY DYNAMICS IN SPIKING NEURAL NETWORKS

Dissertation

for the award of the degree

“Doctor rerum naturalium”

of the Georg-August-Universität Göttingen

within the doctoral program

International Max Planck Research School for

Physics of Biological and Complex Systems

of the Georg-August University School of Science (GAUSS)

submitted by

Jannik Luboeinski

from Wetzlar

Göttingen, 2021

Thesis Advisory Committee:

Dr. Christian Tetzlaff

III. Institute of Physics – Biophysics, University of Göttingen

Prof. Dr. Stefan Klumpp

Institute for the Dynamics of Complex Systems, University of Göttingen

Prof. Dr. Silvio Rizzoli

Department of Neuro- and Sensory Physiology, University Medical Center Göttingen

Members of the examination board:

First referee: Dr. Christian Tetzlaff

III. Institute of Physics – Biophysics, University of Göttingen

Second referee: Prof. Dr. Stefan Klumpp

Institute for the Dynamics of Complex Systems, University of Göttingen

Further members of the examination board:

Prof. Dr. Silvio Rizzoli

Department of Neuro- and Sensory Physiology, University Medical Center Göttingen

Prof. Dr. Alexander Gail

Sensorimotor Research Group, German Primate Center, Göttingen

Dr. Viola Priesemann

Neural Systems Theory, Max Planck Institute for Dynamics and Self-Organization, Göttingen

Prof. Dr. Dr. Oliver Schlüter

Department of Psychiatry and Psychotherapy, University Medical Center Göttingen

Date of the oral examination:

13th July 2021

Abstract

Memory serves to process and store information about experiences such that this information can be used in future situations. The transfer from transient storage into long-term memory, which retains information for hours, days, and even years, is called consolidation. In brains, information is primarily stored via alteration of synapses, so-called synaptic plasticity. While these changes are at first in a transient early phase, they can be transferred to a late phase, meaning that they become stabilized over the course of several hours. This stabilization has been explained by so-called synaptic tagging and capture (STC) mechanisms. To store and recall memory representations, emergent dynamics arise from the synaptic structure of recurrent networks of neurons. This happens through so-called cell assemblies, which feature particularly strong synapses. It has been proposed that the stabilization of such cell assemblies by STC corresponds to so-called synaptic consolidation, which is observed in humans and other animals in the first hours after acquiring a new memory. The exact connection between the physiological mechanisms of STC and memory consolidation remains, however, unclear. It is equally unknown which influence STC mechanisms exert on further cognitive functions that guide behavior. On timescales of minutes to hours (that means, the timescales of STC) such functions include memory improvement, modification of memories, interference and enhancement of similar memories, and transient priming of certain memories. Thus, diverse memory dynamics may be linked to STC, which can be investigated by employing theoretical methods based on experimental data from the neuronal and the behavioral level.

In this thesis, we present a theoretical model of STC-based memory consolidation in recurrent networks of spiking neurons, which are particularly suited to reproduce biologically realistic dynamics. Furthermore, we combine the STC mechanisms with calcium dynamics, which have been found to guide the major processes of early-phase synaptic plasticity *in vivo*. In three included research articles as well as additional sections, we develop this model and investigate how it can account for a variety of behavioral effects. We find that the model enables the robust implementation of the cognitive memory functions mentioned above. The main steps to this are: 1. demonstrating the formation, consolidation, and improvement of memories represented by cell assemblies, 2. showing that neuromodulator-dependent STC can retroactively control whether information is stored in a temporal or rate-based neural code, and 3. examining interaction of multiple cell assemblies with transient and attractor dynamics in different organizational paradigms.

In summary, we demonstrate several ways by which STC controls the late-phase synaptic structure of cell assemblies. Linking these structures to functional dynamics, we show that our STC-based model implements functionality that can be related to long-term memory. Thereby, we provide a basis for the mechanistic explanation of various neuropsychological effects.

Contents

Abstract	iii
Contents	iv
List of abbreviations	vi
1 Introduction	1
1.1 Motivation and aims	1
1.2 Structure of this thesis	3
1.3 Neural networks in the brain	5
1.3.1 Neurons and coding	6
1.3.2 Network structures and their activity	13
1.3.3 Synaptic transmission and integration	15
1.3.4 Plasticity	18
1.3.5 Cell assemblies	29
1.3.6 Neuromodulators and third factor	31
1.4 Neuropsychological view of memory	34
1.4.1 Learning	35
1.4.2 Important brain regions	37
1.4.3 Declarative memory	38
1.4.4 Long-term memory and consolidation	41
1.4.5 Short-term memory	43
1.4.6 Working memory	44
1.4.7 Recall	45
1.4.8 Phenomena important to this thesis	46
2 Consolidation and improvement of a memory representation	53
2.1 Chapter summary	53
2.2 Manuscript: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks	54
Abstract	54
Introduction	55
Results	57
Discussion	74
Methods	78
Additional information	84

2.3	Supplementary Information: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks	87
2.4	Development of the model and regimes of plasticity	93
3	Neuromodulator-dependent consolidation of different types of memory	95
3.1	Chapter summary	95
3.2	Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks	96
	Abstract	96
	Introduction	96
	Results	98
	Discussion	103
	Methods	106
3.3	Supplementary Information: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks	114
4	Interaction of multiple long-term memory representations	117
4.1	Chapter summary	117
4.2	Manuscript: Organization and priming of long-term memory representations with two-phase plasticity	119
	Abstract	119
	Introduction	120
	Results	122
	Discussion	130
	Methods	132
	Additional information	139
4.3	Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity	140
4.4	Activation and transitions of attractor memory representations	148
	Introduction	148
	Results	150
	Discussion	153
	Methods	156
5	Discussion	157
5.1	Model of synaptic memory consolidation	157
5.2	Cognitive phenomena described by the model	160
5.3	Outlook	161
5.4	Final remarks	164
	Bibliography	167
	Acknowledgments	191
	Curriculum Vitae	193

List of abbreviations

AMPA: α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANN: attractor neural network
ATP: adenosine triphosphate
BCM: Bienenstock-Cooper-Munro
BDNF: brain-derived neurotrophic factor
BG: background
CA: cornu ammonis
CaMKII: Ca^{2+} /calmodulin-dependent protein kinase II
cAMP: cyclic adenosine monophosphate
CREB: cAMP response element-binding protein
CS: conditioned stimulus
dep.: dependent
DNA: deoxyribonucleic acid
E-LTD: early-phase long-term depression
E-LTP: early-phase long-term potentiation
EPSP: excitatory postsynaptic potential
ERK: extra-cellular signal-regulated kinase
fEPSP: local-field excitatory postsynaptic potential
fMRI: functional magnetic resonance imaging
G-protein: guanosine triphosphate-binding protein
GABA: γ -aminobutyric acid
IPSP: inhibitory postsynaptic potential
L-LTD: late-phase long-term depression
L-LTP: late-phase long-term potentiation
LC: locus coeruleus
LIF: leaky integrate-and-fire
LTD: long-term depression
LTP: long-term potentiation

MAT2: multi-timescale adaptive threshold(2)
MCPG: (S)- α -methyl-4-carboxyphenylglycine
MEA: multielectrode array
MEG: magnetoencephalography
mRNA: messenger ribonucleic acid
Mut. inform.: Mutual information
NM: neuromodulator amount
NMDA: N-methyl-D-aspartate
num.: numerical
OU: Ornstein-Uhlenbeck
Patt. comp. coeff.: Pattern completion coefficient
PDE4B3: cAMP-specific 3',5'-cyclic phosphodiesterase 4B
PKM ζ : protein kinase M zeta
PP1: protein phosphatase 1
PRP: plasticity-related protein
PTSD: post-traumatic stress disorder
RAM: random-access memory
Rel.: Relative
REM: rapid-eye movement
SLFS: strong low-frequency stimulus
STC: synaptic tagging and capture
STDP: spike-timing-dependent plasticity
STET: strong tetanic stimulus
Suppl.: Supplementary
US: unconditioned stimulus
VIP: vasoactive intestinal peptide
VTA: ventral tegmental area
WLFS: weak low-frequency stimulus
WTET: weak tetanic stimulus

1.1 Motivation and aims

Memory is evolutionary important to process and store information about an experience, such that it can later be recalled and learned from (Bear et al., 2016; Fusi and Wang, 2016). Thereby, memory forms the basis of making predictions for future situations. To maintain this capability for hours, days, and even years, information has to be stored in long-term memory. In other words, it needs to be consolidated.

Although first profound studies of memory consolidation were laid out more than 120 years ago (Müller and Pilzecker, 1900), the exact implementation in the brain is still largely elusive. Two main types of memory consolidation have been identified in the past decades (McGaugh, 2000; Dudai, 2004; Dudai et al., 2015):

1. synaptic consolidation, also referred to as initial or cellular consolidation, is thought to rely on the stabilization of existing synaptic weight changes “in place” (Tonegawa et al., 2015; Okuda et al., 2020);
2. systems consolidation, is thought to redistribute information across brain regions and to occur mostly during sleep and rest (Rasch and Born, 2013; Dudai et al., 2015; Mizuseki and Buzsáki, 2013; Mizuseki and Miyawaki, 2017).

While synaptic consolidation occurs on timescales of hours, systems consolidation occurs on timescales of days to months or even years (Dudai, 2004; Dudai et al., 2015). Based on groundbreaking findings in hippocampal slices, Frey and Morris (1997) proposed a molecular mechanism called “synaptic tagging and capture” (STC) to serve as the physiological basis of synaptic consolidation, which has ever since received substantial further experimental support (Dudai, 2004; Pastalkova et al., 2006; Moncada and Viola, 2007; Wang et al., 2010; Takeuchi et al., 2016) and has become widely accepted (Redondo and Morris, 2011; Okuda et al., 2020). Up to now, several mathematical descriptions of STC have been proposed and were used to reproduce experimental data from typical experiments (Clopath et al., 2008; Barrett et al., 2009; O’Donnell and Sejnowski, 2014; Ziegler et al., 2015; Li et al., 2016).

Recurrent neural networks can feature so-called Hebbian cell assemblies, which are groups of neurons that exhibit highly strengthened internal synaptic connections. These ensembles facilitate the

firing of their neurons upon activation of a fraction of them, which serves to recall information that has been stored in the synaptic connections. Due to this functionality, Hebbian cell assemblies have become widely acknowledged as memory representations at the recurrent network level (Hebb, 1949; Martin et al., 2000; Buzsáki, 2010; Eichenbaum, 2018).

Apart from one study investigating the consolidation of cell assemblies with an abstract model of STC but lacking crucial biological aspects (Päpper et al., 2011), and another study considering a paradigm of behavioral tagging in a feedforward network (Ziegler et al., 2015), the impact of STC on network dynamics remains unknown. Thus, the extensive previous theoretical and experimental work on STC notwithstanding, a most important demonstration has remained undone: showing that biologically grounded STC-based mechanisms account for the consolidation of Hebbian cell assemblies. Yielding a mechanistic explanation for the storing of long-term memory representations on timescales of hours, this would be one of the crucial steps to solidify the link between STC and synaptic consolidation.

There are more phenomena whose theoretical explanation would benefit from a realistic STC-based model of memory consolidation. Everyone knows from experience that knowledge is improved by rehearsal (i.e., by repetitive learning), but sometimes, repetitive recall or even the mere passage of time suffice to improve memories (Erdelyi, 2010; Karpicke et al., 2014; Wallner and Bäuml, 2018). Previous theoretical studies (Elliott and Lagogiannis, 2012; Elliott, 2016) found such improvement of memory strength in abstract models, however, the cases in which improvement of memories occurs in the brain, and particularly the biological mechanisms that lead to this, remain sparsely examined.

Another open question relates to the functional role of neuromodulation for the consolidation of memory representations. While it is known that neuromodulation is generally required for the transfer from short-term to long-term memory (Lisman and Grace, 2005; Wang et al., 2010; Takeuchi et al., 2016) as well as for STC (Navakkode et al., 2007; Wang et al., 2010), the impact of different concentrations of neuromodulator remains unknown. To investigate functional implications of neuromodulation, a model of STC may be extended to account for variable levels of neuromodulation. This would first serve to confirm the capability of STC to solve the distal reward problem (Hull, 1943; Izhikevich, 2007; Päpper et al., 2011; Tetzlaff et al., 2012). Beyond that, a neuromodulator-dependent model of STC would enable to examine whether STC can control which type of information is stored. This question becomes particularly interesting as recent studies suggest that the interplay between different brain areas could adjust the level of neuromodulation in hippocampus and neocortex to facilitate different types of memory (Takeuchi et al., 2016; Duzskiewicz et al., 2019).

In addition to single memories, consolidation on timescales of STC seems to exert substantial influence on many neuropsychological effects that involve multiple memories, while biologically realistic explanations for these effects remain elusive. Such phenomena that occur, inter alia, on timescales of minutes to hours are given by serial-position effects as observed in free recall experiments (Greene, 1986; Brown et al., 2007; de Almeida Valverde Zanini et al., 2012; Recanatesi et al., 2015), interference and enhancement for similar memories (Anderson, 1981; Anderson and Neely, 1996; Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014), or priming (Mongillo et al., 2008; Janiszewski and Wyer, 2014; Elgendi et al., 2018; Was et al., 2019). There are apparently no studies on the functional role of STC in these phenomena, while this role should clearly be investigated to collect further evidence for the hypothesis that STC is indeed the mechanism that describes synaptic consolidation.

Finally, phenomenological and attractor models of associated memories have succeeded in describing various aspects of free recall experiments (Raaijmakers and Shiffrin, 1981; Bradski et al., 1994; Howard and Kahana, 2002; Romani et al., 2013; Recanatesi et al., 2015). Some of these models have pointed toward the importance of the dynamics of overlaps between memory representations (Romani et al., 2013; Recanatesi et al., 2015, also cf. Kropff and Treves, 2006). However, it still remains unknown how corresponding memory representations can be learned and consolidated with biological mechanisms, and how their organization with specific overlaps emerges. Furthermore, the transition dynamics between memory representations in those models have been investigated with rate neurons only (Romani et al., 2013; Recanatesi et al., 2015), whereas spiking neurons can yield distinct and biologically more realistic results (cf. Song et al., 2000; Palm et al., 2014; Brette, 2015). Thus, it remains to be investigated how a spiking model featuring STC can account for the formation and dynamics of long-term memory representations describing characteristics of free recall.

To conclude, there are diverse memory dynamics on timescales of minutes to hours that await to be explained mechanistically. STC-based models of memory consolidation are plausible candidates for this since STC has been recognized as the established mechanism to describe synaptic consolidation.

In a recent review paper, Holtmaat and Caroni (2016) mentioned as one of the important challenges for future research on cell assemblies

“[...] to elucidate the mechanisms and functional logic of consolidation processes, from the time of acquisition to long-term consolidation of memories 12–14 h after acquisition.”

In this thesis, we contribute to closing this gap by developing a biologically well-grounded spiking neural network model that features spike-driven calcium-based synaptic plasticity and STC. Our major aims are to (cf. Fig. 1.1):

- ① employ synaptic plasticity and STC mechanisms to alter the synaptic connections in a recurrent neural network, such that we demonstrate the *formation, consolidation, and improvement of memories* represented by Hebbian cell assemblies;
- ② show that neuromodulator-dependent STC can, in a retroactive manner, not only control *that* information is stored but *which information is stored*;
- ③ relate a variety of behavioral and cognitive effects, occurring *on timescales of minutes to hours*, to STC-based *organizational interactions* of multiple cell assemblies.

By achieving these three aims, we demonstrate that synaptic changes caused by calcium-based, spike-driven synaptic plasticity and STC robustly enable diverse functionality at the network level. Through these network-level functions, we can describe cognitive phenomena of memory in a mechanistic manner.

1.2 Structure of this thesis

This thesis is divided into five chapters. After this introduction chapter, three chapters present and discuss the results and methods of this thesis (chapters 2–4). Each chapter corresponds to one of the three main research aims (see Fig. 1.1, section 1.1). Manuscripts of three articles, one published as

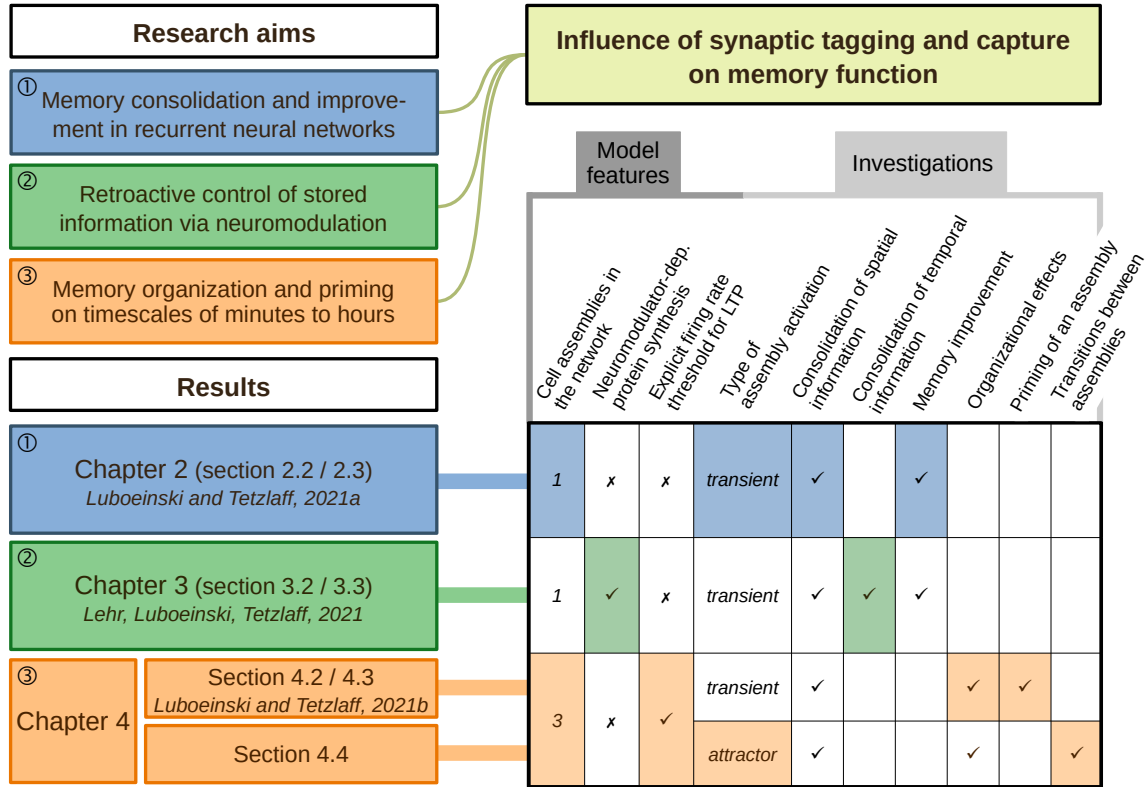


Figure 1.1: **Overview of the chapters and main studies in this thesis.** For each study, the right hand side indicates the variant of the model that was used and the major investigations that were performed. Colored background highlights the aspects that each study adds to the previous.

journal article (Luboeinski and Tetzlaff, 2021a), one under review (Luboeinski and Tetzlaff, 2021b), and one submitted (Lehr[#], Luboeinski[#], Tetzlaff, 2021; [#]shared first authorship), form the main parts of these chapters. Additional sections complement the findings of the manuscripts. Each of the results and methods chapters begins with a summary to put the findings of the chapter into the overall context of the thesis. Please note that since the articles presented in chapters 3 and 4 are based on the article in chapter 2, there is a certain overlap in the methods that are described.

In chapter 2 we develop the main part of our model, based on calcium dynamics for the induction of long-term synaptic plasticity and STC for its maintenance. We show that this robustly enables the consolidation of memory representations across parameters such as assembly size or network inhibition. Moreover, we demonstrate the improvement of memory recall through STC, which we further scrutinize with numerical and analytical methods. Finally, we discuss several approaches for the experimental verification of our model results. The chapter is closed by a section that briefly describes the development of the model and considers the phase space of early-phase plasticity for two different versions of the model. The main part of this chapter has been published in *Communications Biology* (Luboeinski and Tetzlaff, 2021a).

In chapter 3 we examine the impact of neuromodulator-dependent protein synthesis on the type of memory that is consolidated. To this end, we extend the model of memory consolidation that we developed in chapter 2 by introducing a neuromodulator-dependent protein synthesis threshold. We

vary the amount of neuromodulator as we employ STC to form and consolidate cell assemblies. Then, we compare the capability of the network to recall rate-coded and temporal patterns for different amounts of neuromodulator. Thereby, we demonstrate another fundamental property of STC – the selection between different types of neural coding for memory storage. This chapter is based on the submitted manuscript [Lehr et al., 2021](#).

In chapter 4 we consider the learning, consolidation, and interaction of *multiple* memory representations in a network, whereas chapters 2 and 3 have treated a single memory representation. Furthermore, we extend the model by a biologically realistic condition that requires the firing of pre- and postsynaptic neuron for the induction of long-term potentiation (LTP). The first part of chapter 4 consists of a manuscript ([Luboeinski and Tetzlaff, 2021b](#)) in which we investigate the organization and spontaneous activation of three memory representations. The results show that our model can provide a mechanism for neuropsychological findings on free recall, concept similarity, and priming, on long timescales of minutes to hours. The second part of chapter 4 demonstrates the modeling of cell assemblies with attractor behavior, enabled by a slightly adapted version of the model. Although this study mainly constitutes a draft for future investigations, we are able to reproduce findings from the first part of chapter 4 and examine the dynamics of transitions between active attractors.

In chapter 5, this thesis closes with a final discussion of our model, including a review of the main results that we obtained, and an outlook to questions that our model could be employed for in the future.

This introduction chapter shall present an overview over the experimental and theoretical background of the investigations of this thesis. As the manuscripts included in chapters 2–4 contain their own respective introductions, written to address expert readers, the last two sections of this chapter are intended to provide a broader overview. Section 1.3 introduces biological key mechanisms at the cellular and network level, while section 1.4 explains relevant aspects of the neuropsychological side of memory, thereby relating to different brain regions. Experimental findings and theoretical models are discussed side by side. Models are presented here in a manner intended to convey general ideas, whereas detailed technical descriptions have to be found in the respective methods sections in chapters 2–4.

Parts of this thesis are licensed under a Creative Commons CC BY 4.0 license, which can be retrieved from: <https://creativecommons.org/licenses/by/4.0/>. Please note that while the use of most mathematical symbols stays consistent throughout this thesis, in some cases the dimension of a quantity may vary across different chapters, which is a consequence of the different manuscripts included. Also note that we use the singular form here even if a certain brain region consists of two parts. The contribution of the author of this thesis to the presented manuscripts is declared in the beginning of the respective sections. To improve readability, all references are provided at the end of this thesis.

1.3 Neural networks in the brain

Since the late 19th century, it is known that the structure and function of the nervous system are based on nerve cells (neurons). Besides neurons, glial cells (glia) also play a certain functional role, however, they are not the target of the research questions that we pursue here. Neurons are coupled by different types of connections, most importantly, by chemical synapses and electrical synapses. The resulting

overall structure is called a neural network. Theoretical models describing neural networks try to capture aspects that are crucial for functions of the nervous system. Besides these biological neural networks, there are so-called artificial neural networks which are inspired by biology but constitute engineered approaches to accomplish certain computing tasks.

In this thesis we consider biological neural networks, with the main goal to get a better understanding of the mechanisms of the human brain. The similarity to brains of other animals ranges from little to huge, depending on species and brain area. Overall, mammals (other primates, especially) exhibit in many regards close comparability, whereas invertebrates often exhibit tremendous difference (Bear et al., 2016). Therefore, it can be helpful to consider findings from other animals depending on the studied brain area. Note that the mechanisms in the developing brain can be vastly different to those in the adult brain, and we will not discuss them here. Due to their relations to long-term synaptic plasticity and to declarative memory, hippocampus and neocortex are the brain regions that are most relevant to the investigations of this thesis (see sections 1.3.2 and 1.4.2).

1.3.1 Neurons and coding

Neurons are related to the basis of computation in the nervous system. They receive incoming signals from other neurons and process these, which results in discrete responses – the so-called *action potentials* or *spikes*. Spikes are the major way of communication between neurons. Even within a neuron, backpropagation of spikes as well as so-called dendritic spikes play an important role (see section 1.3.3). Following a spike, *excitatory neurons* exert a depolarizing influence on postsynaptic neurons (i.e., causing the net charge of the neuron to become more positive), whereas *inhibitory neurons* exert a hyperpolarizing influence (i.e., causing the net charge of the neuron to become more negative; see section 1.3.3).

The chance to trigger an action potential is called excitability. Early investigations on neuronal excitability were conducted, for instance, by Lapicque (1907), who also provided the first description of a neuron by an RC circuit and a membrane time constant. The first full characterization of action potentials was presented by Gasser and Erlanger (1922). Later, Adrian (1926) found in several species that the rate of spiking varies along with the stimulation strength, which was a seminal indication for *rate coding* (see subsection below). In a second part of the study, Adrian and Zotterman (1926) further investigated the adaptation of the rate after a certain time (cf. “intrinsic plasticity” in section 1.3.4).

Neurons usually consist of three main parts: the *dendrites*, the *soma*, and the *axon* (Fig. 1.2a; nevertheless, some neurons lack dendrites or axon). Neurons typically receive signals from other neurons via the usually highly branching dendrites. At these dendrites, many neurons express so-called *dendritic spines*, which are small protrusions of the membrane (Lee et al., 2005; Sheng and Kim, 2011; Bonilla-Quintana et al., 2021). The soma is the central part or cell body of a neuron and has a diameter of about 10 to 50 μm (Bear et al., 2016; also cf. 1.2b). The interior of the soma is filled with cytosol and contains organelles that are the same as in most other kinds of cells, including ribosomes, which perform *translation* of mRNA into proteins, as well as the nucleus, where *transcription* of DNA to mRNA takes place. Finally, there is maximally one axon, which typically consists of one long section followed by a lot of branches, each of which ends in a so-called axon terminal or synaptic bouton. An illustration showing these main parts of neurons is given in Fig. 1.2a, while microscope images of neurons and dendritic spines are shown in Fig. 1.2b-d.

All parts of a neuron are enclosed by an outer cell membrane, which serves important functions (see the subsection “Biophysics of the membrane” below).

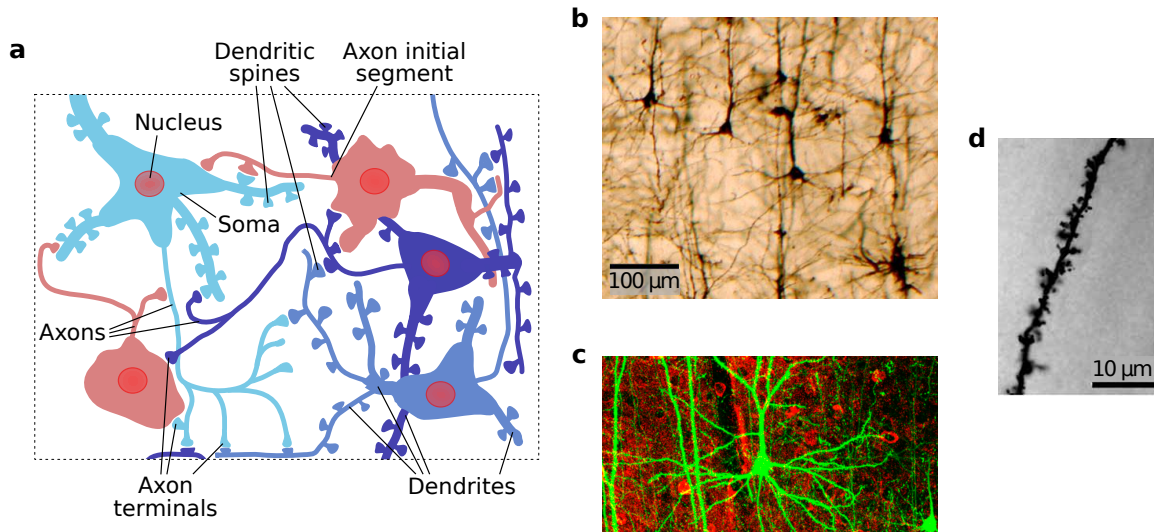


Figure 1.2: **Neurons, connected via synapses.** (a) Illustration of recurrently connected neurons, as they are found in areas of hippocampus and neocortex, with explanation of the different subcellular parts (proportions skewed). In contrast to excitatory neurons (shades of blue), inhibitory neurons (red) usually do not express dendritic spines. Moreover, inhibitory presynaptic neurons rarely connect to dendritic spines, whereas excitatory presynaptic neurons mostly connect to dendritic spines. (b) Image of a Golgi-Cox stained patch of human neocortex with pyramidal neurons (adapted from Džaja et al., 2014). (c) Image of pyramidal neurons (green) and GABAergic interneurons (red) in mouse neocortex, obtained in vivo with two-photon fluorescence microscopy (adapted from Lee et al., 2005). (d) Image of Golgi-Cox stained dendrite with dendritic spines in rat (adapted from Zhong et al., 2019). Black bars in (b) and (d) indicate the scale. All subfigures are licensed under CC BY 4.0.

Neuron types

While there are many different types of neurons in the nervous system, the most frequent neurons in the hippocampus and the neocortex are so-called *pyramidal neurons* (sometimes referred to as “principal” neurons; Freund and Buzsáki, 1996; Bear et al., 2016). These are glutamate-releasing, excitatory neurons with sometimes very long axonal projections, expressing dendritic spines, and with a soma whose shape typically resembles a pyramid. Their firing behavior usually follows a so-called regular-spiking pattern (Connors and Gutnick, 1990). In addition, another substantial fraction of the neurons in the hippocampus and in the neocortex are GABA-releasing inhibitory neurons exhibiting short-range connections, which is why they are often referred to as *interneurons*. Their firing behavior usually follows a so-called fast-spiking pattern (Connors and Gutnick, 1990). Inhibitory neurons are also often called stellate cells, although their shape is (unlike suggested by the term “stellate”) not always star-shaped (Freund and Buzsáki, 1996; Braitenberg and Schüz, 1998; Bear et al., 2016). Furthermore, there are excitatory stellate cells that exhibit dendritic spines (Bear et al., 2016). While excitatory neurons mostly connect via spines to other excitatory neurons (Sheng and Kim, 2011; Gluck et al., 2016), inhibitory neurons usually do not express spines and rarely connect to them (Tremblay et al., 2016).

Apart from spines, postsynaptic sites are also found directly at the dendrites, at the soma, and at the axon (see illustration in Fig. 1.2).

Biophysics of the membrane

The voltage difference that is measured across the neuronal membrane is called the *membrane potential*. The membrane potential is determined by the electric charges of ions on both sides of the membrane. Under resting conditions, the cytosol inside the neuron is negatively charged as compared to the extracellular fluid. Several mechanisms are necessary to actively maintain this dynamic equilibrium. The major contributors to this will be discussed in the following.

The membrane is essentially formed by a phospholipid bilayer. Ion channels, which are composed of one or several subunit proteins, span this bilayer, enabling for ions to move through the membrane. These channels exhibit selectivity for certain ions, for instance, for sodium or for all negatively charged ions (anions). The opening of some channels depends on the surrounding conditions, for example, voltage-gated channels open only if the membrane potential crosses a certain threshold. Ion channels that are open let ions pass passively through the membrane, but there are also ion pumps which actively (under consumption of energy provided by ATP) transport ions from one side of the membrane to the other.

Under resting conditions, the membrane potential is maintained at an equilibrium value by both passive flow of ions through channels and active transport of ions via pumps. The resulting potential V_{rev} is also referred to as *reversal potential* because it represents the value at which the ion flow across the membrane changes directions. The so-called sodium-potassium pump steadily carries three Na^+ ions from the cytosol to the extracellular fluid and two K^+ ions in the opposite direction, thereby consuming one ATP molecule. This active transport creates gradients in electric net charge (more negative inside the cell) and in concentration (excess of potassium inside, excess of sodium outside). Under resting conditions, all sodium channels are closed. However, other kinds of ion channels, such as the so-called potassium leak channels, are open and contribute significantly to the formation of the reversal potential (Dayan and Abbott, 2001). These channels enable both K^+ influx along the electrical gradient and K^+ outflow due to the concentration gradient. Note that the membrane potential is typically in a range which allows the thermal flow of ions along concentration gradients (Dayan and Abbott, 2001).

Membranes of nerve cells and muscle cells are excitable, meaning that they are in principle able to generate action potentials or spikes (Bear et al., 2016). An action potential is a brief process which begins with a rapid depolarization, followed by a rapid repolarization of the membrane potential, which eventually returns to the reversal potential. When an action potential occurs, the neuron is said to “fire”. The depolarization is triggered if the membrane potential at the so-called axon initial segment (see Fig. 1.2a) crosses a certain threshold voltage V_{th} . Next, voltage-gated sodium channels open and sodium ions, which are abundant in large numbers outside the cell, enter the cell. This starts a positive feedback loop that causes more depolarization, closure of the potassium leak channels, and opening of further sodium channels, even leading to the membrane potential changing its direction. At about +40 mV, the membrane potential starts to decrease again because sodium channels have become desensitized while particular voltage-dependent potassium channels are slowly opening. This repolarization phase is ended when the potassium leak channels have re-opened again. After having returned to the reversal potential, a so-called afterhyperpolarization, or undershoot, may occur where

the membrane potential falls below the reversal potential. This is caused by slowly-closing calcium-activated potassium channels.

The neuron is ready to fire again once the so-called *refractory period* has passed, which also determines the maximum firing rate of the neuron. The refractory period consists of two contributions: the time it takes until all sodium channels have returned from desensitization to the closed state, such that they can be activated again, is called the *absolute* refractory period. The *relative* refractory period denotes the time during which firing is hindered, although technically possible, and is caused by additional hyperpolarizing potassium channels that have not closed yet (Bear et al., 2016).

Rate- and spike-based models

Based on the series of action potentials fired by a neuron, its *firing rate* or *activity* can be computed. There are different definitions of the firing rate (see, for example, Dayan and Abbott, 2001). Most generally, the firing rate is defined by the average number of spikes per unit time. Thus, for long time intervals, it can easily be computed dividing the number of spikes that occurred by the time that passed. Other approaches are required if the – approximated – instantaneous firing rate at a certain time is wanted. One possibility is to compute the firing rate in a similar manner as described above, just for very small separate time intervals (this will be utilized in chapter 4). Another method is the convolution of the spike series with a Gaussian-shaped or rectangular window sliding along the time axis (this will be utilized in chapters 2 and 3).

The information processing which explains the stimulus-response relationship of neural systems is referred to as *neural coding*. Note here that it is controversial to use the terms “neural code” or “neural representation” regardless of context for a mere correlation between an experimentally detected property, such as firing rate or spike time, and, for instance, a concept selected by the experimenter (cf. Brette, 2019; for a definition of concept see section 1.4.3).

Coding strategies of neural systems are often divided into spike-based strategies (spike coding, or temporal coding) and strategies based on the firing rate (rate coding). However, a clear separation of such coding paradigms seems inappropriate regarding the large variety of different neural dynamics (Bugmann, 1997; Dayan and Abbott, 2001; Brette, 2015). In this thesis, it suffices mostly to quantify phenomena via firing rates, except for the temporal structures in chapter 3. Nevertheless, spiking dynamics are biologically more realistic and may give rise to distinct dynamics (Song et al., 2000; Dayan and Abbott, 2001; Palm et al., 2014; Brette, 2015), which we will discuss below in more detail. Hence, we simulate in this thesis the dynamics of individual neurons by a spiking model.

There are few cases where the brain actually performs irreducible rate coding (Gerstner and Kistler, 2002). Such rate codes seem to arise, most importantly, from population coding, which refers to the coding by a joint firing rate of a whole population of neurons. This kind of coding typically increases redundancy and thereby stabilizes responses. Georgopoulos et al. (1986) showed in a hallmark study that population coding serves, for instance, to yield precise control of arm movements. Other areas of emergent population coding are the control of eye movements and the encoding of visual stimuli (Averbeck et al., 2006). In summary, there is no strict notion of *the* firing rate. Rather, different kinds of firing rate measures can be useful to describe certain phenomena.

The theoretical advantage of rate-code formulations is that they can be thought to account for a whole population of neurons, which is not possible for spiking neurons as described in Eq. 1.4. A

rate-coding neuron or population i can be modeled in the following way (Wilson and Cowan, 1972, 1973; Knight, 1972):

$$\tau_v \frac{dv_i}{dt} = -v_i + \Phi(I_{\text{syn}}(t)). \quad (1.1)$$

The activation function (or “transfer function”) $\Phi(\cdot)$ is often chosen to be a sigmoidal or rectified linear curve, or, in the simplest case, a Heaviside function (Dayan and Abbott, 2001). The latter yields a binary neuronal response and was used by McCulloch and Pitts (1943) for their seminal neuron model, as well as by Hopfield (1982) for his seminal model of associative memory.

Assuming an exponential synaptic kernel (Dayan and Abbott, 2001; cf. section 1.3.3) the synaptic input current $I_{\text{syn},i}(t)$ is given by:

$$\tau_I \frac{dI_{\text{syn},i}(t)}{dt} = -I_{\text{syn},i}(t) + \sum_j w_{ji} v_j + I_0, \quad (1.2)$$

where w_{ji} is the strength of the connection from neuron (or population) i to neuron (or population) j , v_j is the firing rate of the presynaptic neuron (or afferent population) j , and I_0 is an external input current that accounts, for example, for background noise.

Eq. 1.1 can be explicated as follows: the membrane potential acts as a low-pass filter for the input current, such that the time-dependent firing rate can be regarded as a low-pass filtered outcome of the input current (Dayan and Abbott, 2001). Simplifications of the model can be made depending on the time constants τ_v and τ_I . If $\tau_I \gg \tau_v$ then Eq. 1.1 reduces to $v_i \approx \Phi(I_{\text{syn}}(t))$, whereas if $\tau_v \gg \tau_I$ then Eq. 1.2 reduces to $I_{\text{syn},i}(t) \approx \sum_j w_{ji} v_j + I_0$.

While rate-based models can be sufficient for slower dynamics, it takes spike-based models, which account for continuous-time dynamics, to model very fast dynamics such as stimulus pulses in physiological experiments. Integrate-and-fire models are a class of neuron models that describe the dynamics of the membrane potential as well as the generation of spikes (Gerstner and Kistler, 2002; Brunel, 2016). “Integrate” in this context refers to the summation of postsynaptic currents, which reach the neuron and are used to compute the membrane potential $V(t)$. “Fire” refers to the generation of a spike, which happens each time the membrane potential crosses the threshold value V_{th} . After a spike has been triggered, the membrane potential is usually set back to the reset potential V_{reset} , where it remains for the duration of the refractory period t_{ref} . If the reset potential is chosen to be lower than the reversal potential V_{rev} , it provides – in combination with the refractory period – a simple model of afterhyperpolarization.

The reversal potential is the potential that the neuron would reach in the infinite time limit $t \rightarrow \infty$ (numerically, after a certain time) if no input current is applied to it; it is the potential at which the net charge crossing the membrane is zero (as described in the subsection “Biophysics of the membrane”). However, an input current $I_{\text{syn}}(t)$ accounts for the inputs arriving from other neurons and is usually non-zero.

The so-called *leaky integrate-and-fire* (LIF) model is a widely used model that has often been employed to study neurons in stochastic conditions, to simulate large-scale networks, as well as to perform analytical investigations (cf. Brunel, 2016). It is useful if single-neuron dynamics are negligible compared to collective network phenomena. Consistently, we employ the LIF model in this thesis for simulating large neural networks with noisy input in order to study collective phenomena. “Leaky” refers to the

leakage term $V_{\text{rev}} - V(t)$, which models the conductance of the neuronal membrane, depicted by the membrane-spanning resistor R_m in Fig. 1.3a. Presumably (cf. Brunel and van Rossum, 2007), the LIF model in its modern form was introduced by Stein (1965) and named by Knight (1972).

The leaky Integrate-and-fire (LIF) model is sometimes called *linear* integrate-and-fire model because the underlying differential equation is linear (the reset, however, introduces a nonlinearity). This is done mostly to distinguish the LIF model from other integrate-and-fire models such as the quadratic integrate-and-fire model or the exponential integrate-and-fire model, which feature a quadratic leakage term and an exponential leakage term, respectively (Brunel, 2016).

Integrate-and-Fire models do not model the details of action potentials. Therefore, they cannot transmit information through the shape of the spikes; instead, they transmit information through the timing of the spikes. This can be justified by the fact that experimentally recorded action potentials in the same neuron almost stay constant in shape and proceed very fast (Gerstner et al., 2014; cf. Fig. 1.3b).

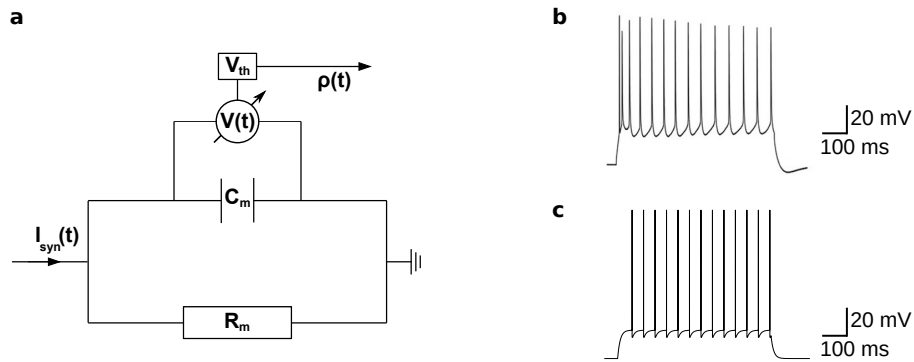


Figure 1.3: **Neuron model and dynamics of the membrane potential.** (a) RC circuit that underlies the leaky integrate-and-fire (LIF) model. The synaptic current $I_{\text{syn}}(t)$ charges the neuron, which responds by generating a spike each time the membrane potential $V(t)$ crosses the threshold V_{th} . The response function $\rho(t)$ represents the series of spikes. The membrane is charged as a capacitor C_m , but at the same time allows charges to pass through as a resistor R_m , which accounts for all ion channels and pumps. (b) Membrane potential of a pyramidal neuron from rat neocortex obtained from whole-cell patch clamp recording (adapted from Elliott et al., 2018), qualitatively compared to (c) membrane potential of a LIF model neuron (parameters: $V_{\text{th}} = -45$ mV, $V_{\text{reset}} = -50$ mV, $I_{\text{stim}} = 2.007$ nA, otherwise as in Table 2.1). Stimulation in (b) and (c) lasts ~ 600 ms. While the LIF model cannot reproduce the firing rate adaptation of regular-spiking pyramidal neurons, it can still approximate the main dynamics (particularly with respect to large populations, see text). All subfigures are licensed under CC BY 4.0.

Fig. 1.3a depicts the electrical circuit that the LIF model is based on. The neuron receives a synaptic current $I_{\text{syn}}(t)$ which accounts for presynaptic inputs from other neurons (see section 1.3.3). The net current that enters the cell by means of ion channels and pumps is modeled by a resistor R_m , while a capacitor C_m models the voltage gradient which additionally builds up across the membrane. Starting from the RC circuit in Fig. 1.3a, we can now derive the equation for the LIF model. Kirchhoff's junction rule tells us that

$$I_{\text{syn}}(t) = I_R(t) + I_C(t), \quad (1.3)$$

where $I_R(t)$ is the current that flows through the resistor R_m and $I_C(t)$ is the current that charges the capacitor C_m .

Using the definitions of capacity and Ohm's law, defining the membrane time constant $\tau_m := R_m C_m$, and performing further transformations, we obtain the following common form of the LIF equation:

$$\tau_m \frac{dV_i(t)}{dt} = -(V_i(t) - V_{\text{rev}}) + R_m I_{\text{syn},i}(t), \quad (1.4)$$

where the index i indicates a specific neuron.

More details of the biophysical dynamics of the membrane potential can be captured, for example, by introducing voltage-dependent conductances (Gerstner and Kistler, 2002). Such a model was developed by Hodgkin and Huxley (1952) based on their pioneering experiments at the giant squid axon. This model contains several differential equations for classes of conductances and can thereby precisely account for the curvature of an action potential. In principle, the Hodgkin-Huxley model adds more resistors to the LIF model and renders them voltage-dependent. Even more detail can be modeled by shifting from a point neuron to a neuron of finite extent, which is done by adding multiple compartments, described by so-called cable theory (Dayan and Abbott, 2001).

Another extension of the LIF model would be to add an adaptation mechanism for the spike rate, which can be achieved by a dynamic threshold voltage for spiking. In the multi-timescale adaptive threshold(2) (MAT2) model, for instance, the threshold is described by two exponential decay processes (Kobayashi et al., 2009; Yamauchi et al., 2011). Although the MAT2 model captures the characteristics of many more neuron types (Gerstner and Naud, 2009; Yamauchi et al., 2011), the LIF model is computationally better tractable because it does not require the computation of additional threshold dynamics for each neuron.

In a review on a whole range of neuron models, Izhikevich (2004) compared models with respect to their computational cost and biological detail. Among these models, the LIF model and the Hodgkin-Huxley model were the cornerstones of computational tractability and biological detail, respectively. Nevertheless, LIF neurons have been shown to account for the frequencies reliably encoded by populations of neocortical pyramidal neurons (Tchumatchenko et al., 2011) and to produce spike sequences that resemble those of fast-spiking cortical interneurons (Jolivet et al., 2004; Yamauchi et al., 2011; cf. Fig. 1.3c). In contrast, the Hodgkin-Huxley model and other conductance-based models have been shown to fail representing the frequency response of populations of pyramidal neurons (Fourcaud-Trocmé et al., 2003; Tchumatchenko et al., 2011). LIF neurons can therefore (still) be considered a standard neuron model for large networks (cf. Brunel, 2016). Nevertheless, using LIF neurons might not suffice if effects of detailed biophysical mechanisms, especially on timescales of the duration of an action potential, are the target of investigation.

On the other side of the neuron model spectrum, rate neurons provide a rather phenomenological description of the dynamics of neural networks (Brette, 2015). They can be useful to model the behavior of whole populations of neurons on long time scales, but do not enable a plausible description of, for example, learning and recall processes on fast timescales. Rate-based models only provide an exact account for large populations of neurons, and they only approximate the neuronal dynamics correctly if the inputs to the neurons in the population are mostly uncorrelated (Dayan and Abbott, 2001). If the inputs to a substantial fraction of the neurons are correlated – as it is the case, for instance, for a strong learning stimulus that is used to form a cell assembly – then the rate approximation does not hold. It is, therefore, necessary to use spiking dynamics to describe the learning and recall of cell assemblies in biologically realistic neural networks. Furthermore, regarding plasticity mechanisms, which will

be discussed in section 1.3.4, there are various self-organizing phenomena that seem to arise from the temporal characteristics of spiking dynamics that rate models cannot capture (Song et al., 2000; Palm et al., 2014).

1.3.2 Network structures and their activity

Nervous tissue exhibits a variety of structural characteristics related to certain functional properties. Traditionally, structure is studied by the discipline of anatomy, whereas function is studied by the discipline of physiology. Nowadays, however, interdisciplinarity has led to approaches that combine both structural and functional investigations (optogenetics or organoid research are prominent examples for this). Networks of interconnected neurons give rise to rich dynamics and emergent behavior. In principle, the nervous system of an organism can be thought to constitute one large neural network. However, the human cerebral cortex holds about 10^{10} , and the whole human brain holds about 10^{11} neurons (Herculano-Houzel, 2009). Thus, even if a suitable model existed, simulations of these structures would currently not be feasible – simply due to the lack of a computing system that would be powerful enough. Network models of smaller size are, nevertheless, essential for linking neuronal and molecular mechanisms to cognition and behavior, since they are able to predict dynamics at a mesoscopic level which is often not amenable to experiments. As the basis of such network models, we will discuss in the following aspects of neural connectivity and activity, while we will consider plastic changes of these in section 1.3.4. Another section (1.3.5) is dedicated to the topic of Hebbian cell assemblies, which are network structures of particular importance to memory. We will focus here on the brain regions of hippocampus and neocortex (cf. Fig. 1.6), as mentioned before, without loss of generality. The model which we introduce in this thesis can be applied to describe dynamics in any other brain region with similar properties.

The *neocortex* covers around 90% of the cerebral cortex and comprises, in terms of evolution, the “latest” areas of it. It only exists in mammals (Bear et al., 2016). The neocortex is vertically organized in six layers of cells (Brodmann, 1909) and is horizontally subdivided into modules (or columns) and larger areas (Buxhoeveden and Casanova, 2002; Bear et al., 2016). The *hippocampus* and parts of the olfactory system belong to the cerebral cortex but not to the neocortex (Braitenberg and Schüz, 1998; Bear et al., 2016). The hippocampus is a neural structure that belongs to the so-called medial temporal lobe. Its name originates from its characteristic appearance, resembling the shape of a small seahorse. The hippocampus is critical for declarative memory (cf. section 1.4) as well as for spatial navigation (O’Keefe and Dostrovsky, 1971; Bear et al., 2016). It is subdivided into the so-called cornu ammonis (CA) areas CA1–CA4, the dentate gyrus, and the subiculum.

In the brain, huge diversity of connectivity is observed, varying across brain areas but also across neuron types (Braitenberg and Schüz, 1998; Barbour et al., 2007; Le Duigou et al., 2014). Feedforward structures typically occur between subregions of hippocampus and neocortex, however, the connectivity within subregions is largely *recurrent* (Le Duigou et al., 2014; Palm et al., 2014; Rolls, 2016). Besides being necessary for the formation of cell assemblies (cf. 1.3.5), this corroborates our choice to investigate recurrent neural networks in this thesis. To get an idea of the orders of magnitude for the connectivity in hippocampus and neocortex, we will consider some experimentally obtained values in the following (usually from brain slices). The connectivity in guinea pig CA3 was found to be around 0.02–0.03 between pyramidal neurons (Miles and Wong, 1986). The connectivity between CA3

and CA1 pyramidal neurons was found 0.06 in guinea pig (Sayer et al., 1990). Another study found a value of 0.019 in rat CA3 as well as values of 0.018 for connections from CA3 to CA1 (Amaral et al., 1990). Organotypic slices, however, were found to exhibit much higher values of 0.3–0.6 (Le Duigou et al., 2014). The connectivity between pyramidal neurons in rat neocortex was found to be 0.1 for layer 2/3 (Holmgren et al., 2003) and layer 5 (Markram et al., 1997a). Another study found 0.11 for rat neocortical layer 5 (Sjöström et al., 2001; Barbour et al., 2007). Other parts of the cortex exhibit values up to 0.3 (Le Duigou et al., 2014). Braitenberg and Schüz (1998) estimated a value of 0.001 for connections between pyramidal neurons in mouse cortex. The probability of connection from pyramidal neurons to fast-spiking interneurons in layer 2 in rat neocortex was found to be as high as 0.5–0.7 (Holmgren et al., 2003). To conclude, values between 0.01 and 0.3 seem reasonable for the in vivo connectivity between pyramidal neurons in local patches of tissue, and we will use a value of 0.1 in this thesis.

The size of individual synapses, of which there can be multiple between each two neurons (Fauth et al., 2015), varies largely and is related to the strength of synaptic transmission (cf. late-phase plasticity in section 1.3.4). The distribution of synaptic strengths typically follows a lognormal distribution and is thus unimodal, although bimodal distribution is often assumed in models of synaptic plasticity (Barbour et al., 2007; Buzsáki and Mizuseki, 2014; also cf. models described in section 1.3.4).

For cortical network models, it is commonly assumed that there is a ratio of 4 : 1 excitatory to inhibitory neurons (Amit and Brunel, 1997a; Brunel, 2000; Gerstner and Kistler, 2002). Experimentally, Braitenberg and Schüz (1998) found a fraction of 5–21% inhibitory *synapses* in the cortex, which can fit this assumed ratio of neurons. Furthermore, to model local patches of tissue that contain not more than several hundreds to several thousands of neurons (which we do in this thesis), range differences of the connections formed by pyramidal neurons and interneurons can be neglected (cf. Braitenberg and Schüz, 1998).

Average neuronal firing rates differ depending on the type of neuron, the brain region, the wake/sleep state of the organism, and other modulatory influences. In rat hippocampus during exploratory wake states, pyramidal neurons have been found to exhibit average firing rates of 0.5–1.0 Hz (Mizuseki and Buzsáki, 2013; Mizuseki and Miyawaki, 2017). Margrie et al. (2002) found regular- and fast-spiking neurons in awake rat neocortex to fire at 0.1–1.0 Hz, and average firing rates of 5–10 Hz were observed in cat cerebral cortex during wake (Burns and Webb, 1976). In both rodents and humans, interneurons typically seem to fire at much higher rates as compared to pyramidal neurons (Frank et al., 2001; Viskontas et al., 2007). Moreover, firing rates during sleep often differ (Mizuseki and Miyawaki, 2017) and shall not be considered here. Based on these accounts, the steady-state firing rate of excitatory neurons in this thesis is adjusted to be ~ 1 Hz.

Similar to synaptic strength, activity in the brain has also been found to follow a lognormal distribution in many situations, while the transition to a scale-free distribution is observed in the presence of so-called neuronal avalanches (Buzsáki and Mizuseki, 2014). Neuronal avalanches are fast activation cascades that involve multiple neurons (Beggs and Plenz, 2003; Plenz and Thiagarajan, 2007; Tetzlaff et al., 2010; Priesemann et al., 2014). They typically occur in a critical network state between different dynamical regimes (cf. Brunel, 2000) and are often considered related to cell assemblies (see section 1.3.5 for definition of cell assemblies). If an avalanche is ongoing for a longer period of time, it can be considered (quasi-)attractor activity. In chapter 4, we will investigate both avalanche and attractor dynamics arising from spontaneous network activity.

1.3.3 Synaptic transmission and integration

Across centuries, the nature of signal transmission in the nervous system had been subject to diverse speculations, most prominently considering fluid flow in tubes and brain ventricles (dating back to Galen, 2nd century CE). Later, hydraulic and other mechanic interactions were considered, until the nervous structure was ultimately found to be realized by networks of cells connected via chemical and electrical synapses (Bear et al., 2016).

Around the late 18th to the mid-19th century, Galvani, Volta, and du Bois-Reymond conducted pioneering experiments on bioelectricity. Based on their work and assisted by his wife, Helmholtz made in frog nerves the hallmark finding that the electrical signal travels not “infinitely” fast but only at a velocity of a few meters per second (Helmholtz, 1850; Schmidgen, 2011), which was confirmed by Bernstein (1868), and precluded a theory of purely electrical transmission.

Later, Hermann (1899) and Bernstein (1902) developed a theory of bioelectrical membranes and conduction, and Foster and Sherrington (1897) proposed the concept of a cleft between axon and dendrite as well as the term “synapse”. The structure and function of chemical synapses was first conclusively characterized by Eccles and colleagues (Brock et al., 1952).

Signal transmission at chemical synapses

A chemical synapse is formed by a presynaptic axon terminal and a part of the postsynaptic neuron that depends on the neuron type, which has already been described in section 1.3.1. Electrical synapses, so-called gap junctions, differ in some properties from chemical synapses (Ostojic et al., 2009) and occur less frequently in the adult mammalian brain; we will thus not consider them here.

Following spike initiation, the action potential is forwarded through the axon to the axon terminals (see Fig. 1.2a). As soon as the action potential reaches an axon terminal, voltage-gated calcium channels are activated, leading to an influx of Ca^{2+} ions into the axon terminal. This triggers a biochemical cascade, eventually causing a process called exocytosis, in which vesicles filled with neurotransmitter fuse with the membrane at the so-called *active zone* and release their content into the synaptic cleft. This is, however, not perfectly reliable, and in some cases no release of neurotransmitter occurs at specific axon terminals (Pun et al., 1986). In this thesis, we will not explicitly model such effects because we anyhow consider the average synaptic connection between two neurons.

Following their release, neurotransmitters diffuse within the cleft and can bind to postsynaptic receptors. It is usually assumed that each neuron releases just one type of neurotransmitter – which is called Dale’s principle – and thereby exerts only excitatory or only inhibitory influence on other neurons. There are, however, some exceptions to Dale’s principle, and even the effect of a specific neurotransmitter depends on the postsynaptic receptor rather than on the neurotransmitter itself. Nevertheless, some neurotransmitters like GABA mostly evoke inhibition, while others like glutamate mostly evoke excitation (Bear et al., 2016).

The main neurotransmitter in excitatory synapses is glutamate. At the postsynaptic site, there are different types of glutamate receptors (cf. Fig. 1.4a), embedded in the so-called postsynaptic density, a densely-packed structure of proteins accounting for signaling and stability of dendritic spines (Sheng and Kim, 2011). While so-called α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors are most important for synaptic transmission, so-called N-methyl-D-aspartate (NMDA) re-

ceptors have been found most important for the induction of long-term synaptic plasticity (see section 1.3.4). These receptor types are *ionotropic*, meaning that they are actually ligand-gated ion channels that open once a neurotransmitter binds to their binding site. Other so-called *metabotropic* receptors trigger a biochemical cascade once a neurotransmitter binds to their binding site. Mediated by so-called G-proteins (guanosine triphosphate-binding proteins), the triggered cascade can lead to the activation of ion channels or enzymes (Bear et al., 2016). AMPA receptors are usually permeable to Na^+ and to K^+ ions. Once open, they lead to a net depolarization of the membrane because Na^+ influx outweighs K^+ outflow. NMDA receptors are also permeable to Na^+ and to K^+ and, in addition, to Ca^{2+} . This aspect is crucial, since Ca^{2+} not only further depolarizes the membrane by its electrical charge – it also serves as a second messenger for many intracellular processes, among them the induction of long-term synaptic plasticity (Artola et al., 1990). Besides this, metabotropic glutamate receptors (mGluRs) have been found to enable long-term synaptic depression independently of NMDA receptors (Bear et al., 2016; cf. section 1.3.4). NMDA receptors, moreover, exhibit voltage dependence and only allow inflow of calcium and other ions if the membrane is sufficiently depolarized (this is also important for plasticity, see section 1.3.4). Other sources of Ca^{2+} are voltage-dependent calcium channels (VDCCs; Malenka and Nicoll, 1999; Graupner and Brunel, 2012) and the rare calcium-permeable AMPA receptors (Hendley and Wilkinson, 2016). The net current that enters the postsynaptic site of a glutamatergic synapse through ion channels upon spike transmission forms a so-called *excitatory postsynaptic potential* (EPSP).

As opposed to this, GABA (γ -aminobutyric acid) is the neurotransmitter that is mostly released by inhibitory neurons (Bear et al., 2016). The ionotropic GABA_A receptors enable the fast inflow of Cl^- ions, thereby leading to a net hyperpolarization of the membrane. On the other hand, GABA_B receptors are metabotropic receptors and provide inhibition on slower timescales. The net current that enters the postsynaptic site by means of GABA receptors forms a so-called *inhibitory postsynaptic potential* (IPSP). Usually, the *strength* (also called *weight* or *efficacy*) of an excitatory synapse refers to the EPSP magnitude. Analogously, the (formally, negative) strength of an inhibitory synapse refers to the IPSP magnitude.

Synaptic transmission is normally ended rapidly due to re-uptake of neurotransmitter molecules by the presynaptic neuron or due to their decomposition by enzymes in the synaptic cleft. The EPSPs and IPSPs travel along the membrane and are finally integrated at the axon initial segment (see Fig. 1.2a). If their sum exceeds the threshold value (cf. section 1.3.1) a new spike will be triggered.

Note that neuromodulation is another type of communication between neurons which shares some commonality with synaptic transmission. The release of neuromodulators happens both at synapses and apart from synapses into the extracellular space (see section 1.3.6).

Models of synaptic interaction

The LIF model presented in Eq. 1.4 describes the neuronal membrane potential depending on a generic synaptic input current $I_{\text{syn},i}(t)$, which is usually composed of multiple contributions. In the scope of this thesis, these are:

- the total current $I_i^{\text{PSP}}(t)$ incoming from other neurons in the simulated network,
- the stochastic background current $I_i^{\text{bg}}(t)$,
- and the stimulus current $I_i^{\text{stim}}(t)$, which depends on the stimulation protocol.

To account for synaptic inputs from other neurons in the network (via $I_i^{\text{PSP}}(t)$), a model of synaptic interaction is needed. That means a model of EPSPs and IPSPs, depending on a spike k of the presynaptic neuron j occurring at time t_j^k . In the following, we will consider several possible models for this, which differ in their biological detail (cf. Gerstner and Kistler, 2002). The simplest variant is a *delta function* which delivers an infinitesimally short (in a numerical simulation: one timestep) current pulse to the postsynaptic neuron i . More realistic synaptic coupling is described by a *monoexponential function*:

$$I_i^{\text{PSP}}(t) = \sum_j \sum_{t_j^k < t'} w_{ji} \cdot \exp\left(-\frac{t - t_j^k - t_{\text{ax,delay}}}{\tau_{\text{syn}}}\right), \quad (1.5)$$

where w_{ji} is the strength of the (average) synaptic connection from neuron j to neuron i (if no connection exists, $w_{ji} = 0$), τ_{syn} is the synaptic time constant, $t_{\text{ax,delay}}$ is the axonal delay time, which in principle accounts for the duration of all the processes from spike initiation at the axon initial segment up to the opening of postsynaptic ion channels. The condition $t_j^k < t'$ with $t' := t - t_{\text{ax,delay}}$ denotes that only spikes that have already occurred are taken into account. Even more realistic synaptic coupling is given by a *biexponential function* with two different synaptic time constants:

$$I_i^{\text{PSP}}(t) = \sum_j \sum_{t_j^k < t'} w_{ji} \cdot \left[\exp\left(-\frac{t - t_j^k - t_{\text{ax,delay}}}{\tau_{\text{syn}}}\right) - \exp\left(-\frac{t - t_{n_j} - t_{\text{ax,delay}}}{\tau_{\text{rise}}}\right) \right], \quad (1.6)$$

where τ_{rise} is the time constant for the rising phase of the EPSP/IPSP. In the limit $\tau_{\text{rise}} \rightarrow 0$, this equation reduces to equation 1.5, which is sufficient for the fast synaptic dynamics mediated by AMPA receptors or GABA_A receptors. For the synaptic time constant, an effective value of $\tau_{\text{syn}} \approx 5$ ms can as well be assumed for both excitatory and inhibitory connections. This value resembles the typical decay time of GABA_A receptor-mediated IPSPs (Gerstner and Kistler, 2002; Roth and van Rossum, 2009) and can also account for EPSPs from a large number of fast AMPA receptors and a small number of slow NMDA receptors (cf. Gerstner and Kistler, 2002). Furthermore, the monoexponential model in Eq. 1.5 has a huge advantage in computational performance compared to the biexponential model in Eq. 1.6 because numerically, inputs can simply be multiplied by a factor for the exponential decay. Hence, in this thesis, we will use the monoexponential model.

Active dendrites

When a spike is triggered in the soma of a neuron, the excitation does not only travel down the axon to be transmitted to other neurons, but also propagates toward the dendritic tree (Stuart et al., 1997; Magee and Johnston, 1997). This so-called *backpropagating action potential* alters the membrane potential in the dendrites (which can be described in more detail by multi-compartment models) and elicits additional calcium influx, which is important for synaptic plasticity (Sabatini et al., 2002; Shouval et al., 2002; Graupner and Brunel, 2012; also see section 1.3.4).

Since voltage-gated ion channels also occur in spines and in dendrites, these structures do not only passively relay the electrical EPSP signal but can actively amplify it via so-called *dendritic spikes* (Häusser et al., 2000; Jahnke et al., 2015). This nonlinear dendritic computation is thought to play an important role in neuronal function, for instance, by introducing positive feedback. Although we do

not employ dendritic spikes in this thesis, we will discuss in chapter 5 their potential benefits to the model.

Modeling background current

In real biological networks, neurons usually receive inputs from a much larger number of presynaptic neurons than we could model. Such deterministic inputs were shown to yield noise-like dynamics (Hartmann et al., 2015; Effenberger et al., 2015), and noise or stochastic dynamics is important to avoid overly synchronous network activity. Thus, it is reasonable to model the external background input $I_i^{\text{bg}}(t)$ (see above) to the neurons in the network by a noisy process.

Poisson processes assume statistically independent spikes and can be used to model spike times of neurons at a certain average firing rate (see Dayan and Abbott, 2001). For exponentially decaying postsynaptic potentials (Eq. 1.5), the so-called diffusion approximation analytically describes the input current that arises from presynaptic Poisson spiking if the spiking activity is sufficiently high and the EPSPs/IPSPs are sufficiently small compared to the threshold V_{th} (Ricciardi, 1977; Moreno-Bote and Parga, 2010). Thereby, the background current $I_i^{\text{bg}}(t)$ is given by an Ornstein-Uhlenbeck (OU) process, described by a Langevin equation (cf. Gillespie, 1996):

$$\tau_{\text{syn}} \frac{dI_i^{\text{bg}}(t)}{dt} = I^0 - I_i^{\text{bg}}(t) + \sigma_{\text{wn}} \cdot \Gamma(t). \quad (1.7)$$

Here, τ_{syn} is the synaptic time constant, I^0 is the mean input current, $\Gamma(t)$ is Gaussian white noise with mean zero and variance $1/dt$, which approaches infinity for $dt \rightarrow 0$, and σ_{wn} is the white-noise standard deviation. To link the background current to a number N_{stim} of putative input neurons with average firing rates of f_{stim} and connecting via synapses of strength w , assuming homogeneous Poisson spiking, the mean current is $I^0 = wN_{\text{stim}}f_{\text{stim}}$ and the standard deviation is $\sigma_{\text{wn}} = w\sqrt{N_{\text{stim}}f_{\text{stim}}}$ (cf. Dayan and Abbott, 2001).

The autocorrelation function of the OU process is given by (cf. Moreno-Bote and Parga, 2010):

$$\left\langle \left(I_i^{\text{bg}}(t) - I^0 \right) \left(I_i^{\text{bg}}(t + \tau) - I^0 \right) \right\rangle = \sigma_{\text{OU}}^2 \cdot e^{-\frac{\tau}{\tau_{\text{syn}}}}, \quad (1.8)$$

where $\sigma_{\text{OU}} := \frac{\sigma_{\text{wn}}}{\sqrt{2\tau_{\text{syn}}}}$ has been introduced as the standard deviation of the OU process. Dynamics with such an exponential autocorrelation function is also referred to as “colored noise”, and it has been shown that an input current evoked by the spiking activity of many neocortical neurons exhibits the same power spectrum (Destexhe et al., 2003). In this thesis, we make use of this and employ the OU process to model biologically realistic background currents. Similarly, we also model the stimulus current $I_i^{\text{stim}}(t)$ mentioned earlier by an Ornstein-Uhlenbeck process.

1.3.4 Plasticity

Neural plasticity generally refers to the property of neural systems to undergo changes. These changes can be of various types and occur on various levels such as the macroscopic structure of the brain (mostly during development; Sanes et al., 2006), the connections between brain areas (Chang, 2014), and local synaptic connectivity (Bliss and Collingridge, 1993; Abraham et al., 2019). This thesis considers *synaptic plasticity* in neural networks, related to patches of tissue in hippocampus or neocortex, which are the most important brain regions for declarative memory (cf. section 1.4.3).

Synaptic plasticity describes changes in the strength of synaptic connections. Following the commonly accepted *synaptic plasticity and memory hypothesis*, the learning of short- and long-term memories occurs through such changes (Martin et al., 2000; Abraham et al., 2019; also see section 1.3.5). The idea that learning involves the strengthening and weakening of synaptic connections dates back to Bain (1873), James (1890), and Ramón y Cajal (1894). It was later refined by Konorski (1948), and, most influentially, by Hebb (1949), who wrote:

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”

The strengthening of a synapse by coincident pre- and postsynaptic activity, and the weakening by activity without coincidence, is therefore also called *Hebbian plasticity* (Sejnowski, 1999; Bliss and Collingridge, 1993; sometimes, however, only the strengthening is referred to as “Hebbian”). In short, this type of plasticity is often described by the saying “neurons that fire together wire together” (Löwel and Singer, 1992; Shatz, 1992).

Synaptic plasticity occurs on different timescales and in different forms. *Long-term (synaptic) plasticity* is the most studied form of synaptic plasticity and has been highly investigated in the mammalian brain at excitatory synapses involving dendritic spines (Abraham et al., 2019). The strengthening and weakening of synapses caused by long-term plasticity are called *long-term potentiation (LTP)* and *long-term depression (LTD)*, respectively. Long-term plasticity exhibits different phases, lasting from minutes up to months (Artola and Singer, 1993; Bear, 1996; Abraham et al., 2019). Modeling this type of plasticity forms a crucial part of this thesis, and will therefore be described in more detail in the subsections “Long-term potentiation and depression” and “Synaptic tagging and capture” below.

While long-term plasticity primarily operates at the postsynaptic site, *short-term plasticity* is a type of synaptic plasticity that operates at the presynaptic site, on timescales of milliseconds to seconds (Stevens and Wang, 1995; Lin and Faber, 2002). There are two types: Short-term facilitation denotes the increased release probability of neurotransmitter caused by enhanced calcium influx, and short-term depression denotes the depletion or lowered release probability of neurotransmitter. An influential model of short-term plasticity was presented by Tsodyks et al. (1998), similar forms of which have been used in many later studies (Mongillo et al., 2008; Hennig, 2013).

Furthermore, there is so-called *structural plasticity*, which is often considered separate from (“functional”) synaptic plasticity and describes the creation and removal of synapses on timescales even longer than those of long-term plasticity (Fauth et al., 2015; Holtmaat and Caroni, 2016). Due to their significantly shorter or longer timescales, short-term plasticity and structural plasticity are not investigated in this thesis.

In addition to the types of synaptic plasticity mentioned above, there are other specific types of synaptic plasticity which will not be discussed here (cf. Bliss and Collingridge, 1993; Bear et al., 2016). Furthermore, as indicated in the beginning of this section, there are types of plasticity which do not mainly act on individual synapses. Some of these types, which are mentioned at some point in this thesis, will be briefly introduced in the following.

Homeostatic plasticity is an umbrella term denoting types of plasticity that counterbalance excessive synaptic changes to maintain stable network dynamics. Prominent forms of homeostatic plasticity are

synaptic scaling, which introduces changes of the synaptic weights toward the restoration of an equilibrium firing rate (Turrigiano et al., 1998; Turrigiano and Nelson, 2000; Tetzlaff et al., 2011), synaptic normalization, which keeps the total synaptic weight within or across neurons constant (von der Malsburg, 1973; Oja, 1982; Lazar et al., 2009), and metaplasticity (“plasticity of plasticity”), which regulates the balance between potentiation and depression (Bienenstock et al., 1982; Bear et al., 2016; Abraham et al., 2019). While all types have been shown to stabilize circuits in theoretical studies, the experimental evidence is ambiguous. Experimental studies on synaptic scaling often reveal slow timescales, while mechanisms to balance runaway synaptic plasticity are required to act on fast timescales (Zenke and Gerstner, 2017). Synaptic scaling does not depend on NMDA receptors (Turrigiano and Nelson, 2000). For non-local synaptic normalization, there is no direct experimental evidence, but there are indications that at least some aspects of it could be implemented by heterosynaptic plasticity (cf. Royer and Paré, 2003; Tetzlaff et al., 2011; Zenke and Gerstner, 2017; Triesch et al., 2018; heterosynaptic plasticity will be treated below). Metaplasticity seems to depend on various factors, among them changes in the subunit composition of NMDA receptors (Bear et al., 2016).

Another type of plasticity is *intrinsic plasticity*, which changes the electrical properties, particularly the excitability, of a neuron (Turrigiano and Nelson, 2000; Cudmore and Desai, 2008). Thus, intrinsic plasticity is not a synaptic but a cellular phenomenon. By reducing the excitability of a neuron, intrinsic plasticity causes neuronal adaptation (in which way it also acts homeostatically). There is also evidence that increased excitability mediated by a CREB (cAMP response element-binding protein) transcription factor plays an important role in memory formation (Holtmaat and Caroni, 2016; Mather et al., 2016; Lisman et al., 2018). Intrinsic plasticity has been theoretically described, for example, in the SORN model (Lazar et al., 2009) and in the MAT2 model (Kobayashi et al., 2009; Yamauchi et al., 2011; cf. section 1.3.1). We will discuss the latter in chapter 5 as a possible extension of our model.

Furthermore, there are hints that besides synaptic and intrinsic plasticity, other cell-intrinsic mechanisms might be relevant for the storage of memories (Abraham et al., 2019). Such could be implemented by non-coding RNAs producing learning-related epigenetic changes in neurons (Landry et al., 2013). However, more data have to be gathered to describe the properties of such mechanisms.

Long-term potentiation and depression

LTP was discovered in the hippocampus by Terje Lømo in the 1960s and was first described by Bliss and Lømo, 1973 and Bliss and Gardner-Medwin, 1973 in rabbit in vivo. In evolution, LTP and LTD have proven to be very successful mechanisms, indicated by the high conservation across vertebrates and invertebrates (Abraham et al., 2019). Before LTP was described in the mammalian hippocampus, a similar phenomenon of long-term synaptic plasticity was investigated by Kandel and colleagues in the sea slug *Aplysia*, which has only about 20,000 neurons, some of them particularly large (in addition, many of the neurons are uniquely identifiable across specimens; Pinsker et al., 1970, 1973). Habituation experiments revealed that these sea slugs have two phases of memory storage, a protein-independent short-term memory lasting for minutes, and a protein-dependent long-term memory lasting for days, with an even longer-lasting subtype. Different phases of synaptic plasticity were found to correspond to these types of memory (Kandel, 2001). This resembles to a remarkable extent the findings on LTP in mammals, however, nailing down the correlates of memory in higher animals is much more complicated due its widely distributed nature (Martin et al., 2000; Abraham et al., 2019). The brain regions

mostly considered to study long-term synaptic plasticity largely correspond to those considered for memory (shown in Fig. 1.6). However, note that for example in the cerebellum, which is important for implicit memory, mechanisms of synaptic plasticity can be vastly different as compared those in the cerebral cortex that we consider here (cf. Smolen et al., 2012; Gallimore et al., 2018).

Since its discovery, experimental studies have identified at least three different phases of LTP, depending on the following processes (Racine et al., 1983; Bliss and Collingridge, 1993; Nguyen et al., 1994; Abraham et al., 2002, 2019):

- | | | |
|--|---|--------------------------|
| 1. pre- and postsynaptic activity at a synapse, | } | early phase (induction) |
| 2. translation of existent mRNA to produce PRPs, | } | late phase (maintenance) |
| 3. transcription in the nucleus to produce mRNA, | | |

where phase 2 is triggered by phase 1, and phase 3 is triggered by phase 2. Similar properties are observed for LTD (Sajikumar et al., 2007; Smolen et al., 2012; Abraham et al., 2019). The latter two phases are dependent on plasticity-related proteins (PRPs) and are often jointly referred to as late-phase LTP/LTD (L-LTP/L-LTD), whereas the first phase is called early-phase LTP/LTD (E-LTP/E-LTD; Frey and Morris, 1997; Abraham et al., 2019). We follow this scheme here, as models featuring two timescales suffice to capture the main characteristics of experimentally induced long-term plasticity (Clopath et al., 2008; Barrett et al., 2009; Li et al., 2016). Furthermore, we will occasionally refer to late-phase LTP/LTD as “late-phase plasticity” and to early-phase LTP/LTD as “early-phase plasticity”.

Traditionally, long-term plasticity has most often been studied in the mammalian hippocampus. In a typical experiment, a transverse slice of hippocampus is used to stimulate the so-called Schaffer collaterals, which are axon bundles that project from pyramidal neurons in the region CA3 to pyramidal neurons in the region CA1 (cf. Bliss and Collingridge, 1993; Frey and Morris, 1997; Sajikumar et al., 2005; Navakkode et al., 2007; Fig. 1.4b). Fiber bundles are stimulated by electrodes at two different locations (S1 and S2), while another electrode in the cell body layer of CA1 records the local-field response (fEPSP), which typically captures the response of several neurons. Changes in synaptic strength are assessed by applying brief test pulses and comparing the response to the baseline value (usually stated in percent of the baseline as %fEPSP). Long-term plasticity has been found to be highly dependent on the stimulation frequency (see Fig. 1.4c). Early- and late-phase plasticity are typically evoked by the following protocols (Bear, 1996; Sjöström et al., 2001; Sajikumar et al., 2005, 2007; also see Fig. 2.10 for detailed protocols, and the inset in Fig. 1.5a for typical outcome of WTET and STET):

- strong tetanic stimulation (STET) to induce late-phase LTP,
- strong low-frequency stimulation (SLFS) to induce late-phase LTD,
- weak tetanic stimulation (WTET) to induce only early-phase LTP,
- weak low-frequency stimulation (WLFS) to induce only early-phase LTD.

For the interaction between such protocols related to late-phase LTP/LTD, see also the subsection “Synaptic tagging and capture” below.

