The role of inhibitory plasticity in the formation and the long-term maintenance of neural assemblies and memories

Raphaël Bergoin

PhD Thesis

to obtain the title of

DOCTOR in Information and Communication Sciences and Technologies

delivered by

CY Cergy Paris University - Economy, Management, Mathematics, Physics and Computer Sciences Doctoral School

Pompeu Fabra University - PhD programme in Information and Communications Technologies

The role of inhibitory plasticity in the formation and the long-term maintenance of neural assemblies and memories

by

Raphaël Bergoin

prepared at the ETIS laboratory, NEURO team and
at the Center for Brain and Cognition, Computational Neuroscience group

defended on December 11, 2023

Thesis directors: Mathias Quoy and Gustavo Deco
Thesis supervisors: Alessandro Torcini and Gorka Zamora-López

Jury:

Reviewers:	Yulia Timofeeva	-	University of Warwick
	Aneta Koseska	-	Max Planck Institute
President:	Nicolas Rougier	-	Inria Bordeaux
Examiners:	Mathias Quoy	-	CY Cergy Paris University
	Gorka Zamora-López	-	Pompeu Fabra University
Invited:	Alessandro Torcini	-	CY Cergy Paris University
	Lou Zonca	-	Pompeu Fabra University
	Simona Olmi	-	Istituto dei Sistemi Complessi, CNR

Acknowledgements

First of all, I would like to thank my thesis director Mathias Quoy, who originally offered this thesis position to me at the end of my internship at the ETIS laboratory and so without whom this thesis would not have been possible. Your advices in terms of organisation and strategies has been a great help to me over these 3 years.

Then, my second thesis director in Barcelona, Gustavo Deco, although our interactions were not that frequent, all the ideas and remarks you suggested to me were extremely pertinent and gave me real openings in my ideas, and for that I thank you.

As my thesis supervisor in Cergy, Alessandro Torcini provided me a real sense of rigour as well as a huge amount of theoretical knowledge, whether from a neuroscience, mathematical or physical point of view, filling in some of the lacunae I had in these domains. I would also like to thank you for all lectures and seminars shared and for your invaluable help in writing the different papers.

Finally, the person who welcomed me from the very first day of my PhD, Gorka Zamora-López, you have been more than just a thesis supervisor during these years. We have spent a lot of time together both inside and outside the lab and I have enjoyed our discussions in various fields. We shared a lot of ideas that were real scientific inspirations for me during this thesis. I sincerely believe that you gave me the motivation to continue and, at end, the taste for scientific research.

As final world, I would like to thank you all four of you for your different contributions, I think you complemented each other perfectly in the supervision of this thesis.

I would also like to thank jury members Nicolas Rougier, Lou Zonca and Simona Olmi, as well as the reviewers of this thesis Yulia Timofeeva and Aneta Koseska who agreed to evaluate my work.

I want to thank all my colleagues from the ETIS laboratory and the Center for Brain and Cognition for these scientifically enriching moments during discussions and other group presentations. I won't forget either the convivial moments spent together in and outside the lab, notably in restaurants, bars and asados.

I thank the CY Cergy Paris University and the Pompeu Fabra University for supporting this PhD thesis. This work was supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific [Grant Agreement No. 945539 (Human Brain Project SGA3)] and by an EUTOPIA funding [EUTOPIA-PhD-2020-0000000066 - NEUROAI].

I'd also like to thank my friends from university, high school, college and even before. Distance and work have kept us apart for a while, but I won't forget the times we spent together that made me who I am today.

Finally, and most importantly, I would like to thank my family, and in particular my close family: Fabien, Elisabeth, Elsa and Jasper. You have always supported me in everything I do, and that's what has kept me going despite all the difficulties life throws at us, merci pour tout.

To Antonio.

Abstract

Brain circuits display modular architecture at different scales of organization. Such neural assemblies are typically associated to functional specialization that favours both the segregation and the integration of information. However, the mechanisms leading to their emergence and consolidation remain elusive.

This PhD thesis aims to understand the formation of modular structures in artificial neural networks and the mechanisms that sustain these memory structures over time while allowing continuous adaptation. In addition, the thesis seeks to validate the architecture and mechanisms using real sensory information and to evaluate their effective integration into cognitive tasks.

In the first chapter, we review the state of the art regarding the formation of multiple clusters in networks of coupled oscillators, focusing specifically on Kuramoto oscillators subjected to adaptation. This chapter also highlights the formation of structural clusters via external stimulation and discusses two approaches: frequency-based and synchrony-based, paving the way for the experiments conducted in the following chapters.

In the second chapter, we focus on the formation and consolidation of modular structures induced by external stimuli in networks of theta-neurons. The results show that inhibitory neurons play a crucial role in the maintenance of these modular architectures. Networks containing both excitatory and inhibitory neurons are able to maintain and consolidate learned memories by avoiding total synchronisation of the network, while networks with only excitatory neurons or networks that do not differentiate between excitatory and inhibitory neurons fail to do so. We also show that the number of inhibitory neurons in the network determines its memory capacity.

In the third chapter, we study the phenomenon of spontaneous memory recall in an asynchronous irregular state and its role in long-term memory consolidation. We consider an excitatory-inhibitory spiking neural network subjected to spike-timing-dependent plasticity. We show that the presence of two groups of inhibitory neurons – one subjected to Hebbian-STDP and one subjected to anti-Hebbian-STDP – is necessary to guarantee the emergence of the modular structures and their spontaneous recall at rest. We prove that these recalls are correlated with a consolidation of the structural modules. In addition, a relationship between the number of inhibitory neurons and the storage capacity is again established.

Finally in the last chapter, we apply this architecture to the learning of audio-visual information. This confirms the possibility of forming and maintaining complex structures over the long-term, but this time with real sensory stimuli. Learning each modality independently results in stable segregated structures, enabling accurate unimodal recognition. The integration of modalities is achieved via hub neurons, facilitating coherent and more efficient processing of multisensory information in recognition and generation tasks.

In summary, this PhD thesis contributes to a better understanding of the impact of inhibition on network dynamics, allowing sustainable memory learning. In addition, these

works highlight the role of this same inhibition in the storage of memories and in their integration and processing in cognitive tasks. In this way, this thesis also provides insights for more bio-realistic artificial intelligence systems, while contributing to a better understanding of neural mechanisms.

Keywords: artificial neural network, spiking neural network, coupled oscillator, learning, adaptation, plasticity, inhibition, synchronization, long-term memory, memory consolidation, modular structure, spontaneous recall, multimodality, recognition, generation

Résumé

Les circuits cérébraux présentent une architecture modulaire à différentes échelles d'organisation. Ces assemblages neuronaux sont généralement associés à une spécialisation fonctionnelle qui favorise à la fois la ségrégation et l'intégration des informations. Cependant, les mécanismes qui conduisent à leur émergence et à leur consolidation restent évasifs.

Cette thèse de doctorat vise à comprendre la formation de structures modulaires dans les réseaux neuronaux artificiels et les mécanismes qui soutiennent ces structures de mémoire dans le temps tout en permettant une adaptation continue. En outre, la thèse cherche à valider l'architecture et les mécanismes en utilisant des informations sensorielles réelles et à évaluer leur intégration efficace dans des tâches cognitives.

Dans le premier chapitre, nous examinons l'état de l'art concernant la formation de modules multiples dans des réseaux d'oscillateurs couplés, en nous concentrant spécifiquement sur les oscillateurs de Kuramoto soumis à l'adaptation. Ce chapitre met également en évidence la formation de modules structurels par stimulation externe et examine deux approches : celle basée sur la fréquence et celle basée sur la synchronie, ce qui ouvre la voie aux expériences menées dans les chapitres suivants.

Dans le deuxième chapitre, nous nous concentrons sur la formation et la consolidation de structures modulaires induites par des stimuli externes dans des réseaux de neurones theta. Les résultats montrent que les neurones inhibiteurs jouent un rôle crucial dans le maintien de ces architectures modulaires. Les réseaux contenant à la fois des neurones excitateurs et inhibiteurs sont capables de maintenir et de consolider les souvenirs appris en évitant une synchronisation totale du réseau, alors que les réseaux contenant uniquement des neurones excitateurs ou les réseaux qui ne font pas la différence entre les neurones excitateurs et inhibiteurs n'y parviennent pas. Nous montrons également que le nombre de neurones inhibiteurs dans le réseau détermine sa capacité de mémoire.

Dans le troisième chapitre, nous étudions le phénomène de rappel spontané de la mémoire dans un état irrégulier asynchrone et son rôle dans la consolidation de la mémoire à long terme. Nous considérons un réseau de neurones à impulsions excitateurs-inhibiteurs soumis à une plasticité dépendante du temps d'occurrence des impulsions. Nous montrons que la présence de deux groupes de neurones inhibiteurs - l'un soumis à une STDP Hebbienne et l'autre à une STDP anti-Hebbienne - est nécessaire pour garantir l'émergence des structures modulaires et leur rappel spontané au repos. Nous prouvons que ces rappels sont corrélés à une consolidation des modules structurels. De plus, une relation entre le nombre de neurones inhibiteurs et la capacité de stockage est à nouveau établie.

Enfin, dans le dernier chapitre, nous appliquons cette architecture à l'apprentissage d'informations audiovisuelles. Cela confirme la possibilité de former et de maintenir des structures complexes à long terme, mais cette fois-ci avec des stimuli sensoriels réels. L'apprentissage indépendant de chaque modalité aboutit à des structures séparées stables, permettant une reconnaissance unimodale efficace. L'intégration des modalités se

fait par l'intermédiaire de neurones hubs, ce qui facilite le traitement cohérent et efficace des informations multisensorielles dans des tâches de reconnaissance et de génération.

En résumé, cette thèse de doctorat contribue à une meilleure compréhension de l'impact de l'inhibition sur la dynamique des réseaux, permettant un apprentissage durable de la mémoire. De plus, ces travaux mettent en évidence le rôle de cette même inhibition dans le stockage des mémoires et dans leur intégration et traitement dans des tâches cognitives. Ainsi, cette thèse fournit des pistes pour des systèmes d'intelligence artificielle plus bio-réalistes, tout en contribuant à une meilleure compréhension des mécanismes neuronaux.

Mots-clés: réseau de neurones artificiels, réseau de neurones à impulsions, oscillateur couplé, apprentissage, adaptation, plasticité, inhibition, synchronisation, mémoire à long terme, consolidation de la mémoire, structure modulaire, rappel spontané, multimodalité, reconnaissance, génération

Resúmen

Los circuitos cerebrales muestran una arquitectura modular a diferentes escalas de organización. La presencia de tales conjuntos neuronales suele asociarse a una organizacion funcional que favorece tanto la segregación como la integración de la información. Sin embargo, los mecanismos que conducen al surgimiento y consolidación de esta organización modular siguen siendo desconocidos.

Esta tésis doctoral trata de comprender la formación de estructuras modulares en redes neuronales artificiales y los mecanismos que sostienen estas estructuras de memoria a lo largo del tiempo, permitiendo así su contínua adaptación. Además, esta tésis pretende validar el surgimiento de tales arquitecturas modulares utilizando para ello información sensorial real, y evaluar su integración efectiva en tareas cognitivas.

El primer capítulo resume el conocimiento actual en el campo con respecto a la formación de architecturas modulares, siguiendo redes de osciladores acoplados. En este caso, nos centramos en redes de osciladores de Kuramoto acoplados sometidos a adaptación. También se destaca la formación de módulos estructurales debido a la aplicación de estimulos externos y se analizan dos enfoques distintos para ello: el basado en la frecuencia y el basado en la sincronía. Estos ejemplos sirven como punto de partida para los experimentos realizados en los capítulos siguientes.

El segundo capítulo se centra en la formación y la consolidación de arquitecturas modulares, inducidas por la presencia de estímulos externos en redes de "neuronas theta". Los resultados muestran que las neuronas inhibitorias son fundamentales para el mantenimiento de estas arquitecturas modulares. Las redes que contienen neuronas excitatorias e inhibitorias son capaces de mantener y consolidar las memorias aprendidas, evitando la sincronización total de la red. Sin embargo, las memories no pueden ser consolidadas en aquellas redes compuestas únicamente de neuronas excitatorias, o en aquellas en las que no se evita diferenciar entre neuronas excitatorias e inhibitorias. También demostramos que la capacidad de memoria de las redes viene determinado por el número de neuronas inhibitorias.

En el tercer capítulo, estudiamos la aparición de recuerdos espontáneos en las redes neuronales, mientras éstas se encuentran en un estado basal caracterizado por la actividad irregular y asíncrona de sus neuronas. También estudiamos el papel de los recuerdos espontáneos en la consolidación de las memorias a largo plazo. Para ello, consideramos una red compuesta por neuronas excitatorias e inhibitorias sometidas a una plasticidad dependiente del tiempo de los picos de actividad neuronal. Demostramos que la presencia de dos grupos de neuronas inhibitorias —uno sometido a Hebbian-STDP y el otro sometido a anti-Hebbian-STDP— es necesaria para garantizar la formación espontánea de las estructuras modulares y la presencia de recuerdos espontáneos en la actividad neuronal. Demostramos también que estos recuerdos espontáneos son necesarios para la consolidación de la architectura modular. Además, establecemos de nuevo una relación entre el número de neuronas inhibitorias y la capacidad de almacenamiento de la red.

Finalmente, en el último capítulo, utilizamos este tipo de arquitecturas para el aprendizaje de información audiovisual. Con ello confirmamos también la formación y mantenimiento de estructuras complejas a largo plazo, pero esta vez empleando estímulos sensoriales reales. El aprendizaje de cada modalidad sensorial de manera independiente da lugar a estructuras segregadas pero estables, lo que a posteriori permite un reconocimiento unimodal preciso. Sin embargo, la integración de las dos modalidades sensoriales se consigue a través de "hubs" neuronales. La presencia de tales neuronas facilita un procesamiento coherente y más eficiente de la información multisensorial en tareas de reconocimiento y generación.

En resúmen, esta tésis doctoral contribuye a la comprensión del papel que juega la inhibición en la dinámica de redes neuronales, permitiendo un aprendizaje sostenible de la memoria. Además, destacamos la importancia de esta inhibición en el almacenamiento de los recuerdos, y en su integración y posterior procesamiento durante tareas cognitivas. De este modo, esta tésis también aporta nuevas ideas para el desarrollo de sistemas de inteligencia artificial con mayor realismo biológico, al tiempo que contribuye a una mayor comprensión de los mecanismos neuronales relacionados con el aprendizaje y la formación de estructuras neuronales.

Palabras clave: red neuronal artificial, red neuronal de espigas, osciladores acoplados, aprendizaje, adaptación, plasticidad, inhibición, sincronización, memoria a largo plazo, consolidación de la memoria, estructura modular, recuerdos espontáneos, multimodalidad sensorial, reconocimiento, generación.

List of Publications

In preparation:

- Matching and mismatching bimodal recognition: a psychological study.
- Recognition and generation of MNIST dataset with spiking neurons continuously adapting: a machine learning study.
- Technical study of the memory capacity of QIF neural networks with two types of inhibition.

Under submission or review:

- R. Bergoin, A. Torcini, G. Deco, M. Quoy and G. Zamora-Lopez. Maintenance and consolidation of long-term memory in asynchronous irregular state through spontaneous recalls mediated by inhibition. *PLoS computational biology*.
- R. Bergoin, S. Boucenna, D. Cohen and A. Pitti. A developmental model of audio-visual attention (MAVA): A premise for bimodal language learning in infants and robots.
- S. Boucenna and R. Bergoin. Bimodal recognition of vowel facilitated by an attention mechanism.

Published:

- R. Bergoin, A. Torcini, G. Deco, M. Quoy and G. Zamora-Lopez. Inhibitory neurons control the consolidation of neural assemblies via adaptation to selective stimuli. *Scientific Reports*, 13(1):6949, 2023.

Contents

List of Tables

List of Figures

List of Acronyms

AI Artificial intelligence . . . 1

ANN Artificial neural networks . . . 1

EEG electroencephalogram . . . 11

STDP Spike-timing-dependent plasticity . . . 1

GWT Global Workspace Theory . . . 2

IIT Integrated Information Theory . . . 3

PDDP Phase-difference-dependent plasticity . . . 11

LTM Long-term memory . . . 53

SNN Spiking neural network . . . 55

QIF quadratic integrate and fire . . . xiii

LTP long-term potentiation . . . 34

LTD long-term depression . . . 34

ISIs interspike intervals . . . 61

E-E excitatory to excitatory . . . 64

E-I excitatory to inhibitory . . . 64

I-E inhibitory to excitatory . . . 66

I-I inhibitory to inhibitory . . . 64

PDF probability density function . . . 60

A-I asynchronous irregular . . . 54

MNIST Mixed National Institute of Standards and Technology . . . 84

FSDD Free Spoken Digit Dataset . . . 84

Introduction

AI and neuroscience

Artificial intelligence (AI) and neuroscience are two fields interconnected by various aspects. On the one hand, computational neuroscience is increasingly incorporating machine learning techniques, including deep neural networks, in response to the increasing amount of precise data to be analysed on brain dynamics and connectivity. On the other hand, current AI systems employing Artificial neural networks (ANN) draw inspiration from some principles derived from neuroscience. In particular, their learning is based on the notion of Hebbian rule, which states that "*cells that fire together, wire together*" [174], which attends to explain the synaptic plasticity of brain neurons during the adaptation process. Also, deep neural networks are based on the principle of convolutional neural network which follows a certain hierarchical organization found in mammals [224]. However, they rely on the back-propagation mechanism which, although having proven its ability to learn complex data effectively [239], deviates from the synaptic weight adaptation observed in biology. In fact, this method relies on non-local information through propagation of error from upper layers. In contrast, biological networks involve local dynamics via the mechanism of Spike-timing-dependent plasticity (STDP), where the time difference between neuron spikes influences the way synaptic connections between neurons strengthen or weaken [46]. In addition to the challenges of efficiently incorporating local plasticity rules into ANN [32], another difference remains. Indeed, the increasingly deep layered architecture of current networks poses a number of optimisation problems that call this type of organisation into question. Indeed, the latest research shows that the brain instead follows a modular connectivity that facilitates both the segregation of information (i.e. the specialisation of brain regions on specific tasks or modalities) and the coherent integration of these information [334, 353, 419, 422].

Despite these differences, the popularization of AI systems such as ChatGPT [289] or Bard [6] capable of efficiently responding to almost any query and even surpassing human in certain cases, raises question on whether these systems are really "intelligent" or just sophisticated machine learning algorithms. Indeed, the latter can also now exhibit a certain level of creativity, as shown by projects like DALL·E [288] or Midjourney [267], creativity that is generally considered to be correlated with intelligence. However, beneath their impressive achievements, these systems remain large neural networks trained on a vast amount of data, far exceeding what an individual can encounter or learn in a lifetime. This makes them extremely efficient for a given task but for a significant computational cost. Moreover, they remain a kind of input/output boxes providing singular answers to specific prompts. In this regard, these models bear a closer resemblance to inert networks without real state of consciousness or thoughts, in contrast to living systems characterized by variable and persistent activity. The human and animal brain, on the other hand, in general is capable of processing and learning information at the same time by being

confronted with it only a few times and then retaining it over a long period of time while remaining active and in constant adaptation. Even if it is difficult to establish a precise definition of "intelligence", this one cannot exist without the autonomous, complex, and coordinated activity of neurons in the brain. These aspects constitute the initial motivation of this thesis to establish a system able of efficiently processing sensory information while coming as close as possible to the biological mechanisms observed in the human brain and, to a lesser extent, to its state of consciousness.

The role of consciousness in learning and integration of information

The brain is a massively parallel system where each specialized area operates relatively independently. These routine execution tasks are most often performed unconsciously. However, in order to overcome unpredictable conditions and process complex information, consciousness seems to be the primary means by which the nervous system relies [10, 21]. In the context of learning, consciousness allows us to direct our attention, make choices, and focus on specific aspects of our environment or mental representations. More precisely, it helps us to focus and select relevant stimuli or information, filtering distractions, which is crucial for effective learning [112, 230]. Although consciousness is a concept with a number of definitions, it is often divided into two main components. Firstly, arousal, which is comparable to wakefulness or vigilance, and secondly, awareness, in the sense of being aware of the environment and self [92, 238]. Plenty of experiments on individuals with brain damage have aimed to distinguish between the different conscious and unconscious states, shedding light on their implications [114, 238, 237]. In particular, it has been shown that the process of consciousness operates on a relatively slow time scale, allowing for the correct integration of information [113, 114]. Therefore, it would seem that these aspects of integration and consciousness are closely linked, as evidenced by studies such as proposed by Laureys [237], which showed that auditory cortices of vegetative patients respond to stimulation, while the higher-order multimodal areas do not, indicating dysfunction in information integration.

In order to explain these observations, the Global Workspace Theory (GWT) suggests that consciousness facilitates to link and integrate separate and independent brain functions [22]. According to this theory, consciousness is necessary to perform complex learning such as that involving novelty or other brain mechanisms (e.g. working memory, voluntary control, attention...) [22, 23]. The GWT proposes that this integration through consciousness is possible thanks to the infrastructure of the brain itself [339, 340]. This theory promotes the idea of functional hubs linking and propagating information from one spatial population to another [23]. Although originally independent, these functional zones can transmit and receive information from other modalities through these hubs. These hubs or workspace nodes are influenced by the competition between these mutually inhibiting cortical populations and by the top-down feedback [21, 339, 340]. Therefore, these neurons could facilitate the flow of information from one region to another through the long-range

connections that link them [339]. In other words, consciousness, thanks to these connections, makes it possible to form associations between different pieces of information.

Another theory approaching the previous one, the Integrated Information Theory (IIT), states that consciousness corresponds to the capacity of a system to integrate information [376, 378]. The main difference here is that we are referring to a quantity of available consciousness (i.e. the amount of information that can be integrated) and no longer to a simple mechanism. To measure this quantity, the theory refers to the causality of the information to be integrated and so to the way in which the elements of a system interact with each other in a unified way, forming a coherent whole [376, 377]. Indeed, the more information are linked (notably spatio-temporally), the more easily they can be integrated. Consequently, IIT also suggests that conscious states are exclusive, meaning that when a particular set of elements within a system enters a state of integrated information, other subsets are excluded [376, 377]. Despite these differences in the approach to deal with consciousness, IIT also evokes the fact that: the presence of different specialized functional regions as in the brain and, the existence of causal structures between elements in a system (as with hubs in the GWT), allow for an ideal integration of information and favours the quality of conscious experience [378]. These theories reflect the aspect of segregation and integration of information, a notion that we will see is present in the brain and will be a fundamental concept throughout this thesis. According to these theories, the architecture of the brain seems to be the basis of the notion of consciousness.

The modular and hierarchical organization of the brain

As stated earlier, numerous studies have shown that the brain's connectivity follows a modular organization at different spatial and functional scales, with neurons and regions associated to common modalities or functions being more strongly connected [404, 403, 179, 266, 422]. These modules are usually associated to particular sensory modality (e.g., vision, audition and motor control) or to specific features within a modality, emerging in an autonomous way [403, 179, 248, 422, 42]. Plastic connection strengths seem to play a significant role in the specification of neural assemblies involved in a particular function under the action of co-activation zones [100, 157]. This highlights the concept of semantic memory where correlated information or functions share a common structure [362, 327]. Some models suggest this concept of semantic memory by having association between mental representations and topology [420, 326]. Also, it has been observed that during rest (i.e., no task activity), small series of sequence activation replays occur, akin to a memory retrieval and consequently to a process of memory consolidation [166, 373]. This memory retrieval process in the dynamics seems based on the activation of particular semantic subgroups [166], again highlighting the impact of the physical organization of the network on these dynamics. In summary, the idea of having segregated specialized functional zones, as stated by the GWT and the IIT, is corroborated by biological findings where the modular and hierarchical organisation of the brain is characterized by randomly densely connected communities of neurons and by sparse connections between modules [80, 278, 340, 421].

However, these inter-module connections appear to be centralised on a limited number

of strongly connected neurons [42, 359, 421, 422]. These hubs form a rich-club at the top of the network hierarchy, providing a central workspace for the integration of multisensory information [217, 256, 278, 359, 335, 421, 422]. All of this directly echoes to the GWT enunciated previously, with the segregation of different sensory information into modules and their integration by the hub neurons [159, 340, 421]. In addition, the rich-club regions could harmonize the brain regions using oscillations to promote integration (or communication) between them according to a particular task [42, 217, 335]. Thus, they are involved in diverse cognitive processes, including attention, memory, decision-making and sensorimotor integration [64]. Therefore, hub neurons are heterogeneous and may differ in their specific properties, connectivity patterns, and functional roles depending on the brain region they belong [271]. In particular, it should be noted that, although hubs are functionally interdependent in terms of inputs and outputs, they don't necessary have strong incoming and outgoing connections [161, 421]. However, their high degree of connectivity and the fact that they are crucial for a normal cognition, they constitute a vulnerable area to damage or dysfunction since their perturbation can have cascading effects on network communication [360, 94]; which ultimately can be at the origin of neurological and psychiatric disorders [64, 359]. Nevertheless, hub neurons are also considered as resilient because they possess redundant connections, which allow compensatory mechanisms and the maintenance of network integrity even in the face of damage or disruptions [360, 390].

Although we have seen that plasticity plays a significant role in the specialization of areas in particular functions, how these modular structures and especially the hubs that connect them emerge remains an open question. However, several key factors seem to contribute to the development of this organisation. First of all, genetic instructions determine the initial organization and development of the brain by guiding the formation, migration, and differentiation of neural progenitor cells, ultimately shaping the overall structure and connectivity of brain regions [287, 310]. Thus, through evolution, genetics provides the foundation of this organization, notably by encoding the pathway of sensory informations to different brain areas. However, it is well known that the frontiers between cortical areas are less distinct in immature brain [286, 309]. This underlines the fact that this structure is not totally innate and the importance of its posterior construction. Indeed, as previous experiments suggest, neural activity shapes brain organization through activity-dependent processes, such as synaptic plasticity, which strengthen specific connections and contribute to efficient information processing [176, 210]. Similarly, this organization is also shaped by sensory experiences and environmental stimuli (including developmental constraint [292, 309]) through experience-dependent plasticity, which enables adaptive modifications in connectivity to optimize information representation and processing based on specific inputs and cognitive demands [137, 219]. In summary, these particular structures appear to be largely induced by the various plasticity phenomena; but their consistent maintenance over time remains elusive.

Problems and objectives

Learning We have seen that the particular structuring of the brain's biological network appears to be largely induced firstly by genetic factors, and then by the different adaptation mechanisms. However from a computational point of view, the emergence of modular structures is still unclear. Some approaches propose that from a totally random network, the topological overlap mechanism can be used to create proto-modules leading to the emergence of more complex modules [101]. Similarly, it is shown that starting from different initial network configurations, local Hebbian plasticity rule can permit a self-(re)organization of connection between units resulting in a modular topology [100, 268, 389]. Also in a different direction, genetic algorithms have been used to evolve the connectivity of neural network models towards modular structures proving the certain optimality of such a structure to maximize network performance and minimize connection costs [88]. From these insights, the first question is to understand how this kind of architecture can emerge through local adaptation in ANN. Consequently, the initial goal of this thesis aims to **highlight the phenomena of segregation in modular neural networks**. To do this, the first step will be to explain the emergence of these structures using specific localized inputs associated with learning rules based on local plasticity. Therefore, it will be necessary to investigate the role of these local dynamics on the excitatory-inhibitory balance and on the network architecture in order to optimize learning capacities. At the same time, an analogy will be drawn between the creation of these structures and that of memory, particularly through their link with learning.

Maintaining and consolidating The second problematic addressed in this thesis is to **understand how the emerging structures can be maintained over time**. Indeed, the information, memories or sensory functions that the learned structures encode, must be maintained so that they can be reused in a coherent way over the long-term. Moreover, unlike conventional AI systems where adaptation is inactive after learning, biological network are in constant adaptation and reorganization, without forgetting what they learn. Going a step further, we saw earlier that the memories contained in these structures can be consolidated [166, 373], once again highlighting this structural maintenance. The objective will be to find the mechanisms that enable this maintenance and consolidation in neural network models while allowing the network to admit continuous activity and adaptation. Thus, these results will provide a better understanding of how **long-term memory** works.

Processing As final purpose, we aim at **evaluating this architecture and the resulting mechanisms using real sensory information**. By applying different cognitive stimuli to the network, we expect to validate the previous objectives by forming stable and maintainable modular structures, each representing a sensory modality. These individual modules should be able to efficiently learn some particular patterns in order to separately process simple recognition or generation tasks proving the modality segregation aspect. Then, the final goal will be to **understand and evidence the effectiveness of the phenomenon of integration**. For this purpose, the modules of the architec-

ture will have to cooperate synergistically to guarantee optimal multimodal processing, and thus demonstrate the gain of such an approach on cognitive tasks. In addition to **contributions to the field of neuroscience**, notably on understanding the formation and maintenance of long-term memory, this thesis will bring some insights for **more bio-realistic AI systems**. These results may lead to more efficient and less energy-consuming learning, thus improving the operation of existing ANN.

Protocol overview

To address the different research questions enunciated, we plan throughout this thesis to base ourselves on the following protocol. The first general idea is to start from a random graph with randomly connected neurons, which can be compared to an immature brain or an untrained network. Then, external inputs are applied to different subsets of neurons in the network. We foresee that these stimulation along with synaptic adaptation will lead to the **formation of modular structures** as represented in Fig. 1.

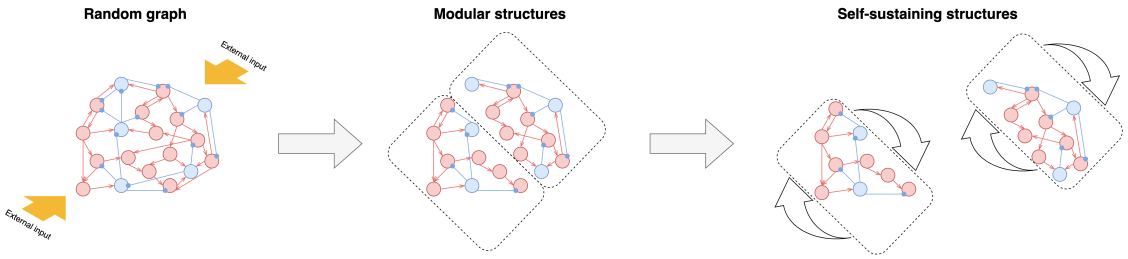


Figure 1: Overview of the general experimental protocol used during this thesis. The network is initially connected randomly, then the application of external inputs leads to the formation of modular structures that are self-maintaining over the long-term. Red (blue) circles represent excitatory (inhibitory) units.

We then let the network evolve, emphasising the mechanisms that enable the structures formed to **maintain themselves over the long-term**. Overall, throughout the process, we focus on the following biological constraints:

- the role of inhibitory neurons
- the maintenance of spontaneous neuronal activity
- the presence of constant synaptic adaptation

Outline

After introducing the original motivation of this thesis by linking **artificial intelligence and neuroscience**, we reviewed some theoretical and biological concepts on consciousness and brain structural organisation allowing to highlight some key topics of this thesis. The

problems and objectives as well as a protocol overview of this study are then enunciated, emphasising on four main themes: learning, maintaining, consolidating and processing memories; paving the way for the four chapters of this thesis.

In the first chapter, we start doing a state of the art of more technical concepts notably the aspects of oscillations, synchronization and oscillator models. Then still in this same chapter, we present and reproduce some state of art results with **Kuramoto oscillators** giving some basis on coupled units subject to synaptic adaptation. These results and this model are then further developed in order to adapt them to our problems and to obtain some preliminary results that will enable us to better define the approaches to be followed and the challenges to be overcome, particularly with regard to the formation and learning of structural clusters via external stimulation.

In the Chapter 2, which constitutes the first content chapter and largely takes up the works of our first article [38], we apply the reflections of the previous chapter on oscillators to **θ-neurons**. In addition to the formation of neuronal assemblies, we focus on the role of inhibition in the maintenance process of these structures. The Chapter 3, linked to the work in our second article [37], extends the previous results to **spiking neurons**. More precisely, we evolve the model adding some realistic constraints in order to make it closer to biology. Here, in addition to the initial problem of formation of memories, we place more emphasis on the richness of neuronal dynamics and their roles, in particular with spontaneous recalls, which favours the consolidation of memory items.

The Chapter 4 constitutes a **multimodal application** of the model established in the previous chapter. We confirm the previous results on learning and memory maintenance using real sensory stimuli. In addition, in this chapter we further develop the aspect of modality integration previously described, by evaluating the model on tasks such as the generation and recognition of audio-visual information.

Finally, in the discussion chapter, we summarize and develop the different notions treated in this thesis, such as **the impact of inhibition on network dynamics, the sustainability of learned memory and the memory capacity of the network**. Lastly, we discuss the perspectives that we envisage for future work, and which may be of interest in the fields of artificial intelligence and machine learning or neuroscience and biology.

Chapter 1

State of the art: formation of multi-clusters with oscillator models

Contents

In this chapter, we revise the state of the art regarding the formation of multiple clusters in networks of Kuramoto oscillators subject to adaptation. In particular, we review concepts such as brain oscillations, synchronization and models of coupled oscillators. We also present preliminary investigations conducted on the formation and learning of structural clusters via external stimulation, which will pave the way for the experiments in the following chapters.

1.1 Background

1.1.1 Oscillations in the brain

The brain exhibits rhythmic and repetitive patterns of neuronal activity known as oscillations. These oscillations play a vital role in various cognitive processes. They facilitate the encoding and processing of information [198, 216], integration and communication between brain regions [71, 146], and coordination of cognitive and physiological functions [68, 343]. They are categorized into different frequency bands based on their characteristic frequencies and functions [216]:

- Delta rhythms (0.5-4 Hz) are associated with deep sleep and certain pathological conditions. They are also thought to be crucial for their developmental and restorative functions, such as memory consolidation, synaptic plasticity, and brain recovery [364].
- Theta rhythms (4-8 Hz) are involved in memory processes, attention, and spatial navigation [328].
- Alpha rhythms (8-12 Hz) are often observed during wakefulness and eye closure. They are also linked to attentional processes, sensory inhibition, and information control [197].
- Beta rhythms (12-30 Hz) are associated with sensorimotor processing and planning, as well as cognitive functions [128].
- Gamma rhythms (30-100 Hz) are involved in higher-order cognitive processes and sensory integration [69, 199].

Brain oscillations emerge from the interplay of neural excitability and inhibition [68, 407], intrinsic properties of neurons [192], network connectivity [148], and neuromodulation [410]. The interaction between excitatory and inhibitory activity, combined with the intrinsic properties of neurons, allows to generate rhythmic patterns [68, 110, 192, 407]. Network connectivity and feedback loops enable the synchronization and coordination of oscillatory activity across brain regions [68, 148]. Neuromodulatory systems influence oscillations by regulating the balance between excitation and inhibition [410]. These factors collectively contribute to the emergence of a wide variety of oscillatory patterns in the brain.

The importance of this mechanism is underlined by the fact that disturbances in brain oscillations have been associated with several neurological and psychiatric disorders. For instance, abnormalities in alpha oscillations have been observed in pathologies such as Alzheimer's disease and attention deficit disorders [24]. Similarly, altered gamma oscillations have been implicated in schizophrenia [31]. These abnormalities have been also linked to impaired sensory integration, disrupted information processing, and impaired cognitive functions such as working memory [84, 387]

1.1.2 Models of coupled oscillators

The brain admits states of collective oscillation dynamics that play a central role in communication and neuronal functions. Therefore, it may be interesting to consider models that focus on this oscillation mechanism in order to investigate the complex dynamics arising from interactions between oscillatory systems. Several fundamental models describe this phenomenon:

- The Kuramoto model is a basic model that demonstrates mutual synchronization among coupled oscillators [225, 226, 228]. It involves a population of phase oscillators that interact through a coupling term. This model has been extensively studied in the

context of synchronization phenomena, including biological rhythms and collective behaviours in complex networks.

- The Stuart-Landau model represents a class of coupled oscillators that exhibit limit cycle oscillations [233, 368, 369, 302]. It describes the dynamics of complex systems, such as chemical reactions or coupled lasers, through a set of differential equations. The model captures the interaction between individual oscillators and the emergence of collective behaviour.

- The Winfree model is used to study the dynamics of coupled oscillators on a two-dimensional grid [129, 270, 412]. It considers the interaction between phase oscillators with nearest-neighbour coupling. The model has been applied to understand the synchronization patterns observed in biological systems, such as firefly flashing or circadian rhythms.

- The Adler model describes the behaviour of coupled oscillators under the influence of weak periodic forcing [5, 45, 293]. It incorporates a forcing term that interacts with the intrinsic dynamics of the oscillators. This model has been instrumental in studying phenomena like frequency entrainment and phase-locking in coupled systems.