A tremendous change in the view of long-term plasticity came about by the finding that it exhibits high temporal precision, also called spike-timing-dependent plasticity (STDP). This was re-

vealed through different studies in the 1990s (Gerstner et al., 1996; Magee and Johnston, 1997; Markram et al., 1997b; Bi and Poo, 1998), showing that the time difference between a pre- and postsynaptic spike determines the induced direction and magnitude of long-term plasticity (see Fig 1.4d; more details are described below in the subsection “Models of LTP and LTD”). After that, many studies have demonstrated that spiking does not only implement irreducible dynamics in static networks (as discussed in section 1.3.1) but is also crucial to model different features of realistic learning, at the fast timescales at which it typically takes place in vivo. Competition between synapses to control the timing of spikes, for example, is inherent to STDP but cannot be captured by purely rate-based models (Song et al., 2000), self-organized synaptic weight structures in cell assemblies have been shown to arise from spiking dynamics and STDP (Palm et al., 2014), and spiking is required to explain various other aspects of temporal dynamics (cf. Harris, 2005; Wörgötter and Porr, 2005; Eichenbaum, 2013; Liang et al., 2020).

Later, it was demonstrated that spike timing and the firing rate jointly determine synaptic plasticity (Sjöström et al., 2001; Froemke and Dan, 2002). Sjöström et al. (2001) pointed out that their results put an end to what they called the “artificial dichotomy between STDP and classical LTP/LTD”, which had dominated earlier research. Nevertheless, Graupner et al. (2016) came to a different conclusion, emphasizing the importance of rate over spike timing for the induction of plasticity.

Morris et al. (1986) showed that NMDA receptors are crucial for both LTP and memory, thereby providing early evidence for the relation between synaptic plasticity and higher-level functions. Later, it was demonstrated that LTD requires NMDA receptors as well (Artola et al., 1990; Dudek and Bear, 1992). Under resting conditions, NMDA receptors are blocked by an Mg^{2+} ion, which only leaves if the postsynaptic membrane is sufficiently depolarized (Ascher and Nowak, 1988; see Fig. 1.4a). Furthermore, as explained in section 1.3.3, glutamate released from the presynaptic axon terminal must bind to the NMDA receptor to finally open it. By these two dependencies, the postsynaptic membrane potential on the one hand and presynaptic glutamate on the other hand, NMDA receptors act as biophysical *coincidence detectors* in the sense of Hebbian plasticity.

Early-phase plasticity (i.e., E-LTP and E-LTD) is thought to depend on the insertion and removal of AMPA receptors into and from the postsynaptic membrane, and potentially on their phosphorylation and subunit composition, while the number of “slots” for AMPA receptors in the membrane seems to stay relatively constant (Redondo and Morris, 2011; Henley and Wilkinson, 2016; Becker and Tetzlaff, 2021). Hence, for the induction of LTP, a synapse will undergo a cascade of molecular events that finally lead to an increased AMPA receptor-mediated response to neurotransmitter. Analogously, a different cascade will lead to a decrease in AMPA receptor-mediated response, yielding LTD (Smolen et al., 2014; Bear et al., 2016). Please note that these biochemical cascades exhibit many details that will not be described here. Nevertheless, they are essentially triggered by different calcium concentrations in the postsynaptic spine (Lynch et al., 1983; Bliss and Collingridge, 1993; Bear and Abraham, 1996; Cummings et al., 1996). This is of particular importance for this thesis because we will model the induction of early-phase plasticity via such calcium dynamics, based on previous models that successfully reproduced experimental data by this approach (Graupner and Brunel, 2012; Li et al., 2016; also see below). It has been found experimentally that high increase in postsynaptic calcium concentration, which follows strong depolarization (cf. section 1.3.3), causes the activation of Ca^{2+} /calmodulin-dependent protein kinase II (CaMKII). This enzyme undergoes autophosphorylation and thereby sustains its activity beyond the presence of the calcium signal (Bhalla and Iyengar,

1999; Lisman and Zhabotinsky, 2001; Smolen et al., 2014), which has made it a candidate for the synaptic tag (Redondo and Morris, 2011; Smolen et al., 2014; Bear et al., 2016; see subsection “Synaptic tagging and capture” below). Subsequently, CaMKII may cause the phosphorylation of AMPA receptors and structural changes (Redondo and Morris, 2011; Bear et al., 2016). While CaMKII autophosphorylation thus seems important for the *induction* of LTP, it has been found not necessary for the *maintenance* of LTP (Smolen et al., 2014). On the other hand, moderate postsynaptic calcium concentration evokes LTD, for which a dephosphorylating enzyme called protein phosphatase 1 (PP1) seems to be important (Bhalla and Iyengar, 1999; Smolen et al., 2012). Furthermore, inter alia, so-called extra-cellular signal-regulated kinases (ERKs) have been found to be involved in both LTP and LTD in different ways (Smolen et al., 2012).

Late-phase weight changes (and also the synaptic tag; see below) have been related to structural changes at the postsynaptic site, possibly leading to the increase or decrease of “slots” for AMPA receptors in the postsynaptic membrane (Redondo and Morris, 2011; Becker and Tetzlaff, 2021). These changes are thought to depend on alterations of the volume and shape of dendritic spines, based on their actin skeleton (Matsuzaki et al., 2001; Redondo and Morris, 2011; Meyer et al., 2014; Pinho et al., 2020; Bonilla-Quintana and Wörgötter, 2021; not to be confused with “structural plasticity” mentioned above). Studies in rodents have demonstrated that a certain fraction of dendritic spines can persist over months (although most have a lifetime of days; Yang et al., 2009), and that also long-term potentiation can last for months (Abraham et al., 2002). Moreover, the dependence of late-phase plasticity on protein synthesis has led to the discovery of heterosynaptic plasticity effects which have elucidated the mechanisms that underlie late-phase plasticity. These mechanisms are described by the so-called synaptic tagging and capture (STC) hypothesis (Frey and Morris, 1997; Pinho et al., 2020; Okuda et al., 2020), which we will discuss in the subsection “Synaptic tagging and capture” in more detail.

The dynamics of plasticity induction and maintenance described above lead to the following key properties of LTP and LTD (cf. Bliss and Collingridge, 1993; Bear and Abraham, 1996; Abraham et al., 2002; Barbour et al., 2007; Eichenbaum, 2011; Abraham et al., 2019): 1. *cooperativity* (sometimes called *associativity*): the membrane needs to be depolarized sufficiently in order to unblock NMDA receptors, therefore, synapses have to “cooperate” and can thereby become associated, which has been related to classical conditioning; 2. *input specificity*: only synapses that have received according input will be potentiated or depressed; 3. *longevity*: lasting for up to a year; 4. *scalability*: a broad range of synaptic strengths can be reached (including near zero, when depression leads to so-called silent synapses without AMPA receptors); 5. *controllability*: regulation through neuromodulators in different ways. Although the common Hebbian notion is that early-phase plasticity is induced by the firing of pre- and postsynaptic neuron, many situations require a so-called third factor (Bliss and Collingridge, 1993; Gerstner et al., 2018; Abraham et al., 2019). This factor is often related to a neuromodulator, and will be discussed in section 1.3.6 (also cf. Fig. 1.4a).

There is strong evidence that LTP and LTD are phenomena which are expressed at the postsynaptic site (Lynch et al., 1983; Artola et al., 1990; Bliss and Collingridge, 1993), but there are also indications for retrograde messengers triggering LTP- and LTD-related changes in the presynaptic site (Bliss and Collingridge, 1993; Regehr et al., 2009). Furthermore, evidence is increasing that astrocytes (a type of glia) play an important role in long-term synaptic plasticity (Palacios-Filardo and Mellor, 2019). Such mechanisms are, however, not part of the investigations presented in this thesis.

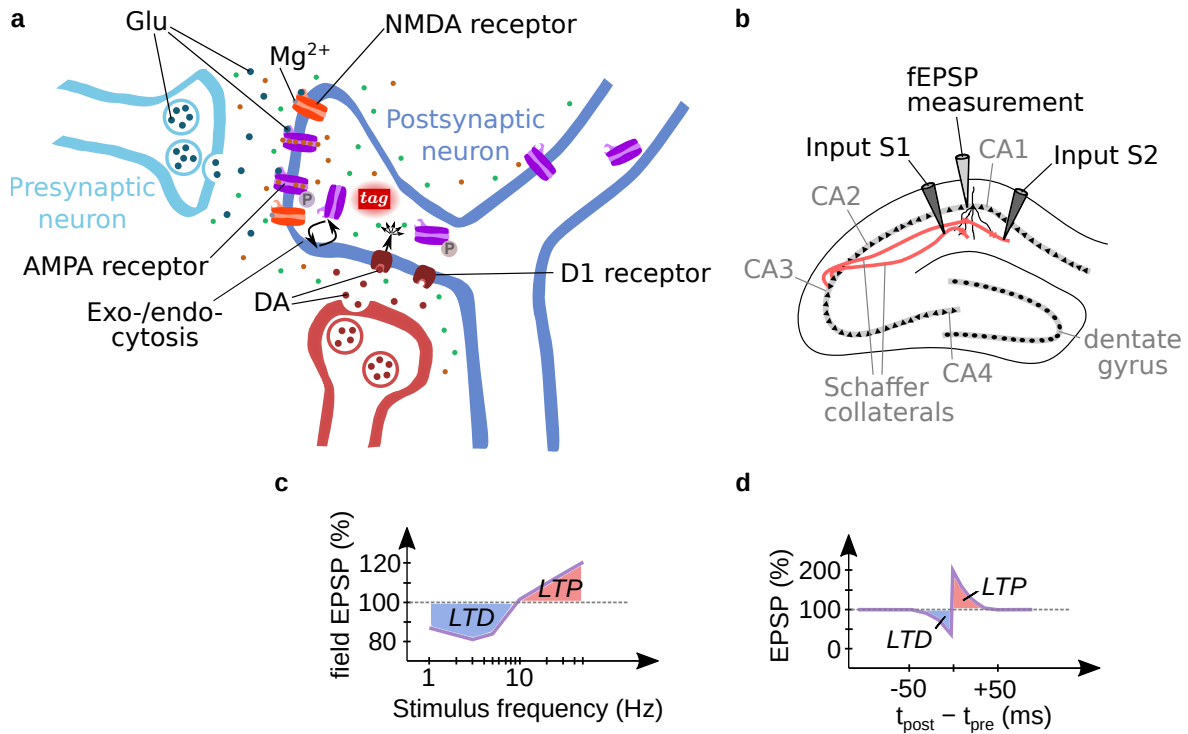


Figure 1.4: Induction of long-term potentiation and depression at an excitatory-excitatory synapse depends on frequency and timing of pre- and postsynaptic spikes, and on neuromodulation. (a) Following induction of long-term potentiation (LTP), a number of AMPA receptors have been phosphorylated (P) or inserted into the membrane by exocytosis, increasing synaptic transmission. A synaptic tag has been set via structural changes or enhanced enzyme activity. Synaptic transmission occurs through binding of glutamate (Glu), released from vesicles at the presynaptic site, to NMDA and AMPA receptors at the postsynaptic site, which in principle enables the flow of ions (small orange and green discs) through the attached channel. As long as the membrane is strongly hyperpolarized, NMDA receptors are blocked by an Mg^{2+} ion that prevents the influx of calcium ions and related synaptic plasticity. As an example for a “third factor”, dopamine (DA) is depicted here acting on metabotropic D1 receptors which elicit biochemical cascades that facilitate synaptic plasticity (cf. section 1.3.6). (b) Sketch of a transverse hippocampal slice with electrodes to stimulate the Schaffer collaterals, and an electrode to measure the local field response which is altered by synaptic plasticity (as in Frey and Morris, 1997; Sajikumar et al., 2005; Navakkode et al., 2007). Small triangles indicate the pyramidal neurons in the CA regions (discs indicate granule cells in the dentate gyrus). Red lines indicate stimulated axons and their synapses in CA1. (c) Depressed (LTD) or potentiated (LTP) response after stimulation of the Schaffer collaterals with different frequencies (data from Dudek and Bear, 1992). (d) Depressed or potentiated response depending on the difference between pre- and postsynaptic spike timing (as typically found in STDP experiments; cf. Markram et al., 1997b; Bi and Poo, 1998; Song et al., 2000). Subfigures (a) and (b) are licensed under CC BY 4.0.

So far we have discussed so-called *homosynaptic plasticity*, which means that the same synapse that receives stimulation undergoes plasticity. It is, however, also possible that long-term plasticity at a synapse is caused by the stimulation of another synapse, which is referred to as *heterosynaptic plasticity*. Heterosynaptic plasticity arises from the property of cooperativity mentioned above, however, the property of input specificity does not hold necessarily. Late-phase heterosynaptic plasticity is explained by synaptic tagging and capture, which we will discuss below, but the mechanisms enabling early-phase heterosynaptic plasticity are not completely understood (Li et al., 2016; Chater and Goda, 2021). An illustration of cross-tagging, which is a form of late-phase heterosynaptic plasticity, can

be found in Fig. 1.5b. Interestingly, for LTD, heterosynaptic plasticity was already discovered in the 1970s (Lynch et al., 1977), whereas it took almost 15 years until homosynaptic LTD was described as well (Dudek and Bear, 1992).

A specific type of long-term plasticity occurs at inhibitory synapses, therefore also called *inhibitory plasticity* (Vogels et al., 2011; Abraham et al., 2019). Its behavior differs, however, from that of plasticity at excitatory synapses described above. As inhibitory plasticity can give rise to interesting functional consequences, it has been employed by various theoretical models (e.g., Vogels et al., 2011; Zenke et al., 2015; Sprekeler, 2017; Herpich and Tetzlaff, 2019). Since we focus here on plasticity between excitatory synapses in the sense of classical Hebbian cell assemblies (cf. section 1.3.5), we will not discuss details of inhibitory plasticity.

Synaptic tagging and capture

Synaptic tagging and capture (STC) describes the transfer from the early phase (lasting up to several hours) to the late phase (lasting up to several days) of LTP/LTD. Strong experimental evidence has also been provided for the assumption that STC accounts for synaptic consolidation of memories, including so-called “behavioral tagging” (Dudai, 2004; Moncada and Viola, 2007; Wang et al., 2010; Takeuchi et al., 2016; Okuda et al., 2020; cf. sections 1.4.4 and 1.4.8).

STC was first described by Frey and Morris (1997), who discovered in hippocampal slices that STET stimulation at one synapse (delivered by an input through an electrode S1, cf. Fig. 1.4b) can cause late-phase LTP in another synapse which only receives moderate WTET stimulation (subsequently delivered by an input through an electrode S2), or which receives STET stimulation but in the presence of the protein synthesis inhibitor anisomycin. This led Frey and Morris to propose the existence of a *synaptic tag* which would occur at sufficiently potentiated synapses, rendering them eligible to the catching or *capture* of plasticity-related proteins (PRPs; the mechanism is illustrated in Fig. 1.5a). Frey and Morris (1997) found the tag to be transient with a lifetime of about 2–3 hours and by itself to be independent of protein synthesis.

While that first study concerned LTP, it was later found that there is also STC for LTD, as well as so-called cross-tagging. This refers to the case in which early-phase LTD is induced at one synapse and late-phase LTP at another, causing the former synapse to be transferred to late-phase LTD (Sajikumar and Frey, 2004a; Sajikumar et al., 2005; see Fig. 1.5b). The reverse case does also occur. Cross-tagging indicates the existence of a common pool of PRPs which are not specific to either LTP or LTD (Sajikumar et al., 2005, 2007).

Furthermore, in addition to this pool of common PRPs, there are indications for the existence of PRPs specifically needed for potentiation or depression. Protein kinase M zeta (PKM ζ), for example, can exhibit bistable dynamics similar to CaMKII and has been proposed to play an important role in maintaining L-LTP (and also long-term memory; Pastalkova et al., 2006; Smolen et al., 2014; Bear et al., 2016; but see Smolen et al., 2019 for how PKM ζ can be replaced by other atypical isoforms of protein kinase C). There are several other candidates for LTP-specific PRPs, including brain-derived neurotrophic factor (BDNF), whereas LTD-specific PRPs are yet to be identified (Barco et al., 2005; Sajikumar et al., 2007; Okuda et al., 2020). A candidate for the common pool is cAMP-specific 3',5'-cyclic phosphodiesterase 4B (PDE4B3; Sajikumar et al., 2005, 2007). Synthesis of mRNA and PRPs happens in the soma, from where these molecules are transported to the dendrites and spines (Abra-

ham et al., 2002; Dudai, 2004; Redondo and Morris, 2011), but PRP synthesis also happens locally in the dendrites (Martin et al., 1997; O'Donnell and Sejnowski, 2014; Glock et al., 2017). Nevertheless, it how exactly PRPs enter into STC mechanisms remains largely unclear (Okuda et al., 2020).

Another important discovery was the resetting of a tag through depotentiation (Sajikumar and Frey, 2004b), which has suggested a strong link between early-phase synaptic strength and synaptic tag. However, other studies also found that tag setting can be blocked using CaMKII- and actin-inhibiting drugs and that despite blocking, early-phase plasticity can be expressed (cf. Redondo and Morris, 2011). To resolve this, as opposed to the original STC hypothesis, Redondo and Morris (2011) proposed in an influential review paper the dissociation of mechanisms of early-phase plasticity and tagging (also cf. Okuda et al., 2020). Such an explanation can, in addition, account for a particular dopamine-induced form of late-phase plasticity called slow-onset potentiation (cf. Navakkode et al., 2007; Ziegler et al., 2015). Nevertheless, in most studies cited above, the induction of E-LTP and the occurrence of the tag have been correlated.

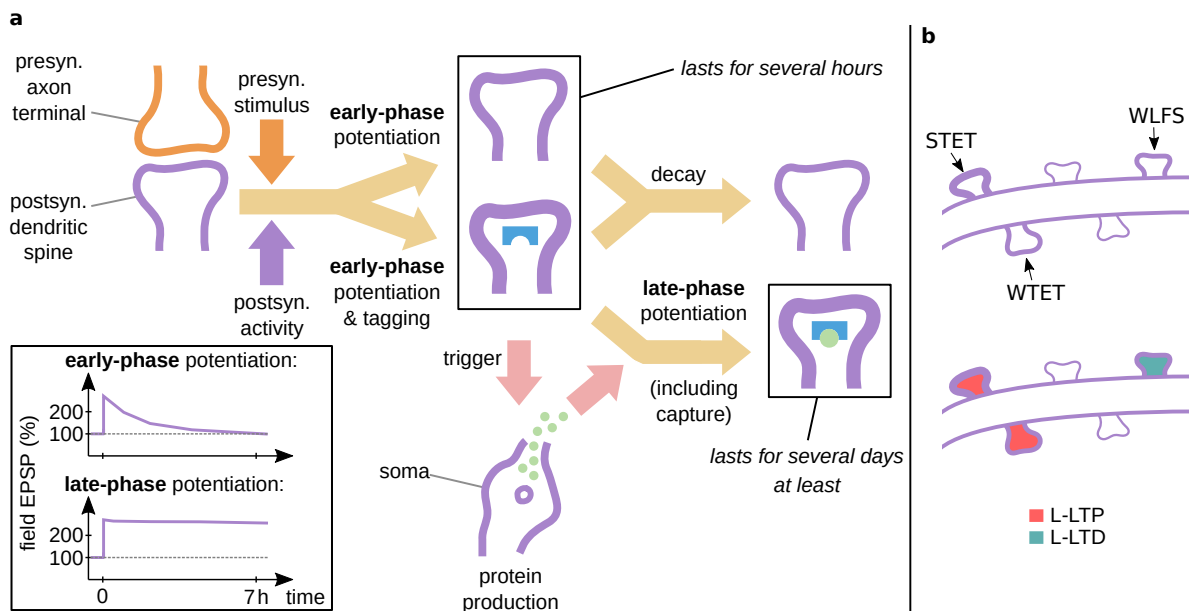


Figure 1.5: **Synaptic tagging and capture (STC).** (a) STC mechanisms for the induction (early phase) and maintenance (late phase) of long-term synaptic potentiation. Analogous mechanisms exist for synaptic depression. Inset: sketched weight dynamics for early- and late-phase potentiation as typically found in experiments (e.g., Frey and Morris, 1997; Sajikumar and Frey, 2004b). (b) Heterosynaptic late-phase plasticity and cross-tagging at different dendritic spines: STET elicits early- and late-phase potentiation by setting a tag and triggering protein synthesis, WTET elicits early-phase potentiation and sets a tag, WLFS elicits early-phase depression and sets a tag. Subsequently, through the protein synthesis triggered by the STET stimulus, also the latter two synapses enter the late phase. All subfigures are licensed under CC BY 4.0.

Models of long-term plasticity

Similar as with neuron models, one has traditionally distinguished between firing-rate models and spike-timing (STDP) models of long-term plasticity, whereas more modern models take both aspects into account.

Purely correlation-based Hebbian plasticity is inherently unstable and thus requires additional mechanisms of stabilization. We have mentioned in the beginning of this section that, for example, metaplasticity or heterosynaptic plasticity can provide such mechanisms. The Bienenstock-Cooper-Munro (BCM) model is a phenomenological, firing-rate-dependent model of synaptic plasticity, which counters strong potentiation or depression by down- or upregulating the firing-rate threshold that decides between LTP and LTD (Bienenstock et al., 1982; cf. Fig. 1.4c). This sliding threshold constitutes a form of metaplasticity. The model presented in this thesis features several mechanisms of homeostatic stabilization, including heterosynaptic late-phase LTP/LTD and adaptation of the weight change rate that can be considered a form of metaplasticity. Other classical firing-rate-dependent learning rules include “Oja’s rule” (Oja, 1982), which features synaptic normalization, and “Differential Hebbian” learning rules (Wörgötter and Porr, 2005). The latter account for temporal aspects by introducing a time derivative of the activity, and thereby resemble STDP.

An early model of STDP was presented by Song et al. (2000), fitting data of STDP experiments (Markram et al., 1997b; Bi and Poo, 1998) by exponential functions (cf. Fig. 1.4d):

$$\Delta w_{ji} = \begin{cases} A_+ \cdot \exp(-|\Delta t|/\tau_+) & \text{if } \Delta t > 0, \\ A_- \cdot \exp(-|\Delta t|/\tau_-) & \text{if } \Delta t \leq 0, \end{cases} \quad (1.9)$$

where $\Delta t := t_{\text{post}} - t_{\text{pre}}$ is the difference between the times of occurrence of pre- and postsynaptic spike. The case $\Delta t > 0$ corresponds to classical Hebbian causation, where firing of the presynaptic neuron is followed by firing of the postsynaptic neuron, which leads to potentiation. The reverse case, accordingly, leads to depression. The constants τ_- and τ_+ describe the decay of the exponentials for increasing difference between spike times. The amplitudes $A_+ > 0$ and $A_- < 0$ describe the maximum LTP and LTD that occur for spike timing close to coincidence. The weight change Δw_{ji} is simply added to the synaptic weights, which makes this an “additive” model (cf. Wörgötter and Porr, 2005).

Later, Gütig et al. (2003) showed that adding a nonlinear weight-dependent term to the model by Song et al. (2000) can substantially improve competition and stability in networks. In the following, models of “Triplet STDP” (Pfister and Gerstner, 2006; Clopath et al., 2010) were developed, accounting for the finding that considering interactions of two spikes alone is not sufficient to reliably reproduce the results of plasticity experiments. Specifically, the model by Clopath et al. (2010) features voltage-based STDP with homeostasis, and can therefore capture a diverse range of dynamics.

Since both firing rates and spike timing have been shown to be important for the induction of plasticity (Sjöström et al., 2001; Froemke and Dan, 2002), biologically more realistic models usually take both into account. As mentioned above, calcium plays a crucial role in the induction of LTP and LTD. Accordingly, by modeling calcium dynamics, theoretical models have achieved good results in reproducing experimental data (Shouval et al., 2002; Graupner and Brunel, 2012; Li et al., 2016). The model by Graupner and Brunel (2012) describes the plasticity of the synaptic weight as follows:

$$\tau_c \frac{dw_{ji}(t)}{dt} = -w_{ji}(t) (1 - w_{ji}(t)) (w_* - w_{ji}(t)) + \gamma_p (1 - w_{ji}(t)) \cdot \Theta [c_{ji}(t) - \theta_p] - \gamma_d w_{ji}(t) \cdot \Theta [c_{ji}(t) - \theta_d] + \xi(t), \quad (1.10)$$

where τ_c is a time constant, $c_{ji}(t)$ is the amount of calcium at the postsynaptic site, $\xi(t)$ is a noise term, γ_p and γ_d are rates of potentiation and depression, and θ_p and θ_d are thresholds for potentiation

and depression, respectively. The constant w_* defines the boundary between the basins of attraction of the two fixed points. $\Theta[\cdot]$ denotes the Heaviside theta function. The noise term and the dynamics of the calcium amount will not be discussed here (for details see Methods in chapters 2–4). Higgins et al. (2014) extended this model by introducing a correction factor for calcium concentrations in vivo (also cf. chapter 2). The dynamics of this model features two stable fixed points (corresponding to an “up” state and a “down” state). Such bimodality of synaptic weights seems, however, biologically implausible (Barbour et al., 2007; Buzsáki and Mizuseki, 2014), the utility of the model notwithstanding.

In particular with regard to modeling early-phase plasticity (for late-phase plasticity see below), the decay to a “center” fixed point seems more reasonable than “up” and “down” states. Li et al. (2016) took this approach and, starting from the model by Graupner and Brunel (2012), removed the bimodal dynamics, while introducing a single stable fixed point. In chapter 2, we will employ a model that is based on this unimodal model by Li et al. (2016). Nevertheless, there is still one drawback: the model does not necessarily account for the assumption of Hebbian plasticity, proposing that both the presynaptic and the postsynaptic neuron should fire to evoke LTP (Hebb, 1949; Bliss and Collingridge, 1993; Abraham et al., 2019). This may become functionally relevant at higher firing rates (Graupner and Brunel, 2012 only considered rates of up to 50 Hz), hence, we will introduce an according additional adaptation of the model in chapter 4.

Regarding the late phase of long-term plasticity, most recent models have focused on the mechanisms of STC. Nevertheless, several theoretical models have proposed different, mostly more abstract, mechanisms of multi-phase plasticity and synaptic memory consolidation (Gardner-Medwin, 1989; Fusi et al., 2005; Elliott and Lagogiannis, 2012; Tetzlaff et al., 2013; Zenke et al., 2015; Elliott, 2016). One of these models is the cascade model by Fusi et al. (2005), which has received a lot of attention regarding its rich computational capability while being more phenomenologically oriented and less biologically realistic than STC.

Considering models of STC, rather phenomenological (Clopath et al., 2008; Barrett et al., 2009; Papper et al., 2011; O’Donnell and Sejnowski, 2014) and biologically more detailed models (Smolen et al., 2012; Ziegler et al., 2015; Li et al., 2016) have been shown to reproduce typical experimental results for populations of synapses. Two studies further investigated memory dynamics at the network level. Papper et al. (2011) examined the consolidation of memories in a recurrent network, yet their STC model lacks important features of the induction of plasticity and of the dynamics of PRPs. Ziegler et al. (2015) presented a sophisticated model to demonstrate memory consolidation related to “behavioral tagging”. However, they only considered feedforward network structures, and additionally, their model only exhibits bistable weight dynamics and PRP dynamics independent of early-phase plasticity. Li et al. (2016) presented another sophisticated model that features, compared to the model by Ziegler et al. (2015), simplified tag dynamics, but instead describes more detailed PRP dynamics and a unimodal weight distribution (cf. early-phase models discussed above). However, the model by Li et al. (2016) does not describe dynamics at the network level. Hence, there is so far no study that has considered the formation and consolidation of memory representations in recurrent networks with biologically realistic dynamics. In this thesis, we therefore adapt the model by Li et al. (2016) to describe the consolidation of cell assemblies in a recurrent neural network (chapter 2). Using this network model, we are then able to investigate diverse memory dynamics related to neuropsychological effects (chapters 2–4).

1.3.5 Cell assemblies

Cell assemblies, or *ensembles*, are groups of neurons that are thought to represent memories at the network level (Palm et al., 2014; Tonegawa et al., 2015). While their exact definition and nature has been discussed for decades (Hebb, 1949; Braitenberg, 1978; Palm, 1981; Sakurai, 1999; Sejnowski, 1999; Buzsáki, 2010; Palm et al., 2014; Russo and Durstewitz, 2017), a common ground is that cell assemblies exhibit particularly strong internal connections and fulfill tasks that serve to store and recall information, such as pattern completion and the reactivation of temporal structures. By virtue of such functional properties, cell assemblies serve as memory representations of input or activity that are related to the strengthening of the internal synaptic connections (and possibly also to an increase in neuronal excitability, cf. section 1.3.4). Thus, cell assemblies provide a link between synaptic and neuronal processes on the one hand and memory on the other hand, which has been formulated in the *synaptic plasticity and memory hypothesis* by Martin et al. (2000):

“Activity-dependent synaptic plasticity is induced at appropriate synapses during memory formation, and is both necessary and sufficient for the information storage underlying the type of memory mediated by the brain area in which that plasticity is observed.”

Martin et al. (2000) further named four criteria that should be fulfilled for synaptic plasticity to play a critical role in memory storage: 1. detectability, learning should yield detectable synaptic changes; 2. mimicry, evoking the same synaptic changes in a naive animal should create the same memory; 3. anterograde intervention, preventing the synaptic changes should prevent learning; 4. retrograde intervention, it should be possible to erase the memory by modification of the synaptic changes. While there is still sparse evidence for mimicry, the other aspects have received strong experimental support (Abraham et al., 2019). Therefore, *memory traces* or *engrams* (Dudai, 2004; Holtmaat and Caroni, 2016; the term “engram” dates back to Semon, 1904) on the cognitive level are usually considered to correspond to cell assemblies on the network level. Cell assemblies thus correspond to features of concepts or even to whole concepts (Quiroga et al., 2008; Kiefer and Pulvermüller, 2012; cf. section 1.4.3 for a definition of concepts). Nevertheless, cell assemblies can also serve as building blocks for non-declarative memories, for example, for motor sequences (Nadh and Huyck, 2010; Tetzlaff et al., 2015).

Since Hebb (1949) provided an influential seminal theory of cell assemblies, one also refers to *Hebbian* cell assemblies. Based on findings by Lorente de Nó (1934), Hilgard and Marquis (1940) discussed reverberating activity in a group of neurons as a potential transient memory trace, which was later conceived by Hebb (1949), who argued that an activity-based trace would be implausible as a correlate for (long-term) memory representations. Hebb proposed that a memory should first be retained by reverberating activity before being consolidated by the strengthening of synaptic connections in the respective group of neurons. This is not far from the current notions of working memory and short-/long-term memory, since working memory indeed seems to be based at least partly on neuronal activity (cf. section 1.4.6), and short-/long-term memory seem to be based on synaptic plasticity (Martin et al., 2000).

Hebb had originally set out to explain how the experience of a stimulus could be represented by networks of neurons (cf. Bear et al., 2016). He argued that this had to be understood prior to the search for exact locations and implementations of memory in the brain, drawing from the misfortune

of Lashley (1929), who had not succeeded in finding the location for engram storage (cf. section 1.4.2). Hebb proposed that following the presentation of a cue, a whole related assembly would become active and stay active. Cell assemblies that exhibit such persistent activity are also called attractor cell assemblies, or attractors, and many theoretical studies have treated the issue of so-called attractor neural networks (e.g., Hopfield, 1982; Amit, 1989; Amit and Tsodyks, 1991a,b; Amit and Brunel, 1997b; Recanatesi et al., 2015; Miller, 2016). A more profound introduction to attractor cell assemblies will be given in section 4.4. Importantly, persistent activation of cell assemblies in theoretical models is usually considered implausible if it is not intrinsically shut down by biological mechanisms (Barbieri and Brunel, 2008). Such mechanisms can be, for example, short-term depression or neuronal adaptation (Kropff and Treves, 2006; Mongillo et al., 2008; Zenke et al., 2015; also cf. section 1.3.4). In the spiking model presented in this thesis, background noise will suffice to end the activation of a cell assembly.

An alternative view to the cell assembly as the neural correlate of memory has been the idea of “grandmother neurons”, proposing that single neurons account for the storage of a memory of, for example, one’s grandmother (cf. Palm et al., 2014). The idea is related to the established notion of neurons serving as sensory feature detectors in low-level brain regions (Barlow, 1953; Hubel and Wiesel, 1962; Eichenbaum, 2018). Single neurons responding to stimuli representing a specific person or another specific concept are indeed found in the medial temporal lobe. However, the idea seems implausible for memory representations because usually, neurons exhibit mixed selectivity to multiple concepts, neurons change their tuning, and memory is too robust to depend on single neurons (Quiroga et al., 2008; De Falco et al., 2016; Eichenbaum, 2018).

For a long time, investigating the theory of cell assemblies in physiological experiments seemed beyond reach (Braitenberg, 1978; Palm et al., 2014; Tonegawa et al., 2015; also cf. section 1.4.2 on brain regions involved in memory). Meanwhile, however, emerging techniques have enabled many new insights into the structure and dynamics of cell assemblies, and research on plasticity processes (see section 1.3.4) has revealed a lot about their possible formation and consolidation. Multielectrode arrays, for example, can be used in vitro to stimulate many neurons simultaneously as well as to measure their activity (Plenz and Thiagarajan, 2007; Liu et al., 2012a). For investigations in vivo, the field of optogenetics offers precise spatial and temporal stimulation of specified neuron genotypes (Chettih and Harvey, 2019; Russell et al., 2019; Luboeinski and Tchumatchenko, 2020) and can be combined with activity measurements through calcium imaging (Akerboom et al., 2013; Stamatakis et al., 2018). Most intriguingly, optogenetic stimulation has enabled to induce the recall of certain memories (Liu et al., 2012b). In addition, besides other genetic approaches, optogenetics has been used to recruit randomly selected neurons into assemblies by increasing the neuronal excitability, mostly via CREB (Holtmaat and Caroni, 2016; Mather et al., 2016; Lisman et al., 2018; cf. section 1.3.4). Furthermore, synchronized activity has been related to cell assemblies and can therefore be used to detect them (Louis et al., 2010; Palm et al., 2014). Another means of assembly detection is the statistical analysis of time lags in neural activity (Russo and Durstewitz, 2017).

Regarding the location of cell assemblies in the brain, the cerebral cortex (including neocortex and hippocampus) has by its anatomical and physiological properties proven to be a most suitable region for computationally rich cell assembly dynamics (Braitenberg, 1978; Palm et al., 2014). These properties are mainly: recurrent connectivity, positive feedback (since excitatory synapses greatly outnumber inhibitory ones), local networks globally connected via long-range connections, and Hebbian plasticity.

Nevertheless, much of the synaptic and neuronal processes underlying the formation, consolidation, and recall of cell assemblies remains elusive (Ziegler et al., 2015; Tonegawa et al., 2015; Holtmaat and Caroni, 2016).

Cell assemblies are subject to the so-called plasticity-stability dilemma, which refers to the issue of forming new memories on the one hand (through plasticity, also called “palimpsest”; Amit, 1989) and maintaining old memories on the other hand (stability; also cf. subsection “Interference” in section 1.4.8). There are different approaches to counter this (Grossberg, 1987; Gardner-Medwin, 1989; Amit, 1989; Fusi et al., 2005; Papper et al., 2011), one being the restriction of inputs and another being the consideration of several phases of plasticity. In this thesis, we follow both, by introducing neuromodulatory signals and by employing the two-phase mechanisms of STC.

The integration of memory representations into larger structures is described by schema theory, which will be treated briefly in section 1.4.3. Accordingly, cell assemblies are thought to form a network of associative memories (Kiefer and Pulvermuller, 2012; Palm et al., 2014) and have been suggested to enable functional interactions such as association, dissociation, and sequential activation (Herpich and Tetzlaff, 2019). Furthermore, multiple experimental studies have demonstrated that cell assemblies can overlap, meaning that neurons can be part of more than one assembly at the same time. Overlaps between cell assemblies related to similar concepts have been found in a range from around 1% up to more than 40% (Sakurai, 1999; De Falco et al., 2016; Cai et al., 2016; Yokose et al., 2017). We investigate the formation and impact of such overlaps in chapter 4, with the goal to connect to neuropsychological findings of similarity between concepts (cf. sections 1.4.3 and 1.4.8) and to describe transitions between attractor cell assemblies as suggested by previous theoretical studies (Romani et al., 2013; Recanatesi et al., 2015).

Moreover, neurons that belong to the periphery of a cell assembly, which means neurons that are either pre- or postsynaptic to other neurons in the assembly, seem to play an important functional role. While it has been shown that such neurons can critically enhance the computational capabilities of cell assemblies (Tetzlaff et al., 2015), we investigate in chapter 3 how STC can control this retroactively.

1.3.6 Neuromodulators and third factor

Neuromodulators are molecules that act either as neurotransmitters or as indirect modulators of neuronal properties. Neuromodulatory systems can exert influence in a global manner on neurons across brain regions. In fact, the same kind of molecule may act as a neurotransmitter both globally by so-called volume transmission and locally by synaptic transmission (Goto et al., 2007; Bear et al., 2016; Schultz, 2019; Liu et al., 2021). Presumably the most important neuromodulators with respect to learning and memory are *dopamine*, *serotonin*, *noradrenaline* (also called norepinephrine), and *acetylcholine* (Doya, 2002; Gerstner et al., 2018; Palacios-Filardo and Mellor, 2019). Here, we will restrict ourselves to properties of these four, focusing especially on dopamine, although neuromodulatory interaction is very complex overall, and several other neuromodulators and hormones also exert influences on memory (cf. Eichenbaum, 2011; Stober et al., 2020). Due to the “diffusive” nature of neuromodulatory systems, neuromodulation reaches many parts of the brain in parallel. Dopaminergic modulation has its origin mostly in the ventral tegmental area (VTA), in the substantia nigra of the basal ganglia, and by a smaller fraction in the locus coeruleus (LC; Bear et al., 2016; Duszkievicz et al., 2019). In turn, most of noradrenergic modulation originates from the LC. Serotonergic modulation mostly originates

from the so-called raphe nuclei, while cholinergic (acetylcholine-based) modulation has its origin in the basal forebrain, among other regions (Bear et al., 2016).

There are many scenarios which make it necessary to extend Hebbian plasticity rules by a so-called *third factor*, meaning that in addition to presynaptic spiking and the postsynaptic state, a third variable controls the direction and magnitude of plasticity. Such three-factor plasticity rules have also been termed “neoHebbian” (Lisman et al., 2011; Gerstner et al., 2018). The third factor can be *reward*, *punishment*, *surprise*, or *novelty*. Third-factor events are often induced by so-called primary reinforcers such as water, food, sex, temperature, or sleep, but they can also be induced by related secondary reinforcers and even by factors completely unrelated to primary reinforcers (cf. Hull, 1943; Gluck et al., 2016; Gerstner et al., 2018).

Phasic signals of the four neuromodulators mentioned above are important candidates to implement third factors (Gerstner et al., 2018; Palacios-Filardo and Mellor, 2019). While hippocampal-based learning has been shown to require dopamine (Pezze and Bast, 2012), striatum-based learning even seems to require simultaneous dopaminergic and glutamatergic input (Lindskog et al., 2006; Yagishita et al., 2014). Furthermore, both late-phase plasticity and synaptic memory consolidation have been shown to depend on dopamine and on the experience of novelty and on dopamine (Lisman and Grace, 2005; Navakkode et al., 2007; Moncada and Viola, 2007; Wang et al., 2010; Brzosko et al., 2015; Takeuchi et al., 2016), and also working memory critically relies on dopamine (Goldman-Rakic, 1995). Note that while most of the mentioned studies have targeted so-called D1-type dopamine receptors which cause an increase in activity of the second messenger cyclic adenosine monophosphate (cAMP), there are also D2-type receptors which cause inhibition of cAMP. Thereby triggering diverse biochemical cascades, dopamine exerts complex influence on synaptic and intrinsic plasticity (Berke and Hyman, 2000; Wörgötter and Porr, 2005; Sanes et al., 2006), and it should be noted that other neuromodulators have similarly complex impact. Moreover, dopamine and noradrenaline are subsequent elements of the same metabolic pathway, such that their dynamics are closely related (cf. Schultz, 1998; Bear et al., 2016). There is evidence that noradrenaline signals attention, at least for a short time, and general novelty, while dopamine can signal differential degrees of novelty based on VTA and LC systems (Lisman and Grace, 2005; Duzskiewicz et al., 2019). Acetylcholine also seems important for attention (Schultz, 2002; Doya, 2002) as well as for working memory (Hasselmo, 2006). Note that an attention-dependent fourth factor has recently been proposed, which could enable powerful computation through error backpropagation (Roelfsema and Holtmaat, 2018; Lillicrap et al., 2020). The role of serotonin as a third factor remains largely elusive, while there is evidence that it may be involved in signaling uncertainty or surprise (Palacios-Filardo and Mellor, 2019). In chapter 3, we abstract from these complex dynamics and consider a generic neuromodulator. Thereby, we are able to present general results and at the same time provide a model that can be fit to describe the dynamics of different neuromodulators. The differential novelty signaling of dopamine, for example, may provide a specific framework for the paradigms of varied neuromodulation that we consider in that chapter.

In three-factor plasticity rules, the pre- and postsynaptic variables give rise to a so-called *eligibility trace*, which can be thought of as a “pending” weight change, requiring an additional third signal to take effect. The third factor can be a global signal or a signal that affects only specific neurons, which

is described by the following equation in a generalized manner (Gerstner et al., 2018):

$$\frac{d}{dt}w_{ji}(t) = e_{ji}(t) M_i(t), \quad (1.11)$$

where w_{ji} is the synaptic weight, e_{ji} is the eligibility trace, and M_i is the third factor specific to the postsynaptic neuron i (for instance, the concentration of neuromodulator affecting the neuron at time t). Thereby, three-factor plasticity rules can describe a variety of neuromodulatory paradigms, which we will exploit in this thesis to enable/disable plasticity at certain times, as well as to model the influence of different neuromodulator levels on synaptic consolidation (cf. further below in this section).

The associative properties of early-phase plasticity/STDP (see section 1.3.4) and of behavioral conditioning paradigms (see section 1.4.1) seem related but act on vastly different timescales, such that a direct link cannot be established (cf. Wörgötter and Porr, 2005; Eichenbaum, 2011; Gerstner et al., 2018). Three-factor plasticity, however, can serve to bridge such time gaps and can thereby solve the *distal reward problem*, also called credit assignment problem, which describes the issue that an associative behavioral signal is sometimes present long after the related experience or learning event. On timescales of minutes to hours, STC is a particularly suitable candidate to solve this (Eichenbaum, 2011; Pöppel et al., 2011; Brzosko et al., 2015, 2017). In STC models, three-factor plasticity can be described by Eq. 1.11 if early- and late-phase plasticity are described by separate mechanisms (cf. Gerstner et al., 2018). For the model in this thesis, we can in fact apply the concept of Eq. 1.11 to both early-phase and late-phase plasticity dynamics, separately. In the former case, a nonlinear function of the calcium concentration constitutes the eligibility trace, whereas in the latter case, the synaptic tag constitutes the eligibility trace. Correspondingly, the third factor is the amount of neuromodulator for the early-phase dynamics (used particularly in chapter 4 to switch early-phase plasticity on and off), while it is the amount of PRPs for the late-phase dynamics (which is investigated depending on neuromodulator in chapter 3).

Reward is one of the most important third factors and plays a crucial role in reinforcement learning paradigms (cf. section 1.4.1). Therefore, we will discuss it here in some more detail. Reward signals seem to be implemented primarily by *phasic* increases of dopamine (Schultz, 1998; Berke and Hyman, 2000; Schultz, 2019). More specifically, the response of certain dopaminergic neurons is elevated by unexpected reward, lowered by unexpected absence of reward, and remains unchanged by expected outcome. Thereby, dopamine signals the so-called reward prediction error. Prior to this, the *tonic* level of dopamine seems to set the expected outcome (Schultz, 2002; Goto et al., 2007). An example for the impact of dopamine on behavior is the transition between controlled behavior and addiction, which is associated with compulsory reward seeking (Berke and Hyman, 2000). To characterize the role of dopamine in this, one has to distinguish between “wanting”, corresponding to the motivation to obtain a reward (conveyed by dopamine, inter alia; also called “incentive salience”), and “liking”, corresponding to hedonic value or pleasure (signaled by endogenous opioids, inter alia). Pure liking does not seem to cause addiction, whereas reward-seeking behavior induced by dopamine can cause addiction (Berridge, 2009; Bear et al., 2016). Caffeine consumption, for example, can create mild physical dependence but usually does not create addiction, as it seems not to be related to substantial stimulation of the dopaminergic system (Nehlig et al., 2010). Addictive drugs, on the contrary, act either directly (cocaine, amphetamine) or indirectly (morphine, nicotine) on dopamine concentra-

tions (Schultz, 2002; Bear et al., 2016), which demonstrates the strong impact of dopamine signals on behavior and psychiatric conditions.

In conclusion, there is plenty of evidence that three-factor plasticity involving neuromodulators is critical for learning and consolidation mechanisms in the brain. As we have pointed out, three-factor plasticity plays a role in different parts of this thesis, most importantly in the investigations of neuromodulator-dependent consolidation in chapter 3 and as on/off switch for early-phase plasticity, particularly in chapter 4. To cope with the vast complexity of neuromodulatory influences in biological systems, we consider generic neuromodulators in our model, which can be adapted to fit the dynamics of, for example, dopamine or noradrenaline.

1.4 Neuropsychological view of memory

While the mechanisms underlying learning and memory on the biophysical and network level were discussed in the previous section, this section shall introduce the basic aspects of memory on higher levels of processing and in behavior.

Around the 1920s to 1950s, the paradigm of behaviorism dominated many theories of learning (Mandler, 2002; Gluck et al., 2016). While radical behaviorism as it was upheld by Skinner (1971) assumes that consciousness and free will are illusions, more general behaviorism simply claims that the principle of action in animals including humans consist of processing external stimuli to result in external behavior (Watson, 1913). Therefore, behaviorist researchers were aiming to study stimulus-response relationships, while disregarding ideas about the information processing in the brain (cognition). To some extent, this was indebted by the lack of techniques to investigate processes inside the brain. In the past decades, many findings have challenged such behaviorism, demanding that in addition to stimulation and behavior, theories of learning should include aspects of cognition. This new paradigm, sometimes referred to as “cognitivism”, has recognized that mental representations significantly contribute to the behavioral response to given stimuli and thus need to be studied. However, unless measurements of neural structure or activity are performed, all outcome of neuropsychological experiments still has to be obtained from some kind of behavior. Therefore, in principle, the cognitivist paradigm does not refute the assertions of behaviorism but rather extends them by acknowledging the existence and importance of mental processes (cf. Mandler, 2002; Gluck et al., 2016).

According to this historical development, behavioral findings on the one hand and physiological findings on the other hand are not disconnected anymore, as both contribute to explain the principles of cognition. Thus, the neuropsychological perspective on learning and memory that is taken in this section and the network perspective conveyed in the previous section necessarily overlap.

Memory has three basic properties: it is *acquired* (encoded, learned), *stored* for a certain time, and *recalled* (Fusi and Wang, 2016). Different types of memory can be distinguished by the functions they serve and by the timescales they operate on. We will briefly discuss such types now, while more detailed information is provided in sections 1.4.3–1.4.6.

Declarative (or explicit) memory denotes memory that can be communicated flexibly by different means, for example verbally, and in formats different to the original format during encoding (Gluck et al., 2016). In another formulation, declarative memory is defined as the memory that is accessible to conscious recall (Squire, 2004). Thus, declarative memory may constitute one of the fundamental

structures by which the brain creates the mind. It features two major subtypes: *episodic memory* and *semantic memory*, a distinction that was first made by Tulving (1972). Short-term (for the notion used here see section 1.4.5) and long-term declarative memories can, possibly through attention, be activated and loaded into working memory, which renders them conscious (cf. James, 1890; Baars, 1997; Oberauer, 2013; but see Soto et al., 2011 and Trübutschek et al., 2017 for counter-statements). Declarative memories can be learned in a single exposure (Gluck et al., 2016). Human declarative memory also features *metamemory*, which enables the reflection and beliefs about one's own memory and common errors related to this, such as feeling-of-knowing effects. Metamemory facilitates "thinking about the thinking" (metacognition), and thereby relates to a theory of mind (Lockl and Schneider, 2007; Gluck et al., 2016).

Non-declarative (or implicit) memory includes procedural and skill memory, which usually require many repetitions to be learned. Essentially, non-declarative types of memory can not easily be verbalized (Gluck et al., 2016), and seem to be fundamentally different from declarative types of memory (Poldrack and Packard, 2003; Squire, 2004). Concepts analogous to working memory have also been proposed for non-declarative memory (Oberauer, 2013). In the following, especially in section 1.4.3, we will put an emphasis on declarative memory because the experimental paradigms that the theoretical investigations of this thesis are intended to fit mostly concern declarative memory (experiments that involve learning by a single trial as well as conscious recall). An exception to this is priming, which seems to be a form of non-declarative memory intermingled with declarative memory (cf. Squire, 2004; also see section 1.4.8).

Memory operates on different timescales in different ways, which suggests the distinction between different stages of memory. There is a consensus that at least three stages of memory exist (Rosenzweig et al., 1993; Dudai, 2004; Tetzlaff et al., 2012; Gluck et al., 2016; Bear et al., 2016). The definition of stages varies greatly, however. A hallmark approach was the Atkinson–Shiffrin or multi-store model (Atkinson and Shiffrin, 1968), proposing a sensory register (for visual, auditory, and haptic information; briefly decaying), a short-term store (the site of working memory operations; retained for up to several minutes through attention and rehearsal), and a long-term store (permanent storage that is accessed via short-term memory; retained for up to a lifetime). The model could, however, not account for many experimental findings. In later models, the sensory register and the short-term store were combined (cf. Raaijmakers, 1990), which roughly corresponds to what we refer to here as working memory (section 1.4.6). More recent models include the two-stage "search of associative memory" (SAM) model (Raaijmakers and Shiffrin, 1980, 1981) that could reproduce many experimental findings, among them certain free recall paradigms. Nowadays, working memory, short-term memory, and long-term memory are often considered the major stages of memory (Tetzlaff et al., 2012; Bear et al., 2016), whereas there is evidence for at least one more intermediate-term stage (Rosenzweig et al., 1993). Long-term memory, short-term memory, and working memory will be discussed in sections 1.4.4–1.4.6 in more detail. Please note that the definitions of these types of memory are not equivalent to the stages in the Atkinson–Shiffrin model.

1.4.1 Learning

Learning in general can be defined as the acquisition or modification of knowledge or skills, hence, the formation of memories. In spite of overlaps with experience-dependent brain development, learning

is usually considered a lifelong process (Tetzlaff et al., 2012; Bear et al., 2016).

Theories of learning date back at least to Aristotle (4th century BCE), who already proposed that memories would become associated by repetitive learning, similarity of the learned entities, and similarity of the spatial and temporal context. Traditionally, learning and memory were often considered separately, caused by the fact that learning had mostly been investigated in studies of animal behavior, while memory had been related to studies of the human mind (Gluck et al., 2016). Thus, forms of learning were classified by observing the behavior of animals upon the performance of a task. This has led to the distinction between *non-associative* forms of learning, namely, habituation and sensitization, and *associative* forms of learning, most importantly, classical conditioning and operant conditioning, among others (Gluck et al., 2016; Bear et al., 2016).

Habituation refers to the decreasing response to a stimulus following repeated exposure. Almost contrary to this, *sensitization* refers to the increasing response to a salient stimulus following repeated exposure. The neural underpinnings of habituation and sensitization were studied most prominently by Kandel and colleagues in the sea slug *Aplysia* (Pinsker et al., 1970, 1973).

Classical conditioning (or Pavlovian conditioning) was most prominently investigated since ~1900 by Pavlov, who found that a primary *unconditioned stimulus* (US) such as the sight of food can, through repeated training sessions, be paired with a *conditioned stimulus* (CS) such as the ringing of a bell. Eventually, the CS alone will evoke the same behavioral response that the US has evoked naturally (Pavlov, 1927). Such a stimulus-response pairing is also referred to as *habit* (Maia, 2009; not be confused with habituation). *Operant conditioning* (or “instrumental” conditioning) is another fundamental learning paradigm, based on reward and penalty instead of a direct pairing of stimuli. In this paradigm, positive and negative reinforcement and punishment influence an agent toward the execution or avoidance of certain actions. It was first described by Thorndike (1911) with his “law of effect” and further developed by Skinner (1938). In addition to these associative forms of learning, there is *extinction learning*, which leads to a decrease in association (also cf. subsection “Reconsolidation” in section 1.4.8). Employing three-factor learning rules (cf. section 1.3.6), the behavioral level of associative learning described here can be linked to mechanisms of synaptic plasticity.

Another classification of forms of learning considers the drivers behind the learning process, given by (also cf. Fig. 1.7):

1. correlations for *unsupervised learning*,
2. reinforcing factors for *reinforcement learning*,
3. instructing signals for *supervised learning*.

The first two of these forms play a particular role in this thesis. Unsupervised learning is the most “self-organized” form of learning and corresponds to Hebbian learning rules (cf. section 1.3.4). Reinforcement learning corresponds to “three-factor” learning rules, where the factor can be reward or penalty, which is typically signaled by neuromodulation (cf. section 1.3.6). Model-free and model-based algorithms of reinforcement learning resemble classical and operant conditioning, respectively (Sutton and Barto, 2018). Supervised learning enables, for example, offspring to learn from teaching or by imitation (Tetzlaff et al., 2012). Also note that this is currently the predominant form of learning in

machine learning, where instructing signals are often conveyed by error backpropagation (for which there may be a biological correlate; cf. section 1.3.6).

1.4.2 Important brain regions

To understand what memory is and how it is implemented, Lashley searched in the 1920s for the place where engrams or memory traces – the neural representations of memories – were stored. He investigated this by removing parts of the brain in rats and observing their behavior. Finally, contrary to the contemporary belief (cf. Bear et al., 2016), he came to propose that every part of the brain might serve to store memories, which became known as equipotentiality hypothesis (Lashley, 1929; Hunter, 1930). More than 100 years earlier, Flourens (1824) had already proposed a similar theory as Lashley’s, stating that the brain as a whole stored memories, which was a response to phrenology. Phrenology was an early attempt to relate different brain areas to particular functions, although being popularized by dubious measurement of skull shapes (Gall and Spurzheim, 1810; Gluck et al., 2016). Neither of the theories proved to hold: While not every part of the brain is equally important for memory, indeed, multiple brain areas are active during the storing and recall of memories. Fig. 1.6 shows the brain structures that are considered most important for memory.

Being a part of the *medial temporal lobe*, the *hippocampus* is essential for declarative memory. Specifically, it seems to add spatial and temporal context information to episodes and is thought to transfer short-term into long-term memory, as well as to provide access to long-term memory (Squire and Zola-Morgan, 1991; Dudai, 2004; Morris, 2006; also cf. sections 1.4.3 and 1.4.4). Discovered by O’Keefe and Dostrovsky (1971) in the rat hippocampus, place cells are a prominent sort of neurons which fire only if the animal is in a certain location. Complementary to place cells, there are speed cells, head direction cells, and grid cells, serving to measure running speed, head direction, and distances crossed. All of these cell types have also been found in humans (Rowland et al., 2016). It has recently been demonstrated, moreover, that neurons in the human hippocampus are activated during free recall (Gelbard-Sagiv et al., 2008; also cf. section 1.4.7). Furthermore, the hippocampus seems to encode temporal information by “time cells” and to disambiguate overlapping memories (MacDonald et al., 2011; Eichenbaum, 2013). The *neocortex* is important for storing long-term memories and for working memory, among many other tasks that are not directly related to memory. The prefrontal cortex, specifically, has been related to working memory (cf. section 1.4.6), by which it controls the effort to remember or to forget a memory, and by which it also aids the hippocampus in processing contextual information (Gluck et al., 2016). The *amygdala* interacts with the hippocampus and the neocortex to form and strengthen memories with emotional context (Squire and Zola-Morgan, 1991; Squire, 2004; Gluck et al., 2016). Other parts of the medial temporal lobe (parahippocampal cortex, the peri- and entorhinal cortex) have also been related to declarative memory (Squire and Zola-Morgan, 1991), as well as the diencephalon, the basal forebrain, and the fornix (Squire, 2004; Gluck et al., 2016). The diencephalon connects hippocampus and neocortex, while the basal forebrain provides GABA and acetylcholine neuromodulation to the hippocampus via the fornix. Further memory-related areas include the striatum (a part of the basal ganglia) and the cerebellum, which are mostly associated with implicit memory (Squire, 2004). Habit formation, for example, seems to be greatly complicated by lesions of the striatum (Bear et al., 2016).

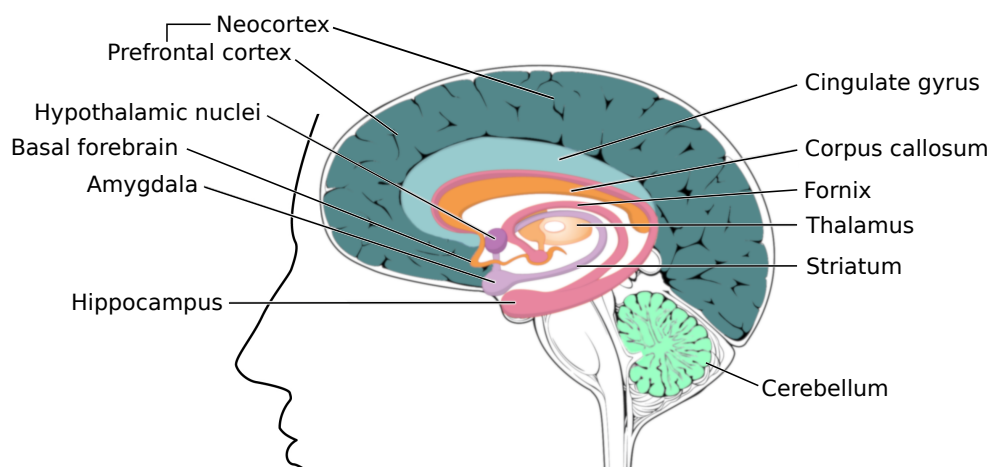


Figure 1.6: **Illustration of brain regions important for different forms of memory** (see text). The hippocampus is surrounded by the amygdala and other medial temporal lobe structures (not shown) and projects via the fornix to thalamus and hypothalamic nuclei (parts of the diencephalon) which project to the neocortex. In turn, neocortex and thalamus send signals via the cingulate gyrus toward the hippocampus. The thalamus is also the main entrance for sensory signals into higher brain regions. The corpus callosum connects the left and right hemispheres of the cerebrum. The amygdala receives input from thalamus, neocortex, and hippocampus, and also projects to the hippocampus. The basal forebrain and cerebellum exhibit diverse connectivity with various brain areas. The striatum (part of the basal ganglia) receives projections from the neocortex and projects back to it via the thalamus. Parts of the basal ganglia lie beneath the thalamus and are not visible here. The description of connectivity provided here demonstrates the complexity of brain systems involved in memory and is not exhaustive (also see [Bear et al., 2016](#); [Gluck et al., 2016](#)). Figure adapted from [Betts et al. \(2013\)](#), licensed under CC BY 4.0.

1.4.3 Declarative memory

As defined above, declarative memory refers to memory that can be communicated flexibly, for instance, verbally. Although the term “declarative” (or “explicit”) is usually used in the context of long-term memory, contents of short-term and working memory can also be classified as declarative or non-declarative ([Oberauer, 2013](#); [Gluck et al., 2016](#); [Bear et al., 2016](#); [Fusi and Wang, 2016](#)).

The two major subtypes of declarative memory are *semantic memory* and *episodic memory*. It is under debate whether semantic or episodic content has to be encoded first for the other to draw from it, or if both are irreducibly interdependent ([Gluck et al., 2016](#)). It has recently been proposed that selective storing of semantic and episodic information could be controlled by dopaminergic input from different brain regions, namely, the ventral tegmental area and the locus coeruleus ([Duszkiewicz et al., 2019](#)). In chapter 3, we will consider a generic paradigm that may relate to this.

Schema theory describes the integration of multiple memories into an overarching web. Schemas are related to semantic memory, although they are also linked to implicit memory processes. We will discuss schemas in a further subsection below.

Episodic memory

Episodic memory contains information about the semantic, temporal and spatial aspects of events: “what” happened, “where” it happened, and “when” ([Tulving, 1972](#)). It is under debate whether non-human animals have an episodic memory, and while there is evidence that they might have an

episodic-like memory, they probably cannot communicate it directly (Gluck et al., 2016). The high capability of storing and deliberately retrieving episodic information, related to the capability of “mental time travel”, might be one of the key factors that distinguish humans from other animals (Tulving, 1985; Klein, 2013).

An important and famous subject for studies of long-term memory was Henry Molaison (until his death abbreviated H.M.), whose hippocampus was bilaterally removed to relieve him from severe epileptic seizures (Scoville and Milner, 1957; Gluck et al., 2016). After the surgery, he suffered from anterograde amnesia affecting the encoding of new episodic information, while he could still encode new semantic information and learn new motor skills. He also suffered from retrograde amnesia, although mostly “only” three years before the surgery were affected. His case contributed significantly to the theory of the transfer of newly formed episodic memories from the hippocampus to the neocortex (cf. “Systems consolidation” in section 1.4.4).

Another patient called E.P. suffered from bilateral damage to large parts of the medial temporal lobe caused by a form of viral encephalitis. He showed similar memory impairment as H.M., but in principle lost any processed information within 15 minutes. Unlike H.M., E.P. was unable to form new semantic memories. Thus, semantic information seems to be encoded primarily in structures of the medial temporal lobe other than the hippocampus (Tramoni et al., 2011; Insausti et al., 2013; Gluck et al., 2016).

Further evidence for the role of the hippocampus in episodic memory is given by its capability to disambiguate overlapping memories and by the existence of cells specialized to code for time (“time cells”; MacDonald et al., 2011; Eichenbaum, 2013).

Semantic memory

Semantic memory is the memory of facts. In the terms used above, semantic memory contains “what” information, but not “when” and “where” it was acquired. Therefore, semantic memory may seem a simpler form of memory as compared to episodic memory. This is, however, hardly the case because semantic memory is shaped by experiencing similar content in different contexts and conditions in an iterative specific-to-generic processing – the information to form a semantic memory is extracted from a variety of different episodes (Tsien et al., 2013; Coutanche et al., 2020). On the other hand, as mentioned before, certain medial temporal lobe damage can leave consolidation of semantic information relatively intact, whereas consolidation of episodic information is prevented (Tramoni et al., 2011). This indicates certain independence of semantic memory from episodic memory. Nevertheless, the medial temporal lobe, and especially the hippocampus, have been shown to be important for the formation of new semantic memories and the transfer of these to the neocortex (Dudai et al., 2015; Vogel et al., 2018; Gastaldi et al., 2021). The hippocampus seems, furthermore, to compress unimportant memory representations and differentiate important ones before the related information is transferred to long-term memory in neocortex and cerebellum (Gluck and Myers, 1993; Gluck et al., 2016).

Semantic memories correspond to mental *concepts*, which can be defined as internal representations that cause the mental state of “having an idea of something” (cf. McClelland et al., 1995; Goldstone et al., 2012). A key property of concepts is generality, meaning that slight variations in appearance of a stimulus that has been linked to a concept will relate that stimulus still to the same concept. Concepts are distinct from categories, which describe a class or division of entities in the world (Goldstone

et al., 2012; Gluck et al., 2016). However, the relationship between concepts and categories is disputed (Goldstone et al., 2012), and a strict differentiation seems questionable, since we perceive all things in the world through our minds.

Humans can, due to their capability of concept generalization, learn to classify given stimuli into categories. Certain aspects of this can be described by classical conditioning (cf. section 1.4.1). The formation of concepts also includes the counterpart of generalization, discrimination, whose key property is the different response to different stimuli (Gluck et al., 2016). Similarity between concepts leads to schema formation (see next subsection), but can also lead to interference (see section 1.4.8). Concepts are often based on a prototype of a category, which serves to make predictions about the world via inductive inference. Nevertheless, in many cases, these predictions fail in the sense that they relate entities to categories they do not fit in, and can thereby have a negative impact on individuals as well as on whole societies (as, for instance, in the case of racism; Goldstone et al., 2012; Gluck et al., 2016).

Not only humans but also other animals can achieve remarkable results through their capability to form concepts. Pigeons, for example, can be trained to discriminate between different styles (that means, different categories) of music (Porter and Neuringer, 1984) and painting (Watanabe et al., 1995), although a changed perspective can in some cases easily disrupt the recognition.

Models of distributed representations such as cell assemblies (cf. section 1.3.5) have proven useful to describe similarity and generalization of concepts, while simpler neuropsychological models require sufficiently distinct stimuli (Gluck et al., 2016). Therefore, it seems likely that functional interactions between cell assemblies, which include association, dissociation, and sequences of activation (Herpich and Tetzlaff, 2019), underlie the dynamics of concepts. Furthermore, concept-related neural computation seems to be facilitated by topographically organized memory representations (Pomi et al., 2018). In chapter 4, we will consider the interaction of cell assemblies in different organizational paradigms to examine how our model can account for neuropsychological effects that involve multiple memory representations or concepts.

Schemas

The overarching web of long-term memories in the brain is described by so-called schemas, which were first proposed by Piaget (1923) and Bartlett (1932). Schemas enable the integration of new entities and the update of old ones by so-called assimilation and accommodation, or cooperation and competition (Piaget, 1923; Preston and Eichenbaum, 2013; Cooper, 2016). Schemas combine features of various semantic knowledge structures and share some similarity with networks of concepts (Ghosh and Gilboa, 2014; cf. previous subsection), however, schemas are typically intended to functionally describe cognitive aspects in a phenomenological manner, complementary to the complex molecular structure of the brain (Arbib, 2003; Cooper, 2016). Nevertheless, schema theory has received substantial support from experimental findings (Tse et al., 2007; Preston and Eichenbaum, 2013). Consistent with findings on semantic memory (cf. previous subsection), the role of the hippocampus in the integration of new memories into an existing schema does not seem critical (Vogel et al., 2018).

Cooper (2016) defines schemas as “abstractions over similar entities” and names as such “perceptual, linguistic, event, action, or motor situations”. Thus, schemas are also linked to non-declarative content. Schemas are, moreover, considered to be organized hierarchically with higher-level schemas (e.g., event and action schemas) based on simpler subschemas (e.g., perceptual and motor schemas).

Therefore, schema theory can be applied to both low- and high-level cognitive functions (Ghosh and Gilboa, 2014; Cooper, 2016). Perceptual schemas serve, for instance, to provide an internal model of the world that enables action planning and execution in combination with other schema types, and is already found in amphibians (Cooper, 2016). For performing routine behavior, schemas are usually accessed unconsciously. Through working memory, however, such procedures can also be consciously modulated (cf. section 1.4.6; Cooper, 2016).

One important function of schemas is to enable the enhanced recall of similar memories (Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014). The opposite, interference, will be discussed in section 1.4.8. In chapter 4 of this thesis, we consider the interaction of multiple memory representations, which can be thought of as a schema. We investigate in that context also the occurrence of enhancement and interference of overlapping/similar memory representations.

1.4.4 Long-term memory and consolidation

Long-term memory denotes the storage of memories for at least several hours to days (McGaugh, 2000; Dudai, 2004; also cf. Nadel and Hardt, 2011), while memories can obviously last up to a lifetime (Fig. 1.7). Memories on such maximum timescales are sometimes considered a separate type, denominated “long-lasting memories” (McGaugh, 2000).

Following acquisition, transient (short-term) memories can be transferred to long-term memory via *consolidation*. Consolidation denotes the stabilization of memory-related plastic changes such that they can last for hours, days, or even years (cf. McGaugh, 2000; Abraham et al., 2002; Dudai, 2004). According to Dudai (2004), the term “consolidation” dates back to Müller and Pilzecker (1900), while early ideas of time-driven consolidation were already conveyed by Quintilian in the 1st century CE. The definition by Dudai (2004) states that memory consolidation

“refers to the progressive postacquisition stabilization of long-term memory, as well as to the memory phase(s) during which such presumed stabilization takes place”.

Put differently, consolidation denotes the maintenance of a memory beyond the usual timescales of short-term memory (cf. section 1.4.5). More methodology-related, Roediger (2008) defines consolidation as

“less interruption from application of some amnesic agent with time since a learning experience rather than [...] any absolute increase in performance over time”,

where the last part indicates that improvement of recall performance is not part of the definition of memory consolidation, although there are cases in which consolidation can improve memory (Roediger, 2008), which we will investigate in chapter 2.

Usually, one distinguishes two main forms of consolidation (Dudai, 2004; Nadel and Hardt, 2011; Dudai et al., 2015): *synaptic consolidation* and *systems consolidation*.

Synaptic consolidation

Synaptic consolidation denotes the stabilization of a memory trace within the first minutes to hours after acquisition and is also called *initial consolidation*. “Initial” refers to the fact that it occurs before

systems consolidation (Okuda et al., 2020; see below). Synaptic consolidation is thought to correspond to the in-place stabilization of synaptic weight changes, that is, stabilization of changes that have been evoked by synaptic plasticity. Another term for synaptic consolidation is *cellular consolidation*, attributed to the fact that it presumably involves neuronal processes beyond those of individual synapses (Dudai, 2004; Redondo and Morris, 2011; cf. section 1.3.4).

The idea of synaptic consolidation has emerged from studies showing that the application of protein synthesis inhibitors prevents the formation of long-term memories, while it also prevents the maintenance of synaptic plasticity, which was suggested as early as the 1960s (Dudai, 2004). The most acknowledged mechanism to describe synaptic consolidation is synaptic tagging and capture (STC; Frey and Morris, 1997; Dudai, 2004; Redondo and Morris, 2011; Okuda et al., 2020), which transfers long-term synaptic plasticity from an early to a late phase, and has been discussed in section 1.3.4. Strong *experimental evidence* for this hypothesis (that STC describes synaptic consolidation) has been presented recently (Pastalkova et al., 2006; Moncada and Viola, 2007; Wang et al., 2010; Takeuchi et al., 2016; cf. “Behavioral Tagging” in section 1.4.8). Investigating the functional impact of STC on memory dynamics *theoretically* in a spiking neural network model is the main goal of this thesis. Other mechanisms of synaptic consolidation may, however, exist, for example the selective strengthening and weakening of synapses by means of sleep-like activity and homeostatic plasticity (Tetzlaff et al., 2013). Also note that synaptic consolidation can comprise a “systems component”, which is the case if neuromodulation from other brain regions is involved (Nader et al., 2000b; Lisman et al., 2011).

Systems consolidation

As opposed to synaptic/initial consolidation, systems consolidation transfers memories across brain regions, most importantly, from the hippocampus to the neocortex, with the purpose of further processing (e.g., extracting semantic information and integrating it in schemas, cf. section 1.4.3) and their stabilization (Nadel and Hardt, 2011; Dudai et al., 2015). This type of consolidation seems to require the presence of activity patterns like shape-wave/ripples, spindles or slow oscillations (Mizuseki and Buzsáki, 2013; Mizuseki and Miyawaki, 2017; Dudai et al., 2015) and is thereby related to slow-wave sleep and rest (Rasch and Born, 2013; Mizuseki and Miyawaki, 2017). Complementary to this, rapid-eye movement (REM) sleep may cause homeostatic weakening of memory traces (Crick and Mitchison, 1983; Mizuseki and Miyawaki, 2017; Li et al., 2017).

From studies of memory, or rather studies of memory deficits (see section 1.4.3), it has become evident that memories are initially hippocampus-dependent but later become relatively independent of the hippocampus. Following the *standard model of systems consolidation*, the hippocampus may serve to link long-term memories in the neocortex (McGaugh, 2000; Dudai, 2004). According to this theory, memories in the neocortex become independent of the hippocampus once they exhibit strong enough links among themselves. Opposed to this is the *multiple trace theory*, which claims that memories not necessarily become independent of the hippocampus. Instead, new encoding take place every time that a memory is reactivated (Nadel and Moscovitch, 1997; Nadel and Land, 2000; Nadel and Hardt, 2011) such that frequently reactivated episodic memories be represented by many traces, and the “gist” of them give rise to semantic memories. This can, better than the standard model, account for some types of amnesia, and furthermore explains findings of reconsolidation (cf. section 1.4.8). However, neuroimaging studies on hippocampal and neocortical activity have yielded ambiguous results, and

it remains to be shown if any of the two theories will hold (Gluck et al., 2016).

Systems consolidation sets in later and acts on much longer timescales as compared to synaptic consolidation. Therefore, it is not subject to the investigations of this thesis. However, in the future, our model may be extended by sleep-like activity patterns (Tetzlaff et al., 2013; Jahnke et al., 2015) to investigate possible mechanisms of systems consolidation on memory improvement (cf. chapter 2) and on the interaction between memory representations (cf. chapter 4).

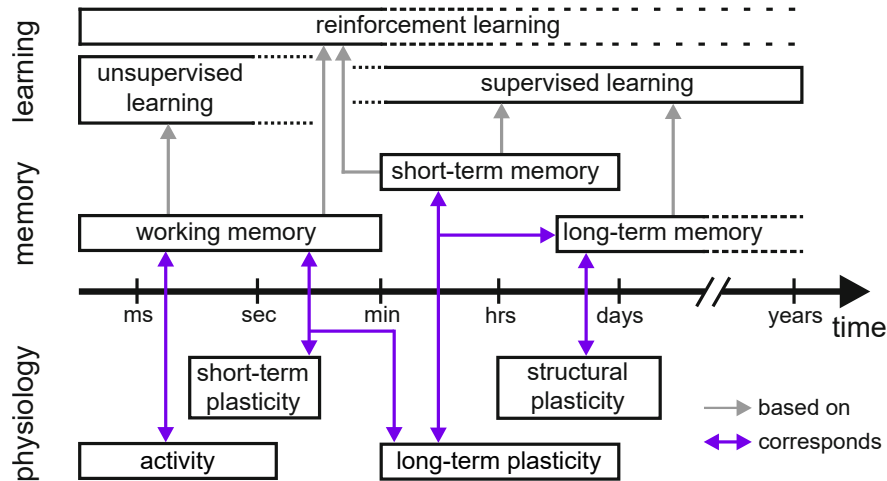


Figure 1.7: **Timescales of learning and memory and their physiological correlates.** The types of learning and memory are described in sections 1.4.1–1.4.6. The different types of plasticity are described in section 1.3.4, and activity related to working memory is treated in section 1.4.6. Further physiological underpinnings of learning and memory are described at various points throughout this whole introduction chapter. Figure adapted from Tetzlaff et al., 2012, licensed under CC BY 2.0.

1.4.5 Short-term memory

Short-term memory is often used as a synonym for the storing (as opposed to the processing) site of working memory (Cowan, 2001; Siegel et al., 2009; Gluck et al., 2016). However, short-term memory is also often used with a different, typically broader meaning (Dudai, 2004; Moncada and Viola, 2007; Tetzlaff et al., 2012; Bear et al., 2016). Contrary to working memory, short-term memory in this notion does not require repetition nor is it strictly limited in storage capacity (Bear et al., 2016). As opposed to long-term memory, it is more susceptible to disruption by behavioral distractors, lesions of the brain, drugs, and seizures (Dudai, 2004). For the consolidation of short-term memories into long-term memory, synaptic consolidation plays an important role (cf. section 1.4.4, and “Behavioral Tagging” in section 1.4.8).

In this thesis, we refer to memory that persists independently of neuronal activity for time spans from several seconds up to several hours as short-term memory (Fig. 1.7). Thereby, the timescale of short-term memory corresponds to the timescale of changes evoked by early-phase long-term synaptic plasticity (cf. section 1.3.4). It should be noted that by this definition, short-term memory acts on a much slower timescale than short-term plasticity, thus, the term “short-term” has to be understood in a relative sense.

Considering the findings on patients' lesions of the medial temporal lobe (cf. section 1.4.3), which still allowed the retention of new memories for up to 15 minutes, and studies on macaque monkeys with similar lesions, which showed memory impairment after a few seconds, it remains an open question in which regions of the brain short-term memory is located (Bear et al., 2016).

1.4.6 Working memory

Working memory describes the capability of retaining and manipulating information on timescales of seconds (Baddeley and Hitch, 1974; Barak and Tsodyks, 2014). Furthermore, working memory has been shown to be important for the *encoding and retrieval of longer-lasting forms of memory* (O'Reilly et al., 1999; Ranganath et al., 2005; Gluck et al., 2016). An important aspect of working memory is that it not only stores but also processes information (unlike short- and long-term memory). Methods such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) have revealed that wide areas of the human brain are active during working memory tasks, while the prefrontal cortex seems to be of particular importance, and different frequencies of oscillatory activity serve as carriers for content (Tesche and Karhu, 2000; Siegel et al., 2009; Constantinidis and Klingberg, 2016). Note that the notion of working memory which we discuss here is similar to that of "short-term memory" in the Atkinson-Shiffrin model, requiring rehearsal to be maintained (Bear et al., 2016). Working memory has been assumed to be the basis for higher-order cognitive functions related to attention, such as reasoning, planning, task switching, selection of stimuli and response control (James, 1890; Gluck et al., 2016).

The average capacity of working memory in humans has been postulated to be around some "magical", fixed number of items, first estimated around 5 to 9 by Miller (1956), later around 4 by Cowan (2001). Note that this number is related to the probability of recall in immediate free recall experiments (for example: Glanzer and Cunitz, 1966; Howard and Kahana, 1999). Nevertheless, there is strong evidence that the capacity of working memory can be expanded by training (Constantinidis and Klingberg, 2016).

Related to working memory is the so-called delay period, which describes a state of retained information, or ongoing activation in a group of neurons, similar to an attractor state (Miyashita and Chang, 1988; Amit and Brunel, 1997b; Mongillo et al., 2008). Such a delay was already described by Lorente de Nó (1934). From this arises the classical view that declarative memories are, upon activation, loaded into working memory, where they persist as long as the neuronal spiking activity is maintained, that is, as long as the putative attractor state persists (Amit and Brunel, 1997b; Barak and Tsodyks, 2014). Accordingly, classical models of working memory either constitute phenomenological descriptions, such as the seminal model by Baddeley and Hitch (1974), or are based on attractor neural networks (Bradski et al., 1994; Amit and Brunel, 1997b). This paradigm was maintained for a long time since there is no evidence that lasting structural changes are involved in working memory (Mongillo et al., 2008; Barak and Tsodyks, 2014).

Nevertheless, the attractor hypothesis has more recently been questioned due to findings that neuronal activity during the delay period is not very high or even absent (Naya et al., 1996; Shafi et al., 2007; Barbieri and Brunel, 2008; Masse et al., 2020). In a much recognized study, moreover, Mongillo et al. (2008) proposed that synaptic facilitation conveyed by presynaptic calcium concentrations should complement attractor-based working memory. Reservoir computation in recurrent neural networks

constitutes another alternative to the classical attractor implementation, however, this alone also seems insufficient to account for all properties of working memory (Barak and Tsodyks, 2014; Nachstedt and Tetzlaff, 2017; Masse et al., 2020).

1.4.7 Recall

Recall refers to the retrieval or recollection of information stored in memory. Therefore, recall is critical for the use of any memory, and it is important to understand the mechanisms that determine the performance of recall. Memory recall in the brain is very different from basic memory operations in common digital computers. While these assign to every chunk of information an address in random-access memory (RAM) by which it is retrievable, memories in the brain are thought to be retrieved by completing a presented part of the pattern of their original content, encoded by neuronal activity. This process is called pattern completion, and this type of memory is referred to as associative or content-addressable memory (Hopfield, 1982; Dayan and Abbott, 2001). The active group of neurons is a cell assembly, which concept has been introduced in section 1.3.5. Such associative recall mechanisms and the preceding formation of related memory representations were already proposed by Hebb (1949) and have been suggested, for example, for the hippocampus (Treves and Rolls, 1992; Martin et al., 2000).

Meanwhile, experiments with human subjects have specifically linked recall to the reactivation of neural activity patterns similar to the patterns observed during initial learning (Ritchey et al., 2013). Furthermore, artificially induced learning and recall of fear memory via stimulation of a population of neurons in the hippocampus of mice has yielded direct experimental evidence for the representation of memories by cell assemblies (Liu et al., 2012b). In addition to spatial patterns, temporal structures can also be subject to recall. This may be implemented either by the spike timing of the neurons in the cell assembly or by subsequent activation of multiple cell assemblies (Harris, 2005; Eichenbaum, 2013). In both cases (spatial patterns and temporal structures), the performance of recalling a memory trace can be better or worse depending on how well the activity resembles the original activity during encoding. This performance is the main measure which we assess in chapter 2 for spatial patterns and in chapter 3 for both spatial and temporal patterns.

Considering psychology, there are cases in which false memories are recalled, just as there are cases where all attempts to recall seem in vain (Gluck et al., 2016). Interestingly, it has been shown that such issues can sometimes be tackled by hypnosis, which is of benefit to psychotherapy and eyewitness testimony of major crimes, although its suggestive influence is considered controversial (Nash and Barnier, 2008; Doyle et al., 2020). Recall performance has been shown to depend on different conditions. The so-called encoding specificity or transfer-appropriate principle frames this by stating that recall works best in the same context in which the memory was formed, which might be related to overlapping cell assemblies at the network level (Tulving and Thomson, 1973; Preston and Eichenbaum, 2013; Gluck et al., 2016).

In neuropsychological experiments, one typically distinguishes between free recall, cued recall, and recognition. Another particular form of recall is serial recall, in which a series of elements has to be recalled in a specific order (Brown et al., 2007). There are indications that serial recall and free recall might – in some sense – also be forms of cued recall. This would be the case if the elements to be

recalled implicitly act as cues for others, while no external cue is presented (cf. Kropff and Treves, 2006; Recanatesi et al., 2017; Fischer et al., 2020).

Recalling memories in free recall is usually much more difficult and takes longer as compared to recall with a provided cue or compared to the recognition of a provided known item (Gluck et al., 2016). This relates to so-called generation-recognition models which claim that to recall, different alternatives might first be generated or *retrieved* before one of them be picked by *recognition* (cf. Bahrnick, 1970; Basden et al., 1993). However, there are cases where recognition performs worse than other forms of recall (Tulving and Thomson, 1973). Due to these differential characteristics, recall and recognition are often considered separate types of *retrieval* processes (Tulving and Thomson, 1973; Basden et al., 1993).

Eventually, recall can render a memory trace unstable and thereby cause its disruption or enhancement. The former is linked to interference and impaired reconsolidation, while the latter is linked to memory improvement (cf. section 1.4.8 on these phenomena). In this thesis, we investigate neural correlates of cued recall (chapters 2 and 3) as well as free recall (chapter 4). In these paradigms, we further examine memory improvement and priming by means of recall stimulation that drives cell assemblies back into a dynamic state. In the final chapter, we also discuss a potential future model of reconsolidation (chapter 5).

1.4.8 Phenomena important to this thesis

Many neuropsychological phenomena are related to learning and memory. In the following, based on the general structures and functions introduced before, we will consider specific phenomena that are modeled in this thesis or that are important to the investigations of this thesis.

Forgetting and improvement

Pioneering work on the retention of memory was done by Ebbinghaus (1885), who performed memory experiments with nonsense syllables on himself and thereby showed the decay of memory over time, often referred to as “forgetting curves”. Further seminal work was presented by Thorndike (1913), who proposed the “law of disuse”, which stated that disused habits decay over time. Thus, memory disuse was traditionally considered to be the primary cause of forgetting. Such hypotheses of merely time-dependent decay have, however, been questioned by Jenkins and Dallenbach (1924) and McGeoch (1932) who argued that interference be the cause for the the loss of memories (also cf. next subsection). Medical conditions that impair the retention of previous and future memories are referred to as retrograde and anterograde amnesia, respectively (also see section 1.4.3).

Despite the long-standing paradigm of forgetting over time, there is also evidence for memory improvement over time, sometimes called hypermnnesia. Phenomena related to memory improvement include spaced learning (Smolen et al., 2016), the levels-of-processing effect (Roediger et al., 1982; Craik and Tulving, 1975) and the testing effect (Roediger and Butler, 2011). An early study by Ballard (1913) studied the retention of poetry, and although there have been issues to reproduce the study, it has had seminal impact and introduced the notions of reminiscence and oblivescence (the regain and loss of items across recall attempts; cf. Erdelyi, 2010). Later studies have shown that recalling learned images for multiple times leads to a gradual increase of recall quality (Payne, 1987; Erdelyi, 2010; Wallner and

Bäuml, 2018), while improvement effects were also observed with lists of words and nonsense syllables (Roediger et al., 1982; Otani and Hodge, 1991). The reduction of item losses across recall attempts (reminiscence), rather than the gain in the number of remembered items across recall attempts, seems to cause such improvement (Erdelyi, 2010; Wallner and Bäuml, 2018). Although there are indications for improvement by the mere passage of time (Erdelyi, 2010), further experimental investigation is necessary to fully prove this.

Furthermore, positive effects of memory disuse are described by the theory of “desirable disuse”, which proposes a long-lasting *storage strength* and a transient *retrieval strength* (Skinner, 1938; Hull, 1943; Estes, 1955; Bjork and Bjork, 1992; Karpicke et al., 2014; Gluck et al., 2016). In this theory, learning causes both storage and retrieval strength to increase. While storage strength undergoes piecemeal increase by repetitive learning, retrieval strength first increases fast but then decreases. The core assertion is that learning will be most effective if the retrieval strength is low, for instance, some time after the first learning phase when the learning-induced retrieval strength has decreased and the storage strength is still low. At this point, attempts to retrieve will take some time but result in a strong increase in storage strength. On the contrary, fast and “easy” retrieval at high retrieval strength will cause little increase in storage strength.

Effects of improved retrieval, as they were discussed here, may depend on the interplay between early- and late-phase plasticity. Using abstract models of multi-phase plasticity based on binary (Elliott and Lagogiannis, 2012) and discrete-valued synaptic weight (Elliott, 2016), Elliott and colleagues predicted an improvement in memory strength with the passage of time, though not linked to specific biological mechanisms. In addition, Papper et al. (2011) employed STC mechanisms to show that these can trigger improvement in memory lifetime. However, they did not find improvement in memory strength. In chapter 2 of this thesis, we will show how improvement in memory strength arises through biologically realistic STC. In chapters 3 and 4, further aspects of this improvement through STC will be considered (a special form being priming, which we investigate in chapter 4).

Interference

Along with the similarity of memories (cf. subsection “Schemas” in section 1.4.3) comes interference. Unlike improved recall as one possible consequence of memory similarity, interference describes competition for retrieval which leads to the disruption of a memory (Wickelgren, 1965; Anderson, 1981; Anderson and Neely, 1996). It was probably first investigated by Bergström (1893) and Müller and Pilzecker (1900). Later, others suggested that interference might be the cause of forgetting rather than a simple decay of the memory trace over time (Jenkins and Dallenbach, 1924; McGeoch, 1932; see previous subsection). There are two types of interference: retroactive and proactive interference (Lohns et al., 2015; Gluck et al., 2016). Retroactive interference refers to the disruption of old memories by newly acquired ones, whereas proactive interference refers to the disruption of new learning by old memories. Proactive interference was shown to be reduced by practicing retrieval (Karpicke et al., 2014). On the other hand, Walker et al. (2003) found that the learning of a new memory leads to destabilization of an old memory if this has been recalled immediately before the learning. This paradoxical behavior has been explained mechanistically by a combination of synaptic plasticity and synaptic scaling (Tetzlaff et al., 2013).

Catastrophic interference/forgetting is a term from neural network theory and refers to the situation

where so many memory representations are stored in a network that no information can be retrieved anymore (Marr, 1970; Amit, 1989; McClelland et al., 1995), relating to the plasticity-stability dilemma that has been discussed in section 1.3.5. This fate has been shown to be countered by two- or multi-phase plasticity (Gardner-Medwin, 1989; Fusi et al., 2005; Päpper et al., 2011). By employing two-phase plasticity, the model presented in this thesis enables to avoid catastrophic interference when learning multiple different memory representations. In chapter 4, we specifically investigate interference in different paradigms of three memory representations.

Serial-position effect in free recall

Free recall is used as a memory test in which subjects are required to reproduce information from memory, for example, a sequence of words or pictures, without receiving external cues for the individual items. It has to be distinguished from other types of memory retrieval such as cued recall (with serial recall being a special case) and recognition (Katkov et al., 2014; Gluck et al., 2016; also see section 1.4.7).

The frequency of how often learned items are recalled in a free recall experiment depends on the order of acquisition. In this regard, typically, free recall experiments exhibit the following three characteristics (Ebbinghaus, 1885; Murdock, 1962; Bjork and Whitten, 1974; Howard and Kahana, 1999; Sederberg et al., 2010):

- increased probability of recalling the items learned first (*primacy*),
- increased probability of recalling neighboring items (*contiguity*),
- increased probability of recalling the most recently learned items (*recency*).

Primacy is commonly thought to be caused by implicit extended rehearsal, in which the recency effect could play an indirect role (Brodie and Murdock, 1977; Davelaar et al., 2005; Brown et al., 2007). Recency is usually much more pronounced as compared to primacy (Murdock, 1962; Bjork and Whitten, 1974). Together, these effects are termed *serial-position effect*.

In neuropsychological experiments, free recall has traditionally been used as a test for working memory (Murdock, 1962; Glanzer and Cunitz, 1966; Bjork and Whitten, 1974; also cf. section 1.4.6). Proposing associations of working memory items, the phenomenological “search of associative memory” model (Raaijmakers and Shiffrin, 1980, 1981) has provided very good matching with free recall data. However, due to its many free parameters, it cannot serve as a functional explanation of the underlying mechanisms (Recanatesi et al., 2015). The “temporal context model” (Howard and Kahana, 2002), also a phenomenological model, has further improved the description of free recall data by accounting for contiguity (the higher probability of transition between neighboring items).

Nevertheless, serial-position effects in free recall were shown not only to occur in working memory but also in longer-lasting forms of memory (Bjork and Whitten, 1974; Greene, 1986; Davelaar et al., 2005; Brown et al., 2007; de Almeida Valverde Zanini et al., 2012). Free recall has, meanwhile, often been linked to attractor neural networks (Bradski et al., 1994; Gelbard-Sagiv et al., 2008; Romani et al., 2013; Katkov et al., 2015; Recanatesi et al., 2015). Based on this, theoretical studies have also shown that rather than working-memory associations as in previous models (Raaijmakers and Shiffrin, 1981; Howard and Kahana, 2002), longer-lasting associations between items seem responsible for free recall

(Romani et al., 2013; Katkov et al., 2015). These studies indicate that associations are given by overlaps between cell assemblies (cf. section 1.3.5). We will draw from these findings in chapter 4 of this thesis, where we seek to explain recency effects with long-term memory representations, both in transient and quasi-attractor dynamical regimes.

While Romani et al. (2013) and Katkov et al. (2015) developed phenomenological models, Recanatesi et al. (2015) provided a model describing similar associations, however, reproducing free recall data with a more mechanistic approach. In that study, oscillating inhibition was used to drive the alternating activation of overlapping attractor memory representations in networks of rate neurons. The authors have argued that such oscillating inhibition be necessary for controlled transitions between attractor states via modulation of the oscillation amplitude (Romani et al., 2013; Recanatesi et al., 2015, 2017). Switching between attractors driven by oscillating inhibition is biologically plausible because of the ubiquitous presence of oscillatory dynamics in the hippocampus and in other brain regions (Tesche and Karhu, 2000; Kahana, 2006; Buzsáki, 2006; Düzel et al., 2010). Nevertheless, there could be alternative mechanisms such as inhibitory plasticity, neuronal adaptation, or synaptic depression (Buzsáki, 2010; Recanatesi et al., 2015). Treves and colleagues, for instance, have shown that the switching between attractor states in a network (which they call “latching”) is enabled by neuronal adaptation combined with overlapping assemblies, exhibiting diverse transition dynamics ranging from deterministic to random (Treves, 2005; Kropff and Treves, 2006).

Priming

Priming denotes a large set of psychological phenomena, all having in common that they facilitate (or attenuate) a later action, the occurrence of an event, or the processing of a stimulus by some kind of pre-activation (cf. Bermeitinger, 2015; also see *ibid.* for an overview of definitions). While this facilitating influence is usually considered temporary, priming phenomena have been shown to occur on a large variety of timescales from seconds to years (Janiszewski and Wyer, 2014; Was et al., 2019).

In psychological experiments, priming is often studied by subliminal stimuli (Bermeitinger, 2015; Elgendi et al., 2018). Although it is therefore often considered to produce a form of non-declarative or implicit memory (cf. Schacter, 1992; Squire, 2004; Elgendi et al., 2018), priming is not strictly independent of conscious awareness (Janiszewski and Wyer, 2014; Bermeitinger, 2015). As prominently proposed by James (1890) and Lashley (1951), priming-like phenomena have been assumed to underlie the formation of word sequences in the context of language (cf. Bermeitinger, 2015).

To describe the various forms of priming phenomena that have been identified so far, different classifications have emerged. Janiszewski and Wyer (2014) refer to consumer-related phenomena and roughly distinguish between content priming and process priming, the former making content more accessible and the latter making processes that operate on content more accessible. The authors further differentiate content priming into semantic priming, goal priming (motivational; a subclass of semantic priming), affective priming (emotion-related), and behavioral priming (mostly related to social behavior), each of which could happen as direct or indirect priming, depending on the relationship (matching) to the stimulus. Following this terminology, the priming of memory representations that we investigate in chapter 4 of this thesis can be considered a form of *direct priming*. As opposed to this, for instance, *indirect semantic priming* would refer to the enhanced accessibility of a concept by the pre-activation of another (e.g., in the way that the word “chicken” would prime the word “duck”;

Herrmann et al., 1993; Becker et al., 1997; Lerner and Shriki, 2014; Bermeitinger, 2015). Note that other literature considers semantic priming generally as indirect and thus distinct from direct (or “repetition”) priming in a narrow sense, which would rely on perception without the need for semantic processing (Schacter, 1992; Was et al., 2019).

Bermeitinger (2015) emphasizes that there are different *levels* of priming, ranging from the most concise experimental priming protocols up to “macro” phenomena such as altered way of thinking. Slightly deviating from the classification described above, she mainly distinguishes between the following types at the “macro” level: perceptual priming, semantic priming, categorization/construal priming (altering how another person is perceived), behavioral priming, and goal/motivational priming. At the “micro” level, Bermeitinger distinguishes between semantic priming and response priming (no semantic processing required, mostly motor-related; also cf. Elgendi et al., 2018). She further differentiates semantic priming into *actual* semantic priming and associative priming, which would be based on the mere co-occurrence of priming stimulus and target, rather than an actual overlap in features. As a subclass for response priming, she names affective priming.

Yet another slightly different classification can be found in Elgendi et al., 2018. For all mentioned paradigms, there also exists the inverse, *negative* priming, which leads to the attenuation of an outcome instead of facilitating it (Tipper, 1985; Bermeitinger, 2015). Although there are many models for the different forms of priming (Janiszewski and Wyer, 2014; Bermeitinger, 2015; Was et al., 2019), the host of priming phenomena remain unexplained in terms of biologically realistic models. Particularly for negative priming, issues are being faced to find a conclusive model (Mayr and Buchner, 2007; Ihrke et al., 2011; Elgendi et al., 2018). In chapter 4, we will investigate direct priming, and we will discuss how the model presented in this thesis could account for (indirect) semantic priming as well as negative priming.

Behavioral tagging

Behavioral tagging is also referred to as “flashbulb memory” and describes the observation that memories, which under “normal” circumstances would not become consolidated, do become consolidated if they happen in close temporal vicinity to an impactful event (Brown and Kulik, 1977; Moncada et al., 2015). Or, as Okuda et al. (2020) put it:

“the retention of trivial everyday memory can be boosted when novel or unexpected experiences happen shortly before or after the time of memory encoding”.

Thus, the “behavioral tag” can be considered a behavioral analog to the synaptic tag (cf. section 1.3.4).

In a seminal study, Li et al. (2003) demonstrated that in the hippocampus of rats, the transfer of E-LTP (induced by weak tetanic stimulation) to L-LTP could be facilitated behaviorally by exposing the rats to a novel environment. The effect was dependent on D1-type dopamine receptors but not on β -adrenoceptors or muscarinic receptors (targeted by noradrenaline and acetylcholine, respectively; also cf. section 1.3.6). Furthermore, activation of D1-type receptors was able to mimic the novelty effect.

The notion of behavioral tagging was eventually established through a study by Moncada and Viola (2007), who showed that the performance of rats in an inhibitory avoidance task was consolidated if they were placed in a novel environment prior to learning. Consolidation did not occur when the

animals were placed in a familiar environment, and it was dependent on dopamine receptors and protein-synthesis in the hippocampus. Later, behavioral tagging was also demonstrated in other learning tasks with both rodents and humans (cf. [Moncada et al., 2015](#)). Employing a feedforward network model, [Ziegler et al. \(2015\)](#) have provided a first theoretical account for behavioral tagging via STC in feedforward networks.

Behavioral tagging is related to several aspects of this thesis. Being a behavioral analog of STC, it may serve to test our predictions at the behavioral level, which applies to the consolidation and improvement of memory representations in chapter 2, the association of memory representations in chapter 4, and due to its dopamine dependence also to the neuromodulator dependence studied in chapter 3.

Reconsolidation

Before new memories become consolidated, there is a time window during which they are particularly vulnerable ([Ribot, 1882](#); [Müller and Pilzecker, 1900](#); [McGaugh, 2000](#); [Dudai, 2004](#); also see sections 1.4.3–1.4.6). They can, for example, be disrupted through electroconvulsive shock or head trauma ([Duncan, 1949](#); [Squire et al., 1981](#); [Ahmed et al., 2000](#)). Nevertheless, even consolidated memories are subject to destabilization *upon being recalled*, which is either followed by disruption or by so-called reconsolidation ([Misanin et al., 1968](#); [Nader et al., 2000a](#); [Sara, 2000](#)). The occurrence of reconsolidation depends on the availability of plasticity-related proteins (cf. sections 1.3.4 and 1.4.4), and is thus disturbed by their absence as well as by other influences such as electroconvulsive shock ([Misanin et al., 1968](#); [Sara, 2000](#)). For an extensive overview over paradigms of reconsolidation experiments, see [Tronson and Taylor \(2007\)](#).

The loss of memory caused by blocked reconsolidation occurs instantaneously and is thereby fundamentally different from extinction learning, which describes the gradual unlearning of associations by the repeated presentation of an unpaired stimulus ([Berman and Dudai, 2001](#); [Furini et al., 2013](#); [Lane et al., 2015](#)). Extinction learning is related to classical and operant conditioning (cf. section 1.4.1) and is typically employed in behavior therapy. Complementary to this, pharmacological blocking of reconsolidation could be a new promising approach to cure mental diseases. However, such pharmacological application in humans still seems far off in the future because the exact mechanisms of reconsolidation remain largely unknown ([Amorim et al., 2021](#)). Nevertheless, reconsolidation underlies the update of existing memory traces ([Kessler et al., 2017](#); [Besnard et al., 2012](#)), and there is evidence that this process also underlies various cases of curing effects by psychotherapy ([Lane et al., 2015](#)). Thus, extinction learning and reconsolidation seem to provide two different ways to treat psychiatric diseases such as post-traumatic stress disorder (PTSD), phobias, and other anxiety-related disorders.

While reconsolidation is not investigated directly in this thesis, it has strong ties to memory consolidation and possibly to the interaction between early- and late-phase synaptic plasticity. Therefore, we will discuss potential future modeling of reconsolidation in chapter 5.

2

Consolidation and improvement of a memory representation

2.1 Chapter summary

This chapter comprises an article (main manuscript: section 2.2, Supplementary Information: section 2.3) in which we follow our first research question: we examine how a model featuring early-phase plasticity based on spike-driven calcium dynamics and featuring late-phase plasticity based on synaptic tagging and capture (STC) can learn, consolidate, and recall a memory trace represented by a cell assembly in a recurrent neural network. To evaluate the recall performance, we apply cue stimulation to half of the neurons in the assembly and measure the firing rates of the neurons in the network. We do this at two different times: 10 s after learning, when late-phase changes are still absent, and 8 h later, when the early phase of the synaptic weights has decayed and the late phase dominates. We refer to these as “10s-recall” and “8h-recall”, respectively. Varying the coupling strength of the connections outgoing from inhibitory neurons in this setup, we find a regime of functional memory recall after learning and after consolidation. We further vary the size of the cell assembly and find that functionality occurs robustly across a large range of assembly sizes. By these results, we have demonstrated that our model can serve to describe synaptic consolidation, which denotes the consolidation of memory traces in the first hours after learning.

The second part of our first research question has concerned the improvement of memory recall through STC mechanisms. Now, we indeed find that in our STC-based model 8h-recall often performs better than 10s-recall, which proves robust across many parameter settings. We further scrutinize this improvement and find that it consists of two contributions that we call passive and active improvement, respectively. Passive improvement occurs by the transfer from the early to the late phase through the mere passage of time, whereas active improvement occurs by fast-rising early-phase weights evoked by 8h-recall, which we find prevented during 10s-recall. We further characterize the occurrence of passive improvement by exploring the phase space of improvement versus deterioration via analytical considerations, finding that passive improvement occurs across a wide range of parameter settings. Moreover, we consider a protocol in which we use an intermediate recall stimulus and show that this improves the final recall even further, consistent with findings of neuropsychological experiments (Erdelyi, 2010; Wallner and Bäuml, 2018). In the end of the article, we discuss different predictions for experiments from the synaptic to the behavioral level, by which our model can be tested.

In addition to the manuscript, this chapter contains a final section 2.4 where we briefly review issues of earlier variants of the model, and where we explore plasticity in the phase space of pre- and postsynaptic firing rates for two different variants of the model. By this, we show that in the model presented here, low background activity does not lead to any substantial plastic changes.

The recurrent network model developed in this chapter will serve as a basis for the investigations in the following chapters 3 and 4. Furthermore, the measures of recall performance will be used again in chapter 3, and the characterization of plasticity regimes will serve to justify approximative computation in chapter 4.

2.2 Manuscript: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks

This section is in content identical to the following published journal article (Luboeinski and Tetzlaff, 2021a):

Luboeinski, J., Tetzlaff, C. 2021. Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks. *Communications Biology* 4, 275. doi: [10.1038/s42003-021-01778-y](https://doi.org/10.1038/s42003-021-01778-y).

The typeset and referencing were adapted to match the format of this thesis. The Supplementary Information of the article is provided in the next section. The references are provided at the end of this thesis. My contribution to this article was about 85%. I developed the software, performed the simulations and data analysis, and produced the figures. I drafted, wrote and reviewed this article jointly with Christian Tetzlaff. This manuscript is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

Abstract

The synaptic-tagging-and-capture (STC) hypothesis formulates that at each synapse the concurrence of a tag with protein synthesis yields the maintenance of changes induced by synaptic plasticity. This hypothesis provides a biological principle underlying the synaptic consolidation of memories that is not verified for recurrent neural circuits. We developed a theoretical model integrating the mechanisms underlying the STC hypothesis with calcium-based synaptic plasticity in a recurrent spiking neural network. In the model, calcium-based synaptic plasticity yields the formation of strongly interconnected cell assemblies encoding memories, followed by consolidation through the STC mechanisms. Furthermore, we show for the first time that STC mechanisms modify the storage of memories such that after several hours memory recall is significantly improved. We identify two contributing processes: a merely time-dependent passive improvement, and an active improvement during recall. The described characteristics can provide a new principle for storing information in biological and artificial neural circuits.

Introduction

In biological neural systems, memories have a wide repertoire of dynamics; most importantly, they can be encoded, stored, recalled, and consolidated. While these dynamics are relatively well-explored at the behavioral and brain-region level (Dudai, 2004; Eichenbaum, 2011; Gluck et al., 2016), the underlying synaptic and neuronal processes remain mainly elusive.

Most generally, learning describes the ability of humans and other animals to obtain knowledge about an entity. This knowledge or information is stored as a memory. The encoding of such a memory in a neural network is commonly assumed to happen in the way as described by Hebb in his seminal work (Hebb, 1949; Martin et al., 2000; Eichenbaum, 2018): A group of recurrently connected neurons that receives the information by an external input starts to fire stronger than the rest of the network. This increased firing yields strengthening of the efficacy of the recurrent synapses within this particular group such that a so-called Hebbian cell assembly is formed, which represents a memory of the input. On the other hand, low firing rates typically cause weakening of connections between neurons, which can lead to either disruption or refinement of a cell assembly. Strengthening and weakening of synapses at timescales relevant to memory is the result of long-term synaptic plasticity (Artola and Singer, 1993; Bear, 1996; Abraham et al., 2019).

Long-term synaptic plasticity creates long-lasting changes of the synaptic efficacy. To become strengthened, synapses undergo a cascade of molecular processes that leads to an increase in the number of postsynaptic receptors, which is called long-term potentiation (LTP; Artola et al., 1990; Bliss and Collingridge, 1993; Smolen et al., 2012; Abraham et al., 2019). Analogously, for weakening, another cascade of processes yields a decrease in the number of receptors, which is called long-term depression (LTD; Artola et al., 1990; Smolen et al., 2012; Abraham et al., 2019). The signaling cascades of both LTP and LTD are triggered by the calcium concentration in the postsynaptic spine. The spiking activities of the pre- and the postsynaptic neurons drive the calcium concentration and, by this, determine whether long-term potentiation or long-term depression of the synaptic efficacy is induced (Artola and Singer, 1993; Hansel et al., 1996; Shouval et al., 2002; Graupner and Brunel, 2012; Li et al., 2016). In general, long-term synaptic plasticity consists of at least two different phases. Changes of the synaptic efficacy in the early phase last for several hours, while efficacy changes in the late phase can be maintained for several days (Bliss and Collingridge, 1993; Abraham, 2003). The transfer from the early to the late phase has been described by the synaptic-tagging-and-capture (STC) hypothesis (Frey and Morris, 1997; Redondo and Morris, 2011). Following the STC hypothesis, the transfer depends on the stimulation of the specific synapse as well as on the stimulation of other synapses at the same postsynaptic neuron. More precisely, the transfer at a synapse occurs if the synapse is tagged, which means that it is primed for transfer, and if proteins necessary for the late phase are abundant or have been synthesized. The tagged synapse then “captures” proteins causing the transfer to the late-phase by increasing for instance the number of receptor slots at the postsynaptic site (Redondo and Morris, 2011). The formation of a tag at a synapse is related to its own early-phase change, while protein synthesis depends on the early-phase state of many synapses (Frey and Morris, 1997; Clopath et al., 2008; Redondo and Morris, 2011; Li et al., 2016). The total or effective synaptic efficacy of a synapse consists of the sum of the early- and late-phase contribution (Clopath et al., 2008; Li et al., 2016).

In general, consolidation of memories means the progressive transfer of memories into a state in which they stay stable over long time intervals (Müller and Pilzecker, 1900; Dudai, 2004; Eichenbaum, 2011). There are two major categories of memory consolidation: systems consolidation and synaptic (or initial) consolidation (Dudai, 2004; Okuda et al., 2020). The basic idea of systems consolidation is that a memory is transiently stored in the hippocampus and possibly transferred to the neocortex, in which it is maintained for a longer period. The question, whether a memory is first encoded in the hippocampus and then transferred to the neocortex or whether the encoding of a memory occurs simultaneously in both regions (multiple trace theory; Nadel and Moscovitch, 1997; Dudai, 2004), is subject to an ongoing debate. In both cases, however, the newly formed memory has to be initially consolidated before systems consolidation sets in. This initial consolidation process is related to local molecular and cellular processes at individual neurons and synapses and is therefore named synaptic consolidation (McGaugh, 2000; Clopath, 2012; Dudai et al., 2015).

The STC hypothesis provides a potential explanation of the neuronal and synaptic processes underlying the synaptic consolidation of memories (Frey and Morris, 1997; Dudai, 2004; Redondo and Morris, 2011), which is supported by several theoretical studies focusing on single synapses or feed-forward networks (Clopath et al., 2008; Barrett et al., 2009; Smolen et al., 2012; Ziegler et al., 2015; Li et al., 2016). However, a clear link between the STC hypothesis and memory consolidation is still missing as the encoding of memories in neural circuits is mainly associated with strongly recurrently connected groups of neurons (cell assemblies). Another study of recurrent networks with multiple cell assemblies already found that cue-triggered recall was only possible after 20 minutes if a type of synaptic consolidation was present (Zenke et al., 2015).

In this study, we developed a theoretical model of recurrently connected spiking neurons with the synaptic efficacies being altered by calcium-dependent synaptic plasticity and the core mechanisms of the STC hypothesis. The individual components of the implemented model reproduce various plasticity phenomena as the ones described above (Frey and Morris, 1997; Bi and Poo, 1998; Sjöström et al., 2001; O'Connor et al., 2005; Wang et al., 2005; Letzkus et al., 2006; Nevian and Sakmann, 2006; Sjöström and Häusser, 2006; Wittenberg and Wang, 2006; Reymann and Frey, 2007; Sajikumar et al., 2007) and for verification we matched the temporal evolution of individual synapses in our model with experimental data. Our network simulations show the synaptic and neuronal dynamics underlying the formation of a memory representation, its recall, and its long-term development. The latter indicates that the STC mechanisms in a recurrent circuit lead to the consolidation of a memory. Finally, the simulations and analytical results suggest a new implication of the STC mechanisms on memory representations, which is the enhancement of the storage of the memory. This enhancement exhibits a new type of memory dynamics, which could be beneficial for biological as well as for artificial memory systems.

Previous theoretical studies (Fusi et al., 2005; Elliott and Lagogiannis, 2012; Elliott, 2016; Benna and Fusi, 2016) investigated general computational principles emerging from multiple-timescale plasticity. For instance, using rather abstract models of multiple-timescale plasticity based on binary (Elliott

and Lagogiannis, 2012) and discrete-valued (Elliott, 2016) synaptic weights, Elliott and colleagues predicted the possibility of the improvement in memory strength with the passage of time. However, the underlying dynamics have not been linked to specific biological mechanisms. While Papper et al. (2011) identified that STC mechanisms can trigger an improvement in memory lifetime, we show that improvement in memory strength can also be related to the biologically plausible mechanisms of synaptic tagging and capture together with calcium-dependent plasticity.

Results

We aimed to set up a biophysically plausible network model of the neuronal and synaptic dynamics underlying synaptic consolidation. For this, based on previous studies (Clopath et al., 2008; Graupner and Brunel, 2012; Li et al., 2016), we utilized a synaptic model that integrates the local calcium dynamics to trigger synaptic plasticity and the mechanisms of synaptic tagging and capture (Fig. 2.1a). We used this synaptic model in conjunction with a leaky integrate-and-fire model, axonal delay, and a finite synaptic time constant to cover the most important aspects of signal generation and transmission in neural networks with parameters based on experimental findings (see Methods; Gerstner and Naud, 2009; Tchumatchenko et al., 2011). The network itself was sparsely coupled and received input such that it resembled connectivity and firing rates as present in the hippocampus during exploratory wake state (Braitenberg and Schüz, 1998; Mizuseki and Buzsaki, 2013; Le Duigou et al., 2014; Mizuseki and Miyawaki, 2017; see also Supplementary Fig. 2.12). By applying an additional, strong input stimulus to a particular subset of excitatory neurons, cell assemblies were formed (Fig. 2.1b). In the following, we will first introduce the key principles of the synapse model and compare its dynamics to experimental data, before we present the characteristics of the network model.

Comparison of synapse model with experimental data

In general, the model assumes that the membrane potential of a neuron determines its spiking dynamics, which drives together with presynaptic spiking the postsynaptic calcium concentration (Fig. 2.1a). The calcium concentration determines the occurrence of early-phase LTP and early-phase LTD, represented by changes in the early-phase component (or weight) of the synaptic efficacy. Large changes of the early phase weight trigger the formation of a synapse-specific tag. A sufficient body of early-phase changes at many synapses of the postsynaptic neuron triggers protein synthesis. Once an adequate amount of proteins is available and the synapse is tagged, the late phase component of the synaptic efficacy is altered; thus, the synapse “captures” proteins. The sum of the early- and late-phase weight yields the total synaptic efficacy that determines the magnitude of postsynaptic potentials arriving at the neuron, influencing its membrane potential. The interplay between these different processes is investigated in standard plasticity induction experiments (Sajikumar and Frey, 2004b; Sajikumar et al., 2005, 2007; Redondo and Morris, 2011). In these experiments a strong tetanus stimulation (STET) is used to induce late-phase potentiation, while weak tetanus stimulation (WTET) is used to induce early-phase potentiation only. For late-phase depression, a strong low-frequency stimulus (SLFS) is used, while for early-phase depression, a weak low-frequency stimulus (WLFS) suffices. As proof of concept of our synaptic model, we reproduced the outcome of these experiments by considering a sin-

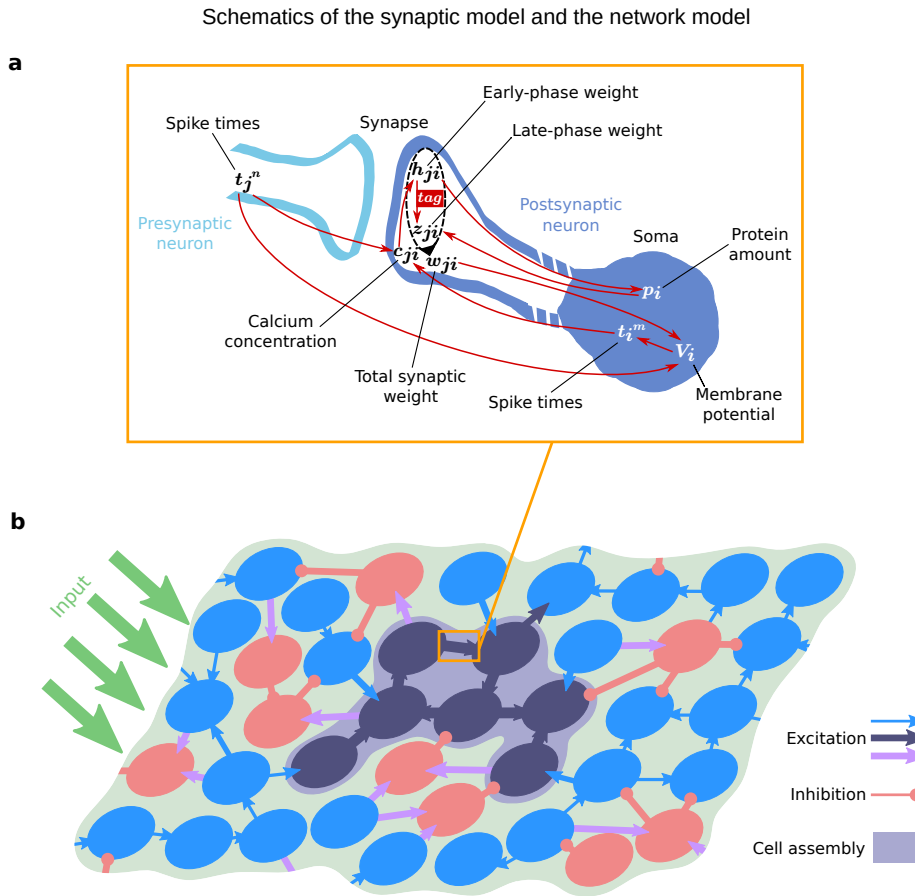


Figure 2.1: Schematics of the synaptic model and the network model. (a) The synaptic model integrates the interplay between various mechanisms of calcium-dependent synaptic plasticity and the STC hypothesis. For more details see the main text. (b) Schematic of a part of the neural network that consists of excitatory (blue and dark blue circles) and inhibitory neurons (red circles) and receives external input from other brain areas (green arrows). Only synapses between excitatory neurons (blue and dark blue arrows) undergo plastic changes by the processes shown in (a). A Hebbian cell assembly represents a memory and consists of a group of strongly interconnected neurons (dark blue).

gle plastic synapse between two neurons and applying similar stimulation protocols to the presynaptic neuron (see Supplementary Fig. 2.10 for details). The resulting time traces of the synaptic efficacy in response to the four induction protocols (Fig. 2.2) match the findings from experiments as discussed above and from previous theoretical studies (Clopath et al., 2008; Barrett et al., 2009; Ziegler et al., 2015; Li et al., 2016). The fluctuations and local maxima in our STET and SLFS simulations are within the common range for late-phase plasticity paradigms (cf. Sajikumar and Frey, 2004b; Sajikumar et al., 2005, 2007). The results indicate that our synaptic model provides a reasonable description of the underlying biomolecular dynamics. Thus, based on this synaptic model, in the next step we will introduce our analysis of the synaptic and neuronal dynamics in a recurrent neural network.

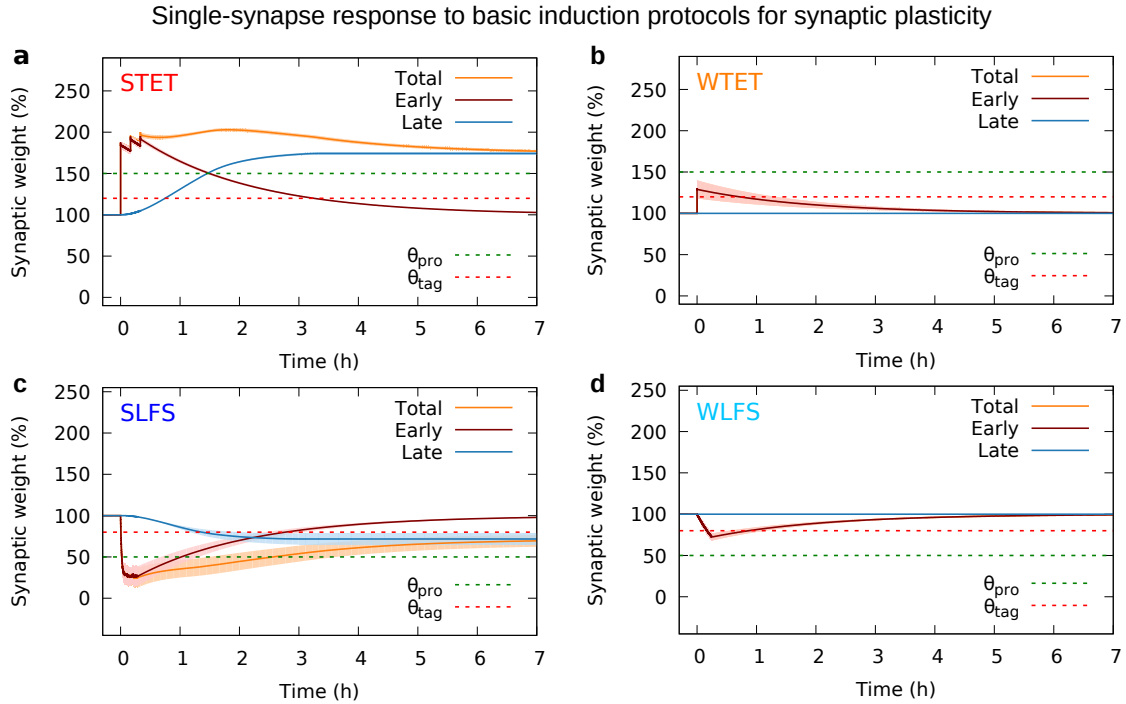


Figure 2.2: **Impact of strong and weak tetanic and low-frequency stimulation protocols at a single synapse.** Each protocol (see Supplementary Fig. 2.10 for details) leads to the induction of a different type of synaptic plasticity: **(a)** late-phase potentiation, **(b)** early-phase potentiation, **(c)** late-phase depression, and **(d)** early-phase depression. The late-phase weight (blue line) is only changed by strong stimulation (STET, SLFS). The early-phase weight (red line) is affected by all protocols. Weak stimulation protocols (WTET, WLFS) suffice to drive the early-phase weight across the threshold of tag formation (θ_{tag} , dashed red line) but not across the threshold of triggering protein synthesis (θ_{pro} , dashed green line). The total weight or synaptic efficacy (orange) is the sum of early- and late-phase weight. Averaged over 100 trials; error bands show the standard deviation. The late-phase weight has been shifted for graphical reasons (cf. Methods).

Network model with synaptic consolidation enables functional memory representations

Employing our synaptic model, we simulated a patch of tissue with hippocampal network characteristics, consisting of 2000 neurons with 20% being inhibitory and a 0.1 average probability of two neurons being connected by a synapse (Fig. 2.1b). Synapses between excitatory neurons (blue, dark blue) feature plasticity mechanisms as described above. Inhibitory neurons provide feedback inhibition; their connections are non-plastic (purple, red). All neurons in the patch received additional inputs from outside the network, which have particular characteristics for the encoding and recalling of a memory. During learning, a specific stimulus is provided to a group of neurons and should trigger the formation of particularly strong connections between these neurons, which then represent a Hebbian cell assembly or memory (Fig. 2.1b, dark blue). During recall, only a subset of the group of neurons that received the learning stimulus received specific external stimulation.

To investigate the synaptic and neuronal dynamics implied by the STC mechanisms, throughout this study, we focused on two time spans. We evaluated the state of the neural network ~ 10 s after the learning stimulus to analyze the short-term dynamics and after ~ 8 h to investigate the long-term

effects of the STC mechanisms.

The learning stimulus consisted of three subsequent pulses of 0.1 s each, strongly activating a random subset of neurons. In experiments, strong tetanus stimuli of ~ 100 Hz are used to evoke late-phase synaptic potentiation (Frey and Morris, 1997; Sajikumar et al., 2005, 2007). To resemble this, we applied stimulation of similar strength via 25 putative input neurons. A necessary consequence of this strong input is the fact that during stimulation, the stimulated neurons spike at their maximum. As expected, the stimulus caused tremendous changes of the synaptic efficacy of diverse synapses in the network (compare Fig. 2.3a with b). Synapses associated with the stimulated neurons (first 150 neurons Fig. 2.3a-c) mainly experienced LTP (red). The synapses between stimulated neurons (black box in Fig. 2.3a-c) are strengthened particularly, indicating the correct formation of a cell assembly. By contrast, synapses between non-stimulated neurons underwent LTP as well as LTD (blue). After 8 h, the synaptic changes between non-stimulated neurons fade, such that mainly changes associated with the stimulated neurons and, thus, with the cell assembly remain (Fig. 2.3c). To validate that the formed cell assembly encodes a memory, next, we tested the ability of recall. For this, we applied a recall stimulus that activates for 0.1 s 50% of the neurons that were stimulated by the learning signal, and analyzed the resulting spread of activation within the network (Fig. 2.3d, e). The externally stimulated neurons ('as') have the highest activity. By the strong recurrent connections in the cell assembly, the average activity of the non-stimulated neurons in the cell assembly ('ans') is significantly higher than the activity of the remaining non-stimulated neurons ('ctrl'). This is not the case for control stimulation applied before learning. In other words, the activity of the recall stimulus spreads via the stimulated neurons to the other cell assembly neurons and yields a type of pattern completion or recall (in the literature sometimes also called retrieval). Please note that, under basal/standby conditions, the mean firing rate of all excitatory neurons resembles the average value of 0.5–1.0 Hz for pyramidal cells in the hippocampus during exploratory wake states (Mizuseki and Buzsáki, 2013; Mizuseki and Miyawaki, 2017; see also Supplementary Fig. 2.12). We further analyzed the spread of activity by considering the mean firing rate correlations between the neuronal subpopulations (Fig. 2.3f). The correlations generally increase with learning and consolidation, which is in line with the previous finding of increased mean firing rates. Furthermore, we found that correlations within the cell assembly ('as', 'ans') are larger than correlations with control neurons ('ctrl'), which again indicates functional memory recall and further leads to the expectation of more LTP for connections involving 'as' and 'ans' neurons. This already suggests an "active improvement" of the cell assembly, which we will investigate in a later section. On the other hand, the fact that correlations between cell assembly neurons and control neurons are also increased by recall stimulation, shows that control neurons indirectly contribute to recall. This is corroborated by previous work (Tetzlaff et al., 2015) and by the distribution of synaptic weights, which exhibits tremendously increased weights from control to cell assembly neurons ("incoming" weights in Fig. 2.3b,c and Supplementary Figs. 2.13 and 2.14). In summary, the results presented in this section show that calcium-dependent synaptic plasticity and the STC mechanisms enable the encoding of memory and its maintenance for several hours in a recurrent neural network. Thereby, our findings support the idea that the STC mechanisms account for the synaptic consolidation of memories. In the next section, specific measures for memory recall will be presented.

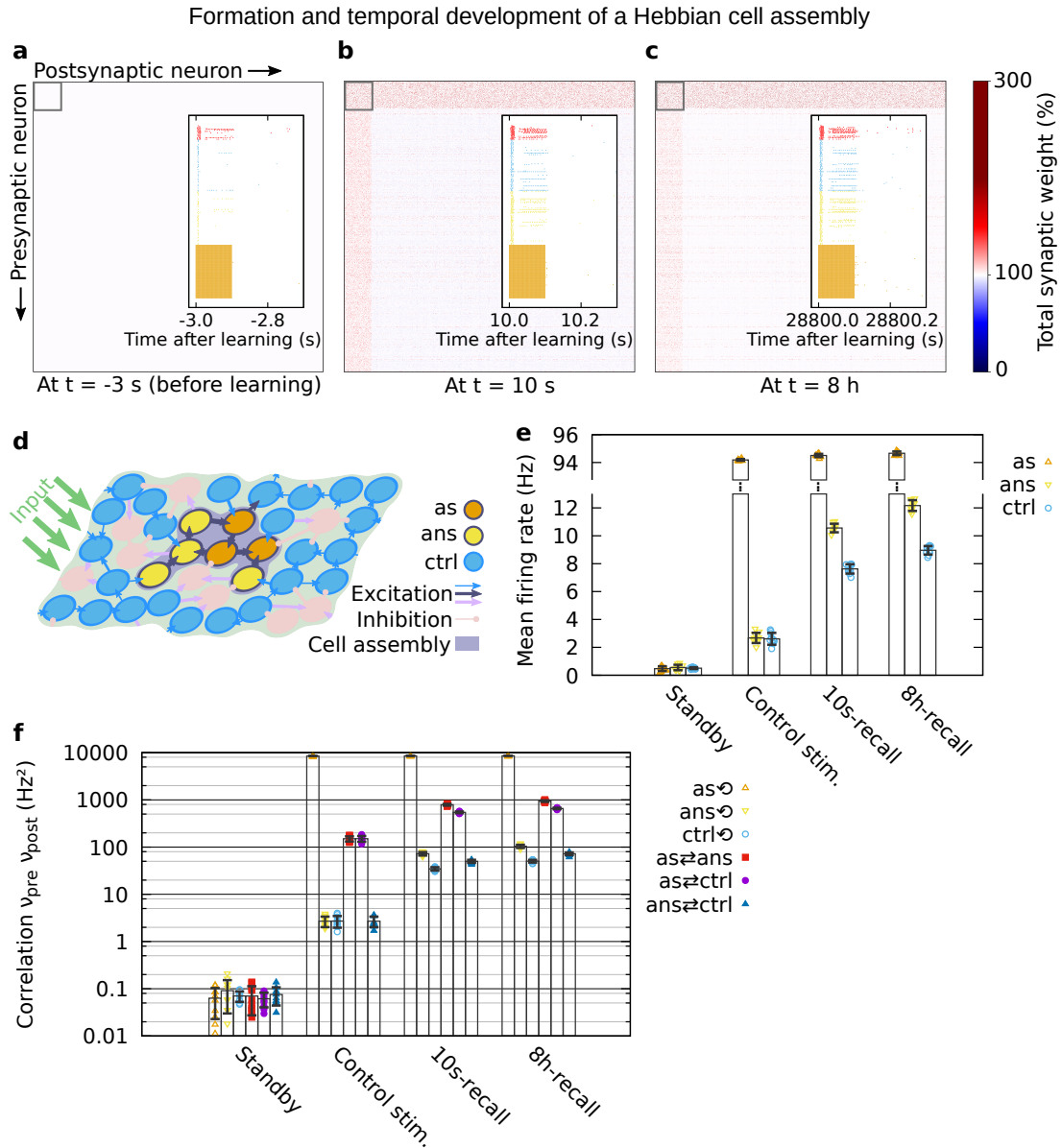


Figure 2.3: Formation and temporal development of a Hebbian cell assembly and its functionality. (a–c) The matrix of total (early- plus late-phase) synaptic weights sorted by the stimulated neurons (first 150): (a) before learning, (b) ~10s after learning, and (c) ~8h after learning. The insets show the spiking during recall stimulation of the neurons in the cell assembly (colors as in (d,e)) and of a random 5% fraction of the neurons in the remaining excitatory (blue) and inhibitory (red) populations. (d) Schematic of the three subpopulations during recall stimulation: the fraction of externally stimulated cell assembly neurons (‘as’), the fraction of cell assembly neurons that are not stimulated (‘ans’), and the remaining excitatory neurons acting as control group (‘ctrl’). (e) The mean firing rates of the three subpopulations defined in (d), before learning and without recall (standby), before learning during control stimulation, upon recall 10s after learning, and upon recall 8h after learning. (f) Mean firing rate correlations (presynaptic multiplied by postsynaptic firing rate, averaged over neuron pairs) within and between subpopulations for the same protocols as in (e). Synaptic weights, mean firing rates and mean correlations were averaged over ten trials. Error bars show the standard deviation. Parameters: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$.

Memory functionality depends on network inhibition and size of the representation

In the following, we investigate the influence of different parameters such as inhibition and cell assembly size on the learning, recall, and consolidation of memory representations.

In general, one role of inhibition in neural networks is to prevent the network activity from runaway or epileptiform dynamics. This kind of dynamics is characterized by neurons constantly exciting each other and firing at the maximum possible rate. On the other hand, if inhibition is excessively strong, neuronal activity is confined, resulting in a more or less silent network. We adjusted the level of inhibition in our network by varying the coupling strength from inhibitory neurons to excitatory neurons (w_{ie}) and the coupling strength from inhibitory to inhibitory neurons (w_{ii}). For every set of these parameter values, we simulated the network dynamics during learning and recall as described before. At two different points in time (10 s and 8 h after learning), we evaluated the memory functionality that consists of learning, consolidation and recall of a memory by measuring the recall quality. To measure the recall quality, we used two different quantities (cf. Methods): on the one hand, the pattern completion coefficient Q describes by how much the activity in the non-stimulated part of the input-defined cell assembly is raised during recall as compared to the activity of the control neurons; on the other hand, the mutual information MI_V describes how similar the network activity state during recall is to the network activity state during learning. We quantitatively defined memory functionality by an average pattern completion coefficient of $Q \geq 0.03$. This criterion demands that, upon stimulation of half of the assembly neurons, the other half of the assembly be activated much stronger than the background neurons, which remain at a low activity level (cf. Fig. 2.3e). Please note that the values of the pattern completion coefficient appear small because we normalize by the activity of the stimulated neurons (cf. Eq. 2.16).

High values of the I→I coupling strength together with low values of the I→E coupling strength imply a low level of inhibition, which impedes memory functionality in our network model (see Fig. 2.4a,c for the recall performance 10 s after learning and Fig. 2.4b,d for 8 h after learning). This impairment goes along with the runaway dynamics discussed before, as indicated by the extreme level of population activity in this parameter regime (Fig. 2.4e,f). In general, if a network is in such an “epileptiform” state, a recall stimulus does not exclusively activate the neurons that belong to the cell assembly, it also activates a large fraction of control neurons, overshadowing memory recall. On the other hand, low values of the I→I coupling strength together with high values of the I→E coupling strength lead to extensive levels of inhibition, suppressing the neuronal activity during recall. In other words, the activity induced by a recall stimulus will immediately be suppressed by the high level of inhibition, also impeding memory recall. As a result, the level of inhibition has to be within a specific regime to enable learning, consolidation and successful recall of memories (indicated by the red box; $Q(10\text{s-recall}) \geq 0.03$). This regime is characterized by medium levels of population activity. The functionality is relatively stable given variations in the mean background current I_0 (Supplementary Fig. 2.15), while also the mean firing rate only weakly depends on the mean background current (Supplementary Fig. 2.12b). Furthermore, the recall quality remains relatively high at low learning and recall stimulus frequencies, except for very low frequencies, where the memory functionality ceases (Supplementary Fig. 2.16). This resilience is partially due to the number of 25 putative input neurons.

In summary, the level of inhibition within the described regime of functionality seems to be sufficient to prevent the overly activation of control neurons, while it remains low enough to allow the desired activation of the non-stimulated neurons within the cell assembly. Please note that such a regime is related to the network state of loosely-balanced excitation and inhibition (Denève and Machens, 2016; also see discussion).

The regime of memory functionality is the same either 10 s or 8 h after providing the learning stimulus. However, the higher values of Q and MI_V after 8 h compared to the 10 s case indicate a positive influence of long-term dynamics on the memory functionality. We further investigate and discuss this result in the next section.

As further parameter, we examined the influence of the size of the cell assembly on the memory functionality (Fig. 2.5a,b). We controlled the size by varying the number of neurons being the sub-group that is stimulated by the learning stimulus. Following our definition of memory functionality, requiring that the coefficient Q be greater or equal than 0.03, we found that learning and recall is only possible if the cell assembly is large enough. In the following, we will focus on a particular set of parameter values for the inhibition ($w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$). All derived conclusions apply to all sets within the specific regime of inhibition discussed before (red box in Fig. 2.4). For small cell assemblies (here ~ 100 neurons), the group of neurons is too small to exhibit functional pattern completion (threshold indicated by the dotted red line in Fig. 2.5a,b). For large cell assemblies (here > 500 neurons), the activity of the cell assembly becomes self-sustained. This means that after learning the neurons within the cell assembly stay active, preventing the typical dynamics of memories we are looking for. Moreover, the measures of the recall quality exhibit local maxima. The occurrence of these maxima cannot be explained by the relationship between the mean firing rate and the cell assembly size, because this relationship is strictly monotonic and does not feature any maximum (Supplementary Fig. 2.12a). Therefore, we reason that the maxima emerge from the additional stimulation by the recall cue. Finally, similar to Fig. 2.4, the pattern completion coefficient and mutual information become higher 8 h after learning compared to 10 s after learning for a large range of cell assembly sizes, which will be examined further in the next section.

Consolidation improves memory recall

Comparing the recall quality of the cell assembly 10 s after learning with the recall quality 8 h after learning for different paradigms (Figs. 2.4 and 2.5a,b), we found that 8 h after learning the system generally exhibits an improved recall quality. In other words, the specific state of the cell assembly after 8 h, resulting from the interplay between calcium-dependent synaptic plasticity and the STC mechanisms, seems to facilitate recall. To elucidate the magnitude of the improvement effect, we computed the relative gain in recall quality between the state 10 s after learning and the state 8 h after learning and present results for the particular inhibition setting that we chose before (Fig. 2.5c,d). It becomes obvious that improvement in recall quality can occur by $> 200\%$ with respect to the pattern completion coefficient and by as much as 30% with respect to the mutual information.

To test the robustness of the improvement effect across the specific inhibition regime identified before, we averaged MI_v and Q over the range of parameter settings of this regime and calculated the relative gain in recall quality (Fig. 2.5e,f). For the chosen inhibition setting $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$ (Fig. 2.5c,d) as well as for the averaged results (Fig. 2.5e,f), we observed a positive gain across all cell assembly sizes. Positive gain means that 8 h after learning the performance was better than 10 s after learning. Thus, for a wide range of parameter settings, the passage of time yields an improvement of the recall quality.

What are the underlying principles that lead to this improvement of the recall quality? We could identify two processes, both being related to the STC mechanisms: an active and a passive improvement.

The *active improvement* is related to the dynamics of the early-phase weights. The recall stimulus provided 10 s after learning only leads to minor variations of the average early-phase weight in the cell assembly (dashed red line at 'R' in Fig. 2.6a). By contrast, 8 h after learning the recall stimulus triggers a substantial increase of the average early-phase weight, which in turn results in a stronger total synaptic weight (dashed red and orange line at 'R' in Fig. 2.6b). Thus, 8 h after learning the average total synaptic weight of the cell assembly increases during the recall, which improves the recall performance (Fig. 2.5). By comparing the state of the neural network between both time points, we find that the average early-phase weight before the 10s-recall resides on a much higher level than before the 8h-recall (Fig. 2.6c). This difference in early-phase weight before recall stimulation could explain the different dynamics of the early-phase weights during the recall. Considering the mathematical formulation of the model, one can see that the change of the early-phase weight depends on the distance of the actual early-phase weight h from its initial or equilibrium value h_0 . Hence, larger changes occur if the early-phase weight is closer to the equilibrium state, while the changes remain smaller if the early-phase weight is close to its maximum value. Thus, the learning stimulus "pushes" the early-phase synaptic weight into a regime in which the subsequent recall cannot trigger further strengthening. However, with the passage of time the early-phase weight decays (while the late-phase weight increases) until it reaches the vicinity of its initial value (Fig. 2.6b,c). In this regime a recall stimulus can again trigger an increase of the early-phase weight, supporting pattern completion. The detailed weight distributions for all cases presented in Fig. 2.6c and additional data revealing the weights between subpopulations of the network are shown in Supplementary Figs. 2.13 and 2.14. To further scrutinize the relation between the early-phase dynamics during recall and the improved recall quality, we performed simulations in which we switched off or blocked early-phase plasticity after learning. Comparing the resulting recall performances with the simulations with early-phase plasticity (Fig 2.7), for both measures Q and MI_v , we did not find an influence of the blockage during 10s-recall but during 8h-recall. These results further support our finding that the dynamics of the early-phase weights within the cell assembly are one essential part underlying the improvement of the recall quality.

The other part of the improvement of the recall quality is related to the dynamics of the late-phase weights within the cell assembly. We refer to this part as *passive improvement*. As expected, 8 h after

learning the STC mechanisms yield a decrease of the average early-phase weight (red line in Fig. 2.6b) accompanied by an increase of the average late-phase synaptic weight in the assembly (blue line) that indicates the process of synaptic consolidation. However, if we compare the average total synaptic weight within the assembly before the 10s-recall with the one before the 8h-recall (Fig. 2.6), we identify an increase that suggests a second component underlying the improvement of recall quality. This is also indicated by the relative gain in Q and MI_V without early-phase plasticity (Fig. 2.7). Hence, although the dynamics of the early-phase weights (active improvement) explains the gain in recall quality to a large extent, there is a remaining part that is related to the dynamics of the late-phase weights (passive improvement).

Parameter dependency of the passive improvement

We have found that the mean total weight of a cell assembly can increase over time, and refer to this phenomenon as passive improvement (cf. Fig. 2.6c). This effect is elicited by the mechanisms of synaptic tagging capture and occurs if the mean normalized late-phase weight $\tilde{z} \cdot h_0$ after 8 h becomes higher than the mean early-phase weight \tilde{h} after learning. To investigate and predict the circumstances under which this effect occurs, namely, different settings of late-phase-related constants, we considered an analytical approach that we present in the following.

To investigate the parameter-dependency of consolidation and improvement of a cell assembly, we analytically solve the linear first-order differential equation describing the dynamics of late-phase LTP (cf. Eq. 2.13) with a variable coefficient introduced by the time-dependent amount of proteins (Eq. 2.14). Since L-LTP is vastly predominant between cell assembly neurons, we neglect here the L-LTD dynamics.

Thus, the analytical expression for the mean late-phase weight after time t after learning in the case of persistent tags and protein synthesis is:

$$z_t(t) = 1 - (1 - z_0) \cdot \exp\left(-\frac{\tau_p(\alpha - p_0) \cdot e^{-\frac{t}{\tau_p}} + \alpha t - \alpha\tau_p + p_0\tau_p}{\tau_z}\right). \quad (2.1)$$

depending on the protein time constant τ_p , the late-phase time constant τ_z , the initial protein amount p_0 , the initial late-phase weight z_0 , and the protein synthesis rate α . Note that this equation only holds if tags have been set and as long as they have not vanished, which we will address later. This solution can be further simplified under the condition that the initial late-phase weight z_0 and the initial protein amount p_0 are zero. The mean increase in late-phase weight due to the STC mechanisms is then:

$$z_s(t) = 1 - \exp\left(-\frac{\alpha \left(t + \tau_p \cdot e^{-\frac{t}{\tau_p}} - \tau_p\right)}{\tau_z}\right) \quad (2.2)$$

For Eq. 2.2, we considered that the synapses within the cell assembly are always tagged, triggering changes of the late-phase weights. However, after learning the early-phase weight decays (Fig. 2.6b) and falls at a certain point in time below the tag-sufficient threshold. Thus, first, we have to calculate

the decay of the mean early-phase weight and, then, use it to determine at what time the tag will have vanished. The average decaying early-phase weight follows

$$h_d(t) = (\tilde{h} - h_0) \cdot \exp\left(-0.1 \frac{t - t_{al}}{\tau_h}\right) + h_0 \quad (2.3)$$

depending on the mean early-phase weight after learning \tilde{h} , the time after learning t_{al} , the initial weight h_0 , and the early-phase time constant τ_h . The time t_{al} is a time after learning (for instance, 10 s) at which \tilde{h} , the mean learning-induced increase in early-phase weight, is measured. Setting $h_d(t)$ to the value of the tagging threshold $\theta_{tag} + h_0$ yields the point in time t_{tag} at which the synaptic tags will have vanished on average:

$$t_{tag} = 10 \tau_h \cdot \ln\left(\frac{|\tilde{h} - h_0|}{\theta_{tag}}\right) + t_{al}. \quad (2.4)$$

The synaptic tags typically vanish before the synthesis of new proteins stops and, thus, t_{tag} determines the end of late-phase changes, such that the average late-phase weight is described by

$$z_{st}(t) = \begin{cases} z_s(t) & \text{if } t \leq t_{tag}, \\ z_s(t_{tag}) & \text{if } t > t_{tag}. \end{cases} \quad (2.5)$$

Hence, for parameter sets yielding $t_{tag} < 8$ h (cf. Eq. 2.4), the mean late-phase weight before the 8h-recall is the same as by the time the tags vanished (see also Fig. 2.6b, blue line):

$$\tilde{z} = z_s(t = t_{tag}). \quad (2.6)$$

In Fig. 2.8a we plot the mean weights within the cell assembly, using our numerical results for early-phase LTP during learning and recall to compute the late phase via the analytical function in Eq. 2.5 and the decay of the early-phase weight as detailed in Eq. 2.3. As expected, we find that this analytical approach predicts very well the mean weight dynamics of our numerical simulations (dashed gray line in Fig. 2.8a, cf. Fig. 2.6b).

When we compare the mean total weight 8 h after learning $w(t = 8 \text{ h})$ with the mean total weight 10 s after learning $w(t = 10 \text{ s})$, we obtain an expression for the passive weight change realized by the STC mechanisms, thus, of the passive improvement:

$$\Delta w = w(t = 8 \text{ h}) - w(t = 10 \text{ s}) = \tilde{z} \cdot h_0 - \tilde{h}. \quad (2.7)$$

If $\Delta w > 0$, the system shows a passive improvement with the passage of time, which is given for a wide range of parameter values. For instance, by considering different values of the protein synthesis parameters, namely the time constant of the protein amount τ_p (Fig. 2.8b) and protein synthesis rate α (Fig. 2.8c), depending on the time scale of the late-phase dynamics τ_z , we can determine the influence of protein synthesis on the passive improvement. We find that, if the protein time scale becomes too high or the synthesis rate becomes too low, the passive weight change switches from improvement to deterioration. However, the protein dynamics can be much slower than in our numerical simulations (red box) and still the STC mechanisms yield an improvement. For the protein synthesis rate, there is

just as well a wide range of values that gives rise to improvement. Thus, the encountered improvement effect is robust to parameter variations.

Our analytical results lead to the following predictions: if the time scale of the protein dynamics is manipulated, the resulting memory behavior should switch between passive improvement and deterioration. In addition, manipulating the speed of signaling cascades that are involved in late-phase LTP should have a similar effect. Taking this into account, our predictions can be useful to narrow down the set of possible proteins and processes involved in synaptic tagging and capture.

Intermediate recall further amplifies memory improvement

There are several psychological studies that investigate hypermnesia/improvement by considering the memory strength after multiple recalls (Erdelyi and Kleinbard, 1978; Payne, 1987; Wallner and Bäuml, 2018). Therefore, we also investigated the impact of an intermediate recall on the memory dynamics. For this, we computed the gain in the pattern completion coefficient Q and in the mutual information MI_v between 10s-recall and 8h-recall after applying an additional, intermediate recall stimulus to randomly drawn neurons.

We varied the time of occurrence of this intermediate stimulus (Fig. 2.9). The gain in recall quality for early intermediate recall (minutes after learning) shows little difference to the gain without intermediate recall (data point at time zero). For late intermediate recall, the gain reaches values that are much higher compared to the case without intermediate recall. Hence, additional recall improves the recall quality even further, which is consistent with the experimental findings (Erdelyi and Kleinbard, 1978; Payne, 1987; Wallner and Bäuml, 2018).

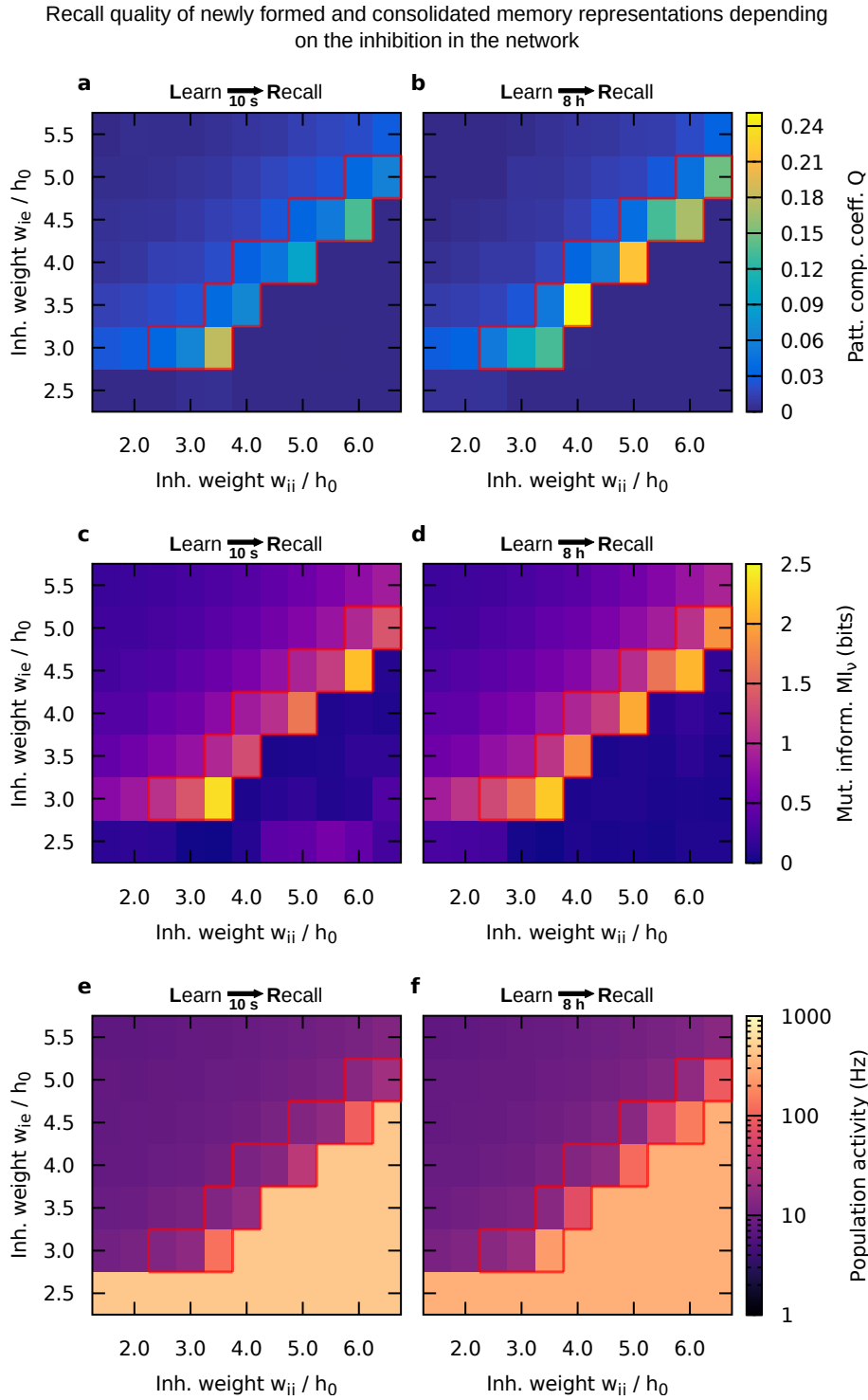


Figure 2.4: **Quality of the recall 10s and 8h after learning as a function of the I→E (w_{ie}) and I→I (w_{ii}) coupling strengths.** Sufficient memory functionality is found within a certain regime, marked by a red box ($Q(10s\text{-recall}) \geq 0.03$). Two measures of recall quality are shown: **(a,b)** the pattern completion coefficient Q (with non-significant values set to zero, see Methods); **(c,d)** the mutual information MI_v between the neuronal activities in the network during learning and during recall. **(e,f)** For comparison, the neuronal activity averaged over the whole excitatory population is shown. All three observables were averaged over ten trials. Parameters: $n_{CA} = 150$, $I_0 = 0.15$ nA.

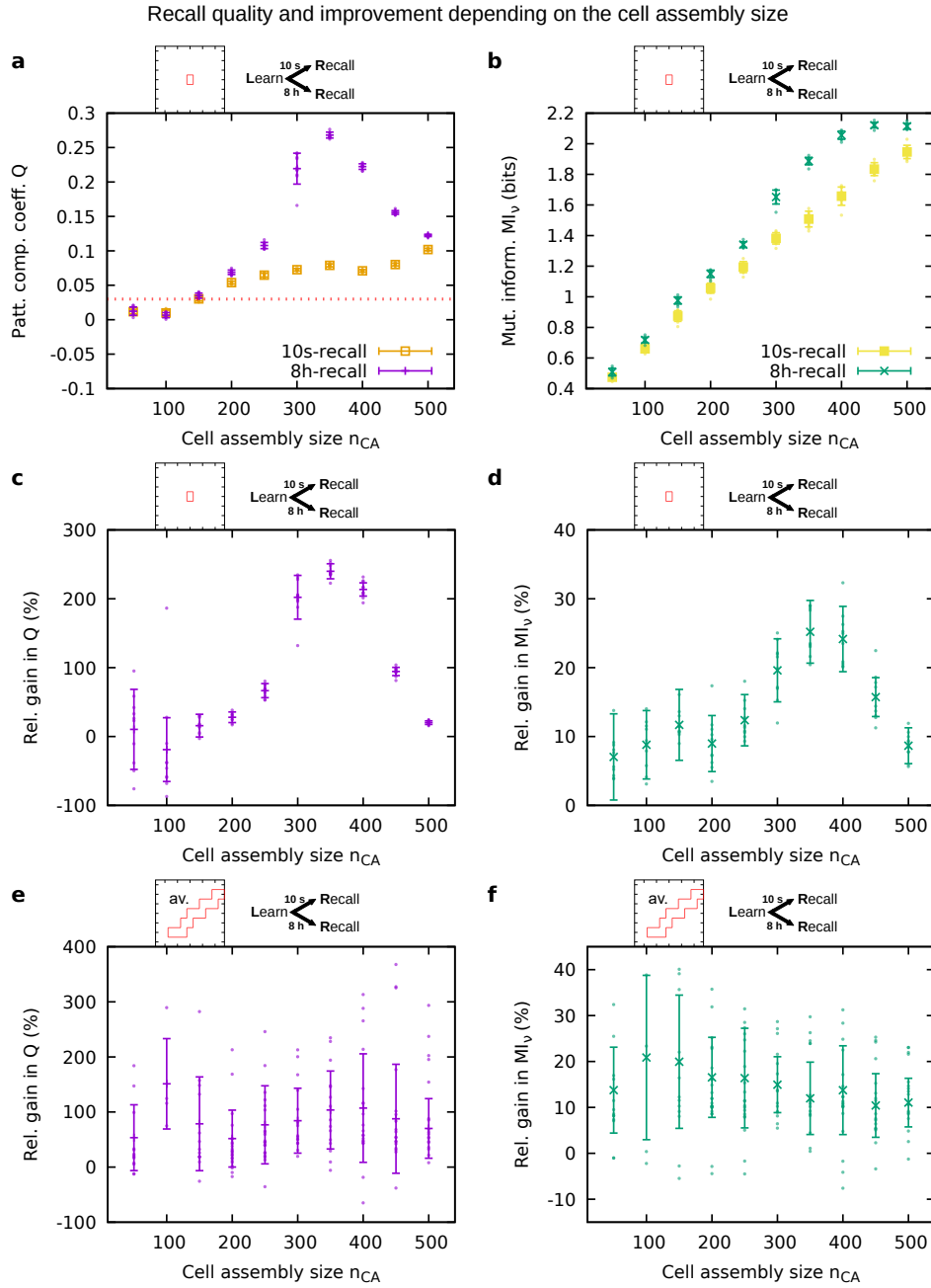


Figure 2.5: **Recall quality 10s and 8h after learning and resulting relative gain as a function of the cell assembly size.** (a) Pattern completion coefficient 10s (orange) and 8h (purple) after learning. The dotted red line indicates the threshold for memory functionality ($Q \geq 0.03$). (b) Mutual information of the neuronal activities 10s (yellow) and 8h (green) after learning. (c) Relative gain in pattern completion coefficient between 10s and 8h after learning. (d) Relative gain in mutual information of the neuronal activities between 10s and 8h after learning. (a-d) Averaged over ten trials; parameter values: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$. (e) Relative gain in pattern completion coefficient, (f) relative gain in mutual information, averaged over the mean gain values (from ten trials) of all inhibition parameter sets that yielded functional memory for the respective cell assembly size (i.e., the sets with $Q(10s\text{-recall}) \geq 0.03$, for size 150 cf. the red boxes in Fig. 2.4). Error bars indicate in (a-b) the standard deviation across ten trials, in (c,d) the error propagated from (a,b), and in (e,f) the error propagated from the standard deviations (from ten trials) of the functional parameter sets.