- The Jansen-Ritt model is a neural mass model that is widely used for simulating electroencephalogram (EEG) oscillatory signals and understanding the dynamics of neural populations in the brain regions [196, 351, 380, 409]. This model is a three-population model (i.e. excitatory pyramidal cells, inhibitory interneurons, and a population of external inputs) that studies EEG phenomena like alpha, beta, and gamma rhythms, as well as neurological disorders such as epileptic seizures.

All the models presented above describe the dynamics and interactions between coupled oscillators. Consequently, the couplings that link them appear to be decisive parameters in their behaviour. Some studies fix these couplings during simulations and eventually adapt them a posteriori [4, 130]. However, similar to the plasticity mechanisms observed in brain neurons, introducing a plasticity mechanism to adapt coupling connections over time can be beneficial. Furthermore, to obtain complex states such as partial synchronization, adaptation could be required in the coupling between oscillators [131, 347]. For this purpose, the STDP rule can be employed to make appear multi-stability of synchronized and desynchronized clusters [253]. More suited to phase models and oscillators, Phase-difference-dependent plasticity (PDDP) establishes a relation between spikes timing and phases [12, 13, 41, 311]. Thus, depending on the phase difference between two oscillators, their connection can be strengthened or weakened depending on the nature of the plasticity function. By employing such learning rules, the adaptation of coupling weights can lead to a variety of collective behaviours in the phases of the oscillators [12, 13, 41]. Consequently, the phase patterns and the weighted network structures mutually influence each other at different time scales, to converge towards stable dynamics [13, 41, 311].

While fully connected networks associated with adaptation can admit complex dynamics, investigating the impact of imposed network topologies provides additional insights. To

begin with, random network topologies with different degrees of connectivity dilution have been analysed [399]. The results reveal that random dilution of links leads to desynchronisation of the originally present clusters in the network dynamics, emphasizing the influence of network topology (learned or imposed) on the states of a system [399]. Similarly, in a modular structure of coupled oscillators, a correlation between highly interconnected units and communities of synchronized oscillators has been demonstrated [18, 17, 15]. Likewise, similar experiments have been carried out on the synchronization of oscillators in scale-free topologies, highlighting the role of the hub neurons [276, 311, 419]. These particular nodes are more stable than weakly linked nodes and therefore synchronize more easily with their neighbours without destroying the latter's synchrony if these hubs are removed [276]. Finally, as in more conventional AI systems, it is possible to implement a reservoir computing architecture with coupled oscillators, enabling cost reduction and efficient processing thanks in particular to the synchronization property of oscillators and their ability to individually represent a large population of units [417].

1.1.3 Synchronization

Oscillations in the brain can facilitate synchronization between multiple neurons or different brain regions by coordinating the timing of neural activity [71, 146]. Synaptic connections between neurons are one of the key factors contributing to the emergence of these synchronised patterns. Indeed, these couplings enable the exchange of electrical and chemical signals between neurons, promoting alignment of oscillatory phases or frequencies between interconnected neurons [8, 407]. Another factor is that neurons with similar resonant frequencies, determined by their intrinsic properties, may synchronize when coupled together, as their oscillatory activity in interactions naturally tends to align [69, 194]. As enunciated earlier, feedback and feed-forward loops in the brain influence synchronization by modulating activity between neuronal populations through feedback or forward connections [33, 108]. Finally, external factors, such as rhythmic sensory input or oscillatory drive from other brain areas, can synchronize neuronal populations by entraining their activity, resulting in coordinated oscillatory patterns [70, 178]. Ultimately, all these factors are mainly linked to the coupling weights between neurons and therefore to the adaptation mechanism that ensures this particular connectivity.

In the field of modelling, we identify the same factors that contribute to the emergence of synchronized patterns in an oscillatory regime [43, 120]. Indeed, synchronization arises as a collective behaviour through interactions and coupling between individual oscillators, which align their phases or frequencies through the exchange of information and energy. The specific mechanisms underlying synchronization vary depending on the model being used and involve factors such as coupling strength, network topology, and oscillator dynamics [50, 302, 367, 16, 270]. For instance, in certain models like the Kuramoto oscillators, synchronisation can easily occur when the coupling strength is sufficiently large. However, in other cases, it may be necessary to modulate the degree of synchronization.

Indeed, although synchronization plays an important role in brain function, global synchronization is not necessarily desirable due to potential information loss [337], reduced

flexibility [147], instability [56], limited functional specialization [354], and association with pathological conditions [386]. Therefore, achieving partial synchronization is more appropriate, allowing for both coordinated activity and functional diversity, enabling the brain to perform complex and adaptive information processing tasks. Various mechanisms can be employed to prevent total synchronization. For example, we can play on the amplitudes of the oscillations (large amplitudes tending to de-phase oscillators) [169], introducing phase conduction delays [93, 322], adding random external noise [227], or having oscillators with non-uniform natural frequencies [227, 322]. Nevertheless, the main factor influencing synchronization in oscillators remains the coupling strength, with stronger coupling increasing the likelihood of neurons to align their phases [227, 322, 414]. Positive couplings promote in-phase relationships, while negative couplings favour anti-phase relationships with the mean field, which can result in fully coherent, incoherent, or multi-stable states [414, 184]. Inhibition plays a crucial role in controlling the levels of synchronization in oscillator models, as it allows for flexible modulation of the balance between synchrony and asynchrony in the collective dynamics of oscillators. By adjusting the strength and timing of inhibitory connections, we can precisely control the emergence and stability of synchronized patterns. Indeed, strong inhibitory interactions can disrupt synchronization by introducing desynchronization or anti-phase relationships between oscillators, leading to more irregular or asynchronous dynamics [408]. Conversely, inhibition can also enhance synchronization by suppressing the activity of some oscillators, favouring the dominance of others, and aligning their phases. This leads to greater coherence and stability in oscillatory patterns while avoiding destabilizing factors [133, 131, 243]. Achieving and maintaining synchronized patterns in oscillator models heavily relies on finding the right balance between excitation and inhibition. Consequently, if we consider stable proportions of excitatory and inhibitory populations, it requires appropriate adjustment of the inhibitory strength relative to excitatory coupling [120, 374]. In particular as enunciated previously, the adaptation of the coupling between the oscillators may be require to obtain this balance [131, 347].

1.2 Emergence of complex dynamical patterns via adaptation

Through the state of the art, we have been able to highlight the importance of oscillation and synchronisation mechanisms in the brain in order to ensure correct cognitive processes. These phenomena can be described fairly accurately using coupled oscillator models, which represent neurons or ensembles of neurons connected by coupling weights. The adaptation mechanism of these connections is decisive for modulating the degree of synchronization and, by extension, for creating complex dynamics. In this section, we illustrate this idea by reproducing previous work using coupled oscillators associated with an adaptation mechanism.

1.2.1 Kuramoto oscillators

The Kuramoto model for coupled phase oscillators has the advantage of being a relatively simple model able to capture different types of dynamics ranging from asynchronous to

partially synchronous states [225, 226, 228]. Some studies have shown the possibility of adding an adaptation mechanism to the coupling leading to more complex dynamical behaviours [12, 13, 40, 39]. For these reasons, we took the decision to start with this model as baseline and reproduce the experiments from Aoki et al. [12, 13]. Therefore, the evolution of the phase θ_i of an oscillator ($i = 1, \ldots, N$) is described by:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{g}{N}\sum_{j=1}^{N} a_{ij}\kappa_{ij}\sin(\theta_j - \theta_i - \alpha). \tag{1.1}$$

The natural frequency of oscillator i is denoted by ω_i, g represents the global coupling strength, while $\kappa_{ij} \in [-1, 1]$ is the relative directed coupling weight from the oscillator j towards the oscillator i. In this context, $\alpha \in [0, \frac{\pi}{2}[$ represents a synaptic transmission delay and $a_{ij} \in \{0, 1\}$ is the connectivity of the adjacency matrix from the oscillator j towards the oscillator i.

It should be noted that the Kuramoto model exhibits a synchronization threshold, where the system transitions from a desynchronized state to synchronization when the coupling strength exceeds a critical value [4]. This coupling threshold is influenced by factors such as the distribution of natural frequencies or the shape of the network connecting the oscillators.

In our case, unless otherwise stated, we consider a fully connected network topology without *autapses* and homogenous and constant natural frequencies $\omega_i = 1$. Thus, with a coupling strength $g = 1$, the network tends to synchronize naturally and its state depends only on α and κ_{ij}.

Concerning the dynamics of the coupling weight κ_{ij} from the pre-synaptic oscillator j to the post-synaptic oscillator i, it is governed by the differential equation Eq. (1.2) of Berner et al. in [41], originally used by Aoki et al. in [12, 13]:

$$\frac{d\kappa_{ij}}{dt} = \varepsilon(-\kappa_{ij} - \sin(\theta_i - \theta_j + \beta)). \tag{1.2}$$

The learning rate is denoted by $\varepsilon \ll 1$ and $\beta \in [-\pi, \pi[$ is the plasticity function parameter. The term κ_{ij} introduced in the equation is a soft bound that keeps $|\kappa_{ij}|$ smaller than one.

The synaptic weights κ_{ij} are reinforced or depressed according the plasticity function $-\sin(\Delta\theta + \beta)$ depending on the instantaneous phase difference between the post-synaptic and pre-synaptic oscillators $\Delta\theta = \theta_i - \theta_j$. The parameter β determines the behaviour of the PDDP rule during weight adaptation. As represented in Fig. 1.1, depending on this value, we obtain three types of learning rules:

- For $\beta = \frac{-\pi}{2}$, we have a Hebbian symmetric rule with a reinforcement of the weights when the oscillators are in phase and a depression when they are in anti-phase.
- For $\beta = 0$, we obtain a Hebbian asymmetric or causal rule with a reinforcement of the weights when oscillator j is in phase advance with respect to oscillator i and a depression when oscillator i has a higher phase. This function can be compared to the

typical STDP [46] rule found in biological neurons with a potentiation (depression) of the weights when the spiking time of the pre-synaptic neuron j precedes (succeeds) that of the post-synaptic neuron.

- For $\beta = \frac{\pi}{2}$, we get a Anti-Hebbian symmetric rule with a reinforcement of the weights when the oscillators are in anti-phase and a depression when they are in phase.

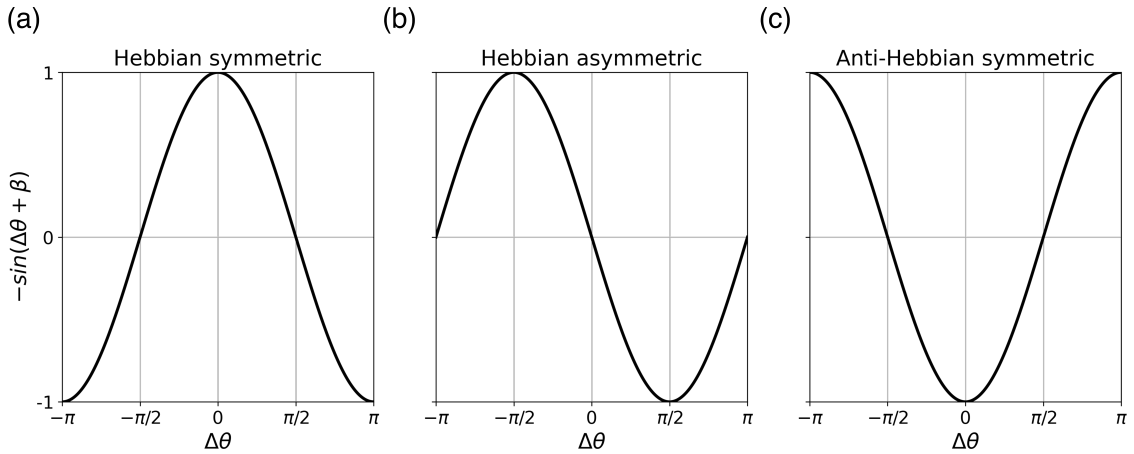


Figure 1.1: The plasticity function $-\sin(\Delta\theta + \beta)$ as a function of $\Delta\theta = \theta_i - \theta_j$ with: (a) for $\beta = \frac{-\pi}{2}$, (b) for $\beta = 0$ and (c) for $\beta = \frac{\pi}{2}$. Figure reproduced from [41].

In order to quantify the degree of synchronization in the network, we introduce the Kuramoto-Daido order parameters $Z_m(t)$ [95, 96, 228]. The m^{th} order parameter is defined as follows:

$$Z_m(t) = R_m(t)\mathrm{e}^{i\Psi_m} = \frac{1}{N}\sum_{j=1}^{N}\mathrm{e}^{im\theta_j}, \tag{1.3}$$

where R_m is the modulus of the complex order parameter Z_m and Ψ_m the corresponding phase. The modulus of Z_1 is employed to characterize the level of phase synchronization in the network: $R_1 > 0$ ($R_1 = 1$) for a partially (fully) synchronized network, while $R_1 \simeq \mathcal{O}(1/\sqrt{N})$) for an asynchronous dynamics. The other modulus R_k with $k > 1$ are used to characterize the emergence of multi-clusters, in particular R_2 is finite whenever two clusters in anti-phase are present in the network.

Then, we can quantify the degree of relation between the phases θ_j of oscillators at time t for a time interval τ by calculating their autocorrelation:

$$Corr(t) = |\frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j(t)}e^{-i\theta_j(t-\tau)}|. \tag{1.4}$$

Finally, to calculate the mean change rate of weights at time t in a network or sub-population of N oscillators, we proceed as follows:

$$K(t) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \kappa_{ij}(t + \Delta t) - \kappa_{ij}(t), \tag{1.5}$$

where $\kappa_{ij}(t)$ and $\kappa_{ij}(t+\Delta t)$ are the coupling weights from oscillator j to i at time t and $t+\delta t$ respectively, Δt being a small time internal. Note that the normalization term takes the form $N(N-1)$ since we consider an *all-to-all* connected network without *autapses*. Thus, the parameter $K(t)$ takes a positive value for an overall increase in weight connectivity and a negative value for an overall decrease.

It is possible to consider a normalized form of this measure by taking the absolute value of the weight difference as follows:

$$\Delta t K(t) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{|\kappa_{ij}(t + \Delta t) - \kappa_{ij}(t)|}{\Delta t}. \tag{1.6}$$

In particular, this measurement shows whether the network has finished learning and converging, or whether its weights are still being adapted.

1.2.2 3 plasticity rules, 3 states

Based on the model described above, the idea here is to study the dynamical and structural states towards which the network converges according to the learning rule employed. To do this, we integrate the differential equations Eqs. (1.1) and (1.2) with the 4th order Runge-Kutta method, and let the network evolve for 6000 iteration time steps, allowing the coupling weights to converge (when this happens).

The networks are initialised with random weights uniformly distributed in the range $\kappa_{ij} \in [-1, 1]$. The initial phases of the oscillators are randomly and uniformly selected in the interval $\theta_i \in [0, 2\pi]$. Concerning the oscillators parameters, $\omega_i = 1$, $g = 1$, $\varepsilon = 0.01$ and $N = 100$ oscillators. The values of the synaptic transmission delay (or phase lag) α and the plasticity parameter β are chosen according to the three cases.

As in Aoki et al. studies [12, 13], using the three learning rules presented in Fig. 1.1, we obtain three different states:

- The Hebbian symmetric rule gives rise to a two-cluster state (antipodal type), see Fig. 1.2.
- The Hebbian asymmetric rule gives rise to a coherent state (splay type), see Fig. 1.3.
- The Anti-Hebbian symmetric rule gives rise to a chaotic state, see Fig. 1.4.

Two-cluster state Using the Hebbian symmetric rule, we observe the formation of two groups of oscillators in anti-phase as shown by the phase patterns in Fig. 1.2 (f), their distribution in Fig. 1.2 (b) and especially in Fig. 1.2 (a) where the 2nd order parameter is dominant and equal to one confirming the presence of two distinct dynamics in the network. These two clusters are also found in the weighted connectivity matrix after convergence

(see Fig. 1.2 (d) and (e)), where synchronized oscillators are connected by positive links and desynchronized ones are connected by negative links. The emergence of these two groups can be seen as a consequence of the two stable fixed points of the plasticity function with oscillators in phase or in anti-phase. Also, the symmetrical nature of this same function can be seen in the symmetry of the weighted connectivity matrix, with the same connections to and from the post-synaptic and pre-synaptic units. We note that although each cluster admits a similar number of oscillators, this size can vary and depends greatly on the phase initialization. For instance, an in-phase initialization (all oscillators have the same initial phase) will lead to one cluster state.

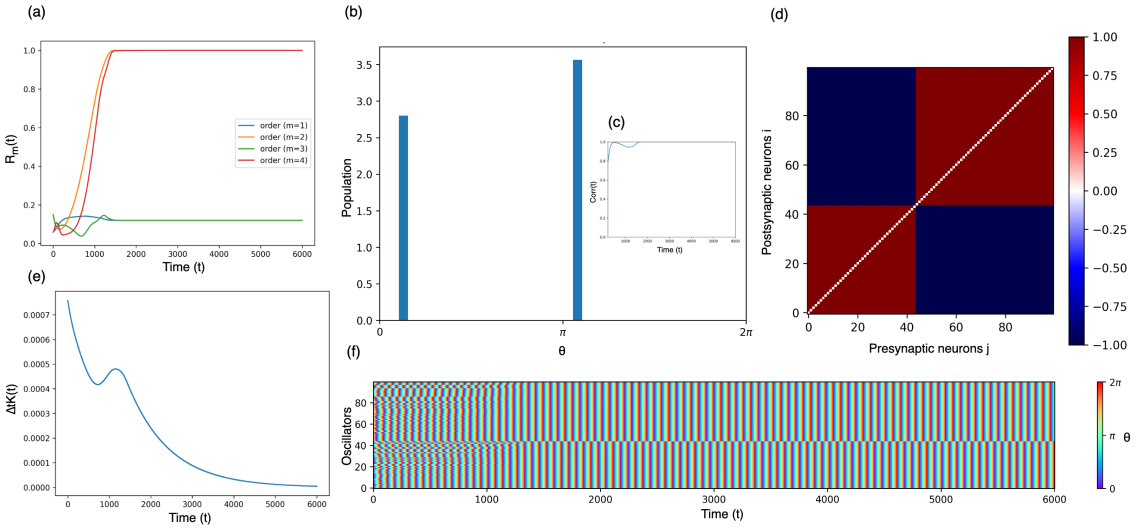


Figure 1.2: Two-cluster state obtained using the Hebbian symmetric rule with: (a) the time development of the order parameters, (b) the distribution of neurons ($N = 100$) by phases (logarithmic scale), (c) the autocorrelations of the phase pattern, (d) the weighted connectivity matrix (sorted by phases), (e) the evolution of the mean change rate of weights and (f) the phase patterns (sorted). Parameters: $\alpha = 0$ and $\beta = -0.5\pi$.

Coherent state Using the Hebbian asymmetric rule, we observe the formation of a "splay" state (see the phase patterns in Fig. 1.3 (f)) where each oscillator maintains a fixed phase difference with its neighbouring oscillators. Although the network is not in a fully synchronized state, as confirmed by the phase distribution in Fig. 1.3 (b) and the fact that no order parameters really dominates (see Fig. 1.3 (a)), the patterns remain coherent as shown by their autocorrelations equal to 1 (see Fig. 1.3 (c)). Concerning the weighted connectivity, we converge towards a more complex structure (see Fig. 1.3 (d) and (e)). In this case, the oscillators project negative links to oscillators with phase advance (by up to π) and positive links to oscillators with phase delay (by up to π maximum). This reflects the causal nature of the plasticity function and its asymmetry, as evidenced by the structure obtained where the links between two oscillators have opposite sign.

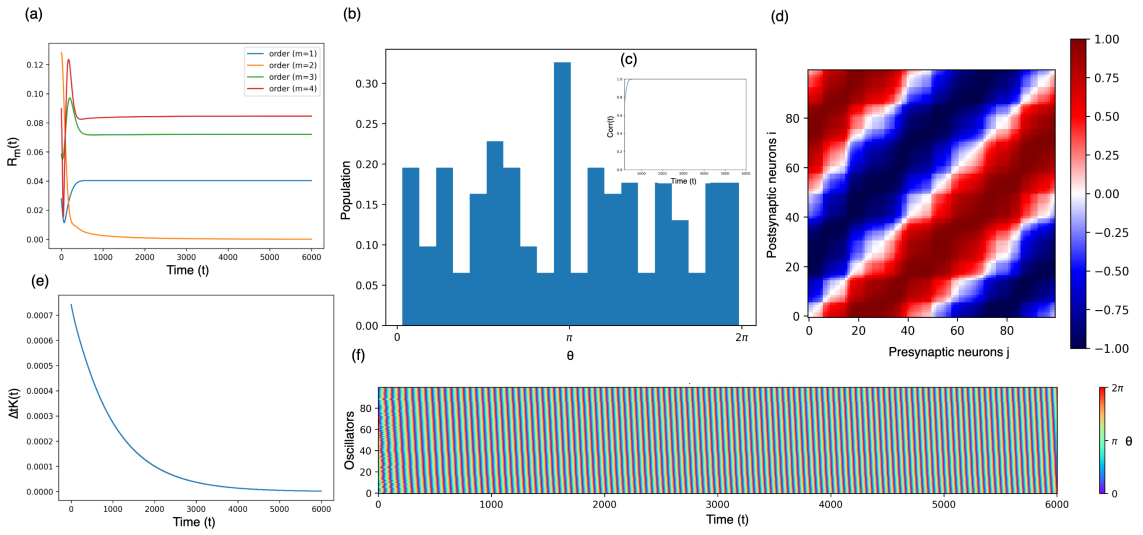


Figure 1.3: Coherent state obtained using the Hebbian asymmetric rule with: (a) the time development of the order parameters, (b) the distribution of neurons ($N = 100$) by phases (logarithmic scale), (c) the autocorrelations of the phase pattern, (d) the weighted connectivity matrix (sorted by phases), (e) the evolution of the mean change rate of weights and (f) the phase patterns (sorted). Parameters: $\alpha = 0.3\pi$ and $\beta = 0$.

Chaotic state Finally, using the Anti-Hebbian symmetric rule, we observe the emergence of chaotic dynamics. Indeed, no oscillators become synchronized as evidenced by the phase patterns in Fig. 1.4 (f), their distribution in Fig. 1.4 (b) and the order parameters, which admit small values in constant variation (see Fig. 1.4 (a)). Moreover, the almost null value of the autocorrelations (see Fig. 1.4 (c)), confirms the non coherence of the patterns. This inconsistency is also found in the weighted connectivity, with random connections and no particular structure formed (see Fig. 1.4 (d)). However, we note that the symmetric nature of the plasticity function is preserved with the same connections to and from post-synaptic and pre-synaptic units in the same manner as with Hebbian symmetric rule. Yet, this structure is not stable, as shown by the mean change rate of weights in Fig. 1.4 (e), which is constantly changing. Intuitively, this can be explained by the fact that oscillators with positive coupling tend to synchronize, but the plasticity function will in this case depress the weights to make them negative. This consequently causes the oscillators to desynchronise, which in turn causes the plasticity function to reinforce the weights, which again leads to synchronisation. This "infinite loop" is at the origin of this chaotic regime and explains the tendency of the weights to be close to 0.

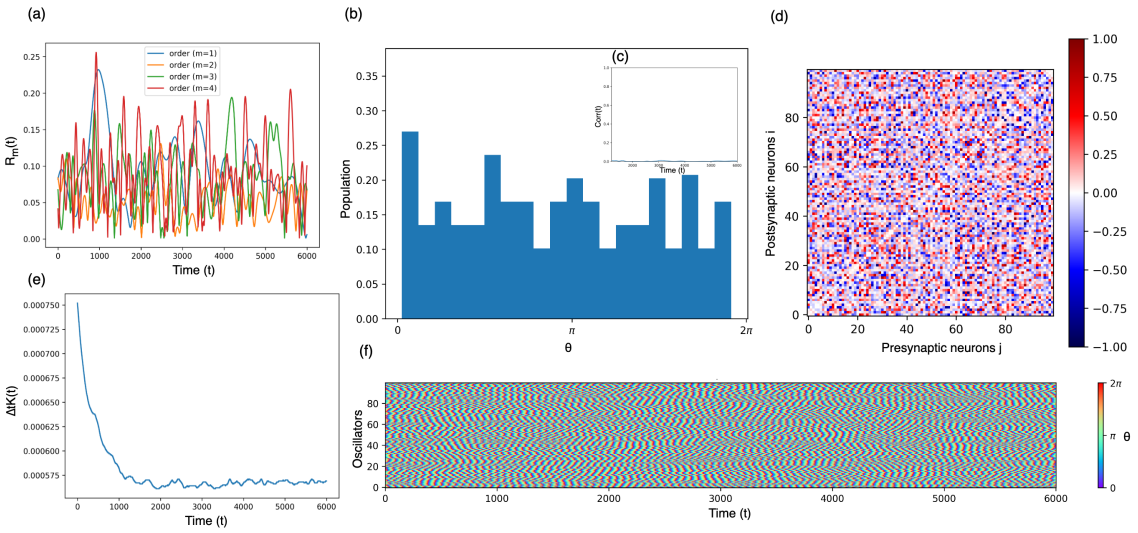


Figure 1.4: Chaotic state obtained using the Anti-Hebbian symmetric rule with: (a) the time development of the order parameters, (b) the distribution of neurons ($N = 100$) by phases (logarithmic scale), (c) the autocorrelations of the phase pattern, (d) the weighted connectivity matrix (sorted by phases), (e) the evolution of the mean change rate of weights and (f) the phase patterns (sorted). Parameters: $\alpha = 0$ and $\beta = 0.5\pi$.

Without going into detail, we have studied the impact of different weight and phase initialisations at the beginning of the simulation. Except in some specific cases (see the paragraph on Two-cluster state), these parameters have a negligible impact. For this reason, we will continue to consider a random distribution to initialize these variables in order to guarantee a certain homogeneity in the results.

We also investigated the effect of the network topology (i.e., the connectivity between neurons given by the adjacency matrix a_{ij} in Eq. (1.1)). Although we have implemented different types of topology (see Fig. A.1 in Appendix A.1), after convergence we always obtain the same states and weight organizations associated with the three learning rules as shown in Fig. A.2 of Appendix A.1. However, we logically observe that the sparser the topology, the longer the system takes to converge to a particular state. We can conclude that in such models, connectivity does not seem to have a significant impact if there is no synaptic propagation delay associated with the particular topology of the network. Therefore, the *all-to-all* topology seems acceptable at this stage of the study.

These experiments confirm that with this model coupled with an adaptation mechanism, it is possible to obtain various collective behaviours and dynamics. Moreover, by playing with the α and β parameters, it is possible to create more complex patterns such as three cluster state, double antipodal and so on [41]. However, we have not developed further in this direction. Indeed, in these cases, the clusters created are only present in the dynamics; structurally, there are no decoupled modules in the connectivity. In addition, the structures obtained are only the consequence of the specific rules applied and not the effect of particular learned information. As a result, other approaches to the formation of

structural clusters need to be investigated.

1.3 Formation of structural clusters via external stimulation

Since, the first goal of this thesis is to highlight the phenomenon of segregation through the creation modular structures, the idea now is to form clusters in weight connectivity of the network by learning some particular patterns. To do this, we modify Eq. (1.1) to incorporate an external input representing the information to be learned.

$$\frac{d\theta_i}{dt} = \omega_i + \frac{g}{N}(\sum_{j=1}^{N} a_{ij}\kappa_{ij}\sin(\theta_j - \theta_i - \alpha)) + I_i(t). \tag{1.7}$$

As shown in Eq. (1.7), the input I_i (associated with the ith oscillator) can be a constant value or a time-varying entry. The immediate effect of this positive external current is to increase the frequency of the oscillator with which it is associated, since this term can be seen as an increase in the natural frequency ω_i. From a more biological point of view, this makes sense if we consider this input as an external stimulus which excites the neuron (or oscillator) and increases its action potential (in this case its phase) and therefore its firing rate. In other words, this entry has a direct impact on the phase state of the oscillator and therefore on the adaptation of the weights.

We note that in this section we will only consider the Hebbian symmetric rule as plasticity function since the main purpose of this part is to develop the approach for learning information rather than how to adapt the weights. As the Hebbian symmetric rule satisfies the notion of "*cells that fire together, wire together*" [174], the results with the asymmetric one are presented in Appendix A.2, whereas the Anti-Hebbian rule is not developed as it is unstable (see Sec. 1.2.2). Furthermore, we will see later that the nature of the plasticity function has no real impact at this stage of the experiment.

1.3.1 Frequency-based adaptation

In order to learn specific patterns and observe the possible emergence of certain dynamics and structures, two approaches are possible. The first technique involves applying inputs of different amplitudes to each oscillator. As a result, the oscillators are stimulated differently and therefore have different frequencies which should imply variability in the weight connections. In other words, it is the frequency of the oscillators (modulated by the external inputs) which create correlations or decorrelations between them and thus encode the information. This approach is summarized in the Fig. 1.5.

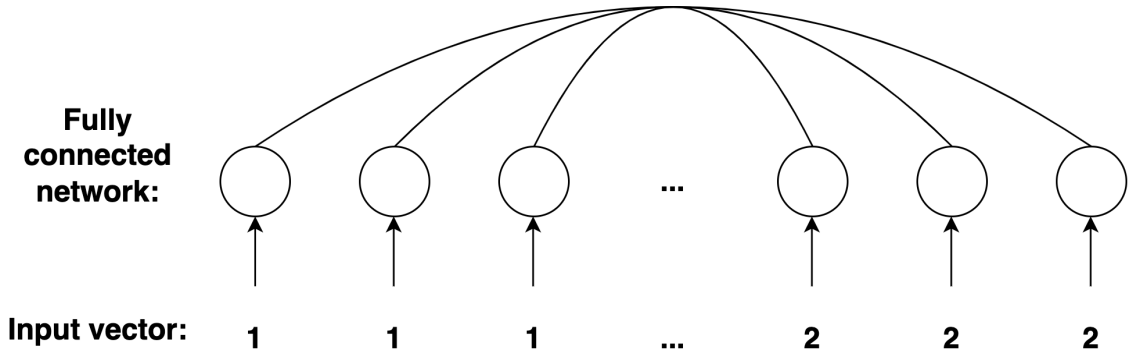


Figure 1.5: Diagram of the frequency-based adaptation with the learning of an input vector composed of 2 different values. Some oscillators receive an external input of 1 while the other part receives an external input of 2.

Thus, several types of input distributions are possible in order to form more or less complex dynamics. To quantify the effective frequency of an oscillator i for a period T at time t, we calculate its mean frequency as follows:

$$\nu_i(t) = \frac{(\theta_i(t) - \theta_i(t-T)) + 2\pi n_i^{sp}}{T}, \tag{1.8}$$

where n_i^{sp} is the spike count emitted by oscillator i in a time interval of duration T. We consider that a spike is emitted each time the phase of an oscillator reaches the value 2π.

Two values input To begin with a simple case, we consider an external input vector I_i composed of two different values (of amplitudes 1 and 2), each applied to a subset of the network as represented in Fig. 1.5. As depicted in Fig. 1.6 with the Hebbian symmetric rule and in Fig. A.3 of Appendix A.2.0.1 with the Hebbian asymmetric rule, we obtain two structural clusters in the weight connectivity matrix (c). These two groups are decoupled while the oscillators constituting them are strongly connected. These clusters clearly correspond to the two input values applied (see the natural frequency distribution in Fig. 1.6 (a)), which implies two populations of neurons with different mean frequencies in panel (d). These different frequency bands are visible in the phase patterns in panel (e), with a group showing faster oscillations than the other. In addition, each cluster exhibits internal dynamics equivalent to that obtained in Fig. 1.2 with a "two-cluster state" as confirmed by the panels (b) and (e). Note that in this case, the external input remains active throughout the simulation. In any case, this experiment confirms our expectations on the possibility of creating modular structures using learning based on input intensity correlation. As a result, this phenomenon can be extended to a larger number of input values in order to obtain more clusters.

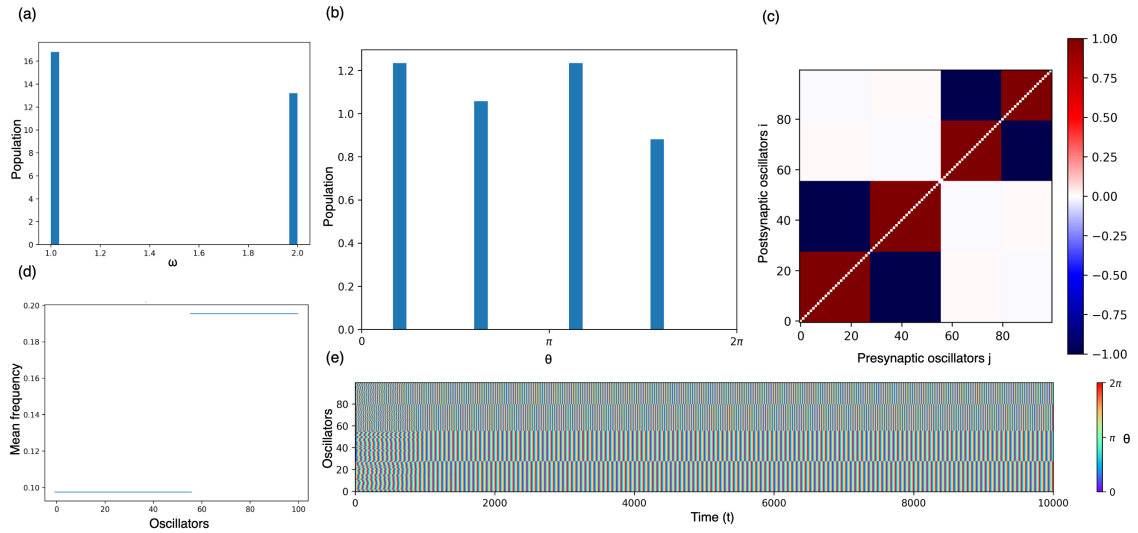


Figure 1.6: Learning of a two values input using a Hebbian symmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the distribution of neurons ($N = 100$) by phases (logarithmic scale), (c) the weighted connectivity matrix (sorted by mean frequencies and phases), (d) the mean frequency of each oscillator and (e) the phase patterns (sorted by mean frequencies and phases). Parameters: $\alpha = 0$ and $\beta = -0.5\pi$.

Gaussian input To extend to more realistic input vectors, we distribute the values using a Gaussian law centred in 2 with a standard deviation of 0.2. The results obtained using the Hebbian symmetric rule are represented in Fig. 1.7 and those using the Hebbian asymmetric rule in Fig. A.4 of Appendix A.2.0.1. In both cases, we observe the formation of several clusters of varying size in the weighted connectivity matrix (c). Each cluster corresponds to a "plateau" in the mean frequency graph (b), induced by the Gaussian distribution of the inputs (a), which explains the presence of larger modules in the centre and smaller ones at the extremities. The interpretation that can be made is that the adaptation of weights is very sensitive to the frequencies of oscillators. Therefore, when neurons receive inputs that are significantly dissimilar, the connections between them tend to approach zero, while conversely, similar stimuli lead to stronger weights near 1 or -1, even if the standard deviation of input values is relatively small. However, there are weights with intermediate values between close clusters, but they remain relatively low. This sensitivity to the nature of the inputs can be seen as a disadvantage if we want to have a good generalisation capability, but on the contrary, it can be seen as an advantage for its ability to well discriminate inputs with a high degree of precision. In any case, this experiment shows the possibility of creating modular structures using heterogeneous inputs, while admitting connections with intermediate coupling between clusters of close frequencies.

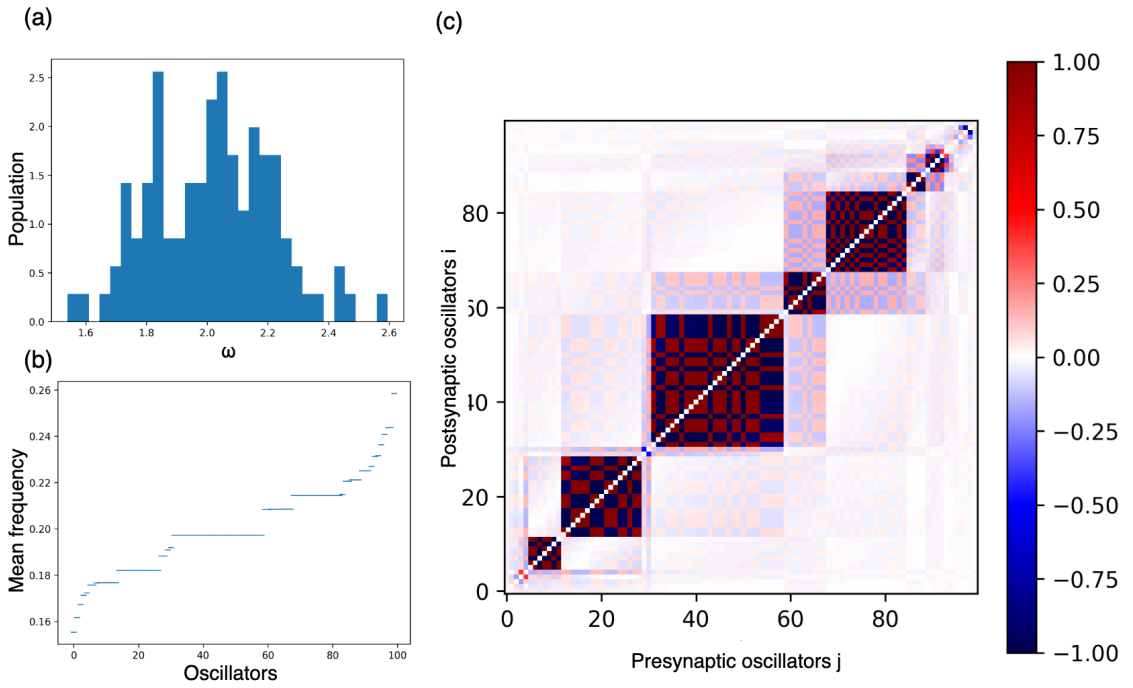


Figure 1.7: Learning of a Gaussian input using a Hebbian symmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the mean frequency of each oscillator and (c) the weighted connectivity matrix (sorted by mean frequencies and phases). Parameters: $\alpha = 0$ and $\beta = -0.5\pi$.

Other types of stimulation have been considered, such as the uniform distribution (see Fig. A.5 and Fig. A.6 of Appendix A.2.0.1), without bringing further relevant findings. Yet, the main limitation of the current approach is the need to keep applied the external input in order to maintain the particular dynamics of each cluster and therefore the associated structural connectivity.

1.3.2 Synchrony-based adaptation

Instead of relying on the frequency of the inputs, a second possibility is to consider the same inputs applied at a particular moment and to a specific subset of oscillators as explained in Fig. 1 in the *Introduction*. In other words, it is the timing and the area stimulated that create correlations rather than the intensity (or frequency) of the input applied as in the previous approach. While the disparity in oscillator frequencies remains the fundamental distinguishing factor, the main focus here is to explore the notion of stimulated and unstimulated regions. In this way, we predict that oscillators that receive a common input at a given time will reinforce each other, while oscillators that are active at different moments will tend to decouple. Thus, this approach is based on the synchrony of the inputs applied. This approach is summarized in Fig. 1.8.

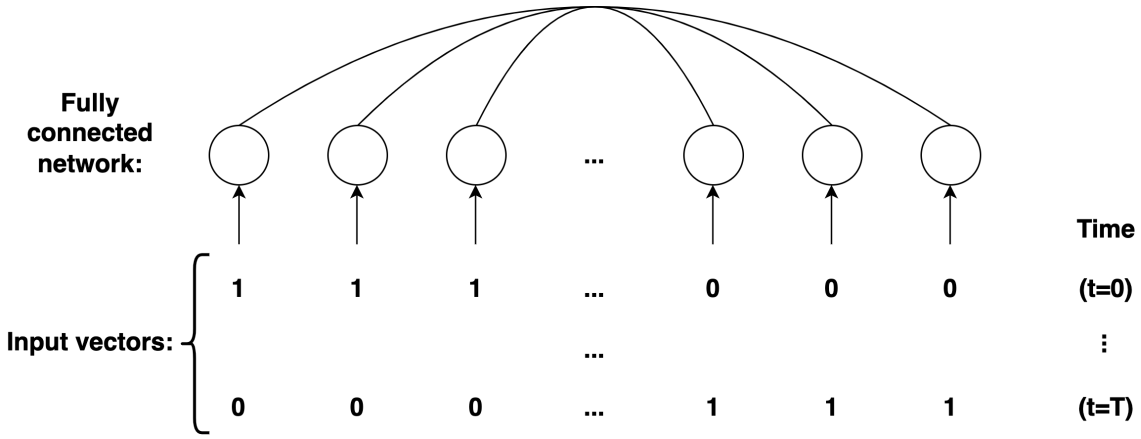


Figure 1.8: Diagram of the synchrony-based adaptation with two different areas stimulated separately. One part of the oscillators is stimulated (with a value of 1) for a duration $T-1$ while the other part is not. After time T, the second part is stimulated while the first one is not.

Unlike the previous approach, here, the external inputs do not remain active all the time. However, when the stimulation is stopped, the coupling weights continue to adapt fairly rapidly, which has the effect of losing the organization of the weights created by the inputs. To overcome this problem, we decided at this stage to freeze the adaptation of the weights when no input is applied, as it is common practice in AI. In this way, we consider that it is the presence of an external input in the pre-synaptic oscillator that initiates learning. This could be justify by the fact that an input directly increase the frequency (or firing rate) of an oscillator as in a biological neuron, which indirectly promotes faster adaptation of the weights. Conversely, for a short time scale, adaptation can be considered negligible in the absence of stimuli. Therefore, the weights dynamics of Eq. (1.2) is now given by Eq. (1.9):

$$\frac{d\kappa_{ij}}{dt} = \varepsilon H(I_j)(-\kappa_{ij} - \sin(\theta_i - \theta_j + \beta)), \tag{1.9}$$

with I_j the input associated to the jth oscillator and the Heaviside function $H(x) = 1$ if $|x| > 0.1$, $H(x) = 0$ otherwise. The Heaviside function acts as a threshold, allowing adaptation only if a sufficiently large input is presented (up to 0.1).

Four inputs To evaluate this approach, we reproduce the protocol described in Fig. 1.8 by considering four different stimulated areas. Thus, the input vectors are composed of binary values localized on four different areas (group of oscillators) and applied alternatively and randomly to the network. After this learning phase, the stimulation are stopped and we observe the dynamics and connectivity of the network. The results obtained with the Hebbian symmetric rule are shown in Fig. 1.9 and those with the Hebbian asymmetric rule in Fig. A.7 of Appendix A.2.0.2. As in Fig. 1.6, we can clearly distinguish structural modules (in this case four) in the weighted connectivity matrix (e) corresponding to the

four different areas stimulated at distinct times as shown in graph (b), where the mean frequencies of clusters of oscillators vary during the learning phase (until iteration 20000) due to the change in stimulation. Within each cluster, we find again the structure found in Fig. 1.2 associated with the Hebbian symmetric rule. Regarding the dynamics of the network after learning (i.e. no external inputs applied), we observe the same dynamics as those observed in the state of the art (see Fig. 1.2) with a two-cluster state (see panels (a), (c) and (d)) brought by the Hebbian symmetric rule. Thus even if the structural clusters are preserved (the weights being frozen), they are not visible in the network dynamics. Indeed, in the absence of inputs, the network converges back to its initial state. Despite this, this experiment confirms our expectations that it is possible to create modular structures using only the spatio-temporal correlations of the inputs, which could pave the way for more complex structures by adjusting the position and duration of activation of the inputs.

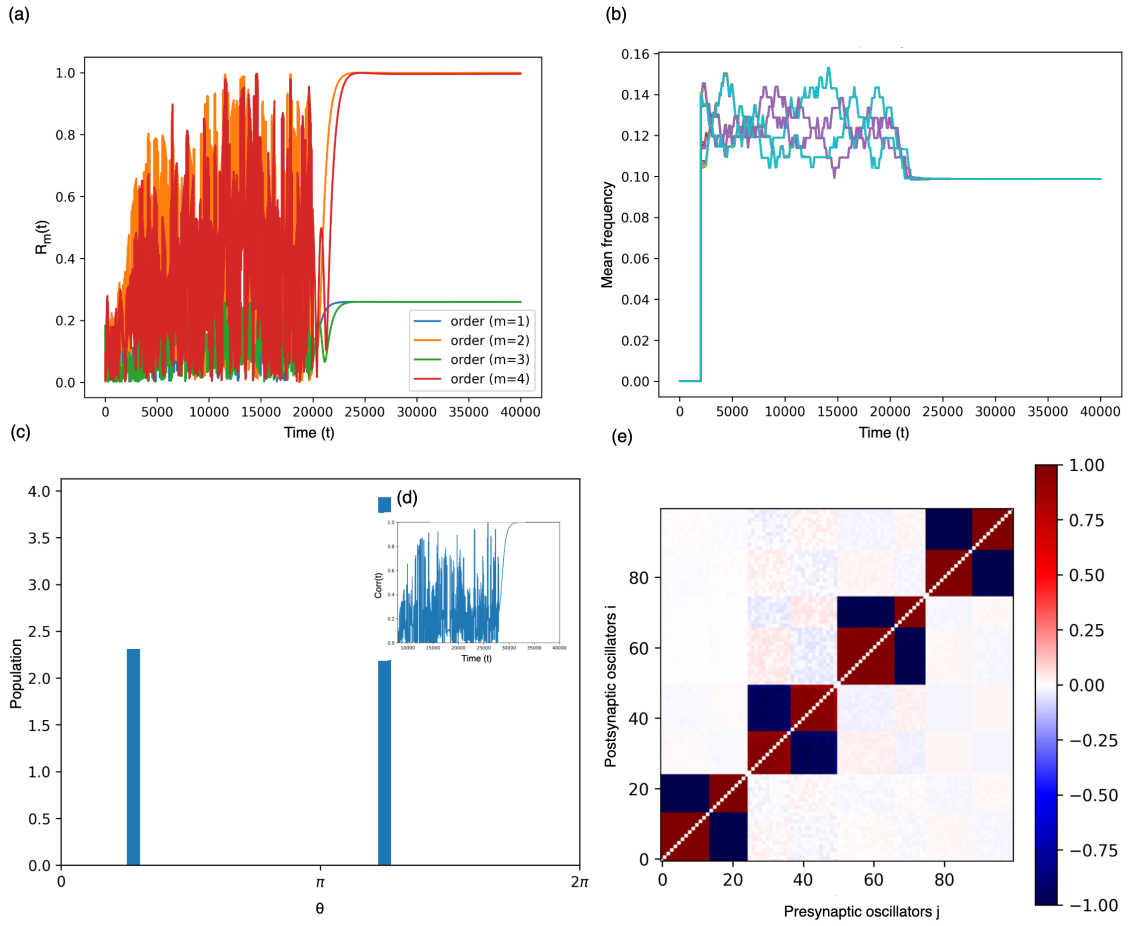


Figure 1.9: Learning of four binary inputs randomly selected using a Hebbian symmetric rule with: (a) the time development of the order parameters, (b) the mean frequencies of each oscillator through the time, (c) the distribution of neurons ($N = 100$) by phases (logarithmic scale) after learning phase, (d) the autocorrelations of the phase pattern and (e) the weighted connectivity matrix (sorted by phases in each cluster). Parameters: duration of learning = 20000 iteration steps, $\alpha = 0$ and $\beta = -0.5\pi$.

Two overlapping inputs So far, we have only considered the case in which the stimuli were independently applied, oscillators were only involved in one of the stimuli. Now, we introduce overlapping stimuli, meaning that some oscillators are targeted by both inputs. Except this small difference, the protocol is analogous to the previous one. The results obtained with the Hebbian symmetric rule are shown in Fig. 1.10 and those with the Hebbian asymmetric rule in Fig. A.8 of Appendix A.2.0.2. As in Fig. 1.9, we observe the formation of structural clusters (here two) in the weighted connectivity matrix (e) corresponding to the two different stimulated areas. However, these two clusters are connected to hub oscillators, an effect of the overlapping stimuli and the constant activation of these particular oscillators during the learning (see mean frequencies in graph (b)). Consequently, these

hubs receive and project connections to both clusters and, in this sense, belong to both groups. We note that, although the weighted connectivity matrix has not been sorted in this case, it seems that the same structure patterns as in Fig. 1.2 emerge within the clusters and the hubs. This is confirmed by the dynamics of the network after learning, which is the same as in the previous experiment, and by extension as in the original case of Fig. 1.2, with a two-cluster state (see panels (a), (c) and (d)). Once again, the structure learned does not seem to have any impact on the dynamics of the network at rest (i.e. without external stimulation). Nevertheless, this experiment provides some insights on how hub nodes could emerge in a network as the result of adaptation, as stated in the *Introduction*, again using the spatio-temporal correlations of the applied inputs.

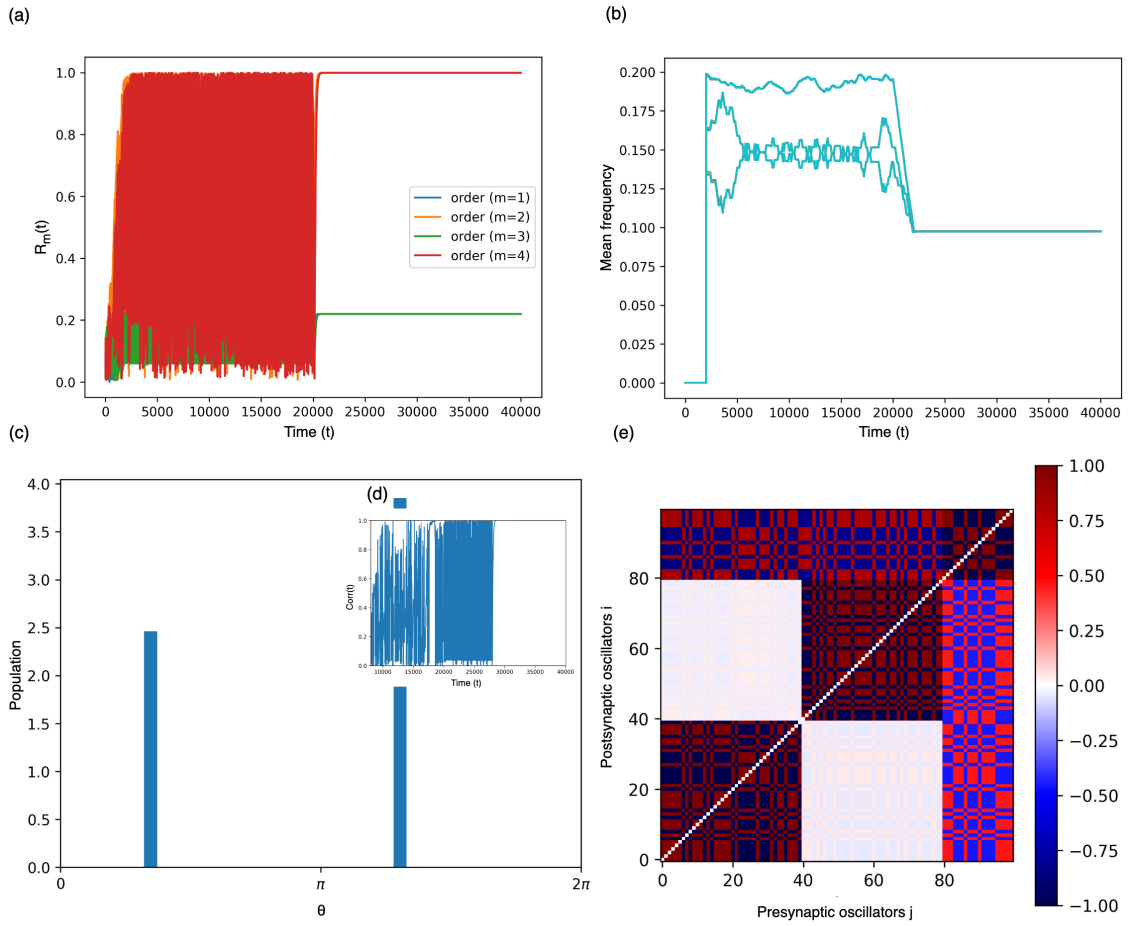


Figure 1.10: Learning of two overlapping binary inputs randomly selected using a Hebbian symmetric rule with: (a) the time development of the order parameters, (b) the mean frequencies of each oscillator through the time, (c) the distribution of neurons ($N = 100$) by phases (logarithmic scale) after learning phase, (d) the autocorrelations of the phase pattern and (e) the weighted connectivity matrix. Parameters: duration of learning = 20000 iteration steps, $\alpha = 0$ and $\beta = -0.5\pi$.

We can go even further and consider this approach for learning to recognize simple binary patterns associated with a label in the same way as in machine learning, as shown in Fig. A.9 of Appendix A.2.0.2. It is also possible to combine the two approaches described above by using input vectors containing values of different amplitudes, as shown in Fig. A.11 of Appendix A.2.0.2.

Through this synchrony-based adaptation, we were able to shape structural clusters and other complex structures in networks made of Kuramoto oscillators thanks to the spatio-temporal correlations of the applied inputs. Unlike the frequency-based adaptation, the stimulation is not always present and can be interrupted after learning. However, these learned clusters are not associated with a particular network dynamics at rest, which may be a problem for the long-term maintenance of these structures if permanent adaptation is considered after the learning phase is finalised.

1.4 Summary and conclusions

In this chapter, we have reviewed and identified the technical constraints we will face and the approaches we will be using in the following chapters. Previous works have shown how to model brain oscillations using oscillator models, emphasizing the role of adaptation and the connectivity induced in these dynamics. These latter, combined with inhibition, are key factors in the emergence and modulation of synchronization, an essential mechanism for good cognitive functions in biological networks. In particular, by reproducing the works of Aoki et al. [12, 13] in Sec. 1.2, we show the role of plasticity in the emergence of collective behaviours in the dynamics of a network of coupled Kuramoto oscillators subject to synaptic adaptation. We then extended these results to form structural clusters through the action of external stimuli. Two adaptation approaches are considered: (1) a frequency-based adaptation where the different amplitudes of the input shape the structures and (2) a synchrony-based adaptation where the timing of the applied inputs forms the structures.

These two strategies can echo the rate-based and spike-based learning found in ANN and biology. Rate-based learning, which relies on the average activity (i.e. firing rate) of neurons, is suitable for time scales of 100ms or longer [212]. To detect input correlations on a time scale of the order of milliseconds or less, learning is based on the temporal coincidences between spikes (i.e. action potentials) of the neurons [212, 213]. Even if the latter provides better stabilization of weights and neuronal information processing [212, 213], stimulation frequency keeps a significant impact, suggesting a link between time- and frequency-dependent approaches [87]. Therefore, although both strategies have their advantages and disadvantages, they are not incompatible as shown in the previous section. Thus, as a first main objective, these two approaches allow us to **highlight the aspect of segregation by proving the emergence of modular and complex structures** in the weighted connectivity through external stimuli.

However, a number of challenges remain to be addressed for a more bio-realistic architecture. Indeed, reservations can be expressed about the realism of the dynamics of the oscillators. When stimulation is stopped, whatever the approach used, **the network returns to its original state** (i.e. two-cluster state or splay state according to the plasticity

rule). Thus, there is no particular dynamics associated with the structural clusters created and therefore no way of observing the learned items this dynamics. This gives rise to the second problem of this model, which is **the maintenance of these structures**. Since the dynamics of the network remain unchanged, the connections between the modules should re-emerge under the effect of plasticity. To address this problem, we decided to freeze the adaption of weights in the absence of external stimuli. However, from a biological point of view, this may raise questions if we consider that adaption always performs in the brain. In this sense, the current model does not display a real active maintenance of the learned structure. Finally, one of the criticism that can be made on the model is that there is **no distinction made between inhibitory and excitatory units** and that, consequently, an oscillator can have positive and negative outgoing weights, which is not consistent with biology [97]. The next chapter will attempt to address all these aspects.

Chapter 2

Maintenance of neural assemblies in networks of θ-neurons through inhibition

Contents

In the previous chapter, we saw how to form structural clusters with Kuramoto oscillators subject to adaptation thanks to the action of external stimuli. In this chapter, we extend these results to θ-neurons while providing additional biological insights in order to overcome certain challenges encountered previously. In particular, the focus will be on the maintenance of these modular structures and the role of inhibition in this process.

2.1 Introduction

Inhibition has an essential role for the dynamics and the adaptation ability of the brain [182, 103]. Inhibitory interactions are mediated by the GABAergic neurotransmitters whose effect is to reduce the probability of their target neurons to emit spikes. This is usually true in the adult brain operating in normal conditions [34]. Inhibition has been shown to be fundamental for perceptual decision making [320] as well for the emergence of brain rhythms [411]. It also plays a crucial role in adaptation processes of the brain, in order to shape and preserve the structure of the network according to the correlations between

inputs [213, 215, 7]. Indeed, it has been proposed that the interconnections between excitatory and inhibitory cells play a relevant role balancing between fast adaptation and long-term conservation of the memories in neural circuits [273, 162, 401].

Many models have been proposed in the literature to study the properties of plastic neural networks that account for excitation and inhibition. For example, as studied in the previous chapter, models that treat neurons as oscillators have been useful to explore the synchronization phenomena of the networks as a function of the ratio between excitation and inhibition [361, 243]. Synaptic plasticity is usually mimicked by introducing adaptive coupling between the oscillators [12, 13, 40], this can promote the emergence of multi-stable regimes (see Sec. 1.2) characterized by synchronized and desynchronized clusters of oscillators [131, 347, 40, 39]; analogous to observations in the presence of STDP [253, 269].

While the simplicity of these models favours their analysis, they typically omit several biological aspects. Two major omissions are shared with common models for ANN employed in AI and machine learning. On the one hand, excitation and inhibition are considered at the level of individual links and thus no distinction is made between excitatory and inhibitory neurons [185, 12, 13, 40]. Despite the fact that from a biological point of view Dale's principle [97] requires that the nature of the neurons should be uniquely defined such that all post-synaptic connections of a neuron are either excitatory or inhibitory. On the other hand, plasticity is only accounted for during the learning phase in which the networks are entrained. After the training, the resulting synaptic weights are generally frozen as we did in Sec. 1.3.2. However, in biological neural networks adaptation is constantly active raising the question of how could memories be consolidated in a network that is susceptible to permanent change.

The aim of this chapter is to study the role of inhibitory neurons in the emergence and consolidation of neural assemblies due to stimulus-driven plasticity. Therefore, we model **networks of excitatory and inhibitory neurons** (represented as θ-neurons [132]) shaped by external stimuli applied to differentiated subsets of neurons (see Sec. 2.2). The formation of assemblies is obtained thanks to a symmetric Hebbian-like PDDP rule allowing the correlated neurons to reinforce their synaptic weights. Our results (see Sec. 2.3) show that **the presence of inhibitory neurons is crucial not only for the formation of clusters of neurons triggered by the stimuli, but also for the consolidation and the recall of memories**. We find that violation of the biological constraints, either by constructing **networks of excitatory-only neurons or by omitting Dale's principle, leads to networks unable to maintain these memories**. Furthermore, we show that **the memory capacity of the network is directly related to the number of inhibitory neurons** and that the conservation of the inhibitory synaptic weights is sufficient for memory recalls even if the synaptic patterns associated to the excitatory neurons are forgotten. The robustness of the results is tested against variations of the entrainment protocol and replacing the θ-neuron model by other models of oscillators such as the Kuramoto and the Stuart-Landau models.

2.2 Methods

This section describes the model employed for the dynamics of the membrane potentials of the neurons in the network, the learning rules governing the weight adaptation and the protocols followed to induce structural patterns and their consolidation, as well as the protocols used to analyse the stability of the emerged patterns.

2.2.1 Neuronal dynamics

To mimic the membrane potential evolution of a neuron we employ the θ-neuron model introduced by Ermentrout and Kopell as the normal form for class I excitable cells [132]. This model has the advantage of reproducing quite faithfully the behaviour of spiking neurons through the description of the evolution of the corresponding "membrane potential" while remaining a phase model [130]. Indeed, it can be put in an one to one correspondence with the Quadratic Integrate and Fire (QIF) spiking neuronal model via a non-linear transformation linking the phase to the membrane potential [153]. Furthermore, the θ model can describe bursting dynamics observable at the level of the membrane potential of specific cells of the Aplysia mollusc [130, 132]. In addition, networks of coupled θ-neurons can exhibit a variety of behaviours ranging from asynchronous regimes to multi-stability, from partial synchronization to chaos [246, 348, 229, 47]. In a network of $N = N_E + N_I$ neurons, where N_E (N_I) is the number of excitatory (inhibitory) neurons, the dynamical evolution of the phase θ_i of a neuron ($i = 1, \dots, N$) is governed by the following equation:

$$\frac{d\theta_i}{dt} = (1 - \cos\theta_i) + (1 + \cos\theta_i)\Big[\eta_i + \frac{g}{N}\Big(\sum_{j=1}^{N} \kappa_{ij} \sin(\theta_j - \theta_i)\Big) + I_i(t) + \xi_i(t)\Big], \qquad (2.1)$$

where η_i is the excitability or the bifurcation parameter controlling the intrinsic frequency of the neuron, g represents a global synaptic strength modulating the overall level of coupling among the neurons, and κ_{ij} is a matrix term measuring the relative synaptic weight from the pre-synaptic neuron j to the post-synaptic neuron i. The latter term is subject to a temporal evolution, therefore we can speak of adaptive coupling, the evolution rules for the terms κ_{ij} will be introduced in the next sub-section. Furthermore, we consider a diffusive sinusoidal coupling function based on the phase differences as in the Kuramoto model [226] employed in Sec. 1.2.1. In other words, the couplings between neurons are considered as electrical gap junctions. The phases θ_i are defined in the range $[-\pi, \pi[$ and we assume that whenever the phase θ_i reaches the threshold π a spike is emitted by the i-th neuron.

In the same manner as in Sec. 1.3 for the formation of structural clusters with oscillators, external inputs are incorporated via the current terms $I_i(t)$ representing time-dependent stimuli. Positive external currents increase the frequency of the neurons, thus imitating the typical biological response of sensory neurons exposed to stimuli of varying intensity. Finally, an independent Gaussian noise $\xi_i(t)$ term is applied to each neuron mimicking background noise always present in real neural circuits.

Simulations are performed by integrating Eq. (2.1) via an Euler scheme, where the multiplicative stochastic term is treated in the Stratonovich sense [284], with an integration time step $dt = 0.01$.

2.2.2 Learning and adaptation

Plasticity function The synaptic weights κ_{ij} are modified according to a symmetric Hebbian-like rule depending on the instantaneous phase differences between the neurons instead of spike timings as observed in real neurons [257]. Since θ-neurons are phase oscillators this choice allows to have a relatively precise temporal adaptation – as compared to a frequency based rule – that can be numerically integrated simultaneously with the neuronal dynamics. A symmetric plasticity facilitates the convergence of the weights and highlights better the correlations with respect to the inputs as observed in Sec. 1.2.2. Nevertheless it shall be noted that a symmetric plasticity is possible in this case because the transmission delays are ignored [250, 251, 135].

We first consider the plasticity function introduced by Aoki et al. [12, 13] and recently employed in Berner et al. [40, 39]. Given the phase difference $\Delta\theta = (\theta_j - \theta_i)$, this plasticity function is defined as:

$$\Lambda_0(\Delta\theta) = \cos(\Delta\theta), \tag{2.2}$$

see Fig. 2.1, blue line. This function is analogous to the function $-\sin(\Delta\theta\frac{-\pi}{2})$ employed in Fig. 1.1 (a) in the previous chapter. This rule follows again the principle of Hebbian learning with a reinforcement (depression) of the weights when the neurons are in phase (in anti-phase). However although simple, the relevance in terms of realism of this function can be questioned due to the symmetry of the positive and negative parts, corresponding to long-term potentiation (LTP) and long-term depression (LTD), respectively. Indeed, in biology LTP and LTD do not act on symmetric time windows [46, 106]. To remedy this, we propose a plasticity function inspired by Lucken et al. [245] and Shamsi et al. [338]. We define the new plasticity function as:

$$\Lambda_1(\Delta\theta) = \begin{cases} e^{\frac{\Delta\theta}{0.1}} - e^{-\frac{\Delta\theta+\pi}{0.5}}, & \text{if } -\pi \leq \Delta\theta < 0 \quad , \\ e^{-\frac{\Delta\theta}{0.1}} - e^{\frac{\Delta\theta-\pi}{0.5}}, & \text{if } 0 \leq \Delta\theta < \pi \quad ; \end{cases} \tag{2.3}$$

where we assume that $\Lambda_1(\Delta\theta)$ is 2π periodic as shown in Fig. 2.1, red curve. This rule follows the same general principle of Hebbian learning as the one in Eq. (2.2) but the cosine is replaced by a combination of exponential functions making the potentiation (where $\Lambda_1(\Delta\theta) > 0$) and the depression (where $\Lambda_1(\Delta\theta) < 0$) phase intervals asymmetric. Since for phase oscillators, the phase evolution can be mapped into a time evolution, these asymmetric phase intervals correspond to asymmetric time windows. Indeed, this asymmetry is more consistent with the biological features: as a matter of fact, in the cortex depression operates over a longer time window with respect to potentiation [136, 345]. Specifically in [106], the authors have shown that for CA3 pyramidal cells in rat hippocampus, the

STDP presents a depression time scale five times longer than the potentation one, analogously to our choice in Eq. (2.3).

In the following, unless stated otherwise, we will employ the function Λ_1 and for simplicity we will refer to it as $\Lambda(\Delta\theta)$.

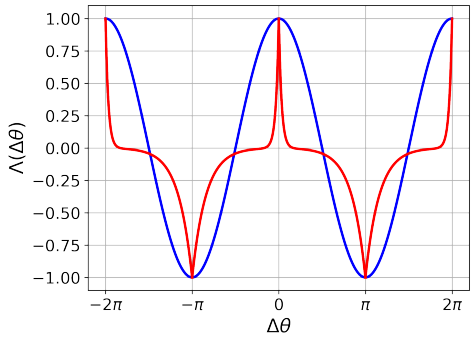


Figure 2.1: Hebbian phase difference-dependent plasticity functions $\Lambda(\Delta\theta)$ versus $\Delta\theta$. In blue the function Λ_0 reported in Eq. (2.2) and in red the function Λ_1 defined Eq. (2.3). Despite having different potentiation and depression phase windows, the two functions attain the same maximum and minimum values.

Slow and fast adaptation We have seen that external stimulations have a direct impact on the adaptation of the synaptic weights. In presence of stimulation the synaptic weights should be modified quite fast to retain the information contained in the stimuli. However, in the absence of further stimuli, a fast adaptation may allow for a rapid relaxation of the weights towards their equilibrium values leading to a loss of the learned input-driven memories. This problem is common to most learning neural network models. The solution typically adopted and implemented in Sec. 1.3.2 consists in freezing the weights once the stimuli have been delivered. There are indications that two learning processes may occur simultaneously or sequentially in the brain. An example, is represented by the complementary mechanisms involved in short-term or long-term depression and potentiation of the synapses [372, 122, 345, 118]. There are also evidences of different learning schemes on proximal and distal dendrites of pyramidal neurons [150] or in the retina [25]. Finally, it is also worth to mention that long-term potentiation may be modulated by homeostatic plasticity [398, 151]. Two learning processes also occur when acquiring new skills: a fast learning in a cortical structure is simultaneously slowly learned in a subcortical structure (habit learning for instance) [89, 301, 279, 275]. Thus, according to previous experiences and the nature of inputs, the rate of learning may need to evolve so that neurons adapt more or less quickly to sensory inputs [25, 139, 89, 85]. More precisely, adaptation in the brain is the result of multiple learning processes acting at distinct time scales allowing more or less rapid learning or relearning of new information without forgetting older memories [424, 89]. To somehow mimic this phenomenon, we introduce two complementary learning time scales:

- a fast time-scale that activates in the presence of an external stimulus presented to the pre-synaptic neuron;
- a slow steady time-scale that has a minimal short-term effect.

The fast and slow time scales have been introduced in this phase model, where adaptation depends on the phase differences and not on the spiking times, to reproduce a fast learning period induced by the higher firing rate of the stimulated neurons and a slower consolidation phase, where the neuronal firing becomes sparse.

From Eq. (1.2) of the previous model, given the plasticity function in Eq. (2.3) and the constraints just mentioned, we define the evolution of the coupling weights $|\kappa_{ij}|\leq 1$ from pre-synaptic neuron j to post-synaptic neuron i to be governed by the following equation:

$$\frac{d\kappa_{ij}}{dt} = \big[\varepsilon_1 + \varepsilon_2 H\left(|I_j(t)| - 0.1\right)\big]\big[-\kappa_{ij} + \Lambda(\Delta\theta)\big], \tag{2.4}$$

where $I_j(t)$ is an external input to neuron j, $H(x)$ is the Heaviside function such that $H(x) = 1$ if $x > 0$ and $H(x) = 0$ otherwise, and $\varepsilon_1 \ll \varepsilon_2 \ll 1$ are the learning rates for the slow and the fast adaptations, respectively. The Heaviside function has a thresholding effect allowing here fast adaptation only if a sufficiently strong stimulus $|I_j(t)| > 0.1$ is presented.

Excitatory and inhibitory neurons In order to account for the excitatory and inhibitory nature of the neurons, we disentangle Eq. (2.4) into two different cases, where the synaptic weights are bounded between $[0, 1]$ ($[-1, 0]$) provided the pre-synaptic neuron is excitatory (inhibitory). The dynamics of the weights are thus governed by the following equations.
If both neurons i and j are excitatory, then:

$$\frac{d\kappa_{ij}}{dt} = \big[\varepsilon_1 + \varepsilon_2 H(|I_j(t)| - 0.1)\big]\,\kappa_{ij}\,(1 - \kappa_{ij})\,\Lambda(\Delta\theta)\ , \tag{2.5}$$

if the pre- and/or post-synaptic neuron is inhibitory the weight evolution is given by:

$$\frac{d\kappa_{ij}}{dt} = \varepsilon_1|\kappa_{ij}|\big(1 - |\kappa_{ij}|\big)\Lambda(\Delta\theta)\ , \tag{2.6}$$

where $0 \leq \kappa_{ij} \leq 1$ ($-1 \leq \kappa_{ij} \leq 0$) if the pre-synaptic neuron is excitatory (inhibitory). The non-linear dependence on κ_{ij} introduced in Eqs. (2.5) and (2.6) is intended to mimic the soft bounds often employed in the implementation of the STDP in spiking neural networks [346] and to maintain $|\kappa_{ij}|$ smaller than one.

Notice that the fast adaptation is only present for synapses connecting excitatory neurons since external inputs are only applied to this type of neurons [30, 274, 208, 415]. Indeed, sensory inputs must be rapidly assimilated in the short-term while being volatile in order to learn new information [273, 89, 242]. Therefore, synaptic weights involving inhibitory neurons will always only evolve on the long time scale. By default, we consider a

ratio of 80% of excitatory neurons and 20% of inhibitory neurons as is commonly accepted in the human cortex [215, 243, 273].

Lastly, we expect that for synapses connecting two excitatory or inhibitory neurons i and j, the weights κ_{ij} and κ_{ji} will become equal in the long run as observed in Sec. 1.2.2 with the Hebbian symmetric rule, due to the symmetry of Eq. (2.5) in the absence of stimulation and of Eq. (2.6). However, this should not be the case if one of the neurons is inhibitory and the other excitatory.

2.2.3 Stimulation protocols

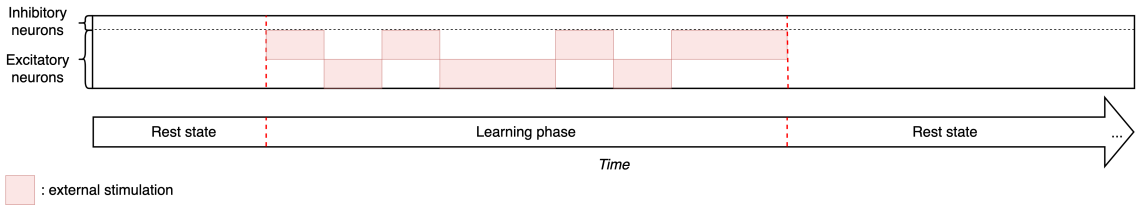


Figure 2.2: Diagram representing the stimulation protocol divided in three phases: a first short phase of spontaneous activity, a learning phase during which two different groups of excitatory neurons are stimulated alternatively over time (the red areas represent the duration and the neurons that received external stimulation), and a final long period of resting-state activity in which the plasticity remains active.

Table 2.1: Parameters for the network of θ-neurons

Parameters	Values
N	100
N_E	80
N_I	20
η	$\mathcal{N}(1.5,\, 0.01)$
g	1
$I(t)$	$\{0, 3\}$
$\xi(t)$	$\mathcal{N}(0.0,\, 0.1)$
ε_1	0.00001
ε_2	0.1
dt	0.01

Main experiments We now summarise the principal protocol followed in the chapter. This one will be based mainly on the synchrony-based adaptation developed in Sec. 1.3.2 in the previous chapter and on the diagram of Fig. 1 in the *Introduction*. Indeed, this approach allowed the easy formation of modular structures thanks to the spatio-temporal correlations of the applied inputs.

The networks are initialised with random weights uniformly distributed in the range $\kappa_{ij} \in [0, 1]$ when the neuron j is excitatory and in the range $\kappa_{ij} \in [-1, 0]$ when the neuron j is inhibitory. We consider heterogeneous neurons with excitabilities η_i randomly distributed according to a normal distribution $\mathcal{N}(1.5,\ 0.01)$. The initial phases of the neurons are randomly selected in the interval $[-\pi, \pi]$. All network parameters are summarized in Table 2.1.