Synaptic weights during memory formation, consolidation and improvement

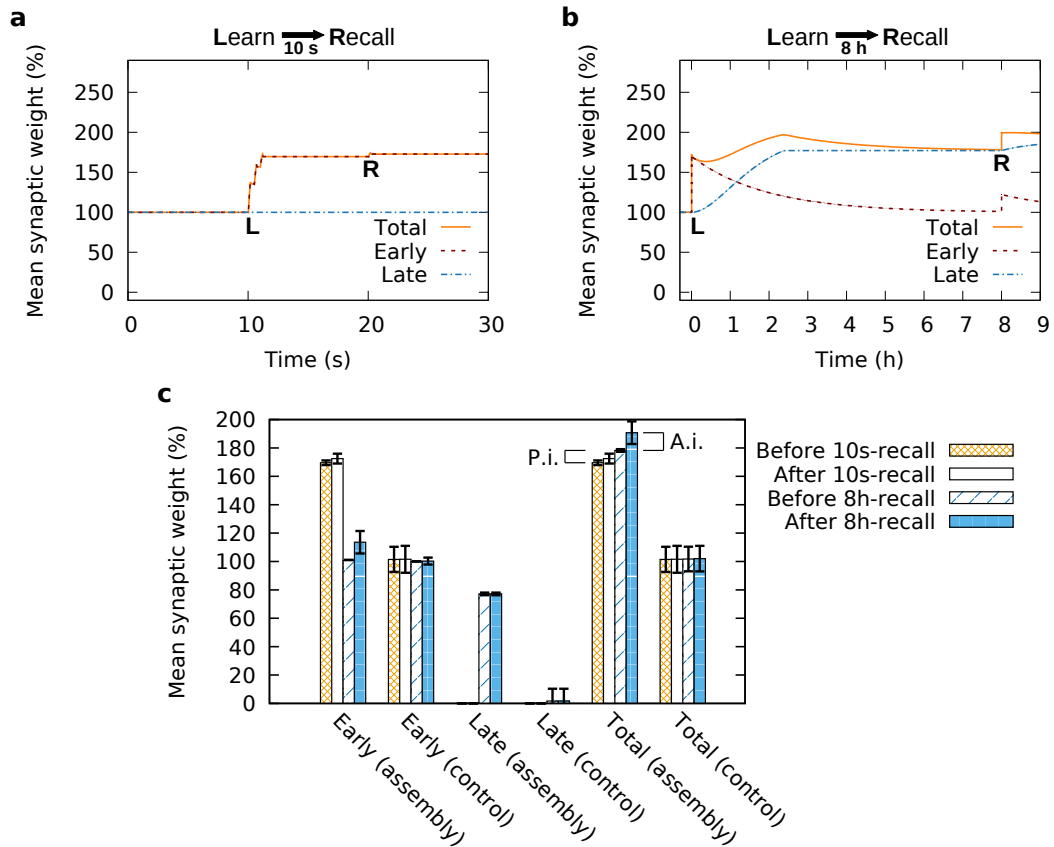


Figure 2.6: **Synaptic dynamics related to memory formation, consolidation and improvement.** (a) During the first seconds after the learning stimulus ('L'), the average early-phase weight (red) of the synapses connecting the stimulated neurons increases, while the average late-phase weight remains constant (blue). A recall stimulus ('R'), provided 10s after the learning stimulus, does not exert a substantial effect on the synaptic weights. (b) Several hours after learning, the average early-phase weight decays while the late-phase increases such that the average total synaptic weight remains on a high level. Providing a recall stimulus 8h after learning triggers an increase of the average early-phase weight and, thus, of the total synaptic weight. (a,b) Averaged over ten trials; the error bands (representing the standard deviation) are too small to be visible. (c) Mean early- and late-phase weights within the assembly and within the non-assembly (control) population immediately before and immediately after the 10s- and 8h-recall. A.i.: active improvement component (difference between solid blue and blue hatches); P.i.: passive improvement component (difference between orange grating and blue hatches). Data from one sample network. Error bars show the standard deviation across the subpopulation. See Supplementary Figs. 2.13 and 2.14 for detailed distributions of the underlying data. Parameter setting: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$, $n_{CA} = 150$.

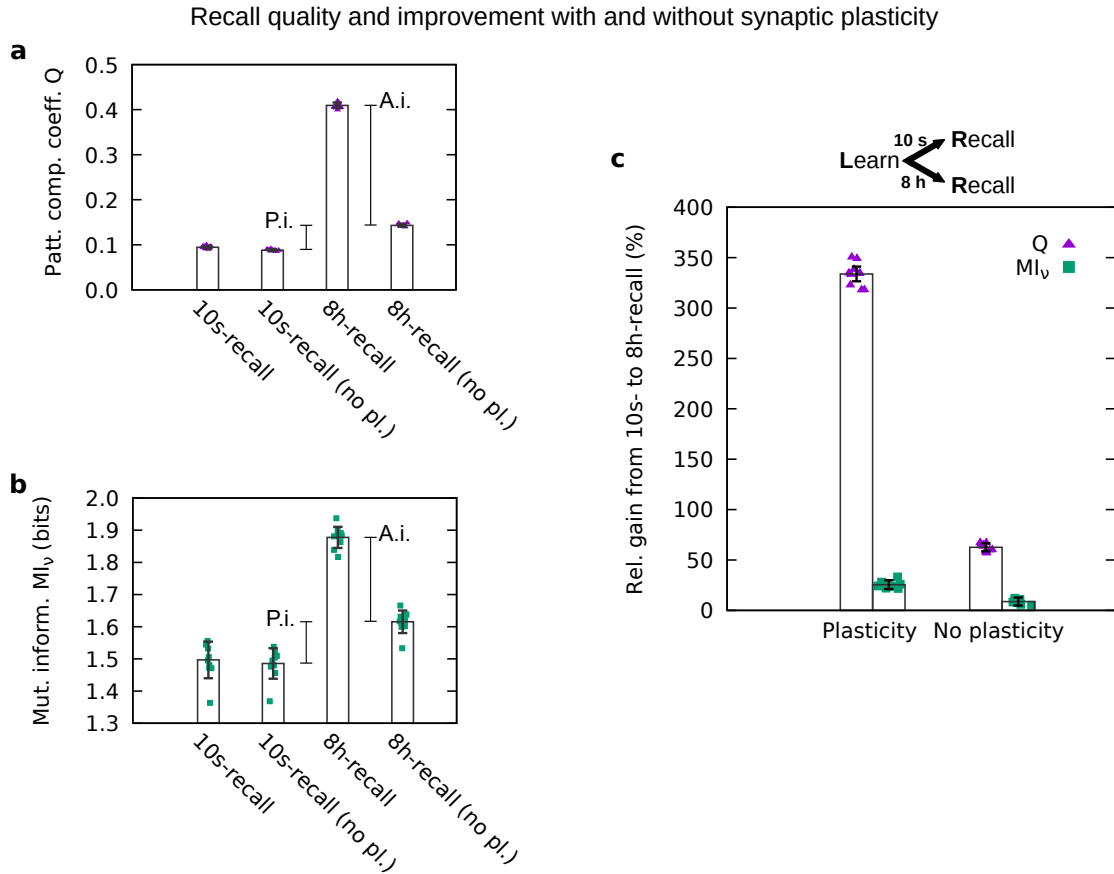


Figure 2.7: Recall quality measured for recall 10 s and 8 h after learning with and without early-phase synaptic plasticity during recall. The effect of early-phase synaptic plasticity on the total improvement is significant (compare 8h-recall with 8h-recall no pl.: A.i., active improvement), but it does not cover the complete level of improvement (compare 10s-recall with 8h-recall no pl.: P.i., passive improvement). (a) Pattern completion coefficient Q ; (b) mutual information MI_v ; (c) relative gain in Q and MI_v from 10s-recall to 8h-recall. All observables averaged across ten trials. Error bars indicate standard deviation across the ten trials in (a,b) and the propagated error in (c). Parameter setting: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$, $n_{CA} = 350$ (value of the maximum in Fig. 2.5c,d).

Analytical calculation of the long-term mean weight dynamics

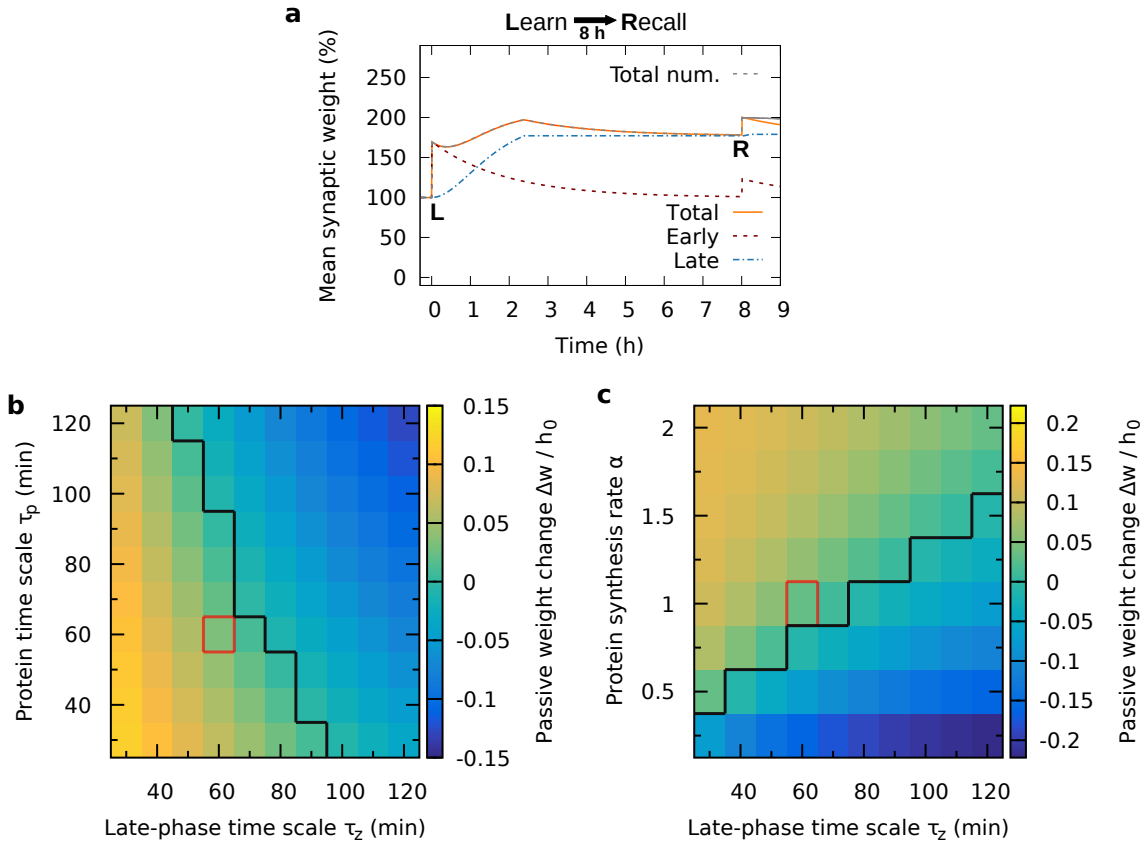


Figure 2.8: **Analytical investigation of the long-term weight dynamics and passive improvement depending on late-phase time constants.** (a) The early-phase values immediately after learning ('L') and after recall ('R') were taken from the average over numerical simulations (Fig. 2.6b) and inserted into the analytical equations for the late-phase dynamics (Eq. 2.5) and the early-phase decay (Eq. 2.3). The curves are in very good match with the simulated results, the total weight of which is shown here for comparison by the dashed gray line (cf. Fig. 2.6b). (b,c) Varying the time constants for late-phase dynamics τ_z and protein amount τ_p (b), as well as the protein synthesis rate α (c), shows that for slower time scales and lower synthesis rate, the passive weight change becomes negative, i.e., there is a passive deterioration instead of improvement. The black separatrix line demarcates the two regimes of improvement and deterioration. The red boxes indicate the parameter values used in our numerical studies.

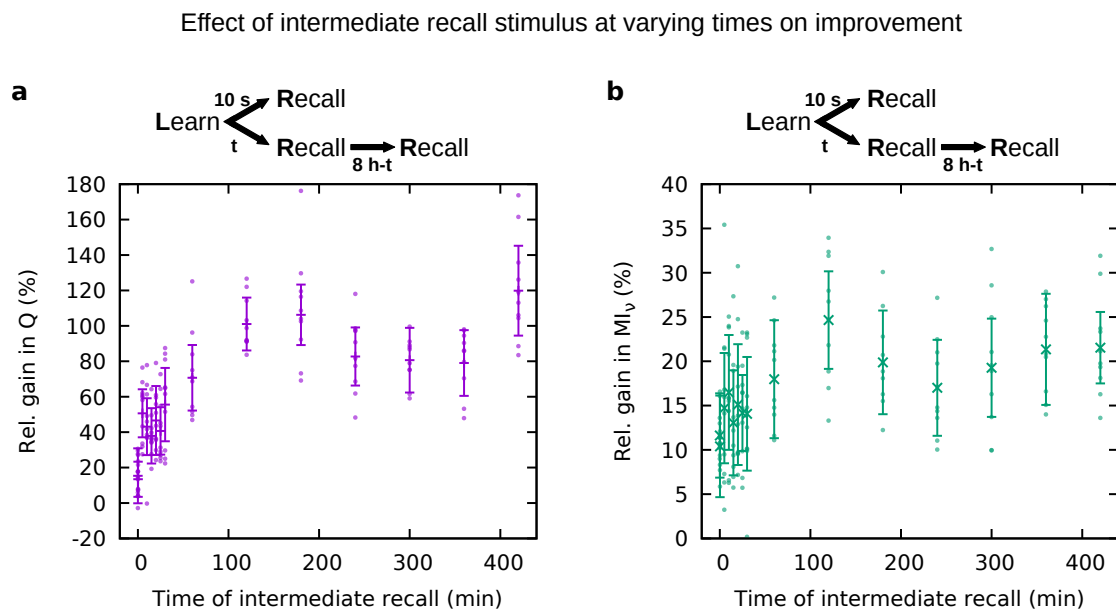


Figure 2.9: **Relative gain in recall quality between 10s and 8h after learning, affected by an intermediate recall stimulus at varying times.** (a) Gain in pattern completion coefficient Q ; (b) gain in mutual information MI_v . Parameter values: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$, $n_{CA} = 150$. Averaged over ten trials. Error bars indicate the error propagated from the standard deviations across ten trials for both 10s-recall and 8h-recall. The data points at time zero show the case without intermediate recall (as in Fig. 2.5c,d).

Discussion

In this study, we developed a recurrent network model to show that the mechanisms of STC (Frey and Morris, 1997; Clopath et al., 2008; Redondo and Morris, 2011; Li et al., 2016) provide a biological basis for encoding, consolidation, and recall of memories. We found that these mechanisms also cause an improvement in the recall quality of memories, which is robust across a large space of parameter settings. While the possibility of such improvement was proposed previously based on an abstract plasticity model featuring multiple timescales (Elliott and Lagogiannis, 2012; Elliott, 2016), we show that memory improvement can arise from the STC mechanisms. Previous theoretical studies of STC focused on molecular pathways (Smolen et al., 2012), single synapses (Clopath et al., 2008; Barrett et al., 2009; Ziegler et al., 2015; Li et al., 2016), and feed-forward networks (Ziegler et al., 2015), and thereby provided fundamental insights into the processes underlying the synaptic consolidation of memories. However, to the best of our knowledge, there have not yet been studies targeting the effects of STC in recurrent networks, which are essential for the encoding of memories. We set out to close this gap and we were able to characterize emergent effects that arise from the conjunction of synaptic tagging and capture with strongly recurrently connected groups of neurons (cell assemblies).

Detailed models describing the molecular dynamics underlying synaptic plasticity and STC have been shown to be compatible to and to offer a complementary view on simpler statistical models (Clopath et al., 2008; Barrett et al., 2009; Graupner and Brunel, 2010, 2012; Smolen et al., 2012). We use a model for synaptic plasticity based on calcium dynamics because this approach captures a wide range of experimental findings, including spike-timing dependent phenomena and rate-coding behavior (Bi and Poo, 1998; Shouval et al., 2002; O'Connor et al., 2005; Wang et al., 2005; Nevian and Sakmann, 2006; Sjöström and Häusser, 2006; Letzkus et al., 2006; Wittenberg and Wang, 2006; Graupner and Brunel, 2012). For STC, we use a model based on generalized plasticity-related proteins that account for the crucial aspects of tagging, capture and cross-tagging (Sjöström et al., 2001; Reymann and Frey, 2007; Sajikumar et al., 2007; Clopath et al., 2008; Redondo and Morris, 2011; Li et al., 2016). Thereby, our model contains the essential features of a biologically more detailed model, while still being computationally tractable.

By varying the size of the cell assembly, we found different recall dynamics of the system. We found that there is a lower bound to the size, below which the amplification of recall stimulation does not suffice for functional recall. On the other hand, large cell assemblies show attractor behavior, meaning that after the application of a recall stimulus they become fully activated and stay activated. Investigating this attractor regime is beyond this study. Nevertheless, considering attractor dynamics in our model could reveal interesting implications, since long-term memory representations exhibiting attractor dynamics can serve as working memory (O'Reilly et al., 1999; Ranganath et al., 2005), possibly in conjunction with additional transient dynamics (Nachstedt and Tetzlaff, 2017). Varying the inhibitory synaptic weights, we found a regime of parameters yielding functional memory dynamics. This regime seems to correspond to 'loosely-balanced' network states, which are characterized by a dynamic equilibrium of excitation and inhibition (Denève and Machens, 2016). In contrast to that, 'tightly-balanced' network states feature inhibition that closely follows excitation, and in some cases seem to enable more efficient coding. By introducing inhibitory plasticity (Vogels et al., 2011; Sprekeler,

2017), our network could possibly enter a tightly-balanced state. This would increase the complexity of our model tremendously, but could be subject to future studies on more efficient memory encoding.

Synaptic plasticity such as long-term potentiation is widely considered to be the main mechanism that underlies memory, as it is expressed by the synaptic plasticity and memory hypothesis (Martin et al., 2000; Abraham et al., 2019). Long-term potentiation of synapses (LTP) has been subdivided into three different types: LTP1, which depends on stimulation only, LTP2, which depends on translation of existent mRNA, and LTP3, which depends on transcription in the soma (Racine et al., 1983; Abraham et al., 2002, 2019). LTP2 and LTP3 are often subsumed as late-phase LTP, whereas LTP1 is called early-phase LTP (Frey and Morris, 1997; Abraham et al., 2019), which paradigm is followed by our study. Similar phenomena and processes as the ones discussed for LTP are found for long-term depression (LTD; Sajikumar et al., 2007; Smolen et al., 2012; Abraham et al., 2019). In addition to synaptic plasticity, there are hints to other, cell-intrinsic mechanisms that might be relevant for the storage of memory (Abraham et al., 2019), such as non-coding RNAs that produce learning-related epigenetic changes in neurons (Landry et al., 2013). However, the potential effects of such non-synaptic memory mechanisms on the results of our study and on memory consolidation in general remain unknown and require further investigations.

Following their encoding, memories can become consolidated to last for hours, days, or even years, depending on conditions like strength of the learning stimulus and neuromodulation (McGaugh, 2000; Abraham et al., 2002; Dudai, 2004). In principal, one can distinguish two different paradigms related to memory consolidation. On the level of brain areas, systems consolidation describes the transfer of newly formed memories, mainly from the hippocampus to the neocortex (McGaugh, 2000; Dudai, 2004), as it was first discovered in the case of the patient H.M. (Scoville and Milner, 1957). This type of consolidation is related to sleep and rest (Rasch and Born, 2013; Tetzlaff et al., 2013; Mizuseki and Miyawaki, 2017). On the other hand, there is synaptic consolidation, which denotes the in-place stabilization of changes in the strength of synaptic connections (McGaugh, 2000; Dudai et al., 2015), and thereby is a synonym for LTP2 and LTP3, or late-phase LTP as discussed before (Abraham et al., 2019). Blocking of hippocampal LTP has been shown to prevent long-term memory during wake in the first hours after learning (Scoville and Milner, 1957; Seidenbecher et al., 1997), which is why the hippocampus seems to be the central region for this first stage of memory consolidation. In our study, we do not consider sleep- or rest-dependent consolidation, which would require the presence of activity patterns like shape-wave/ripples, spindles or slow oscillations (Mizuseki and Buzsáki, 2013; Dudai et al., 2015; Mizuseki and Miyawaki, 2017). Instead, we target the encoding and the in-place synaptic consolidation of memories. This occurs for instance during exploratory wake state in hippocampal networks (Seidenbecher et al., 1997; McGaugh, 2000; Dudai et al., 2015). Although our model resembles some characteristics of hippocampal networks, it could describe a network in any other brain region that features similar processes. Moreover, extending our model by sleep-related activities (Tetzlaff et al., 2013; Jahnke et al., 2015) could yield interesting insights into the interplay between synaptic and systems consolidation.

Sometimes it is being suggested to use the term cellular consolidation instead of synaptic consolida-

tion (Dudai, 2004; Redondo and Morris, 2011), since the processes involved are not only located at the synaptic site. This is mostly due to the fact that mRNA and proteins are synthesized in the soma, from where they are transported to the dendrites and spines (Abraham et al., 2002; Dudai, 2004; Clopath et al., 2008; Redondo and Morris, 2011). Nevertheless, there can also be local protein synthesis in the dendrites (Martin et al., 1997; Jiang and Schuman, 2002; Bramham, 2008; O'Donnell and Sejnowski, 2014), which could confine the here identified protein-dependent dynamics of consolidation and improvement to specific dendritic branches of a neuron. However, due to synaptic clustering (Kastellakis and Poirazi, 2019), local protein synthesis would not necessarily yield different results as compared to our model.

The most important finding of our study is that for many parameter settings, the quality of recall 8 h after learning is much better than the quality of recall 10 s after learning. In psychological literature, such an increase in recall quality over time is referred to as memory improvement or hypermnesia (Erdelyi, 2010). It has been found particularly, but not exclusively, in picture memory (Payne, 1987; Erdelyi, 2010; Wallner and Bäuml, 2018). In our simulations and analyses, we found the improvement effect to occur already at the very first recall attempt, while in psychological experiments the effect of improvement has mostly been measured over the course of several recall attempts. Our model accounts as well for improvement by multiple recalls, as we showed using a protocol with an intermediate recall stimulus. Nevertheless, there are experiments indicating that recall can already be improved in the first attempt (Erdelyi, 2010). It would be desirable to have more experimental data on very early recall as well as on recall after 8 h without previous recall.

Our model explains another interesting, seemingly fundamental feature of memory consolidation. As studies on positive effects of memory disuse have proposed (Bjork and Bjork, 1992; Karpicke et al., 2014), learning will be most powerful after a break, when the storage strength is still quite low and the retrieval strength has again decreased. Attempts to retrieve will then result in a strong increase in storage strength. If the retrieval strength is still high, the gain in storage strength upon this ('easy') retrieval will be low. This theory seems to be partially consistent with our results, considering that the storage strength correspond to the late-phase synaptic weight, and retrieval strength to the early-phase synaptic weight. In our model, increases in early-phase weights upon stimulation are much larger if the early-phase weight is low before the stimulation than if it is high. Since increases in early-phase weight lead to changes in late-phase weight, the gain in late-phase weight (i.e., presumably, the storage strength) indeed inversely depends on the early-phase weight (retrieval strength) before stimulation. Nevertheless, further investigations would have to resolve how our model could match the psychological finding that retrieval at small retrieval strength should take long (Bjork and Bjork, 1992; Karpicke et al., 2014), while at the present our model exhibits fast recall even in the case of low early-phase weights.

Behavioral tagging is a phenomenon that could be the behavioral analog of synaptic tagging, describing the finding that the memory of a weak input perceived around the same time as strong novel stimulation becomes consolidated along with the memory of the strong stimulus (Moncada and Viola, 2007; Ballarini et al., 2009; Moncada et al., 2015). While by now there is plenty of evidence relating

behavioral tagging to synaptic tagging, the knowledge gap is not yet closed (Okuda et al., 2020). Using their theoretical model of a feed-forward network, Ziegler and colleagues (Ziegler et al., 2015) provided a connection between STC mechanisms and behavioral tagging focusing on an inhibitory avoidance paradigm. For this, they considered dopamine modulation to model novelty. By extending our model by such dopamine modulation, for instance via the protein synthesis threshold, more experimental data could be matched to provide further evidence for the relatedness of behavioral tagging and synaptic tagging.

Our model predicts an active and a passive improvement of memories by the synaptic tagging and capture mechanisms for a wide range of parameter values. In the following, we discuss how these predictions can be verified in experiments on the behavioral level, on the network level, and on the synaptic level. At the behavioral or psychological level, our findings can be tested with subjects that need to learn a set of items, like words or pictures, and then recall the items after a very short time (~ 10 s). The performance of this recall should then be compared with the performance of the same subjects in a similar task but with other items, and recall after 8 h. Ideally, to be comparable to our results, there should not be any attempt to recall during the 8 h, which could be hard to manage for the subjects. However, at least trying to avoid intermediate recall would provide some new insights as compared to previous psychological studies that were based on intermediate recall (Erdelyi and Kleinbard, 1978; Payne, 1987; Wallner and Bäuml, 2018). In in-vitro networks, multielectrode arrays (MEAs) can be used to stimulate many neurons at the same time, as well as to measure their activity (Liu et al., 2012a). For our purposes, an MEA could be used to trigger plasticity, forming a strongly interconnected cell assembly. After the formation, a partial recall stimulus should be applied and the resulting activity be measured. The activity during 10s-recall (10 s after learning) and during 8h-recall (8 h after learning) can then be compared as it was shown in Fig. 2.3e. If the activity in the cell assembly following 8h-recall is higher, then there will be experimental evidence for memory improvement. Another promising approach to test memory improvement at the network level would be to use optogenetic stimulation in combination with activity measurements through calcium imaging, which is an established method that is very suitable for in-vivo experiments (Akerboom et al., 2013; Stamatakis et al., 2018), and can be tuned to deliver precise learning and recall stimulation (Chettih and Harvey, 2019; Russell et al., 2019; Luboinski and Tchumatchenko, 2020). Our study also yields predictions for the synaptic level. To test these, first, synaptic potentiation has to be induced applying a strong ‘learning’ stimulus to a nerve fiber. Later application of a shorter ‘recall’ stimulus to the fiber should then cause a response that depends on the time that has passed since learning. Following our predictions, the response, which can be measured as a relative field excitatory postsynaptic potential (often abbreviated “%fEPSP”), should be larger 8 h after learning than 10 s after learning, corresponding to the mean total synaptic weight shown in Fig. 2.6. Furthermore, after the application of two ‘recall’ stimuli, the response should even be higher, as we showed in the section on intermediate recall. Finally, in any experiment, blocking the dynamics that lead to early-phase plasticity should diminish the improvement effect significantly, because we predict that active improvement is triggered by the early-phase dynamics (cf. Fig. 2.7c). Such experiments could be realized by using NMDA receptor antagonists or by blocking exocytosis before the presentation of a recall stimulus. In addition, blocking or slowing down late-phase-related signaling cascades and protein dynamics should, beyond a certain

level, prevent passive improvement and thereby also diminish the total improvement effect (cf. Fig. 2.8b,c).

Please note that for several of the potential extensions of our model discussed above, it would be useful to improve the computational efficiency of the numerical simulation. For that, it is very promising to implement our network model on neuromorphic hardware (Davies et al., 2018; Lin et al., 2018). Neuromorphic hardware enables the simulation of larger networks and thereby the storage of many cell assemblies. In this way, storage capacity and the interference between cell assemblies could be investigated. The neuromorphic approach would not only facilitate scientific investigations but, due to its good performance and energy efficiency (Wang et al., 2017; Lin et al., 2018), also offer great opportunities for the technical application of our model. Implemented on neuromorphic hardware, our model could be used for fast, energy-efficient self-improvement of information storage after a few presentations. In conclusion, our theoretical model provides a further piece of evidence that the STC mechanisms are essential for memory dynamics, and it predicts that these STC mechanisms also allow the improvement of memories, which can be beneficial for biological as well as artificial memory systems.

Methods

Model

To simulate the dynamics of memory representations, we use a network model with hippocampal characteristics that comprises spiking neurons and synapses with detailed plasticity features. In the following, we first present our mathematical description of neurons and synapses, which is depicted in Fig. 2.1a. After that, we explain how our network is structured at the population level. The parameters we use are given in Tables 2.1 and 2.2.

The dynamics of the membrane potential $V_i(t)$ of the leaky integrate-and-fire neuron i is described by (Gerstner and Kistler, 2002):

$$\tau_{\text{mem}} \frac{dV_i(t)}{dt} = V_{\text{rev}} - V_i(t) + R \sum_j \sum_{t_j^k} w_{ji}/s \cdot \exp\left(- (t - t_j^k - t_{\text{ax,delay}}) / \tau_{\text{syn}}\right) + R \left(I_{\text{bg}}(t) + I_{\text{stim}}(t) \right) \quad (2.8)$$

with reversal potential V_{rev} , membrane time constant τ_{mem} , membrane resistance R , synaptic weights w_{ji} , spike times t_j^k , axonal delay time $t_{\text{ax,delay}}$, synaptic time constant τ_{syn} , external background current $I_{\text{bg}}(t)$, and external stimulus current $I_{\text{stim}}(t)$. If V_i crosses the threshold V_{th} , a spike is generated. The spike time t_i^n is then stored and the membrane potential is reset to V_{reset} , where it remains for the refractory period t_{ref} . Apart from learning and recall stimulation, the membrane potential dynamics is mainly driven by a background noise current that accounts for synaptic inputs from outside the network, described by an Ornstein-Uhlenbeck process:

$$\tau_{\text{syn}} \frac{dI_{\text{bg}}(t)}{dt} = I_0 - I_{\text{bg}}(t) + \sigma_{\text{wn}} \cdot \Gamma(t) \quad (2.9)$$

with mean current I_0 and white-noise standard deviation σ_{wn} . Note that in this equation, $\Gamma(t)$ is Gaussian white noise with mean zero and variance $1/dt$ that approaches infinity for $dt \rightarrow 0$ (Gillespie, 1996). The Ornstein-Uhlenbeck process has the same colored-noise power spectrum as the fluctuating input to cortical neurons coming from a large presynaptic population (Destexhe et al., 2003). Therefore, it is well-suited to model background noise in our model. In addition to the background noise, a second Ornstein-Uhlenbeck process is used to model the stimulus current $I_{\text{stim}}(t)$ for learning and recall, which is described in the subsection ‘Learning and recall procedure’ below.

If there is a synaptic connection from some neuron j to neuron i , all spikes k that occur in j are transmitted to i . The postsynaptic current caused by a presynaptic spike depends on the weight of the synapse. The total weight or strength of a synaptic connection from neuron j to neuron i is given by:

$$w_{ji} = \begin{cases} h_{ji} + h_0 \cdot z_{ji} & \text{for E} \rightarrow \text{E}, \\ w_{ei} & \text{for E} \rightarrow \text{I}, \\ w_{ie} & \text{for I} \rightarrow \text{E}, \\ w_{ii} & \text{for I} \rightarrow \text{I}, \end{cases} \quad (2.10)$$

where E and I stand for excitatory and inhibitory neurons, respectively. Thus, in our model, all synaptic connections involving inhibitory neurons are constant. The weight of E→E connections, however, consists of two variable contributions providing the core of the STC mechanisms: the early-phase weight h_{ji} , and the late-phase weight z_{ji} . We followed (Li et al., 2016) and used h_0 as the normalization factor for z to obtain a quantity of the same dimension as h . The value of the normalization factor is in accordance with experiments investigating late-phase LTP/LTD.

The dynamics of the early-phase weight is given by

$$\tau_h \frac{dh_{ji}(t)}{dt} = 0.1 (h_0 - h_{ji}(t)) + \gamma_p (1 \text{ nC} - h_{ji}(t)) \cdot \Theta[c_{ji}(t) - \theta_p] - \gamma_d h_{ji}(t) \cdot \Theta[c_{ji}(t) - \theta_d] + \zeta(t), \quad (2.11)$$

where $\Theta[\cdot]$ is the Heaviside function, τ_h is a time constant, and $c_{ji}(t)$ is the calcium concentration at the postsynaptic site. The first term on the right-hand side describes a relaxation of the early-phase weight back to its initial value h_0 , the second term describes early-phase LTP with rate γ_p for calcium above the threshold θ_p , the third term describes early-phase LTD with rate γ_d for calcium above the threshold θ_d , and the fourth term $\zeta(t) = \sqrt{\tau_h [\Theta(c_{ji}(t) - \theta_p) + \Theta(c_{ji}(t) - \theta_d)]} \sigma_{\text{pl}} \Gamma(t)$ describes calcium-dependent noise-driven fluctuations with standard deviation σ_{pl} and Gaussian white noise $\Gamma(t)$ with mean zero and variance $1/dt$. The calcium concentration $c_{ji}(t)$ at the postsynaptic site depends on all past pre- and postsynaptic spikes n and m , respectively:

$$\frac{dc_{ji}(t)}{dt} = -\frac{c_{ji}(t)}{\tau_c} + c_{\text{pre}} \sum_n \delta(t - t_j^n - t_{c,\text{delay}}) + c_{\text{post}} \sum_m \delta(t - t_i^m), \quad (2.12)$$

where $\delta(\cdot)$ is the Dirac delta distribution, τ_c is a time constant, c_{pre} is the contribution of presynaptic spikes, c_{post} is the contribution of postsynaptic spikes, t_j^n and t_i^m are spike times, and $t_{c,\text{delay}}$ is the delay

of calcium triggered by presynaptic spikes.

The calcium-based plasticity model (Eqs. 2.11 and 2.12) that we use to describe the early phase of long-term potentiation and depression is based on previous theoretical studies (Graupner and Brunel, 2012; Higgins et al., 2014; Li et al., 2016). Similar to Li et al. (2016), the first term on the right-hand side of Eq. 2.11 describes a relaxation to the initial condition to which the early-phase weight returns or decays on a timescale of a few hours. This decay accounts for the fact that early-phase changes are transient providing an accurate description of the experimentally verified dynamics of the STC mechanisms (Frey and Morris, 1997).

The calcium parameters provided by Graupner and Brunel (2012) were obtained by fitting experimental data from in vitro experiments (Wittenberg and Wang, 2006). Since extracellular calcium concentrations are much lower in vivo than in vitro, the parameters need to be corrected for modeling in vivo networks. Following Higgins et al. (2014), the calcium influx into the postsynaptic spine can be assumed to decrease proportionally to the ratio of in vivo and in vitro extracellular calcium concentrations, which leads to a factor of 0.6. Therefore, in our network model we adjust the values provided by Graupner and Brunel (2012) by this factor.

Driven by the calcium-based early-phase dynamics, the late-phase synaptic weight is given by

$$\begin{aligned} \tau_z \frac{dz_{ji}(t)}{dt} = & p_i(t) \cdot (1 - z_{ji}(t)) \cdot \Theta[(h_{ji}(t) - h_0) - \theta_{\text{tag}}] \\ & - p_i(t) \cdot (z_{ji} + 0.5) \cdot \Theta[(h_0 - h_{ji}(t)) - \theta_{\text{tag}}], \end{aligned} \quad (2.13)$$

with the protein amount $p_i(t)$, the late-phase time constant τ_z , and the tagging threshold θ_{tag} . The first term on the right-hand side describes late-phase LTP while the second term describes late-phase LTD. Both depend on the amount of proteins being available. If the early-phase weight change $|h_{ji}(t) - h_0|$ exceeds the tagging threshold, the synapse is tagged. This can be the case either for positive or for negative weight changes. The presence of the tag leads to the capture of proteins (if $p_i(t) > 0$), and thereby gives rise to changes in the late-phase weight.

The synthesis of new proteins depends on the early-phase weights, but the amount of proteins also inherently decays exponentially (Clopath et al., 2008):

$$\tau_p \frac{dp_i(t)}{dt} = -p_i(t) + \alpha \Theta \left[\left(\sum_j |h_{ji}(t) - h_0| \right) - \theta_{\text{pro}} \right]. \quad (2.14)$$

Using the neuron model and the synapse model explained above, we set up a neural network consisting of 1600 excitatory and 400 inhibitory neurons (depicted in Fig. 2.1b). This ratio between excitatory and inhibitory neurons is commonly used for cortical and hippocampal networks (Braitenberg and Schüz, 1998). Some of the excitatory neurons receive specific inputs to learn and recall a memory representation (see next subsection). As described before, only the synapses between excitatory neurons are plastic. The inhibitory population serves to provide realistic feedback inhibition. The overall connectivity across both populations is 10%, meaning that there is a probability of 0.1 that the link

from any neuron to another one in the whole network exists. The value is reasonable for hippocampal region CA3 (Le Duigou et al., 2014).

Learning and recall procedure

Before we stimulate our network, we first let the initial activity settle for 10.0 s. After that, we apply our learning protocol, which delivers three stimulus pulses, of 0.1 s duration each, to the neurons belonging to the desired cell assembly (for instance, to the first 150 neurons in the network). The pulses in our protocol are separated by breaks of 0.4 s. During the pulses, stimulation $I_{\text{stim}}(t)$ enters the neuronal membrane potential given in Eq. 2.8. The stimulation is modeled by the following Ornstein-Uhlenbeck process (also cf. Eq. 2.9):

$$\tau_{\text{syn}} \frac{dI_{\text{stim}}(t)}{dt} = w_{\text{stim}} \cdot N_{\text{stim}} \cdot f_{\text{stim}} - I_{\text{stim}}(t) + w_{\text{stim}} \cdot \sqrt{N_{\text{stim}} \cdot f_{\text{stim}}} \cdot \Gamma(t), \quad (2.15)$$

with Gaussian white noise $\Gamma(t)$ and the synaptic time constant τ_{syn} . Mean and standard deviation of the process are defined by putative input spikes from N_{stim} neurons, occurring at the frequency f_{stim} and conveyed by synapses of weight w_{stim} to the selected neurons of our network (Gillespie, 1996; Dayan and Abbott, 2001). For the reproduction of previous single-synapse investigations (Li et al., 2016; Fig. 2.2), we stimulated the presynaptic neuron with Poisson spikes that were generated using the same number of input neurons N_{stim} and the same frequency f_{stim} as for learning a cell assembly, conveyed with the same synaptic strength w_{stim} .

After 20.0 s, we save the state of the whole network and then apply a recall stimulus, equally modeled by the process given in Eq. 2.15 with the same parameters as the learning stimulus (except for Supplementary Fig. 2.16), for 0.1 s to half of the neurons in the cell assembly (regarding the example above, to 75 randomly-drawn neurons). We refer to this as “10s-recall”. Next, we load the previously saved network state again, such that the network is back in the state where it was immediately before recall. This time, we let the network run for 28810.0 s until we apply the recall stimulus, which we refer to as “8h-recall”. For one part of our study, we apply a further, intermediate, recall stimulus at a time in between 10 s and 8 h. Following this intermediate recall, we do not reset the network by loading an earlier state, such that it in fact affects the later 8h-recall.

Measures of recall performance

To investigate the effects of recall stimulation, we divide the excitatory population into three subpopulations: control neurons that are not directly stimulated neither by learning nor by recall stimulation, cell assembly neurons that are stimulated by both learning and recall stimulus, and cell assembly neurons that are stimulated during learning but not during recall (cf. Fig. 2.3d). The mean activities of these three subpopulations are given by:

- \bar{v}_{as} : mean activity of the neurons stimulated by the learning stimulus and the recall stimulus

as = “assembly, stimulated”,

- \bar{v}_{ans} : mean activity of the neurons stimulated by the learning stimulus but not by the recall stimulus

ans = “assembly, not stimulated”,

- \bar{v}_{ctrl} : mean activity of the neurons stimulated by neither of the two stimuli

ctrl = “control”.

Based on these mean activities, the recall quality can be measured by computing a coefficient for the quality of pattern completion:

$$Q^* := \frac{\bar{v}_{\text{ans}} - \bar{v}_{\text{ctrl}}}{\bar{v}_{\text{as}}}. \quad (2.16)$$

The coefficient typically lies within the interval $(0, 1)$. To achieve pattern completion of the learning-defined cell assembly, the non-stimulated assembly neurons have to be indirectly activated following the activation of the rest of the core neurons, while control neurons should remain relatively silent:

$$\bar{v}_{\text{ctrl}} \ll \bar{v}_{\text{ans}} \leq \bar{v}_{\text{as}}. \quad (2.17)$$

Hence, for good pattern completion or good recall, the value of Q^* will be significantly larger than zero or even approach unity. On the other hand, if it approaches zero, there is either no pattern completion, which means that $\bar{v}_{\text{ctrl}} \approx \bar{v}_{\text{ans}}$, or the network activity diverges, which means that $\bar{v}_{\text{ctrl}} \approx \bar{v}_{\text{ans}} \approx \bar{v}_{\text{as}}$.

In addition to the subpopulation-based quantity Q , we measure the mutual information MI_v between the activity distribution during the recall phase and the activity distribution during the learning phase. The mutual information does not directly relate to pattern completion, but it has the advantage that it is independent of any predefined patterns such as the learning stimulus.

The mutual information of the activity distribution is calculated from the entropy during the learning phase at time $t_{\text{learn}} = 11.0$ s, the entropy during the recall phase at time $t_{\text{recall}} \in \{20.1 \text{ s}, 28810.1 \text{ s}\}$, and the joint entropy between both:

$$MI_v := H(v(t = t_{\text{learn}}, n)) + H(v(t = t_{\text{recall}}, n)) - H(v(t = t_{\text{learn}}, n), v(t = t_{\text{recall}}, n)). \quad (2.18)$$

The firing rate function $v(t, n)$ returns the firing rate of a given neuron n at a given time t , computed using a sliding window of 0.5 s.

Statistics and reproducibility

The pattern completion effect is not necessarily robust across trials for every parameter setting, even though $Q^* \gg 0$ may be given for a single trial. Therefore, we average over trials:

$$Q := \langle Q^* \rangle. \quad (2.19)$$

For Fig. 2.4, due to the lack of error bars, we have to indicate the cases in which no robust pattern completion occurs. Thus, we use the following conditional definition for the pattern completion coefficient:

$$Q := \begin{cases} \langle Q^* \rangle & \text{if } \langle Q^* \rangle > \Delta Q, \\ 0 & \text{else.} \end{cases} \quad (2.20)$$

Put in words, the robustness criterion requires the mean $\langle Q^* \rangle$ to be non-negative, and its absolute value to be larger than the standard deviation ΔQ . If this is not fulfilled, pattern completion is assumed absent and Q is set to zero.

To achieve equal statistics for our results, we also average the mutual information value over multiple trials and use the standard deviation as error.

The relative gain in recall quality after 8 h as compared to recall quality after 10 s is computed by $\frac{Q(8h) - Q(10s)}{Q(10s)}$ for the coefficient Q , and $\frac{\langle MI_V(8h) \rangle - \langle MI_V(10s) \rangle}{\langle MI_V(10s) \rangle}$ for the mutual information. The error of the relative gain is computed by propagation of error from the recall quality at 10 s and at 8 h, such that it is $\sqrt{\left(\frac{\Delta Q(10s) \cdot Q(8h)}{Q(10s)^2}\right)^2 + \left(\frac{\Delta Q(8h)}{Q(10s)}\right)^2}$ for Q and $\sqrt{\left(\frac{\Delta MI_V(10s) \cdot \langle MI_V(8h) \rangle}{\langle MI_V(10s) \rangle^2}\right)^2 + \left(\frac{\Delta MI_V(8h)}{\langle MI_V(10s) \rangle}\right)^2}$ for the mutual information.

Our approach to determine significant effects is to compute and compare standard deviations of observables from at least 10 trials. The number of simulation trials (100 for single-synapse simulations, 10 for learning and recall simulations) was chosen such that the effects that we describe would be significant with $p < 0.006$ in a t-test. This is the case for a single observable being greater than zero if zero is not within the standard deviation of the observable (one-sided test), and also for the inequality of two observables if their standard deviations do not overlap (two-sided test, cf. Payton et al., 2003), because for the chosen number of trials the standard deviation is greater than the 98.8% confidence interval.

The network structure that we used to obtain the results presented here is provided with our code (Luboeinski, 2021). The code automatically generates a different network structure with the same connectivity p_c if no pre-defined structure is used. We did not encounter unexpected deviations with such different network structures.

Computational implementation and software used

We used C++ in the ISO 2011 standard to implement our simulations. To compile and link the code, we employed g++ in version 7.4.0 with boost in version 1.65.1.

Random numbers were generated using the generator ‘minstd_rand0’ from the C++ standard library, while the system time served as the seed. We implemented a loop in our code which ensured that for each distribution a unique seed was used.

For the creation of plots we used gnuplot 5.0.3, as well as Matplotlib 2.0.0 with Python 3.6.8 and NumPy 1.16.4.

The network simulations that we needed to perform for this study were computationally extremely demanding. Fortunately, we had the opportunity to use the computing cluster of the *Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)*, which enables fast computation on a large set of processing units. However, despite this strong computational power and the usage of compiled C++ code, running our spiking network simulations in full detail still took unacceptably long. Thus, to be able to simulate our network faster, we used an approximation that neglects the spiking dynamics in periods without external stimulation. In these periods, we just computed the late-phase dynamics and the exponential decay of the early-phase weights. Supplementary Fig. 2.11 shows that this approach is justified because the weight dynamics of a synapse does not change if sparsely occurring spikes are neglected. Furthermore, the mutual information conveyed by 8h-recall is in the same regime for full and approximating computation.

Additional information

Data availability

The data underlying the results presented in this study can be reproduced using the simulation code and the analysis scripts that we have released (Luboeinski, 2021). Furthermore, we provide the source data for the Figures in this article in Supplementary Data 1 and 2.

Code availability

Our simulation code, along with scripts to analyze the data, has been released with a digital object identifier (DOI) and can be retrieved freely (Luboeinski, 2021).

Acknowledgments

We would like to thank the members of the Department of Computational Neuroscience for many fruitful discussions and various helpful comments on this study. The research was funded by the German Research Foundation (CRC1286, project C1, project #419866478) and by the H2020 - FETPROACT project Plan4Act (#732266). Open Access funding enabled and organized by Projekt DEAL.

Author contributions

Jannik Luboeinski: conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing and editing.

Christian Tetzlaff: conceptualization; funding acquisition; investigation; methodology; project administration; resources; supervision; writing and editing.

Competing interests

The authors declare no competing interests.

Parameters for neuron and static network dynamics

Symbol	Value	Description	Refs.
Δt	0.2 ms	Duration of one time step for numerical computation	this study
τ_{mem}	10 ms	Membrane time constant	Dayan and Abbott (2001); Gerstner and Kistler (2002)
τ_{syn}	5 ms	Synaptic time constant, also for external background current	Dayan and Abbott (2001); Gerstner and Kistler (2002); Roth and van Rossum (2009)
$t_{\text{ax,delay}}$	3 ms	Axonal spike delay	Gerstner and Kistler (2002); Lin and Faber (2002)
t_{ref}	2 ms	Duration of the refractory period	Dayan and Abbott (2001); Kobayashi et al. (2009)
R	10 M Ω	Membrane resistance	Dayan and Abbott (2001)
V_{rev}	-65 mV	Reversal (equilibrium) potential	Dayan and Abbott (2001)
V_{reset}	-70 mV	Reset potential	Dayan and Abbott (2001)
V_{th}	-55 mV	Threshold potential to be crossed for spiking	Dayan and Abbott (2001)
σ_{wn}	0.05 nA s ^{1/2}	Standard deviation for Gaussian noise in external background current	this study
I_0	0.15 nA	Mean of the external background current	this study
N_e	1600	Number of neurons in the excitatory population	this study
N_i	400	Number of neurons in the inhibitory population	this study
p_c	0.1	Probability of a connection existing between two neurons	Le Duigou et al. (2014)
h_0	0.420075 nC	Initial excitatory→excitatory coupling strength	Li et al. (2016)
w_{ei}	$2h_0$	Excitatory→inhibitory coupling strength	this study
w_{stim}	h_0	Coupling strength of synapses from putative input neurons	this study
N_{stim}	25	Number of putative input neurons for stimulation	this study
f_{stim}	100 Hz	Frequency of learning/recall stimulation from putative input neurons	Sajikumar et al. (2005); Li et al. (2016)
r	0.5	Fraction of assembly neurons that are stimulated to trigger recall	this study

Table 2.1: Parameters for neuron and static network dynamics. Values were used as given in this table, unless stated otherwise.

Parameters for synaptic plasticity

Symbol	Value	Description	Refs.
$t_{c,\text{delay}}$	0.0188 s	Delay of postsynaptic calcium influx after presynaptic spike	Graupner and Brunel (2012)
c_{pre}	1 (0.6)	Presynaptic calcium contribution (network adjustment)	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
c_{post}	0.2758 (0.1655)	Postsynaptic calcium contribution (network adjustment)	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
τ_c	0.0488 s	Calcium time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_h	688.4 s	Early-phase time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_p	60 min	Protein time constant	Clopath et al. (2008); Li et al. (2016)
τ_z	60 min	Late-phase time constant	Clopath et al. (2008); Li et al. (2016)
γ_p	1645.6	Potentiation rate	Graupner and Brunel (2012); Li et al. (2016)
γ_d	313.1	Depression rate	Graupner and Brunel (2012); Li et al. (2016)
θ_p	3	Calcium threshold for potentiation	Li et al. (2016)
θ_d	1.2	Calcium threshold for depression	Li et al. (2016)
σ_{pl}	0.290436 nC s ^{1/2}	Standard deviation for plasticity fluctuations	Graupner and Brunel (2012); Li et al. (2016)
α	1	Protein synthesis rate	Clopath et al. (2008); Li et al. (2016)
θ_{pro}	0.210037 nC	Protein synthesis threshold	Li et al. (2016)
θ_{tag}	0.0840149 nC	Tagging threshold	Li et al. (2016)

Table 2.2: **Parameters for synaptic plasticity.** Values were used as given in this table, unless stated otherwise.

2.3. *Supplementary Information: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks*

2.3 Supplementary Information: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks

This section contains the Supplementary Information of the article presented in the previous section (Luboeinski and Tetzlaff, 2021a). The typeset and referencing were adapted to match the format of this thesis. This Supplementary Information is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

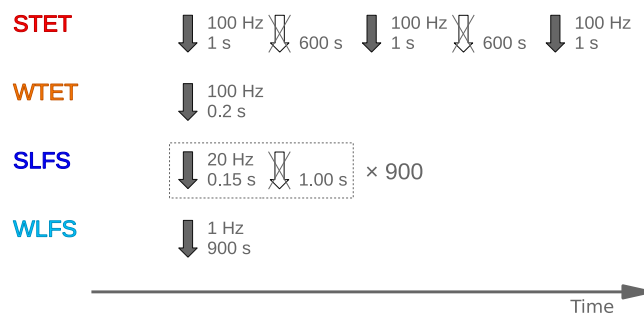


Figure 2.10: **Standard protocols for the the induction of early- and late-phase synaptic potentiation and depression** as used in Li et al., 2016. Filled arrows indicate stimulation at the specified frequency for the specified time, crossed arrows indicate breaks for the specified time.

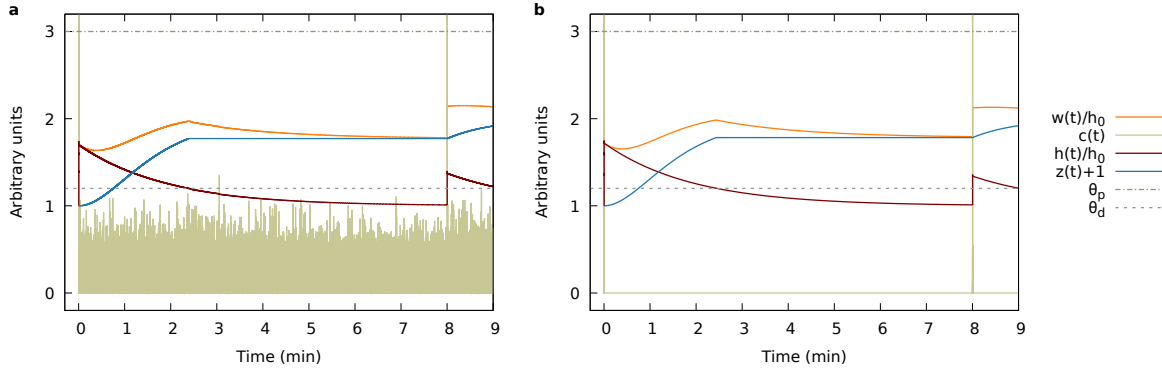


Figure 2.11: **Matching temporal dynamics of the weight of a synapse** in (a) a spiking network simulation run in full detail and (b) an accelerated network simulation with spiking dynamics only during relevant learning- and retrieval-related periods. Outside of these periods, the fully-simulated calcium concentration stays below the plasticity thresholds and thus, can be neglected. Parameter setting: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$, $n_{CA} = 150$.

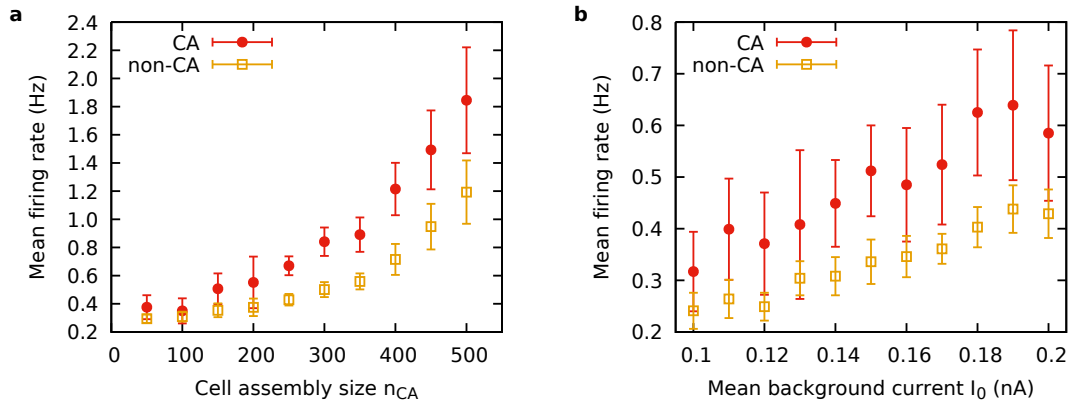


Figure 2.12: **The mean standby firing rate in the excitatory population of a network holding one cell assembly** (a) grows strictly monotonically with the size of the assembly (measurement 10 seconds after learning the assembly). Firing rates within the assembly (CA) are higher than in the rest of the population (non-CA). The firing rates for small- to medium-sized assemblies are around the average physiological value of 0.5–1.0 Hz for pyramidal cells in the hippocampus during exploratory wake states (Mizuseki and Buzsáki, 2013; Mizuseki and Miyawaki, 2017). Mean background current: $I_0 = 0.15$ nA. (b) Variation of the mean background current causes the mean firing rate to rise slightly. Assembly size: $n_{CA} = 150$. Error bars indicate the standard deviation across ten trials. Other parameters: $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$.

Luboevski and Tetzlaff, 2021a

2.3. Supplementary Information: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks

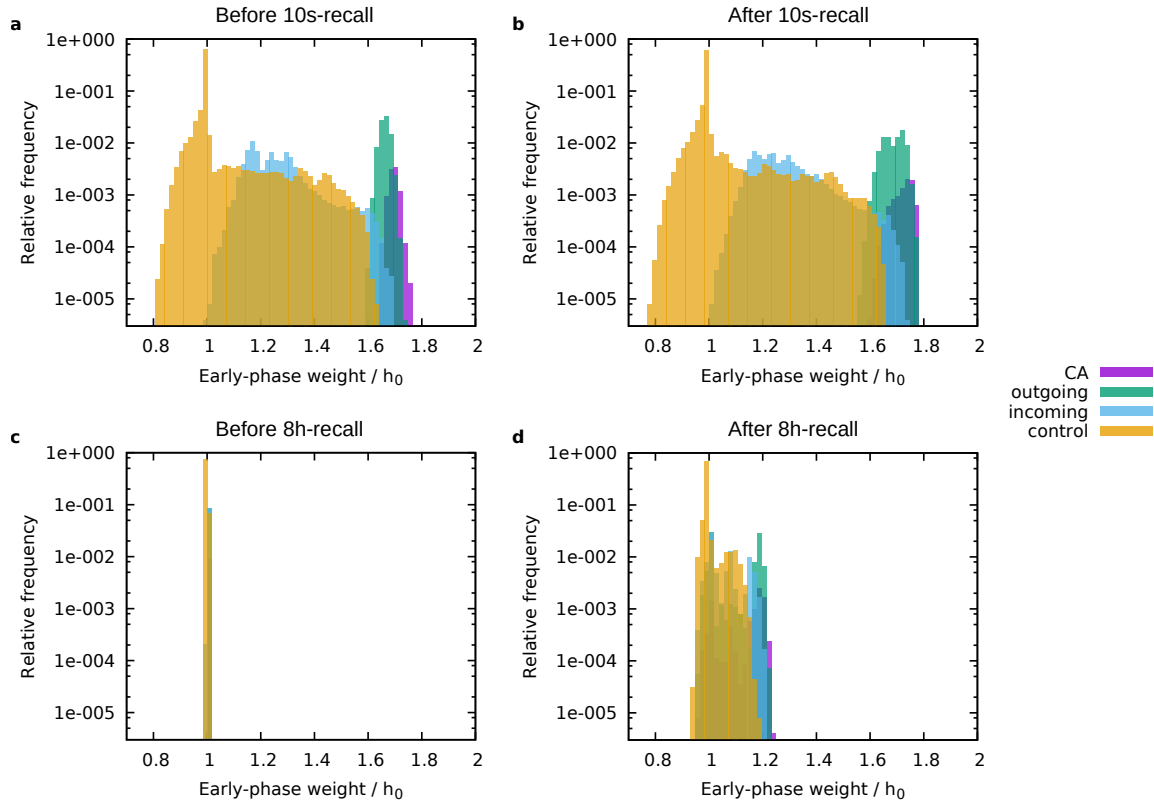


Figure 2.13: Distribution of early-phase weights in different groups of synapses of one sample network. (a,b) Ten seconds after learning, the weights within the cell assembly ('CA') are all increased significantly, as well as the weights outgoing from assembly neurons ('outgoing'). The weights within the control population ('control') stay mostly low, while there is a larger fraction of weights incoming to assembly neurons from control neurons ('incoming') that are increased. **(c,d)** After eight hours, the initially learned structure is not anymore reflected by the early-phase weights (cf. Fig. 2.14 for the late-phase weights). Moreover, the early-phase weights are only slightly affected by 10s-recall (compare **(a)** and **(b)**). By contrast, 8h-recall causes a significant change in the distribution of early-phase weights (compare **(c)** and **(d)**). In each plot, the weights were discretized into 100 bins and the relative frequencies were normalized across all groups of synapses. Note that in the given units, the initial value of early-phase weights is unity. Parameters: $n_{CA} = 150$, $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$.

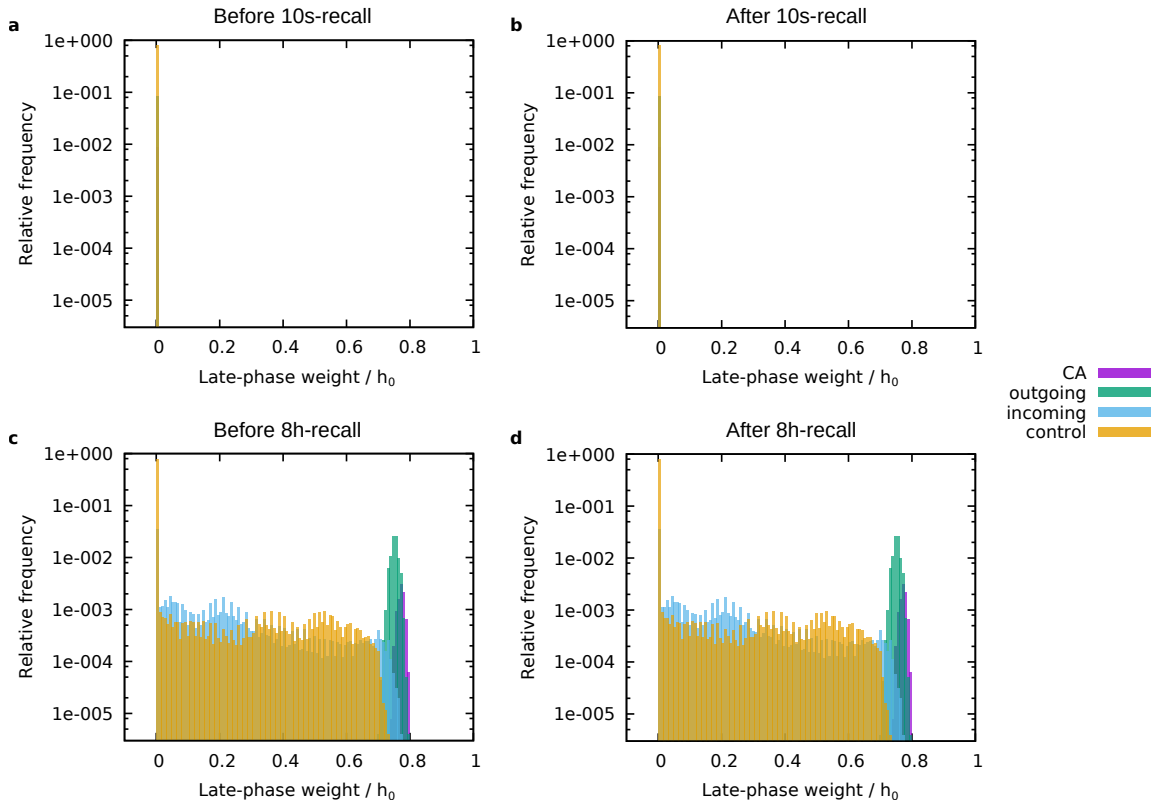


Figure 2.14: **Distribution of late-phase weights in different groups of synapses of one sample network. (a,b)** Due to their long timescales, the late-phase weights are not directly affected by learning or recall (cf. Fig. 2.13). Hence, there is also no substantial change during 10s-recall (compare (a) and (b)) or during 8h-recall (compare (c) and (d)). (c,d) Across eight hours, however, late-phase weights are significantly elevated and reflect the consolidated cell assembly. In each plot, the weights were discretized into 100 bins and the relative frequencies were normalized across all groups of synapses. Note that the initial value of late-phase weights is zero. Parameters: $n_{CA} = 150$, $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$.

2.3. Supplementary Information: Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks

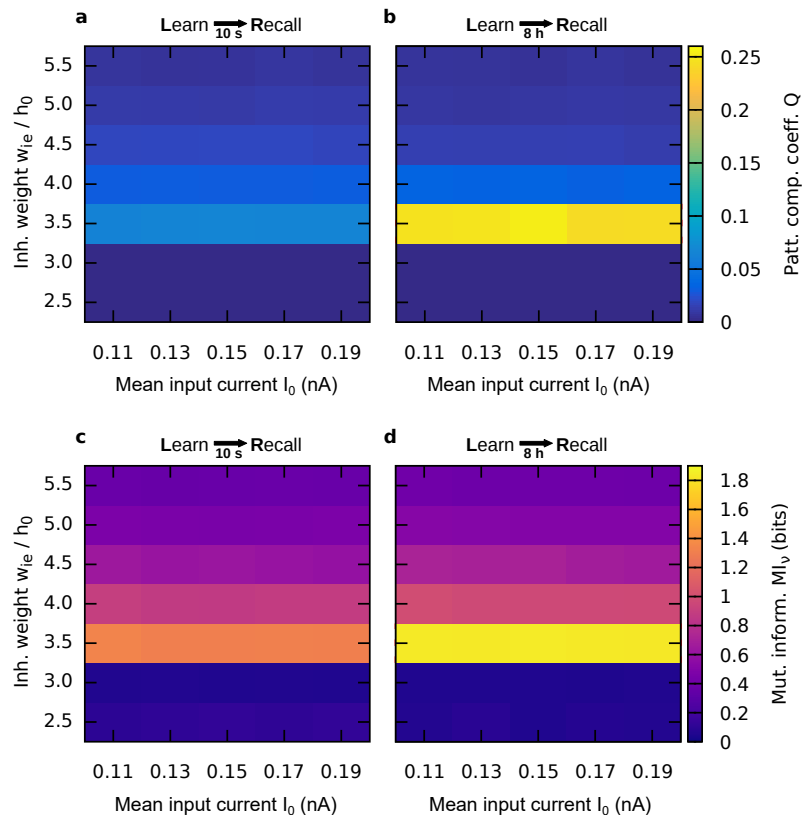


Figure 2.15: Quality of the recall 10s and 8 h after learning as a function of the I→E (w_{ie}) coupling strength and the mean background current I_0 . (a,b) Pattern completion coefficient Q (with non-significant values set to zero, see Methods); (c,d) mutual information MI_v between the neuronal activities in the network during learning and during recall. Values were averaged across ten trials. Parameters: $n_{CA} = 150$, $w_{ii}/h_0 = 4$.

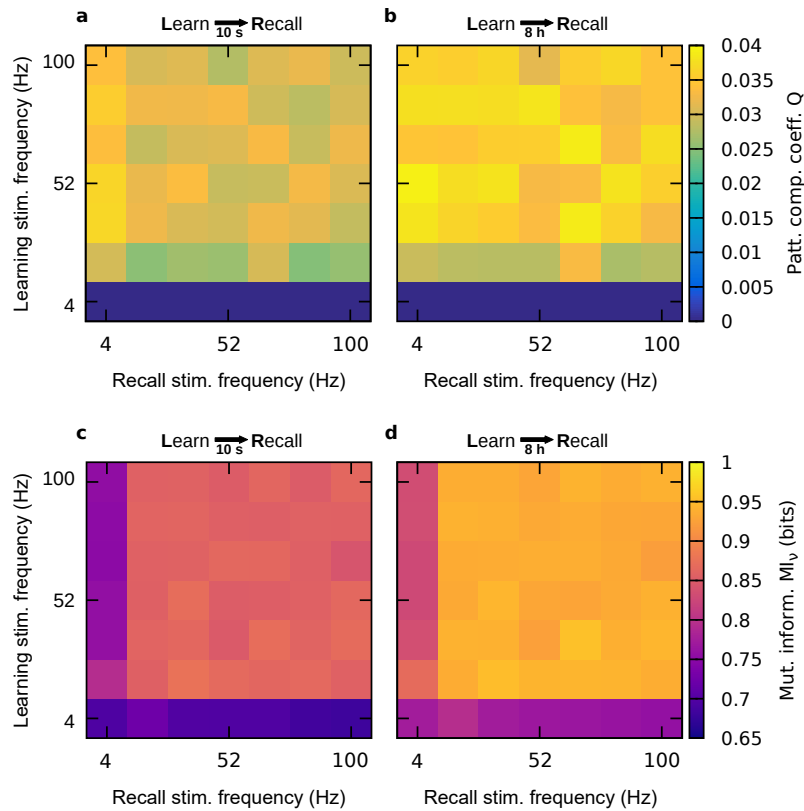


Figure 2.16: **Quality of the recall 10s and 8 h after learning as a function of the stimulation frequency f_{stim} , which is varied separately for learning and recall. (a,b) Pattern completion coefficient Q (with non-significant values set to zero, see Methods); (c,d) mutual information MI_V between the neuronal activities in the network during learning and during recall. Values were averaged across 100 trials. Parameters: $n_{CA} = 150$, $w_{ie}/h_0 = 4$, $w_{ii}/h_0 = 4$.**

2.4 Development of the model and regimes of plasticity

The network model that we used to simulate memory formation and consolidation in section 2.2 (Luboeinski and Tetzlaff, 2021a) is based on a single-neuron model with input synapses that can undergo consolidation via STC. For this, we initially attempted to use the model by Li et al. (2016), whose original paper gives a good description of synaptic plasticity and synaptic consolidation at synapses entering one neuron, using calcium-based early-phase plasticity similar to Graupner and Brunel (2012), as well as a MAT2 model for neuronal dynamics with an adaptive threshold (Kobayashi et al., 2009; Yamauchi et al., 2011). When we wanted to employ the model in a neural network, however, achieving robust recall of previously formed memory representations proved difficult. Increasing the strength of the learning or recall stimulus or the size of the cell assembly only led to divergent activity of the whole network. Achieving robust recall by variation of inhibition was similarly difficult. Finally, we reasoned that the main cause for this difficulty to recall was probably the existence of an adaptive threshold in combination with the lack of inhibitory plasticity. Indeed, disabling the adaptive threshold in the MAT2 model, thereby making it a standard leaky integrate-and-fire (LIF) model, facilitated memory recall in our network model substantially. An alternative resolution could have been the addition of dendritic spikes (Häusser et al., 2000; Jahnke et al., 2015), which may serve to improve the dynamics of pattern completion by providing positive feedback. While inhibitory plasticity and dendritic spikes could be promising complements to an adaptive threshold model, we did not follow these approaches in order to avoid adding even more complexity to our model. Later, we found that a firing-rate threshold for LTP can be another possible solution to enable memory recall even with very large cell assemblies. We will introduce such a threshold in section 4.2 (Luboeinski and Tetzlaff, 2021b).

The theoretical work by Li et al. (2016) has successfully described LTP and LTD following four different, commonly used stimuli (see Fig. 2.10). In experiments, these stimuli evoke the following four phenomena: late-phase LTP, late-phase LTD, only early-phase LTP and only early-phase LTD (Sjöström et al., 2001; Sajikumar et al., 2007, 2005). The model describes late-phase changes that are naturally preceded by early-phase changes and successfully reproduced the experimental responses. Nevertheless, in the original study (Li et al., 2016), the scaling of additive noise in the early-phase plasticity (see Eq. 2.11) was too high (cf. Graupner and Brunel, 2012), which we corrected for the model used in our study (Luboeinski and Tetzlaff, 2021a; see sections 2.2 and 2.3). Furthermore, for biologically more plausible neuronal dynamics, we added a refractory period t_{ref} and an exponential decay of postsynaptic potentials with a finite synaptic time constant τ_{syn} . Employing these three augmentations in conjunction with a LIF neuron model (as mentioned above), we obtained the same qualitative and quantitative results as Li et al., 2016 (see Fig. 2.2).

As described in section 2.2, the model parameters that Graupner and Brunel (2012) obtained from fitting experimental data were based on experiments in vitro (specifically, from hippocampal slices; Wittenberg and Wang, 2006). Since extracellular calcium concentrations are much lower in vivo than in vitro, the calcium contributions need to be corrected for simulations of in vivo processes. Following Higgins et al. (2014), the calcium influx into the postsynaptic spine can be assumed to decrease proportionally to the ratio of in vivo and in vitro extracellular calcium concentrations, which leads to an estimated factor of 0.6. Hence, for network simulations with in-vivo-like dynamics as we consider

them in this thesis, we adjusted our model by this factor after reproducing the results of Li et al. (2016) and of the four basic plasticity-inducing protocols.

The comparison of the regimes of plasticity with and without the adjustment is shown in Fig. 2.17 for a single synapse between neurons. The plots show the largest deviation from the initial value h_0 of the synaptic weight, predominantly determined by Eqs. 2.11 and 2.12 during a simulation time of 1000 s. The pre- and postsynaptic neuron both fire according to Poisson processes, the frequencies of which are varied (ν_{pre} and ν_{post} in Fig. 2.17, respectively). Additional input, including synaptic transmission, is not considered. If the largest deviation of the synaptic weight is positive, this indicates LTP, if it is negative, this indicates LTD. If the value is close to zero, this indicates that no plasticity takes place. Note that typical plasticity-inducing experiments still show depression at very low frequencies (cf. Fig. 1.4c). This is consistent with the results shown in Fig. 2.17a, as those experiments are usually performed in vitro. On the other hand, Fig. 2.17b shows that in our network model presented in section 2.2, input of up to ~ 3 Hz does not elicit early-phase plasticity. Thus, if the background noise in a network evokes firing rates in such a regime, the noise does not substantially change the weight structure of the network.

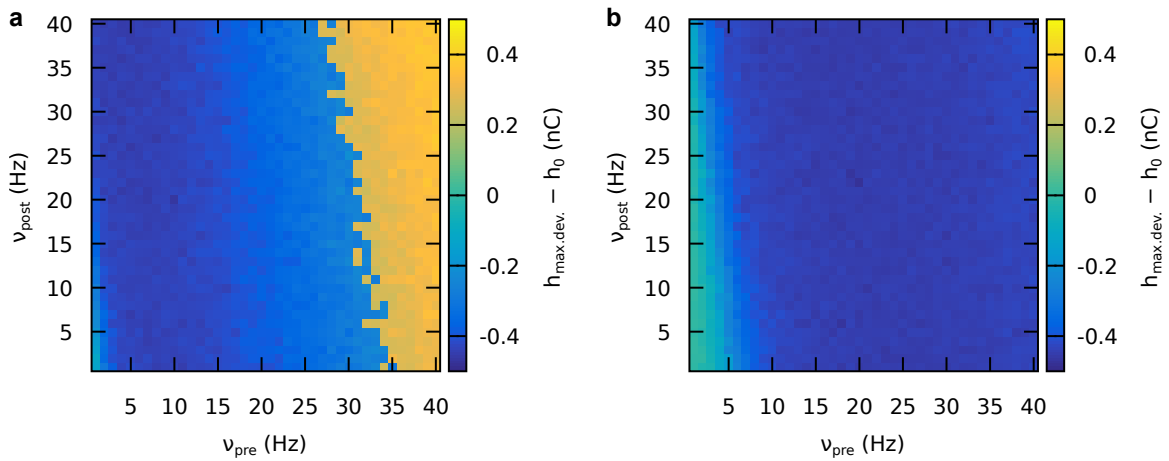


Figure 2.17: **Evoked plasticity over pre- and postsynaptic frequency.** Pre- and postsynaptic neuron spike according to Poisson processes with varied frequencies. The plots show the largest deviation (positive/yellow: LTP, negative/blue: LTD) from the initial value of the synaptic weight. Simulation time: 1000 s. Setting for the calcium contributions: (a) $c_{\text{pre}} = 1$, $c_{\text{post}} = 0.2758$ (before in vivo adjustment); (b) $c_{\text{pre}} = 0.6$, $c_{\text{post}} = 0.1655$ (after in vivo adjustment).

3

Neuromodulator-dependent consolidation of different types of memory

3.1 Chapter summary

This chapter comprises an article (main manuscript: section 3.2, Supplementary Information: section 3.3) in which we consider the impact of neuromodulator-dependent protein synthesis on the consolidation of a memory representation, thereby pursuing our second research aim. Starting from the model developed in chapter 2, we introduce a dependence of the protein synthesis threshold of STC on an abstract neuromodulator, which can be thought of as dopamine or noradrenaline, for instance (cf. Takeuchi et al., 2016; Gerstner et al., 2018). Furthermore, we consider lower average strength of learning and recall stimulation and lower average amount of neuromodulator (which was fixed in chapter 2) as compared to chapter 2, such that we reach a regime where our recall performance measures become sensitive to these parameters. Varying the amount of neuromodulator, we investigate the consolidation of different groups of synapses in networks containing a cell assembly. For high amounts of neuromodulator and strong learning stimulation, we find that the synapses within the cell assembly as well as the synapses outgoing from the cell assembly are highly potentiated (which relates to the findings in chapter 2). For low to moderate amounts of neuromodulator, however, we find that outgoing synapses are much less potentiated compared to core-internal synapses. Most importantly, we show that these differential structural outcomes are related to differential functional neural coding schemes (cf. Dayan and Abbott, 2001; Huxter et al., 2003). Specifically, to consolidate and recall input-defined firing rate patterns (as defined by the coefficient Q , introduced in chapter 2, Eq. 2.16), moderate amounts of neuromodulator yield the best results. On the other hand, to consolidate and recall temporal patterns (defined by the performance in fitting random walks from spiking patterns) and self-organized firing rate patterns (as defined by mutual information; cf. chapter 2, Eq. 2.18), high amounts of neuromodulator yield the best results. With these findings, we demonstrate that STC not only consolidates memory representations but also retroactively controls which type of neural coding is primarily used to store information about a learning experience.

The results in this chapter extend the findings of chapter 2 by characterizing the consolidation and improvement of a memory representation in the parameter space of learning stimulation strength and amount of general neuromodulator. Thereby, the results further support the robustness of consolidation and the improvement effect in our network model. Moreover, after already having appeared in the findings of chapter 2, the functional difference between the recall performance measures for spatial

patterns – coefficient Q and mutual information – now becomes clearer.

In view of chapter 4, the results of this chapter provide insights into the dynamics of assemblies that are particularly confined to their input-defined core, which we find to enhance reactivation of the input-defined pattern. In chapter 4, we will introduce a modified learning rule which gives rise to such confined assemblies, however, independently of the control through neuromodulation that we have assessed here.

3.2 Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

This section is in content identical to the following preprint article, which has been submitted for peer review (Lehr et al., 2021):

Lehr, A.B.[#], Luboeinski, J.[#], Tetzlaff, C. 2021. Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks.