The stimulation protocol is illustrated in Fig. 2.2. Initially, a period of spontaneous activity is considered in order to allow the system to relax to its rest state. Afterwards, a learning phase follows during which two external positive currents are applied randomly and intermittently for constant periods of 20 time units. One current stimulates the first half of the excitatory neurons ($i = 1, 2, \ldots, 40$), and the other one the second half ($i = 41, 42, \ldots, 80$). Inhibitory neurons are never directly stimulated. Finally, after the learning phase the network is left to evolve spontaneously for a long time. During this post-learning phase no neurons are stimulated and therefore only the slow adaptation rate ε_1 remains active, affecting the stabilization of the connectivity patterns formed during the stimulation phase.

This principal experiment was repeated for different variations. In one case, we ignored inhibition and considered networks of only excitatory neurons. In a second case, in order to show the relevance of Dale's principle for pattern consolidation, i.e. the need to preserve the nature of the excitatory and inhibitory neurons during the simulation, we considered unlabelled neurons. In this case neurons were allowed to display both excitatory and inhibitory synapses in the same way as in Chapter 1. Specifically, for the experiment with unlabelled neurons the synaptic adaptation was ruled by Eq. (2.4) with the Hebbian function $\Lambda_0(\Delta\theta)$ in Eq. (2.2).

Additional experiments Thereafter variations of the principal protocol are considered, as follows:

1. three external stimuli during the learning phase instead of two;

2. an overlap of 8 neurons among two groups of excitatory neurons stimulated by two external stimuli;

3. a recall experiment of previously stored items, where the trace of the memory storage is maintained only in the connections involving inhibitory neurons.

Other variations of the main protocol are described in Appendix A.3.

To understand the generality of the results obtained in the additional experiments (1) and (2), we analyse the stability of different configurations. In particular, we perform a stability analysis where the objective is to determine the conditions under which the network connectivity during the resting-state remains stable (i.e. the learned modular structure is preserved) or unstable (i.e. at least one structural cluster is not maintained) over the long term. This stability analysis corresponds in experiment (1) to the maintenance of all the

stored clusters and in experiment (2) to the maintenance of the two stored clusters despite their increasing overlap.

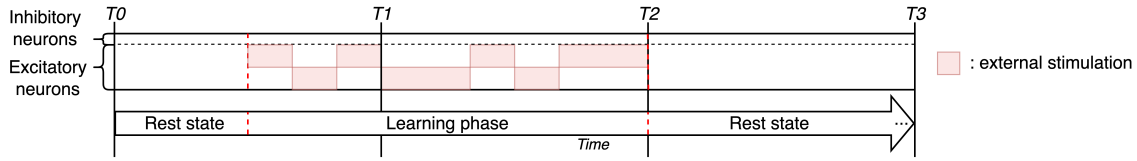
In order to perform this analysis, we start from already trained networks satisfying the desired constraints (i.e. the number of structural clusters or the number of overlapping neurons between clusters versus the number of inhibitory neurons), thus facilitating the replication of experiments under different conditions. Moreover, this pre-training facilitates the analysis of extreme cases where for instance a single inhibitory neuron is associated with each cluster whose emergence is not necessary guaranteed via a random learning. Then we leave the network in a state of spontaneous activity for a long time and compare the final structural state to which it converges with the initially stored one to assess the stability or instability of this latter configuration.

2.3 Results

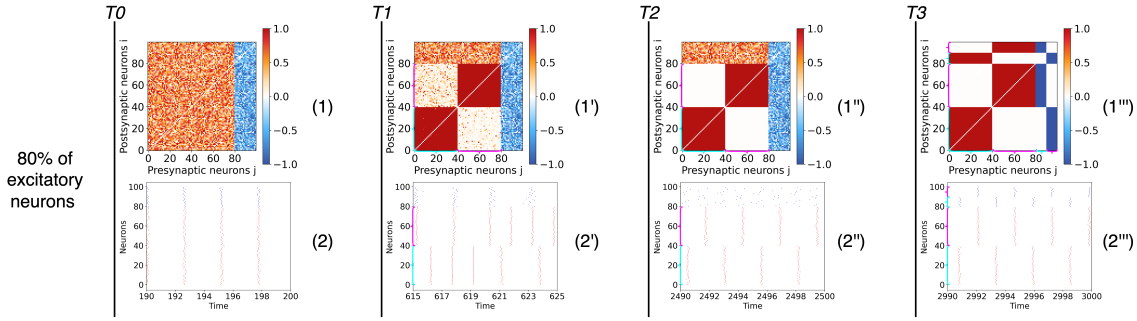
The goal of this chapter is to study the emergence and consolidation of modular architectures induced via a learning process promoted by localised inputs which target distinct subsets of neurons. We will first present results concerning the emergence and consolidation of clustered architecture induced by learning of two external stimuli. This experiment is repeated for networks composed of different types of neurons: namely, purely excitatory neurons, excitatory and inhibitory neurons, and unlabelled neurons. Then, we will study variations of the initial protocol and in particular the learning process promoted (i) by multiple stimuli and (ii) by overlapping stimuli, where neurons can encode for multiple items. Lastly, we investigate the ability of the neural networks to recall the stored memories.

2.3.1 Learning and consolidation of modular assemblies

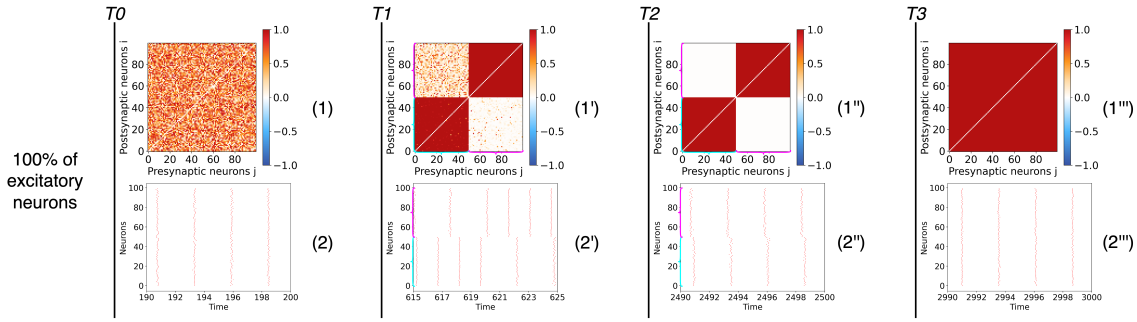
We start by investigating the emergence of two modules due to stimulations of non-overlapping neural populations. We first study this phenomenon in networks containing excitatory and inhibitory neurons and then we explore cases with only excitatory neurons or with unlabelled neurons whose connections can be either excitatory or inhibitory as in the first chapter.


Inhibitory neurons
Excitatory neurons
T0
T1
T2
T3
: external stimulation
Rest state
Learning phase
Rest state
Time

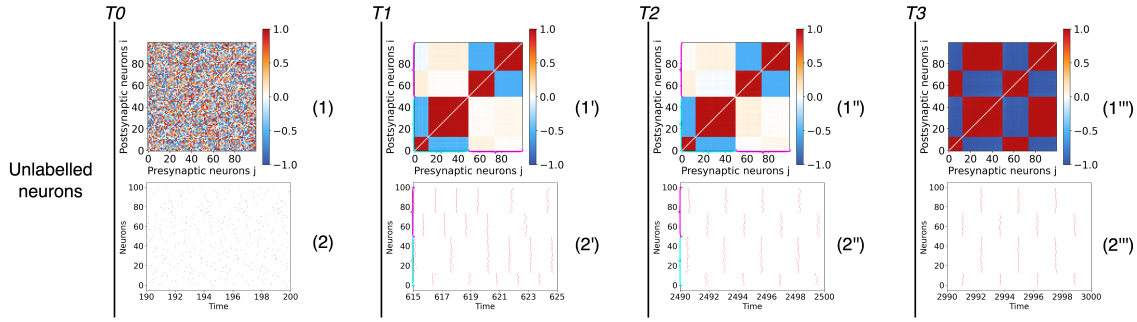
(a)


80% of excitatory neurons
T0
T1
T2
T3
Postsynaptic neurons i
Presynaptic neurons j
Neurons
Time
(1)
(1')
(1'')
(1''')
(2)
(2')
(2'')
(2''')

(b)


100% of excitatory neurons
T0
T1
T2
T3
Postsynaptic neurons i
Presynaptic neurons j
Neurons
Time
(1)
(1')
(1'')
(1''')
(2)
(2')
(2'')
(2''')

(c)


Unlabelled neurons
T0
T1
T2
T3
Postsynaptic neurons i
Presynaptic neurons j
Neurons
Time
(1)
(1')
(1'')
(1''')
(2)
(2')
(2'')
(2''')

(d)

Figure 2.3: Entrainment of networks of θ-neurons defined by different combinations of excitatory and inhibitory connections. (a) Experimental protocol consisting of the stimulation of two non-overlapping neuronal populations of θ-neurons with plastic synapses. Stimuli are presented in temporal alternation. (b) Results for a network with 80% excitatory and 20% inhibitory neurons. (c) Results for a network of only excitatory neurons and (d) for unlabelled neurons projecting both excitatory and inhibitory post-synaptic connections. The results are reported at different moments of the stimulation. The time T0 corresponds to the beginning of the simulation before the learning phase, after a transient period $t_transient = 190$ has been discarded. The time T1 corresponds to an early moment during the learning phase. Time T2 corresponds to the end of the learning phase and the beginning of the resting-state. The time T3 corresponds to the end of a long period of spontaneous activity during which synaptic weights are consolidated. Panels labelled (1), (1'), (1") and (1"') represent the weight matrices at times T0, T1, T2 and T3: the color denotes if the connection is excitatory (red) or inhibitory (blue) or absent (white). Panels (2), (2'), (2") and (2"') are raster plots at times T0, T1, T2 and T3, displaying the firing times of excitatory (red dots) and inhibitory (blue dots) neurons. Note that the inhibitory neurons are sorted by phases at time T3 in the weight matrices and raster plots reported in (b) and (c); while the neurons are sorted by phase in the weight matrices and raster plots within each cluster at times T1, T2 and T3 of the row (d). The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

Networks of excitatory and inhibitory neurons We consider a network of $N = 100$ θ-neurons subdivided in $N_E = 80$ excitatory and $N_I = 20$ inhibitory ones, initially connected via a random weighted matrix (uniform distribution between $[0, 1]$ or $[-1, 0]$ for pre-synaptic excitatory or inhibitory neurons, respectively), see *Methods* in Sec. 2.2. The network is entrained using two independent stimuli applied to separate subsets of neurons following the protocol shown in Fig. 2.3(a). The phases of the neurons are randomly initialised (uniform distribution between $[-\pi, \pi]$). The results of this experiment are reported in Fig. 2.3(b). During the initial rest phase the network is left to evolve spontaneously and it converges into a state close to synchrony despite the fact that the weight matrices at time T0 are randomly distributed, see raster plot (1). Both excitatory and inhibitory neurons tend to synchronize, as shown in the raster plot (2) at T0. This is due to the value of the chosen coupling strength g, which is sufficiently large to favour a synchronized phase.

After the period of spontaneous activity the learning phase starts. The resulting connectivity matrix and the network dynamics are shown at two intervals: during the learning phase – time T1 – and at the end of this phase – time T2. As we see, the presence of the two stimuli leads to the emergence of two modular structures (weight matrices (1') and (1")) among the excitatory neurons while the weights involving inhibitory neurons do

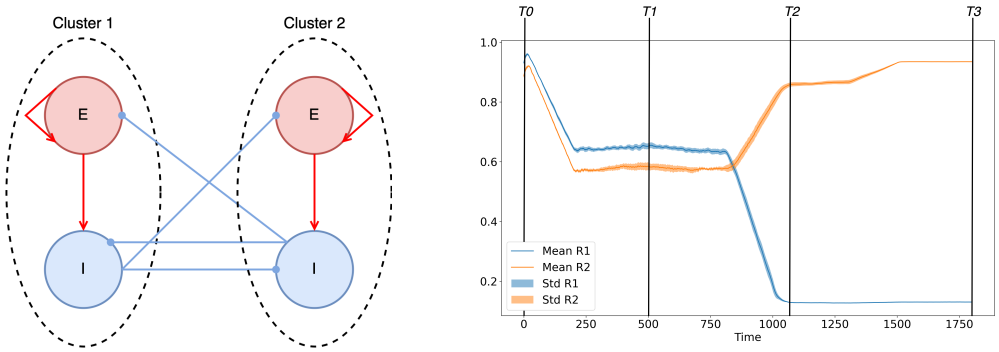
not evolve much due to the separation of fast and slow learning rates. The presence of an input leads to an increase in the firing rate of the stimulated neurons, see raster plot (2'). At the end of the learning phase, two disconnected clusters of excitatory neurons firing in anti-phase emerge in the network, while the inhibitory neurons remains uncorrelated (see raster plot (2")).

Following the learning period the network is left to evolve driven by its spontaneous activity. In this stage the adaptation is governed by the slow adaptation rate. The long-term results are shown in the time column T3. As seen, the learned structure is consolidated. The two excitatory clusters are maintained, see weighted connectivity matrix (1"'), and besides this the input/output inhibitory weights continue adapting such that the modular structure is reinforced. Finally, also the inhibitory neurons split in two structural clusters. The final consolidated connectivity structure is reported in Fig. 2.4(a). From this figure it emerges that the neurons are organised in two clusters, each one involving a group of excitatory neurons that projects in a feed-forward manner on a group of inhibitory neurons. Therefore the inhibitory neurons within the first cluster synchronise with the excitatory neurons driving them, as shown in panel (2"'). Furthermore, the inhibitory neurons of one cluster projects on the excitatory and inhibitory neurons of the other cluster. This induces a repulsion of the dynamics of the two clusters that adjust in anti-phase one with respect to the other, both clusters displaying exactly the same period of oscillation, see raster plot in panel (2"') in Fig. 2.3(b).

In order to validate the robustness of the results, the experiment was repeated several times (as all the following experiments) by randomly varying the presentation order of the stimuli to the two populations. At the same time, the initial values of the phases of the neurons and of the connection weights were also randomly assigned in each realization. From these data we have estimated the evolution in time of the modulus of the 1^{st} and 2^{nd} Kuramoto-Daido order parameters (as defined in Eq. (1.3) of Chapter 1). These indicators are useful to analyse the level of synchronization in the network and to test for the emergence of clusters at different stages during the experiment. The temporal evolution of the average order parameters and their standard deviations are shown in Fig. 2.4(b) for the previous experiment displayed in Fig. 2.3(b).

At time T0, we observe with a negligible variation that the network is almost totally synchronized, as testified by the fact that $R_1 \simeq 1$ and $R_2 \simeq 1$. During the learning phase (i.e. at time T1), the value of the two parameters drop and their standard deviations increase, thus testifying for a larger variability and a desynchronization of the network. After the learning at time T2, R_2 (R_1) increases (decreases) tending toward a large (small) value. These results indicate the emergence of two main synchronized clusters in the network corresponding to the two learned modules. Finally, after the long term consolidation (i.e. at time T3), the network converges towards two clusters in anti-phase where $R_2 \simeq 1$ ($R_1 \simeq 0$) and the corresponding standard deviations are negligible, thus confirming the stability of this final state, despite the different stimulation sequences employed in the performed experiments. These observations put in evidence that the formation and consolidation of input-driven modules are a robust outcome of the model, which depends mainly on the characteristics of the stimulated populations and on the number of inhibitory neurons in

the network, and is not significantly affected by the variability of the initial conditions nor of the particular sequence of the presented stimuli.



(a) (b)

Figure 2.4: (a) Schematic diagram representing the connectivity matrix emerging after the consolidation phase for the stimulation of two non-overlapping populations reported in Fig. 2.3. Red (blue) circles represent excitatory (inhibitory) populations, dashed circles identify clusters of synchronized neurons. (b) Evolution of the modulus of the 1^{st} (in blue) and 2^{nd} (in orange) Kuramoto-Daido order parameters Eq. (1.3) averaged over 10 different realizations randomised initial conditions and stimulation sequences. The mean and the corresponding standard deviation are shown versus time during the realization of the experiment in Fig. 2.3. The time labels have the same meaning as in Fig. 2.3.

In addition, further experiments were performed to validate the robustness of these results. In particular, we investigated the case where partially trained areas were not completely formed due to partial random stimulations and the case where a sub-group of neurons remains untrained (unstimulated), see Figs. A.13 and A.14 of Appendix A.3.

Networks of excitatory neurons Now we consider the case in which all neurons are excitatory and no inhibition is present. As before, the neurons synchronise during the initial stage of spontaneous activity and the training phase leads to the formation of two structural clusters (see, Fig. 2.3(c)). However, from the dynamical perspective the neurons in the two clusters are no longer firing in anti-phase, as in the previous case. Instead, at the end of the learning phase they present a small phase-shift and they indeed tend to synchronize as evidenced from the raster plot (2") at time T2. As a consequence, during the stabilization phase, all the synapses are reinforced leading to an all-to-all connectivity matrix (see matrix (1"') at time T3). Therefore the two structural modules that emerged during the learning are now completely forgotten.

Networks of unlabelled neurons As a last case, we consider the typical scenario of artificial neural networks, where neurons are unlabelled, meaning that each neuron can display both excitatory and inhibitory post-synaptic connections violating Dale's principle.

This case is very similar to the approach considered in Chapter 1 with Kuramoto oscillators, only the model for the dynamics and the double time scales for the adaptation are different. The results of this experiment are reported in Fig. 2.3(d). This time the initial stage of spontaneous activity leads to an asynchronous neuronal state (raster plot (2)) since every neuron randomly attracts and repulses its neighbours via excitatory and inhibitory links. This initial state resembles the chaotic state found in Fig. 1.4 with a similar weighted connectivity matrix. After the learning phase (time T2), two symmetric structural clusters emerge in the weighted connectivity matrix (1"). More precisely, the Hebbian rule in Eq. (2.2) leads to the creation within each structural module of two phase clusters in anti-phase, depending on the initial phases as explained in Aoki et al. [12, 13] and in the previous chapter. Once again, this result is analogous to the one described in Fig. 1.9 with the formation of structural clusters thanks to the synchrony-based adaptation, but without any proper dynamics associated with these clusters (i.e., only a two-cluster state provided by the plasticity function).

In contrast with the previous cases, the structural modules are now not well defined. Weak connections among the two clusters are still present as shown in (1"), allowing the possibility for the two clusters to synchronize, see the spike trains in (2"). During the consolidation phase these weak cross-modular connections increase, leading to a fully coupled matrix with no structural modules, see (1"'). However, even in the absence of structural modules one can logically observe clustered phases in the temporal evolution of the neurons as also reported in [40]. The neurons connected through excitatory (inhibitory) pre-synaptic links tend to fire in phase (anti-phase). In the end, the network return to its original stable state shown in Fig. 1.2. Nevertheless, it is important to bear in mind that these phase clusters do not reflect the two structural modules previously stored in the network.

In summary, we have shown that the presence of both excitatory and inhibitory neurons is needed for the formation and the consolidation of structural modules driven by the learning of independent stimuli. These results can be reproduced with oscillator models (i.e. Kuramoto and Stuart-Landau model) as shown in Fig. A.17 of Appendix A.4. In the following, we generalize these results to cases in which the number of stimuli – and therefore of expected modules – is arbitrary and to the case where the stimuli act on overlapping neuronal populations.

2.3.2 Learning multiple clusters

In this experiment, the previous learning protocol is repeated for networks with excitatory and inhibitory neurons but considering now three inputs as illustrated in Fig. 2.5(a). The results are analogous to the previous ones, apart that in this case three clusters emerge. As before during the final consolidation stage each inhibitory neuron ends associated to one of these clusters, connectivity matrix (1"') in Fig. 2.5(b). From the dynamical point of view, the neurons of the different clusters no longer fire in anti-phase, but instead at regular intervals corresponding to one-third of the firing period of the single cluster, see the raster plot (2"') in Fig. 2.5(b).

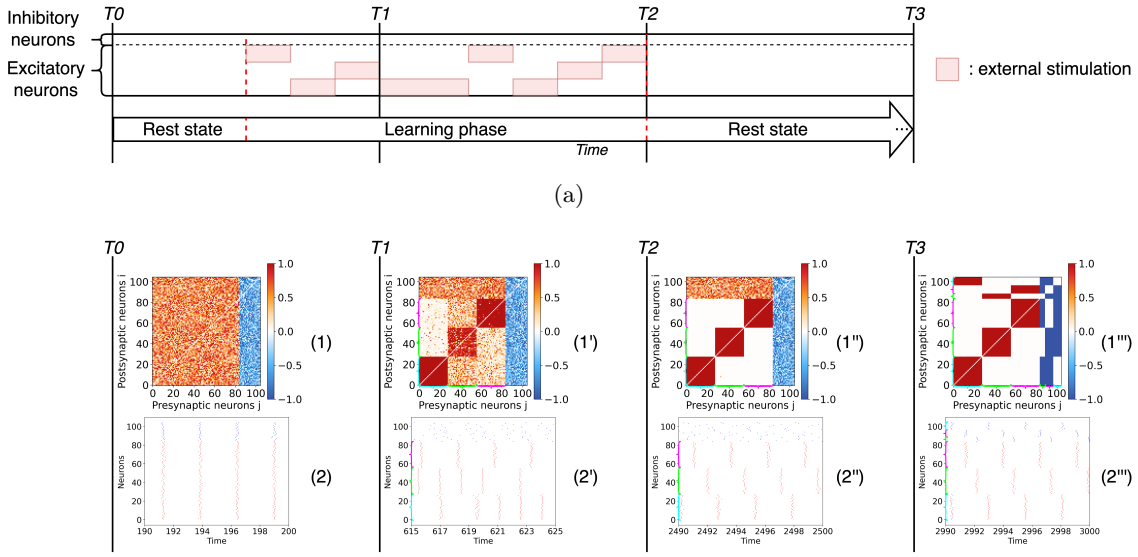


Figure 2.5: Entrainment of a networks of excitatory and inhibitory θ-neurons with three stimuli. (a) Stimulation protocol of the excitatory neurons by three different and non-overlapping stimuli. (b) Entrainment results at different instances of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. Inhibitory neurons are sorted by phases at time T3 for visualization purposes. The cyan, green and magenta brackets represent clusters 1, 2 and 3 respectively when they are visible in weight matrices and raster plots.

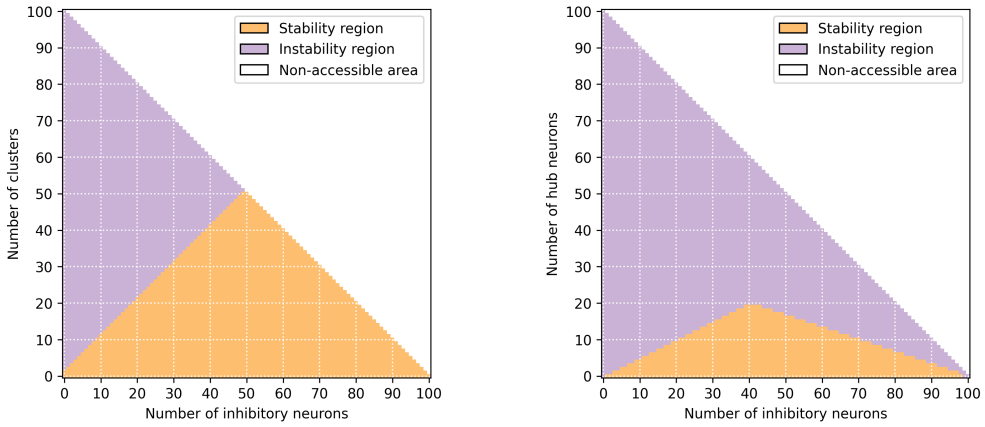
These results generalize by increasing the number of independent inputs. However, when there are too many inputs, we observe that it is more difficult for the network to sustain separate dynamics for each cluster and therefore the long-term consolidation is compromised. The reason is that, as the number of inputs increases, the clusters are made of fewer neurons and it becomes less likely that a sufficient number of inhibitory neurons will associate with each cluster allowing for its consolidation. This opens up the question about how many inhibitory neurons are necessary in the network in order to maintain clusters that fire at distinct times.

To answer this question, we perform the following experiment: we artificially prepare weight matrices made of M modules of excitatory neurons similar to those found at the post-learning time T2. Then, we associate each available inhibitory neuron to a single structural module such that it receives excitatory inputs from this cluster and inhibits the excitatory neurons of all the other clusters. Thus generalizing the architecture reported in Fig. 2.4(a) to many clusters, each containing a single inhibitory neuron. Starting from these weight matrices, we let the network spontaneously evolve for a long period, similarly to what was previously done during the post-training phase, and at the end of the simulation we count the number of structural modules that survive. We consider $N = 100$ neurons in the network and repeat the experiment for an increasing number M of modules (namely,

from $M = 2$ to $M = 100$) controlled by varying the number of inhibitory neurons.

The results are shown in Fig. 2.6(a). We find that in order to maintain M independent modules the network needs to contain at least $M - 1$ inhibitory neurons. An example of the effect of this limitation is shown in Fig. A.15 of Appendix A.3, where the experiment of Fig. 2.5 is repeated for a network containing a single inhibitory neuron. In this case even if the original weight matrices contained three structural clusters only two independent clusters could be maintained in the end.

In the same way as in Fig. 2.4(b), these experiments have been reproduced several times with randomised initial conditions and stimulation sequences. Again thanks to the Kuramoto-Daido order parameters, we were able to quantify over different realizations the number of clusters in the network during each phase of the protocol according to the number of inputs learned and inhibitory neurons. As before, these analyses gave negligible variations confirming the robustness of the results.



(a) (b)

Figure 2.6: (a) Regions of stability (orange) and instability (purple) of the structural modules versus the number of clusters initially present in the weighted connectivity matrix and the number of inhibitory neurons in the network. The line separating the two regions represents the upper limit to observe M independent clusters versus the number of inhibitory neurons. This limit corresponds to $M - 1$ inhibitory neurons. (b) Regions of stability (orange) and instability (purple) of the two clusters versus the number of excitatory neurons encoding for the two stimuli (hub neurons) and the number of inhibitory neurons in the network. The lines separating the two regions represent the minimal number of inhibitory neurons N_I required to observe two clusters with N_H hubs. For $N_I \leq \frac{2}{5}N$ the line is given by $N_I = 2N_H + 1$. For larger number of inhibitory neurons the limit is given by $N_I = N - 3N_H$. Note that both graphs have been realized for a network made of $N = N_E + N_I = 100$ neurons in total, of which N_E are excitatory and N_I inhibitory, this constraint is at the origin of the non-accessible areas (white regions).

In practice, when the networks are randomly initialized, it is rather difficult to reach

the optimal configuration representing the upper limit in which each inhibitory neuron controls for one of the M clusters. However, we show that by preparing the weight matrices as explained before and by considering a network with 50% of excitatory and inhibitory neurons it is possible even to obtain a "splay state" [427] (similar to that found in Fig. 1.3), characterized by $N/2$ clusters, each formed by a pair of excitatory and inhibitory neurons, spiking one after the other (see Fig. A.12 of Appendix A.3).

Another important aspect to consider is that the time interval (the phase shift) between the population bursts associated to each single cluster reduces with the number of clusters present in the network. Therefore, since the time (phase) potentiation window of the plasticity function $\Lambda(\Delta\theta)$, shown in Fig. 2.1, has a finite duration (width) this can induce correlations among clusters characterized by time (phase) shifts smaller than the duration (width) of such a window. These correlations then lead to a merging of initially independent clusters into a single one on the long run. This phenomenon is clearly visible in Fig. A.12, where the "splay state" merges into few clusters over the long term. Therefore, the time (phase) potentiation window should be carefully selected to be sufficiently narrow to reinforce neurons in phase, but also sufficiently broad to allow for some tolerance to the exact phase matching due to noise and variability. To summarize, we can safely affirm that there is a time difference (phase-shift) limit below which nearby clusters cannot remain dynamically independent. At the same time, we should keep in mind that the tolerance of the adaptation process is also a parameter affecting the stabilization of the clusters in the long term.

2.3.3 Neurons encoding for multiple stimuli

So far we have only considered cases in which external inputs were spatially segregated, meaning that each input targets a separate group of neurons. We now consider the possibility that a sub-group of neurons can encode two stimuli at the same time, thus exhibiting a simple form of "mixed selectivity" often observed in neurons in the prefrontal cortex of primates during the performance of memory related tasks [317, 296]. In particular, we study the case in which a sub-group of neurons responds to two distinct stimuli. For illustration we start by considering two stimuli presenting an overlap over eight neurons, see protocol in Fig. 2.7(a). To facilitate the formation of the connections we now strictly alternate the areas stimulated and keep a short resting period between each stimulation. This experiment is similar to the one carried out in the previous chapter in Fig. 1.10.

The results depicted in Fig. 2.7(b) are similar to the observations reported before with the formation of two clusters due to the adaptation to the stimuli (see matrix (1")). The only difference is that in this case a few neurons – those who receive overlapping inputs – become structural hubs. During the consolidation stage, these putative hub neurons associate with one of the two clusters decoupling from the other one but tend to remain connected with the other hubs as shown in matrix (1"'). Regarding the spiking dynamics, at the end of the training period, the hubs behave as a third independent cluster, by firing at an intermediate time between the two clusters, see raster plot (2"). Although at the end of the consolidation stage they finally become synchronized with one of the two clusters,

see raster plot (2"'). Therefore, this may question on the possibility of exhibiting persistent mixed selectivity with such a model even if the hubs still present structural connections among them. However when increasing the size of the overlap, the two clusters are more likely to become synchronized with each other. This consequently raises the question of what is the largest size of the overlap that allows the two clusters to remain distinct at the end.

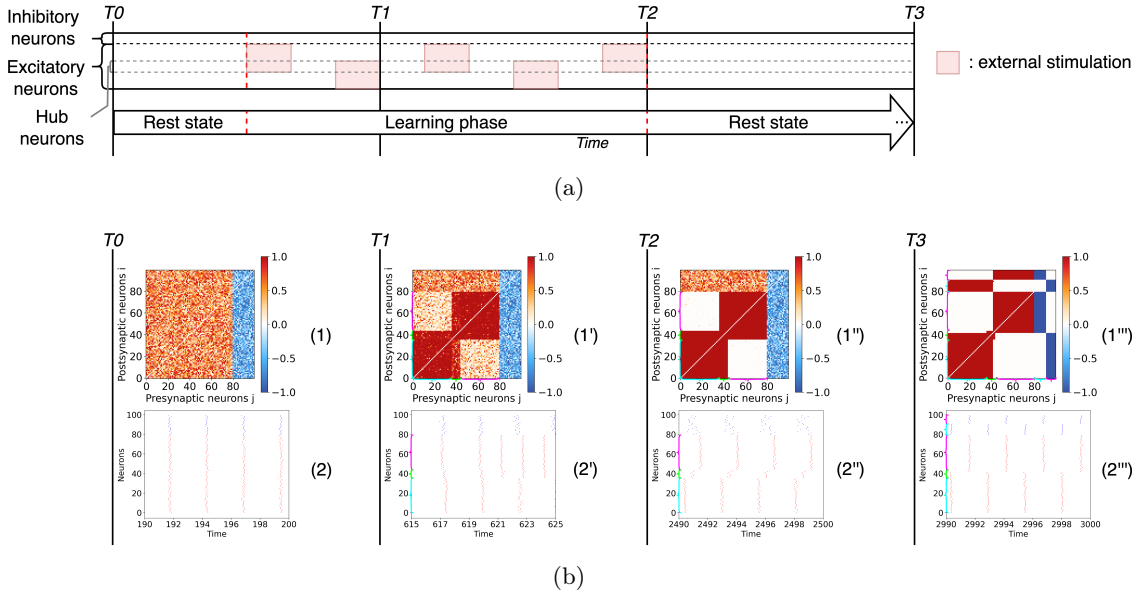


Figure 2.7: Entrainment of a networks of excitatory and inhibitory θ-neurons with two overlapping stimuli. (a) Schema of the experiment protocol showing that the two presented stimuli involve 8 shared neurons. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. Note that the inhibitory neurons and the putative hub neurons are sorted by phases at time T3 for visualization purposes. The cyan, magenta and green brackets represent clusters 1, 2 and the hubs respectively when they are visible in weight matrices and raster plots.

To address this question, we generalize the protocol by increasing the number of hubs N_H (neurons that receive both inputs) and the number of inhibitory neurons N_I in the network. For each combination of overlap and number of inhibitory neurons, we observe whether the final network still exhibits two spatially segregated clusters, or if they merge into an all-to-all connected network. The results of this analysis are reported in Fig. 2.6(b): the orange (purple) region refers to the stability (instability) region of the two clusters. As a first constraint observed, the number of excitatory neurons in each cluster cannot be smaller than the number of hubs, i.e. of the number of excitatory neurons shared with the other cluster. Otherwise the hubs would become the dominant cluster leading to synchrony between all excitatory neurons. Furthermore, the number of inhibitory neurons required to maintain the two clusters segregated should be at least twice the number of hubs ($N_I \geq$

$2\,N_H$). The reason is that to avoid global synchrony, each module needs to compensate for the excitatory links shared with the other cluster via a similar amount of inhibitory links. Therefore the minimal number of inhibitory neurons required to stabilize the two clusters with N_H overlapping neurons is $N_I = 2N_H + 1$, where an extra inhibitory neuron is needed to maintain the repulsion between the two excitatory modules as previously observed in Fig. 2.6(a). This argument can be generalized to M clusters, giving rise to the following rule $N_I = (N_H + 1) * M - 1$, where the N_H hubs encode for all the M stimuli. However, for sufficiently large N_I, namely $N_I > \frac{2}{5}N$, due to the first constraint discussed above, the minimal number of inhibitory neurons needed to observe the two clusters becomes $N_I = N - 3N_H$. For an arbitrary number of clusters M, this rule generalizes as $N_I = N - (M + 1) * N_H$. This limit is further confirmed by the analysis reported in Fig. A.16 of Appendix A.3, where we repeat the experiment shown Fig. 2.7 by reducing the number of inhibitory neurons from $N_I = 20$ to $N_I = 16$. In this case, since the minimal number of required inhibitory neurons should be 17, to maintain 2 clusters with 8 hubs, the two clusters indeed merge into an unique one.

Similarly to Fig. 2.4(b), these experiments have been reproduced several times with randomised initial conditions. Here the Kuramoto-Daido order parameters allow to quantify over different realizations if the two clusters remain stable in the network during and after the learning depending on the number of hubs and inhibitory neurons. Again, these analyses gave negligible variations confirming the robustness of the results.

2.3.4 Memory storage and recall are controlled by the inhibitory neurons

In this last experiment, we start from previous results showing that the connections between excitatory neurons are quite volatile in the absence of learning and that inhibitory connections control for the long-term storage of memories in the cortex [273]. We consider the scenario in which the excitatory clusters formed in the first experiment of Fig. 2.3 have disappeared due to the volatility of the excitatory connections (e.g. due to new stimulation patterns), while the connections involving inhibitory neurons are maintained since they are more stable. Specifically, we select random connections between the excitatory neurons, while the connectivity involving the inhibitory neurons conserves the values induced by the original learning process. Then we examine the response of the network with this altered connectivity matrix to a brief recall of the stored memory patterns. The schema summarizing the stimulation protocol and the results of this experiment are shown in Figs. 2.8(a) and 2.8(b).

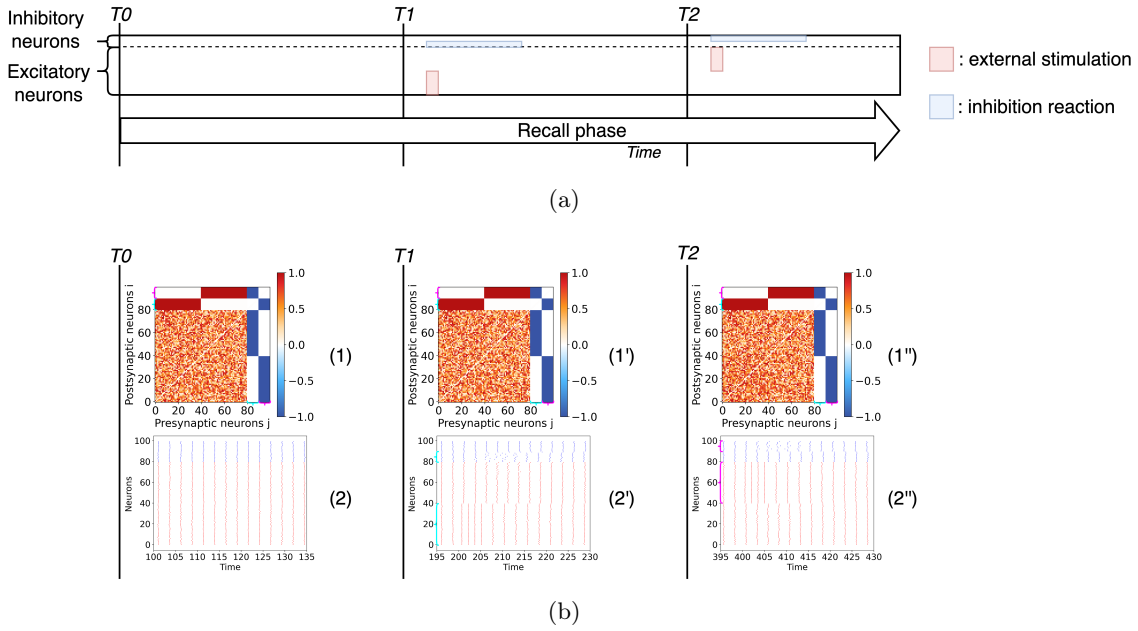


Figure 2.8: Memory recall experiment for a network of excitatory and inhibitory θ-neurons. (a) Schema of the pattern recall experiment. (b) The results are given at different instants of the simulation. The time T0 corresponds to the beginning of the simulation before the stimulations, after a transient period $t_transient = 100$ has been discarded. Time T1 corresponds to the stimulation of the first pattern (from time $t = 200$ to time $t = 205$) followed by the recall of this one. Time T2 corresponds to the stimulation of the second pattern (from time $t = 400$ to time $t = 405$) followed by the recall of this one. The graphs have the same significance as in Fig. 2.3. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

First of all we observe that when no stimulation is applied to the system, the network is in a relatively synchronized state as shown by the raster plot (2) at time T0. Therefore, the information stored in the network, corresponding to two segregated memory patterns, is not evident from the temporal evolution of the neurons. However, when one of two patterns is presented at times T1 and T2 this induces an initial increase in the spiking rate of the stimulated neurons and shortly after, the desynchronization of the inhibitory neurons associated to the stimulus from the rest of the population for a brief time interval (see raster plots (2') and (2")). This leads to a transient re-emergence of the stored memory pattern, reflecting an activity that is almost in anti-phase with the activity of the excitatory neurons associated to other patterns. These results show that for a specific pattern to be recognized it is sufficient that memories are stored in the connections associated with the inhibitory neurons.

2.4 Summary and conclusions

The neural networks of animals, from the simple nervous system of the worm *Caenorhabditis elegans* to the mammalian brain, display modular architectures at different scales of organization [403, 179, 248, 278, 266, 396, 422, 217, 42, 44]. In this chapter we have investigated the formation and the consolidation of neural assemblies as driven by the entrainment to different inputs in networks of oscillatory θ-neurons. Previous analyses, in particular in Chapter 1, have shown that learning and adaptation could generate modular networks [350, 100, 389, 39], but as it is common practice in artificial neural networks, those studies overlooked some fundamental biological constraints. Here, we have shown that satisfying Dale's principle – i.e. the distinction between excitatory and inhibitory neurons – is crucial for the consolidation of stimulus-driven neural assemblies. Furthermore, at variance with other popular models of learning where the synaptic weights and the neural activity are frozen once the training phase is finalized, we allowed for a spontaneous activity of the adaptive network even after the training was finished, thus mimicking a more biologically realistic scenario. We have found that during this post-training phase the learned memories are consolidated, provided the network is made of both excitatory and inhibitory neurons. Indeed, if the network contains only excitatory neurons, then the learned memories are lost during to the post-learning phase. Moreover, if the network contains excitatory and inhibitory synapses, but Dale's principle is violated, the network ends in a state in which the emerged synaptic connectivity does not correspond to the encoded stimuli.

In conclusion, the formation of modular structures is induced by the spatio-temporal correlations of the stimuli applied to the network and by the resulting adaptation mechanism, thus maintaining a certain synchrony within them. By preserving a specific dynamics to each cluster, **the inhibitory neurons allow to maintain and consolidate the learned structure over the long-term**. Also in Fig. 2.8(b), we have proved that the learned items are stored in the connections involving the inhibitory neurons and so these links are sufficient to recall or recognize information as in AI systems.

Furthermore, we have shown that **the number of inhibitory neurons controls the memory capacity of the network**. Indeed, Fig. 2.6(a) shows that the number of inhibitory neurons is related to the number of different neural assemblies that can keep independent dynamics and, therefore, be stored over the long-term. Similarly, we have shown in Fig. 2.6(b) that the number of inhibitory neurons also controls the maximal number of neurons which can code different items at the same time.

Despite the efforts performed in this chapter to fulfil several biologically plausible constrains, some aspects of the model could be further improved. For example, a more realistic scenario could replace the θ-neuron model here employed by spiking neurons with STDP plasticity. The learning rule considered here corresponds to the basic paradigm of an Hebbian rule implying that "*cells that fire together, wire together*" [174] and depends on the difference of the phases associated to each neuron in a continuous manner. By employing this rule, since synaptic potentiation occurs only when two phases match within a quite narrow window, we have shown that a lack of precision in the measure-

ments of the phase difference can have a real impact. In particular, we have shown how this lack of precision can limit the number of memory items that can be effectively stored in the long term, in Fig. A.12 of Appendix A.3. Another crucial element of the model is the presence of two time scales associated to the evolution of the synaptic weights, a faster adaptation during the learning phase and a slower plasticity evolution during rest. While this choice is justified by several indications that learning and consolidation processes may occur simultaneously or sequentially in the brain characterized by different time-scales [122, 345, 372, 89, 118, 301, 279, 275], it is still unclear how the brain itself solves this problem and avoids the loss of learned memories given the fact that synapses are permanently susceptible to adaptation. Better understanding of the biological mechanisms will allow to define more realistic adaptation rules in the models. The following chapter will attempt to bring further biological constraints to the model while resolving these issues.

Chapter 3

Long-term memory consolidation in spiking neural networks

Contents

The previous chapter emphasized the role of inhibitory neurons to maintain neural assemblies in networks of θ-neurons. In this chapter, we aim to reproduce these findings with spiking neural networks admitting more realistic dynamics and closer to biological networks. We will notably show how long-term memory consolidates thanks to spontaneous recall provided by dual STDP inhibition, in other words two populations of inhibitors subject to different plasticity functions.

3.1 Introduction

Long-term memory (LTM) refers to the ability of the brain to store and retrieve information over an extended period of time, ranging from a few minutes to several decades [262, 357]. These memories are stored in various regions of the brain, including the hippocampus which may be responsible for consolidating new memories [145, 222], and the cortex which stores long-term memories for a lifetime [51]. The process of encoding LTM is complex and involves various neural mechanisms such as synaptic plasticity and neurochemical

changes [49, 254]. In particular, as demonstrated in previous chapters, inhibitory STDP plays a crucial role in adaptation processes of the brain in order to shape and preserve the structure of the network according to the correlations between inputs [7, 213, 215].

Specifically, the process of maintenance and consolidation of LTM involves a complex set of neural and molecular mechanisms that allow new information to be retained over extended periods of time. First of all, the new encoded information, initially fragile and subject to interference, must be stabilized to be maintained over time [262]. To do this, a reorganization of neuronal connections can occur, characterized by the synthesis of new proteins and the modification of gene expression [207]. Thus, different stages of consolidation occurring over different time frames (i.e. minutes to hours, days to weeks, and even months to years) are carried out in different brain regions, notably in the hippocampus, which plays a particularly important role in the early stages of consolidation [356, 124]. Finally, other factors influence memory consolidation, such as sleep, stress, emotions and attention, by modulating the activity of neurotransmitters, hormones and other signaling molecules [263, 365, 201].

In particular during sleep and rest (i.e., no task activity), short phases of partially synchronous activation or series of sequence activation, which can be likened to spontaneous recall or replay of learned information, occur promoting memory consolidation [166, 365, 373]. Spontaneous or internal recall refers to the retrieval of information without intentional effort to remember or the action of external cues. It can occur in response to internal triggers such as thoughts or emotions. These triggers can activate memory associations that lead to the retrieval of relevant information [330, 28]. These internal recalls of long-term memory can also be influenced by factors such as mood and attention [55, 91]. In contrast, we define the conscious or external recall, which is generally associated with the intentional retrieval of information from long-term memory, often facilitated by external retrieval cues (e.g. sensory inputs) [384, 203]. This process may refer to explicit memory, also known as declarative memory, involving the conscious recall of past events and facts [357]. In particular, the capacity to express accurate external recalls is influenced by the specificity of the retrieval cues, the level of prior knowledge and the strength of the memory trace [319, 261].

Although these short epochs of partially synchronous activity associated with memory recall may occur, in contrast to the partially synchronous and regular dynamics observed with the phase models in the previous chapters, neurons of the cortex generally admit asynchronous irregular (A-I) dynamics during the resting-state of the brain [363]. This state, often observed in the neocortex, is thought to facilitate flexible information processing by allowing the simultaneous integration of multiple sensory streams and the formation of new associations between pieces of information [2, 68, 146].

However, precisely how inhibition enables the creation, maintenance and consolidation of LTM as well as how this relates to spontaneous recall, is still unclear. Although such recall has been observed in several studies using spiking neural models [9, 62, 272], it is generally associated with working memory and does not necessarily occur in a spontaneous way. Moreover, it remains questionable how the A-I state, although also observed independently in several computational studies [119, 60, 313, 395, 400], can coexist with these

spontaneous recalls. The aim of this chapter is first to investigate **how inputs can be learned to shape modular structures through synaptic plasticity and inhibition** in a similar way to the previous chapters, but now **with spiking neurons**. Then, the objective is to understand **how to maintain these LTM in an A-I state dynamics** by adding a few biological constraints. In particular, we focus on **the role of inhibition on the emergence of spontaneous internal recall**, promoting the consolidation of the structures.

The chapter is organized as follows. After introducing the subject, a description of the methods employed for the modelling of the neural network and its quantification is given in Sec. 3.2. Then, we investigate the conditions necessary for learning modular structures and admitting spontaneous recalls of them during resting-state, emphasizing the role of inhibition in this process in Sec. 3.3.1.1 and 3.3.1.2. We also generalize these results to different number of clusters. In a second part, we focus on the effective role of these spontaneous recalls, especially for the maintenance and consolidation of the learned structure in Sec. 3.3.2. It appears again that inhibitory links are decisive for the emergence and re-emergence the memory items. Thereafter, we also study formation and consolidation of overlapping memories in Sec. 3.3.3. The last results are devoted to the provocation of induced external recalls in particular to better understand how recall appears in Sec. 3.3.4.

3.2 Methods

This section describes the spiking neuronal network model governing the dynamics of the neurons, as well as the learning rules used for the adaptation of synaptic weights. Then, a number of measurements are given to quantify the states of the network throughout the different simulations.

3.2.1 Spiking neuronal network model

The action potentials of biological neurons, accurately described by the Hodgkin-Huxley model [181], can also be represented by a wide range of Spiking neural network (SNN) models. These models are described as the third generation of neural networks [249, 425], due to their increased efficiency in computational processing of spatial and temporal information compared to the first two generations (i.e., the perceptron [323] and non-linear neural networks [325]) commonly used in machine learning. Moreover, a major difference in this type of network is that each neuron admits a potential which evolves as a function of the sum of the spikes coming from the pre-synaptic neurons, and when it reaches a threshold, the neuron fires [249]. In the previous chapters, we modelled this potential as the phase of the neurons/oscillators. Finally, it should be noted that SNN with both excitatory and inhibitory populations are more biologically realistic and can exhibit more complex dynamics [249, 361].

Consequently, throughout this chapter, we consider an excitatory-inhibitory neural network made of heterogeneous QIF neurons [61, 170, 236]. This model is a type I neuron, characterized by a saddle-node bifurcation on a invariant circle (unlike the Hodgkin-Huxley

model which is a type II neuron with a sub-critical Hopf bifurcation) and which only produces action potential patterns and omits certain aspects such as gating variables [144]. Although it is less physiologically accurate, it is significantly less computationally expensive and easier to compute and study. The θ-neuron model from Chapter 2 can be mapped to each other, which facilitates comparisons.

Unless specified, the network is composed of 80% of excitatory neurons and 20% of inhibitory neurons as is commonly accepted in the human cortex [215, 243, 273]. This network is composed of three different populations of neurons:

- A population of excitatory neurons following a Hebbian learning as in Chapter 2.
- A population of inhibitory neurons following a Hebbian learning as in Chapter 2.
- A population of inhibitory neurons following an anti-Hebbian learning.

The fact of having two populations of inhibitory neurons can be supported first by the large variety of subclasses of GABAergic cells present in the brain [1] and secondly by the studies highlighting the predominant role of both feedback [232] and feed-forward inhibitions [231] in learning. Therefore, the evolution of the membrane potential V_i of each neuron ($i = 1, ..., N$) is described by the following equation:

$$\tau_m \dot{V_i} = V_i^2(t) + \eta_i + g_e S_i^e(t) + g_{hi} S_i^{hi}(t) + g_{ai} S_i^{ai}(t) + I_i(t) + \xi_i(t), \tag{3.1}$$

where the synaptic inputs $S_i^e(t)$, $S_i^{hi}(t)$ and $S_i^{ai}(t)$ (for excitatory, Hebbian inhibitory and anti-Hebbian inhibitory inputs respectively) of the neuron i are given by:

$$\tau_d^e \dot{S_i^e} = -S_i^e + \frac{\tau_d^e}{N_e} \sum_j^{N_e} w_{ij} \delta(t - t_j), \tag{3.2}$$

$$\tau_d^i \dot{S_i^{hi}} = -S_i^{hi} + \frac{\tau_d^i}{N_{hi}} \sum_j^{N_{hi}} w_{ij} \delta(t - t_j), \tag{3.3}$$

$$\tau_d^i \dot{S_i^{ai}} = -S_i^{ai} + \frac{\tau_d^i}{N_{ai}} \sum_j^{N_{ai}} w_{ij} \delta(t - t_j), \tag{3.4}$$

where $\tau_m = 0.02$s is the membrane time constant, $\tau_d^e = 0.002$s and $\tau_d^i = 0.005$s the time decay for the excitatory and inhibitory neurons, $\eta_i \sim \mathcal{N}(0.0, (\pi\tau_m)^2)$ (to maintain an average neuron frequency at rest of around 1Hz) the excitability parameter, $N = N_e + N_{hi} + N_{ai} = 100$, $N_e = 80$, $N_{hi} = 10$ and $N_{ai} = 10$ respectively the number of excitatory and Hebbian and anti-Hebbian inhibitory neurons, $g_e = 100$, $g_{hi} = 200$ and $g_{ai} = 400$ the global coupling strength for the excitatory neurons and Hebbian and anti-Hebbian inhibitory neurons. The coupling weights from neuron j to i is depicted by w_{ij}, t_j is the time of spike of the j-th neuron, and $\delta(t)$ is the Dirac delta function (corresponding to δ-spikes). Finally, $I_i(t) = \{0, (50\pi\tau_m)^2\}$ (to obtain neurons with a frequency of 50Hz when stimulated) is an external input current and $\xi_i(t) \sim \mathcal{N}(0.0, (4\pi\tau_m)^2)$ (to have a neuron

frequency variability around 4Hz) is a Gaussian noise. We note that these noise and excitability values allow for relatively heterogeneous neurons admitting δ waves rhythms (0.5–4 Hz) to θ waves rhythms (4–8 Hz) typically characteristic of sleep states [3, 104, 140]. On the contrary, when an external input current is applied to the neurons, they admit high frequency oscillations of γ rhythm characterizing a high-order perceptual activity [177, 371].

As before, we consider a fully connected network without self-connections. We integrate the Eqs. (3.1), (3.2), (3.3) and (3.4) using the Euler method. Whenever $V_i(t)$ reaches the peak value $V_p = 10$, the neuron i emits a spike (after a time of $\frac{1}{V_i}\tau_m$ seconds) and its membrane voltage is reset to $V_r = -10$ after a certain refractory period $\frac{2}{V_i}\tau_m$ seconds to compensate with the approximation $V_p = -V_r = 10$. These periods are equivalent to the theoretical time intervals to reach $V_i(t) = \infty$ from $V_i(t) = 10$ and to pass from $V_i(t) = -\infty$ to $V_i(t) = -10$ [370].

3.2.2 Plasticity functions

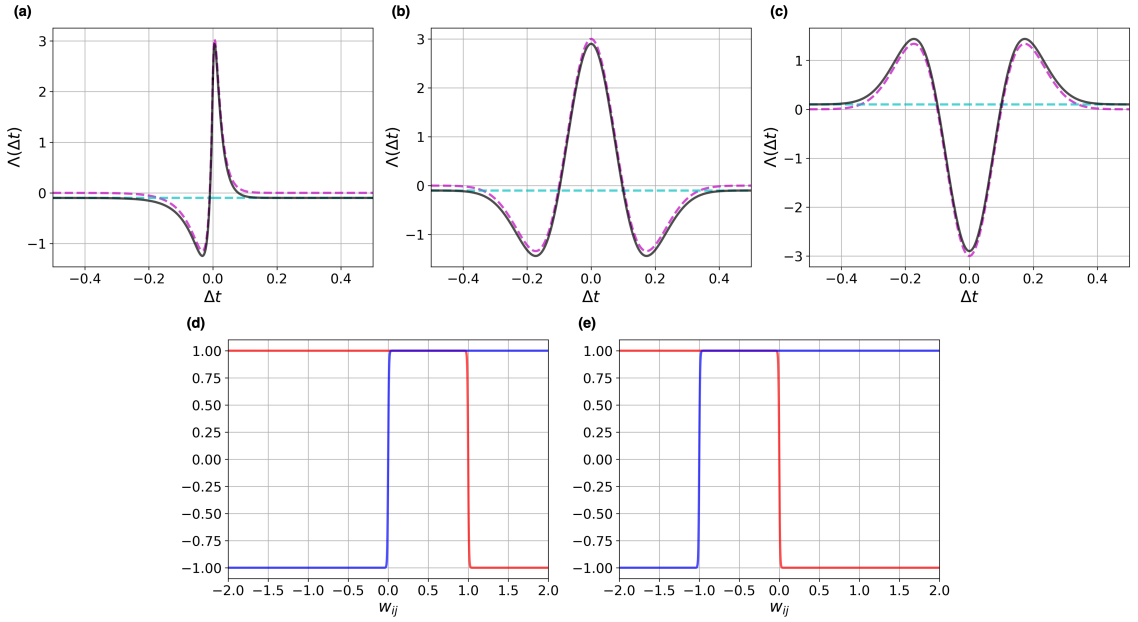


Figure 3.1: Plasticity and soft bound functions. (a) Plasticity function for Hebbian STDP asymmetric (in dashed magenta), the forgetting function (in dashed cyan) and the combination of the two functions defined in Eq. (3.7) (in black). (b) Plasticity function for Hebbian STDP symmetric (in dashed magenta), the forgetting function (in dashed cyan) and the combination of the two functions defined in Eq. (3.8) (in black). (c) Plasticity function for anti-Hebbian STDP symmetric (in dashed magenta), the forgetting function (in dashed cyan) and the combination of the two functions defined in Eq. (3.9) (in black). (d) The soft bound functions for excitatory neurons with potentiation function in red ($\tanh(\lambda(1-w_{ij}))$) and depression function in blue ($\tanh(\lambda w_{ij})$) with $\lambda = 100$. (e) The soft bound functions for inhibitory neurons with potentiation function in red ($\tanh(\lambda(-w_{ij}))$) and depression function in blue ($\tanh(\lambda(w_{ij}+1))$ with $\lambda = 100$.

Regarding the adaptation of the weights w_{ij}, we use STDP rules depending on the time difference $\Delta t = t_i - t_j$ between the last spikes of the post-synaptic neuron i and pre-synaptic neuron j neurons. The plasticity functions $\Lambda^+(\Delta t)$ and $\Lambda^-(\Delta t)$ from Eqs. (3.5) and (3.6) for potentiation and depression respectively, depend on the nature of the pre-synaptic neuron.

$$\Lambda^+(\Delta t) = \begin{cases} \Lambda(\Delta t), & \text{if } \Lambda(\Delta t) \geq 0, \\ 0, & \text{if } \Lambda(\Delta t) < 0, \end{cases} \tag{3.5}$$

$$\Lambda^-(\Delta t) = \begin{cases} 0, & \text{if } \Lambda(\Delta t) \geq 0, \\ \Lambda(\Delta t), & \text{if } \Lambda(\Delta t) < 0. \end{cases} \tag{3.6}$$

For excitatory neurons we use a Hebbian STDP asymmetric function commonly used in the literature [46, 160, 212, 215, 247, 329] described by Eq. (3.7) and by Fig. 3.1 (a). This function potentiates the weights of the neurons spiking in a causal way. In other words, when the pre-synaptic neuron emits a spike before (after) the post-synaptic neuron, the synaptic weight increases (decreases) [76].

$$\Lambda(\Delta t) = \begin{cases} A_+ e^{-\frac{\Delta t}{\tau_+}} - A_- e^{-\frac{4\Delta t}{\tau_+}} - f, & \text{for } \Delta t \geq 0, \\ A_+ e^{\frac{4\Delta t}{\tau_-}} - A_- e^{\frac{\Delta t}{\tau_-}} - f, & \text{for } \Delta t < 0, \end{cases} \tag{3.7}$$

with the time constants $\tau_+ = 0.02$s and $\tau_- = 0.05$s, the amplitudes $A_+ = 5.296$ and $A_- = 2.949$. The forgetting term $f = 0.1$ allows to have a constant small depression of the weights whatever the spike timing difference. It models the natural, constant and slow forgetting of memories [172, 413].

For Hebbian inhibitory neurons we use a Hebbian STDP symmetric function [215, 231, 247, 298, 401] described by Eq. (3.8) and by Fig. 3.1 (b). This function takes the form of a Ricker wavelet (or Mexican hat) function, potentiating (depressing) weights of neurons spiking in a correlated (uncorrelated) way.

$$\Lambda(\Delta t) = A(1 - (\frac{\Delta t}{\tau})^2) e^{\frac{-\Delta t^2}{2\tau^2}} - f, \tag{3.8}$$

with the time constant $\tau = 0.1$s, the amplitude $A = 3$ and the forgetting term $f = 0.1$.

For anti-Hebbian inhibitory neurons we use an anti-Hebbian STDP symmetric function [143, 220, 215, 232, 304] described by Eq. (3.9) and by Fig. 3.1 (c). This function takes the form of a reverse Ricker wavelet (or Mexican hat) function, potentiating (depressing) weights of neurons spiking in a uncorrelated (correlated) way.

$$\Lambda(\Delta t) = -A(1 - (\frac{\Delta t}{\tau})^2) e^{\frac{-\Delta t^2}{2\tau^2}} + f, \tag{3.9}$$

with the time constant $\tau = 0.1$s and the amplitude $A = 3$. In this case, the forgetting term $f = 0.1$ allows to have a constant small potentiation (the rule being anti-Hebbian) of the weights whatever the spike timing difference.

3.2.3 Adaptation of synaptic weights

Now that we have defined the plasticity functions according to the nature of the neurons, we describe the equations governing the evolution of synaptic weights. We update the weights accordingly when one of the two neurons spike.
If the pre-synaptic neuron j is excitatory, the weights evolve according to:

$$\tau_l \dot{w_{ij}} = [\tanh(\lambda(1 - w_{ij})) * \Lambda_e^+(\Delta t) + \tanh(\lambda w_{ij}) * \Lambda_e^-(\Delta t)]. \tag{3.10}$$

If the pre-synaptic neuron j is Hebbian inhibitory, the weights evolve according to:

$$\tau_l \dot{w_{ij}} = [\tanh(\lambda(-w_{ij})) * \Lambda_{hi}^+(\Delta t) + \tanh(\lambda(w_{ij} + 1)) * \Lambda_{hi}^-(\Delta t)]. \tag{3.11}$$

If the pre-synaptic neuron j is anti-Hebbian inhibitory, the weights evolve according to:

$$\tau_l \dot{w_{ij}} = [\tanh(\lambda(-w_{ij})) * \Lambda_{ai}^{+}(\Delta t) + \tanh(\lambda(w_{ij}+1)) * \Lambda_{ai}^{-}(\Delta t)], \tag{3.12}$$

with $\tau_l = 0.2$s the learning time scale for the adaptation, Λ_e^+ and Λ_e^- the plasticity functions associated with the excitatory neurons for potentiation and depression respectively, Λ_{hi}^+ and Λ_{hi}^- the plasticity functions associated with the Hebbian inhibitory neurons for potentiation and depression respectively and Λ_{ai}^+ and Λ_{ai}^- the plasticity functions associated with the Hebbian inhibitory neurons for potentiation and depression respectively.

The parameter $\lambda = 100$ controls the slope of the soft bound functions tanh. These two-sided non-linear bounds represented in Fig. 3.1 (d) and (e) allow in a first place to maintain weights $|w_{ij}|$ between zero and one by having a synaptic depression dominating the potentiation for large weights and reciprocally [346, 394]. In other words, these functions work as a saturation of the weights depending on the current state of the weights [200]. In a second place, the steep slope of the function (controlled by λ) guarantee a sufficiently important dynamics even for large weights [158]. In this sense, the functions used are at the limit between a soft and a hard bound.

We note that in this case, contrary to the models in the previous chapters, **the weight adaption does not admit two learning timescales**. Here, the firing rate of the neurons indirectly impacts the adaption rate. Nevertheless, synaptic weights remain continuously subject to adaptation, unlike in conventional AI systems.

We integrate the equations Eqs. (3.10), (3.11) and (3.12) for the weights with the Runge-Kutta fourth order method. All network parameters are summarized in Table 3.1.

3.2.4 Quantification measurements

To quantify the degree of synchronisation of the network or parts of it, we use the Kuramoto-Daido order parameters defined in Eq. (1.3). To calculate the phase θ_j of neuron j at time t, we proceed in the following way:

$$\theta_j(t) = \frac{2\pi(t - t_j(n))}{(t_j(n+1) - t_j(n))}, \tag{3.13}$$

with $t_j(n)$ the last firing time of neuron j at time t and $t_j(n+1)$ the next firing time of neuron j at time t (i.e. $t_j(n) \leq t \leq t_j(n+1)$).

To study the metastability of the network we compute the free energy landscape as follow:

$$F(R) = -\log P(R), \tag{3.14}$$

where $P(R)$ is the probability density function (PDF) of the Kuramoto order parameter R.

The frequency of a neuron j is calculated as its mean firing rate in the following manner:

$$\nu_j = \frac{n_j^{sp}}{T}, \tag{3.15}$$

Table 3.1: Parameters for the network of QIF neurons.

Parameters	Values
N	100
N_e	80
N_{hi}	10
N_{ai}	10
g_e	100
g_{hi}	200
g_{ai}	400
η	$\mathcal{N}(0.0,\, (\pi\tau_m)^2)$
$I(t)$	$\{0, (50\pi\tau_m)^2\}$
$\xi(t)$	$\mathcal{N}(0.0,\, (4\pi\tau_m)^2)$
V_p	10
V_r	-10
τ_m	0.02s
τ_d^e	0.002s
τ_d^i	0.005s
τ_l	0.2s
τ_+	0.02s
τ_-	0.05s
τ	0.1s
dt	0.001s
A_+	5.296
A_-	2.949
A	3
f	0.1
λ	100

with n_j^{sp} the spike count in an time interval of duration T. For a population of N_p neurons, we can compute its average activity during the time interval T:

$$A = \frac{1}{N_p} \sum_{j=1}^{N_p} \nu_j. \tag{3.16}$$

Finally, to quantify the degree of variation of a neuron j, we calculate its coefficient of variation in the following manner:

$$CV_j = \sigma_j / \mu_j, \tag{3.17}$$

with the mean (μ_j) and standard deviation (σ_j) of the interspike intervals (ISIs) of neuron j. ISIs are the time intervals between successive spikes of this neuron.

3.3 Results

The aim of this chapter was to investigate the formation and long-term consolidation of modular architectures associated with memory items in spiking neural networks induced by learning to selective external stimuli. We first study the conditions necessary for the emergence and spontaneous recall of neural assemblies driven by the learning to selective stimuli. Therefore, we compare the efficacy of Hebbian and anti-Hebbian learning rules for the inhibitory neurons in an example of two-stimulus. These results are then generalized to an arbitrary number of stimuli. Secondly, we investigate the properties of spontaneous neural activity after the learning phase and the role of spontaneous recall in order to consolidate and maintain the learned memories. These spontaneous recalls are also studied in networks without structural modules and in networks admitting overlapping memories. Lastly, we investigate the reactivation of memory by externally induced short recalls of the learned memories.

3.3.1 How to learn and recall modular structures: the need for dual STDP inhibition

We begin by studying the conditions in terms of the number of inhibitory neurons and their plasticity rule (i.e. Hebbian and anti-Hebbian) to on the one hand, guarantee a correct learning. In other words, the creation of well-formed and clearly decoupled modules. On the other hand, these conditions must ensure the possibility in the resting-state (i.e. spontaneous activity without external stimulation), to observe realistic activity with low firing frequency, asynchronous irregular spikes and with spontaneous recalls of learned patterns characterized by random epochs close to synchrony in clusters.

We consider networks made of 100 heterogeneous neurons with excitabilities η_i randomly distributed according to a normal distribution $\mathcal{N}(0.0, (\pi\tau_m)^2)$. The initial membrane potential of the neurons are randomly selected in the interval $[-10, 10]$. The weights are randomly distributed in the range $w_{ij} \in [0, 1]$ when the neuron j is excitatory and in the range $w_{ij} \in [-1, 0]$ when the neuron j is inhibitory according to a normal distribution $\sim \mathcal{N}(0.0, 0.2)$. At this stage, the experiment consists in learning 2 distinct memory items (modules).

3.3.1.1 Learning protocol

The typical stimulation protocol is illustrated in Fig. 3.2 (a). Initially, a period of spontaneous activity of 5s is considered in order to allow the system to relax to low frequency resting-state. Then, a learning phase follows during which one of the two groups is randomly chosen among 2 populations and stimulated by a constant external positive current (to obtain neurons with a frequency of 50 Hz) during a fixed period of 0.8s while the rest of the network remain unstimulated. The stimulation is stopped during a transient period of 0.2s and the process restart for a duration of 35s. This transient period allows for more qualitative learning by preventing temporal correlations when changing the stimulated population. The first population correspond to the first half of the excitatory neurons (i

= 1, 2, . . . , 40) and the first half of the inhibitory neurons (i = 81, 82, . . . , 90), and the other one to the second half (i = 41, 42, . . . , 80) and (i = 91, 92, . . . , 100). Unless otherwise specified, each population of inhibitory neuron is composed of an half of Hebbian inhibitors and an half of anti-Hebbian inhibitors. The external stimulation of the excitatory neurons can be seen as incoming sensory inputs, while the stimulation of the inhibitory neurons as higher-level signals from the hippocampus promoting consolidation of memories [356, 124].

Finally, after the learning phase the network is left to evolve spontaneously. During this post-learning phase no neurons are stimulated but the adaptation remains active. In addition to the variation of the previously described parameters subject to stocasticity, the main experiment is repeated for six different variations:

- A network of only excitatory neurons
- A network made of excitatory neurons and inhibitory neurons (only Hebbian)
- A network made of excitatory neurons and inhibitory neurons (only anti-Hebbian)
- A network made of excitatory neurons and inhibitory neurons (Hebbian and anti-Hebbian)
- A network made of excitatory neurons and inhibitory neurons (Hebbian and anti-Hebbian) where only excitatory neurons are stimulated
- A network made of excitatory neurons and inhibitory neurons (Hebbian and anti-Hebbian) where only inhibitory neurons are stimulated

During the initial period of spontaneous activity (see time column T0 in Fig. 3.2), we observe asynchronous irregular states with low firing rates in raster plots (2) at time T0 in all cases except in panel (b) where there is no inhibition. Indeed, the random and weak connectivity of matrices (1) at time T0 favours this random dynamics in addition to the noise and heterogeneity present. However, without inhibition, the spikes of the excitatory neurons induce the other neurons to discharge and so on, without the possibility of reducing their action potential using the balancing effect of inhibition.

During the learning phase at time T1, without any inhibition (in Fig. 3.2 (b)), we observe that when one group of excitatory neurons is stimulated, the neurons of the other group continue to spike (see raster plot (2)). In addition, due to the weak synaptic links present between the two populations at the beginning of the simulation, the stimulation of one group causes the firing of the other one. This reinforces the links between these populations leading to a fully coupled matrix without structural modules (see weighted connectivity matrix (1) at times T1 and T2). Therefore the formation of modular structures is impossible under these conditions.

Simulations of Fig. 3.2 (c) appear to show that adding inhibitory neurons subjected to Hebbian plasticity overcome this issue. During the stimulation of the populations, this inhibition allows to shut down the firing of the neurons of the opposite cluster while the

firing rate of the stimulated neurons increase, and so not implying correlations between different items (see raster plot (2)). This leads to the emergence of two modular structures (see weight matrices (1) at times T1 and T2) among the excitatory neurons where each group projects in a feed-forward manner on a group of inhibitory neurons. Thus, the inhibitory neurons split also in two groups and each projects on the excitatory and inhibitory neurons of the opposite cluster. Therefore, these results show that each module need to be associated to at least one inhibitory Hebbian neuron to be correctly learned.

We see during simulations of Fig. 3.2 (d) that with inhibitory anti-Hebbian neurons, we can have somewhat similar results. During stimulation these inhibitors seems to perform the same role by reducing the firing of the not targetted neurons (see raster plot (2)). Nevertheless, this is only due to the weak connections present between the two populations allowing these neurons to inhibit both populations. With the formation of the two modular structures among the excitatory neurons as before (see weight matrices at times T1 and T2), the inhibitory neurons are also divided in two modular structures where each one projects on the excitatory neurons of its group contrary to the previous case. However, there are still small connections between these modules and so this condition does not guarantee complete decoupling between modules.

Using both types of inhibition in Fig. 3.2 (e), we achieve the emergence of two modular structures as in panel (c). The main difference is the presence of two populations of inhibitory neurons associated with each cluster (see weight matrices at times T1 and T2), corresponding to the Hebbian inhibitory neurons as in panel (c) and the anti-Hebbian inhibitory neurons as in panel (d). This confirms that at this stage of the experiment, inhibitory Hebbian neurons seem to be the necessary ingredient to ensure a correct learning.

In order to test the necessity of stimulating both populations of neurons, the inhibitory neurons are not stimulated in Fig. 3.2 (f). We observe similar results to those in Fig. 3.2 (b) when inhibitory neurons are not present. Indeed, although present, very few inhibitory neurons fire and with a low firing rate. This is notably due to the inhibitory to inhibitory (I-I) links that reduce their activity and the fact that the small excitatory to inhibitory (E-I) links originality present rapidly disappear under the effect of plasticity. This lack of inhibition leads to a high firing of the unstimulated groups for the reasons stated above (see raster plot (2)). As a consequence, the network ends up with excitatory neurons all-to-all coupled (see weighted connectivity matrix at times T1 and T2). This supports the choice of applying stimuli on the inhibitors to balance the action of the external current on the exciters and to facilitate the associations between the different populations in each cluster.

In the opposite case, when only inhibitory neurons are stimulated, in Fig. 3.2 (g), we observe somewhat the inverse phenomenon. Indeed, when a group of inhibitory neurons is stimulated during the learning phase, their firing shuts down all the other neurons in the network as depicted in raster plot (2). This allows to create correlations between stimulated inhibitors and thus the formation of structures between them (see weighted connectivity matrix at times T1 and T2). However, the excitatory neurons do not split into modules (i.e., in excitatory to excitatory (E-E) links) and E-I connections are non-existent. Therefore, the learning of modular structures with this model seems impossible without stimulating both types of neuron populations.