The typeset and referencing were adapted to match the format of this thesis. The Supplementary Information of the article is provided in the next section. The references are provided at the end of this thesis. My contribution to this article was about 45% ([#]shared first authorship; these authors contributed equally). I developed the software and performed the simulations. I performed the data analysis and produced the figures jointly with Andrew Lehr. All authors drafted, wrote and reviewed this article.

Abstract

Important events trigger neuromodulator release in brain areas responsible for cognitive and behavioral function. It is well known that sufficient amount of neuromodulator is required for a memory to be stored in long-term memory. However, how different amounts of neuromodulator may affect the type of information committed to long-term memory is less understood. In a recurrently connected spiking neural network model featuring neuromodulator-dependent synaptic tagging and capture, we study how synaptic consolidation depends on neuromodulator amount in the hours after learning. We find that whether the network consolidates rate-coded patterns (input-defined vs. self-organized) or temporal sequences can be controlled simply via the amount of neuromodulator present. The selection of different groups of synapses for consolidation mediates this effect. Our results suggest that the amount of neuromodulator in the hours after learning may select the later functional state of neural activity during long-term memory recall, ultimately controlling which type of information structure is committed to long-term memory.

Introduction

The brain learns via adjustments at its trillions of synapses. It is thought that neural activity changes synaptic connections to support the “neural code” for an experience. Yet how the brain selects neurons and synapses to optimally store a certain memory still remains elusive.

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

The relationship between neural activity, synaptic plasticity, and the formation and consolidation of memory representations crucially depends on neuromodulation. The presence of neuromodulator is required for memory formation (Tran et al., 2008; Pezze and Bast, 2012; Johansen et al., 2014) as well as for initial memory consolidation (synaptic consolidation). For the latter, neuromodulator has to be present during a certain time window of up to hours (Moncada and Viola, 2007; Wang et al., 2010; Grogan et al., 2015; Brzosko et al., 2017; Kim et al., 2020). Following the synaptic plasticity and memory hypothesis (Martin et al., 2000; Abraham et al., 2019), supported by our previous findings on the consolidation of memory representations by synaptic tagging and capture (STC; Luboeinski and Tetzlaff, 2021a), the neuromodulator dependence of synaptic consolidation has to be understood by the neuromodulator dependence of STC.

The mechanisms of STC describe the transfer from early- to late-phase synaptic plasticity (Frey and Morris, 1997; Redondo and Morris, 2011). The STC hypothesis states that through sufficient activity at a synapse, a so-called synaptic tag is set, which enables the capture of plasticity-related proteins, leading to the stabilization of synaptic strength. There is strong experimental evidence that dopamine and other neuromodulators lower the threshold for the synthesis of these proteins in neurons (Frey et al., 1990; Sajikumar and Frey, 2004a; Navakkode et al., 2007; Mather et al., 2016). Thus, neuromodulator can facilitate the late-phase maintenance of synaptic plasticity, consistent with the requirement of neuromodulator not only during learning, but also after the experience. However, previous theoretical models of STC have considered fixed protein synthesis thresholds, reflecting fixed neuromodulator levels (Clopath et al., 2008; Barrett et al., 2009; Ziegler et al., 2015; Li et al., 2016; Luboeinski and Tetzlaff, 2021a). Therefore, the connection between neuromodulatory influence on late-phase plasticity and memory functionality has remained unexplored.

On the network level, the established theory for memory storage is based on cell assemblies, which are groups of neurons that exhibit strong internal synaptic connections (Hebb, 1949; Martin et al., 2000; Harris, 2005; Palm et al., 2014; Eichenbaum, 2018). A cell assembly is formed when a group of random neurons receives strong learning stimulation which increases the synaptic strengths within the group. Later, if part of the original stimulus is presented to the network, the strong internal connections elicit a spatial firing rate pattern (“spatial” in the sense of network topology) similar to the pattern that was originally present during learning. This similarity is thought to support rate coding, that is, the storing of information by firing rates (Harris, 2005). However, the learning stimulus can also cause the strengthening of synapses connecting to neurons outside the input-defined assembly, recruiting them as “support” neurons. The activity patterns of these support neurons have been shown to enrich the assembly dynamics and to provide a basis for a temporal memory component (Tetzlaff et al., 2015), in addition to the rate-based spatial component.

In this study, we investigate how consolidation of synaptic strength via neuromodulator-dependent STC gives rise to network structures supporting different types of neural coding. Specifically, we augment an existing recurrent network model of synaptic consolidation (Luboeinski and Tetzlaff, 2021a) to study how neuromodulation during consolidation promotes either rate-based or temporal codes for the same memory representation learned earlier.

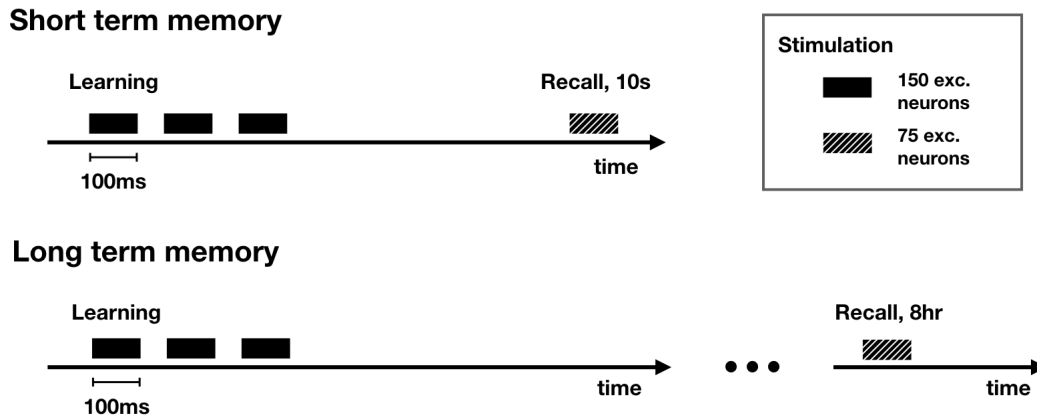


Figure 3.1: **Tasks.** Stimulation to learn and recall a memory representation. Recall stimulation is applied either 10s or 8h after learning, corresponding to short-term and long-term memory, respectively.

Results

In this study, we employed the model of memory consolidation by synaptic tagging and capture (STC) from Luboinski and Tetzlaff (2021a) and introduced a neuromodulator dependence for the protein synthesis threshold as suggested by Clopath et al. (2008). We obtained results that provide insights into emergent network dynamics evoked by synaptic tagging and capture upon variation of the amount of neuromodulator. We show that in our model, the consolidation of synapses connecting to neurons outside an input-defined cell assembly, and thereby the extent of cell assembly outgrowth, critically depends on neuromodulation in the time after learning. Based on this, we show that neuromodulation does not only gate whether or not consolidation occurs, but also controls which type of neural coding is used. The main result of our study is that if the amount of neuromodulator is high enough, cell assembly outgrowth is extensive and the rich dynamics of the support neurons enable information storage by temporal structure. On the other hand, with low amounts of neuromodulator, and hence no outgrowth, only core cell assemblies are formed, featuring improved storing of information in rate-based input-defined structures.

Neuromodulation in the hours after learning determines which synapses get consolidated

To form a cell assembly in our network, we applied a strong learning stimulus to a subset of 150 excitatory neurons (see Fig. 3.1). We call the neurons receiving this direct external stimulation during learning the *core assembly*, and synapses between core assembly neurons are denoted *core-internal synapses*. Synapses from the core assembly into the rest of the network are called *outgoing synapses*. If the outgoing synapses are significantly strengthened, we refer to the neurons they connect to as *support neurons* or *outgrowth* (cf. Fig. 3.2b).

Our first aim was to demonstrate how neuromodulation controls the consolidation of outgoing synapses to enable long-term outgrowth, and to compare this to the impact of the strength of the learning stimulus. To this end, we conducted simulations with various stimulus strengths and amounts of neuromodulator during and after learning (see Fig. 3.2e,f).

Depending on the amount of neuromodulator present after learning, different patterns of consolidated synapses emerge. As expected, if neuromodulator is not present, then no consolidation of synaptic weights takes place, regardless of the strength of the learning stimulus. Similarly, no consolidation takes place if the frequency of the learning stimulus is too low (Fig. 3.2a,e,f). With increasing neuromodulator amount and stimulus strength, selected subpopulations of synapses enter the late phase. Core-internal synapses become consolidated already at low levels of neuromodulation after learning (see Fig. 3.2a,e). Higher levels of neuromodulation, however, lead to consolidation of both core-internal synapses as well as outgoing synapses (see Fig. 3.2a,f). The distribution of the incoming synapses entering the core assembly, as well as the control synapses outside the core assembly, are shown in Supplementary Fig. 3.5a,b.

Next, we applied a recall stimulus to half of the neurons that had received the learning stimulus (Fig. 3.1). Measuring the spiking activity outside of the core assembly during recall after 8 h, we found that it increases drastically with the amount of neuromodulator that was present during consolidation (Fig. 3.2c,d), but approaches saturation at moderate to high amounts. The distribution of the activity inside the core assembly is shown in Supplementary Fig. 3.5c.

Considering the amount of protein synthesis that has taken place in the network by 1 hour and 2 hours after learning (Supplementary Fig. 3.6), one sees that most of the consolidation of outgoing synapses has to happen early, since there are much lower protein levels in non-core postsynaptic neurons around 2 hours after learning as compared to 1 hour after learning. As opposed to this, the amount of protein in core neurons even further increases beyond the first hour after learning (“After 2 hours”, Supplementary Fig. 3.6e), indicating that here, consolidation exerts more impact at later times. Independent of proteins, consolidation finally stops in all parts of the network around 3 hours after learning, by which time all synaptic tags have vanished.

Taken together, neuromodulation gates both whether consolidation via STC takes place and which type of synapses are consolidated. While low neuromodulator concentration facilitates consolidation of core-internal synapses, higher amounts of neuromodulator promote the late phase of synapses within the core assembly and from the core assembly to neurons throughout the rest of the network. A natural question is whether there are functional consequences of the different patterns of neuromodulator-dependent late-phase synaptic weights.

Neuromodulator-dependent late phase has differential effects on recall of input-defined and self-organized firing rate patterns

To investigate the functional consequences of the neuromodulator dependence of late-phase synaptic plasticity, we started from the firing rates in our network upon recall stimulation. We derived two measures from the distribution of firing rates, describing the performance on different pattern completion tasks (cf. Luboinski and Tetzlaff, 2021a): the completion of an input-defined pattern and the completion of a self-organized pattern (Fig. 3.3g). We compared the performance on these tasks directly after learning with the performance 8 hours later, that means, after consolidation. During learning we administered stimulation to 150 neurons in three pulses of 100 ms each, whereas to recall, only 75 of these neurons received one 100 ms pulse (cf. Fig. 3.1). We measured the completion of an input-defined pattern as the extent Q to which recall stimulation activated the 75 non-stimulated neurons that had previously received the learning stimulus compared to background activity in control neu-

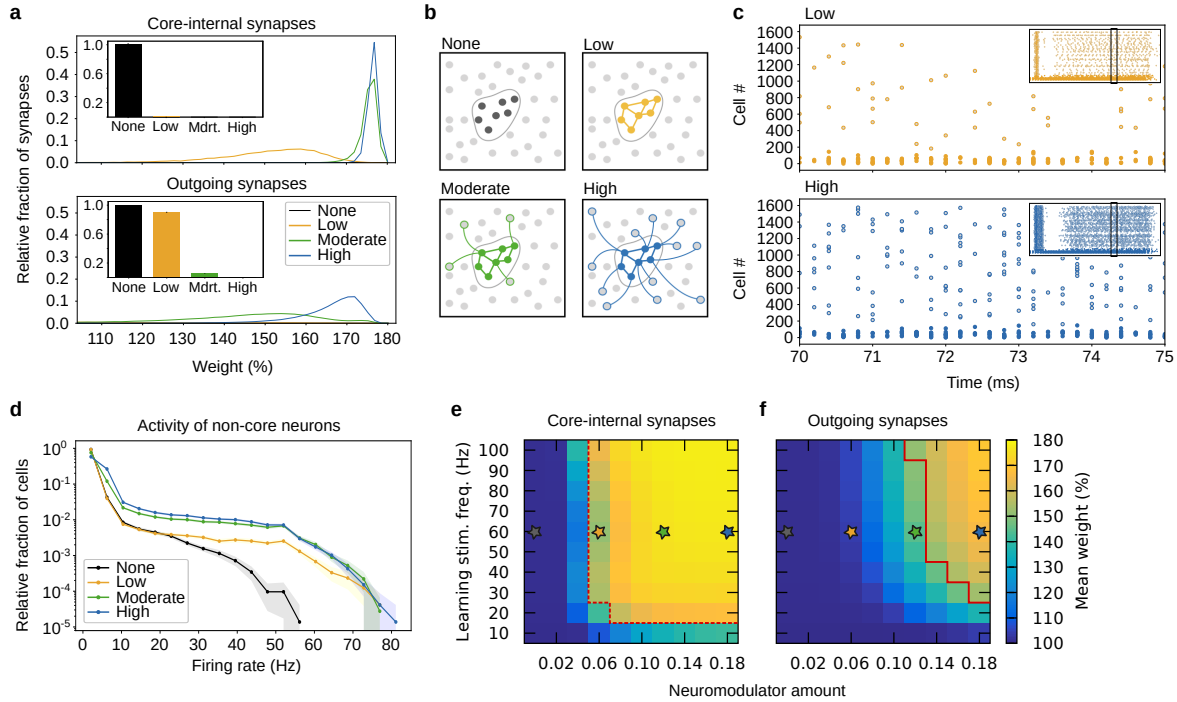


Figure 3.2: Neuromodulation controls the extent of structural outgrowth and the activity outside the core assembly. Stars in panels (e,f) indicate the four different levels of neuromodulation shown in (a-d). **(a)** Distribution of the weights of core-internal synapses and outgoing synapses at recall after 8h, averaged via bins of $\sim 1.64\%$. Insets show the fraction of synapses with no substantial late phase ($< 3.26\%$ potentiation, first two bins). **(b)** Sketches of core assembly and outgrowth at four different levels of neuromodulation. **(c)** Spike raster plots showing the spiking of the excitatory neurons during recall stimulation to the first 75 neurons, 8 h after learning. **(d)** Distribution of the firing rates of the excitatory neurons outside the core assembly during recall 8 h after learning. Averaged via bins of 4.16 Hz. **(e,f)** Raster plots showing the mean weight of the core-internal synapses and the outgoing synapses. The dashed and solid red lines demarcate the regimes where the weights are above 150%. All data except the spike raster plots were averaged over 50 networks. Error bands show the 95% confidence interval.

rons. On the other hand, for self-organized patterns, we measured the success of pattern completion by the resemblance (mutual information) of the distribution of firing rates in the network at recall and the distribution during learning (MI).

As expected, we found that 10 seconds after learning only the strength of the learning stimulation had an influence on pattern completion (Fig. 3.3a,d). However, we found that consolidation through STC led to different results 8 hours after learning (Fig. 3.3b,e). Scrutinizing the difference between the two times, we found that there are regimes of deterioration and improvement, depending on learning stimulation and neuromodulation (Fig. 3.3c,f). Remarkably, the completion of an input-defined pattern functions best in an intermediate regime, while it deteriorates at higher neuromodulator concentrations. The completion of a self-organized pattern, however, behaves almost contrary – it exhibits moderate performance for intermediate amounts of neuromodulator, but strongly benefits from higher neuromodulator concentrations. The dashed and solid red lines emphasize that these findings are correlated with the outgrowth shown in Fig. 3.2e,f. The difference in the two measures at the level of firing rates becomes evident from Fig. 3.3h, which shows that after consolidation, a high neuromodulator amount gives rise to much more activity in the control population, whose neurons are recruited into

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

the self-organized pattern. In the next section, we will examine how these rich dynamics can serve to store temporal structures.

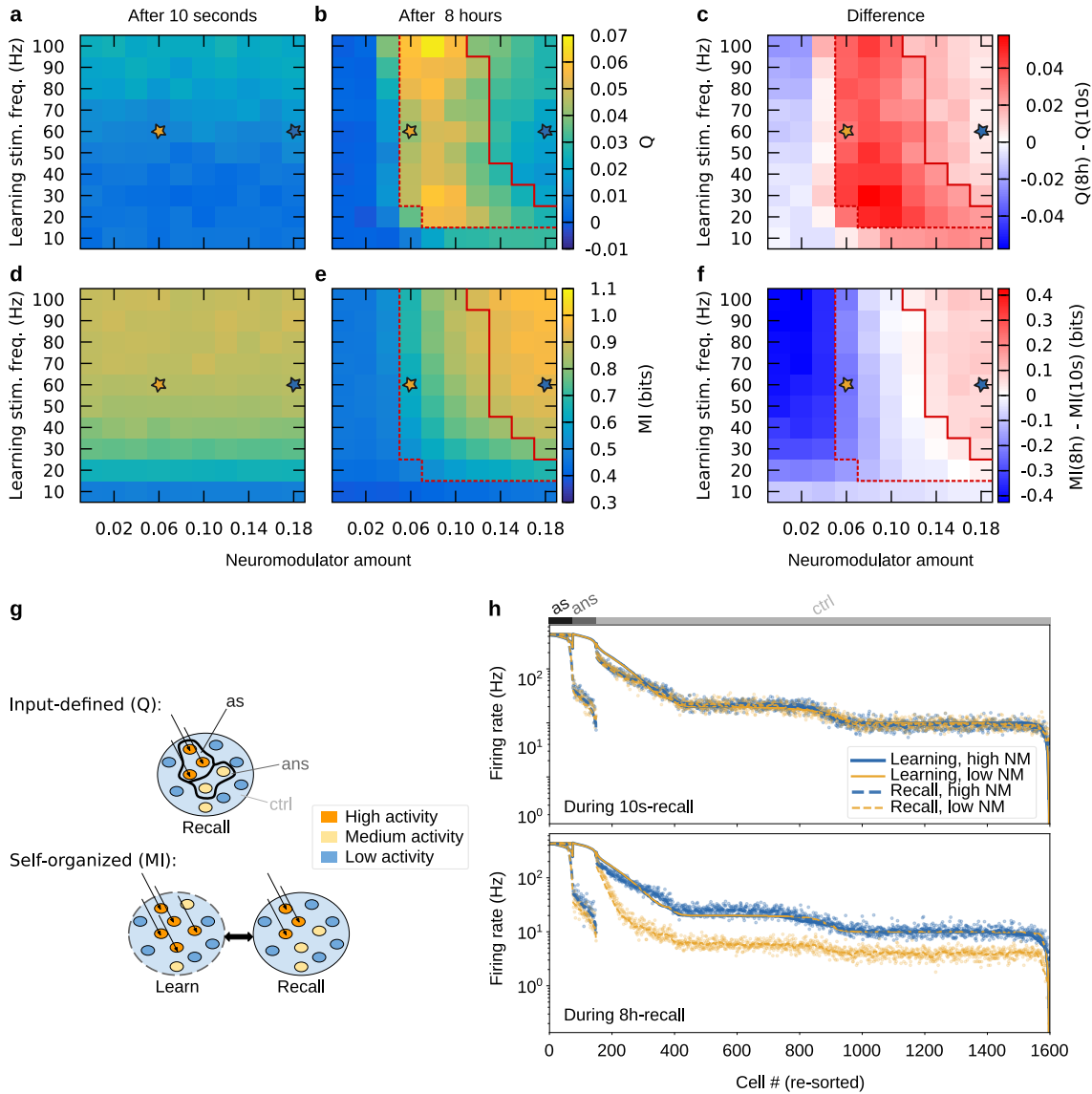


Figure 3.3: Consolidation has differential effects on completion of input-defined and self-organized patterns. The performance on completion of an input-defined pattern (**a**) 10 s, and (**b**) 8 h after learning, as measured by the coefficient Q , demonstrates the impact (**c**) of neuromodulator-dependent STC. Analogously, panels (**d–f**) show the performance on completion of a self-organized pattern, measured by mutual information MI . Dashed and solid red lines demarcate where the late-phase weight of core-internal and outgoing synapses reached 150% potentiation, respectively (cf. Fig. 3.2e,f). (**g**) depicts the definitions of the two performance measures – Q depends on the activities during recall in the subpopulations “as”, “ans”, and “ctrl”, which are defined by the inputs for learning and recall, while the mutual information MI depends on the relationship between the activities during learning and during recall. Colors represent neuronal firing rates. (**h**) shows the firing rates of the excitatory neurons grouped by subpopulation. Each subpopulation has been sorted according to their firing rate during learning. Data in (a–f) were averaged over 50 networks.

High neuromodulator concentration promotes outgrowth supporting temporal structure

Next we asked how temporal information in neural activity patterns depends on the presence of neuromodulator during the consolidation of synaptic weights. We hypothesized that consolidation of temporal patterns would improve with increased levels of neuromodulator for several reasons. First, as we showed in the previous section, low neuromodulator levels facilitate recruitment of an input-defined cell assembly, with the remaining neurons firing at low rates (Fig. 3.3h). However with increased neuromodulator levels, consolidation of outgoing synapses was improved and neurons outside of the input region were recruited, leading to improved consolidation of a self-organized rate code.

The self-organized code consisted of higher and more diverse firing rates outside the input region (Fig. 3.3h), which we hypothesized should increase temporal information. Second, recruitment of neurons that did not directly receive input stimulation has previously been associated with increased performance on a motor sequence task (Tetzlaff et al., 2015). If activity patterns from the network are to be used to reliably control a sequence of motor outputs, then the temporal progression of the activity is important. This suggests recruitment of support neurons via the consolidation of synapses from the assembly to the rest of the network should increase, and possibly stabilize, temporal structure in the activity patterns.

To test whether temporal structure was increased after consolidation specifically under high neuromodulator presence, we first estimated the dimensionality of the neural activity using principal component analysis (PCA). PCA was computed on the 1-ms binned spike data and the number of principal components required to explain 70% of the variance were counted. As expected, during the 10 s recall, dimensionality of the spike data increased with stimulation strength, but neuromodulator concentration had no effect (see Supplementary Fig. 3.7). However during the 8 h recall, after neuromodulator-dependent consolidation, the dimensionality of the neuronal activity increased compared with the 10 s recall for high neuromodulator concentrations, consistent with the consolidation of outgrowth (Fig. 3.2) and similar to the effect for self-organized rate codes (Fig. 3.3). For low neuromodulator levels, the dimensionality of the activity patterns decreased significantly after consolidation.

Notwithstanding the PCA results – for the increased temporal structure to be of actual use, temporal information induced by the input during learning should be available during recall. To test the extent of shared temporal information during learning and recall we trained a linear read out on the spike data to produce an arbitrary 1-dimensional trajectory, a random walk (see Fig. 3.4e). We used a simple linear read out as we were interested in the amount of temporal information in the activity data, and not in maximizing performance. In particular, we learned a set of output weights using ridge regression on the three learning pulses and recall simultaneously (see Methods; cf. Fig. 3.1), to test whether there is sufficient shared temporal information. Using these learned weights, we then computed the predicted output on the recall trials and measured how well it fit the desired output. For recall after 10 s, as expected there was no effect of neuromodulator, and output functions could be learned well, with a decrease in performance only for very low stimulation frequencies (Fig. 3.4a). For 8 h recall, only under high neuromodulator concentration did temporal information survive the transfer to the late phase (Fig. 3.4b). Then, however, high levels of neuromodulator even increased the shared temporal information between learning and 8 h recall compared with 10 s recall (Fig. 3.4c).

Further, we subsampled the network to evaluate how many output projections were required to learn an arbitrary output sequence. This analysis sheds light on the number of neurons that participate

in conveying temporal information and the extent to which the information they provide is independent from other neurons. While for low neuromodulator levels performance increased approximately linearly with the percent connectivity, maxing out at $R^2 \approx 0.6$, when neuromodulator levels were high, performance increased supralinearly reaching $R^2 \approx 0.6$ with only 30% connectivity and $R^2 \approx 0.9$ for 100% connectivity (Fig. 3.4d). Thus, the number of output projections required to read out a temporal sequence from the network depends crucially on neuromodulator levels during consolidation.

Taken together, these results suggest that for temporal coding to emerge under constant firing rate input in a recurrently connected network, sufficient neuromodulator or plasticity-related proteins should be present during the consolidation of synaptic weights. The emergent temporal structure could support memories with a temporal component, like learning a motor sequence, or it could simply code for static information using spatiotemporal code, akin to “odor sequences” in the hippocampus (MacDonald et al., 2013; Terada et al., 2017). Here we have shown that high neuromodulator levels during consolidation give rise to stable temporal structure when a subset of 150 neurons in the network receive a constant firing rate input. Even more, there is shared temporal information between learning and recall when only half of the neurons are stimulated at recall 8 hours later.

Discussion

We have shown that neuromodulation can retroactively modify the type of neural coding for a recent learning experience. For low levels of neuromodulation after learning, an input-defined rate code is transferred to the late phase and preferentially recalled 8 hours later. For high levels of neuromodulation, instead, a self-organized rate code emerges. Looking deeper, we found that this self-organized rate code contains temporal structure that supports temporal coding. The neuromodulator-dependent effect on coding was accompanied by a parallel effect on the consolidation of synaptic weights. In particular, while under low levels of neuromodulation STC only transferred core-internal synapses to the late phase, under higher levels of neuromodulation both core-internal and outgoing synapses were transferred. Importantly, we have shown that for a single learning event the type of neural coding and the related activity pattern at recall depends on the neuromodulator concentration during consolidation.

Across the brain, neural correlates for both rate and temporal coding have been observed in many tasks (Georgopoulos et al., 1986; Huxter et al., 2003; MacDonald et al., 2011; Dragoi, 2020). Depending on the behavioral state or task, the same brain region can show indications of both rate and temporal codes. Place cells and grid cells are prototypical examples of putative rate-based codes in the hippocampus and entorhinal cortex (O’Keefe and Dostrovsky, 1971; Fyhn et al., 2004). However embedded in these rate codes are so-called theta sequences, a temporal coding scheme where neighboring place/grid cells fire in sequence, dependent on the local theta rhythms (Foster and Wilson, 2007). Interestingly, temporal and rate codes may be independent (Huxter et al., 2003), and disrupting the temporal code can leave the rate code preserved in some instances (Wang et al., 2015; Lepperød et al., 2021). Flexible switching between type of coding depending on the task also appears to be possible. During the delay of a working memory task, when a rodent is running on a wheel, stable sequences of neural activity arise in the hippocampus and these predict subsequent decisions. Sequences however do not arise when the animal runs on the wheel in a non-memory task (Czurkó et al., 1999; Hirase

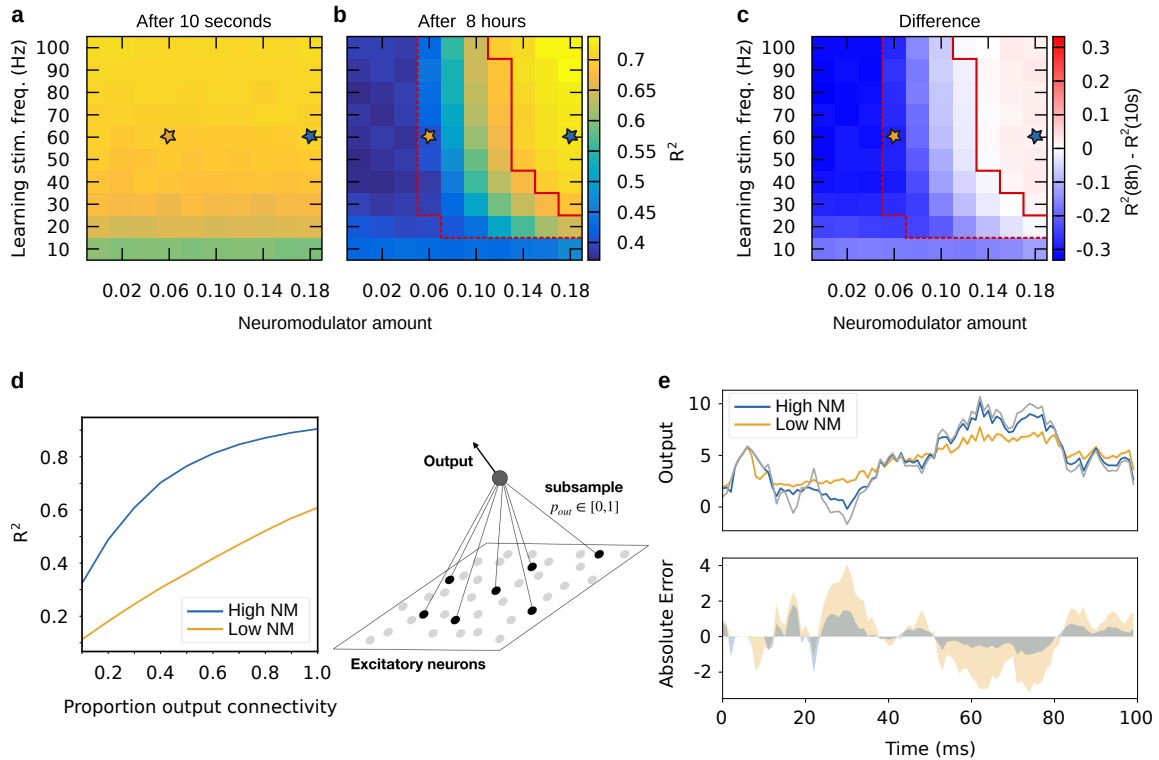


Figure 3.4: High level of neuromodulation during consolidation supports temporal structure. The goodness of fit (R^2) of the regularized linear regression model is shown for recall (a) 10 s and (b) 8 h after learning. Yellow star represents low neuromodulator condition, blue star high neuromodulator. (c) The impact of neuromodulator-dependent STC on the shared temporal information during learning and recall is measured as the difference between the 8 h and 10 s results. (d) The goodness of fit is shown for high (blue) and low (yellow) across different percent output connectivities. A depiction of the subsampling process is shown on the right. (e) The predicted trajectory at 8 h recall in the low (yellow) and high (blue) neuromodulator conditions are shown for one example target function (gray). The bottom panel shows the absolute error (with sign). Data in (a,b,c) were averaged over 5000 instances: 50 networks with 10 target functions, across 10 values of output connectivity. In (d), 500 instances were averaged for each percent connectivity: 50 networks, 10 target functions. In (e), the average of 50 networks is shown for an example target function and $p_{out} = 1$. Error bands in (d,e) show the 95% confidence interval.

et al., 1999; Pastalkova et al., 2008), in which case a subset of neurons show sustained firing, akin to a rate code.

But why might a single network in the brain need to switch between rate coding and temporal coding retroactively? A reasonable answer is that depending on the task, a particular type of code may be better suited. For example, in the working memory task already mentioned, stable temporal sequences are thought to hold information about the upcoming decision for the duration of the delay period (Pastalkova et al., 2008). Disrupting these sequences indeed disrupts performance. In other cases, for example to signal a particular location in space, a rate-based code for position (e.g. place cells) may suffice. The critical point, which is quite familiar in our everyday experience, is that the utility of an event and the nature of its consequences are often not clear until after it is over, when reward or other task-related information becomes available (Wörgötter and Porr, 2005). Thus, during the early stages of learning a task, both types of coding should be possible, with task-specific neuromodulator release subsequently promoting a particular type of code to be consolidated into long term

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

memory. This is consistent with the finding that dopamine in hippocampal CA1 after stimulation can retroactively modify the effect of a learning stimulus, converting synaptic depression into potentiation (Brzosko et al., 2015; Gerstner et al., 2018).

Our results suggest that for the retroactive modification of coding, there should be mechanisms to flexibly release the appropriate amount of neuromodulator after an experience. This could perhaps happen via the type of activity patterns in the neuromodulator releasing brain area (e.g. tonic vs. phasic; Liu et al., 2021; Sonneborn and Greene, 2021). Furthermore, this could be guided by the recruitment of different neuromodulatory systems. In particular, both the locus coeruleus (LC) and the ventral tegmental area (VTA) are known to release dopamine into the hippocampus, with LC also releasing norepinephrine. Vastly more dopaminergic projections to the dorsal hippocampus come from the LC than from VTA (Takeuchi et al., 2016), suggesting that the concentration of neuromodulator released upon activation of LC should be significantly higher than for VTA. Duzkiewicz et al. (2019) proposed that these two neuromodulatory systems are recruited for different types of novelty processing (VTA when event has similarities to past ones, LC when experience is distinct) and result in different types of memory storage. Specifically, the authors suggest that VTA promotes semantic memory via systems memory consolidation while LC promotes episodic memory via synaptic consolidation. Interestingly, while episodic memory consists of a vivid depiction of a sequence of events, semantic memory is not considered to contain a temporal component. Our results suggest that these effects might be explained simply by the amount of neuromodulator release after learning, consistent with the projection densities of VTA and LC inputs and with the proposed functional distinction between the type of memory that ends up getting stored.

The spatial and temporal profiles of neuromodulator release exist across multiple orders of magnitude (Liu et al., 2021). Release can take place very locally with the substance being cleared up in milliseconds, however, it can also be widespread and last for hours. Mather et al. (2016) suggest that spatially localized norepinephrine release in cortex and hippocampus triggers local protein synthesis to enhance memory consolidation in selected cell populations. In hippocampal CA2, selective release of neuromodulators has been suggested to prioritize important experiences for replay via neuromodulator-dependent plasticity (dependent e.g. on vasopressin, oxytocin, or substance P; Stöber et al., 2020). Our results here suggest that widespread release coming after an experience can still selectively determine the neuronal subpopulation recruited to store a memory. Even more, the coding dynamics of this subpopulation is controlled by the amount of neuromodulator.

We investigated how neuromodulation as well as learning stimulation influence the functionality of memory representations. The principal difference in the impact of neuromodulation and learning stimulation is given by the nonlinear nature of the synaptic tag. Once the tag of a particular synapse has vanished, this synapse no longer undergoes stabilization through STC, even if neuromodulation is drastically increased. Learning stimulation, on the contrary, is able to modify the synaptic weight and thereby controls the lifetime of the synaptic tag, which defines the time window for STC. The following thought experiment demonstrates why this difference matters. Consider learning stimulation that evokes the formation of a cell assembly with moderate outgrowth (as in the ‘moderate’ case in Fig. 3.2). Right after learning, the core of the assembly will exhibit strong internal weights, while the outgoing weights will mostly be of moderate strength. This corresponds to long lifetimes of the tags of core-internal synapses, and medium lifetimes of the tags of outgoing synapses. Hence, shortly

after most of the tags of outgoing synapses have vanished, most core synapses will still be tagged. If neuromodulation is drastically increased at this point in time, the core of the assembly will undergo enhanced stabilization, while the outgrowth will not experience any more stabilization. Learning stimulation, for instance a second learning stimulus occurring after the initial one, can affect any synapse and therefore does not yield this type of selective retrograde modification.

We have constrained our investigations here to the simplest possible situation that can shed light on our question. Of course, in the brain, a host of neuromodulatory influences can be present simultaneously (Eichenbaum, 2011; Roelfsema and Holtmaat, 2018). In addition to synaptic consolidation, they may also affect excitability (Mather et al., 2016; Duzskiewicz et al., 2019) and early-phase synaptic plasticity at the time of encoding (Otmakhova and Lisman, 1996; Lindskog et al., 2006; Zhang et al., 2009; Nakano et al., 2010; Navakkode et al., 2012). Different molecules, including receptor subunits, CaMKII, PKM ζ , and actin have been suggested to differentially affect synaptic plasticity as plasticity-related proteins or as correlate for the synaptic tag, although the exact mechanisms remain elusive (Matsuzaki et al., 2001; Meyer et al., 2014; Smolen et al., 2019; Becker and Tetzlaff, 2021; Bonilla-Quintana and Wörgötter, 2021). Furthermore, many different types of neurons exist with diverse connection profiles (Freund and Buzsáki, 1996; Braitenberg and Schüz, 1998; Le Duigou et al., 2014). Our results show that a generic mechanism with abstract neuromodulator, proteins, and tag can control the consolidation of synaptic subpopulations and thereby determine functional dynamics for the recall of long-term memory representations. We could relate these functional dynamics to rate codes and temporal sequences, which have been extensively treated in the literature (Dayan and Abbott, 2001; Huxter et al., 2003; Harris, 2005). In networks with more diverse neural populations under the control of a number of neuromodulatory influences, we would expect additional functional dynamics.

Please note that while we considered networks holding a single cell assembly in this study, it has been shown that networks holding multiple cell assemblies can be described with a related model (Luboeinski and Tetzlaff, 2021b). In addition, the simulation code that we provide already enables to do this.

Taken together, neuromodulator release after an event controls retroactive modification of the network structure and resulting dynamics associated with that event. In order to optimally store memories, this is a crucial feature for a brain to possess.

Methods

The model, the stimulation protocols, and the measures used in this study are described in this section. Unless stated elsewhere, the parameters that we used are shown in Tables 3.1 and 3.2.

Network model

To simulate the dynamics of long-term memory representations, we use the network model of memory consolidation with synaptic tagging and capture from (Luboeinski and Tetzlaff, 2021a) (for more details, please see that paper). The network has hippocampal characteristics regarding firing rate and connectivity and comprises spiking neurons and synapses with detailed plasticity features. As we use leaky integrate-and-fire neurons, the membrane potential of neuron i is given by the following equation (cf. Gerstner and Kistler, 2002):

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

$$\tau_m \frac{dV_i(t)}{dt} = V^{\text{rev}} - V_i(t) + V_i^{\text{PSP}}(t) + R_m \left(I_i^{\text{bg}}(t) + I_i^{\text{stim}}(t) \right) \quad (3.1)$$

where V^{rev} is the reversal potential, τ_m is the membrane time constant, R_m is the membrane resistance, and $I_i^{\text{bg}}(t)$ and $I_i^{\text{stim}}(t)$ are the external background current and external stimulus current, respectively. The synaptic weights and presynaptic spike times are given by w_{ji} and t_j^k , respectively, yielding the postsynaptic voltage

$$V_i^{\text{PSP}}(t) := \sum_j \sum_{t_j^k} w_{ji} \cdot \exp\left(-\frac{(t - t_j^k - t_{\text{ax,delay}})}{\tau_{\text{syn}}}\right) \Theta(t - t_j^k - t_{\text{ax,delay}}), \quad (3.2)$$

where $t_{\text{ax,delay}}$ is the axonal delay time, τ_{syn} is the synaptic time constant, and $\Theta(\cdot)$ is the Heaviside step function.

If the membrane potential V_i crosses the threshold V^{th} , a spike is generated and its time of occurrence is stored to serve as input to other neurons and to be used for computation of firing rates and temporal traces. Subsequently, the membrane potential is reset to V^{reset} , and remains fixed at this value for the refractory period t_{ref} . In the absence of learning and recall stimulation, only a background noise current enters the membrane potential dynamics, described by an Ornstein-Uhlenbeck process:

$$\tau_{\text{syn}} \frac{dI_i^{\text{bg}}(t)}{dt} = I^0 - I_i^{\text{bg}}(t) + \sigma_{\text{wn}} \cdot \Gamma_i(t) \quad (3.3)$$

with mean current I^0 , white-noise standard deviation σ_{wn} , and Gaussian white noise $\Gamma_i(t)$ with mean zero (Gillespie, 1996). Since the power spectrum of the Ornstein-Uhlenbeck process resembles that of fluctuating input from a large presynaptic neuronal population in the cortex (Destexhe et al., 2003), it accounts for synaptic inputs from outside our network. Besides the background noise, we also model the stimulus current for learning and recall by an Ornstein-Uhlenbeck process:

$$\tau_{\text{syn}} \frac{dI_i^{\text{stim}}(t)}{dt} = w_{\text{stim}} \cdot N_{\text{stim}} \cdot f_{\text{stim}} - I_i^{\text{stim}}(t) + w_{\text{stim}} \cdot \sqrt{N_{\text{stim}} \cdot f_{\text{stim}}} \cdot \Gamma_i(t). \quad (3.4)$$

Here, mean and standard deviation of the process are defined by putative input spikes to the stimulated neuron, conveyed by N_{stim} neurons at the frequency f_{stim} through synapses of weight w_{stim} (Gillespie, 1996; Dayan and Abbott, 2001).

Our neural network consists of 1600 excitatory and 400 inhibitory neurons, the ratio between excitatory and inhibitory neurons being typical for cortical and hippocampal networks (Braitenberg and Schüz, 1998). A spike occurring in neuron j is transmitted to neuron i (cf. Eq. 3.1) if there is a connection from j to i . The probability of connection across our whole network is 10%, which is a reasonable value for hippocampal region CA3, for instance (Le Duigou et al., 2014). The postsynaptic current that is evoked in neuron i by a presynaptic spike from neuron j depends on the weight of the corresponding synapse, which is given by:

$$w_{ji} = \begin{cases} h_{ji} + h_0 \cdot z_{ji} & \text{for E} \rightarrow \text{E}, \\ w_{\text{ei}} & \text{for E} \rightarrow \text{I}, \\ w_{\text{ie}} & \text{for I} \rightarrow \text{E}, \\ w_{\text{ii}} & \text{for I} \rightarrow \text{I}, \end{cases} \quad (3.5)$$

where E relates to excitatory and I relates to inhibitory neurons. While the synaptic connections involving inhibitory neurons are constant, the total synaptic weight w_{ji} of E→E connections consists of two variable contributions, which is a critical feature of STC mechanisms. The first contribution is the early-phase weight h_{ji} and the second is the late-phase weight z_{ji} . We used h_0 to normalize z , such that it has the same dimension as h . This factor is in accordance with experimental data (Li et al., 2016; Luboeinski and Tetzlaff, 2021a).

The early-phase weight is described by the following differential equation:

$$\tau_h \frac{dh_{ji}(t)}{dt} = 0.1 (h_0 - h_{ji}(t)) + \gamma_p (1 - h_{ji}(t)) \cdot \Theta[c_{ji}(t) - \theta_p] - \gamma_d h_{ji}(t) \cdot \Theta[c_{ji}(t) - \theta_d] + \zeta(t), \quad (3.6)$$

where τ_h is a time constant, $c_{ji}(t)$ is the calcium concentration at the postsynaptic site. Early-phase LTP occurs with rate γ_p if the calcium concentration is above the threshold θ_p , and early-phase LTD occurs with rate γ_d if the calcium concentration is above the threshold θ_d . In addition, a noise term is added if LTD or LTP occurs:

$\zeta(t) = \sqrt{\tau_h} [\Theta(c_{ji}(t) - \theta_p) + \Theta(c_{ji}(t) - \theta_d)] \sigma_{pl} \Gamma(t)$, with a scaling factor σ_{pl} and Gaussian white noise $\Gamma(t)$ with mean zero and variance $1/dt$.

The calcium-based model of early-phase plasticity is based on the calcium-driven plasticity model by Graupner and Brunel (2012), using parameters fitted on hippocampal slice data (Wittenberg and Wang, 2006), and corrected by a factor of 0.6 to account for in vivo conditions (Higgins et al., 2014). The relaxation term was introduced by (Li et al., 2016) to account for synaptic tagging and capture, along with the late-phase weight dynamics described below.

The calcium concentration depends on pre- and postsynaptic spikes at times t_j^n and t_i^m , and is described by the following equation:

$$\frac{dc_{ji}(t)}{dt} = -\frac{c_{ji}(t)}{\tau_c} + c_{pre} \sum_n \delta(t - t_j^n - t_{c,delay}) + c_{post} \sum_m \delta(t - t_i^m), \quad (3.7)$$

where τ_c is a time constant, c_{pre} is the increase in calcium evoked by presynaptic spikes, c_{post} is the increase in calcium by postsynaptic spikes, $t_{c,delay}$ is a delay for presynaptic signals, and $\delta(\cdot)$ is the Dirac delta distribution.

The late-phase synaptic weight depends on the early-phase weight and is given by (Li et al., 2016):

$$\tau_z \frac{dz_{ji}(t)}{dt} = p_i(t) \cdot (1 - z_{ji}(t)) \cdot \Theta[(h_{ji}(t) - h_0) - \theta_{tag}] - p_i(t) \cdot (z_{ji} + 0.5) \cdot \Theta[(h_0 - h_{ji}(t)) - \theta_{tag}], \quad (3.8)$$

where τ_z is a time constant, $p_i(t)$ is the amount of plasticity-related proteins, and θ_{tag} is the tagging threshold. Both late-phase LTP and late-phase LTD depend on the amount of plasticity-related proteins and on the presence of the synaptic tag. The synapse is tagged if the change in early-phase weight $|h_{ji}(t) - h_0|$ exceeds the tagging threshold θ_{tag} . The late-phase weight changes when the synapse is

tagged and plasticity-related proteins are abundant ($p_i(t) > 0$).

New plasticity-related proteins are synthesized if the sum of changes in early-phase weight across the whole neuron exceeds the protein synthesis threshold θ_{pro} (Clopath et al., 2008):

$$\tau_p \frac{dp_i(t)}{dt} = -p_i(t) + \alpha \Theta \left[\left(\sum_j |h_{ji}(t) - h_0| \right) - \theta_{\text{pro}} \right]. \quad (3.9)$$

While Luboeinski and Tetzlaff (2021a) used a fixed protein synthesis threshold, we consider it variable, in order to model the impact of neuromodulation. The protein synthesis threshold depends on the concentration of neuromodulator NM in the following way (Clopath et al., 2008):

$$\theta_{\text{pro}}(\text{NM}) = \frac{1}{\text{NM} + 0.001}. \quad (3.10)$$

Protocol to learn and recall spatial patterns

Initially, we let the activity settle for 10.0 seconds, before applying the learning protocol. The learning protocol included three stimulus pulses, each lasting for 0.1 seconds, applied to 150 neurons in the network (see Fig. 3.1). The stimulus pulses were separated by breaks of 0.4 seconds. During the pulses, the stimulus current $I_{\text{stim}}(t)$, described by Eq. 3.4, enters the neuronal membrane potential, given by Eq. 3.1.

We applied a recall stimulus either 10 seconds or 8 hours after the end of the learning stimulus. The recall stimulus consisted of one pulse, lasting for 0.1 seconds, to half of the neurons that had earlier received the learning stimulus. The stimulus current for the recall stimulus was also modeled by Eq. 3.4.

The network simulations that we needed to perform for this study were computationally extremely demanding. Although it was very helpful that we could use the computing cluster of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) and although we used compiled C++ code, this study would still not have been feasible if we had had to run our spiking network simulations in full detail. Thus, in those simulations with recall stimulation after eight hours, we refrained from computing the spiking dynamics during the consolidation phase in between learning and recall stimulus. In this phase, we just computed the late-phase dynamics and the exponential decay of the early-phase weights. We have shown in Luboeinski and Tetzlaff (2021a) that the weight dynamics of a synapse does not change if such sparsely occurring spikes are neglected.

Measuring spatial recall

We used two measures to determine the performance in recalling a spatial pattern (note that these measures were also used in Luboeinski and Tetzlaff, 2021a). The input-defined measure Q describes the degree to which the pattern defined by the learning stimulus (which can be described in short by “150 neurons ‘on’, the rest ‘off’”) is completed upon activation of half of the originally stimulated neurons.

To this end, the excitatory population is divided into three subpopulations: assembly neurons that are stimulated by both recall and learning stimulus (“as”), assembly neurons that are not stimulated

by recall but were stimulated by learning stimulus (“ans”), and control neurons that are not stimulated by neither recall nor learning stimulus (“ctrl”). The mean firing rate in these three subpopulations is given by \bar{v}_{as} , \bar{v}_{ans} , and \bar{v}_{ctrl} , respectively.

Based on these mean activities, the quality of input-defined recall is computed by the following equation:

$$Q := \frac{\bar{v}_{\text{ans}} - \bar{v}_{\text{ctrl}}}{\bar{v}_{\text{as}}}. \quad (3.11)$$

Perfect pattern completion of the input-defined pattern would thus be achieved for $Q = 1$ (which is hard to achieve in a non-attractor system), while $Q = 0$ describes the case with no pattern completion at all.

In addition to measuring the recall of the input-defined pattern, we measured the performance in recalling a self-organized pattern. To do this, we considered the mutual information MI_v between the distribution of firing rates at the end of recall and the distribution of firing rates at the end of learning. Specifically, this is calculated from the entropy at time $t_{\text{learn}} = 11.0$ s (during learning), the entropy at time $t_{\text{recall}} = 20.1$ s or $t_{\text{recall}} = 28810.1$ s (during recall), and the joint entropy between both:

$$MI_v := H(v(t = t_{\text{learn}}, n)) + H(v(t = t_{\text{recall}}, n)) - H(v(t = t_{\text{learn}}, n), v(t = t_{\text{recall}}, n)), \quad (3.12)$$

where $v(t, n)$ describes the firing rate of a given neuron n at a given time t , computed using a sliding window of 0.5 s.

Dimensionality estimation

We aimed to estimate the amount of temporal information available to a downstream readout from the spiking activity of the 1600 excitatory neurons during recall both at 10 s and 8 h. For this, we estimated the dimensionality of the spike data using principal component analysis (PCA). The data was first binned into 1 ms bins before we applied PCA to the binned spike data from the 10 s and 8 h recall trials using `sklearn.decomposition.PCA` from Python’s scikit-learn library. The data is stored in a matrix containing ones when a neuron spiked, and otherwise zeros (despite binning given the refractory period a neuron can only firing one spike in a 1ms bin). Its dimension is 1600 excitatory neurons by 100 time points (100 ms stimulus with 1 ms temporal bins).

To estimate dimensionality we computed the number of principal components required to retain 70% of the variance. This was done for each neuromodulator amount and learning stimulus frequency for 50 randomly initialized networks. Averaging over the data from the 50 networks leads to the matrices for 10 s and 8 h recall shown in Supplementary Fig. 3.7a,b.

Temporal sequence learning

To test whether temporal structure present during learning is also present during recall, we trained the network to produce temporal sequences. We binned the data into 1 ms bins and generated 10 Gaussian random walks given by

$$y[i + 1] = y[i] + \mathcal{N}(0, 1)$$

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

with $y[0] = 1$ and length of 100 time bins. We then computed linear least squares with L2 regularization (using `sklearn.linear_model.Ridge`), which minimizes the objective function

$$\|y - Xw\|_2^2 + \alpha \|w\|_2^2$$

where y is the target function and X is the data matrix of dimension (number of samples, number of features). Here the target is our random walk and the data are the binned spike rasters (transposed). In practice, to train the model, y and X are concatenated over the three learning trials and one recall trial, such that $y \in \mathbb{R}^{400}$, with 400 time bins, and $X \in \mathbb{R}^{400 \times p_{\text{out}} \cdot 1600 + 1}$ again with 400 time bins and $p_{\text{out}} \cdot 1600$ neurons plus one dimension for the intercept. The output weight vector is $w \in \mathbb{R}^{p_{\text{out}} \cdot 1600 + 1}$, a weight for each neuron plus one dimension for the intercept. Here $\alpha = 0.1$ is the regularization parameter which acts to keep the weights w from becoming too large. For each of the 10 random walk target functions we used 10 different values of p_{out} , namely $p_{\text{out}} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. The linear regression was computed in this way for each of the 10 s and 8 h recalls. Then, the resulting models were used to predict the target function only based on the data from the appropriate recall trial. This was repeated for each neuromodulator amount and learning stimulation frequency for each of the 50 networks.

To produce Fig. 3.4a,b,c, we averaged over the R^2 scores for the predictions on the 10 target functions, 10 percent output connectivities, and 50 networks.

Software

For data analysis and the creation of plots we used Python 3.7.3, NumPy 1.20.1, scikit-learn 0.21.3, Matplotlib 3.1.1, as well as gnuplot 5.0.3.

We used C++ in the ISO 2011 standard to implement our simulations. To compile and link the code, we employed g++ in version 7.5.0 with boost in version 1.65.1. Our code has been released under the Apache-2.0 license and can be retrieved from: <https://github.com/jlubo/memory-consolidation-stc>.

Random numbers were generated using the generator 'minstd_rand0' from the C++ standard library, while the system time served as the seed. We implemented a loop in our code which ensured that for each distribution a unique seed was used.

Symbol	Value	Description	Refs.
Δt	0.2 ms	Duration of one time step for numerical computation	this study
τ_m	10 ms	Membrane time constant	Dayan and Abbott (2001); Gerstner and Kistler (2002)
τ_{syn}	5 ms	Synaptic time constant, also for external input current	Dayan and Abbott (2001); Gerstner and Kistler (2002); Roth and van Rossum (2009)
$t_{\text{ax,delay}}$	3 ms	Axonal spike delay	Gerstner and Kistler (2002); Lin and Faber (2002)
t_{ref}	2 ms	Duration of the refractory period	Dayan and Abbott (2001); Kobayashi et al. (2009)
R_m	10 M Ω	Membrane resistance	Dayan and Abbott (2001)
V^{rev}	-65 mV	Reversal (equilibrium) potential	Dayan and Abbott (2001)
V^{reset}	-70 mV	Reset potential	Dayan and Abbott (2001)
V^{th}	-55 mV	Threshold potential to be crossed for spiking	Dayan and Abbott (2001)
I^0	0.15 nA	Mean of the external background current	Luboeinski and Tetzlaff (2021a)
σ_{wn}	0.05 nA s ^{1/2}	Standard deviation for Gaussian noise in the background current	Luboeinski and Tetzlaff (2021a)
N_e	1600	Number of neurons in the excitatory population	Luboeinski and Tetzlaff (2021a)
N_i	400	Number of neurons in the inhibitory population	Luboeinski and Tetzlaff (2021a)
p_c	0.1	Probability of a connection existing between two neurons	Le Duigou et al. (2014)
h_0	4.20075 mV	Initial excitatory→excitatory coupling strength	Li et al. (2016); Luboeinski and Tetzlaff (2021a)
w_{ei}	$2h_0$	Excitatory→inhibitory coupling strength	Luboeinski and Tetzlaff (2021a)
w_{ie}	$4h_0$	Excitatory→inhibitory coupling strength	Luboeinski and Tetzlaff (2021a)
w_{ii}	$4h_0$	Inhibitory→inhibitory coupling strength	Luboeinski and Tetzlaff (2021a)
f_{recall}	100 Hz	Frequency of recall stimulation	Sajikumar et al. (2005); Li et al. (2016)
N_{stim}	4	Number of input neurons for stimulation	this study
r	0.5	Fraction of assembly neurons that are stimulated to trigger recall	Luboeinski and Tetzlaff (2021a)

Table 3.1: **Parameters for neuron and static network dynamics.** Values were used as given in this table, unless stated otherwise.

3.2. Manuscript: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

Symbol	Value	Description	Refs.
$t_{c,\text{delay}}$	0.0188 s	Delay of postsynaptic calcium influx after presynaptic spike	Graupner and Brunel (2012)
c_{pre}	0.6	Presynaptic calcium contribution, in vivo adjusted	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
c_{post}	0.1655	Postsynaptic calcium contribution, in vivo adjusted	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
τ_c	0.0488 s	Calcium time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_h	688.4 s	Early-phase time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_p	60 min	Protein time constant	Clopath et al. (2008); Li et al. (2016)
τ_z	60 min	Late-phase time constant	Clopath et al. (2008); Li et al. (2016)
γ_p	1645.6	Potentiation rate	Graupner and Brunel (2012); Li et al. (2016)
γ_d	313.1	Depression rate	Graupner and Brunel (2012); Li et al. (2016)
θ_p	3	Calcium threshold for potentiation	Li et al. (2016)
θ_d	1.2	Calcium threshold for depression	Li et al. (2016)
σ_{pl}	0.290436 s ^{1/2}	Standard deviation for plasticity fluctuations	Graupner and Brunel (2012); Li et al. (2016)
α	1	Protein synthesis rate	Clopath et al. (2008); Li et al. (2016)
θ_{tag}	0.0840149	Tagging threshold	Li et al. (2016)

Table 3.2: Parameters for synaptic plasticity. Values were used as given in this table, unless stated otherwise.

3.3 Supplementary Information: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

This section contains the Supplementary Information of the article presented in the previous section (Lehr et al., 2021). The typeset and referencing were adapted to match the format of this thesis.

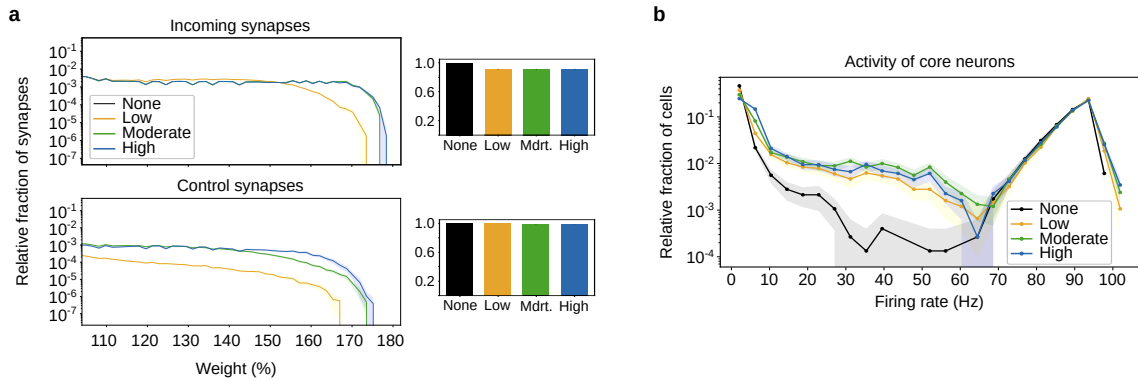


Figure 3.5: **Further characteristics of the impact of neuromodulation on consolidation of a cell assembly.** (a) Distribution of the weights of incoming and control synapses at recall after 8 h, averaged via bins of $\sim 1.64\%$. Bar plots on the right show the fraction of synapses that did not undergo substantial late-phase potentiation (< 3.26% potentiation, first two bins). (b) Distribution of the firing rates of the excitatory neurons inside the core assembly during recall 8 h after learning. Averaged via bins of 4.16 Hz across 50 networks.

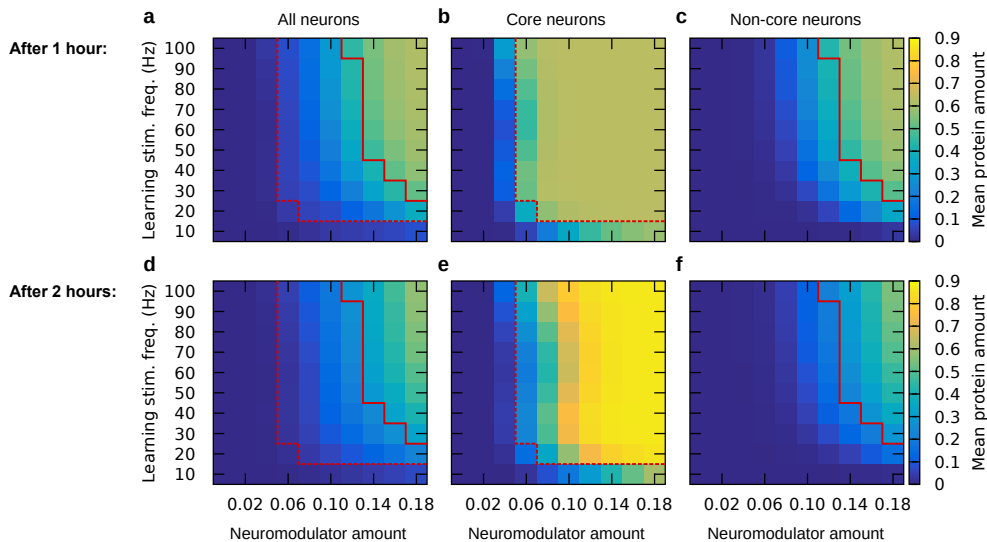


Figure 3.6: **Mean protein amount in postsynaptic neurons.** (a) One hour after learning, mean of all excitatory neurons; (b) one hour after learning, mean of all core neurons; (c) one hour after learning, mean of all non-core neurons; (d) two hours after learning, mean of all excitatory neurons; (e) two hours after learning, mean of all core neurons; (f) two hours after learning, mean of all non-core neurons. All data were averaged over 50 networks.

3.3. Supplementary Information: Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks

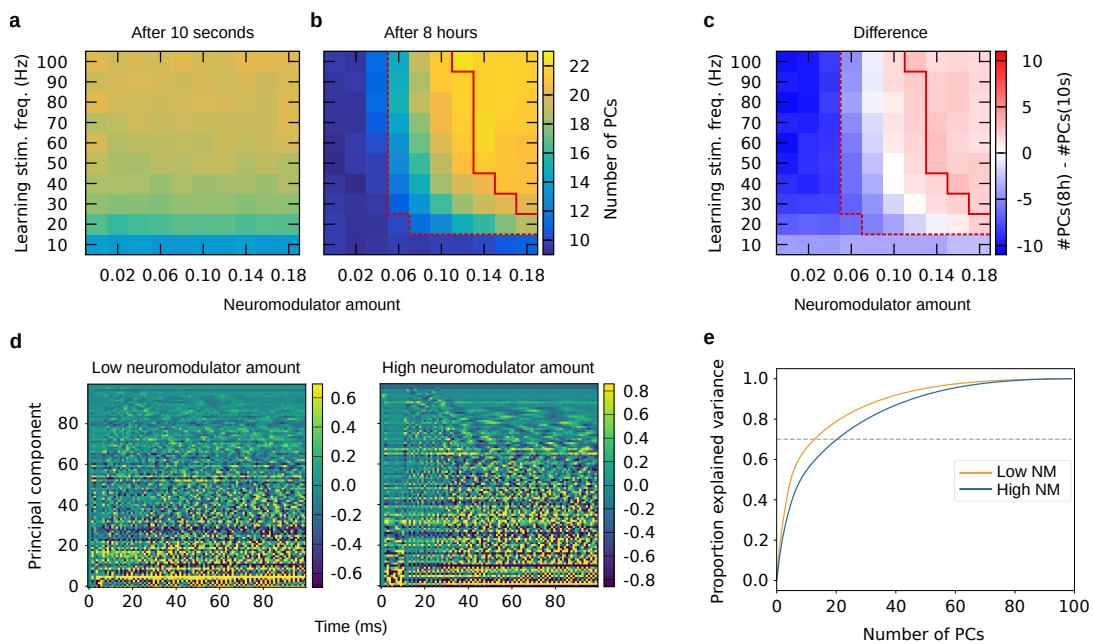


Figure 3.7: High neuromodulator levels during consolidation increases dimensionality of neuronal activity at 8h-recall. Number of principal components (PCs) required to explain 70% of variance in 1 ms-binned spike data during (a) 10s-recall and (b) 8h-recall. (c) The difference between 10s-recall and 8h-recall. (d) Examples of the transformed data (projection onto principal components) at 8h-recall, for visualization purposes clipped at the standard deviation ($\sigma_{\text{low}} = 0.69$ and $\sigma_{\text{high}} = 0.86$, $n = 10000$, 100 principal components \times 100 time bins). (e) The cumulative variance plot for low (yellow) and high (blue) neuromodulator conditions (stars in a,b) at 8h-recall, averaged over 50 networks, error bands 95% confidence interval. Gray dashed line shows 70% of variance.

4

Interaction of multiple long-term memory representations

4.1 Chapter summary

This chapter comprises an article (main manuscript: section 4.2, Supplementary Information: section 4.3) in which we investigate the capability of an STC-based network model to account for cognitive effects that involve *multiple* memory traces, following the considerations of networks with a single memory representation in chapters 2 and 3. To this end, we start from the model developed in chapter 2 and extend the size of the recurrent network to learn and consolidate three large cell assemblies (each consisting of 600 neurons). The cell assemblies need to be large because we aim to examine free recall properties here and therefore consider spontaneous reactivation rather than cued reactivation as in the previous chapters. To prevent divergent network activity when learning three large assemblies, we extend the mechanisms of early-phase plasticity in our model by a threshold that allows potentiation only in the case of sufficient pre- and postsynaptic activity, which also reflects the notion of Hebbian plasticity better than our original model (Hebb, 1949; Bliss and Collingridge, 1993; Gerstner et al., 2018). To quantify the spontaneous reactivation of cell assemblies driven by background noise, we consider neuronal avalanche characteristics (Plenz and Thiagarajan, 2007; Tetzlaff et al., 2010; Priesemann et al., 2014) in the spiking activity of the three cell assemblies. Note that although precisely-timed events such as avalanches seem to be more important for cognitive function (cf. Jezek et al., 2011; Russo and Durstewitz, 2017), our results remain similar if we consider plain average firing rates. Determining the likelihood of avalanche occurrence in different paradigms of structural assembly organization, we relate the structural interactions to functional properties that may underlie behavioral effects. Specifically, we examine *paradigms of different overlap* between different assemblies, corresponding to similarity of memories at the cognitive level (Anderson and Neely, 1996; Preston and Eichenbaum, 2013; Recanatesi et al., 2015). To form these organizational structures, we consider different simulation protocols, the most important of which are: 1. learning all three assemblies sequentially before consolidating all three, and 2. learning the three assemblies sequentially with intermediate consolidation after each individual assembly. Our first finding is that the learning order leaves its mark on the likelihood of activation of the cell assemblies, with more recently learned assemblies being activated significantly more often. While this is already the case for non-overlapping assemblies, the effect increases for assemblies that overlap. Furthermore, the effect occurs before and after consolidation, and can be related to retroactive interference, which means the disruption of older memories

by such that are newly acquired. Seeking the cause of this, we find that learning new representations induces long-term depression (LTD) in older ones, as blocking of LTD prevents the effect. This can provide a mechanistic explanation for the so-called recency effect found in free recall experiments (cf. Bjork and Whitten, 1974; Howard and Kahana, 1999; Lohnas et al., 2015). We thus approach our third research aim in this thesis: to demonstrate functional implications of STC with regard to neuropsychological effects involving multiple memory traces. Depending on the organizational paradigm, we find that the activation of a memory representation is either decreased or enhanced as compared to the non-overlapping standard case. Such differential outcomes have also been observed by neuropsychological studies on similarity between memories (cf. Anderson and Neely, 1996; Preston and Eichenbaum, 2013). Regarding hub-like organization, where one “hub” assembly connects the other two, we find that the recency/learning order effect is partially countered. However, most importantly, we find that the described effects significantly depend on the simulation protocol and thereby on the consolidation procedure, which substantiates a functional role of STC in related neuropsychological phenomena. Comparing the outcome of the two simulation protocols mentioned above, we find that STC acts by filtering out weak plastic changes (mostly LTD; cf. Figs. 4.2d, 4.6b, and 4.9b) and can thereby control the functional dynamics that arise from a certain organizational paradigm in combination with a certain simulation protocol.

We further demonstrate that our framework can be used for priming of memory representations that persists for several hours, for which no mechanistic model seems to exist so far (cf. Janiszewski and Wyer, 2014; Bermeitinger, 2015). By applying a short stimulus to the cell assembly that is to be primed, we trigger early-phase weight changes that give rise to enhanced activation of that assembly. We show that as the early-phase weight changes decay on a timescale of minutes to hours, the functional effect of priming decays on a similar timescale. In the discussion of the article, we reason that this mechanism accounting for direct (positive) priming can be extended to account for semantic or negative priming. Note that passive improvement, which we have characterized in chapter 2, occurs throughout the investigations of this chapter. Active improvement, on the other hand, depends on additional stimulation which we do not apply in our investigations of spontaneous activation. The investigations on priming, however, are closely related to active improvement: The priming stimulus causes a fast rise of early-phase weights, which is only possible after cell assemblies have been consolidated through STC mechanisms – these are the same conditions that give rise to active improvement (see chapter 2).

In chapter 3, we enabled the confinement of cell assemblies to their input-defined core by low to moderate neuromodulator amounts. In this chapter, we have enabled such confinement by introducing a firing-rate-dependent threshold for early-phase plasticity. Core confinement of cell assemblies serves here to facilitate the spontaneous reactivation of assemblies, as well as their detection, since pattern completion performed by such assemblies gives rise to attractor-like behavior (cf. Amit, 1989; also see chapter 3 and the discussion in chapter 5).

While we have so far considered activation states that are rather transient and less attractors, we explore properties of our model regarding attractor dynamics in section 4.4 of this chapter. In this section, we adapt the model from the article in section 4.2 to enable the learning and consolidation of memory representations exhibiting attractor behavior, thereby connecting to the large number of studies on the organization of memories as attractors. In particular, rate-based attractor neural networks in

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

conjunction with oscillatory input have recently been considered for the reproduction of data from free recall experiments (Romani et al., 2013; Recanatesi et al., 2015). To take this idea further with a biologically more realistic model, we present a draft of a similar approach with our model, which accounts for spiking dynamics as well as for the processes of learning and consolidating memory representations. Then, we abandon oscillatory input to again consider noisy background input. By this, we reproduce, using the attractor model, our previous findings on organizational effects (section 4.2) and thereby demonstrate that attractor dynamics may account for a mechanistic description of neuropsychological effects of recency and similarity as well as transient dynamics. Essentially, however, the attractor model enables us to examine the dynamics of transitions between activated memory representations. Using our plasticity model extended by the firing-rate threshold mentioned above, we are able to organize the memory representations such that they exhibit overlaps with strong, discretely elevated synaptic weights. We eventually find hints that such overlaps between memory representations play a significant role in noise-driven transitions between attractor states, which has been indicated by previous studies (Treves, 2005; Kropff and Treves, 2006; Recanatesi et al., 2017).

4.2 Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

This section is in content identical to the following preprint article, which has been submitted for peer review (Luboeinski and Tetzlaff, 2021b):

Luboeinski, J., Tetzlaff, C. 2021. Organization and priming of long-term memory representations with two-phase plasticity. bioRxiv doi: [10.1101/2021.04.15.439982](https://doi.org/10.1101/2021.04.15.439982).

The typeset and referencing were adapted to match the format of this thesis. The Supplementary Information of the article is provided in the next section. The references are provided at the end of this thesis. My contribution to this article was about 85%. I developed the software, performed the simulations and data analysis, and produced the figures. I drafted, wrote and reviewed this article jointly with Christian Tetzlaff. This manuscript is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

Abstract

Synaptic tagging and capture (STC) is a molecular mechanism that accounts for the consolidation of synaptic changes induced by plasticity. To link this mechanism to long-term memory and thereby to the level of behavior, its dynamics on the level of recurrent networks have to be understood. To this end, we employ a biologically detailed neural network model of spiking neurons featuring STC, which models the learning and consolidation of long-term memory representations. Using this model, we investigate the effects of different organizational paradigms of multiple memory representations, and demonstrate a proof of principle for priming on long timescales. We examine these effects considering the spontaneous activation of memory representations as the network is driven by background noise. Our first finding is that the order in which the memory representations are learned significantly biases the likelihood of spontaneous activation towards more recently learned memory representations. Secondly, we find that hub-like structures counter this learning order effect for representations

with less overlaps. We show that long-term depression is the mechanism underlying these findings, and that intermediate consolidation in between learning the individual representations strongly alters the described effects. Finally, we employ STC to demonstrate the priming of a long-term memory representation on a timescale of minutes to hours. As shown by these findings, our model provides a mechanistic synaptic and neuronal basis for known behavioral effects.

Introduction

Synaptic tagging and capture (STC) provides a well-established hypothesis of the molecular processes underlying synaptic consolidation, which has been investigated by a multitude of experimental and theoretical studies (Frey and Morris, 1997; Sajikumar and Frey, 2004a; Sajikumar et al., 2007; Clopath et al., 2008; Barrett et al., 2009; Redondo and Morris, 2011; Smolen et al., 2012; Ziegler et al., 2015; Luboeinski and Tetzlaff, 2021a). However, the dynamics resulting from the interplay between STC and recurrent network structures, which are characteristic for brain structures like the hippocampus and the neocortex (Amaral et al., 1990; Le Duigou et al., 2014), and their relation to long-term memory still remain unclear.

Neuronal activity causes the increase or decrease of the transmission strength of synapses via long-term potentiation (LTP) and long-term depression (LTD). The early phase of these phenomena typically lasts for up to a few hours (Bliss and Collingridge, 1993; Abraham, 2003). To be maintained, the synaptic changes have to undergo a consolidation process, which is described by the STC hypothesis (Frey and Morris, 1997; Redondo and Morris, 2011). When a group of neurons in a recurrent neural network experiences strong (learning) stimulation, this causes the formation of a Hebbian cell assembly, which thence represents a memory of the stimulation. Hebbian cell assemblies are groups of neurons exhibiting particularly strong synaptic connections, giving rise to facilitated activation of the assembly neurons (Hebb, 1949; Martin et al., 2000; Buzsáki, 2010; Eichenbaum, 2018). In our previous study (Luboeinski and Tetzlaff, 2021a) we developed a computational model and showed that a memory representation is first encoded by the early-phase of long-term plasticity, while STC subsequently transfers the synaptic changes to the late phase, yielding a long-term memory representation. Here we develop this model further and investigate the link between STC, properties of *multiple* long-term memories and their behavioral implication.

In the brain, spontaneous activity is ubiquitous and has been shown to inherently arise from deterministic dynamics in neural networks (Hartmann et al., 2015; Effenberger et al., 2015). In theoretical models spontaneous activity is usually elicited by background noise. This approach is biologically well-grounded through matching power spectra of in vivo neuronal activity on the one hand and of the stochastic processes used to model background noise on the other hand (Amit and Brunel, 1997b; Destexhe et al., 2003). Spontaneous activity causes the reactivation of existent cell assemblies via neuronal avalanches (Plenz and Thiagarajan, 2007; Tetzlaff et al., 2010) and thereby becomes relevant for cognition. In the current study, we use the concept of reactivation by spontaneous activity to link the reactivation dynamics of multiple long-term memories organized by STC to behavioral experiments. Namely, we compare our modeling results to two different classes of experiments from cognitive psy-

chology: free recall experiments, and priming experiments.

Free recall is a memory test that requires subjects to reproduce information from memory, for example a sequence of words, without cuing. While several models have been developed to describe free recall (Raaijmakers and Shiffrin, 1981; Bradski et al., 1994; Howard and Kahana, 2002; Romani et al., 2013; Katkov et al., 2015; Recanatesi et al., 2015), its biological underpinnings remain elusive. With respect to the serial position of the input items, free recall experiments typically exhibit three prominent characteristics (Ebbinghaus, 1885; Murdock, 1962; Bjork and Whitten, 1974; Howard and Kahana, 1999; Sederberg et al., 2010): 1. primacy, the better recall of the items learned first, 2. contiguity, the increased probability of recalling neighboring items, and 3. recency, the better recall of the most recently learned items. Primacy is commonly related to extended rehearsal, in which the recency effect could play an indirect role (Brodie and Murdock, 1977; Davelaar et al., 2005; Brown et al., 2007). Recency was first thought to be a solely short-term memory-dependent effect, but has then also been found to persist in long-term memory (Bjork and Whitten, 1974; Greene, 1986; Davelaar et al., 2005; de Almeida Valverde Zanini et al., 2012). The recency effect is usually much stronger than the primacy effect (Murdock, 1962; Bjork and Whitten, 1974). These effects are often referred to together as serial-position effect. The serial-position effect is found to be expressed differently in experiments with different delay times before recall (Glanzer and Cunitz, 1966; Bjork and Whitten, 1974; Pacheco and Verschure, 2017). Given these diverse findings, free recall experiments have essentially contributed to distinguish between a short-term and a long-term memory storage in the human brain (Howard and Kahana, 1999). Two recent studies (Romani et al., 2013; Katkov et al., 2015) have shown that long-term associations between items account for free recall rather than short-term associations as in previous models (Raaijmakers and Shiffrin, 1981; Howard and Kahana, 2002), but there is also evidence that familiarity might not essentially influence free recall (Coutanche et al., 2020). With our model, we link neuronal and synaptic dynamics including STC to emergent network principles to provide a mechanistic explanation of the recency effect in long-term memory representations.

Priming (not to be confused with primacy) describes a large set of psychological phenomena across a multitude of timescales, which can occur either consciously or unconsciously (Janiszewski and Wyer, 2014; Bermeitinger, 2015; Elgendi et al., 2018). Priming phenomena have in common that the recall of a certain memory is enhanced, following its reactivation by a “priming stimulus”. Studies on priming often target processes in the intersection between psychology and economics, for instance, bias in deciding to buy a certain product. As an example, if we see an advertisement of some brand of orange juice that we have had before, and soon afterwards go to the supermarket to do the weekly shopping, the likelihood of buying exactly that brand seen in the advertisement will be increased. Here, the advertisement acts as a priming stimulus (Nedungadi, 1990; Janiszewski and Wyer, 2014). For priming on a timescale of seconds, Mongillo et al. (2008) provided a model featuring the interplay of short-term plasticity and long-term plasticity. The authors also discussed the possibility of transferring this concept to the processes on the next, longer timescales: early- and late-phase long-term plasticity. We take on this idea and show that the STC mechanisms can account for a priming effect on a timescale of hours.