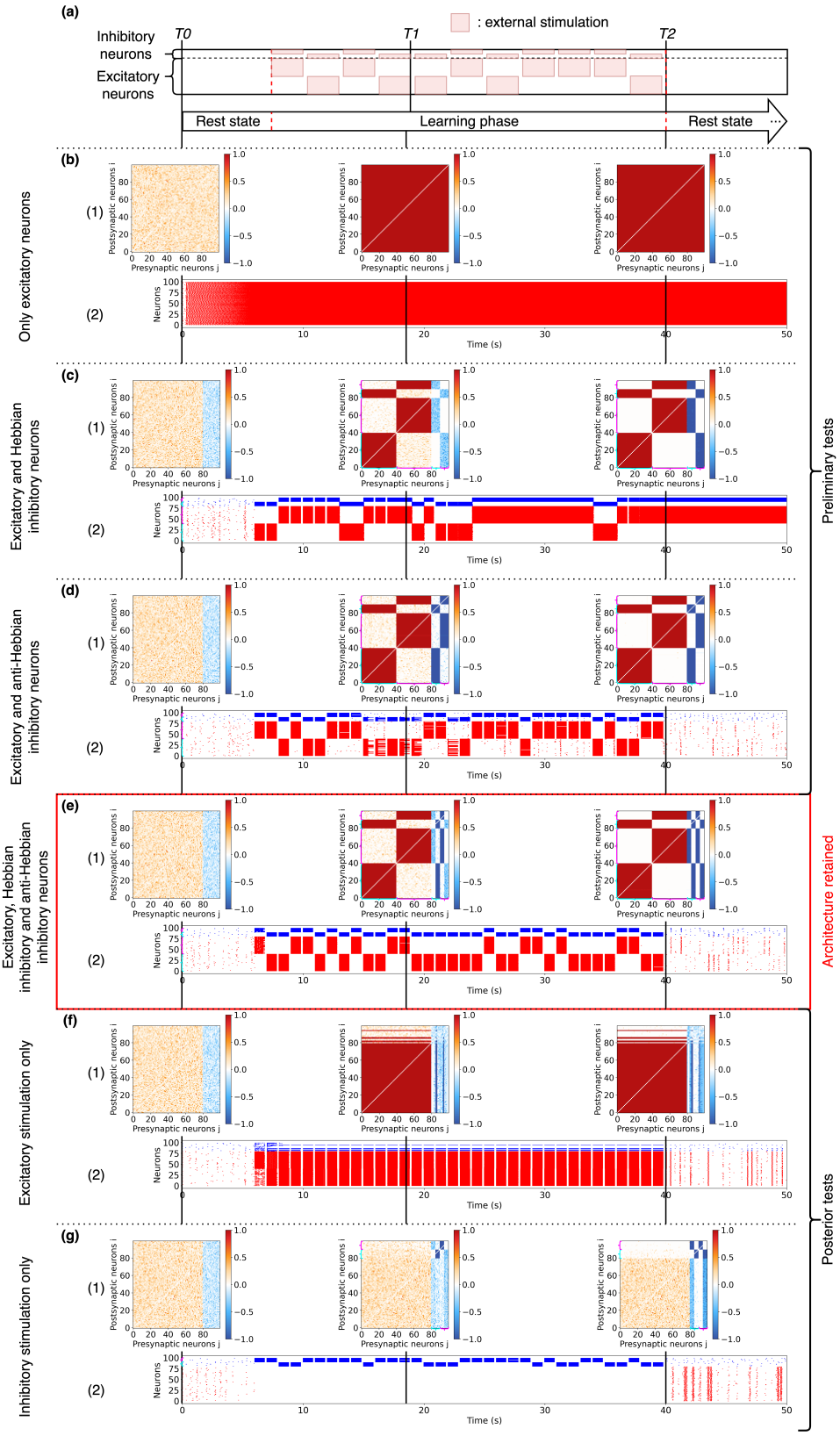

(a)
: external stimulation
Inhibitory neurons
Excitatory neurons
T0
T1
T2
Rest state
Learning phase
Rest state
(b)
Only excitatory neurons
(c)
Excitatory and Hebbian inhibitory neurons
(d)
Excitatory and anti-Hebbian inhibitory neurons
(e)
Excitatory, Hebbian inhibitory and anti-Hebbian inhibitory neurons
(f)
Excitatory stimulation only
(g)
Inhibitory stimulation only
(1)
(2)
Postsynaptic neurons i
Presynaptic neurons j
Neurons
Time (s)
Preliminary tests
Architecture retained
Posterior tests

Figure 3.2: Learning of 2 stimuli in networks of QIF neurons. (a) Experimental protocol consisting of the stimulation of two non-overlapping neuronal populations of QIF neurons with plastic synapses. Stimuli are presented in temporal alternation. (b) Results for a network with 100% of excitatory neurons. (c) Results for a network with 80% of excitatory and 20% of Hebbian inhibitory neurons. (d) Results for a network with 80% of excitatory and 20% of anti-Hebbian inhibitory neurons. (e) Results for a network with 80% of excitatory and 20% of inhibitory neurons (including 10% of Hebbian and 10% of anti-Hebbian). (f) Results for a network with 80% of excitatory and 20% of inhibitory neurons (including 10% of Hebbian and 10% of anti-Hebbian) where only excitatory neurons are stimulated. (g) Results for a network with 80% of excitatory and 20% of inhibitory neurons (including 10% of Hebbian and 10% of anti-Hebbian) where only inhibitory neurons are stimulated. The results are reported at different moments of the stimulation. The time T0 corresponds to the beginning of the simulation before the learning phase. The time T1 corresponds to an early moment during the learning phase. Time T2 corresponds to the end of the learning phase and the beginning of the resting-state. Panels (1) represent the weight matrices at times T0, T1 and T2: the color denotes if the connection is excitatory (red) or inhibitory (blue) or absent (white). Panels (2) are raster plots during the simulation, displaying the firing times of excitatory (red dots) and inhibitory (blue dots) neurons. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

3.3.1.2 Post-learning resting-state

During the post-learning resting-state (time column T2 in Fig. 3.2), the inhibitory anti-Hebbian neurons play a major role. First, we observe that they significantly reduce the firing rate of neurons in their cluster, thus preventing an explosion of network activity (see raster plot (2) of panel (d)). On the contrary with Hebbian inhibitory neurons in panel (c) or without inhibition in panel (b), we clearly observe an explosion of the firing rate of the neurons in raster plots (2). Also, this "feedback inhibition" allows the neurons of the module to admit a certain degree of synchronization between them (degree highly depending on the associated global coupling).

Instead, in panels (f) and (g), it is the Hebbian inhibitory neurons that reduce excitatory firing. Indeed, since most excitatory neurons are not associated with inhibitors (i.e. no E-I connections), they receive inhibition from all Hebbian ones through maintained inhibitory to excitatory (I-E) connections. In this sense, in both cases, we can observe states of A-I firing with some epoch of synchrony of all excitatory neurons (since no structural modules are formed) in raster plot (2). The synchrony is logically more global in panel (f) than in panel (g) because the coupling weights are stronger. Finally, while in panel (f) the structural modules were not present in any connections (neither in E-E, E-I, I-E or I-I connections), the experiment of panel (g) shows that I-I connections are not sufficient to induce spontaneous recalls of both clusters.

Coming back to panel (c), the presence of Hebbian inhibitory neurons in each module allows them not to spike at the same time by repulsing the others when a module spikes and thus not to have synchronization between the modules at rest. However, alone this inhibition tends to completely stop the firing of the other modules and causes a single module to spike with a high firing in the manner of a "winner-takes-all" (see raster plot (2) of panel (c)). Therefore, in addition to the points described above, the addition of inhibitory anti-Hebbian neurons in each module allows a temporal alternation of activating clusters through the application of a strong self-inhibition. Note that the presence of a certain amount of noise in the dynamics of the neurons also facilitate this random alternation.

Thus, the combination of these two inhibitions compensates for their imperfections while maintaining their strengths. The presence of these two types of inhibitory neurons in each module (see Fig. 3.2 (e)) allows to display neurons firing in an asynchronous irregular way with a low firing rate while admitting in a totally random way spontaneous recalls of the learned items, characterized by relatively synchronous firing of only the neurons belonging to the clusters. Moreover, this particular dynamics could favour a certain maintenance of the learned structure in the long-term thanks to the alternating recalls of the patterns in the dynamics. This notion of maintenance and consolidation will be further developed in the next section.

3.3.1.3 Generalization

We have performed the experiment for different numbers of inhibitory neurons and we have been able to conclude that to satisfy all the constraints during learning and resting for two modules, we need at least 4 inhibitory neurons (1 inhibitory anti-Hebbian neuron and 1 inhibitory Hebbian neuron per module) as summarised in Fig. 3.3 (a). Indeed, it appears that the number of excitatory and inhibitory neurons present in a cluster is not an important parameter, but the presence of one neuron of each type seems sufficient (i.e., one excitatory, one inhibitory Hebbian and one inhibitory anti-Hebbian). We note that more inhibitory units are needed to stabilise $M > 2$ modular structures compared with the oscillator models of the previous chapter ($M - 1$ inhibitory neurons previously versus $2M$ inhibitory neurons here).

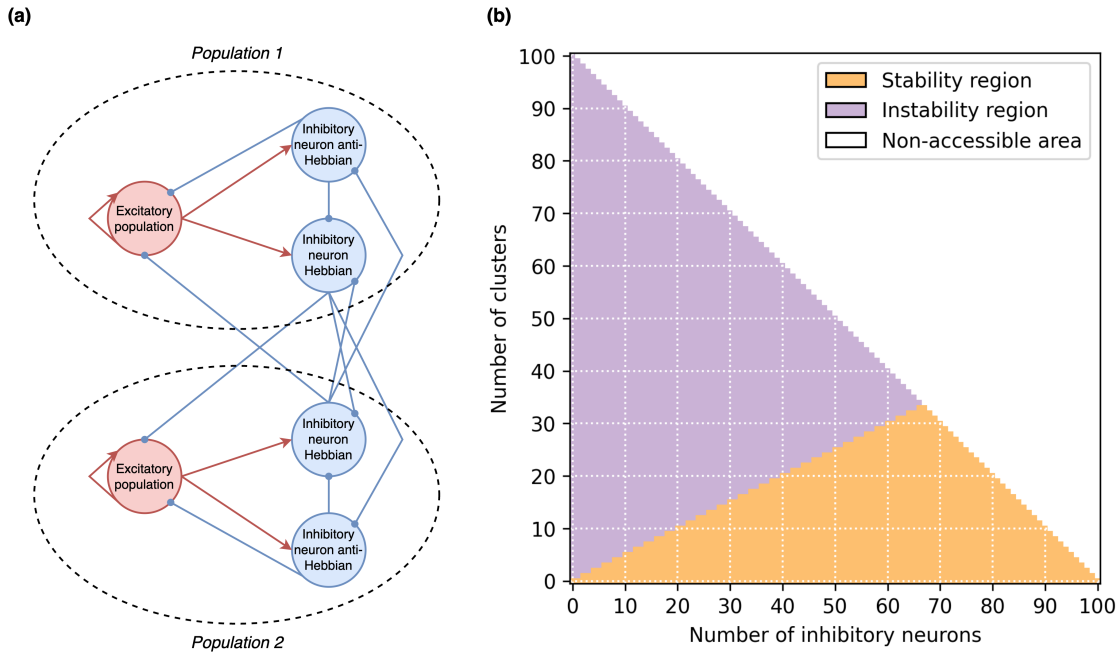


Figure 3.3: Minimal conditions to get stable modular structures. (a) Schematic diagram representing the connectivity of 2 stable populations of neurons after learning. Each population is composed at least of a population of excitatory neurons (in red), one anti-Hebbian inhibitory neuron and one Hebbian inhibitory neuron (in blue). Dashed circles identify groups of neurons admitting synchronization epochs. (b) Regions of stability (orange) and instability (purple) of structural modules versus the number of initial structural clusters in the weighted connectivity matrix and the number of inhibitory neurons present in the network. The line separating the two regions represents the upper limit to observe M independent clusters versus the number of inhibitory neurons. This limit corresponds to $2M$ inhibitory neurons. Note that the graph has been realized for a network made of $N = N_e + N_{hi} + N_{ai} = 100$ neurons in total, of which N_e are excitatory and N_{hi} and N_{ai} inhibitory, this constraint is at the origin of the non-accessible areas (white regions).

Subsequently, we can generalize this experiment to M modules. To do so, we studied the stability of different configurations of amount of inhibition and structural clusters in the network. Specifically, we perform a stability analysis where the objective is to determine the conditions under which the network during the resting-state admits stable dynamics (i.e. all the learned modular structures can be spontaneously recalled properly in an A-I state) or unstable dynamics (i.e. at least one structural cluster is not recalled properly).

Thus, we found that to ensure satisfactory learning and recall of them, at least $2M$ inhibitory neurons (M Hebbian and M anti-Hebbian, one per module) are required in the network, as summarized in Fig. 3.3 (b). Furthermore, for a network of N neurons, in theory a maximum number of $M = \frac{N}{3}$ modules (each one made of 1 excitatory neuron and 2 inhibitory ones) can be learned and recalled. For a more realistic network of 80% excitatory neurons and of 20% inhibitory neurons (i.e.,10% for each inhibitory neuron

population), the theoretical limit is fixed to $M = 0.1N$ modules.

Finally, it should be noted that up to now and in the remaining of this study, the input current applied to the neurons is considered to be constant (to give neurons with a frequency of 50 Hz). Nevertheless, tests have been set up considering: other frequency values, time-varying input signal (sinusoidal), randomly stimulated neurons (see Fig. B.5 of Appendix B.1), and random input amplitudes (see Fig. B.6 of Appendix B.1). Most of these alternative stimulation protocols produced somewhat more complex structures within the modules, but the main results remain identical, showing that the most important aspect is the averaged spatio-temporal correlations of the inputs applied and not so much their amplitude.

3.3.2 Role of spontaneous recall: Maintain and consolidate the learned structure

In the previous section we demonstrated that the need of two types of plasticity for inhibitory neurons (Hebbian and anti-Hebbian) results in coherent spontaneous recalls of learn items at rest (see panel (e) of Fig. 3.2). However, in panels (c) and (d) of Fig. 3.2 with only one type of inhibition, we have seen that even if the recalls are not guaranteed, the structure remains relatively learnable and maintainable for some time. Therefore, this may raise questions about the real interest of these spontaneous recalls in the long-term. Therefore, the objective is now to understand the function and impact of these spontaneous recalls observed during post-learning resting-state.

3.3.2.1 Spontaneous recalls after learning

To further investigate the possible function of these recalls, we start from an already learned structure of two modules obtained in the same way as explained in the previous part. The network is allowed to evolve spontaneously for a certain time without applying any stimulation (resting-state). We consider that the weights of the connectivity matrix at the beginning of the simulation (see Fig. 3.4 (a)), have not yet reached their maximum value (0.7 for excitatory links and -0.7 for inhibitory links) to observe any reinforcement in weight connectivity. Apart from this aspect, the two modules largely satisfy the minimal conditions of inhibition previously described.

The first thing we observe at the end of the simulation is that the structure is maintained as it was originally (two well-formed and completely decoupled modules) and that it also seems consolidated. Indeed, if we observe the weights matrix at the end of simulation (Fig. 3.4 (a)) and its distribution (Fig. 3.4 (b)), the network admits stronger links than before (both inhibitory and excitatory) with weights close to their maxima. Despite this, the modules remain decoupled, showing that it is only the connections already present that are reinforced.

Concerning the dynamics of the network (see Fig. 3.4 (c)), we observe that the spontaneous recalls, characterized by neurons from the same cluster spiking at a similar instant, lead to an increase in the mean firing rate and of the order parameter (i.e. synchrony) of the corresponding cluster. More importantly, these recalls are correlated with an increase

in the mean change rate of weights of that module. In other words, a spontaneous recall of a cluster in the dynamics is directly related to a consolidation from a structural point of view of this same cluster. We also notice that during a recall, the more important the firing rate and the order parameter of a cluster are, the less important those of the other module are, and consequently the same goes for the change rate of weights. In particular, this allows the modules not to consolidate with each other, which is explained by the action of the Hebbian inhibitory neurons. The rest of the time, when no recalls are observed (e.g. from times 2768s to 2770s), the network is in A-I state characterized notably by a low and irregular firing rate, a low degree of synchrony and coefficient of variation close to 1 as a theoretical poisson process (see Fig. 3.4 (d)). Thus, in the absence of recalls, the connectivity of the modules appears to decrease slowly (i.e. characterized by a negative change rate) due to the forgetting term in the weight adaptation (see Sec. 3.2.2).

In addition to appearing randomly, spontaneous recalls can be partial when only a part of the pattern emerges, which consequently has less impact on the consolidation process. Furthermore, despite the randomness appearance of these synchronized patterns, it appears that spontaneous recalls, whether partial or complete, occur with a very similar probability for both clusters, as evidenced by the probability density functions of the order parameters in Fig. 3.4 (e). This results in a similar and consistent consolidation of the different learn items. The PDF of the order parameter also confirms that the network never reaches a fully synchronous regime and is on average in a fairly asynchronous regime (see green curve Fig. 3.4 (e) with a maxima at $R = 0.2$). On the same idea, the free energy associated with the Kuramoto order parameter in Fig. 3.4 (f), reveals relatively high values for high R indicating a less stable state or uncertainty in synchronisation for the network and the clusters even if no metastability is clearly observed at this stage. In fact, the network exhibit transient episodes of synchronous activity, then transit to asynchronous states [107, 110, 111, 149]. As a result, the system is never fully synchronized, and can explore different states characterized by a continuous level of different synchronization values. In particular, this allows the network to maintain a certain level of flexibility and adaptability in its representations and responses [375], as evidenced by the richness of recall dynamics observed.

It is quite complicated to quantify the amplitude of the change rate of weights in a cluster during a recall and a forgetting. Indeed, this rate depends on the current value of the weights w_{ij} and the number of neurons spiking at the time t. Thus, for maximum weight capacity, a potentiation (and so a recall) has almost no impact, whereas a depression due to forgetting is more impactful (see soft bound functions in Fig. 3.1 (d) and (e)). Also, the neurons spiking in a A-I manner, the weights don't adapt in a regular way. Therefore, in this state, we observe small random epochs of decrease in very few weights in the cluster. On the contrary, in recalls, a larger number of neurons spike in a similar instant, implying the adaptation of a large number of weights in the cluster. Moreover, the potentiation logically depends on the precise timing between neurons' spikes (STDP), so the more synchronized the neurons are in the recall, the more impactful the recall will be. In spite of this, we can get a general idea of the impact of a recall on a synaptic weight compared to a forgetting thanks to the plasticity function (see Fig. 3.1 (a)). Indeed, for uncorrelated

spikes in the A-I state (i.e. for $|\Delta t| > 0.5$), the weight depress with an amplitude of $\Lambda(\Delta t) \approx -0.1$ (due to the forgetting term), if they are fully synchronized (i.e. $\Delta t = 0$), $\Lambda(\Delta t) = 2.347$. Thus, Eq. (3.10) becomes:

$$\tau_l \dot{w_{ij}} = \begin{cases} 2.347 \tanh(\lambda(1 - w_{ij})), & \text{for synchronized spikes during recall,} \\ -0.1 \tanh(\lambda w_{ij}), & \text{for uncorrelated spikes during A-I state.} \end{cases} \tag{3.18}$$

If we consider an average weight $w_{ij} = 0.5$, the soft bound functions become $\tanh(\lambda(1 - w_{ij})) = \tanh(\lambda w_{ij}) \approx 1$, and we obtain:

$$\tau_l \dot{w_{ij}} = \begin{cases} 2.347, & \text{for synchronized spikes during recall,} \\ -0.1, & \text{for uncorrelated spikes during A-I state.} \end{cases} \tag{3.19}$$

Dividing these two constant quantities of potentiation and depression, we can conclude that to compensate for such recall, $2.347/0.1 = 23.47$ uncorrelated spikes inducing forgetting would be required. If we consider relatively regular firing of 2 Hz over a given time window, $23.47/2 = 11.735$ seconds of forgetting are required to compensate for the potentiation of a recall in a weight. This shows that the formation of recalls from time to time is sufficient for the maintenance of the structure.

This consolidation experiment was repeated for longer simulations (10000 seconds) and observed similar results when reaching the maximum values of the weights, confirming that no long-term forgetting is observable in this resting-state (i.e. memory not deteriorated over time) and that this state makes it possible to complete the learning previously initiated. These findings on spontaneous recall and consolidation were also obtained with different structure configurations such as multi-clusters (see Figs. B.1 and B.2 of Appendix B.1) and partially learned modules (see Figs. B.5 and B.6 of Appendix B.1). We conclude that these alternating recalls of the patterns in the dynamics have the role to sustain the learned memory items over the long-term and if necessary to consolidate them in the face of the natural tendency of the network to forget.

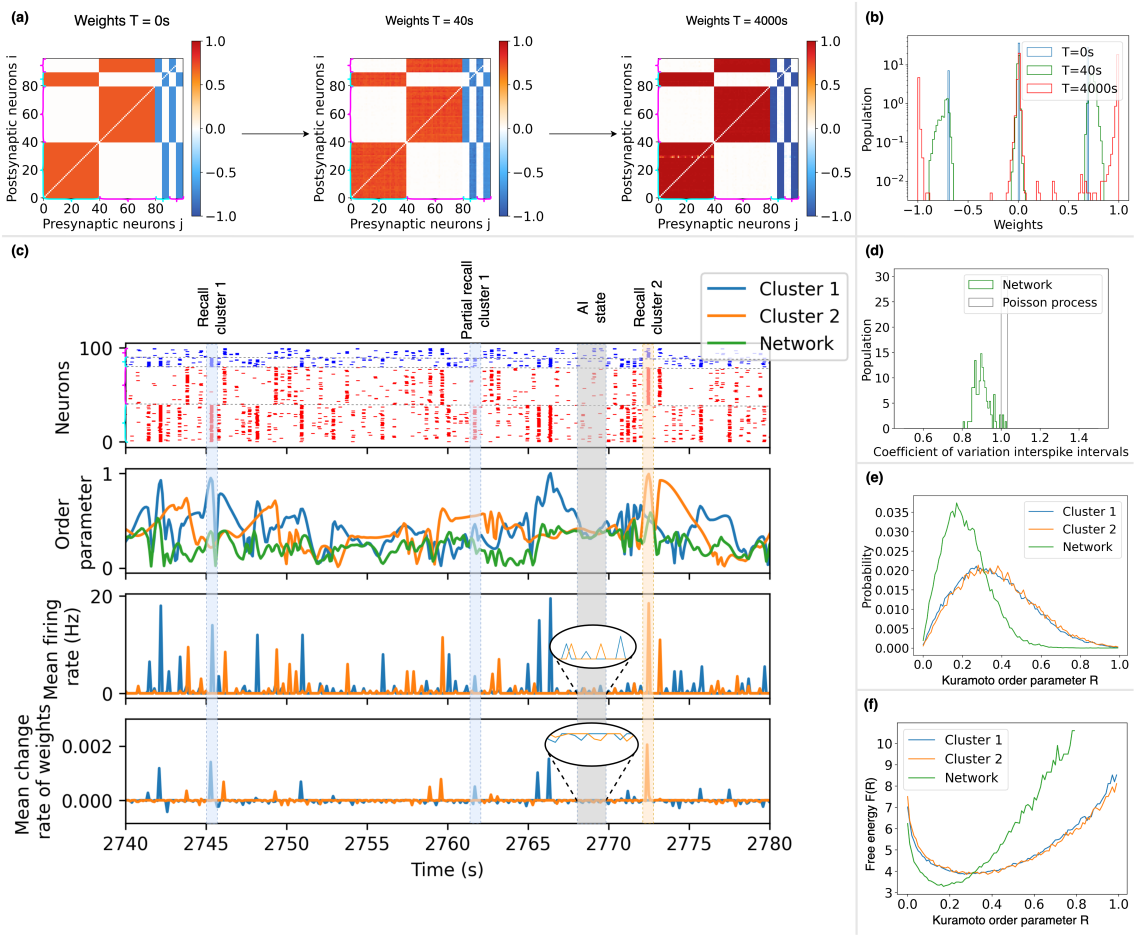


Figure 3.4: Spontaneous activity of a network made of 2 structural modules during resting-state. (a) Weights matrices at the different instant of the simulation (t=0s, T=40s, T=4000s). (b) Distributions (in log scale) of the weights in the network at the different instant of the simulation (t=0s (in blue), T=40s (in green), T=4000s (in red)). (c) Network dynamics during resting-state with respectively: the spike raster plot with excitatory neurons in red, inhibitory in blue, cluster 1 with neurons 0-39 and 80-89 and cluster 2 with neurons 40-79 and 90-99, the Kuramoto order parameter of the clusters, with cluster 1 in blue, cluster 2 in orange and the network in green, the mean firing rate over time of the clusters, with cluster 1 in blue, cluster 2 in orange and the mean change rate of weights of the clusters (positive: increase in the weights, negative: decrease in the weights; highlighted by the loupe), with cluster 1 in blue, cluster 2 in orange. (d) Distributions of the coefficient of variation of individual neurons interspike intervals during the simulation (in green) and of the theoretical poisson process (in grey). (e) The probability density function during the simulation as a function of the Kuramoto order parameter R for cluster 1 (in blue), cluster 2 (in orange) and the network (in green). (f) The free energy $F(R)$ during the simulation as a function of the Kuramoto order parameter R for cluster 1 (in blue), cluster 2 (in orange) and the network (in green). The cyan and magenta brackets represent clusters 1 and 2 respectively in the raster plot and the weight matrices.

3.3.2.2 Spontaneous recalls without structural modules

In order to better understand how memories are encoded, we replicated the previous experiment with the initial weights matrix of Fig. 3.4 (a) but this time considering random weights between excitatory neurons as represented in Fig. 3.5 (a). In other words, the structural modules are no longer present in the connectivity, only connections involving inhibitory neurons remain unchanged. This experiment resembles the one carried out in the previous chapter in Fig. 2.8.

Surprisingly, in the resting-state, unlike with oscillator models, we can still observe the previously described spontaneous recalls associated to reinforcement of the connectivity (see Fig. 3.5 (c)). Indeed, although the excitatory neurons all stimulate each other with different degrees of connectivity, when a large enough number of neurons belonging to a previously learned cluster spike, they reactivate the associated inhibitors allowing partial recall of the learned items. More precisely, once activated, inhibitory neurons associated with a cluster, prevent the firing of the other clusters (with feed-forward inhibition) and promote synchronization within the cluster (with feedback inhibition).

Moreover, the connections already present to and from the inhibitors seem to reinforce, while the modules (i.e. the links between excitatory neurons) seem to partially reappear and so reconstruct themselves in the long-term as shown in Fig. 3.5 (b). We can therefore predict that, over a sufficiently long period of time, the original learned structure of Fig. 3.4 (a) will somewhat reform, or at least partially. We thus observe results similar to those with oscillator models in the previous chapter with the possibility of recalling information only through links associated to inhibitory neurons [38].

Recall of an item is even possible with the reactivation of the associated inhibitory neurons when the pattern is presented as shown in Fig. 3.5 (c) at time $t = 40s$. However, external recall of the modules by partial stimulation of some excitatory neurons seems impossible since connections between these entities are no longer fully present. More details on external recall will be given in the next section.

These results confirm that **long-term memories are stored in connections involving inhibitory neurons**. Thus, the modular structures can admit weaker connectivities (or completely deconstructed). As long as the modules or the memory items are sufficiently learned and encoded in weights involving inhibitory neurons (i.e. E-I and I-E connections), the recovery and consolidation of the learned memories is possible.

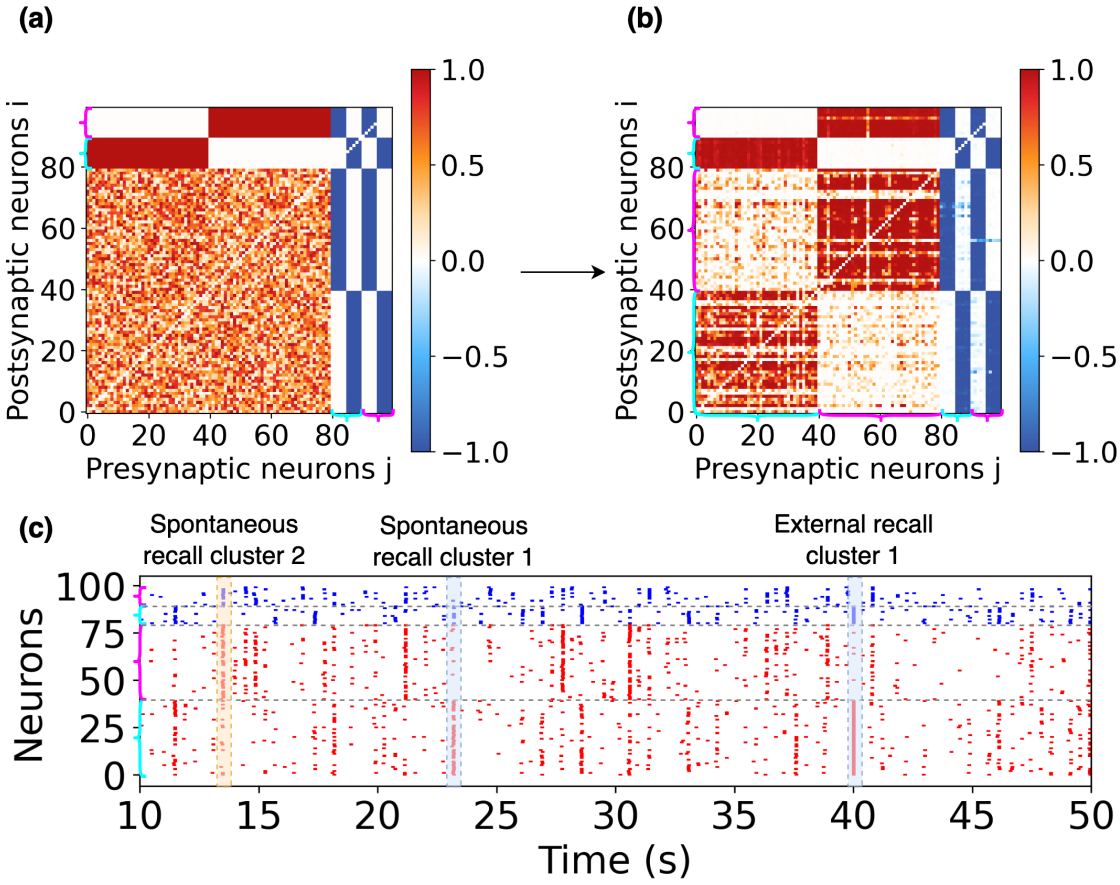


Figure 3.5: Resting-state of a modular network of 100 QIF neurons with only links associated with inhibitory neurons. (a) Weighted connectivity matrix at the beginning of the resting stage. (b) Weighted connectivity matrix after a period of rest of 4000s. (c) Raster plot of spikes train (excitatory neurons in red and inhibitory neurons in blue) during resting-state with an external recall at time 40s (i.e. stimulation of one excitatory module). The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

3.3.3 Formation and consolidation of overlapping memories

To complete this study, we consider the case where there is an overlap between the two applied stimuli (i.e. the 2 modules share some neurons in common). The idea is to reproduce the experiment of Sec. 2.3.3 of the previous chapter including neurons showing a certain "mixed selectivity" by encoding two stimuli at the same time. For this purpose, we follow the same protocol as in Fig. 3.2 (a), but this time, during the learning phase, eight neurons are always stimulated, regardless of the selected area, as illustrated in the protocol in Fig. 3.6 (a). In this case, to facilitate the formation of the connections, we strictly alternate the areas stimulated.

The results depicted in Fig. 3.6 (b) are similar to previously reported observations in Sec. 3.3.1 with the formation of two clusters due to the adaptation to the stimuli (see matrices (1) at times T1 and T2). The difference and direct effect of this overlap at the end of the learning is the creation of structural "hub" neurons in the connectivity (see matrices (1) at times T1 and T2). These hubs are strongly connected to both modules (input and output excitatory links). They are weakly affected by the Hebbian inhibition since they are considered to belong to both populations but consequently, but they are affected by the two anti-Hebbian inhibitions as represented in the diagram in Fig. 3.6 (c).

Regarding the spiking dynamics during post-learning, we observe richer patterns, see raster plot (2). Indeed in addition to the omnipresence of the A-I state as before, the spontaneous recalls are more varied with:

- Synchronous spiking of one cluster together with the hub neurons
- Synchronous spiking of one cluster without the hub neurons
- Synchronous spiking of the hub neurons alone

We also sometimes encounter combinations of these basic patterns with: synchronous spiking of one cluster and the hub neurons directly followed by the synchronous spiking of the other cluster, synchronous spiking of one cluster without the hub neurons directly followed by the synchronous spiking of the hub neurons, synchronous spiking of the hub neurons alone followed by the synchronous spiking of one cluster and synchronous spiking of the hub neurons alone followed by the synchronous spiking of both clusters.

We can conclude that despite the wide variety of behaviours that the network can display, they seem to respect a certain coherence thanks again to the inhibition allowing the uncorrelated items not to synchronize. Therefore, we can say that these recall patterns promote the long-term maintenance of the modules and the hubs present in the structure as in the previous experiment. Thus, this model can exhibit **neurons admitting persistent mixed selectivity** and so is able of encoding multiple stimuli in contrast to our study with oscillators [38] in the previous chapter.

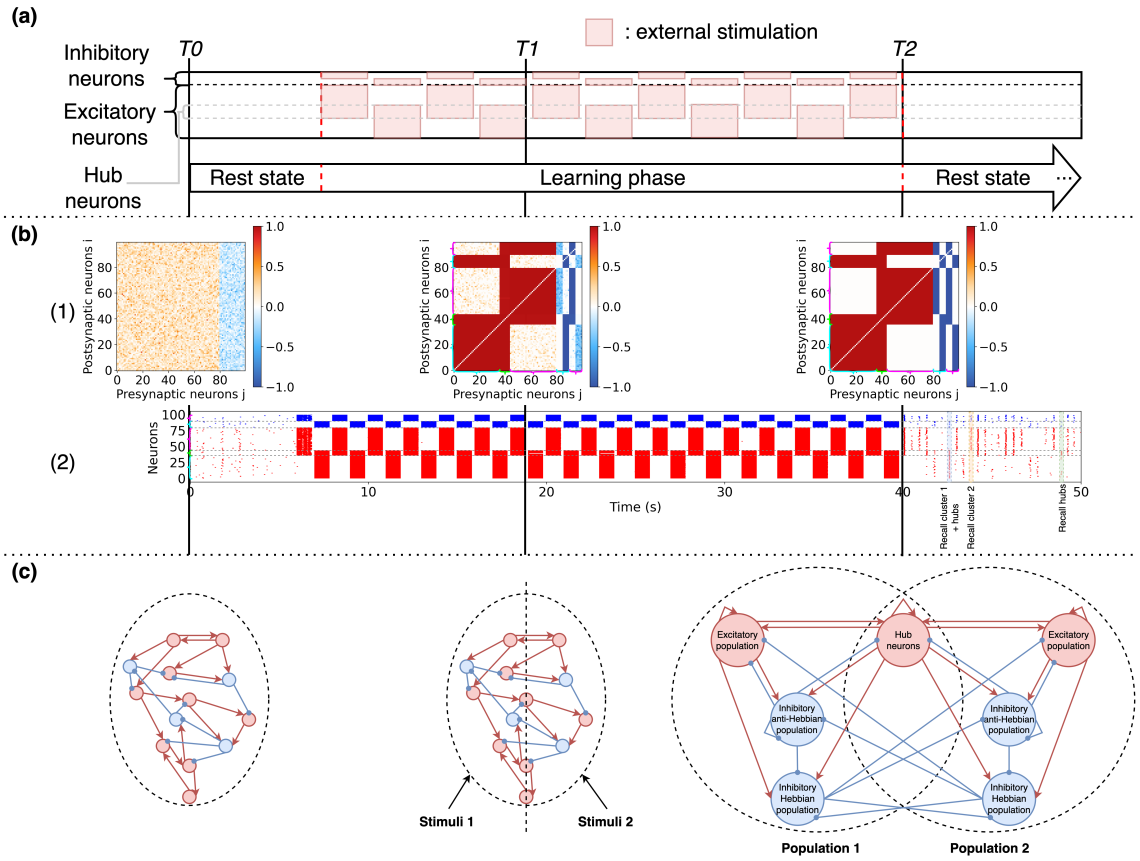


Figure 3.6: Learning of 2 overlapping stimuli in a network 100 QIF neurons. (a) Schema of the experiment protocol showing that the two presented stimuli involve 8 shared neurons. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 3.2. The three typical spontaneous recalls observed are indicated during rest state. (c) Schematic diagrams representing the formation after learning of the connectivity of 2 stable populations of neurons and hub neurons. Each population is composed at least of a population of excitatory neurons and a population of hub neurons (in red), one population of anti-Hebbian inhibitory neurons and one population of Hebbian inhibitory neurons (in blue). Dashed circles identify groups of neurons admitting synchronization epochs. The cyan, magenta and green brackets represent clusters 1, 2 and the hubs respectively when they are visible in weight matrices and raster plots.

As in the previous experiment, we also considered different amounts of inhibition and overlapping neurons in order to quantify the conditions under which the two modules admit stable dynamics. Surprisingly, unlike the study with oscillators in Sec. 2.3.3, we found no correlation between these two quantities and the stability of the dynamics. We conclude that despite the connections between the modules, the strong inhibition compensates and makes it very unlikely that the two modules will discharge simultaneously. Nevertheless,

one rule holds true in this part: the size of the overlap must be smaller than the size of the individual cluster. If the overlap becomes larger than the clusters, it becomes dominant over the others.

3.3.4 Provoke external induced recall

To complete this study, we focus on the elements that provoke recall in addition to the inhibition. More precisely, we focus on the number of neurons firing at the same instant necessary to cause a total recall of a memory item. To do so, the idea is to induce recall through the action of an external input applied to some excitatory neurons. Thus, we expect that, very shortly after stimulating some neurons, the other neurons of the associated cluster will spike.
We identify three main parameters influencing the impact of the stimulation on the network:

- The duration of the stimulation (in s)
- The intensity of the stimulation (to control the frequency of neurons in Hz or spikes/s)
- The number of neurons stimulated

In addition to these values quantifying the external input, other parameters such as the global couplings (for excitation and inhibition inputs), the local couplings (i.e. weights), the amount of noise and the excitability parameters, impact the reaction of a neuron to a stimulus. Thus, for the sake of simplicity, we will first consider homogenous neurons not subjected to noise in order to quantify only the impact of the stimuli; we will return to normal conditions afterwards. Also the frequency and the duration of the stimulation being directly linked, we decide to fix them in order to have 1 spike per stimulation per neuron. Thus, the protocol consists, in a resting-state with the same initial connectivity of Fig. 3.4 (a), to apply an external input which gives a neuron firing rate of 100 Hz for 0.01s (inducing 1 spike for each stimulated neuron) to a variable number of neurons at the time 17s of the simulation.

In a first step, we experiment in the condition of a totally silent network (i.e. without noise and with excitability parameters of 0). In this case, the stimulation of a single neuron is sufficient to provoke a total recall of the associated module. When using constant negative excitability parameters, one additional neuron to be stimulated was needed per Hz (e.g. for η involving a firing rate of -1 Hz: 2 neurons needed, for $\eta =$ involving a firing rate of -2 Hz: 3 neurons needed). When the excitability parameters are distributed and noise is present, it becomes much more complicated to quantify a precise rule in such a complex network. Indeed, the stochasticity and heterogeneity of the neurons make them react very differently to a stimulus, and the randomly spiking of inhibitory neurons make the recalls even more unpredictable.

Returning in the normal conditions of the previous experiments (i.e. with heterogeneity and noise), we found a fairly satisfactory value with 8 stimulated neurons (as shown in Fig. 3.7 (a)). This one is obtained using the lower bounds of the excitability and noise

distributions (i.e. involving firing rates of -2 Hz and -5 Hz) and by relying on observations with constant values. However, it should be bear in mind that this value remains purely heuristic and does not guarantee a recall in all circumstances, for the reasons mentioned above.

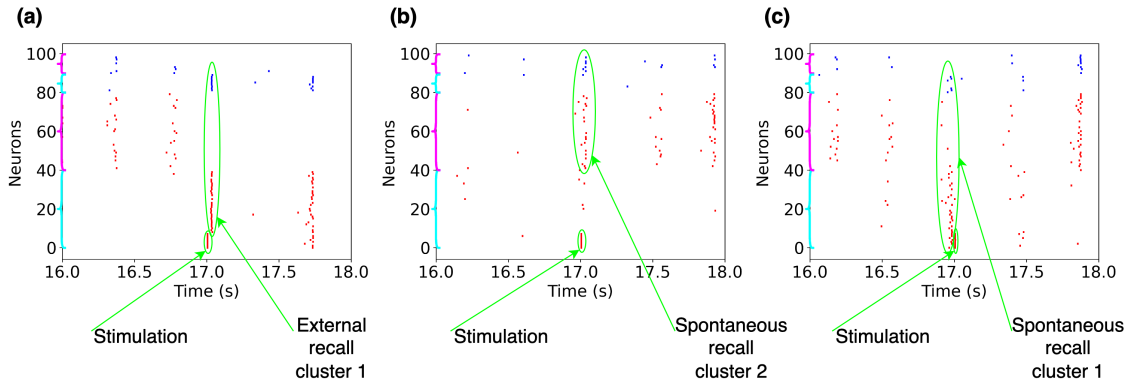


Figure 3.7: Partial stimulation of a cluster in a network of 100 QIF neurons during resting-state. (a) Case where an external recall of cluster 1 is provoked after stimulation. (b) Case where a spontaneous recall of cluster 2 occurs during stimulation and no external recall is observed. (c) Case where a spontaneous recall of cluster 1 occurs before stimulation and no external recall is observed. In each case, a stimulation which gives neuron firing rate of 100 Hz is applied to 8 neurons (index 0-7) of cluster 1 for 0.01s at time 17.0s under the same condition as in Fig. 3.4. The cyan and magenta brackets represent clusters 1 and 2 respectively in raster plots.

Thus, we observe in this case that the external recall is characterized by the spiking of the excitatory and inhibitory neurons of the target cluster at a similar time (almost in synchrony) shortly after the spiking of the stimulated neurons. This phenomenon is very similar to the spontaneous recall observed in the previous section with the complete reactivation of a memory item accompanied by the extinction of the activity of the neurons not belonging to this item. The difference lies in the voluntary and controlled aspect of this reactivation and the small temporal lag between the trigger and the action. We observe in all experiments that there is a clear correlation between the number of stimulated neurons and the shortness of this decay of few milliseconds between stimulation and recall (i.e. more spikes induce less delay). Also, it is logically confirmed that the higher the number of stimulated neurons, the higher the chances of provoking a recall. Finally, this experiment shows the potential of the network to perform tasks found in AI systems such as recognition, generation or reconstruction of a memory item (i.e. module) thanks to incomplete information provided. Indeed, here the recall is not spontaneously induced but is a fast response to a particular request.

It should be noted that the directly stimulated neurons admit only one spike after the stimulation, despite the cascade of spikes from neurons of the same cluster. This is explained firstly by the refractory period of the neurons, which makes them insensitive for

a certain time to other sensory inputs after their discharge. Secondly, the self-inhibition produced by the anti-Hebbian inhibitory neurons greatly discharges their potential, making it impossible to spike immediately after the recall. The only way to provoke further recalls and so spikes would be to present the external stimuli for a longer duration. For similar reasons, we observe cases where the external recall does not occur in Fig. 3.7 (b) and (c). In panel (b), a spontaneous recall of the other cluster occurs at a similar time of the stimulation. As a result, the Hebbian inhibitory neurons associated with this other cluster inhibit the other one and make the action of the external stimuli inefficient. In panel (c), it is a spontaneous recall of the targeted cluster that occurs just before the stimulation. This induces a strong self-inhibition to the cluster's neurons (action of the anti-Hebbian inhibitory neurons) making the external input not effective. After the induced external recall, the network returns to its initial state of spontaneous activity. Nevertheless, these behaviours remain coherent according to our architecture by avoiding the spiking of different clusters at a similar instant and the explosion of the network activity.

3.4 Summary and conclusions

Spontaneous recall of memories during sleep, allowing their consolidation, is a phenomenon extensively studied [218, 291, 365, 366, 393], although rarely implemented in modelling. In this chapter, we have investigated how inputs can be learned to shape modular structures and how to maintain these structures in the long-term in an excitatory-inhibitory neural network made of QIF neurons subject to STDP. Analyses of chapters 1 and 2 with phase models, provided insight into how temporal alternating external stimuli form structural clusters and on how inhibition modulates the degree of synchrony of the network to maintain these assemblies. The present chapter has extended these findings to spiking neurons admitting A-I dynamics. In particular, we focused on the role of inhibition in the emergence of spontaneous internal recall promoting the consolidation of long-term memory.

In summary, we have shown that inhibition in general first enables to regulate network activity by establishing an excitatory-inhibitory balance that prevents runaway excitation and promotes input selectivity. Secondly, feedback inhibition, provided by anti-Hebbian-STDP, promotes synchronized dynamics within the structural modules. Finally, feed-forward inhibition, brought by Hebbian-STDP, leads to desynchronization between neurons of different clusters. As a result, we have shown that the presence of these two groups of inhibitory neurons is necessary to guarantee the emergence of the modular structures and their spontaneous recall at rest.

In a second part, we have demonstrated that these **spontaneous recalls in the resting network are correlated with strengthening of synaptic connectivity** of the associated structural clusters, which can be compared to an autonomous memory consolidation process. Thus, for a long period of spontaneous activity, the network maintains its structure despite the constant adaptation and decay of its weights, and therefore prevents long-term forgetting. These internal replays can take various forms to support the memorization of complex information, such as those with mixed selectivity, which enables several sources of information to be represented and processed simultaneously. Moreover,

these recalls of information can also be induced by the action of an external stimulus.

As a final conclusion, we again established the **relationship between the number of inhibitory neurons in the network and the number of neural assemblies that can sustain independent spontaneous recall and so be maintained**. Thus, for a network composed of 20% inhibitory neurons, the storage capacity of a network of N neurons is $0.1N$ memory items. Moreover, this relationship is further highlighted in the experiment in Fig. 3.5, which shows that information is stored in the links associated with the inhibitory units (E-I and I-E links) and that these links are sufficient for spontaneous recall and recognition tasks.

The impact of parameters on network dynamics, particularly on spontaneous recalls, was not extensively discussed in this study. Reproducing experiments under various conditions revealed that as soon as the three populations of neurons are present in a cluster with a dominance of global inhibition couplings, spontaneous recalls can occur in the network dynamics. Increasing the excitability and noise amplitude of neurons subsequently increases the frequency of recalls, while introducing more parameter heterogeneity leads to a more asynchronous and irregular network, necessitating increased anti-Hebbian coupling to enhance recall synchrony. Ultimately, the network remains quite flexible on its dynamics and this tuning is more a matter of a balance between the A-I states and the frequency of the spontaneous recalls while keeping relatively low firing rates, which justifies the chosen parameters. Similarly, the forgetting term in the plasticity functions strikes a trade-off between recalls frequency discussed above and potentiation amplitude. While one may question the choice of such a forgetting parameter favouring or not the consolidation of clusters by recalls, the study of Fig. 3.4 revealed that a recall largely compensate for the forgetting generated during the A-I state.

In these experiments, we focused on using simple stimuli to be learned through constant external inputs. The amplitude of this external input being directly linked to the adaptation rate (like the amplitudes of the plasticity functions and the learning time scale), we set it to a relatively large value in order to ensure a rather rapid convergence of the weights (learning lasting 35s in Fig. 3.2). Indeed, the primary purpose of this study was not to show how to encode complex inputs but rather to explore their storage mechanisms. However, these results can generalized to more complex inputs. Thus, in Figs. B.1 and B.2 of Appendix B.1, we show the possibility to apply our paradigm to larger number of clusters. In Fig. B.3 of Appendix B.1, we show the possibility of sharing hubs between 4 different clusters and thus exhibit even more complex dynamics at rest. The results of Fig. B.4 of Appendix B.1 highlight the effect of a group of untrained neurons in the network. Finally, Figs. B.5 and B.6 of Appendix B.1 show the impact of a learning where (1) neurons in clusters are randomly stimulated and (2) neurons in clusters receive inputs of random values. All of these alternative stimulation protocols provide insight into how to encode complex inputs. Nevertheless, in the following chapter we will apply this model to real sensory inputs in order to study the benefits of such learning and resting dynamics in recognition and generation tasks.

Chapter 4

Recognition and generation of audio-visual information

Contents

In the previous chapters, we modelled a spiking neural network with the ability to efficiently learn and retain long-term memories. In this chapter, we apply this architecture to encode real sensory information. The idea is to learn audio-visual features for recognition and generation tasks that can be found both in the field of AI and cognitive sciences. In particular, the focus will be on the integration of these two modalities to achieve an efficient multimodal process.

4.1 Introduction

The human brain processes and interprets sensory information from the environment using specific sensory receptors associated with a particular sense, which detect stimuli and transmit signals to the brain for processing. These senses include vision and audition, on which we will focus on in this part. The sense of vision is mediated by the eyes and the visual cortex in the brain. Specialized cells in the eyes, called photoreceptors, detect light and convert it into electrical signals [259]. More precisely, these photoreceptor cells are of two main types: (1) rods that are responsible for low-light vision and (2) cones that are responsible for color vision and visual acuity [221]. The signals are then transmitted through the optic nerve to the visual cortex, primarily located in the brain's occipital lobe [190, 155]. The visual cortex processes the signals and interprets them into the perception of colors, shapes, depth, and motion [165, 405]. On the side of audition,

it involves detecting sound waves in the environment using the three main parts of the ear: the outer ear, middle ear, and inner ear [300, 307]. Sound waves enter the outer ear, pass through the auditory canal and reach the eardrum in the middle ear [168, 307]. The eardrum vibrates and transmits these vibrations via tiny bones in the middle ear to the cochlea [168, 307], a spiral-shaped structure in the inner ear lined with hair cells [98, 59]. The hair cells convert mechanical vibrations into electrical signals, responding to specific sound frequencies [98, 59]. These electrical signals are transmitted by the auditory nerve to the auditory cortex in the brain's temporal lobe [312]. The auditory cortex processes the signals, allowing us to perceive and interpret different aspects of sound, such as pitch, volume, and timbre [206, 423].

As a result, the pathways from sensory receptors to sensory cortices appear to be fairly separated from each other, which could explain the modular architecture of the brain [419]. Yet, within theses modules (i.e. primary sensory cortices), sensory information can shape neural connections through synaptic plasticity, in order to learn and form memories [156, 264]. This learning is facilitated by the repetition of sensory stimuli, which leads to the strengthening of neural connections in the relevant sensory cortices [125, 187]. This leads to enhanced perception and recognition of specific sensory inputs. Thus, after extracting relevant features and patterns from the input, the networks are able to identify or interpret learned sensory stimuli, such as objects, faces, sounds, speech, etc [27, 388]. Conversely, the brain has the ability to "generate" sensory information internally by creating a mental representation of it, through spontaneous activation of sensory brain areas, without direct external input [297, 223].

Thus, separately, these sensory cortices can process more or less complex features through the multiple hierarchical layers of brain regions. However, for more complex cognitive processes and sensory experiences, sensory informations need to be processed and integrated in higher-order brain regions [138, 265]. Higher-order brain regions, including association cortices and multimodal integration areas, are essential for integrating information from different sensory modalities to form associations between memories, emotions, and cognitive functions, resulting in a coherent perception of the world [19, 138, 265, 341]. Finally, feedback loops are involved, where higher-order brain regions also send signals back to sensory cortices, refining sensory perception, enabling the brain to process sensory information in both bottom-up and top-down manners, and contributing to cognitive functions such as attention, expectation, and memory retrieval [14, 138, 230, 191, 344].

However, how exactly the brain merges these modalities to **integrate multisensory information** in a coherent way is still open to debate [294, 382]. Several studies attempt to model this integration mechanism emphasizing the role of memory association [73, 163, 234], the way modalities fuse and accumulate effectively [35, 154], or the interest of this multimodality for machine learning tasks [54, 82, 282, 283, 418]. In our case, we have demonstrated the capacity of our spiking neural model to form and maintain over the long-term complex structures composed of modules and hub neurons, thanks in particular to the different inhibition mechanisms. We foresee that these hubs and inhibitions will play a key role in the coherent integration between the different learned items. Thus, a first objective is to illustrate the **emergence of modular structures** as before but this

time **through localized cognitive stimuli** where one part of the network learns visual features in the manner of the visual cortex and another part learns auditory features in the manner of the auditory cortex to show the aspect of modality segregation. At the same time, the aspect of integration should be highlighted by the **formation of hubs between these two modalities**. Then, the final objective is to demonstrate the effectiveness of this approach and architecture for cognitive tasks of recognition and generation, in both unimodal (i.e. the segregation aspect) and multimodal way (i.e. the integration aspect).

This last chapter is organized as follows. After this brief introduction to the process of sensory information, from its pathways to its integration passing through its learning for recognition and generation, the methods employed are presented in Sec. 4.2. We first describe the audio-visual dataset used, then explain the experimental protocol set up. Then, in the results in Sec. 4.3, we explore the operation and implications of the learning process and apply the trained network to recognition and generation tasks. Despite the simplicity of the data learned, the network appears to perform the different cognitive tasks efficiently, paving the way for more complex experiments.

4.2 Methods

This section presents the audio-visual dataset used and the pre-processing applied to extract the features to be learned by the neural network. Then, the architecture of the network is explained in order to understand the experimental protocol set up to learn the sensory informations and test them on recognition and generation tasks. Throughout this chapter, we consider the same model of QIF neurons subject to STDP described in Sec. 3.2 of the previous chapter. Unless otherwise stated, the characteristics and parameters of these neurons are analogous.

4.2.1 Extraction of audio-visual features

Although the main focus of this thesis is not to study the way to code informations but rather how to store and integrate them efficiently, in this chapter we have to question on the way to represent the sensory information we want to learn and how to interpret them. We saw in the introduction that some neurons respond selectively to specific features of stimuli and therefore emit spikes to their detection [105, 189]. Instead of relying only on individual neurons, the brain often uses populations of neurons to represent sensory information [305, 308]. This directly echoes to the concept of synergy, according to which the efficiency of the neural code depends both on the independent activity of individual cells in response to stimuli, and on synchronized redundancy among nearby neurons, leading to synergistic interactions that create efficient symbol encoding [57, 235, 260, 331, 332]. Then, neurons can encode information using two main strategies: rate coding and temporal coding. In rate coding, the intensity or strength of a stimulus is represented by the firing rate (frequency) of action potentials [66, 57, 116, 235, 316, 349]. In temporal coding, the timing and patterns of action potentials (spikes train) carry additional information about the stimulus [53, 57, 66, 79, 235, 249, 316, 425].

Thus, in this experiment, we consider a dataset of ten digits (i.e. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9). For the visual modality, we use the handwritten digits images from the well-known machine learning database Mixed National Institute of Standards and Technology (MNIST) [117, 240]. Concerning the auditory modality, we use spoken digits from the Free Spoken Digit Dataset (FSDD) [195] database. In this experiment, we will only consider one digit of each for each database (i.e. a total of 20 data items with 10 different images and 10 different sounds). At this stage, the aim is to highlight the learning and processing capacity of our architecture for a limited set of data, not to learn and test the entire database and compete with other machine learning algorithms. In order to learn these data in our network, it is necessary to pre-process them in order to extract the relevant features.

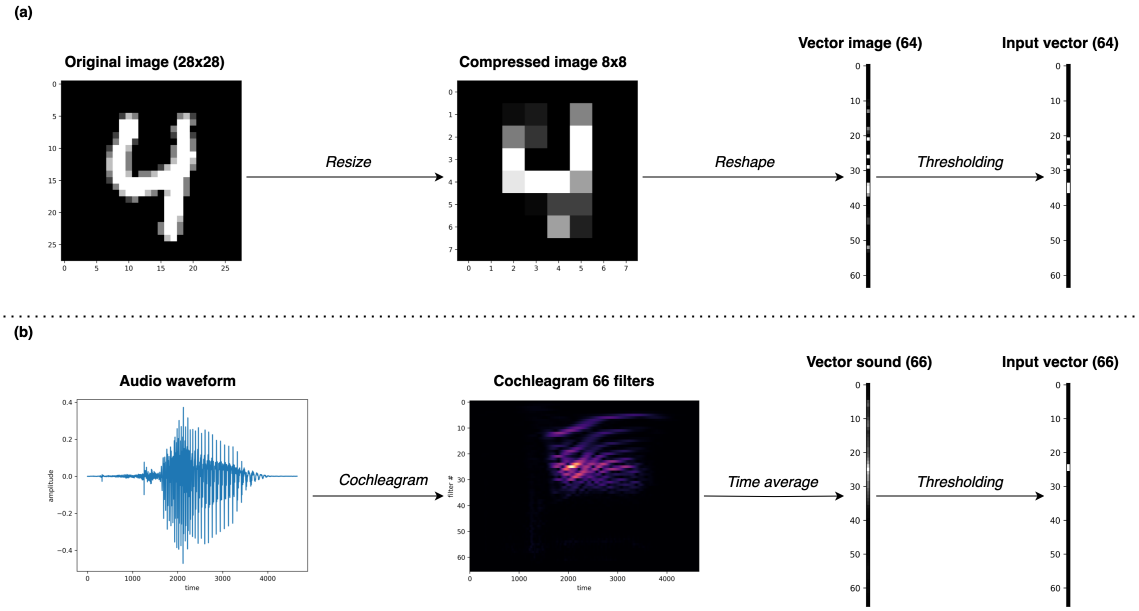


Figure 4.1: Pre-processing of audio-visual data. (a) Pre-processing applied to MNIST images to obtain neural network input vectors. (b) Pre-processing applied to FSDD sounds to obtain neural network input vectors.

For the images, we decide to use only pixels intensity as features. From a machine learning point of view, using such features can be questioned in terms of robustness (instead of using characteristic points or edge detections), but given that the aim of this study is not to obtain the "best" recognition accuracy, but simply to trivially characterize the data and create associations between neurons, these simple informations should be sufficient. The biological aspect of these inputs can also be criticised for shape recognition tasks, but can nevertheless make sense if we consider each pixel as a degree of luminance associated to a specific area of the visual space. In this way, we can imagine associations between these simple low-level features and a higher-level label. Therefore, using the original images, we resize them (i.e. from 28x28 images to 8x8 images for 64 input pixels) to reduce

the dimension of the network inputs for computational reasons, and then we perform a thresholding operation (i.e., pixels below a threshold are set to 0, those above are set to 255) to keep the most relevant information from the resulting image vector (see Fig. 4.1 (a)).

For the sounds, we emulate the cochlea by extracting the frequency bands of the audio signal using the cochleagram. This provides a fairly bio-realistic representation of the auditory sensory inputs by mimicking the capture of the sound by the cochlea itself connected to the primary areas of the auditory cortex as stated in the introduction. Moreover, these features should be quite robust considering the simplicity of the sounds treated (i.e. small words) and more precise than a simple spectrogram since the cochleagram takes into account the non-linear sensitivity of our ears to different frequencies (i.e. high frequency components tend to be represented with less precision than middle and low frequencies). Therefore, we consider cochleagrams of 66 filters as inputs. We then average over time the cochleagram and threshold it to retain relevant information characterizing the sound (see Fig. 4.1 (b)). Although averaging loses the temporal information of the data, it simplifies the process and the association with visual data that are not time-dependent (i.e. images). The complete audio-visual database is represented in Fig. B.7 of Appendix B.2.

It should be noted that during the thresholding operations in both modalities, given that the vectors are normalized between 0 and 255, we use a relatively high threshold of 192 in order to correctly discriminate the different data. Then, the remaining active features are binarised (at 255), stimulating neurons with the same amplitude, favouring associations. Preliminary tests with analog values were performed without significantly improving performances. As a result, during the learning process and therefore the encoding of information, we will rely on a strictly **temporal coding of the information**, as enunciated earlier, where only the timing of the applied input will create the association between neurons. This directly echoes to the "synchrony-based" adaptation considered throughout this thesis, where the spatio-temporal correlations of inputs shape the structures. Nevertheless, rate coding is employed during information decoding processes, where the activity of neurons (i.e. firing rate) is used to make decisions or reconstruct the signal over a given time window.

4.2.2 Experimental protocols

Architecture The experiment consists of a learning phase and a testing phase (where adaptation is still present as before) in a network of 200 neurons. In this network, we have 64 excitatory neurons dedicated to the visual modality (i.e. image inputs) and 66 excitatory neurons dedicated to the auditory modality (i.e. sound inputs). These excitatory neurons allow to receive sensory inputs from the receptors in the same manner as primary sensory cortices and thus represent the particular features to be encoded. Then, we have 30 excitatory neurons representing global labels of the digits (3 neurons for each 10 digits). These neurons aim to represent higher-order brain regions and a higher level of abstraction allowing to integrate and bridge between the two modality areas. This group of neurons admits a certain selectivity and redundancy (3 specific neurons associated to one digit),

allowing synergistic integration of signals from sensory neurons and inhibitions for a more robust decision making capability. Finally, the network contains 40 inhibitory neurons enabling to satisfy the minimal conditions to learn and retain 20 memory items with a pair of Hebbian and anti-Hebbian inhibitory neurons for each data (see Sec. 3.3.1.3). These pairs of inhibitory neurons can also be considered as local labels, each representing a memory concept to be encoded and stored at the local level of the modality. This architecture is summarised in the diagram in Fig. 4.2. The only purpose of this architecture representation is to provide a better overview of where the different signals are applied in the network. Indeed, as throughout this thesis, the network is considered to have an all-to-all connectivity with initially random weights and this structure is more expected to emerge naturally through plasticity.

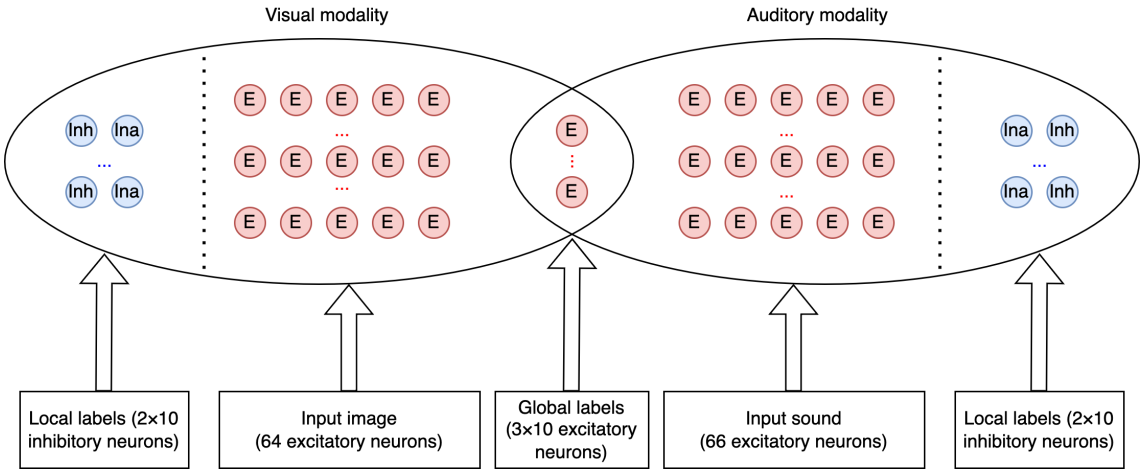


Figure 4.2: Network architecture for learning of audio-visual data. The network is divided into two areas, one for the visual modality and the other for the auditory modality. Each zone contains: excitatory neurons (red circles) receiving sensory inputs (images or sounds), pairs of Hebbian and anti-Hebbian inhibitory neurons (blue circles) representing local memory labels (for the 10 digits). At the intersection of these two areas, 10 groups of 3 excitatory neurons represent the global labels for the 10 digits. The size of the inputs are given in the brackets.

The network parameters are summarized in Table 3.1 in the previous chapter. Those which admit new values, such as the learning time scale (set to $\tau_l = 1$s to allow a slower learning rate and avoid overlearning the data), are represented in Table 4.1.

Learning phase During the learning phase (which lasts 100 seconds), we randomly alternate every second the modality and the digit selected to stimulate the corresponding area with the associated input vector for 0.8 seconds followed by 0.2 seconds of rest. In this way, the fact that each modality acts in distinct area of the network with the stimulation of specific neurons (as represented in Fig. 4.2), should lead to the creation of pseudo-modular structures induced by the spatio-temporal correlations of the inputs in the same manner

Table 4.1: Parameters for the network of QIF neurons during multimodal experiments.

Parameters	Values
N	200
N_e	160
N_{hi}	20
N_{ai}	20
$I(t)$	$\{0, (255\pi\tau_m)^2\}$
τ_l	1.0s

as in Fig. B.5 of Appendix B.1.

In addition to the sensory input, a supervision signal (external inputs of the same amplitude as sensory inputs) is added at the local level by stimulating a pair of inhibitory neurons associated with the label and modality of the data in order to maintain the memory item (see Sec. 3.3.1). In addition to the justifications given in the previous chapter emphasizing the role of higher-level signals from the hippocampus to the inhibitory neurons promoting memory consolidation [356, 124], it has been shown that sensory inputs also target inhibitory inter-neurons [83, 209].

Finally, another supervision signal (of the same amplitude as sensory inputs) is applied at the global level to the corresponding excitatory label neurons associated with the selected digit. This should form some hub neurons as in the previous chapter, allowing connections between the two modalities. More specifically, characteristic features of each digit from both modalities should connect to the corresponding hubs. Thus, during recognition task, we could read the activity of these hub neurons as we did in the preliminary experiment of Fig. A.10 of Appendix A.2.0.2. Although this process makes learning rather supervised, this mechanism is used in almost all current machine learning algorithms and this experiment remains a first step towards less supervised learning in the future.

We note that the two modalities are learned separately, although a scenario in which sensory inputs from both modalities are given at the same time could be of real interest. Some experiments have been conducted in this direction in Sec. B.2.2 of Appendix B.2. Nevertheless, the separation of the modalities during learning should facilitate the structural segregation of each modality and therefore their individual performances.

Testing phase During the testing phase, we evaluate the recognition capacity of the network. Firstly, we experiment with each modality alone by giving the sensory input of each digit and modality to the network without stimulating the local and global labels. In this way, we expect the neurons of the corresponding labels (local and global) to respond accordingly to the given input. Thus, we can quantify the ability of each sub-network to individually recognize the previously learned patterns. Then, we move on to the bimodal scenario in which we present both visual and auditory inputs to the network. Thereby, we can expect bimodality to bring some benefits to the recognition capability of the system by integrating these different informations. Finally, we test a bimodal mismatch scenario

where the two digits given to the two modalities are not in congruence. In this case, the network should not be able to correctly make a decision about which digit it recognizes.

In the second testing experiment, we evaluate the capacity of the network to generate learned information. It corresponds to the reverse process of the recognition task, since we only stimulate the global label neurons with the aim of regenerating the associated sensory inputs (auditory and visual). Thus, the fact that these higher-order label neurons connect both modalities, should make it possible to generate all the sensory outputs at the same time. This experiment highlights the advantages of this architecture and model, which enable both bottom-up (recognition) and top-down (generation) processing.

4.3 Results

The aim of this final chapter is to apply the model and paradigms described in the previous chapters to a more cognitive relevant scenario. Firstly, we describe the learning process of audio-visual data leading to the formation of a complex structure made up of sparse modular structures and hubs connecting them. We also study in this part the maintenance of this structure in the long-term thanks to the inhibition mechanisms described in the last chapter. Then, we probe the efficiency of this learning with the unimodal and bimodal recognition of sensory inputs. Finally, we investigate the reverse process with the generation the two sensory informations from their labels.

4.3.1 Learning leading to complex structure

As explained in the experimental protocol in Sec. 4.2.2, during the learning phase represented in Fig. 4.3, we observe that each digit learned is encoded by the stimulation of certain excitatory neurons in the corresponding modality area (see panel (a)). At the same time, groups of neurons at (1) the local level (inhibitory neurons) and at (2) the global level (excitatory neurons) also increase their firing according to the nature of the digit. These coordinated activations of different zones of the network are even more visible in the graphs in panel (b) where the spiking patterns of the neurons are translated into mean firing rates in the different regions. This highlights the associations between the supervisory signals applied in the local areas to retain memory items and in the global area for integration, and the inputs applied in the sensory areas.

This learning process allows us to move from the random and weakly connected weights matrix in (c) to the structurally connected matrix in (d). Thus, we can observe the formation of sparse connections between excitatory neurons belonging to the same sensory modality (visual or auditory), emphasizing the phenomenon of segregation of the modalities into distinct modular structures (green and magenta areas in the weighted matrices). These two excitatory structural modules are themselves composed of small memory patterns associated with groups of inhibitory neurons maintaining these learned items over time, as explained in Sec. 3.3.1. In other words, in each modality area, a pair of Hebbian and anti-Hebbian inhibitory neurons stores and represents a digit at the local level and the

connections to/from them encode the information (i.e., E-I and I-E connections). These connections therefore establish a link between sensory features and labels.

However, some feature neurons may encode for different digit memories (e.g., neuron 50), again highlighting the notion of mixed selectivity. This makes the patterns of E-I and I-E connections even more complex with overlapping memories and multiple feedback inhibitions in the same manner as in Fig. 3.6. On contrast, some excitatory neurons may encode no information at all, characterized by the absence of reinforcement of their weights and their slow disappearance. Indeed, most of these neurons are associated with unused frequency bands in sounds or with borders of images, and are therefore neurons that never receive sensory stimulation.

A similar process is performed at the global level, with groups of excitatory neurons each representing a digit label and receiving (projecting) connections from (to) both modalities. Thus, a label is associated (and therefore connected) to some degree with certain features of the two modalities. These global label neurons are also connected with inhibitory neurons from both modalities (e.g., neuron 150), which ensures their long-term structural maintenance. As a result, these neurons are densely connected to all areas of the network, forming hubs. In this way, we foresee for the rest of the results that these high-level hubs will form a module for multisensory integration [421].

Finally, the results of the previous chapter seem to be confirmed with the long-term maintenance of the learned complex structure after 4000 seconds of spontaneous activity at rest (see weights matrix (e)). Indeed, by satisfying the constraints described in Sec. 3.3.1 with the double inhibition associate with each memory item, modules, hubs or mixed selective neurons are conserved. This consolidation is highlighted by the distribution of weights in panel (f), where strong (weak) links $> |0.5|$ ($< |0.5|$) appear to be more (less) numerous between the end of the learning (in green at time T = 100s) and the end of the simulation (in red at time T= 4000s). Nevertheless, the number of null connections continue to increase, showing the decoupling of modules and of non-coding neurons.

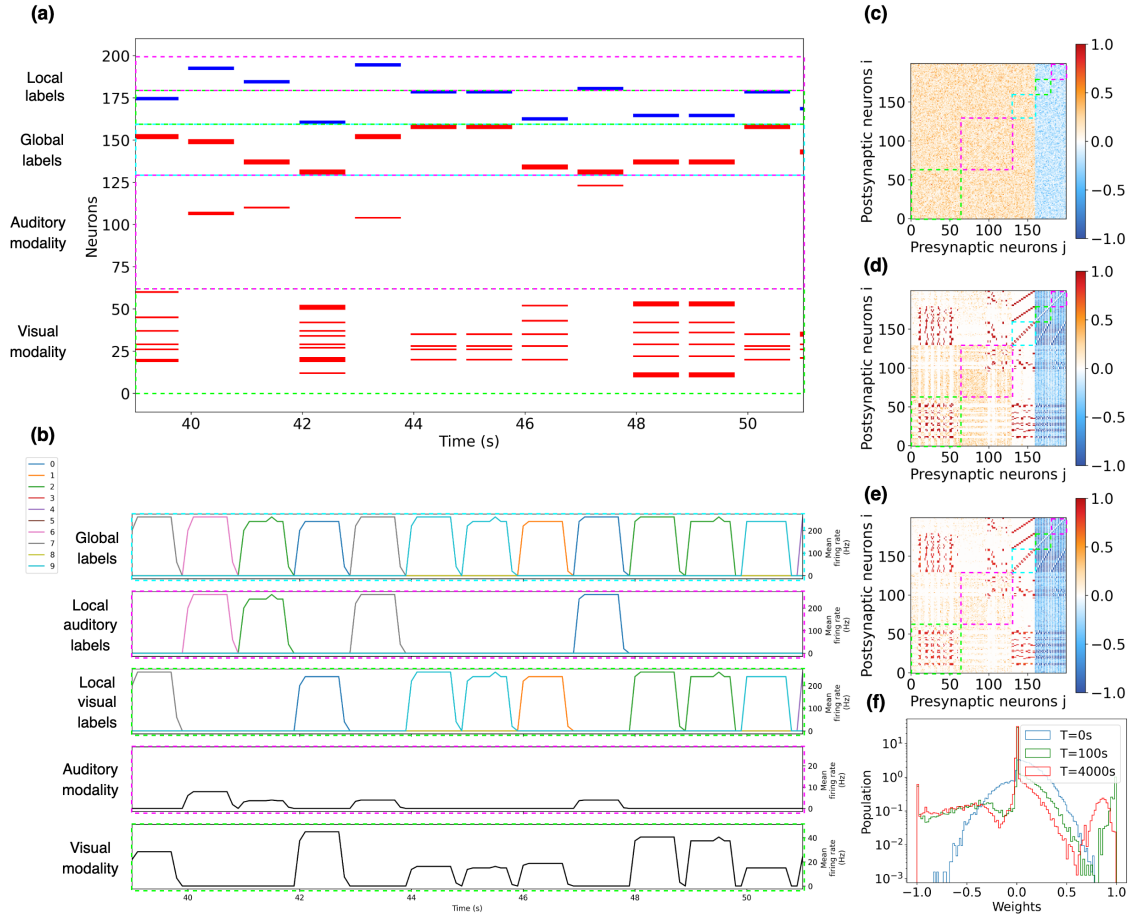


Figure 4.3: Simulation of a neural network of 200 QIF neurons learning bimodal digits. Neurons 0-63 code visual information, neurons 64-129 code auditory information, neurons 130-159 code global labels (3 consecutive neurons per label), neurons 160-179 code local labels for visual modality (2 consecutive neurons per label) and neurons 180-199 code local labels for auditory modality (2 consecutive neurons per label). The green dashed rectangles highlight the visual areas, the magenta dashed rectangles highlight the auditory areas and the cyan dashed rectangles highlight the global label area. (a) The spike raster plot during the learning phase with excitatory neurons in red, inhibitory in blue. (b) Mean firing rates of the different areas of the network during the learning phase. The colours in the label areas represent the groups of neurons associated to each digit label. (c) Weights matrix at the beginning of the simulation, before learning. (d) Weights matrix after learning (T=100s). (e) Weights matrix at the end of the simulation (T=4000s). (f) Distributions (in log scale) of the weights in the network at the different instant of the simulation (before learning t=0s (in blue), after learning T=100s (in green), at the end of the simulation T=4000s (in red)).