Results

In this study, we employ a biologically plausible spiking recurrent neural network model, based on a previous model (Luboeinski and Tetzlaff, 2021a) of memory consolidation via synaptic tagging and capture (STC). In a proof-of-principle approach, we investigate multiple paradigms of the formation, consolidation, and interaction of three cell assemblies in the network. Measuring the likelihood of activation of these three assemblies, we show that two factors mainly guide their activation: 1. the learning order and 2. the type of overlaps between the assemblies. By mechanistically describing the learning order effect, our model provides a possible neural correlate for the recency effect in free recall experiments. Furthermore, we encounter differential effects for hub-like overlap structures, indicating that our model may provide neural correlates for the diverse behavior observed for similarity of memories, such as retroactive interference and enhanced recall (Wickelgren, 1965; Anderson, 1981; Anderson and Neely, 1996; Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014). A major focus of our study is on the role of consolidation via STC. We show that the emergence of the learning order effect and the effects in hub-like structures are strongly biased by a protocol with altered consolidation phases. Thereby, we make predictions on the memory dynamics on different time scales, which can be tested in behavioral experiments. As a final step, we show that the interaction between early- and late-phase long-term plasticity provides a molecular mechanism for priming phenomena lasting from minutes to hours.

The neural network that we use consists of 2500 excitatory and 625 inhibitory leaky integrate-and-fire neurons. To form a cell assembly, we apply strong stimulation to a subset of the excitatory neurons in the network and thereby elicit synaptic plasticity, which is modeled by a spike-timing-dependent calcium model and a firing rate condition. This early-phase plasticity is followed by consolidation via STC, which transfers weight changes to the long-lasting late phase. The synaptic and neuronal level of the model is depicted in Fig. 4.1a, linked to a schematic of the network level in Fig. 4.1b (for more details, see Methods and Luboeinski and Tetzlaff, 2021a).

We investigated the activation of the assemblies by considering neuronal avalanches within the assemblies (Plenz and Thiagarajan, 2007). To consider an assembly being active, we utilized the border value of the 99% quantile for the number of spikes within a time window. That means, we considered the 1% of the time windows with most spikes (for more details, see Methods and Suppl. Fig. 4.15). Comparing the results for the likelihood of avalanche occurrence to the average firing rates in corresponding paradigms, we found similar behavior (Suppl. Fig. 4.10).

Learning and consolidation of multiple long-term memory representations

Before investigating the spontaneous activation of long-term memory representations, we will demonstrate how the network learns and consolidates them. We start from an unbiased network in which all weights have the same value. Then, learning proceeds according to the protocol which is shown in Fig. 4.1d, followed by consolidation and the investigation of the spontaneous activity, shown in Fig. 4.1e. Substantially extended breaks between learning stimuli, such that consolidation already oc-

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

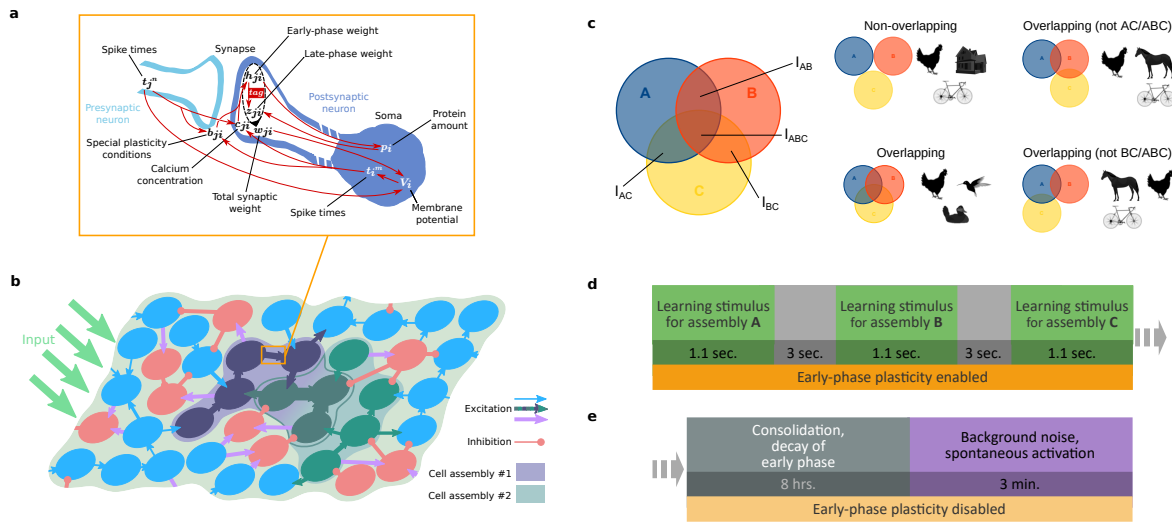


Figure 4.1: Model and considered protocols. (a,b) Schematic of the synaptic and network model (adapted from Luboinski and Tetzlaff, 2021a). The synaptic model integrates the interplay between different mechanisms of calcium-dependent and rate-dependent synaptic plasticity and the STC hypothesis. The neural network consists of excitatory (light and dark blue and green disks) and inhibitory neurons (red disks) and receives external input from other brain areas. Only synapses between excitatory neurons undergo plastic changes via the processes shown in a. Hebbian cell assemblies representing memories consist of groups of strongly interconnected neurons (dark blue, green). The assemblies in b overlap by a fraction of their neurons (dark green). For more details see the main text. (c) Different organizational paradigms are defined by different overlap relationships between the three cell assemblies. Pictures show real-life examples (all adapted from pixabay.com). (d) Protocol of the learning phase: the three memory representations are learned sequentially, with three-seconds breaks in between. (e) Protocol for the consolidation and recall phase during which consolidation takes place for eight hours after learning. Then, spontaneous activity in the network, driven by background noise, is observed for three minutes. Time spans in d and e are similar to those in free recall experiments (Murdock, 1962; Glanzer and Cunitz, 1966; Howard and Kahana, 1999).

curs in those phases, are considered in a later part of this study. In the main text, we consider four organizational paradigms, shown in Fig. 4.1c. The differences between these paradigms are visualized by small pictures. The different paradigms will be examined in the next subsection, while we restrict ourselves to the non-overlapping case in this subsection.

As a first step, after learning and consolidating three non-overlapping memory representations, we tested whether all three assemblies served as functional memory representations. To this end, we examined the ability to recall the learned and consolidated assemblies by applying stimulation to 20% of the neurons of each assembly, subsequently (Fig. 4.2a,b). We found the response caused by this stimulation to be stronger and to last longer than the response to the same stimulation in a naive network before learning (the example for A is shown in Fig. 4.2a,b in the two leftmost panels), which confirms the functionality of the memory representations. Moreover, we observed avalanches within the assemblies during spontaneous activity after learning and consolidation (Fig. 4.2a,b). Note that the synchronicity in the activity is also increased after learning and consolidation (Fig. 4.2a). Since the average excitatory synaptic weight in our network after learning and consolidation is increased (103.8%, integrated over all weights in Fig. 4.2c), this can be explained by the findings of Brunel (2000), stating that strengthening of excitatory connections in the network causes a shift towards a regime

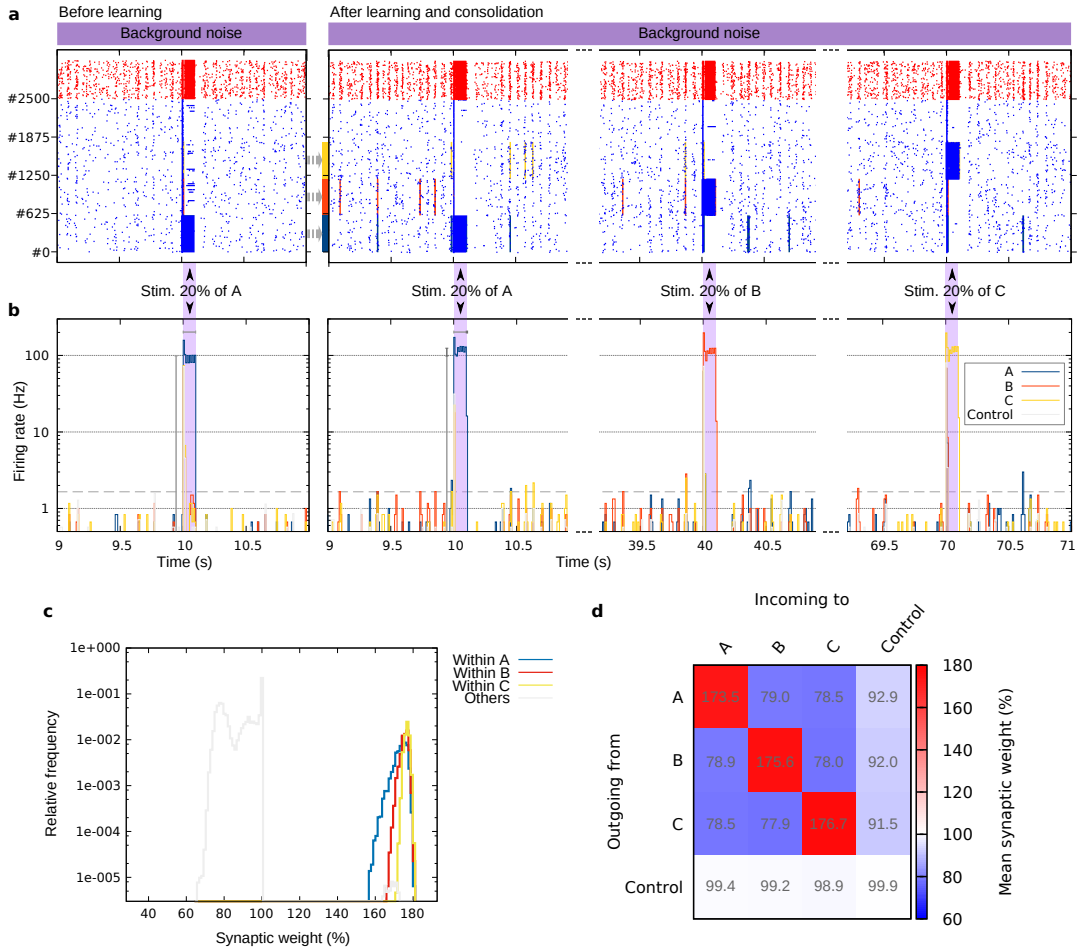


Figure 4.2: Three non-overlapping memory representations are successfully learned and consolidated. Learning order: *A-B-C*. **(a)** Raster plot showing the spiking in the whole network before and after learning and consolidation. At specified times, additional input is applied to 20% of the neurons of one assembly. Each assembly comprises 600 neurons. Spikes of excitatory and inhibitory neurons are shown by blue and red dots, respectively. Avalanches occurring in an assembly are indicated by bars in the respective color. **(b)** Firing rates corresponding to the raster plots (computed via bins of 10 ms). Shaded areas indicate the time frame of additional stimulation. The dashed line indicates the detection threshold for avalanches. In the two leftmost panels, gray bars next to the peaks emphasize the increase after consolidation in the mean value of the firing rate of *A* during the induced activation (+26.3%) and in the duration of the activation (+11.1%). **(c)** Weight distribution of the synapses within the assemblies, revealing the spread of LTP and LTD. **(d)** Abstract weight matrix showing the mean weight within and between all subpopulations of the excitatory population. The asymmetric nature of the matrix reveals the directionality of the coupling strengths. Data in **c** and **d** were averaged over 10 trials after learning and consolidation, and are relative to the initial value before learning.

of synchronized regular activity. Thereby, we conclude that our network operates close to the phase transition, and such criticality is typical for the occurrence of avalanches (Plenz and Thiagarajan, 2007; Tetzlaff et al., 2010).

In addition to the activity, we scrutinized the weight structure in the network after learning and consolidation. As expected, we found that the weights within each assembly are much higher than the weights in the remainder of the network (Fig. 4.2c). Remarkably, however, the weight distributions of

the three assemblies differ. This is already a hint to the bias in assembly activation introduced by the learning order, which we will investigate in the following subsection. Complementary to the weight distribution, the abstract weight matrix in Fig. 4.2d reveals the directionality of the coupling strengths within and between subpopulations of the network after learning and consolidation. The matrix reflects the highly potentiated weights within each cell assembly, but also the depressed weights between the assemblies. Presumably, this depression arises from activation of the other assemblies upon learning a new assembly. The weights involving control neurons, on the other hand, exhibit only slight plasticity. Remarkably, synapses incoming from assembly neurons to control neurons are on average slightly depressed, while synapses outgoing from control neurons remain unchanged. Comparing with the weights after learning but before consolidation (Suppl. Fig. 4.6) reveals that at that time there is also depression in the synapses outgoing from control neurons, which is yet too weak to become consolidated. The depression of the weights presynaptic to the control neurons probably also emerges during learning of the assemblies, while the weights postsynaptic to the control neurons only receive indirect activation during learning. The higher calcium increase through presynaptic spikes in our model might increase this effect further.

For the overlapping paradigm, the weight distribution and the abstract weight matrix are given in Suppl. Figs. 4.7 and 4.8. Those results reveal further asymmetries for the weights of the intersections between assemblies, exemplifying the richness of our model.

Learning order and overlaps guide spontaneous activity

After the learning and consolidation of the three assemblies, we systematically investigated the spontaneous reactivation of the assemblies driven by background noise. We collected data from three minutes, which is a time span similar to the recall period in free recall experiments (Murdock, 1962; Glanzer and Cunitz, 1966; Howard and Kahana, 1999). A sample of spontaneous activity and the detected avalanches is shown in Fig. 4.3a,b.

Overlaps of memory representations on the network level are of huge relevance because they form the basis of memory similarity on the cognitive level (Anderson and Neely, 1996; Romani et al., 2013; Preston and Eichenbaum, 2013; Recanatesi et al., 2015). Here, we consider four different paradigms of overlaps, as they are shown in Fig. 4.1c (also see Methods). Small images picture real-life examples of non-overlapping and overlapping memory representations: representations of entities from three different categories, like chicken, house, and bicycle, do not overlap, while representations from the same category, like three types of birds, overlap. Mixed cases occur if two entities are in the same category while there is no common category for all three, for instance, bicycle and horse are both means of transportation, and chicken and horse are both farm animals. In the Supplementary Information to our study, we present results on more paradigms, including such with overlaps of varied percentage (Suppl. Figs. 4.10 and 4.11). For an overview over these paradigms, see Suppl. Table 4.3. The most important result from these additional paradigms is that in the majority of cases, larger overlaps further increase the effects that we describe in the following.

Considering the mean likelihood of activation in each of the four paradigms from Fig. 4.1c with

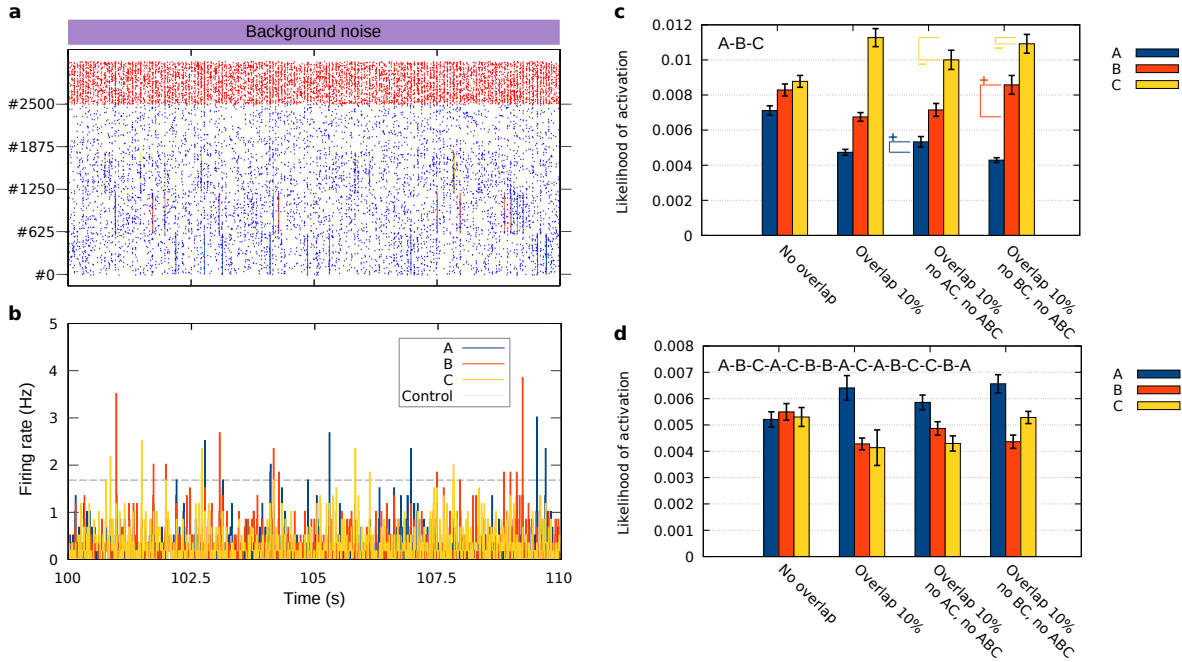


Figure 4.3: Spontaneous activity of long-term memory representations, driven by background noise. (a) Raster plot showing the spontaneous spiking activity in the whole network induced by background noise. Each assembly comprises 600 excitatory neurons (spikes indicated by blue dots). Spikes of inhibitory neurons are shown by red dots. Avalanches occurring in the assemblies are indicated by bars in the assembly colors. Assemblies do not overlap and are learned in the order: *A-B-C*. **(b)** Firing rates corresponding to the raster plot (computed via bins of 10 ms). The dashed line indicates the threshold for the detection of an avalanche. **(c)** Overview of the likelihood of avalanches in different organizational paradigms, learned with standard protocol (order: *A-B-C*). Averaged over mean likelihood from 10 networks. Error bars show the 95% confidence interval. Braces and plus/minus signs indicate differences to the “Overlap 10%” paradigm. **(d)** Overview of the likelihood of avalanches as in **c**, but learned with extended, interleaved protocol (order: *A-B-C-A-C-B-B-A-C-A-B-C-C-B-A*). Data in **c** and **d** were averaged over 10 trials.

assemblies that were learned in the order *A-B-C*, we encounter the first main result of our study: the learning order significantly biases the spontaneous activity, such that the activation of more recently learned assemblies is promoted (shown in Fig. 4.3c). While this is already the case without overlap, overlap between assemblies increases the effect even further.

Our second finding is indicated in Fig. 4.3c by braces with a plus or minus sign. The braces demonstrate differences between the paradigms with less overlaps and the “Overlap 10%” paradigm. The lack of an overlap between two assemblies (in the abundance of the other overlaps) creates a hub-like structure, where the assembly that overlaps with the two other assemblies is the hub. We find that such hub-like structure reduces the learning order effect for the two assemblies with less overlaps. This means, the activation of the assembly that is disadvantaged by the learning order effect (*A* in the “no AC” paradigm, *B* in the “no BC” paradigm) is increased, while the activation of the assembly favored by the learning order effect (*C*) is decreased. As we will show below, this effect as well as the learning order effect itself are essentially caused by LTD.

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

For gathering further knowledge about the learning order effect and to examine other effects without the bias of the learning order, we tried to eradicate the learning order effect (also cf. the following subsection on priming). To this end, we employed a learning protocol for the interleaved learning of the three assemblies, using an extended sequence with the order $A-B-C-A-C-B-B-A-C-A-B-C-C-B-A$ (random with the constraint that every assembly occurs equally often). Compared to Fig. 4.1d, the interleaved protocol simply contains more steps, but otherwise stays the same. In the non-overlapping paradigm, the interleaved learning successfully eradicates any significant impact of the learning order (Fig. 4.3d). In the other paradigms, however, a certain impact remains. Note that the most recently learned assembly in the extended sequence is A , which is also the mostly reactivated assembly.

In the next step, we considered a control case for the learning order $A-B-C$ in which we blocked LTD, and found what had already been hinted by the distribution of weights resulting from the standard protocol (inset in Fig. 4.4a): the learning order effect vanishes if there is no LTD (Fig. 4.4a,c). Furthermore, the effect in hub-like structures that we discussed previously, vanishes as well. Instead, the activation of the hub assembly (B in the “no AC” paradigm, A in the “no BC” paradigm) clearly stands out, as it is expected for an LTP-only model. In addition, the overall likelihood of activation increases.

Next, we investigated the activation of assemblies that had undergone intermediate consolidation after being learned, meaning that the breaks shown in Fig. 4.1d lasted for 8 hours instead of 3 seconds. Although LTD was switched on for these investigations, much of the learning order effect vanishes as well (Fig. 4.4b,d). The effect of hub-like structures is also similar to the no-LTD case. It seems that here, LTD is not significantly expressed in the synapses within the assemblies, even though LTD occurs in other synapses (cf. Fig. 4.4b and inset in Fig. 4.4a). Apparently, if only one assembly is learned before the next one follows after hours, much of the depression which evokes the learning order effect is too weak to become consolidated (also cf. Suppl. Figs. 4.6 and 4.9).

These results indicate that diverse effects arise from the interaction of plasticity processes on different timescales, triggered by different learning/consolidation protocols. While the standard protocol causes both late-phase LTP and substantial late-phase LTD, leading to the described strong learning order effect and its reduction for assemblies with less overlaps, the intermediate consolidation protocol rather causes operation in the regime of late-phase LTP, leading to a weaker learning order effect and promoted activation of hub assemblies.

Priming on long timescales

After investigating the functional implications of different organizational paradigms, in this subsection, we provide a proof of principle to show that our two-phase plasticity model is able to describe priming on long timescales. To this end, we use a protocol that includes a brief additional stimulus delivered to all neurons of one selected assembly after learning and consolidation (Fig. 4.5a). This stimulation causes the respective assembly to be transiently primed for later reactivation, which we investigated by considering the likelihood of avalanches under background noise, as it was done in the previous section. Our results demonstrate a mechanism for priming enabled by the interplay of

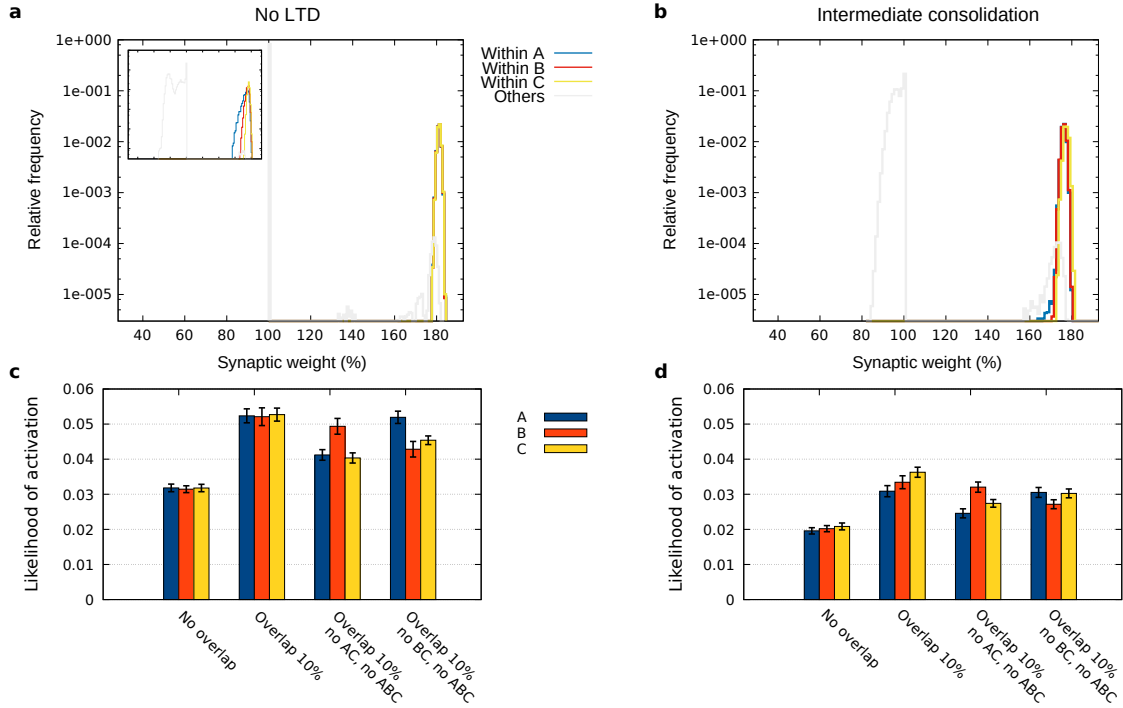


Figure 4.4: Scrutinizing the underpinnings of the learning order and overlap effects. (a) Weight distribution of the synapses within the assemblies, revealing that if LTD is blocked, the synapses belonging to the three assemblies are equally potentiated (non-overlapping case). The inset repeats the plot from Fig. 4.2c for comparison. (b) Weight distribution of the synapses within the assemblies learned with a modified standard protocol where the breaks between learning *A*, *B*, and *C* last 8 hours instead of 3 seconds (intermediate consolidation). The plot reveals a similar structure as in the no-LTD case. (c) Overview of the likelihood of avalanches in different organizational paradigms, learned with standard protocol (*A*-*B*-*C*), with LTD blocked. (d) Overview of the likelihood of avalanches in different organizational paradigms for the intermediate consolidation case (order *A*-*B*-*C*). Data were averaged over 10 trials. Error bars show the 95% confidence interval. Weight values are relative to the initial value before learning.

early- and late-phase long-term plasticity (Fig. 4.5b-e). The priming decays on a timescale of minutes to hours. Interestingly, in the non-overlapping paradigm, the priming stimulation causes the primed memory representation to behave in an almost attractor-like manner (Fig. 4.5b, where the likelihood of activation reaches values of more than 0.4).

For investigating priming in the absence of other effects, we sought to attenuate the learning order effect and therefore used the extended interleaved protocol introduced in the previous section. For the non-overlapping case, this approach works fine as it almost entirely eliminates the influence of the learning order (Fig. 4.5b). However, in the overlapping case (Fig. 4.5c), the learning order effect still plays a role, since in the basal state, *A* is much more active than *B* and *C*. The priming of *B* shows a significant effect in the first hour, but then the effect is overshadowed by the high activity of assembly *A*. Finally, we have to mention that in rare cases, priming stimulation after learning with the interleaved protocol seems to disrupt the recallable memory structure and render reactivation impossible (this occurred in 1 of 10 trials).

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

The mean weight plotted in Fig. 4.5d,e shows roughly the same qualitative time course as the activation (Fig. 4.5b,c). It should be noted that in Fig. 4.5d, the mean weight in assembly *B* decreases to a lower value than it had before priming, which can be explained by the strong depression of a few synapses that thus enter the late phase, while most synapses are subject to early-phase potentiation.

Data for more paradigms are shown in Suppl. Fig. 4.12. The case where only 50% of the assembly neurons receive the priming stimulus is shown in Suppl. Fig. 4.13 (more paradigms in Suppl. Fig. 4.14). This case leads to the counterintuitive but interesting outcome of priming by differential degree of LTD. Further investigations on this are, however, beyond the scope of this study.

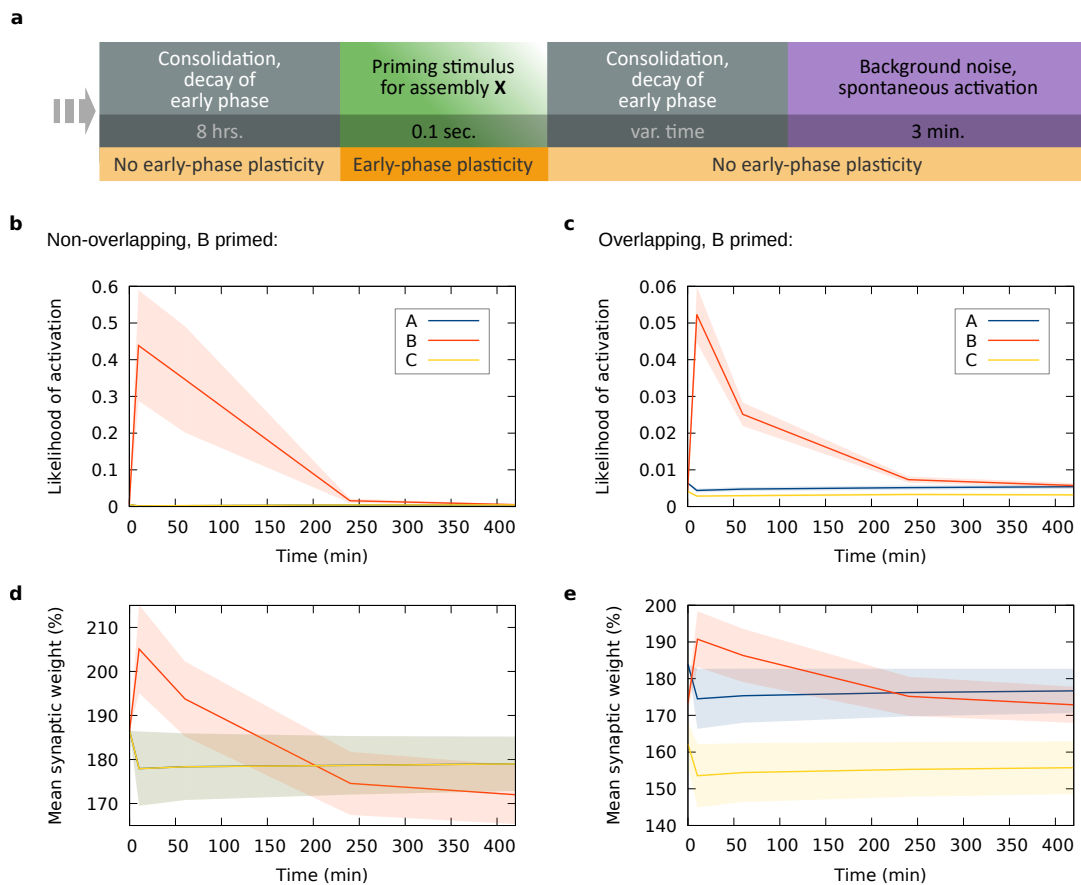


Figure 4.5: Priming on long timescales is enabled by synaptic tagging and capture. (a) Consolidation and priming phase (following learning as sketched in Fig. 4.1d, with order *A-B-C-A-C-B-B-A-C-A-B-C-C-B-A*): consolidation takes place for eight hours after learning. After that, a priming stimulus is applied to one of the assemblies, followed by another period of consolidation and early-phase decay. Then, driven by background noise, spontaneous activity of the memory representations is observed for three minutes. (b,c) Resulting likelihood of avalanches in the three assemblies at different times after priming *B*. The values at time zero show the case before (without) priming. (d,e) Temporal development of the mean synaptic weight in the assemblies, for the same cases as in b and c. Data in panels b-e are averaged over 10 networks. Error bands show standard error of the mean. Weight values are relative to the initial value before learning.

Discussion

In this study, we have shown by proof of principle that the molecular mechanisms of calcium-based synaptic plasticity and STC enable to robustly learn, consolidate, organize, and prime long-term memory representations. Our results demonstrate that these features give rise to specific characteristics as they are found in neuropsychological experiments.

Measuring the likelihood of spontaneous activation of three cell assemblies in a spiking recurrent neural network after learning and consolidation, we found two factors that guide the activation: 1. the learning order, and 2. the type of overlaps between the assemblies. As an experimental study has shown, overlap accounts for memory association but is not necessary for the recall of the original separate memories, which matches our findings for the non-overlap and overlap paradigms (Yokose et al., 2017). Moreover, there is experimental evidence that memories can be potentiated and depotentiated independently of each other (Abdou et al., 2018), which also seems to be possible in our model. Data from free recall experiments typically exhibit primacy (high probability of activating the first learned items), recency (high probability of activating the last learned items), and contiguity (higher probability of transition between neighboring items; Murdock, 1962; Howard and Kahana, 1999). By our model, we provide a mechanistic explanation for the recency effect. The direct reproduction of data from free recall experiments (Sederberg et al., 2010; Lohnas et al., 2015; Long et al., 2015) with our model would, however, require a larger number of cell assemblies, which was beyond the capacity of this study. Since our detailed network model had a high demand of computational power, we had to restrict our investigations to a medium-sized network of 3125 neurons and three assemblies. Nevertheless, direct matching of a larger network based on our model and data from free recall experiments could be the goal of a future study. Extending our model to more assemblies would also be very helpful for investigating and modeling contiguity, complementary to previous models (Katkov et al., 2015; Recanatesi et al., 2015). Another interesting approach in a future study could be to adjust our network such that the memory representations constitute attractors. This would enable the study of the transitions between attractor states, which have been proposed to account for the switching between concepts (Russo and Treves, 2012; Recanatesi et al., 2015, 2017), and would build on the vast amount of previous studies on attractor neural networks.

In addition to recency in free recall, our model may provide a mechanistic basis for retroactive interference (Wickelgren, 1965; Anderson, 1981; Anderson and Neely, 1996) and enhanced recall of similar memories (Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014). Our findings show that the activation of memory representations featuring hub-like overlaps is enhanced or attenuated, depending on the consolidation protocol (Figs. 4.3c and 4.4d). Moreover, for longer intervals between learning stimuli but otherwise equal parameters, “memory-improving” interference occurs (Suppl. Fig. 4.11) instead of the “memory-deteriorating” interference that results from the standard protocol (Suppl. Fig. 4.10). From our theoretical investigations, we made the prediction that the effects described above are essentially caused by LTD (Fig. 4.4a,c). This can be tested experimentally by a protocol that blocks LTD, for example, via the mGlu receptor antagonist MCPG, or via BDNF (Cho et al., 2001; Ikegaya et al., 2002). How our findings relate to experiments showing that the acquisition of overlapping assemblies sufficiently close in time facilitates recall, while memories acquired further apart in time tend

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

to be separated (Cai et al., 2016; Rashid et al., 2016), remains to be shown in future investigations. Furthermore, the experimental investigation of our “intermediate consolidation” protocol could yield a large gain in knowledge. The difference between this protocol and typical free recall protocols is that after learning a memory representation we wait until it is consolidated (i.e., for about 8 hours), and then learn the next representation, whereas in free recall experiments, the whole list of items is usually learned in short time before consolidation may take place (similar to our “standard” protocol). The experimental investigation of the “intermediate consolidation” protocol would enable to test our predictions and yield further interesting insights in the interaction of short- and long-term memory for free recall. An issue could be that the subjects would have to stay awake for the whole duration of the protocol of about 24 hours, to avoid a biasing effect from sleep consolidation. However, similar experiments have already been performed to test the impact of sleep deprivation on free recall (de Almeida Valverde Zanini et al., 2012; Ren and Coutanche, 2020).

In our simulations, we assumed early-phase plasticity to be disabled during certain phases. As we pointed out in the Methods section, this is justified by a lack of novelty or attention, which goes along with lowered neuromodulator concentrations (Lisman et al., 2011; Duzskiewicz et al., 2019), and thereby prevents learning (Lindskog et al., 2006; Pezze and Bast, 2012). If, however, early-phase plasticity were enabled all the time, spontaneous activity could exert the following influence on the structure of a cell assembly: either the assembly would stay unaffected, as we have shown for smaller cell assemblies in our previous study (Luboeinski and Tetzlaff, 2021a), or, LTD would cause a degradation by downregulating the synaptic weights, or, LTP would cause reinforcement by increasing the synaptic weights. While the latter is not possible due to the firing rate threshold in our model, we assume that spontaneous occurrence of LTD, if any, would not qualitatively change our results.

In the second part of our study, we demonstrated the priming of one of the three previously learned and consolidated memory representations, on a timescale of minutes to hours. This effect is essentially caused by the interplay between early- and late-phase long-term plasticity. After consolidation, the structure of the memory representations is permanently stored by the late-phase weights, while additional transient information, like priming information, is stored by the early-phase weights, which decay across minutes to hours. However, as our results also indicate, priming information may “spill over” to the late phase, causing permanent changes in the structure of a memory representation. In our study, the priming effect is strongest if priming is only 10 minutes ago, and declines until it has vanished almost completely after 7 hours (Fig. 4.5b,c). This time dependence, as well as the impact of priming in general, may be affected by the strength and duration of the priming stimulus which “loads the representation into short-term memory”. Moreover, the effect of the time delay may be influenced by a wide range of further experimental conditions, making it difficult to draw a general conclusion, which was pointed out by others earlier (Lerner and Shriki, 2014). As our results constitute a proof of principle, we leave it to further studies to investigate the influence of different kinds of stimulation. Furthermore, while we showed that our model can account for direct positive priming, our model can likely be extended with relatively low effort to account for other types of priming. We already encountered predominance of LTD in the case where only 50% of the assembly neurons received priming stimulation (Suppl. Figs. 4.13 and 4.14). While this paradigm still causes a slightly positive priming

effect, lowering this fraction even further or simply applying background activity in the presence of early-phase plasticity should revert the outcome and cause a negative priming effect. Such further results would be of great importance, especially because a mechanistic biological theory for negative priming still remains to be found (Mayr and Buchner, 2007; Ihrke et al., 2011; Elgendi et al., 2018). Moreover, a tuned version of our model could possibly account for semantic priming, which would be the case if the stimulation of an assembly that overlaps with a second assembly would cause primed activation of the second assembly (e.g., in the way that the word “chicken” would prime the word “duck”; Herrmann et al., 1993; Becker et al., 1997; Lerner and Shriki, 2014; Bermeitinger, 2015). Finally, the fact that the interplay of early- and late-phase plasticity enables priming is another hint that they might constitute correlates for the concepts of retrieval strength and storage strength (Bjork and Bjork, 1992; Luboinski and Tetzlaff, 2021a).

To summarize, we could show that LTD on the one hand and the interplay of early- and late-phase long-term plasticity on the other hand exert crucial influence on memory function. Our results provide neural correlates to – at least partially – explain different behavioral effects. In the future, our model may serve as a mechanistic basis for the detailed reproduction of findings obtained in behavioral experiments.

Methods

Model

To simulate the dynamics of memory representations, we used a network model that comprises spiking neurons and synapses with detailed plasticity features. In this section, we first present our mathematical description of neurons and synapses, which is depicted in Fig. 4.1a, as well as the structure of our network at the population level. After that, we explain the different overlap paradigms and stimulation protocols. The parameters that we used are given in Tables 4.1 and 4.2. Parts of this section have been adapted from our previous study (Luboinski and Tetzlaff, 2021a), which provided the basis for our model.

Neuron model. The dynamics of the membrane potential $V_i(t)$ of the Leaky Integrate-and-Fire neuron i is described by Gerstner and Kistler (2002):

$$\tau_m \frac{dV_i(t)}{dt} = V_{\text{rev}} - V_i(t) + \sum_j \sum_{t_j^k} w_{ji} \cdot \exp\left(-\frac{(t - t_j^k - t_{\text{ax,delay}})}{\tau_{\text{syn}}}\right) + R_m \left(I_{\text{bg}}(t) + I_{\text{stim}}(t)\right) \quad (4.1)$$

with reversal potential V_{rev} , membrane time constant τ_m , membrane resistance R_m , synaptic weights w_{ji} , spike times t_j^k , axonal delay time $t_{\text{ax,delay}}$, synaptic time constant τ_{syn} , external background current $I_{\text{bg}}(t)$, and external stimulus current $I_{\text{stim}}(t)$. If V_i crosses the threshold V_{th} , a spike is generated. The spike time t_i^l is then stored and the membrane potential is reset to V_{reset} , where it remains for the refractory period t_{ref} . Under basal conditions, the membrane potential dynamics is mainly driven by an external background current that accounts for synaptic inputs from outside the network, described

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

by an Ornstein-Uhlenbeck process:

$$\tau_{\text{syn}} \frac{dI_{\text{bg}}(t)}{dt} = I_0 - I_{\text{bg}}(t) + \sigma_{\text{wn}} \cdot \Gamma(t) \quad (4.2)$$

Symbol	Value	Description	Refs.
Δt	0.2 ms	Duration of one time step for numerical computation	this study
τ_m	10 ms	Membrane time constant	Dayan and Abbott (2001); Gerstner and Kistler (2002)
τ_{syn}	5 ms	Synaptic time constant, also for external input current	Dayan and Abbott (2001); Gerstner and Kistler (2002); Roth and van Rossum (2009)
$t_{\text{ax,delay}}$	3 ms	Axonal spike delay	Gerstner and Kistler (2002); Lin and Faber (2002)
t_{ref}	2 ms	Duration of the refractory period	Dayan and Abbott (2001); Kobayashi et al. (2009)
R_m	10 M Ω	Membrane resistance	Dayan and Abbott (2001)
V_{rev}	-65 mV	Reversal (equilibrium) potential	Dayan and Abbott (2001)
V_{reset}	-70 mV	Reset potential	Dayan and Abbott (2001)
V_{th}	-55 mV	Threshold potential to be crossed for spiking	Dayan and Abbott (2001)
I_0	0.15 nA	Mean of the external background current	Luboeinski and Tetzlaff (2021a)
σ_{wn}	0.05 nA s ^{1/2}	Standard deviation for Gaussian noise in the background current	Luboeinski and Tetzlaff (2021a)
N_e	2500	Number of neurons in the excitatory population	this study
N_i	625	Number of neurons in the inhibitory population	this study
p_c	0.1	Probability of a connection existing between two neurons	Le Duigou et al. (2014)
h_0	4.20075 mV	Initial excitatory→excitatory coupling strength	Li et al. (2016); Luboeinski and Tetzlaff (2021a)
w_{ei}	2 h_0	Excitatory→inhibitory coupling strength	Luboeinski and Tetzlaff (2021a)
w_{ie}	3.5 h_0	Excitatory→inhibitory coupling strength	this study
w_{ii}	3.5 h_0	Inhibitory→inhibitory coupling strength	this study
f_{stim}	100 Hz	Frequency of stimulation	Sajikumar et al. (2005); Li et al. (2016)
N_{stim}	25	Number of input neurons for stimulation	Luboeinski and Tetzlaff (2021a)
n_{thresh}	10	Detection threshold for an avalanche	this study

Table 4.1: Parameters for neuron and static network dynamics. Values were used as given in this table, unless stated otherwise.

with mean current I_0 and white-noise standard deviation σ_{wn} . In this equation, $\Gamma(t)$ is Gaussian white noise with mean zero and variance $1/dt$ that approaches infinity for $dt \rightarrow 0$. The Ornstein-Uhlenbeck process has the same colored-noise power spectrum as the fluctuating input to cortical neurons coming from a large presynaptic population, and is therefore well-suited to model the background noise in our model (Destexhe et al., 2003).

In addition to the background noise, we use another Ornstein-Uhlenbeck process to model the stimulus current $I_{stim}(t)$ for learning, recall, and priming (cf. the subsection ‘Learning, recall, and priming stimulation’ below):

$$\tau_{syn} \frac{dI_{stim}(t)}{dt} = w_{stim} \cdot N_{stim} \cdot f_{stim} - I_{stim}(t) + w_{stim} \cdot \sqrt{N_{stim} \cdot f_{stim}} \cdot \Gamma(t). \quad (4.3)$$

Symbol	Value	Description	Refs.
$t_{c,delay}$	0.0188 s	Delay of postsynaptic calcium influx after presynaptic spike	Graupner and Brunel (2012)
c_{pre}	0.6	Presynaptic calcium contribution, in vivo adjusted	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
c_{post}	0.1655	Postsynaptic calcium contribution, in vivo adjusted	Graupner and Brunel (2012); Higgins et al. (2014); Li et al. (2016)
τ_c	0.0488 s	Calcium time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_h	688.4 s	Early-phase time constant	Graupner and Brunel (2012); Li et al. (2016)
τ_p	60 min	Protein time constant	Li et al. (2016); Clopath et al. (2008)
τ_z	60 min	Late-phase time constant	Clopath et al. (2008); Li et al. (2016)
ν_{th}	40 Hz	Firing rate threshold for LTP induction	Bear (1996); Sajikumar et al. (2007)
γ_p	1645.6	Potential rate	Graupner and Brunel (2012); Li et al. (2016)
γ_d	313.1	Depression rate	Graupner and Brunel (2012); Li et al. (2016)
θ_p	3	Calcium threshold for potentiation	Li et al. (2016)
θ_d	1.2	Calcium threshold for depression	Li et al. (2016)
σ_{pl}	$0.290436 \text{ s}^{1/2}$	Standard deviation for plasticity fluctuations	Graupner and Brunel (2012); Li et al. (2016)
α	1	Protein synthesis rate	Clopath et al. (2008); Li et al. (2016)
θ_{pro}	0.210037	Protein synthesis threshold	Li et al. (2016)
θ_{tag}	0.0840149	Tagging threshold	Li et al. (2016)

Table 4.2: Parameters for synaptic plasticity. Values were used as given in this table, unless stated otherwise.

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

Mean and standard deviation of this process are defined by putative input spikes from N_{stim} neurons, occurring at the frequency f_{stim} and conveyed by synapses of weight w_{stim} to the neurons of our network that are to be stimulated.

Model of synapses and STC. All spikes k that occur in neuron j are transmitted to i , if there is a synaptic connection from neuron j to neuron i . The postsynaptic current caused by a presynaptic spike then depends on the total weight or strength of the synapse, which is given by:

$$w_{ji} = \begin{cases} h_{ji} + h_0 \cdot z_{ji} & \text{for E} \rightarrow \text{E}, \\ w_{ei} & \text{for E} \rightarrow \text{I}, \\ w_{ie} & \text{for I} \rightarrow \text{E}, \\ w_{ii} & \text{for I} \rightarrow \text{I}. \end{cases} \quad (4.4)$$

While in the model, all synaptic connections involving inhibitory neurons are constant, the weight of E→E connections comprises two variable contributions to account for STC: the early-phase weight h_{ji} , and the late-phase weight z_{ji} . The dynamics of the early-phase weight is given by

$$\tau_h \frac{dh_{ji}(t)}{dt} = 0.1 (h_0 - h_{ji}(t)) + b_{ji}(t) \cdot (\gamma_p (1 - h_{ji}(t)) \cdot \Theta[c_{ji}(t) - \theta_p] - \gamma_d h_{ji}(t) \cdot \Theta[c_{ji}(t) - \theta_d] + \zeta(t)), \quad (4.5)$$

where $\Theta[\cdot]$ is the Heaviside function, τ_h is a time constant, $c_{ji}(t)$ is the calcium concentration at the postsynaptic site, and $b_{ji}(t)$ describes special additional conditions for the induction of plasticity. The first term on the right-hand side describes the decay of the early-phase weight to its initial value h_0 , the second term describes early-phase LTP with rate γ_p , given that calcium is above the threshold θ_p , the third term describes early-phase LTD with rate γ_d , given that calcium is above the threshold θ_d , and the fourth term $\zeta(t) = \sqrt{\tau_h [\Theta(c_{ji}(t) - \theta_p) + \Theta(c_{ji}(t) - \theta_d)]} \sigma_{pl} \Gamma(t)$ describes calcium-dependent fluctuations with standard deviation σ_{pl} and Gaussian white noise $\Gamma(t)$ with mean zero and variance $1/dt$. The calcium concentration $c_{ji}(t)$ at the postsynaptic site is driven by the pre- and postsynaptic spikes n and m , respectively:

$$\frac{dc_{ji}(t)}{dt} = -\frac{c_{ji}(t)}{\tau_c} + c_{\text{pre}} \sum_n \delta(t - t_j^n - t_{c,\text{delay}}) + c_{\text{post}} \sum_m \delta(t - t_i^m), \quad (4.6)$$

where $\delta(\cdot)$ is the Dirac delta distribution, τ_c is a time constant, c_{pre} is the contribution of presynaptic spikes, c_{post} is the contribution of postsynaptic spikes, t_j^n and t_i^m are spike times, and $t_{c,\text{delay}}$ is the delay of calcium triggered by presynaptic spikes.

In addition to the calcium dependence of early-phase plasticity, which is based on previous studies (Graupner and Brunel, 2012; Higgins et al., 2014; Li et al., 2016; Luboinski and Tetzlaff, 2021a), we imposed special conditions for the occurrence of plasticity to enable the stable formation, consolidation, and activation of large cell assemblies. In their original study, Graupner and Brunel (2012) had only considered firing rates up to 50 Hz. A drawback in their calcium-based model is the chance that spiking at high rates on only one side of the synapse would suffice to trigger LTP, which is physiologically

not plausible. To account for the evidence that the firing of only one neuron (i.e., on one side of the synapse) at high rates should not be sufficient to evoke LTP (Hebb, 1949; Bliss and Collingridge, 1993; Abraham et al., 2019), we introduced an additional threshold ν_{th} which allows the induction of LTP only if both pre- and postsynaptic firing rate cross this threshold. Furthermore, we blocked early-phase plasticity outside the learning and priming phases to enable fast computation and stable activation dynamics. This is explained by a lack of novelty outside the learning and priming phases, which goes along with lowered neuromodulator concentrations (Lisman et al., 2011; Duzskiewicz et al., 2019) and thereby prevents learning, as experimental studies have demonstrated (Pezze and Bast, 2012; Lindskog et al., 2006). In summary, the two “special plasticity conditions” that we added to the model can be put down formally in the following way:

$$b_{ji}(t) = (\Theta[c_{ji}(t) - \theta_p] \cdot \Theta[v_i(t) - \nu_{th}] \cdot \Theta[v_j(t) - \nu_{th}] + \Theta[c_{ji}(t) - \theta_d] \cdot \Theta[\theta_p - c_{ji}(t)]) \cdot NM, \quad (4.7)$$

where the first term in the round brackets describes the LTP induction depending on the calcium concentration c_{ji} and the firing rate threshold ν_{th} , and the second term describes the LTD induction which only depends on calcium. The firing rate function $v_i(t)$ describes the firing rate of neuron i at time t , computed using a sliding window of 0.5 s. The factor NM is the neuromodulatory impact, which we assume unity during learning and priming phases and zero otherwise.

Driven by the calcium-based early-phase dynamics, the late-phase synaptic weight is given by:

$$\begin{aligned} \tau_z \frac{dz_{ji}(t)}{dt} = & p_i(t) \cdot (1 - z_{ji}(t)) \cdot \Theta[(h_{ji}(t) - h_0) - \theta_{tag}] \\ & - p_i(t) \cdot (z_{ji} + 0.5) \cdot \Theta[(h_0 - h_{ji}(t)) - \theta_{tag}]. \end{aligned} \quad (4.8)$$

with the protein amount $p_i(t)$, the late-phase time constant τ_z , and the tagging threshold θ_{tag} . The first term on the right-hand side describes late-phase LTP, while the second one describes late-phase LTD. Both depend on the amount of proteins being available. If the early-phase weight change $|h_{ji}(t) - h_0|$ exceeds the tagging threshold, the synapse is tagged. This can be the case either for positive or for negative weight changes. The presence of the tag leads to the capture of proteins (if $p_i(t) > 0$), and thereby gives rise to changes in the late-phase weight.

The synthesis of new proteins requires sufficient changes in early-phase weights, while there is also an inherent decay of the protein amount (Clopath et al., 2008):

$$\tau_p \frac{dp_i(t)}{dt} = -p_i(t) + \alpha \Theta \left[\left(\sum_j |h_{ji}(t) - h_0| \right) - \theta_{pro} \right]. \quad (4.9)$$

Population structure. Using the neuron model and the synapse model explained above, we set up a neural network consisting of 2500 excitatory and 625 inhibitory neurons (depicted in Fig. 4.1b). The ratio of 4:1 between excitatory and inhibitory neurons is typical for cortical and hippocampal networks (Braitenberg and Schüz, 1998). Some of the excitatory neurons receive specific inputs for the formation of cell assemblies (see the subsection ‘Learning, recall, and priming stimulation’). The inhibitory population provides feedback inhibition, while not directly being subject to plastic changes. The probability for the existence of a synapse between two neurons is 0.1, across the whole network. This is a

plausible value for hippocampal region CA3 (Le Duigou et al., 2014). For every trial, a new network structure was drawn.

Learning, recall, and priming stimulation

Before we apply the first learning stimulus to our network, we let the initial activity settle for 10.0 seconds. Then, we apply our learning protocol, which delivers three stimulus pulses, of 0.1 seconds duration each, to the neurons dedicated to the first cell assembly (A , the first 600 neurons in the network). Breaks of 0.4 seconds separate the pulses. The stimulation is delivered by an Ornstein-Uhlenbeck process (see Eq. 4.3 introduced above). After the last pulse has ended, we let the network run under basal conditions for 3.0 seconds (Fig. 4.1d). For the case of intermediate consolidation, which we consider in one part of our study, we extend this period to 28800.0 seconds, such that the synaptic weights may undergo consolidation through synaptic tagging and capture. Every time that stimulation ends, we disable early-phase plasticity (except for the decay of early-phase changes) after a time window of 12.5 seconds. This is reasonable considering the impact of novelty and neuromodulation on encoding (Lindskog et al., 2006; Pezze and Bast, 2012).

Then, we apply another learning stimulus of the same kind as the first one to the neurons dedicated to the second cell assembly B . In the non-overlapping case, this set of neurons is completely disjoint from the first set. In the cases with overlaps, 10% of the neurons in the assembly (in the main paper; also cf. Suppl. Table 4.3) are shared with the other assembly. The size of the overlaps is of the same order of magnitude as experimental findings, which range from around 1% to more than 40% (Sakurai, 1999; Cai et al., 2016; De Falco et al., 2016; Yokose et al., 2017). After the learning stimulus, we let the network again run under basal conditions for 3.0 seconds (28800.0 seconds in the intermediate consolidation case).

Finally, we apply a third learning stimulus of the same kind, this time to the neurons dedicated to the third cell assembly C . In the non-overlapping case, this set of neurons is completely disjoint from the first and second set. For the overlap cases, still, 10% of the neurons between each two assemblies are shared. We denote the exclusive intersections between each two assemblies by I_{AB} , I_{AC} , I_{BC} (see Fig. 4.1c), which make up for half of each overlap between two assemblies. The intersection between all three assemblies I_{ABC} makes up for the other half of each overlap between two assemblies (cf. Suppl. Table 4.3). Specifically, the different subpopulations are set to occupy the following neuron indices in the network (example for overlapping case/“OVERLAP10”):

- I_{ABC} : 540–569,
- I_{AC} : 0–29,
- \tilde{A} : 30–539,
- I_{AB} : 570–599,
- \tilde{B} : 600–1109,

- I_{BC} : 1110–1139,
- \tilde{C} : 1140–1649.

By \tilde{A} , \tilde{B} , and \tilde{C} we denote the subsets of neurons that exclusively belong to one of the assemblies. Note that in the non-overlapping case, $\tilde{A} = A$, $\tilde{B} = B$, and $\tilde{C} = C$.

At two points in our study, we used interleaved learning of the three memory representations. For this, instead of the three steps A-B-C, we employed the following scheme of 15 learning steps: A-B-C-A-C-B-B-A-C-A-B-C-C-B-A.

After learning, we let the network first run under basal conditions for 3.0 seconds and then consolidate for 28800.0 seconds (eight hours, see Fig. 4.1e). This caused all early-phase weight changes to either decay or be transferred to the late phase, thereby yielding long-term memory representations. After that, we started measuring the spontaneous activity (see section below).

For the testing of memory function in Fig. 4.2, we applied a recall stimulus to randomly drawn 20% of the neurons of an assembly, delivered by one pulse 0.1 seconds long, with otherwise the same parameters as for the learning stimulus. During this, we left early-plasticity switched off.

For the priming of a memory representation, which we did in the last part of our study, we applied a stimulus to all neurons of one selected assembly (alternatively, to randomly drawn 50% of the assembly neurons, the results of which are only shown in the Supplementary Information). The stimulus was delivered via one pulse 0.1 seconds long, with otherwise the same parameters as for the learning stimulus (Fig. 4.5a). The priming stimulus was followed by a period of consolidation and early-phase decay of varied duration, before we started measuring the spontaneous activity.

Determining avalanches in the spontaneous activity

After the learning, consolidation, and possibly priming of the cell assemblies, we measured the network activity across three minutes to evaluate the spontaneous activity of the assemblies. To this end, we computed the likelihood that an avalanche with a certain minimum size occurred in a specific assembly, averaged across the whole period of three minutes. In addition, we measured the plain mean firing rate of each assembly and the control population across the whole period of three minutes.

To determine if an avalanche occurs, we counted the number of spikes that were produced by the neurons of each cell assembly in a time frame of 10 ms. Then, we compared these numbers to the threshold value n_{thresh} and considered an assembly active if the threshold was reached. Multiple assemblies could be active at the same time. Then, we computed the likelihood of an avalanche in assembly A , for instance, by dividing the number of periods $N_{\text{act}}(A)$ in which the assembly was active by the total number of time frames (given by the total duration of the simulation $t_{\text{max}}/10$ ms):

$$l(A) := \frac{N_{\text{act}}(A)}{t_{\text{max}}/10 \text{ ms}}. \quad (4.10)$$

4.2. Manuscript: Organization and priming of long-term memory representations with two-phase plasticity

The likelihoods for avalanches in B and C are computed analogously. The threshold value n_{thresh} was chosen to be about the border value of the 99% quantile for the number of spikes of an assembly in one time frame (Suppl. Fig. 4.15).

Computational implementation and software used

We used C++ in the ISO 2011 standard to implement our simulations. To compile and link the code, we employed g++ in version 7.5.0 with boost in version 1.65.1.

For further data analysis we used Python 3.7.3 with NumPy 1.20.1 and pandas 1.0.3. For the creation of plots we used gnuplot 5.0.3, as well as Matplotlib 3.1.0.

We generated random numbers using the generator 'minstd_rand0' from the C++ standard library, while the system time served as the seed. We implemented a loop in our code which ensured that for each distribution a unique seed was used. The simulations that had to be performed for this study were computationally extremely demanding. Thus, it was very fortunate that we had the opportunity to run most of our jobs on the computing cluster of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG).

Additional information

Data availability

All data underlying the results presented in this study are contained in this manuscript and the Supplementary Information. The data can be reproduced using the simulation code and the analysis scripts that we have released (Luboeinski, 2021).

Code availability

Our code is freely available and has been released under the Apache-2.0 license (Luboeinski, 2021). As it continues to be developed, the latest version can be retrieved from <https://github.com/jlubo/memory-consolidation-stc>.

Acknowledgments

We thank the members of the Department of Computational Neuroscience, especially Sebastian Schmitt, for many helpful comments on this study. The research was funded by the German Research Foundation (CRC1286, project C1, project #419866478) and by the H2020 - FETPROACT project Plan4Act (#732266).

Author contributions

Conceptualization, methodology, and investigation: CT and JL; software and simulations: JL; data curation, formal analysis, and visualization: JL; funding acquisition and supervision: CT; writing and editing: CT and JL.

4.3 Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity

This section contains the Supplementary Information of the article presented in the previous section (Luboeinski and Tetzlaff, 2021b). The typeset and referencing were adapted to match the format of this thesis. This Supplementary Information is licensed under [CC BY 4.0](#).

Label	Appears in main paper	Description
NOOVERLAP	x	No overlaps
OVERLAP10	x	Each two assemblies exclusively overlap by 5%, and all three assemblies commonly overlap by 5%
OVERLAP10 no ABC		Each two assemblies exclusively overlap by 10%, there is no common overlap between all three assemblies
OVERLAP10 no AC, no ABC	x	Both A, B and B, C overlap by 10%
OVERLAP10 no BC, no ABC	x	Both A, B and A, C overlap by 10%
OVERLAP15		Each two assemblies exclusively overlap by 7.5%, and all three assemblies commonly overlap by 7.5%
OVERLAP15 no ABC		Each two assemblies exclusively overlap by 15%, there is no common overlap between all three assemblies
OVERLAP15 no AC, no ABC		Both A, B and B, C overlap by 15%
OVERLAP15 no BC, no ABC		Both A, B and A, C overlap by 15%
OVERLAP20		Each two assemblies exclusively overlap by 10%, and all three assemblies commonly overlap by 10%
OVERLAP20 no ABC		Each two assemblies exclusively overlap by 20%, there is no common overlap between all three assemblies
OVERLAP20 no AC, no ABC		Both A, B and B, C overlap by 20%
OVERLAP20 no BC, no ABC		Both A, B and A, C overlap by 20%

Table 4.3: All organizational paradigms referred to by this Supplementary Information.

4.3. Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity

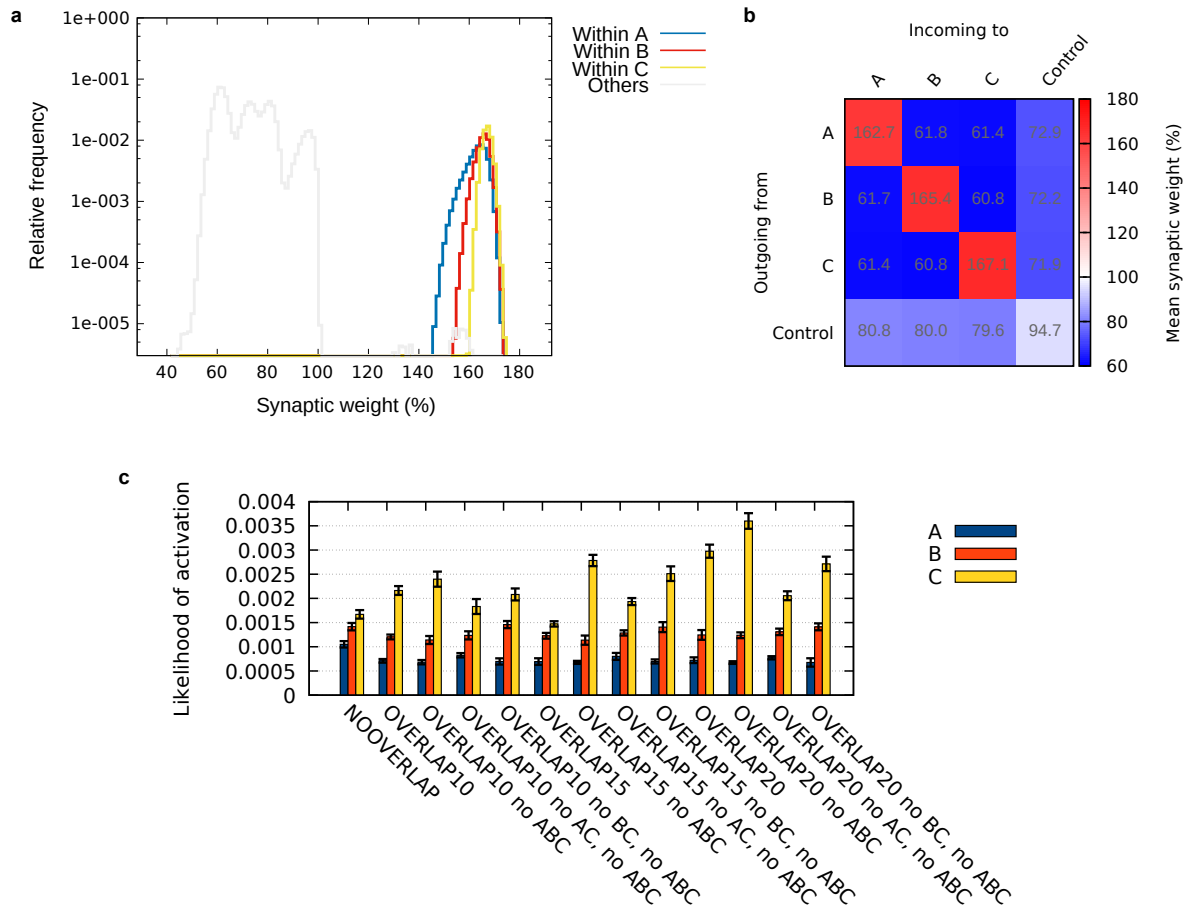


Figure 4.6: **NOOVERLAP** paradigm; after learning but before consolidation (standard protocol, learning order: A-B-C). **(a)** Weight distribution revealing the spread of LTP and LTD. **(b)** Abstract weight matrix showing the mean weight within and between all subpopulations of the excitatory population: the exclusive parts of the cell assemblies, the exclusive intersections, and the control population. The asymmetric nature of the matrix reveals the directionality of the coupling strengths. **(c)** During spontaneous activity, driven by background noise: overview of likelihood of avalanches in the memory representations across different organizational paradigms, including overlaps of varying size. Error bars show the 95% confidence interval. Data in all panels were averaged over 10 trials. Weight values are relative to the initial value before learning.

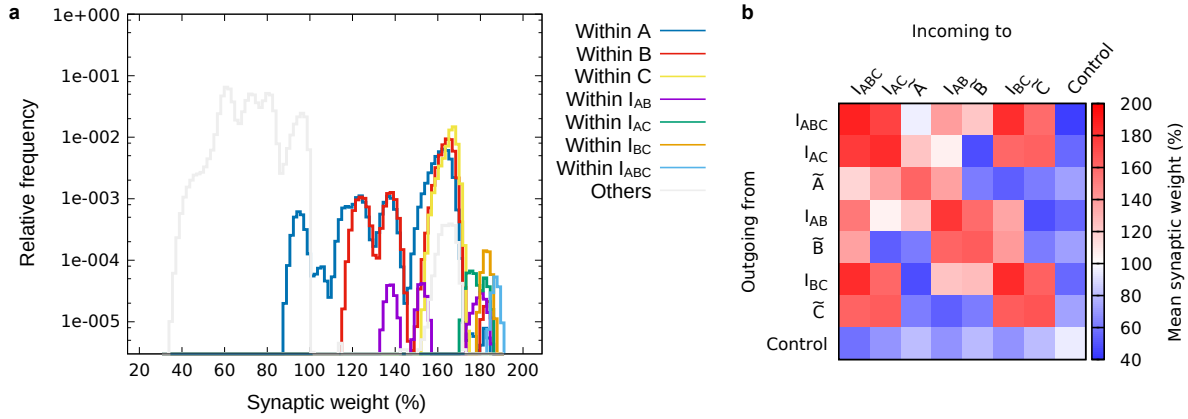


Figure 4.7: **OVERLAP10 paradigm; after learning but before consolidation** (standard protocol, learning order: A-B-C). (a) Weight distribution revealing the spread of LTP and LTD. (b) Abstract weight matrix showing the mean weight within and between all subpopulations of the excitatory population: the exclusive parts of the cell assemblies, the exclusive intersections, and the control population. The asymmetric nature of the matrix reveals the directionality of the coupling strengths. Data in both panels were averaged over 10 trials. Values are relative to the initial value before learning.

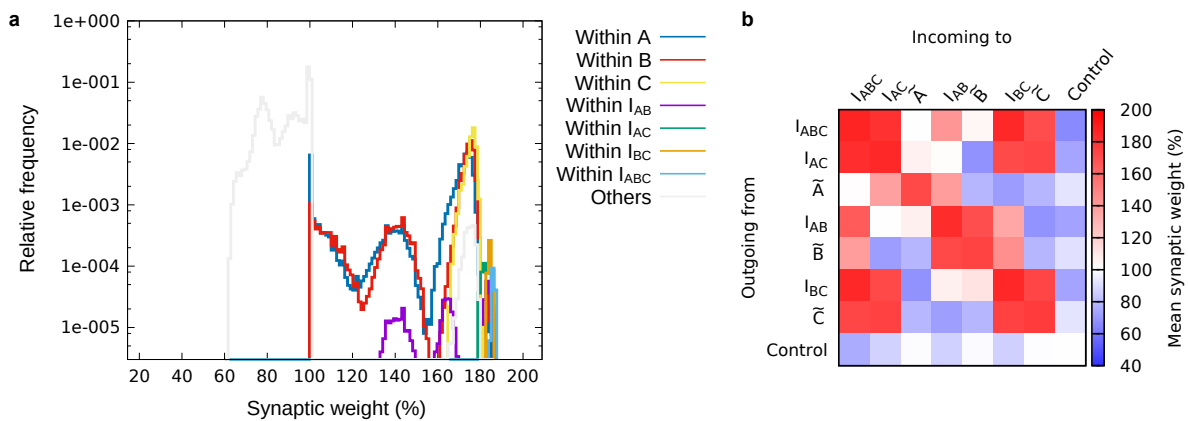


Figure 4.8: **OVERLAP10 paradigm; after learning and consolidation** (standard protocol, learning order: A-B-C). (a) Weight distribution revealing the spread of LTP and LTD. (b) Abstract weight matrix showing the mean weight within and between all subpopulations of the excitatory population: the exclusive parts of the cell assemblies, the exclusive intersections, and the control population. The asymmetric nature of the matrix reveals the directionality of the coupling strengths. Data in both panels were averaged over 10 trials. Values are relative to the initial value before learning.

4.3. Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity

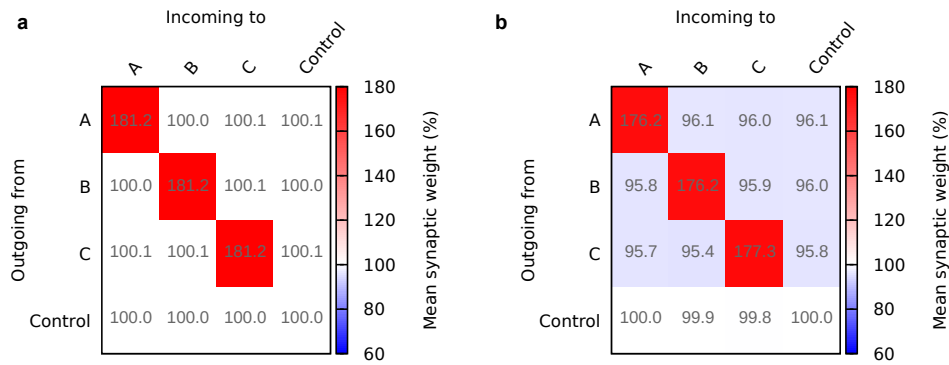


Figure 4.9: Abstract weight matrices showing the mean weight within and between all subpopulations of the excitatory population after learning and consolidation (order: A-B-C). (a) NOOVERLAP paradigm with blocked LTD, (b) NOOVERLAP paradigm with intermediate consolidation of 8 hours in between learning A, B, and C. Data in both panels were averaged over 10 trials. Values are relative to the initial value before learning.

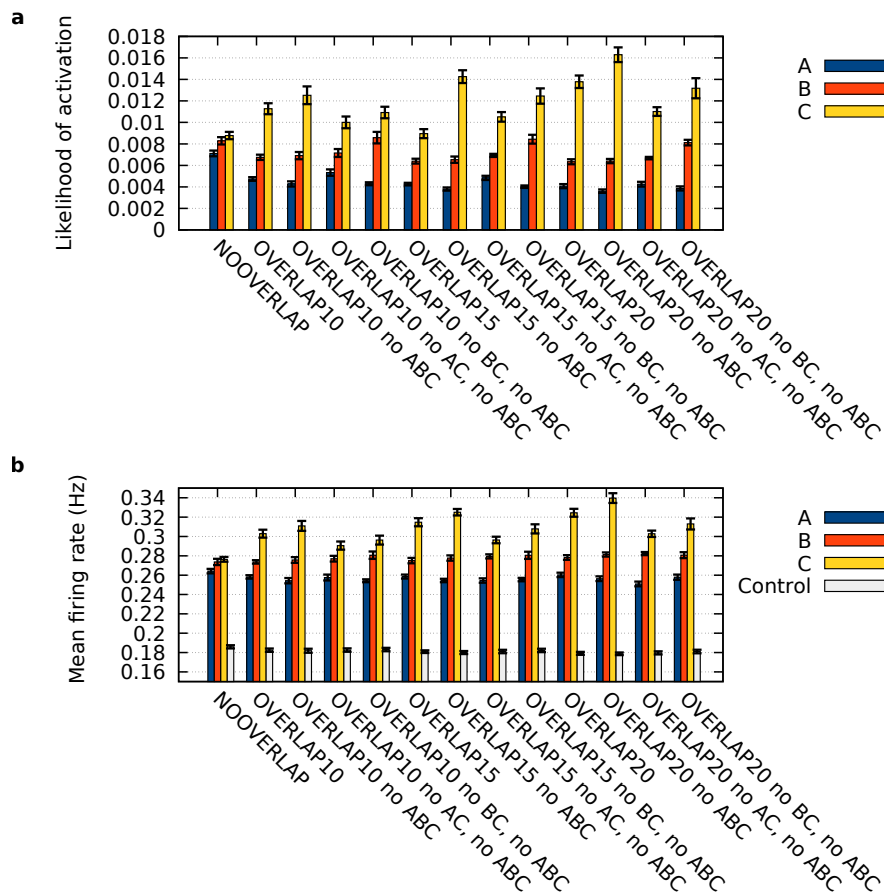


Figure 4.10: Spontaneous activity of long-term memory representations, driven by background noise. Standard protocol for learning/consolidation (order: A-B-C). (a) Overview of likelihood of avalanches across different organizational paradigms, including overlaps of varying size. (b) Overview of mean firing rates in the same organizational paradigms as in a. Averaged over 10 trials. Error bars show the 95% confidence interval.

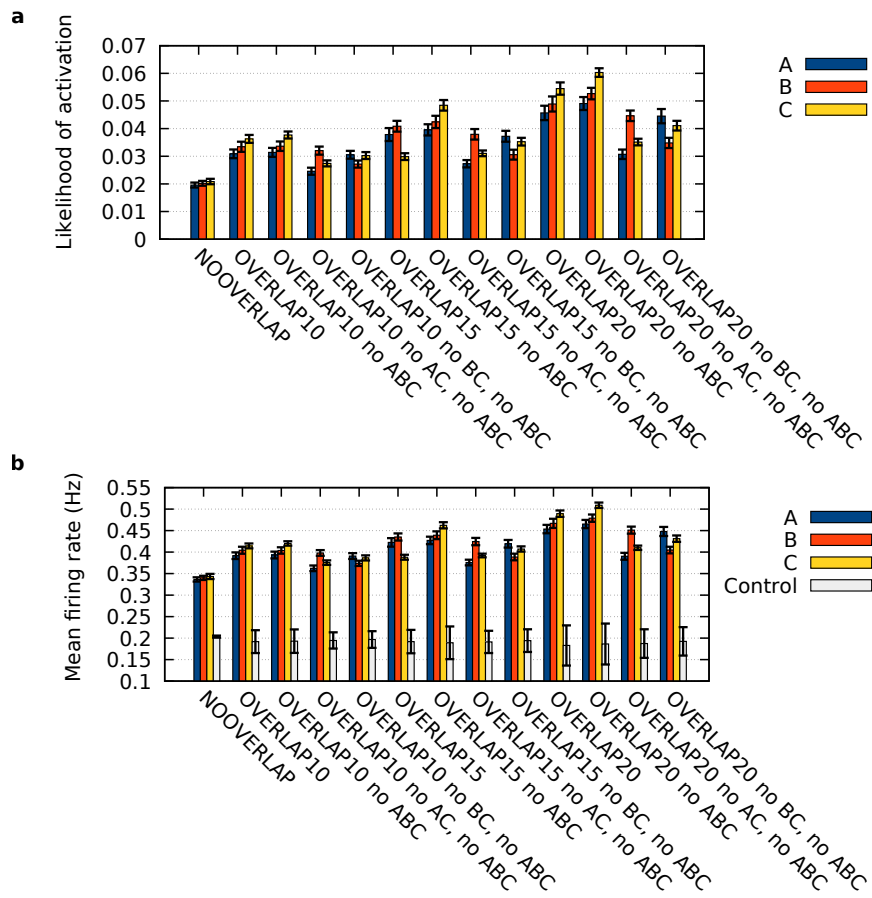


Figure 4.11: Spontaneous activity of long-term memory representations, driven by background noise. Intermediate consolidation protocol (order: A-B-C). (a) Overview of likelihood of avalanches across different organizational paradigms, including overlaps of varying size. (b) Overview of mean firing rates in the same organizational paradigms as in a. Averaged over 10 trials. Error bars show the 95% confidence interval.

4.3. Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity

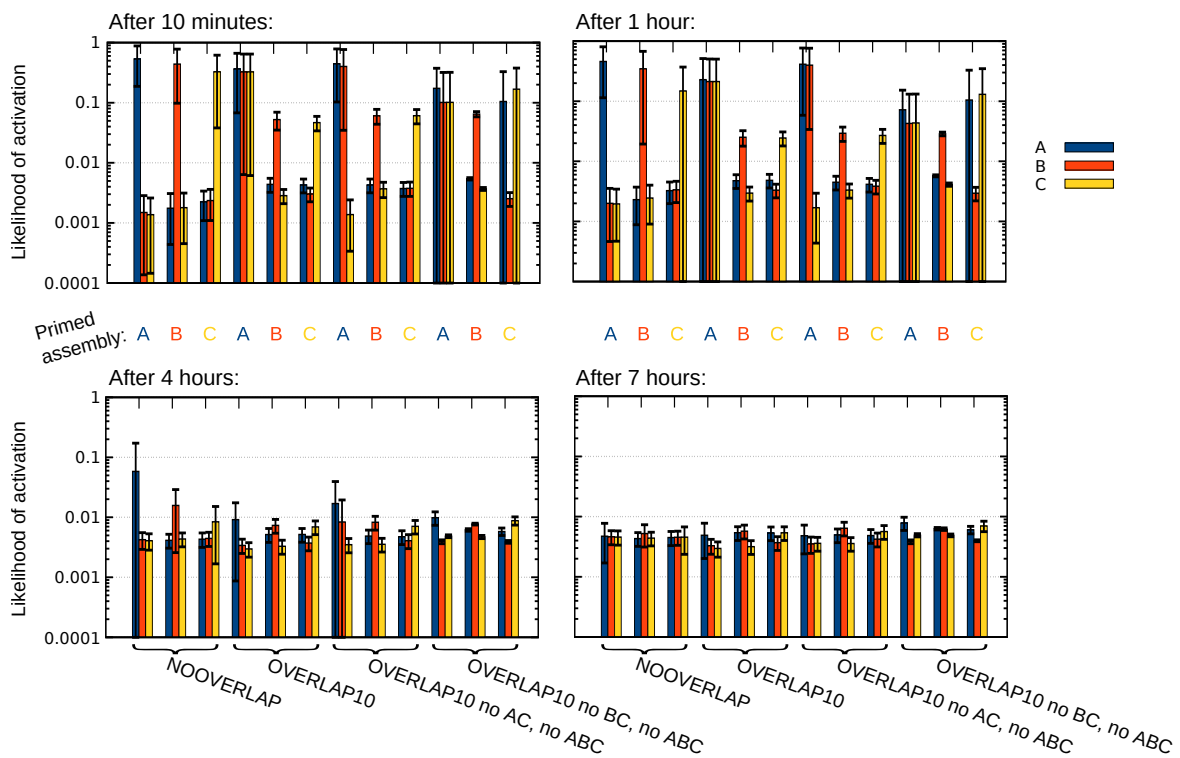


Figure 4.12: Spontaneous activity of long-term memory representations, driven by background noise, after priming stimulus to all neurons of the specified cell assembly. Shown is the likelihood of avalanches across different organizational paradigms at different times after priming. Note the logarithmic scale. Learned with interleaved protocol (order: $A-B-C-A-C-B-B-A-C-A-B-C-C-B-A$). Averaged over 10 trials. Error bars show the 95% confidence interval.

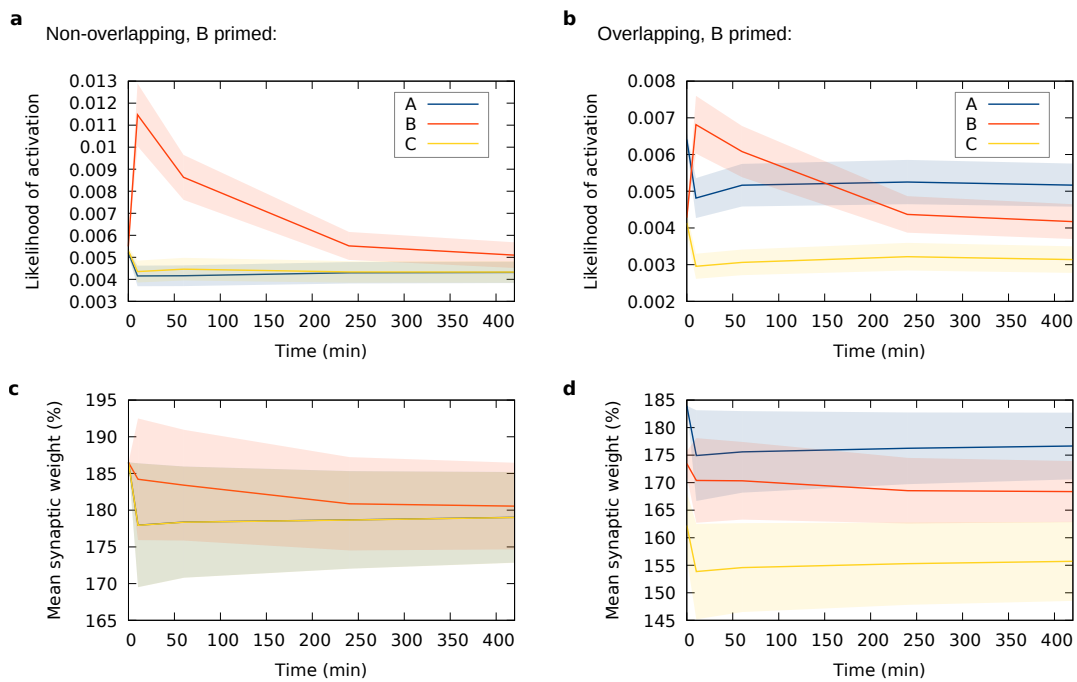


Figure 4.13: **Priming on long timescales, enabled by varying degrees of long-term depression, caused by priming stimulation applied to 50% of the neurons of assembly B.** (a,b) Resulting likelihood of avalanches in the three assemblies at different times after priming B. The values at time zero show the case before/without priming. NOOVERLAP paradigm in a, OVERLAP10 paradigm in b. (c,d) Temporal development of the mean synaptic weight in the assemblies, for the same cases as in a and b. Averaged over 10 trials. Error bands show standard error of the mean. Weight values are relative to the initial value before learning. Learned with interleaved protocol (order: A-B-C-A-C-B-B-A-C-A-B-C-C-B-A).

4.3. Supplementary Information: Organization and priming of long-term memory representations with two-phase plasticity

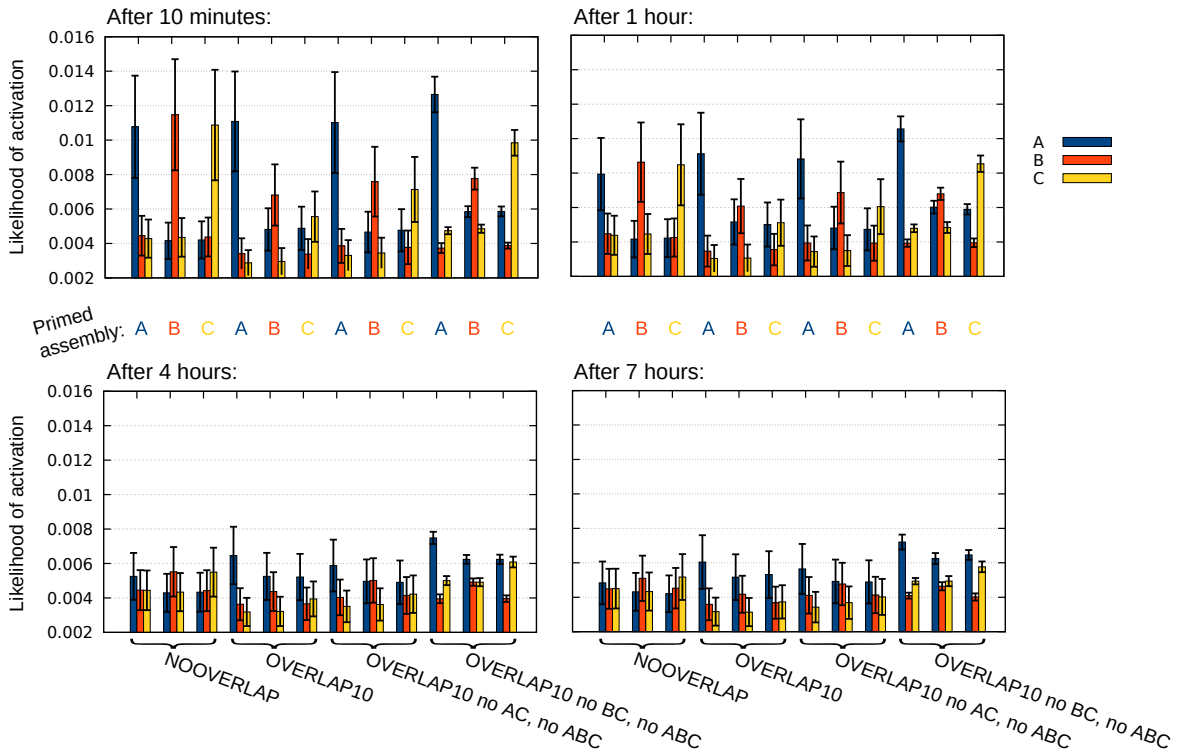


Figure 4.14: Spontaneous activity of long-term memory representations, driven by background noise, after priming stimulus to 50% of the neurons of the specified cell assembly. Shown is the likelihood of avalanches across different organizational paradigms at different times after priming. Learned with interleaved protocol (order: A-B-C-A-C-B-B-A-C-A-B-C-C-B-A). Averaged over 10 trials. Error bars show the 95% confidence interval.

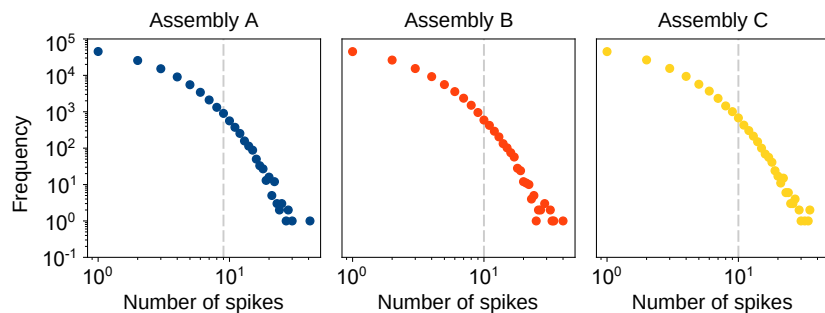


Figure 4.15: Distribution of avalanche size (number of spikes per 10 ms bin), within the three assemblies, after learning and consolidation with standard protocol in the NOOVERLAP paradigm. The dashed lines indicate the border of the 99% quantile for each individual distribution. Data from 10 simulations, each lasting 3 minutes, of the same network.

4.4 Activation and transitions of attractor memory representations

This section presents work that exhibits some similarity with, and can be considered complementary to the previously presented article [Luboeinski and Tetzlaff, 2021b](#) (sections 4.2 and 4.3). After considering cell assemblies that are transiently active, we now turn to another specific kind of cell assemblies – attractor cell assemblies, or attractors.

Introduction

Phenomenological and attractor models of associated memories have successfully described many aspects of free recall experiments ([Raaijmakers and Shiffrin, 1981](#); [Bradski et al., 1994](#); [Howard and Kahana, 2002](#); [Romani et al., 2013](#); [Recanatesi et al., 2015](#)). Several theoretical studies have, moreover, indicated a crucial role of overlaps between attractor memories to explain transitions between recalled items ([Kropff and Treves, 2006](#); [Romani et al., 2013](#); [Recanatesi et al., 2015](#)). However, an explanation of the formation, consolidation, and biologically realistic spike-based activation of such memory representations in free recall paradigms remains elusive.

Attractor behavior commonly relates to fixed points in the phase space of a dynamical system. There are, nevertheless, also cyclic, bump, line, or chaotic attractors. A vast amount of literature has studied attractor neural networks (ANNs), mostly with rate-coding neurons, and thereby featuring fixed point attractors ([Amit, 1989](#); [Fusi and Wang, 2016](#)). However, networks of spiking neurons can also express persistent, self-sustained activation of cell assemblies, which lasts until it is silenced by external stimulation, and thereby essentially constitutes a cyclic attractor ([Miller, 2016](#)). Attractor states that last for a limited time span are also referred to as quasi-attractor states ([Amit, 1989](#)).

ANNs have been investigated since the 1970s and '80s with seminal studies by [Little \(1974\)](#) and [Hopfield \(1982\)](#). Importantly for neuroscience, such networks constitute realizations of auto-associative (content-addressable) and hetero-associative memory, similar to that found in hippocampal areas CA3 and CA1 ([Hasselmo et al., 1995](#)). There are further indications that the hippocampal area CA3, for instance, could constitute an attractor network ([McNaughton and Morris, 1987](#); [Treves and Rolls, 1992](#)). By now, much of the pertinent literature on ANNs has focused on storage capacity, working memory, or engineering applications ([Tsodyks and Feigel'man, 1988](#); [Gardner, 1988](#); [Amit, 1989](#); [Amit and Tsodyks, 1991a,b](#); [Amit and Brunel, 1997b](#); [Adachi and Aihara, 1997](#); [Treves, 2005](#); [Kurikawa and Kaneko, 2012](#); [Miller, 2016](#)).

Upon stimulation, attractor cell assemblies exhibit pattern completion in an all-or-none manner ([Braitenberg, 1978](#); [Fusi and Wang, 2016](#)). The resulting pattern of activity in the network is characterized by elevated activity of the assembly neurons and more quiet activity in the rest of the network. The attractor activity is typically self-stabilizing, such that the pattern persists even in the absence of small distortions ([Amit, 1989](#); [Amit and Tsodyks, 1991a](#)).

Due to this capability to sustain information, attractors have been under debate as a neural correlate for working memory ([Amit and Brunel, 1997b](#); [Fusi and Wang, 2016](#); see also section 1.4.6). Nevertheless, concerns about this hypothesis have been raised, particularly with respect to rate-coding neurons ([Fusi and Wang, 2016](#); [Nachstedt and Tetzlaff, 2017](#)). Yet, there have also been some attempts to model ANNs with biologically more realistic spiking neurons. [Amit and Tsodyks \(1991a\)](#) provided in an early study a theoretical description for the conversion of currents to firing rates, grounded on

4.4. Activation and transitions of attractor memory representations

integrate-and-fire dynamics. In a second part of the study, the authors used this framework to investigate memory retrieval at low firing rates in a network with symmetric synaptic connections (Amit and Tsodyks, 1991b). Gerstner et al. (1993) went beyond this and showed that spiking enables learning and retrieval of irreducible spatio-temporal memory representations. Nevertheless, while ANNs with spiking dynamics remained relatively unattended for a long time (cf. Eliasmith, 2005), persistent firing characteristics have been questioned (Shafi et al., 2007; Barbieri and Brunel, 2008). Thus, it seems reasonable to employ biologically detailed spiking and plasticity dynamics to enable the learning and retrieval of quasi-attractor rather than attractor dynamics. This can then be the basis for a model that holds for the longer timescales of consolidation.

A recent study modeled attractor dynamics in a spiking neural network with plasticity and an abstract consolidation mechanism for synaptic weights, and found good agreement with several experiments (Zenke et al., 2015). However, this model describes bistable plasticity dynamics and thereby cannot give rise to the diversity in synaptic strength that is required to describe differential connectivity between cell assemblies. Overlaps with differential synaptic strength are necessary to distinguish cell assemblies from each other, as opposed to becoming one large assembly if the weights are binary. These and more functional implications of differential synaptic strength have been shown in various paradigms (Romani et al., 2013; Katkov et al., 2015; Recanatesi et al., 2015; Herpich and Tetzlaff, 2019), some of which will be discussed in the following.

ANNs in conjunction with oscillatory input have recently been considered for the modeling of memory recall (Romani et al., 2013; Recanatesi et al., 2015). Using such ANNs, Recanatesi et al. (2015) were able to reproduce data from free recall experiments. They found that intersections between memory representations play a crucial role in the transitions between attractor states by acting as cues, which has also been suggested by other theoretical (Treves, 2005; Kropff and Treves, 2006; Recanatesi et al., 2017) and experimental studies (Fischer et al., 2020). Moreover, oscillations in the theta-frequency range have often been related to memory functions (Tesche and Karhu, 2000; Osipova et al., 2006; Kahana, 2006) and specifically to switching between discrete network states, besides slow-wave oscillations (Pedreschi et al., 2020). Recanatesi and colleagues argued that oscillations serve to control the switching between attractors, which they achieved by modulating the oscillation amplitude (Recanatesi et al., 2015, 2017). Nevertheless, other models have described transitions between memory representations through non-oscillatory background noise in conjunction with neuronal adaptation (Kropff and Treves, 2006; Russo and Treves, 2012; Zenke et al., 2015; Roach et al., 2016) or short-term depression (Mongillo et al., 2008; Zenke et al., 2015). Similar ideas were already formulated by Braitenberg (1978). Experimental evidence also exists for such spontaneous switching (Jezek et al., 2011; Monasson and Rosay, 2015).

Here, we employ a biologically realistic model with spiking neurons and synaptic plasticity described by calcium dynamics and synaptic tagging and capture (STC), which enables us to provide a mechanism for the formation and consolidation of overlapping long-term attractor memory representations. Using these memory representations and drafting a connection to an existing rate-based model of free recall, we explore the possibility that oscillatory inhibition in the slow-wave and theta regime give rise to attractor activation. After that, we examine the spontaneous activation of attractors and the transitions between them as induced by background noise, to obtain insights into the processes that guide switching between recalled attractor memories. By these investigations, we further pursue

our third research aim in this thesis, which is to find a mechanistic basis for the STC-based description of behavioral effects that involve multiple memory representations. Considering similar paradigms as in section 4.2, we complement the findings of that section by characterizing the impact of attractor dynamics. In addition, considering attractor dynamics enables us now to examine transitions between activated memories.

Results

This section presents adaptations of our model previously introduced in section 4.2 (Luboeinski and Tetzlaff, 2021b), with the objective to model overlapping memory representations that enter (quasi-) attractor states, as well as transitions between them. To obtain the results described here, we used the model with different parameter values (see Methods), and with a protocol featuring breaks of 8 h between learning stimuli (as in the intermediate consolidation protocol in section 4.2).

Learning of attractor assemblies and activation by oscillatory input

To demonstrate that our model is able to form and consolidate overlapping attractor structures and to investigate the impact of overlaps on the transitions between attractors, as in section 4.2, we first let our network subsequently learn three partially overlapping (or non-overlapping) assemblies. Each learning process was followed by a phase of eight hours to consolidate the newly formed cell assembly (as in the intermediate consolidation protocol in section 4.2). During consolidation, early-phase potentiation within the respective cell assembly was transferred to the late phase. The weight distributions after learning and consolidating the first, second, and third assembly without overlaps are shown in Fig. 4.16a,b,d. For overlapping assemblies, every consolidation phase caused a large, relatively discrete increase in the mean synaptic weight of the intersections (Fig. 4.16c,e,f), enabled by the extended range of values for the late-phase weight (as compared to the model in section 4.2). We assumed this to facilitate transitions and to let intersections act as attractors themselves (cf. Recanatesi et al., 2015, 2017). At the same time, parts of the previously learned assemblies A and B underwent slight depotentiation, which is probably caused by LTD, as demonstrated in our previous investigations (cf. section 4.2). On the other hand, upon learning the third assembly, some synapses that were not fully located within an assembly became potentiated, which is presumably due to the high activity of the whole network (note that the postsynaptic neuron has to fire at a certain frequency to undergo LTP, as described by Eq. 4.7). No substantial depotentiation occurred in the non-overlapping case, which is consistent with the findings from our previously model (see Fig. 4.4b).

After we had the cell assemblies learned and consolidated, we wanted to explore the capability of our biologically detailed model to reproduce the findings Recanatesi et al. (2015) had made with rate-coding neurons. To this end, we applied in addition to the background noise the oscillatory input

$$I_{\text{stim}}(t) = \hat{a} \cdot \sin(2\pi f_{\text{osc}} \cdot t) + b \quad (4.11)$$

to all inhibitory neurons in our neural network (cf. Recanatesi et al., 2015). We considered frequencies of 1 Hz and 5 Hz, for which the oscillation period is much larger than the membrane time constant, such that the network adapts to the oscillations in an adiabatic manner. We found that such input triggers the activation of attractor cell assemblies (Fig. 4.17a,b,d,e). The fact that we considered a slow

4.4. Activation and transitions of attractor memory representations

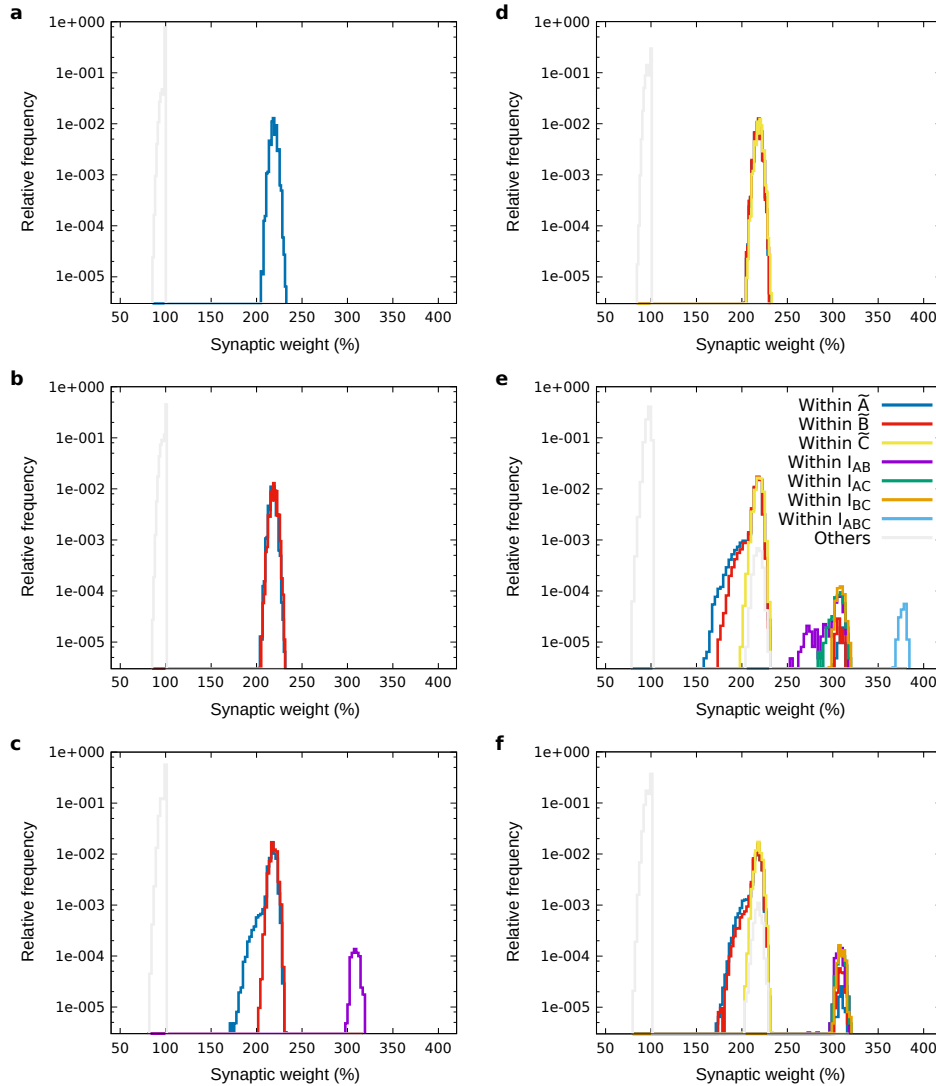


Figure 4.16: **Distribution of the synaptic weights in a network;** with the cell assemblies A , B , and C , the intersections of which exhibit discrete increases in synaptic weight. **(a)** After learning and consolidation of A ; **(b)** after learning and consolidation of A - B , no overlap; **(c)** after learning and consolidation of A - B , overlap 10%; **(d)** after learning and consolidation of A - B - C , no overlap; **(e)** after learning and consolidation of A - B - C , overlap 10%; **(f)** after consolidation of A - B - C , overlap 10% but without common intersection of all three assemblies.

frequency of 1 Hz, and a theta frequency of 5 Hz may support the notion that slow-wave and theta oscillations play an important role in memory dynamics (Tesche and Karhu, 2000; Osipova et al., 2006; Kahana, 2006; Pedreschi et al., 2020). However, while the oscillatory activity of the inhibitory population enabled the activation of attractors at defined times, not in every oscillation period an attractor was active. More refinement of the parameters seems necessary to enable the detailed study of transitions between attractors in this paradigm as it was done by Recanatesi et al. (2015). Also note that applying oscillatory input to the excitatory neurons, instead of the inhibitory neurons, may constitute an alternative which will be much easier to tune.

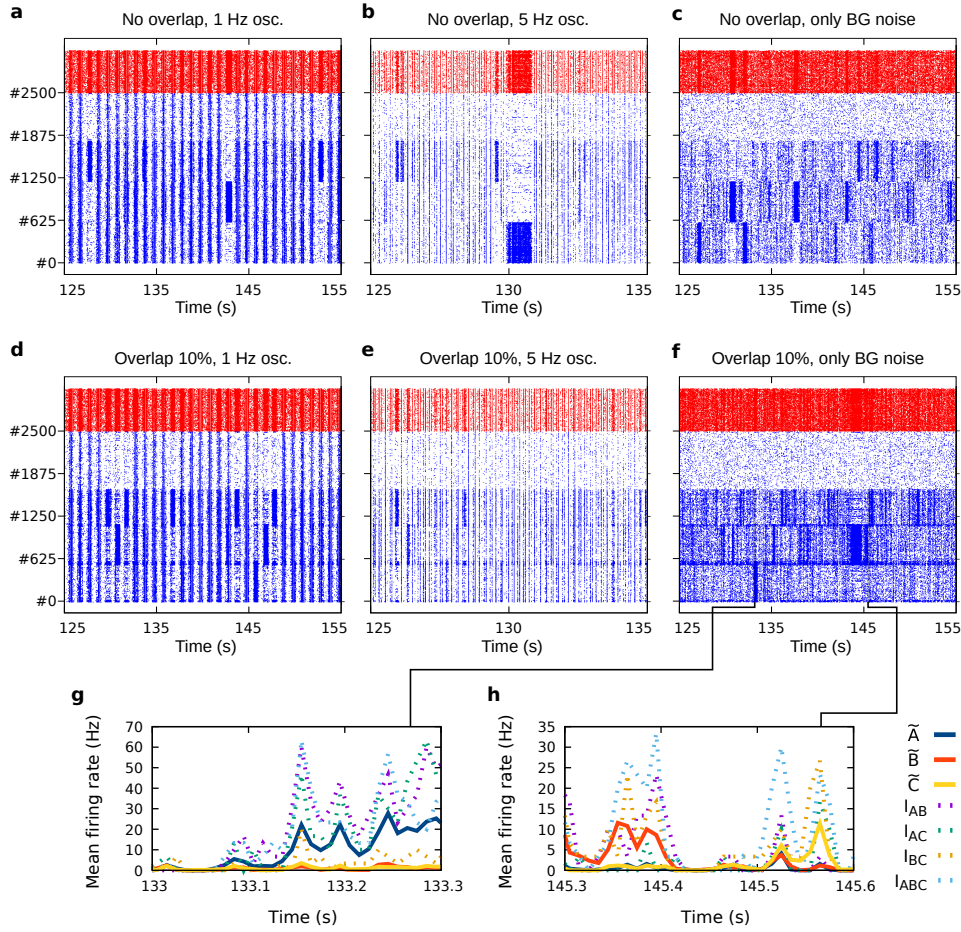


Figure 4.17: Neuronal activity in networks with three attractor cell assemblies, activated by different types of input. (a-c) No overlap between assemblies. Each assembly comprises 600 excitatory neurons, shown in blue. Inhibitory neurons are shown in red. For the exact placement of overlaps, see Methods in section 4.2. Network is the same as in Fig. 4.16d. In (a), oscillatory input of $f_{osc} = 1$ Hz, and in (b), oscillatory input of $f_{osc} = 5$ Hz are applied to the inhibitory neurons. Amplitude of oscillatory input: $\hat{a} = 0.3$ nA, baseline: $b = 0$. (c) is without oscillatory input. (d-f) as in (a-c) but with overlaps of 10% between assemblies. Network is the same as in Fig. 4.16e. (g,h) Detailed view of the firing rates in the different subpopulations (exclusive fractions of the assemblies, and intersections) of the network in (f). Firing rates were computed via bins of 10 ms.

Activation of attractor assemblies by background noise

Even in the absence of oscillations, the interplay of background noise, spiking dynamics, and inhibitory feedback in our model naturally enabled spontaneous activation and switching of attractors. This is shown in Fig. 4.17c,f. Two time intervals are magnified in Fig. 4.17g,h, which also highlights the differences between the activities of the exclusive parts of the assemblies and the intersections (cf. Methods in section 4.2 of this thesis for the exact definitions). Next, we systematically investigated the probability of activation for each of the attractors in different organizational paradigms (see Methods in this section for the computation of the probability, and sections 4.2 and 4.3 for details on the organizational paradigms). The probability of activation is presented in Fig. 4.18a. We observed an evident impact of the learning order, similar to the effect described previously in section 4.2 (Luboeinski and

Tetzlaff, 2021b). As it is to be expected, the probability for the attractor memory representations to be active was much higher than for non-attractor memory representations (compare Fig. 4.18a and Fig. 4.3c). This property, however, may render the attractor system also more prone to instabilities. In one trial, we found that during the learning procedure the cell assemblies were disrupted. Including this trial would have distorted the results substantially, such that we had to exclude the trial and repeated it with another network.

Considering the details of the dynamics at short timescales (Fig. 4.17g,h), we found as another result that the intersections between cell assemblies act as attractors themselves, meaning that their activation can also switch spontaneously. Such behavior was found previously in rate-based attractor neural networks (Recanatesi et al., 2015, 2017). The demonstration of this behavior with our model provides further evidence for an improved mechanistic description of free recall data based on spiking neural networks.

The spontaneously switching activation of attractors driven by background noise has been found in experiments (Jezek et al., 2011; Monasson and Rosay, 2015) and reminds of the latching dynamics that have been shown to occur in attractor neural networks (Treves, 2005; Kropff and Treves, 2006; Russo and Treves, 2012). In the next subsection, we will scrutinize this further and examine the transitions between the spontaneously active attractors in our network model.

Transitions between active attractors

As mentioned above, we increased the range of late-phase weight values to enable discrete, large differences between the mean weights of the exclusive parts of the assemblies and the intersections between assemblies. Following the approach of earlier studies that employed rate neurons (Recanatesi et al., 2015, 2017), we did this in order to investigate the capability of intersections to constitute attractors themselves. While this was shown in the previous subsection, we now investigate if the intersections may facilitate transitions between attractors.

The probabilities of transitions between different activation states are presented in Fig. 4.18b-f (see the Methods section in this chapter for the computation of the probabilities). We find that the transition matrices are asymmetric, which is probably chiefly caused by the learning order effect. In previous studies, activity-related effects have also been shown to cause asymmetry in transitions (Kropff and Treves, 2006). The matrix plots indicate that the overlaps between the assemblies in our network indeed give rise to such asymmetry, as can be seen, for example, comparing the “From A to C” transition in the “No overlap” and the “Overlap 10%” paradigms. However, the overlaps do not only increase the richness of the dynamics, but also enhance the learning order effect, thereby shifting the dynamics toward monopoly activation of C (see Fig. 4.18a, and rows “From C” and columns “To C” in Fig. 4.18b-f).

Discussion

In this additional study, we presented an adapted version of the model from section 4.2 (Luboeinski and Tetzlaff, 2021b) with the goal to describe long-term (quasi-)attractor memory representations. Not only does our model account for the processes of learning and consolidating these memory representations, it also enables the spontaneous activation of these attractors as well as directed transitions

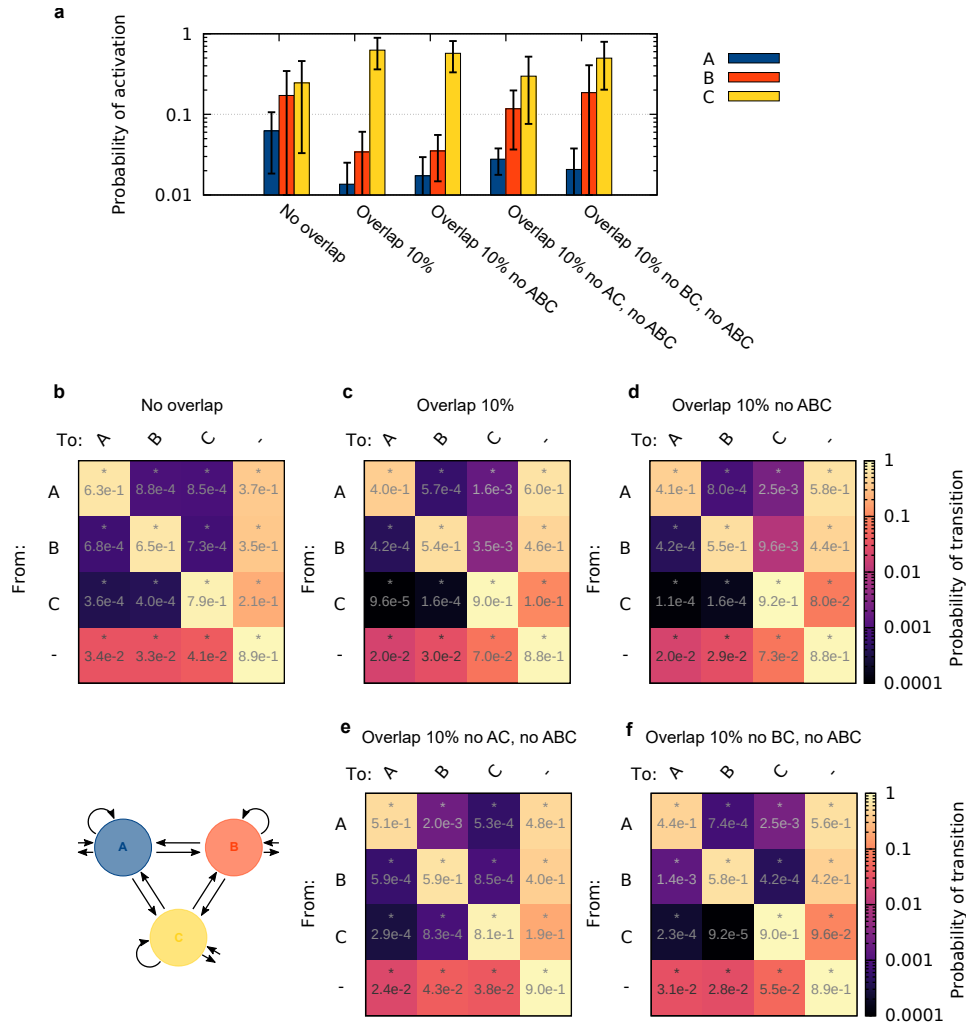


Figure 4.18: **Spontaneous activation of attractor cell assemblies, driven by background noise.** (a) Overview of the probability of activation in different organizational paradigms, learned with intermediate consolidation (order: A-B-C). Mean probability across three minutes of biological time, averaged over 10 networks. Error bars show the 95% confidence interval. (b-f) Matrices showing the probabilities of transitions between attractors (A, B, C) and void (-) for different organizational paradigms (normalized across each row). Mean probabilities were obtained from three minutes of simulated biological time, averaged over 10 networks. Asterisks (*) indicate significant occurrence of transitions ($p < 0.05$, one-sided t-test).

between them. In an effort to establish a link to previous studies, we could also show that oscillatory input of certain frequency enables the activation of attractors, which may serve as a basis for a future study.

Recanatesi et al. (2015) grounded their work on analytical investigations, by which they predicted the existence of stable states apart from single attractors (also see Recanatesi et al., 2017). While it would have been equally desirable to have analytical predictions for the behavior of our computational model, this was prevented by multiple nonlinearities in our biologically more realistic equations. Furthermore, Recanatesi et al. (2015) suggested that a more realistic model of free recall by oscillation should comprise features of vasoactive intestinal peptide (VIP) interneurons, which account for dis-

4.4. Activation and transitions of attractor memory representations

inhibition (Pi et al., 2013). Since our model already features couplings from inhibitory to inhibitory neurons, it can easily be adapted for investigating the impact of such interneurons. To conclude, in a future study, our model driven by oscillatory input may serve to reproduce free recall data in the manner of previous studies (Romani et al., 2013; Recanatesi et al., 2015). The gain would be a biologically realistic spiking neural network model that accounts for the learning and consolidation of attractor memories as well as for free recall experiments.

Our finding that larger overlaps or higher connectivity between assemblies exert an influence on the transitions is consistent with earlier findings that transitions become more deterministic and less random as connectivity is increased (Treves, 2005; Kropff and Treves, 2006). Nevertheless, the exact underpinnings for the impact of overlaps on the transitions remain to be determined in a future study. An important factor may be that the activity in the intersections acts as a “starter” for the activation of an assembly (cf. Fig. 4.17g,h), similar to the mechanism in Recanatesi et al., 2015. Thereby, the activation of one assembly could act as a cue for the activation of the next assembly, as it was shown in earlier theoretical studies and experiments (Kropff and Treves, 2006; Romani et al., 2013; Recanatesi et al., 2015; Fischer et al., 2020). Recently, a theoretical study employing experimental data targeted the associative properties of overlaps between cell assemblies in the medial temporal lobe (Gastaldi et al., 2021). The findings of this study support our finding that overlaps have to be in a certain functional range to evoke transitions between attractors. Furthermore, Treves (2005) proposed that deterministic transitions between attractors occur if the number of assemblies or the connectivity is high but that otherwise transitions are relatively random. Meanwhile, Katkov et al. (2015) found a correlation between the size of an assembly and the probability of its activation. Investigating such dependencies on the number or the size of the assemblies in our model could be the subject of a future study with a larger network and more computing power. Moreover, the model by Zenke et al. (2015), which seems to be the most realistic attractor model so far, does not account for depotentiation (LTD) following low-frequency stimulation. This, however, plays an important role in sequential learning of memories, as we have demonstrated here and in sections 4.2 and 4.3. Therefore, our model provides an important contribution to the understanding of the interplay between attractor cell assemblies.

Employing attractor memory representations, it seems difficult to model priming in the manner in which we demonstrated it in section 4.2. The reason for this is that the strong activation of an attractor state could cause the early-phase weight changes elicited by a priming stimulus to be transferred to the late phase (cf. Figs. 4.5 and 4.12). Therefore, if the paradigm of attractor memories shall not be abandoned, careful tuning would be necessary, or additional mechanisms such as neuronal adaptation (Russo et al., 2008; Zenke et al., 2015), short-term plasticity (Mongillo et al., 2008), or inhibitory plasticity (Vogels et al., 2011; Zenke et al., 2015) would have to be introduced.

As it was demonstrated by others (Russo et al., 2008; Romani et al., 2013), neuronal adaptation could also serve to prevent falling back into the same attractor state upon switching between memory representations, for example during free recall. A promising neuron model for this could be the MAT2 model (Kobayashi et al., 2009; Yamauchi et al., 2011). However, as mentioned in section 2.4, a model with neuronal adaptation might also complicate the learning and recall of attractor memories, at least without additional inhibitory plasticity. For the formation and recall of attractor memories, inhibitory plasticity has been targeted by several studies, revealing that such plasticity stabilizes the network activity whilst enabling the robust formation of memory representations. While it was shown that

inhibitory plasticity improves the self-organization of network activity (Vogels et al., 2011; Zenke et al., 2015) and enables functional discrimination between memories (Herpich and Tetzlaff, 2019), it does not seem essential for memory recall (Zenke et al., 2015). Since these mechanisms were not needed for the successful formation and organization of memory representations in our model, we decided against including inhibitory plasticity and neuronal adaptation and left this to further studies.

Methods

The model was the same as in Luboinski and Tetzlaff, 2021b (see section 4.2), with the different parameter values $\nu_{\text{th}} = 80$ Hz and $z_{\text{max}} = 5$. The late-phase weight dynamics were described by the following augmented equation (cf. Eq. 4.8):

$$\begin{aligned} \tau_z \frac{dz_{ji}(t)}{dt} = & p_i(t) \cdot \frac{z_{\text{max}} - z_{ji}(t)}{z_{\text{max}}} \cdot \Theta[(h_{ji}(t) - h_0) - \theta_{\text{tag}}] \\ & - p_i(t) \cdot (z_{ji} + 0.5) \cdot \Theta[(h_0 - h_{ji}(t)) - \theta_{\text{tag}}]. \end{aligned} \quad (4.12)$$

To determine if an attractor was active within a period of 10 ms, we counted the number of spikes fired by the neurons of each assembly during that period. Then, we attributed the period to the assembly that had fired the most spikes, given that the spike count of the assembly exceeded 10, and that it was by at least 1.5 times greater than the spike count of the other two assemblies. Here, it is valid to consider the activation of an assembly to be exclusive because unlike the assembly activation considered in section 4.2, the activation of an attractor typically causes the suppression of activity in the rest of the network by evoking strong inhibition (see Fig. 4.17). Based on this, we obtained the probability of activation of the attractor A by dividing the number of periods $N_{\text{act}}(A)$ that were attributed to this attractor by the total number of simulated periods, given by the total duration of the simulation t_{max} and the duration of one period:

$$p(A) := \frac{N_{\text{act}}(A)}{t_{\text{max}}/10 \text{ ms}}. \quad (4.13)$$

The probability of activation for the attractors B and C was obtained analogously. Note that because activation is exclusive here, we can indeed compute probability instead of likelihood.

To compute the probabilities of transitions, we considered the activation of attractors in each two subsequent periods of 10 ms. As described above, in each period, the network can be in one of four states: A is active, B is active, C is active, or none of the attractors is active (-). The probability of a transition from state $X \in \{A, B, C, -\}$ to state $Y \in \{A, B, C, -\}$ is then given by:

$$p_X(Y) := \frac{N_{\text{trans}}(X \rightarrow Y)}{N_{\text{act}}(X)}. \quad (4.14)$$

with $N_{\text{trans}}(X \rightarrow Y)$ the number of transitions from X to Y and $N_{\text{act}}(X)$ the total number of periods in state X (cf. Eq. 4.13). Note that thereby, the probability is normalized as $\sum_Y N_{\text{trans}}(X \rightarrow Y) = N_{\text{act}}(X)$.

Memories in the brain are thought to be retained by the strength of synaptic connections (Martin et al., 2000; Abraham et al., 2019). Recurrently connected groups of neurons with strong synaptic connections, so-called Hebbian cell assemblies, are particularly important for memory storage at the network level (Palm et al., 2014; Holtmaat and Caroni, 2016). In this thesis, we have developed a biologically realistic model for the formation and synaptic consolidation of such memory representations in recurrent spiking neural networks, allowing the recall of information by pattern completion or by reactivation of temporal structures. We analyzed the dependence of recall function on various conditions such as inhibitory input, cell assembly size (chapter 2), the concentration of an abstract neuromodulator (chapter 3), and pre-existing cell assemblies in the network (chapter 4). By our results, we could demonstrate that the presented model can account for diverse cognitive and behavioral phenomena, while we provided predictions that can be tested in further experiments ranging from the neuronal to the neuropsychological level.

Throughout this thesis, presented results have been accompanied by specific discussion in the respective sections. Furthermore, many assumptions made for our model have already been treated in chapter 1 (sections 1.3 and 1.4), and we have thereby captured several possible caveats. Thus, in this chapter, we will take a more general perspective on the model that has been developed in this thesis (section 5.1), and on the results that have been – and can be – obtained with it (sections 5.2 and 5.3).

5.1 Model of synaptic memory consolidation

Part of the aims of this thesis was to find a biologically realistic model of synaptic (initial) memory consolidation, based on spiking neurons and the two-phase-plasticity mechanisms of synaptic tagging and capture (STC). To this end, in chapter 2, we developed a recurrent neural network model using an adaptation of the theoretical model of synaptic plasticity and STC by Li et al. (2016). To achieve better biological plausibility of synaptic currents, we introduced a finite synaptic time constant and synaptic delay; to obtain better computational tractability, we changed from MAT2 neurons to LIF neurons; and finally, we corrected the noise term for early-phase plasticity, which had been too strong in the implementation of Li et al., 2016 (cf. section 2.4). Following these modifications, we still obtained the same qualitative and approximately the quantitative results as Li et al. (2016), which match well-established experimental findings (e.g., Frey and Morris, 1997; Sajikumar and Frey, 2004b).

As to previous models of synaptic consolidation via STC, there are to our knowledge only two candidates: the model by Papper et al. (2011) and the one by Ziegler et al. (2015). The model by Papper et al. (2011) conclusively describes prolonged lifetime of memories selected for consolidation and implements memory consolidation in recurrent networks via STC. Nevertheless, the model overall remains on a very abstract level, consisting of binary McCulloch-Pitts neurons and only allowing discrete values of synaptic strength. Furthermore, Papper et al. (2011) have mentioned that their model cannot account for the direct transition into late-phase states following weak stimulation, which is a basic property of STC (cf. Frey and Morris, 1997) that our model enables inherently. According to these aspects, our model seems to provide a much more realistic account of neuronal and synaptic processes, especially on fast timescales (by virtue of spiking dynamics). The feedforward network model by Ziegler et al. (2015) features various biological details, and reproduces a wide range of data from in vitro and even from behavioral tagging experiments. It has, however, not yet been extended to recurrent neural networks. Multi-phase plasticity mechanisms other than STC shall not be considered here, since there is strong experimental evidence that synaptic consolidation occurs through mechanisms of STC (Dudai, 2004; Pastalkova et al., 2006; Moncada and Viola, 2007; Wang et al., 2010; Takeuchi et al., 2016). In conclusion, our model based on spiking, calcium, and protein dynamics seems to be the first to bridge the gap between biologically realistic synaptic dynamics on the one hand and memory consolidation at the level of recurrent networks on the other hand.

There is, however, one drawback: albeit not being entirely impossible (cf. Brunel, 2000; Fourcaud and Brunel, 2002), analytical approaches to investigate the dynamics of spiking models are limited because nonlinearities in differential equations often prevent the analytical expression of solutions. This has also been the case for parts of our detailed two-phase plasticity model. While we could derive analytical expressions to investigate the parameter space for passive improvement and deterioration by solving the equations for L-LTP dynamics (see Eqs. 2.1 & 2.2), variable coefficients in the non-homogeneous differential equations prevented the derivation of analytical expressions for early-phase dynamics, as well as for L-LTD dynamics with simultaneous protein synthesis. Hence, for further investigations, we had to rely on the numerical exploration of emergent dynamics arising from our model. To this end, we investigated the phase space of model parameters in different paradigms, aiming to show by proof of principle that the mechanisms of our model serve to explain experimental findings on cognitive and behavioral phenomena.

Along with our model, we also developed methods to measure the performance in recalling a memory, the recall quality. One of these measures is mutual information – a concept that was first introduced by Shannon (1948) in his hallmark paper on information theory – which is a widely used tool in neuroscience and describes the information that one probability distribution conveys about another (in other words: by which it reduces the uncertainty of another). With respect to our application (chapters 2 and 3), mutual information can be considered as a measure of self-organized pattern structure. The other measure of recall quality that we introduced in chapter 2 is what we termed the “pattern completion coefficient” Q . We have designed this measure such that it describes for an input-defined pattern/cell assembly by how much the activity of indirectly activated assembly neurons (“ans” neurons) differs from the activity of control (“ctrl”) neurons, in relation to the activity of directly stimulated neurons (“as” neurons). We found that while both measures may often yield similar qualitative descriptions of recall quality, they are fundamentally different and can reveal functional differences

in network dynamics (see chapters 2 and 3). In chapter 4, we did not make use of these measures, particularly because the activation of cell assemblies was strong enough to be detected by the mere spike count within an assembly during a small time window. In addition, for the host of investigations in that chapter, which target non-cued spontaneous reactivation (with the aim to relate to free recall experiments), the pattern completion coefficient Q would mathematically not be defined. In a quick check we observed that, as expected, values of Q during recall with $r = 0.5$ for parameters from section 4.2 were high compared to those from chapters 2 and 3, and we observed even higher values for parameters from section 4.4, i.e., for quasi-attractors (results not shown). In addition to Q and mutual information to measure the recall of spatial patterns (spatial in the sense of network topology), we introduced a measure for the recall of temporal patterns in chapter 3. For this measure, we used the goodness of fit (R^2) achieved by fitting a projection of the temporal structure during recall to the temporal structure that was present during learning.

To scrutinize the capability of STC to retroactively control memory traces, in chapter 3, we extended our model by a neuromodulator dependence for the protein synthesis threshold. The existence of a dopamine-dependent threshold has previously been proposed based on experimental findings with dopamine blocking (Navakkode et al., 2007; Wang et al., 2010) and been formulated in a general theoretical form by Clopath et al. (2008) and Ziegler et al. (2015). Here, we employed the formulation by Clopath et al. (2008), assuming an abstract neuromodulator (cf. section 1.3.6 on candidate neuromodulators).

Regarding the numerical implementation of our model, there has been one possible caveat: to increase the computing performance, we disabled early-phase plasticity in the network during periods without external stimulation (cf. Methods in chapters 2–4), which may have altered the results. Biologically, this approach can be explained, for example, by a novelty signal and related neuromodulation, as we discussed it in section 1.3.6. The approach can also be justified computationally, however. For networks as in chapters 2 and 3, with cell assemblies of 150 neurons, we showed that the approach seems to approximate the full simulation comprising spiking and calcium dynamics very well (Fig. 2.11). In addition, we found that for any two pre- and postsynaptic neurons with firing rates below ~ 3 Hz, synaptic strength will not be altered (see Fig. 2.17). Hence, switching off early-phase plasticity in periods without external stimulation is a biologically and computationally justified feature of our model.

To investigate spontaneous transient activation as well as attractor behavior in chapter 4, we needed to consider larger cell assemblies than in the previous chapters. Specifically, we considered assemblies of 600 neurons, as compared to 50–500 neurons in chapter 2 and 150 neurons in chapter 3. To maintain stable network dynamics in spite of these large assemblies, we had to adapt our model so as to avoid that learning gave rise to divergent activity. To this end, we introduced a firing rate threshold for the induction of LTP, based on the Hebbian assumption that both pre- or postsynaptic neuron have to fire sufficiently to evoke LTP (Hebb, 1949; Bliss and Collingridge, 1993; Gerstner et al., 2018). This is further supported by experimental evidence (Sjöström et al., 2001). Note that Päpper et al. (2011) used a similar approach: to control for their cue-target associations in an attractor paradigm, they used an alternative learning rule to connect neurons within a random pattern by LTP and to disconnect these neurons from the rest of the network by LTD.

In conclusion, we have developed a model that serves to describe synaptic consolidation. The model

enables the variation of a generic neuromodulator signal for late-phase plasticity, and approximates the weight dynamics in periods without external stimulation. Furthermore, it extends previous plasticity models (Graupner and Brunel, 2012; Li et al., 2016) to account for the Hebbian postulate that presynaptic and postsynaptic neuron should fire sufficiently to evoke LTP. In addition to this model, we have presented methods to assess the recall quality for memories of spatial and temporal patterns before and after consolidation.

5.2 Cognitive phenomena described by the model

In this thesis, we implemented the dynamics of two-phase long-term synaptic plasticity with synaptic tagging and capture (STC) in a mechanistic model, well-founded on previous experimental results. Investigating these dynamics in a recurrent spiking neural network, we showed, by proof of principle, that the model describes both learning and subsequent consolidation of memory representations. The timescales of this STC-based approach match those of synaptic consolidation, that is, minutes to hours. We further characterized this consolidation and found that it depends on conditions such as inhibition and cell assembly size (see chapter 2). Besides this, our first aim comprised to assess improvement of memory representations by STC, or more specifically, improvement of memory recall. We found indeed that STC enables improvement of memory recall in two different ways (see chapter 2). First, we found passive improvement to arise from the transfer of early-phase weights to the late phase. We analyzed the regime of this improvement analytically and found that there is a wide range of late-phase- and protein-related parameter settings in which passive improvement would take place, while there is also a wide range in which the memory representation would deteriorate. Second, we observed active improvement, occurring during the recall process itself and caused by increasing early-phase weights. This property is an inherent feature of our two-phase plasticity mechanism and does not appear in recall before consolidation when early-phase weights are still high.

The second of our main aims was to explore the capacity of STC to control stored information retroactively, which we pursued in chapter 3. To do this, we extended the model by a neuromodulator-dependent protein synthesis threshold. This enabled, as expected, the modeling of the prevention of memory consolidation through blocked neuromodulator (as shown by experiments; Moncada and Viola, 2007; Wang et al., 2010). On the other hand, we found that the amount of neuromodulator could retroactively control whether the recall of input-defined patterns or the recall of self-organized patterns and temporal structures would be enhanced. While the subject of this theoretical examination are abstract processes that still require rigorous experimental characterization, the results demonstrate that STC yields more than just consolidation of information: it also shapes the information that is consolidated. Adding support to this, recent experimental studies have pointed into the direction of differential processes to store episodic and semantic memories, driven by different amounts of dopamine, in the hippocampus (Takeuchi et al., 2016; Duzskiewicz et al., 2019).

Our third major aim was to show that our model serves to explain findings from neuropsychological experiments that involve multiple memory traces (again, the focus was on the timescales of STC, i.e., minutes to hours). Therefore, in chapter 4, we considered networks that were larger than those in chapters 2 and 3, in order to describe multiple cell assemblies. The actual cell assemblies were also larger, because we aimed to observe spontaneous activation rather than cued recall as before. To

account for such large structures, we had to extend our model by a threshold for Hebbian LTP (see discussion in previous section 5.1). Using this setup, we were able to show that our model accounts for recency effects similar to such found in free recall experiments (cf. Bjork and Whitten, 1974; Howard and Kahana, 1999; Lohnas et al., 2015). Considering overlaps between memory representations, we found retroactive interference or enhanced recall, depending on the protocol. This may explain findings from neuropsychological experiments in which effects related to the similarity between memories have been observed (cf. Anderson and Neely, 1996; Preston and Eichenbaum, 2013). By performing control simulations in which we blocked LTD, we demonstrated that LTD is the process underlying these effects. Furthermore, comparing the findings of two protocols, either learning the three assemblies first and then consolidating them (“standard” protocol), or consolidating after learning each individual assembly (“intermediate consolidation” protocol), we discovered significant differences. This reveals a functional role of STC in organization of memory representations, which besides our findings in chapters 2 and 3 once more emphasizes the importance of STC for the implementation of cognitive function. Next, we strove to employ STC to describe yet another phenomenon, namely, direct priming (cf. Janiszewski and Wyer, 2014; Bermeitinger, 2015). Following the application of a short stimulus to one of the cell assemblies in the network, we could enhance the likelihood of its reactivation for several hours. We showed that this enhanced activation is due to raised early-phase weights, which decay on a timescale of several hours. Thus, the enhanced recall caused by the priming stimulus shares its origin with the active improvement discussed before (cf. chapter 2) – both are enabled by the capacity of early-phase weights to undergo tremendous plastic changes once the memory representation has been consolidated. The description of priming effects provided by our model can likely be extended to account for other forms of priming, such as semantic priming or negative priming (Bermeitinger, 2015; Was et al., 2019). Finally, in further investigations in chapter 4, we set up our model to give rise to attractor dynamics, which has been considered in many previous studies to explain aspects of working memory (Amit and Brunel, 1997b; Kropff and Treves, 2006; Barak and Tsodyks, 2014; Fusi and Wang, 2016) and to explain findings from free recall experiments in relation to long-term memory (Romani et al., 2013; Recanatesi et al., 2015). Seeking to connect to the latter, we considered transitions between activated, long-term attractor cell assemblies and found significant differences between overlapping and non-overlapping assemblies, while we also reproduced our previous results (obtained with transient activation dynamics) on the recency effect. Moreover, our results indicate that activity within an intersection between cell assemblies may guide transitions by acting as a cue, as it was suggested earlier (Kropff and Treves, 2006; Recanatesi et al., 2017; Fischer et al., 2020).

In conclusion, the results contained within this thesis have shown that synaptic memory consolidation based on spike-driven synaptic calcium dynamics and synaptic tagging and capture (STC) can account for diverse cognitive effects as they have been observed in neuropsychological experiments. Furthermore, the predictions of our model can be tested in new experiments on the behavioral as well as on the neuronal level.

5.3 Outlook

In chapter 3, we have discussed the impact of the timing of a neuromodulatory signal, mentioning that there is a fundamental difference between the impact of neuromodulator (acting on protein synthesis)

and the impact of the learning stimulation caused by the nonlinear nature of the synaptic tag. Quantitatively examining this influence of the timing of neuromodulation would complement our theoretical consideration and enable to make precise predictions that can be tested in experimental setups. Furthermore, more detailed description of the action of specific neuromodulators such as dopamine and noradrenaline would enable even more precise predictions, and could lead to interesting dynamics: for example, different neuromodulators could gate plasticity on different timescales and thereby give rise to a larger diversity of functional subpopulations of synapses than we observed for one neuromodulator in chapter 3. There is already a lot of experimental evidence for different neuromodulators acting on different timescales (Brzosko et al., 2017; Duzskiewicz et al., 2019; Kim et al., 2020; Liu et al., 2021). In section 5.1, we discussed analytical approaches with respect to our model. Analytical investigations of the dynamics of neuromodulator-dependent L-LTP could, in the future, extend our current numerical results which were presented in chapter 3. This could be of particular use for examining the impact of the timing of a neuromodulatory signal.

In this thesis, we considered plasticity-related proteins (PRPs) that were unspecific to potentiation and depression, for which there is solid evidence from experimental and theoretical findings (Frey and Morris, 1997; Sajikumar and Frey, 2004a; Sajikumar et al., 2007; Clopath et al., 2008; Li et al., 2016). There are, however, also indications for the existence of additional, specific proteins for potentiation or depression (Sajikumar et al., 2007; cf. section 1.3.4). Therefore, it will be of great interest for further studies to examine the impact of such different protein pools, which may reveal additional functional implications of the interplay between potentiation and depression. We already tried to do this by setting up a relatively simple phenomenological model (not presented in this thesis). This model stochastically describes the transitions between different discrete early- and late-phase weight states of individual synapses. All synapses together form the population of synapses belonging to a neuron, which holds up to three different protein pools to influence the late-phase transitions. However, after multiple attempts of optimization to find a set of model parameters that would reproduce the outcome of typical plasticity-inducing experiments (cf. Fig. 2.2), we came to the conclusion that this model was probably not sufficient to explain the basics of synaptic consolidation. Hence, a more sophisticated, ideally mechanistic model seems necessary. The model presented in this thesis is such a model, and it accounts for the outcome of typical plasticity-inducing experiments even with a single protein pool, which we have shown in chapter 2. Nevertheless, the present software implementation of our model (Luboeinski, 2021) already features the necessary variables and functions to simulate three different protein pools. Thus, computational investigations, and more experiments that yield data to adjust the parameters, await to be done. There are plenty of consequences that can arise from the existence of specific protein pools. As an example, one could think of a mechanistic explanation for reconsolidation, which we will consider in the following.

After several hours, the weights in our model eventually enter a steady state describing, in principle, static memory representations. If, however, external learning or recall stimulation is applied, plasticity can occur once again, which we considered at several points throughout this thesis (cf. active improvement in chapter 2; priming in chapter 4). There is also experimental evidence that consolidated memories become labile upon recall, which has most dramatically been demonstrated by studies on *reconsolidation* and related memory disruption (cf. section 1.4.8). The experimental findings of such dynamic behavior leads to the question: How could our model account for the disruption of already

consolidated memories upon recall? First, this could be accounted for by a stimulus that primarily elicits LTD, which would weaken the connections within the cell assembly. Induced by early-phase depression, the disruption would become permanent by being transferred to the late phase. Nevertheless, reconsolidation experiments typically show disruption of the memory trace through blocked protein synthesis, whereas the mere recall stimulation does not evoke disruption. In the current state, our model does not yet capture the disruption of blocked reconsolidation by blocked protein synthesis. However, a certain balance of early-phase LTP and LTD in combination with differential consolidation dynamics for LTP and LTD may explain such disruption/impaired reconsolidation. In the future, this could be modeled by introducing LTP- and LTD-specific protein pools with different properties (which is, as mentioned above, already implemented in our simulation code). This could complement another recently proposed mechanism for reconsolidation which is based on synaptic scaling (Amorim et al., 2021).

Synthesis of plasticity-related proteins (PRPs) does not only occur in the soma, but also locally in the dendrites (Martin et al., 1997; Glock et al., 2017). Since late-phase plasticity depends not only on PRPs but also on transcription in the nucleus (Bliss and Collingridge, 1993; Abraham et al., 2019), it can be justified to model late-phase plasticity as a somatic process. Clustering of synapses with similar response characteristics, for example, can provide a situation in which STC based on local protein synthesis would yield exactly the same results as if it were based on somatic protein synthesis (cf. Kastellakis and Poirazi, 2019; Pinho et al., 2020). Whether the results would be different or not would then depend on the particular parameter set of the model. If, however, the translation- and transcription-dependent phases of long-term plasticity are to be considered separately (cf. section 1.3.4), local protein synthesis in the dendrites becomes important. There is a multitude of functional consequences that can arise from spatial compartmentalization, one of which is the competition of synapses for proteins (cf. Fonseca, 2015). To account for this, our model would need to be modified such that proteins be “consumed” by synapses instead of decaying globally, and it would have to be extended by spatial compartmentalization. Spatial compartmentalization of protein synthesis has been investigated in computational models by O’Donnell and Sejnowski (2014) and by Fonkeu et al. (2019), which could serve as a basis for studying the impact of local dendritic protein synthesis on synaptic consolidation with our model.

As discussed in section 1.3.4, there is a variety of plasticity mechanisms in the brain. In this thesis, we focused on long-term plasticity between excitatory synapses, by which we were able to model memories as Hebbian cell assemblies. Nevertheless, other mechanisms might drastically alter the dynamics of our network model and give rise to enhanced functionality that the brain might use. Inhibitory plasticity, for example, has been shown to enhance self-organization (Vogels et al., 2011; Zenke et al., 2015) and the discrimination between memories (Herpich and Tetzlaff, 2019). Such organizational properties may be examined in our model in similar investigations as those presented in chapter 4. As the model presented in this thesis does not feature neuronal adaptation (or intrinsic plasticity), extending the model by such mechanisms would be especially interesting with regard to the attractor dynamics that we discussed in section 4.4. Neuronal adaptation could exert influence by turning off an attractor, thereby initiating the activation of another attractor. This would also enable more exact reproduction of earlier findings on such dynamics (cf. Kropff and Treves, 2006; Russo and Treves, 2012). A good candidate for a neuron model featuring neuronal adaptation is, for instance, the MAT2 model (see

section 1.3.1). This model has already been implemented in our simulation code (Luboeinski, 2021), and first attempts to employ it within our framework have been made (see section 2.4). Similarly to neuronal adaptation, introducing short-term depression into our model could be of special interest for investigations of transitions between attractors, as we considered them in section 4.4. In the sense of Mongillo et al., 2008 (also cf. section 1.4.6), short-term depression could turn off an attractor after some seconds and thereby enable the activation of another attractor.

To summarize, employing one or several of these different plasticity mechanisms could yield new functional properties that our model thus far cannot account for. Beyond that, it would be interesting to explore how effects such as memory improvement, which was considered in chapters 2 and 3, or priming, which was investigated in chapter 4, would be altered by homeostatic processes such as neuronal adaptation and short-term depression.

Dendritic spikes enable nonlinear computation in dendrites and have been related to neuronal activity patterns during sleep (Häusser et al., 2000; Jahnke et al., 2015). The simulation code that we used for this thesis (Luboeinski, 2021) already enables the utilization of conductance-based synapses and dendritic spikes (as in Jahnke et al., 2015) with our model. Thus, investigations that would consider, for example, the improvement of memory recall through positive feedback conveyed by dendritic spikes, may be performed at relatively low effort. Employing dendritic spikes to consider sleep-related activities in our model could, furthermore, yield interesting insights into the interplay between synaptic consolidation and systems consolidation (see section 1.4.4).

Artificial neural networks, essential for machine learning and data science, mostly feature units similar to rate-based neurons at the present time. However, the application of spiking mechanisms seems to become increasingly important in this field (Gupta and Long, 2007; Soni, 2018). At the same time, neuromorphic hardware promotes the use of spike coding by offering energy-efficient computing solutions (Wang et al., 2017; Lin et al., 2018; Michaelis et al., 2020). Thereby, biologically-inspired spiking neural networks may become the basis for more fruitful exchange between neuroscience and engineering.

In addition, computation with cell assemblies is not only of outstanding importance to understand the cognitive processes in the brain, but also enables a large variety of applications with artificial neural networks (cf. Nadh and Huyck, 2010; Tetzlaff et al., 2015; Kreiser et al., 2018). While the opportunities of application continue to multiply, two-phase plasticity models describing the dynamics of cell assemblies can become increasingly important for efficient storing of information in neural networks. Our findings of improvement, retroactive modification, and priming may contribute to this.

5.4 Final remarks

Any basic research should – in principle – be acceptable, assuming that “every thought possible will be thought” (cf. Dürrenmatt, 1962). This would include the investigation of the neural underpinnings of memory. At the same time, however, there are issues of privacy and liberty: Will it at some point be feasible to read memories from the human brain, or, to implant new memories by artificial alteration of synapses (see “mimicry” in Martin et al., 2000; Abraham et al., 2019)? The answer is probably, yes. Nevertheless, it remains unclear under which circumstances this will be feasible. It might be relieving that reading or implanting memories *noninvasively* in a live human being will probably stay impossi-

ble. However, reading memories via spectroscopy after injections of a certain dye, or posthumously from brain slices, might become possible in the future. We should hope that strong ethical reservations will accompany such applications in humans and we should establish according constraints. On the other hand, we should keep in mind the chances, for clinical treatment of severe diseases as well as for the development of technology based on artificial intelligence, which arise from understanding how the brain implements memory.

Bibliography

- ABDOU, K., SHEHATA, M., CHOKO, K., NISHIZONO, H., MATSUO, M., MURAMATSU, S.-I., AND INOKUCHI, K. 2018. Synapse-specific representation of the identity of overlapping memory engrams. *Science* 360:1227–1231.
- ABRAHAM, W. C. 2003. How long will long term potentiation last? *Philosophical Transactions of the Royal Society B* 358:735–744.
- ABRAHAM, W. C., JONES, O. D., AND GLANZMAN, D. L. 2019. Is plasticity of synapses the mechanism of long-term memory storage? *NPJ Science of Learning* 4:1–10.
- ABRAHAM, W. C., LOGAN, B., GREENWOOD, J. M., AND DRAGUNOW, M. 2002. Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *Journal of Neuroscience* 22:9626–9634.
- ADACHI, M. AND AIHARA, K. 1997. Associative dynamics in a chaotic neural network. *Neural Networks* 10:83–98.
- ADRIAN, E. D. 1926. The impulses produced by sensory nerve endings: Part I. *The Journal of Physiology* 61:49–72.
- ADRIAN, E. D. AND ZOTTERMAN, Y. 1926. The impulses produced by sensory nerve-endings: Part II. The response of a single end-organ. *The Journal of Physiology* 61:151–171.
- AHMED, S., BIERLEY, R., SHEIKH, J. I., AND DATE, E. S. 2000. Post-traumatic amnesia after closed head injury: a review of the literature and some suggestions for further research. *Brain Injury* 14:765–780.
- AKERBOOM, J., CARRERAS CALDERÓN, N., TIAN, L., WABNIG, S., PRIGGE, M., TOLÖ, J., GORDUS, A., ORGER, M. B., SEVERI, K. E., MACKLIN, J. J., ET AL. 2013. Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics. *Frontiers in Molecular Neuroscience* 6:2.
- AMARAL, D. G., ISHIZUKA, N., AND CLAIBORNE, B. 1990. Neurons, numbers and the hippocampal network, pp. 1–11. In J. Storm-Mathisen, J. Zimmer, and O. P. Ottersen (eds.), *Understanding the Brain Through the Hippocampus*, volume 83 of *Progress in Brain Research*, chapter 1. Elsevier, Amsterdam, Netherlands.
- AMIT, D. J. 1989. *Modeling Brain Function: The World of Attractor Neural Networks*. Cambridge University Press, Cambridge, UK.
- AMIT, D. J. AND BRUNEL, N. 1997a. Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Computation in Neural Systems* 8:373–404.
- AMIT, D. J. AND BRUNEL, N. 1997b. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex* 7:237–252.
- AMIT, D. J. AND TSODYKS, M. V. 1991a. Quantitative study of attractor neural network retrieving at low spike rates: I. Substrate—spikes, rates and neuronal gain. *Network: Computation in Neural Systems* 2:259–273.
- AMIT, D. J. AND TSODYKS, M. V. 1991b. Quantitative study of attractor neural networks retrieving at low spike rates: II. Low-rate retrieval in symmetric networks. *Network: Computation in Neural Systems* 2:275–294.
- AMORIM, F. E., CHAPOT, R. L., MOULIN, T. C., LEE, J. L. C., AND AMARAL, O. B. 2021. Memory destabilization during reconsolidation – a consequence of homeostatic plasticity? *Preprint at bioRxiv* 2021.02.05.429950.
- ANDERSON, J. R. 1981. Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory* 7:326.

Bibliography

- ANDERSON, M. C. AND NEELY, J. H. 1996. Interference and inhibition in memory retrieval, pp. 237–313. *In* E. L. Bjork and R. A. Bjork (eds.), *Memory*. Academic Press, San Diego/CA, USA.
- ARBIB, M. A. 2003. *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge/MA, USA, 2nd edition.
- ARTOLA, A., BRÖCHER, S., AND SINGER, W. 1990. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* 347:69–72.
- ARTOLA, A. AND SINGER, W. 1993. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends in Neurosciences* 16:480–487.
- ASCHER, P. AND NOWAK, L. 1988. The role of divalent cations in the N-methyl-D-aspartate responses of mouse central neurones in culture. *The Journal of Physiology* 399:247–266.
- ATKINSON, R. C. AND SHIFFRIN, R. M. 1968. Human memory: A proposed system and its control processes, pp. 89–195. *In* K. Spence and J. T. Spence (eds.), *Psychology of Learning and Motivation*, volume 2. Academic Press, New York/NY, USA.
- AVERBECK, B. B., LATHAM, P. E., AND POUGET, A. 2006. Neural correlations, population coding and computation. *Nature Reviews Neuroscience* 7:358–366.
- BAARS, B. J. 1997. Some essential differences between consciousness and attention, perception, and working memory. *Consciousness and Cognition* 6:363–371.
- BADDELEY, A. D. AND HITCH, G. 1974. Working memory, pp. 47–89. *In* G. H. Bower (ed.), *Psychology of Learning and Motivation*, volume 8. Academic Press, New York/NY, USA.
- BAHRICK, H. P. 1970. Two-phase model for prompted recall. *Psychological Review* 77:215.
- BAIN, A. 1873. *Mind and Body: The Theories of their Relation*, volume 4. D. Appleton & Co., New York/NY, USA.
- BALLARD, P. B. 1913. Obliviscence and Reminiscence, volume 1 of *British Journal of Psychology – Monograph Supplements*. Cambridge University Press, London, UK.
- BALLARINI, F., MONCADA, D., MARTINEZ, M. C., ALEN, N., AND VIOLA, H. 2009. Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences of the USA* 106:14599–14604.
- BARAK, O. AND TSODYKS, M. 2014. Working models of working memory. *Current Opinion in Neurobiology* 25:20–24.
- BARBIERI, F. AND BRUNEL, N. 2008. Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex? *Frontiers in Neuroscience* 2:3.
- BARBOUR, B., BRUNEL, N., HAKIM, V., AND NADAL, J.-P. 2007. What can we learn from synaptic weight distributions? *Trends in Neurosciences* 30:622–629.
- BARCO, A., PATTERSON, S., ALARCON, J. M., GROMOVA, P., MATA-ROIG, M., MOROZOV, A., AND KANDEL, E. R. 2005. Gene expression profiling of facilitated L-LTP in VP16-CREB mice reveals that BDNF is critical for the maintenance of LTP and its synaptic capture. *Neuron* 48:123–137.
- BARLOW, H. B. 1953. Summation and inhibition in the frog's retina. *The Journal of Physiology* 119:69–88.
- BARRETT, A. B., BILLINGS, G. O., MORRIS, R. G. M., AND VAN ROSSUM, M. C. W. 2009. State based model of long-term potentiation and synaptic tagging and capture. *PLOS Computational Biology* 5:e1000259.
- BARTLETT, F. C. 1932. *Remembering: A study in Experimental and Social Psychology*. Cambridge University Press, London, UK.
- BASDEN, B. H., BASDEN, D. R., AND GARGANO, G. J. 1993. Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:603.
- BEAR, M. F. 1996. A synaptic basis for memory storage in the cerebral cortex. *Proceedings of the National Academy of Sciences of the USA* 93:13453–13459.

- BEAR, M. F. AND ABRAHAM, W. C. 1996. Long-term depression in hippocampus. *Annual Review of Neuroscience* 19:437–462.
- BEAR, M. F., CONNORS, B. W., AND PARADISO, M. A. 2016. Neuroscience: Exploring the Brain. Wolters Kluwer, Philadelphia/PA, USA, 4th edition.
- BECKER, M. F. P. AND TETZLAFF, C. 2021. The biophysical basis underlying the maintenance of early phase long-term potentiation. *PLOS Computational Biology* 17:e1008813.
- BECKER, S., MOSCOVITCH, M., BEHRMANN, M., AND JOORDENS, S. 1997. Long-term semantic priming: a computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23:1059.
- BEGGS, J. M. AND PLENZ, D. 2003. Neuronal avalanches in neocortical circuits. *Journal of Neuroscience* 23:11167–11177.
- BENNA, M. K. AND FUSI, S. 2016. Computational principles of synaptic memory consolidation. *Nature Neuroscience* 19:1697–1706.
- BERGSTRÖM, J. A. 1893. Experiments upon physiological memory by means of the interference of associations. *The American Journal of Psychology* 5:356–369.
- BERKE, J. D. AND HYMAN, S. E. 2000. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 25:515–532.
- BERMAN, D. E. AND DUDAI, Y. 2001. Memory extinction, learning anew, and learning the new: dissociations in the molecular machinery of learning in cortex. *Science* 291:2417–2419.
- BERMEITINGER, C. 2015. Priming, pp. 42–88. In M. Khosrow-Pour et al. (eds.), *Psychology and Mental Health: Concepts, Methodologies, Tools, and Applications*. IGI Global, Hershey, PA/USA.
- BERNSTEIN, J. 1868. Ueber den zeitlichen Verlauf der negativen Schwankung des Nervenstroms. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 1:173–207.
- BERNSTEIN, J. 1902. Untersuchungen zur Thermodynamik der bioelektrischen Ströme. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 92:521–562.
- BERRIDGE, K. C. 2009. Wanting and liking: Observations from the neuroscience and psychology laboratory. *Inquiry* 52:378–398.
- BESNARD, A., CABOCHE, J., AND LAROCHE, S. 2012. Reconsolidation of memory: A decade of debate. *Progress in Neurobiology* 99:61–80.
- BETTS, J. G., YOUNG, K. A., WISE, J. A., JOHNSON, E., POE, B., KRUSE, D. H., KOROL, O., JOHNSON, J. E., WOMBLE, M., AND DESAIX, P. 2013. *Anatomy and Physiology*. OpenStax, Houston/TX, USA. Access for free at <https://openstax.org/books/anatomy-and-physiology/pages/1-introduction>.
- BHALLA, U. S. AND IYENGAR, R. 1999. Emergent properties of networks of biological signaling pathways. *Science* 283:381–387.
- BI, G.-Q. AND POO, M.-M. 1998. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience* 18:10464–10472.
- BIENENSTOCK, E. L., COOPER, L. N., AND MUNRO, P. W. 1982. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2:32–48.
- BJORK, R. A. AND BJORK, E. L. 1992. A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, and R. M. Shiffrin (eds.), *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes*, volume 2, chapter 2. L. Erlbaum, Hillsdale/NJ, USA.
- BJORK, R. A. AND WHITTEN, W. B. 1974. Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology* 6:173–189.
- BLISS, T. V. P. AND COLLINGRIDGE, G. L. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361:31.

Bibliography

- BLISS, T. V. P. AND GARDNER-MEDWIN, A. R. 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the unanaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology* 232:357–374.
- BLISS, T. V. P. AND LØMO, T. 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology* 232:331–356.
- BONILLA-QUINTANA, M. AND WÖRGÖTTER, F. 2021. Exploring new roles for actin upon LTP induction in dendritic spines. *Scientific Reports* 11.
- BONILLA-QUINTANA, M., WÖRGÖTTER, F., D'ESTE, E., TETZLAFF, C., AND FAUTH, M. 2021. Reproducing asymmetrical spine shape fluctuations in a model of actin dynamics predicts self-organized criticality. *Scientific Reports* 11.
- BRADSKI, G., CARPENTER, G. A., AND GROSSBERG, S. 1994. Store working memory networks for storage and recall of arbitrary temporal sequences. *Biological Cybernetics* 71:469–480.
- BRAITENBERG, V. 1978. Cell assemblies in the cerebral cortex, pp. 171–188. In R. Heim et al. (eds.), *Theoretical Approaches to Complex Systems*. Springer, Berlin, Germany.
- BRAITENBERG, V. AND SCHÜZ, A. 1998. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin, Germany, 2nd edition.
- BRAMHAM, C. R. 2008. Local protein synthesis, actin dynamics, and LTP consolidation. *Current Opinion in Neurobiology* 18:524–531.
- BRETTE, R. 2015. Philosophy of the spike: Rate-based vs. spike-based theories of the brain. *Frontiers in Systems Neuroscience* 9:151.
- BRETTE, R. 2019. Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences* 42.
- BROCK, L. G., COOMBS, J. S., AND ECCLES, J. C. 1952. The recording of potentials from motoneurons with an intracellular electrode. *The Journal of Physiology* 117:431–460.
- BRODIE, D. A. AND MURDOCK, JR., B. B. 1977. Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior* 16:185–200.
- BRODMANN, K. 1909. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth, Leipzig, Germany.
- BROWN, G. D., NEATH, I., AND CHATER, N. 2007. A temporal ratio model of memory. *Psychological Review* 114:539.
- BROWN, R. AND KULIK, J. 1977. Flashbulb memories. *Cognition* 5:73–99.
- BRUNEL, N. 2000. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience* 8:183–208.
- BRUNEL, N. 2016. Basic neuron and network model, pp. 73–99. In M. A. Arbib (ed.), *From Neuron to Cognition via Computational Neuroscience*, chapter 2. MIT Press, Cambridge/MA, USA.
- BRUNEL, N. AND VAN ROSSUM, M. C. W. 2007. Lapicque's 1907 paper: from frogs to integrate-and-fire. *Biological Cybernetics* 97.
- BRZOSKO, Z., SCHULTZ, W., AND PAULSEN, O. 2015. Retroactive modulation of spike timing-dependent plasticity by dopamine. *eLife* 4:e09685.
- BRZOSKO, Z., ZANNONE, S., SCHULTZ, W., CLOPATH, C., AND PAULSEN, O. 2017. Sequential neuromodulation of hebbian plasticity offers mechanism for effective reward-based navigation. *eLife* 6:e27756.
- BUGMANN, G. 1997. Biologically plausible neural computation. *Biosystems* 40:11–19.
- BURNS, B. D. AND WEBB, A. C. 1976. The spontaneous activity of neurones in the cat's cerebral cortex. *Proceedings of the Royal Society B* 194:211–223.

- BUXHOEVEDEN, D. P. AND CASANOVA, M. F. 2002. The minicolumn hypothesis in neuroscience. *Brain* 125:935–951.
- BUZSÁKI, G. 2006. *Rhythms of the Brain*. Oxford University Press, Oxford, UK.
- BUZSÁKI, G. 2010. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron* 68:362–385.
- BUZSÁKI, G. AND MIZUSEKI, K. 2014. The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience* 15:264–278.
- CAI, D. J., AHARONI, D., SHUMAN, T., SHOBE, J., BIANE, J., SONG, W., WEI, B., VESHKINI, M., LA-VU, M., LOU, J., ET AL. 2016. A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534:115–118.
- CHANG, Y. 2014. Reorganization and plastic changes of the human brain associated with skill learning and expertise. *Frontiers in Human Neuroscience* 8:35.
- CHATER, T. E. AND GODA, Y. 2021. My neighbour hetero—deconstructing the mechanisms underlying heterosynaptic plasticity. *Current Opinion in Neurobiology* 67:106–114.
- CHETTIH, S. N. AND HARVEY, C. D. 2019. Single-neuron perturbations reveal feature-specific competition in V1. *Nature* 567:334–340.
- CHO, K., AGGLETON, J. P., BROWN, M. W., AND BASHIR, Z. I. 2001. An experimental test of the role of postsynaptic calcium levels in determining synaptic strength using perirhinal cortex of rat. *The Journal of Physiology* 532:459–466.
- CLOPATH, C. 2012. Synaptic consolidation: an approach to long-term learning. *Cognitive Neurodynamics* 6:251–257.
- CLOPATH, C., BÜSING, L., VASILAKI, E., AND GERSTNER, W. 2010. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nature Neuroscience* 13:344.
- CLOPATH, C., ZIEGLER, L., VASILAKI, E., BÜSING, L., AND GERSTNER, W. 2008. Tag-trigger-consolidation: a model of early and late long-term potentiation and depression. *PLOS Computational Biology* 4:e10000248.
- CONNORS, B. W. AND GUTNICK, M. J. 1990. Intrinsic firing patterns of diverse neocortical neurons. *Trends in Neurosciences* 13:99–104.
- CONSTANTINIDIS, C. AND KLINGBERG, T. 2016. The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience* 17:438.
- COOPER, R. P. 2016. Schema theory and neuropsychology, pp. 433–456. In M. A. Arbib (ed.), *From Neuron to Cognition via Computational Neuroscience*, chapter 14. MIT Press, Cambridge/MA, USA.
- COUTANCHE, M. N., KOCH, G. E., AND PAULUS, J. P. 2020. Influences on memory for naturalistic visual episodes: sleep, familiarity, and traits differentially affect forms of recall. *Learning & Memory* 27:284–291.
- COWAN, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:87–114.
- CRAIK, F. I. M. AND TULVING, E. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104:268.
- CRICK, F. AND MITCHISON, G. 1983. The function of dream sleep. *Nature* 304:111–114.
- CUDMORE, R. H. AND DESAI, N. S. 2008. Intrinsic plasticity. *Scholarpedia* 3:1363. revision #129344.
- CUMMINGS, J. A., MULKEY, R. M., NICOLL, R. A., AND MALENKA, R. C. 1996. Ca²⁺ signaling requirements for long-term depression in the hippocampus. *Neuron* 16:825–833.
- CZURKÓ, A., HIRASE, H., CSICSVARI, J., AND BUZSÁKI, G. 1999. Sustained activation of hippocampal pyramidal cells by ‘space clamping’ in a running wheel. *European Journal of Neuroscience* 11:344–352.

Bibliography

- DAVELAAR, E. J., GOSHEN-GOTTSTEIN, Y., ASHKENAZI, A., HAARMANN, H. J., AND USHER, M. 2005. The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review* 112:3.
- DAVIES, M., SRINIVASA, N., LIN, T.-H., CHINYA, G., CAO, Y., CHODAY, S. H., DIMOU, G., JOSHI, P., IMAM, N., JAIN, S., ET AL. 2018. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro* 38:82–99.
- DAYAN, P. AND ABBOTT, L. F. 2001. *Theoretical Neuroscience*. MIT Press, Cambridge/MA, USA, 1st edition.
- DE ALMEIDA VALVERDE ZANINI, G., TUFIK, S., ANDERSEN, M. L., DA SILVA, R. C. M., BUENO, O. F. A., RODRIGUES, C. C., AND POMPEIA, S. 2012. Free recall of word lists under total sleep deprivation and after recovery sleep. *Sleep* 35:223–230.
- DE FALCO, E., ISON, M. J., FRIED, I., AND QUIROGA, R. Q. 2016. Long-term coding of personal and universal associations underlying the memory web in the human brain. *Nature Communications* 7:1–11.
- DENÈVE, S. AND MACHENS, C. K. 2016. Efficient codes and balanced networks. *Nature Neuroscience* 19:375.
- DESTEXHE, A., RUDOLPH, M., AND PARÉ, D. 2003. The high-conductance state of neocortical neurons in vivo. *Nature Reviews Neuroscience* 4.
- DOYA, K. 2002. Metalearning and neuromodulation. *Neural Networks* 15:495–506.
- DOYLE, J. M., LOFTUS, E. F., DYSART, J. E., AND NEWIRTH, K. A. 2020. *Eyewitness Testimony: Civil and Criminal*. LexisNexis, Dayton/OH, USA, 6th edition.
- DRAGOI, G. 2020. Cell assemblies, sequences and temporal coding in the hippocampus. *Current Opinion in Neurobiology* 64:111–118. Systems Neuroscience.
- DUDAI, Y. 2004. The neurobiology of consolidation, or, how stable is the engram? *Annual Review of Psychology* 55:51–86.
- DUDAI, Y., KARNI, A., AND BORN, J. 2015. The consolidation and transformation of memory. *Neuron* 88:20–32.
- DUDEK, S. M. AND BEAR, M. F. 1992. Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proceedings of the National Academy of Sciences of the USA* 89:4363–4367.
- DUNCAN, C. P. 1949. The retroactive effect of electroshock on learning. *Journal of Comparative and Physiological Psychology* 42:32.
- DÜRRENMATT, F. 1962. *Die Physiker: eine Komödie in zwei Akten*. Verlag der Arche, Zürich, Switzerland.
- DUSZKIEWICZ, A. J., MCNAMARA, C. G., TAKEUCHI, T., AND GENZEL, L. 2019. Novelty and dopaminergic modulation of memory persistence: a tale of two systems. *Trends in Neurosciences* 42:102–114.
- DÜZEL, E., PENNY, W. D., AND BURGESS, N. 2010. Brain oscillations and memory. *Current Opinion in Neurobiology* 20:143–149.
- DŽAJA, D., HLADNIK, A., BIČANIĆ, I., BAKOVIĆ, M., AND PETANJEK, Z. 2014. Neocortical calretinin neurons in primates: increase in proportion and microcircuitry structure. *Frontiers in Neuroanatomy* 8:103.
- EBBINGHAUS, H. 1885. *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Duncker & Humblot, Leipzig, Germany.
- EFFENBERGER, F., JOST, J., AND LEVINA, A. 2015. Self-organization in balanced state networks by STDP and homeostatic plasticity. *PLOS Computational Biology* 11:e1004420.
- EICHENBAUM, H. 2011. *The Cognitive Neuroscience of Memory: an introduction*. Oxford University Press, New York/NY, USA, 2nd edition.
- EICHENBAUM, H. 2013. Memory on time. *Trends in Cognitive Sciences* 17:81–88.
- EICHENBAUM, H. 2018. Barlow versus Hebb: When is it time to abandon the notion of feature detectors and adopt the cell assembly as the unit of cognition? *Neuroscience Letters* 680:88–93.
- ELGENDI, M., KUMAR, P., BARBIC, S., HOWARD, N., ABBOTT, D., AND CICHOCKI, A. 2018. Subliminal priming—state of the art and future perspectives. *Behavioral Science* 8:54.

- ELIASMITH, C. 2005. A unified approach to building and controlling spiking attractor networks. *Neural Computation* 17:1276–1314.
- ELLIOTT, M. C., TANAKA, P. M., SCHWARK, R. W., AND ANDRADE, R. 2018. Serotonin differentially regulates L5 pyramidal cell classes of the medial prefrontal cortex in rats and mice. *eNeuro* 5.
- ELLIOTT, T. 2016. The enhanced rise and delayed fall of memory in a model of synaptic integration: Extension to discrete state synapses. *Neural Computation* 28:1927–1984.
- ELLIOTT, T. AND LAGOIANNIS, K. 2012. The rise and fall of memory in a model of synaptic integration. *Neural Computation* 24:2604–2654.
- ERDELYI, M. H. 2010. The ups and downs of memory. *American Psychologist* 65:623.
- ERDELYI, M. H. AND KLEINBARD, J. 1978. Has Ebbinghaus decayed with time? The growth of recall (hypermnnesia) over days. *Journal of Experimental Psychology: Human Learning and Memory* 4:275.
- ESTES, W. K. 1955. Statistical theory of distributional phenomena in learning. *Psychological Review* 62:369.
- FAUTH, M., WÖRGÖTTER, F., AND TETZLAFF, C. 2015. Formation and maintenance of robust long-term information storage in the presence of synaptic turnover. *PLOS Computational Biology* 11:e1004684.
- FISCHER, C., CZOSCHKE, S., PETERS, B., RAHM, B., KAISER, J., AND BLEADOWSKI, C. 2020. Context information supports serial dependence of multiple visual objects across memory episodes. *Nature Communications* 11:1–11.
- FLOURENS, M. J. P. 1824. Recherches Experimentales sur les Propriétés et les Fonctions du Système Nerveux dans les Animaux Vertébrés. Crevot, Paris, France.
- FONKEU, Y., KRAYNYUKOVA, N., HAFNER, A.-S., KOCHEN, L., SARTORI, F., SCHUMAN, E. M., AND TCHUMATCHENKO, T. 2019. How mRNA localization and protein synthesis sites influence dendritic protein distribution and dynamics. *Neuron* 103:1109–1122.
- FONSECA, R. 2015. Synaptic cooperation and competition: two sides of the same coin?, pp. 29–44. In S. Sajikumar (ed.), *Synaptic Tagging and Capture*. Springer, New York/NY, USA.
- FOSTER, D. J. AND WILSON, M. A. 2007. Hippocampal theta sequences. *Hippocampus* 17:1093–1099.
- FOSTER, M. AND SHERRINGTON, C. S. A. 1897. A Textbook of Physiology, volume 3: The Central Nervous System. MacMillan & Co, London, UK, 7th edition.
- FOURCAUD, N. AND BRUNEL, N. 2002. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Computation* 14.
- FOURCAUD-TROCMÉ, N., HANSEL, D., VAN VREESWIJK, C., AND BRUNEL, N. 2003. How spike generation mechanisms determine the neuronal response to fluctuating inputs. *Journal of Neuroscience* 23:11628–11640.
- FRANK, L. M., BROWN, E. N., AND WILSON, M. A. 2001. A comparison of the firing properties of putative excitatory and inhibitory neurons from CA1 and the entorhinal cortex. *Journal of Neurophysiology* 86:2029–2040.
- FREUND, T. AND BUZSÁKI, G. 1996. Interneurons of the hippocampus. *Hippocampus* 6:347–470.
- FREY, U. AND MORRIS, R. G. M. 1997. Synaptic tagging and long-term potentiation. *Nature* 385:533–536.
- FREY, U., SCHROEDER, H., AND MATTHIES, H. 1990. Dopaminergic antagonists prevent long-term maintenance of posttetanic LTP in the CA1 region of rat hippocampal slices. *Brain Research* 522:69–75.
- FROEMKE, R. C. AND DAN, Y. 2002. Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* 416:433–438.
- FURINI, C. R. G., MYSKIW, J. C., BENETTI, F., AND IZQUIERDO, I. 2013. New frontiers in the study of memory mechanisms. *Brazilian Journal of Psychiatry* 35:173–177.

Bibliography

- FUSI, S., DREW, P. J., AND ABBOTT, L. F. 2005. Cascade models of synaptically stored memories. *Neuron* 45:599–611.
- FUSI, S. AND WANG, X. J. 2016. Short-term, long-term, and working memory, pp. 319–344. In M. A. Arbib (ed.), *From Neuron to Cognition via Computational Neuroscience*, chapter 11. MIT Press, Cambridge/MA, USA.
- FYHN, M., MOLDEN, S., WITTER, M. P., MOSER, E. I., AND MOSER, M.-B. 2004. Spatial representation in the entorhinal cortex. *Science* 305:1258–1264.
- GALL, F. J. AND SPURZHEIM, J. G. 1810. *Anatomie et Physiologie du Système Nerveux en Général et du Cerveau en Particulier*, volume 1. F. Schoell, Paris, France.
- GALLIMORE, A. R., KIM, T., TANAKA-YAMAMOTO, K., AND DE SCHUTTER, E. 2018. Switching on depression and potentiation in the cerebellum. *Cell Reports* 22:722–733.
- GARDNER, E. 1988. The space of interactions in neural network models. *Journal of Physics A* 21:257.
- GARDNER-MEDWIN, A. R. 1989. Doubly modifiable synapses: a model of short and long term auto-associative memory. *Proceedings of the Royal Society B* 238:137–154.
- GASSER, H. S. AND ERLANGER, J. 1922. A study of the action currents of nerve with the cathode ray oscillograph. *American Journal of Physiology-Legacy Content* 62:496–524.
- GASTALDI, C., SCHWALGER, T., DE FALCO, E., QUIROGA, R. Q., AND GERSTNER, W. 2021. When shared concept cells support associations: theory of overlapping memory engrams. *Preprint at bioRxiv* 2021.03.12.434964.
- GELBARD-SAGIV, H., MUKAMEL, R., HAREL, M., MALACH, R., AND FRIED, I. 2008. Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* 322:96–101.
- GEORGOPOULOS, A. P., SCHWARTZ, A. B., AND KETTNER, R. E. 1986. Neuronal population coding of movement direction. *Science* 233:1416–1419.
- GERSTNER, W., KEMPTER, R., VAN HEMMEN, J. L., AND WAGNER, H. 1996. A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383:76–78.
- GERSTNER, W. AND KISTLER, W. M. 2002. *Spiking Neuron Models*. Cambridge University Press, Cambridge, UK, 1st edition.
- GERSTNER, W., KISTLER, W. M., NAUD, R., AND PANINSKI, L. 2014. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, Cambridge, UK, 1st edition.
- GERSTNER, W., LEHMANN, M., LIAKONI, V., CORNEIL, D., AND BREA, J. 2018. Eligibility traces and plasticity on behavioral time scales: experimental support of neoHebbian three-factor learning rules. *Frontiers in Neural Circuits* 12:53.
- GERSTNER, W. AND NAUD, R. 2009. How good are neuron models? *Science* 326:379–380.
- GERSTNER, W., RITZ, R., AND VAN HEMMEN, J. L. 1993. Why spikes? Hebbian learning and retrieval of time-resolved excitation patterns. *Biological Cybernetics* 69:503–515.
- GHOSH, V. E. AND GILBOA, A. 2014. What is a memory schema? a historical perspective on current neuroscience literature. *Neuropsychologia* 53:104–114.
- GILLESPIE, D. T. 1996. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Physical Review E* 54.
- GLANZER, M. AND CUNITZ, A. R. 1966. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior* 5:351–360.
- GLOCK, C., HEUMÜLLER, M., AND SCHUMAN, E. M. 2017. mRNA transport & local translation in neurons. *Current Opinion in Neurobiology* 45:169–177.
- GLUCK, M. A., MERCADO, E., AND MYERS, C. E. 2016. *Learning and Memory: From Brain to Behavior*. Worth Publishers, New York/NY, USA, 3rd edition.

- GLUCK, M. A. AND MYERS, C. E. 1993. Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus* 3:491–516.
- GOLDMAN-RAKIC, P. S. 1995. Cellular basis of working memory. *Neuron* 14:477–485.
- GOLDSTONE, R. L., KERSTEN, A., AND CARVALHO, P. F. 2012. Concepts and categorization. In I. B. Weiner (ed.), *Handbook of Psychology*, volume 4. Wiley Online Library, Hoboken/NJ, USA, 2nd edition.
- GOTO, Y., OTANI, S., AND GRACE, A. A. 2007. The Yin and Yang of dopamine release: a new perspective. *Neuropharmacology* 53:583–587.
- GRAUPNER, M. AND BRUNEL, N. 2010. Mechanisms of induction and maintenance of spike-timing dependent plasticity in biophysical synapse models. *Frontiers in Computational Neuroscience* 4:136.
- GRAUPNER, M. AND BRUNEL, N. 2012. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences of the USA* 109:3991–3996.
- GRAUPNER, M., WALLISCH, P., AND OSTOJIC, S. 2016. Natural firing patterns imply low sensitivity of synaptic plasticity to spike timing compared with firing rate. *Journal of Neuroscience* 36:11238–11258.
- GREENE, R. L. 1986. Sources of recency effects in free recall. *Psychological Bulletin* 99:221.
- GROGAN, J., BOGACZ, R., TSIVOS, D., WHONE, A., AND COULTHARD, E. 2015. Dopamine and consolidation of episodic memory: timing is everything. *Journal of Cognitive Neuroscience* 27:2035–2050.
- GROSSBERG, S. 1987. Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science* 11:23–63.
- GUPTA, A. AND LONG, L. N. 2007. Character recognition using spiking neural networks. In *IEEE International Joint Conference on Neural Networks*, pp. 53–58. IEEE.
- GÜTIG, R., AHARONOV, R., ROTTER, S., AND SOMPOLINSKY, H. 2003. Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *Journal of Neuroscience* 23:3697–3714.
- HANSEL, C., ARTOLA, A., AND SINGER, W. 1996. Different threshold levels of postsynaptic [Ca²⁺]_i have to be reached to induce LTP and LTD in neocortical pyramidal cells. *Journal of Physiology, Paris* 90:317–319.
- HARRIS, K. D. 2005. Neural signatures of cell assembly organization. *Nature Reviews Neuroscience* 6:399–407.
- HARTMANN, C., LAZAR, A., NESSLER, B., AND TRIESCH, J. 2015. Where’s the noise? Key features of spontaneous activity and neural variability arise through learning in a deterministic network. *PLOS Computational Biology* 11:e1004640.
- HASSELMO, M. E. 2006. The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology* 16:710–715.
- HASSELMO, M. E., SCHNELL, E., AND BARKAI, E. 1995. Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* 15:5249–5262.
- HÄUSSER, M., SPRUSTON, N., AND STUART, G. J. 2000. Diversity and dynamics of dendritic signaling. *Science* 290:739–744.
- HEBB, D. O. 1949. *The Organization of Behavior*. Wiley, New York/NY, USA, 1st edition.
- HELMHOLTZ, H. 1850. Messungen über den zeitlichen Verlauf der Zuckung animalischer Muskeln und die Fortpflanzungsgeschwindigkeit der Reizung in den Nerven. *Archiv für Anatomie, Physiologie und wissenschaftliche Medicin* pp. 276–364.
- HENLEY, J. M. AND WILKINSON, K. A. 2016. Synaptic AMPA receptor composition in development, plasticity and disease. *Nature Reviews Neuroscience* 17:337–350.
- HENNIG, M. H. 2013. Theoretical models of synaptic short term plasticity. *Frontiers in Computational Neuroscience* 7:45.
- HERCULANO-HOUZEL, S. 2009. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience* 3:31.

Bibliography

- HERMANN, L. 1899. Zur Theorie der Erregungsleitung und der elektrischen Erregung. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 75:574–590.
- HERPICH, J. AND TETZLAFF, C. 2019. Principles underlying the input-dependent formation and organization of memories. *Network Neuroscience* 3:606–634.
- HERRMANN, M., RUPPIN, E., AND USHER, M. 1993. A neural model of the dynamic activation of memory. *Biological Cybernetics* 68:455–463.
- HIGGINS, D., GRAUPNER, M., AND BRUNEL, N. 2014. Memory maintenance in synapses with calcium-based plasticity in the presence of background activity. *PLOS Computational Biology* 10:e1003834.
- HILGARD, E. R. AND MARQUIS, D. G. 1940. *Conditioning and Learning*. Appleton-Century, New York/NY, USA.
- HIRASE, H., CZURKÓ, A., CSICSVARI, J., AND BUZSÁKI, G. 1999. Firing rate and theta-phase coding by hippocampal pyramidal neurons during ‘space clamping’. *European Journal of Neuroscience* 11:4373–4380.
- HODGKIN, A. L. AND HUXLEY, A. F. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117:500–544.
- HOLMGREN, C., HARKANY, T., SVENNENFORS, B., AND ZILBERTER, Y. 2003. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of Physiology* 551:139–153.
- HOLTMAAT, A. AND CARONI, P. 2016. Functional and structural underpinnings of neuronal assembly formation in learning. *Nature Neuroscience* 19:1553.
- HOPFIELD, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA* 79:2554–2558.
- HOWARD, M. W. AND KAHANA, M. J. 1999. Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25:923.
- HOWARD, M. W. AND KAHANA, M. J. 2002. A distributed representation of temporal context. *Journal of Mathematical Psychology* 46:269–299.
- HUBEL, D. H. AND WIESEL, T. N. 1962. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of Physiology* 160:106–154.
- HULL, C. L. 1943. *Principles of Behavior*. Appleton-Century-Crofts, New York/NY, USA.
- HUNTER, W. S. 1930. A consideration of Lashley’s theory of the equipotentiality of cerebral action. *The Journal of General Psychology* 3:455–468.
- HUXTER, J., BURGESS, N., AND O’KEEFE, J. 2003. Independent rate and temporal coding in hippocampal pyramidal cells. *Nature* 425:828–832.
- IHRKE, M., BEHRENDT, J., SCHROBDSORFF, H., HERRMANN, J. M., AND HASSELHORN, M. 2011. Response-retrieval and negative priming. *Experimental Psychology* 58:154–161.
- IKEGAYA, Y., ISHIZAKA, Y., AND MATSUKI, N. 2002. BDNF attenuates hippocampal LTD via activation of phospholipase C: implications for a vertical shift in the frequency–response curve of synaptic plasticity. *European Journal of Neuroscience* 16:145–148.
- INSAUSTI, R., ANNESE, J., AMARAL, D. G., AND SQUIRE, L. R. 2013. Human amnesia and the medial temporal lobe illuminated by neuropsychological and neurohistological findings for patient E.P. *Proceedings of the National Academy of Sciences of the USA* 110:E1953–E1962.
- IZHIKEVICH, E. M. 2004. Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks* 15.
- IZHIKEVICH, E. M. 2007. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex* 17:2443–2452.

- JAHNKE, S., TIMME, M., AND MEMMESHEIMER, R.-M. 2015. A unified dynamic model for learning, replay, and sharp-wave/ripples. *Journal of Neuroscience* 35:16236–16258.
- JAMES, W. 1890. *The Principles of Psychology*. Henry Holt and Co., New York/NY, USA.
- JANISZEWSKI, C. AND WYER, JR., R. S. 2014. Content and process priming: A review. *Journal of Consumer Psychology* 24:96–118.
- JENKINS, J. G. AND DALLENBACH, K. M. 1924. Obliviscence during sleep and waking. *The American Journal of Psychology* 35:605–612.
- JEZEK, K., HENRIKSEN, E. J., TREVES, A., MOSER, E. I., AND MOSER, M.-B. 2011. Theta-paced flickering between place-cell maps in the hippocampus. *Nature* 478:246–249.
- JIANG, C. AND SCHUMAN, E. M. 2002. Regulation and function of local protein synthesis in neuronal dendrites. *Trends in Biochemical Sciences* 27:506–513.
- JOHANSEN, J. P., DIAZ-MATAIX, L., HAMANAKA, H., OZAWA, T., YCU, E., KOIVUMAA, J., KUMAR, A., HOU, M., DEISEROTH, K., BOYDEN, E. S., ET AL. 2014. Hebbian and neuromodulatory mechanisms interact to trigger associative memory formation. *Proceedings of the National Academy of Sciences of the USA* 111:E5584–E5592.
- JOLIVET, R., LEWIS, T. J., AND GERSTNER, W. 2004. Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy. *Journal of Neurophysiology* 92.
- KAHANA, M. J. 2006. The cognitive correlates of human brain oscillations. *Journal of Neuroscience* 26:1669–1672.
- KANDEL, E. R. 2001. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294:1030–1038.
- KARPICKE, J. D., LEHMAN, M., AND AUE, W. R. 2014. Retrieval-based learning: An episodic context account, pp. 237–284. In B. Ross (ed.), *Psychology of Learning and Motivation*, volume 61. Academic Press, New York/NY, USA.
- KASTELLAKIS, G. I. AND POIRAZI, P. 2019. Synaptic clustering and memory formation. *Frontiers in Molecular Neuroscience* 12:300.
- KATKOV, M., ROMANI, S., AND TSODYKS, M. 2014. Word length effect in free recall of randomly assembled word lists. *Frontiers in Computational Neuroscience* 8:129.
- KATKOV, M., ROMANI, S., AND TSODYKS, M. 2015. Effects of long-term representations on free recall of unrelated words. *Learning & Memory* 22:101–108.
- KESSLER, H., BLACKWELL, S. E., AND KEHYAYAN, A. 2017. Reconsolidation and trauma memory, pp. 369–379. In N. Axmacher and B. Rasch (eds.), *Cognitive Neuroscience of Memory Consolidation*. Springer, Cham, Switzerland.
- KIEFER, M. AND PULVERMÜLLER, F. 2012. Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex* 48:805–825.
- KIM, H. R., MALIK, A. N., MIKHAEL, J. G., BECH, P., TSUTSUI-KIMURA, I., SUN, F., ZHANG, Y., LI, Y., WATABE-UCHIDA, M., GERSHMAN, S. J., ET AL. 2020. A unified framework for dopamine signals across timescales. *Cell* 183:1600–1616.
- KLEIN, S. B. 2013. The complex act of projecting oneself into the future. *Wiley Interdisciplinary Reviews: Cognitive Science* 4:63–79.
- KNIGHT, B. W. 1972. Dynamics of encoding in a population of neurons. *Journal of General Physiology* 59.
- KOBAYASHI, R., TSUBO, Y., AND SHINOMOTO, S. 2009. Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Frontiers in Computational Neuroscience* 3.
- KONORSKI, J. 1948. Conditioned reflexes and neuron organization.
- KREISER, R., CARTIGLIA, M., MARTEL, J. N. P., CONRADT, J., AND SANDAMIRSKAYA, Y. 2018. A neuromorphic approach to path integration: a head-direction spiking neural network with vision-driven reset. In *IEEE Intern. Symp. Circuits Syst.*, pp. 1–5. IEEE.

Bibliography

- KROPFF, E. AND TREVES, A. 2006. The complexity of latching transitions in large scale cortical networks. *Natural Computing* 6:169–185.
- KURIKAWA, T. AND KANEKO, K. 2012. Associative memory model with spontaneous neural activity. *EPL (Europhys. Lett.)* 98:48002.
- LANDRY, C. D., KANDEL, E. R., AND RAJASETHUPATHY, P. 2013. New mechanisms in memory storage: piRNAs and epigenetics. *Trends in Neurosciences* 36:535–542.
- LANE, R. D., RYAN, L., NADEL, L., AND GREENBERG, L. 2015. Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science. *Behavioral and Brain Sciences* 38.
- LAPICQUE, L. 1907. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Générale* 9.
- LASHLEY, K. S. 1929. *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain*. University of Chicago Press, Chicago/IL, USA.
- LASHLEY, K. S. 1951. The problem of serial order in behavior, pp. 112–136. In L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior*, volume 21. Wiley, New York/NY, USA.
- LAZAR, A., PIPA, G., AND TRIESCH, J. 2009. SORN: a self-organizing recurrent neural network. *Frontiers in Computational Neuroscience* 3:23.
- LE DUGOU, C., SIMONNET, J., TELEŃCZUK, M., FRICKER, D., AND MILES, R. M. 2014. Recurrent synapses and circuits in the CA3 region of the hippocampus: an associative network. *Frontiers in Cellular Neuroscience* 7:262.
- LEE, W.-C. A., HUANG, H., FENG, G., SANES, J. R., BROWN, E. N., SO, P. T., AND NEDIVI, E. 2005. Dynamic remodeling of dendritic arbors in gabaergic interneurons of adult visual cortex. *PLOS Biology* 4.
- LEHR, A. B., LUBOEWSKI, J., AND TETZLAFF, C. 2021. Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks. *Preprint in this thesis*.
- LEPPERØD, M. E., CHRISTENSEN, A. C., LENSJØ, K. K., BUCCINO, A. P., YU, J., FYHN, M., AND HAFTING, T. 2021. Optogenetic pacing of medial septum parvalbumin-positive cells disrupts temporal but not spatial firing in grid cells. *Science Advances* 7:eabd5684.
- LERNER, I. AND SHRIKI, O. 2014. Internally-and externally-driven network transitions as a basis for automatic and strategic processes in semantic priming: theory and experimental validation. *Frontiers in Psychology* 5:314.
- LETZKUS, J. J., KAMPA, B. M., AND STUART, G. J. 2006. Learning rules for spike timing-dependent plasticity depend on dendritic synapse location. *Journal of Neuroscience* 26:10420–10429.
- LI, S., CULLEN, W. K., ANWYL, R., AND ROWAN, M. J. 2003. Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nature Neuroscience* 6:526–531.
- LI, W., MA, L., YANG, G., AND GAN, W.-B. 2017. REM sleep selectively prunes and maintains new synapses in development and learning. *Nature Neuroscience* 20:427–437.
- LI, Y., KULVICIUS, T., AND TETZLAFF, C. 2016. Induction and consolidation of calcium-based homo- and heterosynaptic potentiation and depression. *PLOS One* 11:e0161679.
- LIANG, Q., ZENG, Y., AND XU, B. 2020. Temporal-sequential learning with a brain-inspired spiking neural network and its application to musical memory. *Frontiers in Computational Neuroscience* 14:51.
- LILLICRAP, T. P., SANTORO, A., MARRIS, L., AKERMAN, C. J., AND HINTON, G. 2020. Backpropagation and the brain. *Nature Reviews Neuroscience* 21:335–346.
- LIN, C.-K., WILD, A., CHINYA, G. N., CAO, Y., DAVIES, M., LAVERY, D. M., AND WANG, H. 2018. Programming spiking neural networks on Intel’s Loihi. *Computer* 51:52–61.

- LIN, J.-W. AND FABER, D. S. 2002. Modulation of synaptic delay during synaptic plasticity. *Trends in Neurosciences* 25:449–455.
- LINDSKOG, M., KIM, M., WIKSTRÖM, M. A., BLACKWELL, K. T., AND KOTALESKI, J. H. 2006. Transient calcium and dopamine increase PKA activity and DARPP-32 phosphorylation. *PLOS Computational Biology* 2:e119.
- LISMAN, J., COOPER, K., SEHGAL, M., AND SILVA, A. J. 2018. Memory formation depends on both synapse-specific modifications of synaptic strength and cell-specific increases in excitability. *Nature Neuroscience* 21:309–314.
- LISMAN, J., GRACE, A. A., AND DUZEL, E. 2011. A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences* 34:536–547.
- LISMAN, J. E. AND GRACE, A. A. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46:703–713.
- LISMAN, J. E. AND ZHABOTINSKY, A. M. 2001. A model of synaptic memory: a CaMKII/PP1 switch that potentiates transmission by organizing an AMPA receptor anchoring assembly. *Neuron* 31:191–201.
- LITTLE, W. A. 1974. The existence of persistent states in the brain. *Mathematical Biosciences* 19:101–120.
- LIU, C., GOEL, P., AND KAESER, P. S. 2021. Spatial and temporal scales of dopamine transmission. *Nature Reviews Neuroscience* pp. 1–14.
- LIU, M.-G., CHEN, X.-F., HE, T., LI, Z., AND CHEN, J. 2012a. Use of multi-electrode array recordings in studies of network synaptic plasticity in both time and space. *Neuroscience Bulletin* 28:409–422.
- LIU, X., RAMIREZ, S., PANG, P. T., PURYEAR, C. B., GOVINDARAJAN, A., DEISSEROTH, K., AND TONEGAWA, S. 2012b. Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* 484:381–385.
- LOCKL, K. AND SCHNEIDER, W. 2007. Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development* 78:148–167.
- LOHNAS, L. J., POLYN, S. M., AND KAHANA, M. J. 2015. Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review* 122:337.
- LONG, N. M., DANOFF, M. S., AND KAHANA, M. J. 2015. Recall dynamics reveal the retrieval of emotional context. *Psychonomic Bulletin & Review* 22:1328–1333.
- LORENTE DE NÓ, R. 1934. Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *Journal für Psychologie und Neurologie* 46:113–177.
- LOUIS, S., BORGELT, C., AND GRÜN, S. 2010. Complexity distribution as a measure for assembly size and temporal precision. *Neural Networks* 23:705–712. Analysis and Modeling of Massively Parallel Neural Signals.
- LÖWEL, S. AND SINGER, W. 1992. Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. *Science* 255:209–212.
- LUBOEINSKI, J. 2021. Simulation code and analysis scripts for recurrent spiking neural networks with memory consolidation based on synaptic tagging and capture. <https://doi.org/10.5281/zenodo.4429195>.
- LUBOEINSKI, J. AND TCHUMATCHENKO, T. 2020. Nonlinear response characteristics of neural networks and single neurons undergoing optogenetic excitation. *Network Neuroscience* 4:852–870.
- LUBOEINSKI, J. AND TETZLAFF, C. 2021a. Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks. *Communications Biology* 4:275.
- LUBOEINSKI, J. AND TETZLAFF, C. 2021b. Organization and priming of long-term memory representations with two-phase plasticity. *Preprint at bioRxiv* 2021.04.15.439982.
- LYNCH, G., LARSON, J., KELSO, S., BARRIONUEVO, G., AND SCHOTTLER, F. 1983. Intracellular injections of EGTA block induction of hippocampal long-term potentiation. *Nature* 305:719–721.

Bibliography

- LYNCH, G. S., DUNWIDDIE, T., AND GRIBKOFF, V. 1977. Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. *Nature* 266:737–739.
- MACDONALD, C. J., CARROW, S., PLACE, R., AND EICHENBAUM, H. 2013. Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *The Journal of Neuroscience* 33:14607–14616.
- MACDONALD, C. J., LEPAGE, K. Q., EDEN, U. T., AND EICHENBAUM, H. 2011. Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron* 71:737–749.
- MAGEE, J. C. AND JOHNSTON, D. 1997. A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* 275:209–213.
- MAIA, T. V. 2009. Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective, & Behavioral Neuroscience* 9:343–364.
- MALENKA, R. C. AND NICOLL, R. A. 1999. Long-term potentiation—a decade of progress? *Science* 285:1870–1874.
- MANDLER, G. 2002. Origins of the cognitive (r)evolution. *Journal of the History of the Behavioral Sciences* 38:339–353.
- MARGRIE, T. W., BRECHT, M., AND SAKMANN, B. 2002. In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflügers Archiv – European Journal of Physiology* 444:491–498.
- MARKRAM, H., LÜBKE, J., FROTSCHER, M., ROTH, A., AND SAKMANN, B. 1997a. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *The Journal of Physiology* 500:409–440.
- MARKRAM, H., LÜBKE, J., FROTSCHER, M., AND SAKMANN, B. 1997b. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275:213–215.
- MARR, D. 1970. A theory for cerebral neocortex. *Proceedings of the Royal Society B* 176:161–234.
- MARTIN, K. C., CASADIO, A., ZHU, H., YAPING, E., ROSE, J. C., CHEN, M., BAILEY, C. H., AND KANDEL, E. R. 1997. Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell* 91:927–938.
- MARTIN, S. J., GRIMWOOD, P. D., AND MORRIS, R. G. M. 2000. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual Review of Neuroscience* 23:649–711.
- MASSE, N. Y., ROSEN, M. C., AND FREEDMAN, D. J. 2020. Reevaluating the role of persistent neural activity in short-term memory. *Trends in Cognitive Sciences* 24:242–258.
- MATHER, M., CLEWETT, D., SAKAKI, M., AND HARLEY, C. W. 2016. Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences* 39.
- MATSUZAKI, M., ELLIS-DAVIES, G. C. R., NEMOTO, T., MIYASHITA, Y., IINO, M., AND KASAI, H. 2001. Dendritic spine geometry is critical for AMPA receptor expression in hippocampal CA1 pyramidal neurons. *Nature Neuroscience* 4:1086–1092.
- MAYR, S. AND BUCHNER, A. 2007. Negative priming as a memory phenomenon. *Zeitschrift für Psychologie/Journal of Psychology* 215:35–51.
- MCCLELLAND, J. L., MCNAUGHTON, B. L., AND O'REILLY, R. C. 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* 102:419.
- MCCULLOCH, W. S. AND PITTS, W. 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* 5:115–133.
- MCGAUGH, J. 2000. Memory - a century of consolidation. *Science* 287:248–251.
- MCGEOCH, J. A. 1932. Forgetting and the law of disuse. *Psychological Review* 39:352.

- MCNAUGHTON, B. L. AND MORRIS, R. G. M. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences* 10:408–415.
- MEYER, D., BONHOEFFER, T., AND SCHEUSS, V. 2014. Balance and stability of synaptic structures during synaptic plasticity. *Neuron* 82:430–443.
- MICHAELIS, C., LEHR, A. B., AND TETZLAFF, C. 2020. Robust trajectory generation for robotic control on the neuromorphic research chip Loihi. *Frontiers in Neurobotics* 14:97.
- MILES, R. AND WONG, R. K. 1986. Excitatory synaptic interactions between CA3 neurones in the guinea-pig hippocampus. *The Journal of Physiology* 373:397–418.
- MILLER, G. A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63:81.
- MILLER, P. 2016. Itinerancy between attractor states in neural systems. *Current Opinion in Neurobiology* 40:14–22.
- MISANIN, J. R., MILLER, R. R., AND LEWIS, D. J. 1968. Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science* 160:554–555.
- MIYASHITA, Y. AND CHANG, H. S. 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70.
- MIZUSEKI, K. AND BUZSÁKI, G. 2013. Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. *Cell Reports* 4:1010–1021.
- MIZUSEKI, K. AND MIYAWAKI, H. 2017. Hippocampal information processing across sleep/wake cycles. *Neuroscience Research* 118:30–47.
- MONASSON, R. AND ROSAY, S. 2015. Transitions between spatial attractors in place-cell models. *Physical Review Letters* 115:098101.
- MONCADA, D., BALLARINI, F., AND VIOLA, H. 2015. Behavioral tagging: a translation of the synaptic tagging and capture hypothesis. *Neural Plasticity* 2015.
- MONCADA, D. AND VIOLA, H. 2007. Induction of long-term memory by exposure to novelty requires protein synthesis: evidence for a behavioral tagging. *Journal of Neuroscience* 27:7476–7481.
- MONGILLO, G., BARAK, O., AND TSODYKS, M. 2008. Synaptic theory of working memory. *Science* 319:1543–1546.
- MORENO-BOTE, R. AND PARGA, N. 2010. Response of integrate-and-fire neurons to noisy inputs filtered by synapses with arbitrary timescales: Firing rate and correlations. *Neural Computation* 22.
- MORRIS, R. G. M. 2006. Elements of a neurobiological theory of hippocampal function: the role of synaptic plasticity, synaptic tagging and schemas. *European Journal of Neuroscience* 23:2829–2846.
- MORRIS, R. G. M., ANDERSON, E., LYNCH, G. S., AND BAUDRY, M. 1986. Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* 319:774–776.
- MÜLLER, G. E. AND PILZECKER, A. 1900. Experimentelle Beiträge zur Lehre vom Gedächtniss [sic]. Verlag von J. A. Barth, Leipzig, Germany.
- MURDOCK, JR., B. B. 1962. The serial position effect of free recall. *Journal of Experimental Psychology* 64:482.
- NACHSTEDT, T. AND TETZLAFF, C. 2017. Working memory requires a combination of transient and attractor-dominated dynamics to process unreliably timed inputs. *Scientific Reports* 7:1–14.
- NADEL, L. AND HARDT, O. 2011. Update on memory systems and processes. *Neuropsychopharmacology* 36:251–273.
- NADEL, L. AND LAND, C. 2000. Memory traces revisited. *Nature Reviews Neuroscience* 1:209–212.

Bibliography

- NADEL, L. AND MOSCOVITCH, M. 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology* 7:217–227.
- NADER, K., SCHAFE, G. E., AND LEDOUX, J. E. 2000a. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406:722–726.
- NADER, K., SCHAFE, G. E., AND LEDOUX, J. E. 2000b. The labile nature of consolidation theory. *Nature Reviews Neuroscience* 1:216–219.
- NADH, K. AND HUYNH, C. R. 2010. A pong playing agent modelled with massively overlapping cell assemblies. *Neurocomputing* 73:2928–2934.
- NAKANO, T., DOI, T., YOSHIMOTO, J., AND DOYA, K. 2010. A kinetic model of dopamine-and calcium-dependent striatal synaptic plasticity. *PLoS Computational Biology* 6:e1000670.
- M. R. Nash and A. J. Barnier (eds.) 2008. *The Oxford Handbook of Hypnosis: Theory, Research, and Practice*. Oxford University Press, Oxford, UK.
- NAVAKKODE, S., SAJIKUMAR, S., AND FREY, J. U. 2007. Synergistic requirements for the induction of dopaminergic D1/D5-receptor-mediated LTP in hippocampal slices of rat CA1 in vitro. *Neuropharmacology* 52:1547–1554.
- NAVAKKODE, S., SAJIKUMAR, S., KORTE, M., AND SOONG, T. W. 2012. Dopamine induces LTP differentially in apical and basal dendrites through BDNF and voltage-dependent calcium channels. *Learning & Memory* 19:294–299.
- NAYA, Y., SAKAI, K., AND MIYASHITA, Y. 1996. Activity of primate inferotemporal neurons related to a sought target in pair-association task. *Proceedings of the National Academy of Sciences of the USA* 93:2664–2669.
- NEDUNGADI, P. 1990. Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of Consumer Research* 17:263–276.
- NEHLIG, A., ARMSPACH, J.-P., AND NAMER, I. J. 2010. Spect assessment of brain activation induced by caffeine: no effect on areas involved in dependence. *Dialogues in Clinical Neuroscience* 12:255.
- NEVIAN, T. AND SAKMANN, B. 2006. Spine Ca^{2+} signaling in spike-timing-dependent plasticity. *Journal of Neuroscience* 26:11001–11013.
- NGUYEN, P. V., ABEL, T., AND KANDEL, E. R. 1994. Requirement of a critical period of transcription for induction of a late phase of LTP. *Science* 265:1104–1107.
- OBERAUER, K. 2013. The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in Human Neuroscience* 7:673.
- O’CONNOR, D. H., WITTENBERG, G. M., AND WANG, S. S.-H. 2005. Dissection of bidirectional synaptic plasticity into saturable unidirectional processes. *Journal of Neurophysiology* 94:1565–1573.
- O’DONNELL, C. AND SEJNOWSKI, T. J. 2014. Selective memory generalization by spatial patterning of protein synthesis. *Neuron* 82:398–412.
- OJA, E. 1982. Simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology* 15:267–273.
- O’KEEFE, J. AND DOSTROVSKY, J. 1971. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* 34:171–175.
- OKUDA, K., HØJGAARD, K., PRIVITERA, L., BAYRAKTAR, G., AND TAKEUCHI, T. 2020. Initial memory consolidation and the synaptic tagging and capture hypothesis. *European Journal of Neuroscience* 00:1–24.
- O’REILLY, R., BRAVER, T. S., AND COHEN, J. D. 1999. A biologically-based computational model of working memory. In A. Miyake and P. Shah (eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, chapter 11. Cambridge University Press, New York/NY, USA.

- OSIPOVA, D., TAKASHIMA, A., OOSTENVELD, R., FERNÁNDEZ, G., MARIS, E., AND JENSEN, O. 2006. Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience* 26:7523–7531.
- OSTOJIC, S., BRUNEL, N., AND HAKIM, V. 2009. Synchronisation properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of Computational Neuroscience* 26.
- OTANI, H. AND HODGE, M. H. 1991. Does hypermnesia occur in recognition and cued recall? *The American Journal of Psychology* pp. 101–116.
- OTMAKHOVA, N. A. AND LISMAN, J. E. 1996. D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. *Journal of Neuroscience* 16:7478–7486.
- PACHECO, D. AND VERSCHURE, P. F. M. J. 2017. Long-term spatial clustering in free recall. *Memory* 26:798–806.
- PALACIOS-FILARDO, J. AND MELLOR, J. R. 2019. Neuromodulation of hippocampal long-term synaptic plasticity. *Current Opinion in Neurobiology* 54:37–43.
- PALM, G. 1981. Towards a theory of cell assemblies. *Biological Cybernetics* 39:181–194.
- PALM, G., KNOBLAUCH, A., HAUSER, F., AND SCHÜZ, A. 2014. Cell assemblies in the cerebral cortex. *Biological Cybernetics* 108:559–572.
- PÄPPER, M., KEMPTER, R., AND LEIBOLD, C. 2011. Synaptic tagging, evaluation of memories, and the distal reward problem. *Learning & Memory* 18:58–70.
- PASTALKOVA, E., ITS KOV, V., AMARASINGHAM, A., AND BUZSÁKI, G. 2008. Internally generated cell assembly sequences in the rat hippocampus. *Science* 321:1322–1327.
- PASTALKOVA, E., SERRANO, P., PINKHASOVA, D., WALLACE, E., FENTON, A. A., AND SACKTOR, T. C. 2006. Storage of spatial information by the maintenance mechanism of LTP. *Science* 313:1141–1144.
- PAVLOV, I. P. 1927. *Conditioned Reflexes: an Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press, London, UK. Translated and edited by G.V. Anrep.
- PAYNE, D. G. 1987. Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin* 101:5.
- PAYTON, M. E., GREENSTONE, M. H., AND SCHENKER, N. 2003. Overlapping confidence intervals or standard error intervals: what do they mean in terms of statistical significance? *Journal of Insect Science* 3.
- PEDRESCHI, N., BERNARD, C., CLAWSON, W., QUILICHINI, P., BARRAT, A., AND BATTAGLIA, D. 2020. Dynamic core periphery structure of information sharing networks in entorhinal cortex and hippocampus. *Network Neuroscience* 4:946–975.
- PEZZE, M. AND BAST, T. 2012. Dopaminergic modulation of hippocampus-dependent learning: blockade of hippocampal D1-class receptors during learning impairs 1-trial place memory at a 30-min retention delay. *Neuropharmacology* 63:710–718.
- PEISTER, J.-P. AND GERSTNER, W. 2006. Triplets of spikes in a model of spike timing-dependent plasticity. *Journal of Neuroscience* 26:9673–9682.
- PI, H.-J., HANGYA, B., KVITSIANI, D., SANDERS, J. I., HUANG, Z. J., AND KEPECS, A. 2013. Cortical interneurons that specialize in disinhibitory control. *Nature* 503:521–524.
- PIAGET, J. 1923. *The Origin of Intelligence in the Child*. Routledge & Kegan Paul, London, UK. Translated by M. Cook.
- PINHO, J., MARCUT, C., AND FONSECA, R. 2020. Actin remodeling, the synaptic tag and the maintenance of synaptic plasticity. *IUBMB Life* 72:577–589.
- PINSKER, H., KUPFERMANN, I., CASTELLUCCI, V., AND KANDEL, E. 1970. Habituation and dishabituation of the GM-withdrawal reflex in Aplysia. *Science* 167:1740–1742.
- PINSKER, H. M., HENING, W. A., CAREW, T. J., AND KANDEL, E. R. 1973. Long-term sensitization of a defensive withdrawal reflex in Aplysia. *Science* 182:1039–1042.

Bibliography

- PLENZ, D. AND THIAGARAJAN, T. C. 2007. The organizing principles of neuronal avalanches: cell assemblies in the cortex? *Trends in Neurosciences* 30:101–110.
- POLDRACK, R. A. AND PACKARD, M. G. 2003. Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41:245–251.
- POMI, A., MIZRAJI, E., AND LIN, J. 2018. Tensor representation of topographically organized semantic spaces. *Neural Computation* 30:3259–3280.
- PORTER, D. AND NEURINGER, A. 1984. Music discriminations by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 10:138.
- PRESTON, A. R. AND EICHENBAUM, H. 2013. Interplay of hippocampus and prefrontal cortex in memory. *Current Biology* 23:R764–R773.
- PRIESEMANN, V., WIBRAL, M., VALDERRAMA, M., PRÖPPER, R., LE VAN QUYEN, M., GEISEL, T., TRIESCH, J., NIKOLIĆ, D., AND MUNK, M. H. J. 2014. Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Frontiers in Systems Neuroscience* 8:108.
- PUN, R. Y., NEALE, E. A., GUTHRIE, P. B., AND NELSON, P. G. 1986. Active and inactive central synapses in cell culture. *Journal of Neurophysiology* 56:1242–1256.
- QUIROGA, R. Q., KREIMAN, G., KOCH, C., AND FRIED, I. 2008. Sparse but not ‘grandmother-cell’ coding in the medial temporal lobe. *Trends in Cognitive Sciences* 12:87–91.
- RAAIJMAKERS, J. G. 1990. The two-store model of memory: past criticisms, current status, and future directions. Technical report, TNO Institute for Perception, Soesterberg, Netherlands.
- RAAIJMAKERS, J. G. AND SHIFFRIN, R. M. 1981. Search of associative memory. *Psychological Review* 88:93.
- RAAIJMAKERS, J. G. W. AND SHIFFRIN, R. M. 1980. SAM: A theory of probabilistic search of associative memory, pp. 207–262. In G. H. Bower (ed.), *Psychology of Learning and Motivation*, volume 14. Academic Press, New York/NY, USA.
- RACINE, R. J., MILGRAM, N. W., AND HAFNER, S. 1983. Long-term potentiation phenomena in the rat limbic forebrain. *Brain Research* 260:217–231.
- RAMÓN Y CAJAL, S. 1894. The Croonian lecture – La fine structure des centres nerveux. *Proceedings of the Royal Society* 55:444–468.
- RANGANATH, C., COHEN, M. X., AND BROZINSKY, C. J. 2005. Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. *Journal of Cognitive Neuroscience* 17:994–1010.
- RASCH, B. AND BORN, J. 2013. About sleep’s role in memory. *Physiological Reviews* 93:681–766.
- RASHID, A. J., YAN, C., MERCALDO, V., HSIANG, H.-L. L., PARK, S., COLE, C. J., DE CRISTOFARO, A., YU, J., RAMAKRISHNAN, C., LEE, S. Y., ET AL. 2016. Competition between engrams influences fear memory formation and recall. *Science* 353:383–387.
- RECANATESI, S., KATKOV, M., ROMANI, S., AND TSODYKS, M. 2015. Neural network model of memory retrieval. *Frontiers in Computational Neuroscience* 9.
- RECANATESI, S., KATKOV, M., AND TSODYKS, M. 2017. Memory states and transitions between them in attractor neural networks. *Neural Computation* 29:2684–2711.
- REDONDO, R. L. AND MORRIS, R. G. M. 2011. Making memories last: the synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience* 12:17–30.
- REGHEH, W. G., CAREY, M. R., AND BEST, A. R. 2009. Activity-dependent regulation of synapses by retrograde messengers. *Neuron* 63:154–170.

- REN, X. AND COUTANCHE, M. N. 2020. Sleep reduces the semantic coherence of memory recall: An application of latent semantic analysis to investigate memory reconstruction. *Preprint at PsyArXiv*.
- REYMANN, K. G. AND FREY, J. U. 2007. The late maintenance of hippocampal LTP: requirements, phases, 'synaptic tagging', 'late-associativity' and implications. *Neuropharmacology* 52:24–40.
- RIBOT, T. 1882. Diseases of Memory: An Essay in the Positive Psychology, volume 41 of *The International Scientific Series*. D. Appleton, New York/NY, USA.
- RICCIARDI, L. M. 1977. Diffusion Processes and Related Topics in Biology. Springer, Berlin, Germany, 1st edition.
- RITCHEY, M., WING, E. A., LABAR, K. S., AND CABEZA, R. 2013. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cerebral Cortex* 23:2818–2828.
- ROACH, J. P., SANDER, L. M., AND ZOCHOWSKI, M. R. 2016. Memory recall and spike-frequency adaptation. *Physical Review E* 93:052307.
- ROEDIGER, III., H. L. 2008. Relativity of remembering: Why the laws of memory vanished. *Annual Review of Psychology* 59:225–254.
- ROEDIGER, III., H. L. AND BUTLER, A. C. 2011. The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences* 15:20–27.
- ROEDIGER, III., H. L., PAYNE, D. G., GILLESPIE, G. L., AND LEAN, D. S. 1982. Hypermnnesia as determined by level of recall. *Journal of Verbal Learning and Verbal Behavior* 21:635–655.
- ROELFSEMA, P. R. AND HOLTMAAT, A. 2018. Control of synaptic plasticity in deep cortical networks. *Nature Reviews Neuroscience* 19:166.
- ROLLS, E. T. 2016. Hippocampus: from spatial processing to episodic memory, pp. 533–558. In M. A. Arbib (ed.), *From Neuron to Cognition via Computational Neuroscience*, chapter 18. MIT Press, Cambridge/MA, USA.
- ROMANI, S., PINKOVIEZKY, I., RUBIN, A., AND TSODYKS, M. 2013. Scaling laws of associative memory retrieval. *Neural Computation* 25:2523–2544.
- ROSENZWEIG, M. R., BENNETT, E. L., COLOMBO, P. J., LEE, D. W., AND SERRANO, P. A. 1993. Short-term, intermediate-term, and long-term memories. *Behavioural Brain Research* 57:193–198.
- ROTH, A. AND VAN ROSSUM, M. C. W. 2009. Modeling synapses. In E. De Schutter (ed.), *Computational Modeling Methods for Neuroscientists*, chapter 6. MIT Press, Cambridge/MA, USA, 1st edition.
- ROWLAND, D. C., ROUDI, Y., MOSER, M.-B., AND MOSER, E. I. 2016. Ten years of grid cells. *Annual Review of Neuroscience* 39:19–40.
- ROYER, S. AND PARÉ, D. 2003. Conservation of total synaptic weight through balanced synaptic depression and potentiation. *Nature* 422:518–522.
- RUSSELL, L. E., YANG, Z., TAN, P. L., FISEK, M., PACKER, A. M., DALGLEISH, H. W., CHETTIH, S., HARVEY, C. D., AND HÄUSSER, M. 2019. The influence of visual cortex on perception is modulated by behavioural state. *Preprint at bioRxiv* 706010.
- RUSSO, E. AND DURSTEWITZ, D. 2017. Cell assemblies at multiple time scales with arbitrary lag constellations. *eLife* 6:e19428.
- RUSSO, E., NAMBOODIRI, V. M. K., TREVES, A., AND KROPFF, E. 2008. Free association transitions in models of cortical latching dynamics. *New Journal of Physics* 10:015008.
- RUSSO, E. AND TREVES, A. 2012. Cortical free-association dynamics: distinct phases of a latching network. *Physical Review E* 85:051920.
- SABATINI, B. L., OERTNER, T. G., AND SVOBODA, K. 2002. The life cycle of Ca²⁺ ions in dendritic spines. *Neuron* 33:439–452.

Bibliography

- SAJIKUMAR, S. AND FREY, J. U. 2004a. Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiology of Learning and Memory* 82:12–25.
- SAJIKUMAR, S. AND FREY, J. U. 2004b. Resetting of ‘synaptic tags’ is time- and activity-dependent in rat hippocampal CA1 in vitro. *Neuroscience* 129:503–507.
- SAJIKUMAR, S., NAVAKKODE, S., AND FREY, J. U. 2007. Identification of compartment- and process-specific molecules required for “synaptic tagging” during long-term potentiation and long-term depression in hippocampal CA1. *Journal of Neuroscience* 27:5068–5080.
- SAJIKUMAR, S., NAVAKKODE, S., SACKTOR, T. C., AND FREY, J. U. 2005. Synaptic tagging and cross-tagging: the role of protein kinase M ζ in maintaining long-term potentiation but not long-term depression. *Journal of Neuroscience* 25:5750–5756.
- SAKURAI, Y. 1999. How do cell assemblies encode information in the brain? *Neuroscience & Biobehavioral Reviews* 23:785–796.
- SANES, D. H., REH, T. A., AND HARRIS, W. A. 2006. Development of the Nervous System. Academic Press, Burlington/MA, USA, 2nd edition.
- SARA, S. J. 2000. Retrieval and reconsolidation: toward a neurobiology of remembering. *Learning & Memory* 7:73–84.
- SAYER, R. J., FRIEDLANDER, M. J., AND REDMAN, S. J. 1990. The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice. *Journal of Neuroscience* 10:826–836.
- SCHACTER, D. L. 1992. Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience* 4:244–256.
- SCHMIDGEN, H. 2011. Passagenwerk 1850. Bild und Zahl in den physiologischen Zeitexperimenten von Hermann von Helmholtz. *Berichte zur Wissenschaftsgeschichte* 34:139–155.
- SCHULTZ, W. 1998. Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* 80:1–27.
- SCHULTZ, W. 2002. Getting formal with dopamine and reward. *Neuron* 36:241–263.
- SCHULTZ, W. 2019. Recent advances in understanding the role of phasic dopamine activity. *F1000Research* 8.
- SCOVILLE, W. B. AND MILNER, B. 1957. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry* 20:11–21.
- SEDERBERG, P. B., MILLER, J. F., HOWARD, M. W., AND KAHANA, M. J. 2010. The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition* 38:689–699.
- SEIDENBECHER, T., REYMANN, K. G., AND BALSCHUN, D. 1997. A post-tetanic time window for the reinforcement of long-term potentiation by appetitive and aversive stimuli. *Proceedings of the National Academy of Sciences of the USA* 94:1494–1499.
- SEJNOWSKI, T. J. 1999. The book of Hebb. *Neuron* 24:773–776.
- SEMON, R. 1904. Die Mneme als erhaltendes Prinzip des organischen Geschehens. Verlag von Wilhelm Engelmann, Leipzig, Germany.
- SHAFI, M., ZHOU, Y., QUINTANA, J., CHOW, C., FUSTER, J., AND BODNER, M. 2007. Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* 146:1082–1108.
- SHANNON, C. E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27:379–423.
- SHATZ, C. J. 1992. The developing brain. *Scientific American* 267:60–67.
- SHENG, M. AND KIM, E. 2011. The postsynaptic organization of synapses. *Cold Spring Harbor Perspectives in Biology* 3.
- SHOUVAL, H. Z., BEAR, M. F., AND COOPER, L. N. 2002. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences of the USA* 99:10831–10836.

- SIEGEL, M., WARDEN, M. R., AND MILLER, E. K. 2009. Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences of the USA* 106:21341–21346.
- SJÖSTRÖM, P. J. AND HÄUSSER, M. 2006. A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron* 51:227–238.
- SJÖSTRÖM, P. J., TURRIGIANO, G. G., AND NELSON, S. B. 2001. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32:1149–1164.
- SKINNER, B. F. 1938. *The Behavior of Organisms: an Experimental Analysis*. Appleton-Century-Crofts, New York/NY, USA.
- SKINNER, B. F. 1971. *Beyond Freedom and Dignity*. Pelican Books, Harmondsworth, UK.
- SMOLEN, P., BAXTER, D. A., AND BYRNE, J. H. 2012. Molecular constraints on synaptic tagging and maintenance of long-term potentiation: a predictive model. *PLOS Computational Biology* 8:e1002620.
- SMOLEN, P., BAXTER, D. A., AND BYRNE, J. H. 2019. How can memories last for days, years, or a lifetime? proposed mechanisms for maintaining synaptic potentiation and memory. *Learning & Memory* 26:133–150.
- SMOLEN, P., ZHANG, Y., AND BYRNE, J. H. 2016. The right time to learn: mechanisms and optimization of spaced learning. *Nature Reviews Neuroscience* 17:77.
- SMOLEN, P. D., BAXTER, D. A., AND BYRNE, J. H. 2014. Modeling and analysis of intracellular signaling pathways, pp. 175–205. In J. H. Byrne, R. Heidelberger, and M. N. Waxham (eds.), *From Molecules to Networks*. Elsevier, London, UK, 3rd edition.
- SONG, S., MILLER, K. D., AND ABBOTT, L. F. 2000. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience* 3:919–926.
- SONI, D. 2018. Spiking neural networks, the next generation of machine learning. <https://towardsdatascience.com/spiking-neural-networks-the-next-generation-of-machine-learning-84e167f4eb2b>. Retrieved 26 March 2021.
- SONNEBORN, A. AND GREENE, R. W. 2021. Norepinephrine transporter antagonism prevents dopamine-dependent synaptic plasticity in the mouse dorsal hippocampus. *Neuroscience Letters* 740:135450.
- SOTO, D., MÄNTYLÄ, T., AND SILVANTO, J. 2011. Working memory without consciousness. *Current Biology* 21:R912–R913.
- SPREKELER, H. 2017. Functional consequences of inhibitory plasticity: homeostasis, the excitation-inhibition balance and beyond. *Current Opinion in Neurobiology* 43:198–203.
- SQUIRE, L. R. 2004. Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory* 82:171–177.
- SQUIRE, L. R., SLATER, P. C., AND MILLER, P. L. 1981. Retrograde amnesia and bilateral electroconvulsive therapy: Long-term follow-up. *Archives of General Psychiatry* 38:89–95.
- SQUIRE, L. R. AND ZOLA-MORGAN, S. 1991. The medial temporal lobe memory system. *Science* 253:1380–1386.
- STAMATAKIS, A. M., SCHACHTER, M. J., GULATI, S., ZITELLI, K. T., MALANOWSKI, S., TAJIK, A., FRITZ, C., TRULSON, M., AND OTTE, S. L. 2018. Simultaneous optogenetics and cellular resolution calcium imaging during active behavior using a miniaturized microscope. *Frontiers in Neuroscience* 12:496.
- STEIN, R. B. 1965. A theoretical analysis of neuronal variability. *Biophysical Journal* 5.
- STEVENS, C. F. AND WANG, Y. 1995. Facilitation and depression at single central synapses. *Neuron* 14:795–802.
- STÖBER, T. M., LEHR, A. B., HAFTING, T., KUMAR, A., AND FYHN, M. 2020. Selective neuromodulation and mutual inhibition within the CA3–CA2 system can prioritize sequences for replay. *Hippocampus* 30:1228–1238.
- STUART, G., SPRUSTON, N., SAKMANN, B., AND HÄUSSER, M. 1997. Action potential initiation and backpropagation in neurons of the mammalian CNS. *Trends in Neurosciences* 20:125–131.

Bibliography

- SUTTON, R. S. AND BARTO, A. G. 2018. Reinforcement Learning: An Introduction. MIT Press, Cambridge/MA, USA.
- TAKEUCHI, T., DUSZKIEWICZ, A. J., SONNEBORN, A., SPOONER, P. A., YAMASAKI, M., WATANABE, M., SMITH, C. C., FERNÁNDEZ, G., DEISSEROTH, K., GREENE, R. W., AND MORRIS, R. G. M. 2016. Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* 537:357–362.
- TCHUMATCHENKO, T., MALYSHEV, A., WOLF, F., AND VOLGUSHEV, M. 2011. Ultrafast population encoding by cortical neurons. *Journal of Neuroscience* 31:12171–12179.
- TERADA, S., SAKURAI, Y., NAKAHARA, H., AND FUJISAWA, S. 2017. Temporal and rate coding for discrete event sequences in the hippocampus. *Neuron* 94:1248–1262.
- TESCHE, C. D. AND KARHU, J. 2000. Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences of the USA* 97:919–924.
- TETZLAFF, C., DASGUPTA, S., KULVICIUS, T., AND WÖRGÖTTER, F. 2015. The use of Hebbian cell assemblies for nonlinear computation. *Scientific Reports* 5:12866.
- TETZLAFF, C., KOŁODZIEJSKI, C., MARKELIC, I., AND WÖRGÖTTER, F. 2012. Time scales of memory, learning, and plasticity. *Biological Cybernetics* 106:715–726.
- TETZLAFF, C., KOŁODZIEJSKI, C., TIMME, M., TSODYKS, M., AND WÖRGÖTTER, F. 2013. Synaptic scaling enables dynamically distinct short- and long-term memory formation. *PLOS Computational Biology* 9:e1003307.
- TETZLAFF, C., KOŁODZIEJSKI, C., TIMME, M., AND WÖRGÖTTER, F. 2011. Synaptic scaling in combination with many generic plasticity mechanisms stabilizes circuit connectivity. *Frontiers in Computational Neuroscience* 5.
- TETZLAFF, C., OKUJENI, S., EGERT, U., WÖRGÖTTER, F., AND BUTZ, M. 2010. Self-organized criticality in developing neuronal networks. *PLOS Computational Biology* 6:e1001013.
- THORNDIKE, E. L. 1911. Animal Intelligence: Experimental Studies. The Macmillan Company, New York/NY, USA.
- THORNDIKE, E. L. 1913. The Psychology of Learning, volume 2 of *Educational Psychology*. Teachers College, Columbia University, New York/NY, USA.
- TIPPER, S. P. 1985. The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology* 37:571–590.
- TONEGAWA, S., PIGNATELLI, M., ROY, D. S., AND RYAN, T. J. 2015. Memory engram storage and retrieval. *Current Opinion in Neurobiology* 35:101–109.
- TRAMONI, E., FELICIAN, O., BARBEAU, E. J., GUEJ, E., GUYE, M., BARTOLOMEI, F., AND CECCALDI, M. 2011. Long-term consolidation of declarative memory: insight from temporal lobe epilepsy. *Brain* 134:816–831.
- TRAN, A. H., UWANO, T., KIMURA, T., HORI, E., KATSUKI, M., NISHIJO, H., AND ONO, T. 2008. Dopamine D1 receptor modulates hippocampal representation plasticity to spatial novelty. *Journal of Neuroscience* 28:13390–13400.
- TREMBLAY, R., LEE, S., AND RUDY, B. 2016. GABAergic interneurons in the neocortex: From cellular properties to circuits. *Neuron* 91:260–292.
- TREVES, A. 2005. Frontal latching networks: a possible neural basis for infinite recursion. *Cognitive Neuropsychology* 22:276–291.
- TREVES, A. AND ROLLS, E. T. 1992. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2:189–199.
- TRIESCH, J., VO, A. D., AND HAFNER, A.-S. 2018. Competition for synaptic building blocks shapes synaptic plasticity. *eLife* 7:e37836.
- TRONSON, N. C. AND TAYLOR, J. R. 2007. Molecular mechanisms of memory reconsolidation. *Nature Reviews Neuroscience* 8:262–275.

- TRÜBUTSCHEK, D., MARTI, S., OJEDA, A., KING, J.-R., MI, Y., TSODYKS, M., AND DEHAENE, S. 2017. A theory of working memory without consciousness or sustained activity. *eLife* 6:e23871.
- TSE, D., LANGSTON, R. F., KAKEYAMA, M., BETHUS, I., SPOONER, P. A., WOOD, E. R., WITTER, M. P., AND MORRIS, R. G. M. 2007. Schemas and memory consolidation. *Science* 316:76–82.
- TSIEN, J. Z., LI, M., OSAN, R., CHEN, G., LIN, L., WANG, P. L., FREY, S., FREY, J., ZHU, D., LIU, T., ZHAO, F., AND KUANG, H. 2013. On initial brain activity mapping of episodic and semantic memory code in the hippocampus. *Neurobiology of Learning and Memory* 105:200–210.
- TSODYKS, M., PAWELZIK, K., AND MARKRAM, H. 1998. Neural networks with dynamic synapses. *Neural Computation* 10:821–835.
- TSODYKS, M. V. AND FEIGEL'MAN, M. V. 1988. The enhanced storage capacity in neural networks with low activity level. *EPL (Europhys. Lett.)* 6:101.
- TULVING, E. 1972. Episodic and semantic memory, pp. 381–403. In E. Tulving and W. Donaldson (eds.), *Organization of Memory*, chapter 10. Academic Press, New York/NY, USA.
- TULVING, E. 1985. Memory and consciousness. *Canadian Psychology/Psychologie Canadienne* 26:1.
- TULVING, E. AND THOMSON, D. 1973. Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80:352–373.
- TURRIGIANO, G. G., LESLIE, K. R., DESAI, N. S., RUTHERFORD, L. C., AND NELSON, S. B. 1998. Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391:892–896.
- TURRIGIANO, G. G. AND NELSON, S. B. 2000. Hebb and homeostasis in neuronal plasticity. *Current Opinion in Neurobiology* 10:358–364.
- VISKONTAS, I. V., EKSTROM, A. D., WILSON, C. L., AND FRIED, I. 2007. Characterizing interneuron and pyramidal cells in the human medial temporal lobe in vivo using extracellular recordings. *Hippocampus* 17:49–57.
- VOGEL, S., KLUEN, L. M., FERNÁNDEZ, G., AND SCHWABE, L. 2018. Stress leads to aberrant hippocampal involvement when processing schema-related information. *Learning & Memory* 25:21–30.
- VOGELS, T. P., SPREKELER, H., ZENKE, F., CLOPATH, C., AND GERSTNER, W. 2011. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* 334:1569–1573.
- VON DER MALSBERG, C. 1973. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik* 14:85–100.
- WALKER, M. P., BRAKEFIELD, T., HOBSON, J. A., AND STICKGOLD, R. 2003. Dissociable stages of human memory consolidation and reconsolidation. *Nature* 425:616–620.
- WALLNER, L. A. AND BÄUML, K.-H. T. 2018. Hypermnesia and the role of delay between study and test. *Memory & Cognition* 46:878–894.
- WANG, H.-X., GERKIN, R. C., NAUEN, D. W., AND BI, G.-Q. 2005. Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nature Neuroscience* 8:187–193.
- WANG, Q., LI, Y., SHAO, B., DEY, S., AND LI, P. 2017. Energy efficient parallel neuromorphic architectures with approximate arithmetic on FPGA. *Neurocomputing* 221:146–158.
- WANG, S.-H., REDONDO, R. L., AND MORRIS, R. G. M. 2010. Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proceedings of the National Academy of Sciences of the USA* 107:19537–19542.
- WANG, Y., ROMANI, S., LUSTIG, B., LEONARDO, A., AND PASTALKOVA, E. 2015. Theta sequences are essential for internally generated hippocampal firing fields. *Nature Neuroscience* 18:282–288.

Bibliography

- WAS, C., WOLTZ, D., AND HIRSCH, D. 2019. Memory processes underlying long-term semantic priming. *Memory & cognition* 47:313–325.
- WATANABE, S., SAKAMOTO, J., AND WAKITA, M. 1995. Pigeons' discrimination of paintings by monet and picasso. *Journal of the Experimental Analysis of Behavior* 63:165–174.
- WATSON, J. B. 1913. Psychology as the behaviorist views it. *Psychological Review* 20:158.
- WICKELGREN, W. A. 1965. Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior* 4:53–61.
- WILSON, H. R. AND COWAN, J. D. 1972. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal* 12:1–24.
- WILSON, H. R. AND COWAN, J. D. 1973. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13:55–80.
- WITTENBERG, G. M. AND WANG, S. S.-H. 2006. Malleability of spike-timing-dependent plasticity at the CA3–CA1 synapse. *Journal of Neuroscience* 26:6610–6617.
- WÖRGÖTTER, F. AND PORR, B. 2005. Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Computation* 17:245–319.
- YAGISHITA, S., HAYASHI-TAKAGI, A., ELLIS-DAVIES, G. C. R., URAKUBO, H., ISHII, S., AND KASAI, H. 2014. A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* 345:1616–1620.
- YAMAUCHI, S., KIM, H., AND SHINOMOTO, S. 2011. Elemental spiking neuron model for reproducing diverse firing patterns and predicting precise firing times. *Frontiers in Computational Neuroscience* 5.
- YANG, G., PAN, F., AND GAN, W.-B. 2009. Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462:920–924.
- YOKOSE, J., OKUBO-SUZUKI, R., NOMOTO, M., OHKAWA, N., NISHIZONO, H., SUZUKI, A., MATSUO, M., TSUJIMURA, S., TAKAHASHI, Y., NAGASE, M., WATABE, A. M., MASAKIYO, S., FUSAO, K., AND INOKUCHI, K. 2017. Overlapping memory trace indispensable for linking, but not recalling, individual memories. *Science* 355:398–403.
- ZENKE, F., AGNES, E. J., AND GERSTNER, W. 2015. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications* 6:1–13.
- ZENKE, F. AND GERSTNER, W. 2017. Hebbian plasticity requires compensatory processes on multiple timescales. *Philosophical Transactions of the Royal Society B* 372:20160259.
- ZHANG, J.-C., LAU, P.-M., AND BI, G.-Q. 2009. Gain in sensitivity and loss in temporal contrast of stdp by dopaminergic modulation at hippocampal synapses. *Proceedings of the National Academy of Sciences of the USA* 106:13028–13033.
- ZHONG, F., LIU, L., WEI, J.-L., AND DAI, R.-P. 2019. Step by step golgi-cox staining for cryosection. *Frontiers in Neuroanatomy* 13:62.
- ZIEGLER, L., ZENKE, F., KASTNER, D. B., AND GERSTNER, W. 2015. Synaptic consolidation: from synapses to behavioral modeling. *Journal of Neuroscience* 35:1319–1334.

Acknowledgments

First of all, I would like to thank my supervisor Christian Tetzlaff for providing me with exciting scientific ideas and for very helpful advice on the issues of this thesis. I never had the feeling that I needed more support from him. Besides Christian, I would like to thank the further members of my Thesis Advisory Committee, Stefan Klumpp and Silvio Rizzoli, for providing me with great feedback and motivating me in pursuing my research questions.

Moreover, I would like to express my appreciation to the members of the Tetzlaff group (the “Tetzlab”) and to the members of the Department of Computational Neuroscience, which is led by Florentin “Flo” Wörgötter. I very much enjoyed the nice and friendly environment of our group as well as the good discussions and scientific comments. I have to thank Flo for employment, for good scientific input, and for keeping up a great atmosphere. Special thanks also go to Juliane Herpich, Timo Nachstedt, Michael Fauth, Moritz Becker, Yinyun Li, Mayte Bonilla Quintana, Tomas Kulvicius, Carlo Michaelis, David Kappel, Andrew Lehr, Timo Lüddecke, Simon Reich, Sebastian Schmitt, and Johannes Auth.

Furthermore, I would like to thank Ursula “Ursel” Hahn-Wörgötter, Thomas Geiling, and the other staff of the III. Institute of Physics, as well as Antje Erdmann and Frauke Bergmann from the IMPRS-PBCS office, for always being available for any kind of issue. I would also like to thank the organizers, tutors, and participants of the OIST Computational Neuroscience Course 2019 for making it such a great experience.

For reading and commenting on parts of this thesis, I would like to thank Ruben Luboeinski, Peter Kuhn, Marius Schneider, and Andrew Lehr.

Last but not least I want to express my gratitude to my family (especially to my parents, Antje and Michael), to my friends, and to my partner Sonja for supporting me in all kinds of situations and for bringing so much happiness into my life.

Curriculum Vitae

Personal Information

Date of birth: 30th October 1990
Place of birth: Wetzlar, Germany
E-mail: mail@jlubo.net

Education

- 2017 – 2021 Doctorate, University of Göttingen
Program: International Max Planck Research School for Physics of Biological and Complex Systems
Thesis: “The role of synaptic tagging and capture for memory dynamics in spiking neural networks” (supervisor: Christian Tetzlaff)
- 2014 – 2016 Master of Science, Goethe University, Frankfurt/Main
Major subject: Physics
Minor subject: Philosophy
Thesis: “Modeling of a neural network with integrated light-sensitive channels” (at Max Planck Institute for Brain Research, Frankfurt/Main; supervisor: Tatjana Tchumatchenko)
- 2010 – 2013 Bachelor of Science, Technical University of Darmstadt
Major subject: Physics
Minor subject: Computer Science
Thesis: “Entropy study in neuronal networks with chemical and electrical synapses” (at Max Planck Institute for Brain Research, Frankfurt/Main; supervisor: Tatjana Tchumatchenko)
- 2001 – 2010 Abitur (university entrance diploma), Weidigschule, Butzbach
Focus subjects: Physics, Chemistry
- 1997 – 2001 Primary education, Mittelpunktschule “Oberer Hüttenberg”, Butzbach