Without going into details, in Fig. B.8 of Appendix B.2, we reproduced the same learning protocol but this time learning the visual and auditory modalities at the same

time. The results appear to be similar, except that the structural segregation of modalities is compromised.

4.3.2 Unimodal and bimodal recognition task

Now that the consequences of the learning phase on the formation of a complex multimodal structure have been described, we can quantify the effectiveness of this same learned structure in recognition tasks. This recognition tests evaluate the ability of the network to activate the right group of global label neurons in response to the sensory stimuli presented. The Fig. 4.4 recapitulates this recognition phase with, in particular, in panel (a) the activity (in terms of mean firing rate) of the different regions of the network and so the reaction of label neurons to: visual inputs alone (from times 102 to 111), auditory inputs alone (times 112 to 121) and audio-visual inputs (times 122 to 131). In each cases (unimodals and bimodal), the ten digits are presented in order (i.e. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9) for 0.1 seconds and separated by a rest period of 0.9 seconds. It is within this 0.1 seconds window that the mean firing rates of global labels is taken for each input presented and reported in confusion matrices (b), (c) and (d). Thus, these matrices highlight the network's reactions to stimuli and assess classification accuracy. We consider that the global label with the highest activity makes the decision (i.e., to decide which digit is recognized). For equal activities, the decision cannot be made. The activities of local labels (visual and auditory) are given for information but will not be taken into account here.

Thus, we observe that with vision alone (in panel (b)) and sound alone (in panel (c)), the network has an average recognition success of the different digits with rates of 70% and 40% respectively. The recognition errors are quite diverse, with for example, sometimes the detection of completely different digits (high activity on the wrong label), sometimes the detection of no digits at all (no label activity), or sometimes the same activity on different labels and therefore the impossibility of differentiating. Indeed, as previously enunciated, some digits share common features with mixed selective input neurons, making the decision sometimes uncertain. To this must be added the natural noise and heterogeneity of the network, making its reactions variable over time from one stimulation to another. However, these errors are not necessarily the same in both modalities. In fact, they can be complementary: when one modality fails to recognize a digit, the other succeeds and vice versa. Yet, some digits are correctly recognized by both modalities, and other by neither. In this sense, we can say that these hub labels also exhibit a form of mixed selectivity, since they react to different sensory modalities.

This complementary aspect seems somehow confirmed by the bimodality (in panel (d)) where the network recognises 100% of the digits. Thus the two modalities seem to complement each other, although it is interesting to note that in some cases the individual modalities are unable to recognize a digit, but together they can. This bimodality is therefore not just a mater of adding recognition capacities, but a real cooperation between modalities for coherent decision making. Thus, the hub neurons allow the correct integration of these sensory inputs by retaining relevant features and filtering out irrelevant ones. This is due in particular to the competition between the different inhibition processes,

which makes the correct memory item to arise. Consequently, as expected, the use of bimodal information brings a real gain in the recognition capacity of the network compared with unimodal one.

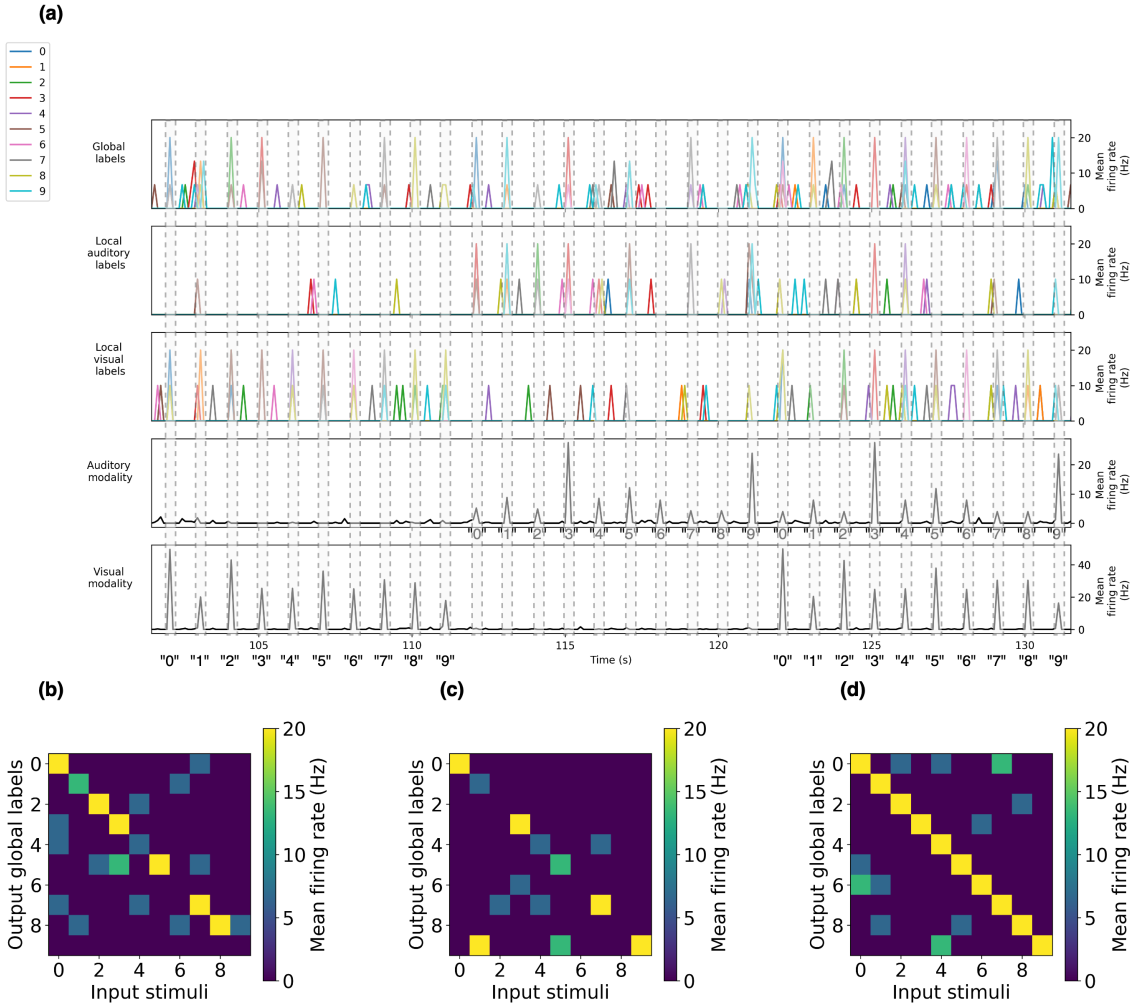


Figure 4.4: Unimodal and bimodal recognition of digits. (a) Mean firing rates of the different areas of the network during the recognition phase. Colours in the label areas represent the groups of neurons associated to each digit label. (b) Confusion matrix for recognition of the visual modality alone. The input stimuli for images 0, 1, 2, 3, 4, 5, 6, 8, 9 are applied at times 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 respectively. (c) Confusion matrix for recognition of the auditory modality alone. The input stimuli for sounds 0, 1, 2, 3, 4, 5, 6, 8, 9 are applied at times 112, 113, 114, 115, 116, 117, 118, 119, 120, 121 respectively. (d) Confusion matrix for recognition of the visual and auditory modality together. The input stimuli for images and sounds 0, 1, 2, 3, 4, 5, 6, 8, 9 are applied at times 122, 123, 124, 125, 126, 127, 128, 129, 130, 131 respectively. In all confusion matrices, the mean firing rates of global label neurons is taken into account at the moment of the stimulation (represented by the dotted grey areas).

From a machine learning point of view, this could make sense as we increased the dimensionality of the inputs by considering two sensory information. However, as said pre-

viously, these information must be complementary and relevant in order to be beneficial and, on the contrary, not to penalise the process. For instance, in the mismatch modality recognition experiments in Fig. B.9 of Appendix B.2, giving inputs not in congruence (i.e. not the same digit in both modalities) makes the recognition very complicated with the stimulation of features promoting the activation of two different global labels which are themselves inhibited by the local inhibition of each modality. These two contradictory information go against the associations created during learning and logically makes the network confused and most of the time unable to make a decision. This particular experiment tend to show some limitations of this kind of architecture, which is subject only to all external cues without any real internal motivation, such as an attention mechanism that filters relevant information according to a specific task, as in [36].

4.3.3 Audio-visual information generation task

In this last experiment, we perform the reverse process by stimulating the global labels in order to obtain the associated sensory inputs; all this after the same learning phase of Sec. 4.3.1. So, as represented in Fig. 4.5 (a), we activate the ten global labels in order (0 to 9), each during 1 second followed by 1 second of rest. This has the effect of increasing the activity (in terms of mean firing rates) of the different regions of the network. More precisely, activation of a specific global label triggers the corresponding inhibitory local labels of both modalities as well as the feature neurons (visual and auditory) associated with the selected digit. These sensory reactivations allow to generate the original input signals by taking the average activity of the feature neurons during this one-second time window, as represented in panels (c) and (e).

Thus, by comparing panel (b) (i.e. the images effectively learned) and panel (c) (i.e. the generated images), we can conclude that the network is able of recreating the images fairly faithfully according to the label activated. The same applies to the auditory modality by comparing panel (d) (i.e. the sounds effectively learned) and panel (e) (i.e. the generated sounds). In this way, from a chosen label, the network can generate the associated information in the visual and auditory domains at the same time.

However, one criticism that can be made at the current stage of the experiment is that the input data learned are highly simplified (thresholded, binarized and averaged) and therefore logically also those generated. This explains why we only partially reconstruct the original data, which sometimes makes their interpretation complicated. This is particularly the case for the auditory modality, where we average the cochleagram over time, so a lot of information is lost (loss of the temporal dimension, which is not present for images).

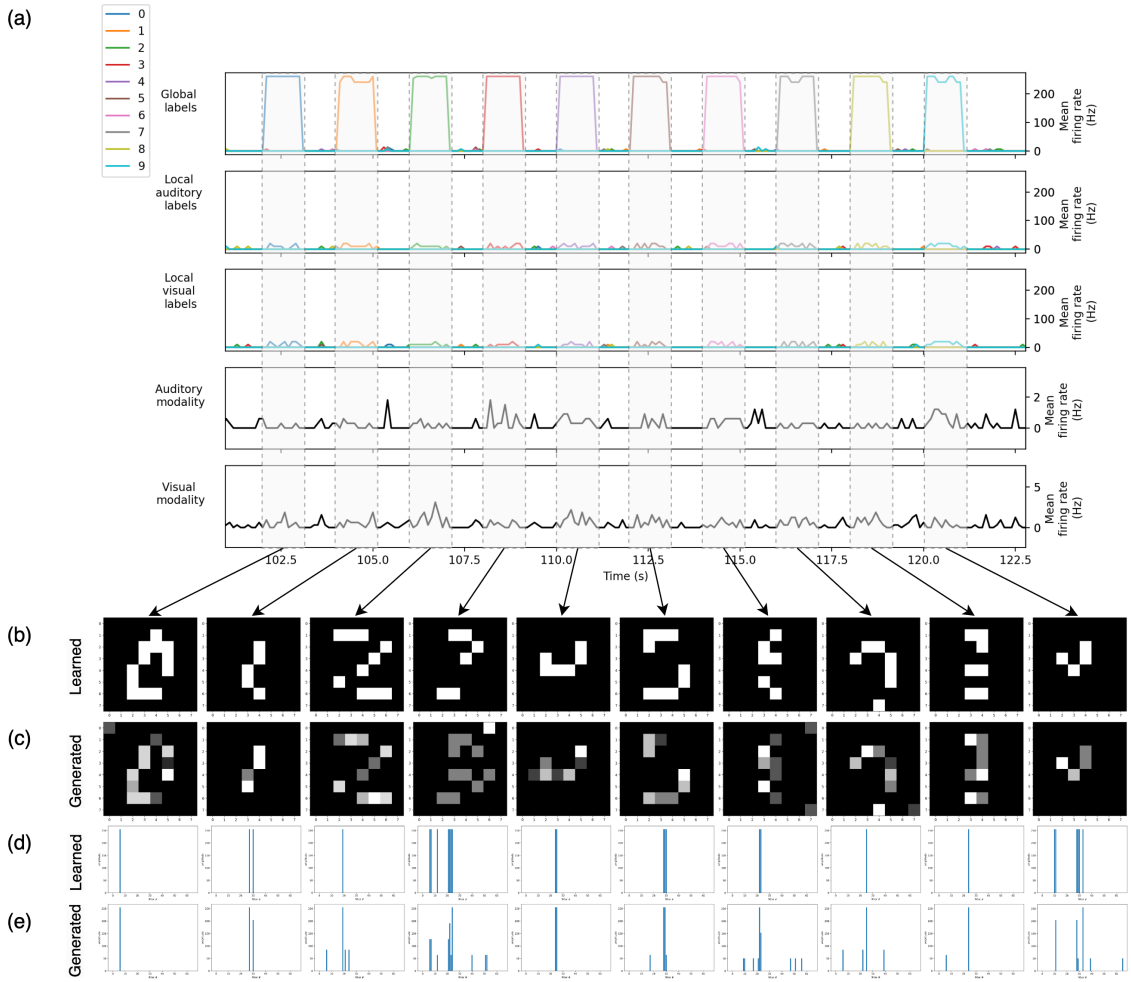


Figure 4.5: Generation of audio-visual digits. (a) Mean firing rates of the different areas of the network during the generation phase. Colours in the label areas represent the groups of neurons associated to each digit label. (b) Input images (0, 1, 2, 3, 4, 5, 6, 8, 9) learned during the learning phase. (c) Output images generated by stimulating global labels 0, 1, 2, 3, 4, 5, 6, 8, 9 at times 102-103, 104-105, 106-107, 108-109, 110-111, 112-113, 114-115, 116-117, 118-119, 120-121 respectively. (d) Input sounds/cochleagrams (0, 1, 2, 3, 4, 5, 6, 8, 9) learned during the learning phase. (e) Output sounds/cochleagrams generated by stimulating global labels 0, 1, 2, 3, 4, 5, 6, 8, 9 at times 102-103, 104-105, 106-107, 108-109, 110-111, 112-113, 114-115, 116-117, 118-119, 120-121 respectively. In all generation outputs, the mean firing rates of each sensory neuron during the one-second time windows (represented by the dotted grey areas) are used to reconstruct the data (images and instantaneous cochleagrams).

Similarly, as with the recognition task, the heterogeneity and the noise present in the network lead to a certain variability in the output generated. In other words, the network will never respond and generate exactly the same data, while keeping the main features

of them. This can be a disadvantage from a strict accuracy point of view, with the risk of losing some information. But it can be advantage if we consider it as a capacity for generalization or even "creativity". Moreover, during spontaneous activity in the resting-state, we seem to observe some epochs of reactivation of learned items (images, sounds or both) like the spontaneous recalls observed in Sec. 3.3.2, which, in the same way, certainly allow the maintenance and consolidation of learned items over the long-term.

In fact, this generation procedure can also be considered as an external recall as in Sec. 3.3.4, where here the global label neurons serve as parts of the memory items initiating the recall of the rest of the information. In other words, this process can be seen as data reconstruction where incomplete information (i.e. the global label) is given in order to obtain the complete data (i.e. the sensory digit). The same analogy can be made in the recognition task, but in the reverse process.

Anyway, the results show that the generation remains very accurate with respect to the information learned, so in the future we can envisage experimenting with more numerous and complex data, and therefore expect more complex generation.

4.4 Summary and conclusions

Learning systems generally behave like black boxes trained on a certain amount of data whose parameters are then frozen, in order to mechanically process the inputs to give fixed answers. In this chapter, we applied classical machine learning tasks of digits recognition and generation to a more biologically realistic model with the spiking neural network of Chapter 3. In addition to confirming the possibility to form and maintain over the long-term complex structures made of modules and hubs, but this time with real sensory stimuli, we demonstrate the efficiency of such an architecture for the cognitive tasks described above emphasising the aspect of segregation and integration of modalities.

Indeed, we have shown that **learning each modality independently allows to form two stable segregated structures** representing the areas of each modality. Within these modules, memory items are encoded and maintained by inhibitory neurons in the E-I and I-E links as in the previous chapter. As a result, each area can recognize in a unimodal way sensory inputs with a acceptable success rate.

We have also demonstrated how to structurally merge these modalities employing a similar approach as in Sec. 3.3.3, notably with mixed selectivity neurons. These particular hub neurons respond to both modalities and can process their information simultaneously to integrate them. Thus, in the recognition task, bimodality admits largely better performances than the modalities alone, confirming the coherent integration of multisensory information. Similarly, in the generation task, the activation of hubs triggers the sensory information related to that item, making an analogy with the concept of associative memory. **This integration, in addition to being performed at the level of the hubs, is balanced by the different inhibitory links** which inhibit non-congruent information in order to keep synergistic ones. This once again highlights the predominant role of inhibition, from the storage of memories to their processing.

The main point subject to debate in this chapter is the simplicity of the inputs learned. However, it should be borne in mind that these experiments remain preliminary results, constituting a proof of concept for more ambitious works. Thus, the previous tests have demonstrated the reliability of the architecture and the learning process, while retaining a degree of flexibility in the dynamics induced either intentionally or spontaneously. In particular, with very few resources (only 200 neurons) and a short learning time (100 seconds), the network can achieve competitive performances. This way, these foundations pave the way for more complex experiments and perspectives.

The first thing to experiment with would be to increase the dimension of the features learned and therefore their complexity. In particular, it would be interesting to explore further the temporal aspect in the data (as with sounds) and thus try to learn sequences. Similarly, the network should be able to learn several data associated with the same digit, in order to show its generalisation capacity and, why not, to compete with other machine learning algorithms. From another perspective, we would like to explore the possibility of obtaining a similar architecture and performance without supervision (i.e. without direct stimulation of the labels). In addition to the basic recognition and generation tasks, more complex experiments could be carried out. For example, we can imagine a more complex generation task by activating different labels at the same time in order to obtain a mixture of learned information and thus create new data. Another idea might be to investigate the capacity of the network to minimise the surprise of sensory inputs, as in predictive coding [32, 214, 355], by sequentially presenting two identical or similar inputs and determining whether the network predicts its future state based on its expectations in relation to its past experiences.

Discussion

Contents

The initial objective of this thesis was to understand the emergence of modular structures in ANN through local adaptation. Then, the second problem addressed was to unravel the mechanisms that sustain these structures over time, shedding light on long-term memory functioning. The final purpose was to validate the proposed architecture and mechanisms using real sensory data and to evaluate their integration in multimodal processing.

Throughout the different chapters of this thesis, various biological constraints have been implemented in our models in order to meet these objectives. Among these constraints, we note:

- the presence of **distinct excitatory and inhibitory populations**, admitting different synaptic plasticities (i.e., asymmetric Hebbian for excitation, symmetric Hebbian and anti-Hebbian for inhibition)
- the fact to have a **persistent spontaneous neuronal activity** in the network, unlike conventional AI and machine learning systems
- the **constant adaptation of synaptic weights**, whereas most learning systems freeze the process at a given moment

Based on these constraint, we found that the common denominator of these objectives was the predominant role of inhibition in their proper functioning. In particular, in addition to the impact of this inhibition on network dynamics, we have shown that:

1. **sustainable memory learning** is supported by inhibition, which promotes the recall of learned items in the network dynamics at rest for their long-term maintenance and consolidation

2. the number of inhibitory neurons present in the network is correlated with the number of structural clusters that can be formed and stabilized over time, and thus with the **memory storage capacity of the network**

3. inhibition enhances **memory processing**, notably by facilitating integration of learned information

This discussion chapter will develop and analyse these findings, while extending the implications and perspectives they provide.

The impact of inhibition on network dynamics

Before going into detail of the main findings of this thesis about the direct consequences of inhibition from a purely functional point of view, it is interesting to understand the mechanisms by which inhibition impacts the network dynamics. The diagrams in Fig. 4.6 give an overview of the architectures found for two stable populations in the case of electrical coupling (phase models) and chemical coupling (spiking neuron models).

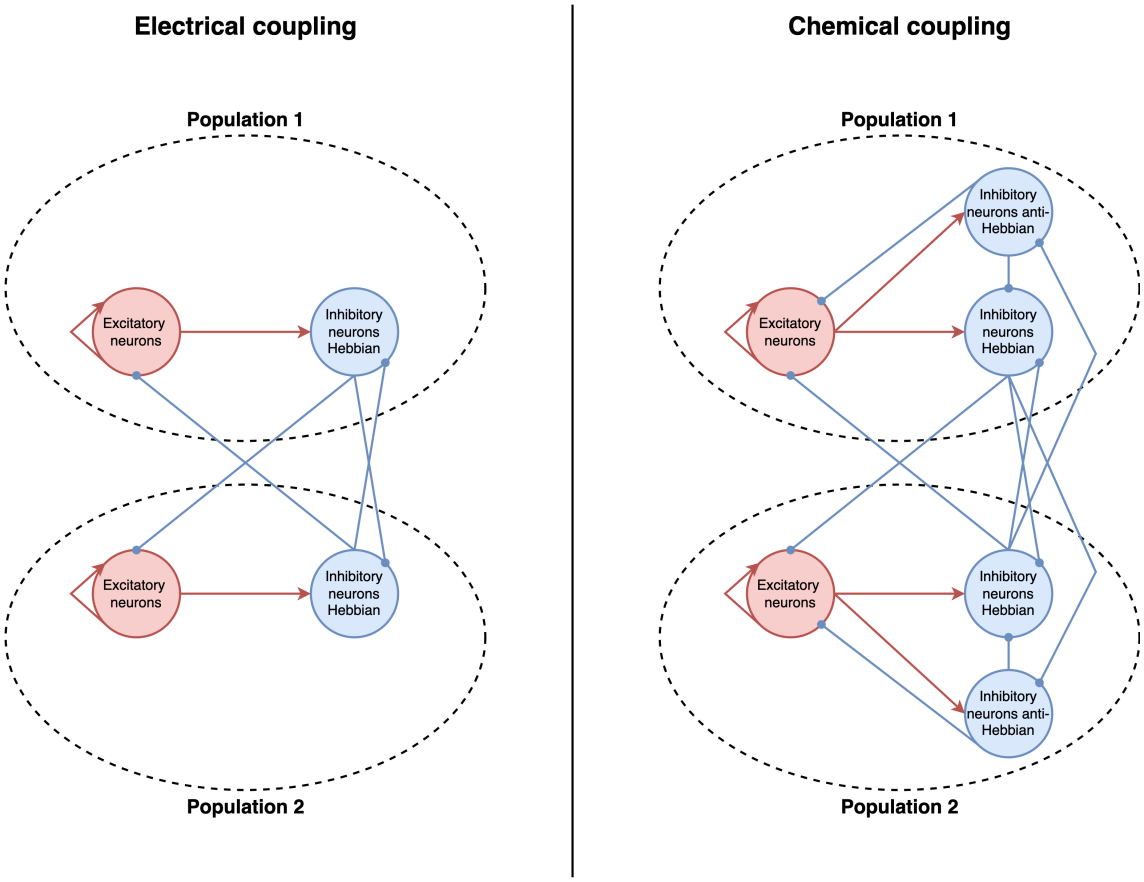


Figure 4.6: Diagrams of the architectures selected for units with electrical coupling (phase models) and chemical coupling (spike neuron models).

Inhibition regulates network activity

In SNN and biological neurons in general, inhibitory neurons are known as **regulator cells**, allowing to reduce the action potential of target neurons. Thus in Chapter 3, we have seen that excitatory neurons not receiving inhibition admit high firing rates, especially when these neurons are strongly connected (see Fig. 3.2 (b) and (c)). Indeed, several studies have shown that an excitatory-inhibitory balance is crucial in order to regulate network activity [20, 285]. In particular, disruption of this balance can bring an abnormal network activity leading to disorders such as seizures [167], altered visual processing [416], autism or schizophrenia [385].

Moreover, during the learning phase and the stimulation of subgroups of neurons by an external input, we have seen in Fig. 3.2 (d) and especially in Fig. 3.2 (f) of Chapter 3 that this regulation brought by inhibition on excitatory neurons is required to distinguish activities of uncorrelated neurons. This idea is supported by studies in the visual cortex demonstrating that **selective activation of inhibitory neurons** decreasing the activity of some excitatory neurons and so reducing their sensitivity to sensory stimuli, improves

feature selectivity and perception [211, 241]. This reinforces our choice in Chapter 4 to model excitatory stimulation as sensory inputs [67, 188] and inhibitory stimulation as consolidation signals from the hippocampus [356, 124].

With regard to phase models in Chapters 1 and 2, inhibition does not necessarily reduce the action potential (i.e. the phase) of target neurons. Indeed, the electrical couplings of these particular neurons, tend to push them in anti-phase with respect to the pre-synaptic neuron, which refers more to a mechanism for modulating synchronisation.

Inhibition modulates the degree of synchronisation

In our spiking neural model developed in Chapter 3, we observed in Fig. 3.2 (b) and (c) that without inhibition, excitatory neurons have difficulties to exhibit synchronized dynamics. On the contrary, in Fig. 3.2 (d) and (e) when anti-Hebbian inhibitory neurons are present in the network, the structural modules can present synchronized epochs. Indeed, this particular plasticity allows the formation of a feedback loop between the excitatory and inhibitory populations of a cluster, where the discharge of the first population activates the feedback inhibition of the second (see Fig. 3.3). Thus, this feedback inhibition, in addition to regulate the activity of the cluster as we previously discussed, favours synchronisation within the cluster. This feedback inhibition phenomenon is observed in the brain, where activation of certain inhibitory interneurons can lead to synchronized oscillations of a particular group of excitatory and inhibitory neurons [285, 383, 342, 410]. Specifically, in a balanced network, in addition to regulating the flow of information [383], this feedback reduces the variability of the excitatory neurons responses [285] (variability characteristic of the A-I state), and it prevents runaway excitation [342, 167] (problem observed in Fig. 3.2 (b) and (c)). Also, this feedback promoting synchronization between excitatory and inhibitory neurons resembles the PING (Pyramidal-Interneuron Network Gamma) rhythm [52, 69, 410]. However, in our case, the synchronization epochs appear at rather low frequencies relative to the gamma frequency range (30-80 Hz) for PING rhythm in the brain [69, 75].

Nevertheless, we observed in Fig. 3.2 (c) and (e) that though Hebbian-STDP inhibition, we can have the opposite effect and promote desynchronization between neurons of different clusters. Indeed, this plasticity leads to the formation of a feed-forward inhibition targetting all neurons not belonging to its population as shown in Fig. 3.3. In this sense, these inhibitors can be considered as an output layer sending signals only outside the cluster. As a result, when excitatory neurons in one population spike, they trigger the Hebbian inhibitory neurons in that same population and they inhibit all other populations. We have seen previously that this effect is decisive to regulate the activity of other populations and to promote a correct learning. At this stage of learning, we can interpret that this feed-forward inhibition prevents the others neurons to synchronize with the stimulated neurons. This is all the more visible at rest in Fig. 3.2 (e) where the clusters spike at distinct instants. Thus, in the brain, it is demonstrated that feed-forward inhibition plays a critical role in desynchronization of excitatory neurons [81, 193, 167, 303]. Above all, this desynchronization allows to encode and recall multiple stimuli promoting separate

dynamics for each [303].

We can conclude that **feedback inhibition (brought by anti-Hebbian-STDP) promotes synchronization** for correlated information from the same cluster and that **feed-forward inhibition (brought by Hebbian-STDP) promotes desynchronization** for uncorrelated information from different clusters. The inhibition found in this kind of model (using chemical couplings) behaves differently from the inhibition used in phase models. Indeed, using electrical couplings, positive couplings associated with excitatory synapses tend to favour synchronicity and so have an attractive effect [171, 243, 361, 407]. On the contrary, negative couplings associated with inhibitory synapses promote desynchronization by repulsing units [171, 243, 407]. In SNN, the synaptic inputs between neurons can be considered more as pulse coupling where excitatory units increase the firing rate of target neurons and inhibitors decrease it. Despite this, as for phase models, we have demonstrated the impact of inhibition to modulate the degree of synchronization in a SNN and so to exhibit partially synchronized dynamics. More precisely, while the impact of E-E and I-I connections appears negligible (see Fig. 3.5 and Fig. 3.2 (g)), E-I and I-E connections emerge as decisive for the formation of spontaneous recalls.

Sustainable memory learning

The problem of maintaining over the long-term the memories a network has learned, while keeping an on-going adaptation process that avoids forgetting past memories by new ones, is a relevant issue encountered in models from various fields, from biology to machine learning and AI. As the main purpose of this thesis, we investigate how the different synchronized and desynchronized regimes brought by inhibitions described in the previous section, allow the creation, maintenance and consolidation of these memories from phase models to more realistic SNN.

The creation of memory

Throughout the different chapters of this thesis, we have seen that **the creation of modular structures is driven by the adaptation mechanism** in which excitatory neurons receiving **the same input at the same time** reinforce their connections while connections between uncorrelated neurons are essentially suppressed. This echoes the principle of Hebbian learning: "*cells that fire together, wire together*" [174], developed with the "Synchrony-based" adaptation in Chapter 1. This approach was preferred to the "Frequency-based" adaptation for simplicity of implementation. Moreover, it appeared that with sensory stimuli, it was the most natural and efficient way to create associations between neurons.

Indeed, in Chapter 4, we demonstrated the possibility of forming these **modular structures with the application of localised audio-visual sensory stimuli**. The structure emerging in the network following the learning phase resembles the modular and hierarchical organisation of the brain where each module represents a sensory modality [403, 179, 278, 421, 422]. In addition to representing high-level cognitive functions

with the segregation of sensory areas, neural assemblies can be associated with memory items encoded by plasticity at a lower level [100]. Since these modular structures are the result of encoding learned stimuli, they can be perceived as memory items [127, 362].

By using θ-neurons in Chapter 2, the learning process allows these connected neurons to maintain a certain degree of synchrony among them thanks to electrical couplings, as explained, while exhibiting some variability in their dynamics due to the external noise and the heterogeneity in their parameters. Differently, by using spiking neurons in Chapters 3 and 4, neurons present in the modules admit a A-I behaviour punctuated by sporadic epochs of transient synchrony. These modules interact via inhibitory connections which play a decisive role in their long-term maintenance.

The maintenance and consolidation of memory

In the case of the maintenance of structural modules in networks composed of θ-neurons (see Chapter 2), we observed that the inhibitory neurons play a fundamental role in preventing a global resynchronization of the whole network after the learning phase. Therefore, **each cluster exhibits an independent dynamics, preventing a long-term forgetting** of the structural neural assemblies, each coding for a different memory item (stimulus). In contrast to what has been reported in studies with chemical synapses where sufficiently strong inhibitory connections and strong E-I coupling favour the synchronization in the network [52, 315], in this model the inhibitory neurons are also synchronized by the excitatory ones, but their activity tends to maintain the excitatory clusters desynchronized. As shown in Sec. 4.4, this difference is explained by noticing that in this model, the couplings among the neurons are essentially electrical gap junctions, which for moderate strength and in absence of delay are known to promote in-phase (anti-phase) dynamics among excitatory (inhibitory) neurons [26] similarly to what is observable for phase oscillators.

On the other hand, we have shown the possibility of consolidating structures despite the learning did not involve all excitatory neurons. In particular, we have studied the case in which partially trained areas were not completely formed due to partial random stimulations and we have analysed the case where a sub-group of neurons remains untrained (unstimulated), see Figs. A.13 and A.14 of Appendix A.3. From these analyses, we can conclude that even if the system partially learns (but at a sufficient degree) the given stimulation patterns, then the connections will be anyway reinforced during the post-learning phase. Here, an analogy could be drawn with memory consolidation during sleep [365, 242], when the memories of the daytime experiences are recalled for their reinforcement. Finally, we have shown that the reported results are general by replicating the experiments with two other oscillator models (i.e. Kuramoto and Stuart-Landau model) as shown in Fig. A.17.

In the case of spiking neurons in Chapters 3 and 4, the maintenance process is somewhat more complicated given the non-deterministic nature of their dynamics at rest. Yet, we have shown in Fig. 3.4 that spontaneous recalls of learned items are directly correlated with a strengthening of the synaptic connectivity of the associated structural cluster. Indeed, each memory seems to consolidate autonomously thanks to the particular

dynamics obtained at rest. More precisely, in the absence of recall, the A-I state supports small forgetting, particularly due to the slow and uncorrelated firing of the neurons. When neurons synchronize, this natural forgetting is largely compensated, preventing the learned structure to be lost in the long-term. These elements confirm our idea that the **spontaneous recall allows the maintenance of long-term memory** in a process of structure-dynamics positive feedback. Furthermore, this idea is confirmed in the Chapter 4 with the learning and retention of sensory stimuli.

Our spiking neural network admits relatively low mean firing rate with δ (0.5-4 Hz) to θ (4–8 Hz) waves rhythms oscillations at rest in the A-I state. In contrast, during full spontaneous recalls, it admits a higher frequency of the range of α (8–12 Hz) to β (12–35 Hz), rhythms that may occur in certain stages of sleep and process of memory [141, 216, 333]. This phenomenon may refer to the different stages of sleep, with low wave activities in the NREM stages and rather high activities in the REM stages (i.e. synchronisation epochs in our case). In particular, an analogy can be made with dreams, which are characterized by a large variety of frequencies from theta and alpha waves, but also by occasional bursts of faster activity from beta to gamma waves [126, 258]. This process can also refer to the up and down state in sleep [134, 379, 406] for memory consolidation. It should be noted that in our model, there is no frequency adaptation or manual modulation of the neuron's excitability to switch from one state to another. These state changes, generally induced by gene expression in the brain [86, 152, 314], are in our case only due to: the reception of external inputs or to the natural fluctuation of the network (i.e. noise) which can lead to a cascade effect characterized by recalls and increase in network frequency.

In addition to these complete recalls, experiments from Fig. 3.6 and especially from Fig. 4.3 with multimodal data, highlight the fact that internal replays can take various forms and so sustain the memorization of more complex information. Indeed, this experiment determined the possibility for the network to learn and retain items admitting mixed selectivity (i.e. neurons responding to multiple inputs). The notion of mixed selectivity appears to be decisive in complex cognitive tasks [58, 173]. Specifically, it enables the brain to represent or process multiple sources of information simultaneously and to integrate them [290, 318]. These neurons could be seen also as hub neurons allowing for connections between the stored clusters and facilitating the ability to potentially transmit and integrate information [340, 390, 422, 421]. In Fig. B.3 of Appendix B.1, we go further considering hubs shared between 4 different clusters. This allows even more complex dynamics, even if the general idea remains the same. In Fig. 4.3, hub labels only connect to particular clusters (or memory items), creating particular places on the information path where the information are integrated in a different way and, in other words, causing different reactions to external stimuli. As well, we observed that even **partial recalls** (i.e. when only a few neurons of a cluster spike), characterised by lower activity, also **promote structure consolidation**. In the end, a multitude of states participate in this maintenance. However, whatever their characteristics and frequency of occurrence, the network seems to remain stable in its asynchronous state and does not converge in one of these (partially) synchronized states. By comparison, previously with θ-neurons, the network was able to maintain memory through strictly alternating epochs of synchronization of each cluster,

which can be linked to a limit cycle. On the contrary, with our spiking neural model, the network fluctuates and remains in a A-I state while stored items naturally pop-up as multiple attractors (balanced fixed points) induced by noise. These dynamic transitions clearly echo the noise-driven state switching found in neural activity [65, 109, 280, 321].

The memory capacity of the network

In addition to be directly related to the formation and maintenance of long-term memories, here we discuss how inhibition correlates with the memory capacity of the network. We find relations between this inhibition and the number of memories stored and the way in which they are organized and processed.

The storage of memory

In our study with θ-neurons in Chapter 2, we have shown that **the number of inhibitory neurons is linked to the number of different neural assemblies that may keep an independent (not synchronised) dynamics and consequently be consolidated** (see Fig. 2.6(a)). As each cluster is considered as a stored memory item, we can easily link the number of inhibitory neurons to the capacity of a network to learn and store information. We can safely affirm that the number of inhibitory neurons is clearly related to the memory capacity of the network [162]. In particular, we have shown that for non-overlapping memories the maximal capacity is proportional to the number of inhibitory neurons, thus we expect that the maximal storage capacity will grow as $\simeq 0.20N$ by assuming that 20% of the neurons are inhibitory as observed in the cortex. This capacity is definitely larger than that of the Hopfield model in which the nature of excitatory and inhibitory neurons is not preserved and whose memory capacity can grow at most as $\simeq 0.14N$ [185]. Furthermore, we have shown in Fig. 2.6(b) that the number of inhibitory neurons controls also the maximal number of neurons which can code for different items at the same time, i.e. exhibiting a simple form of "mixed selectivity" [317]. These neurons could be seen also as hub neurons allowing for connections between the stored clusters and facilitating the ability to potentially transmit and integrate information [340, 421, 422, 391]. Once again we have shown that the amount of inhibition can be related with the cognitive capability of the network.

These relations were confirmed for spiking neurons in Chapter 3. Since memory is linked to recall, by extension it is directly correlated to inhibition. Indeed, in addition to being at the origin of the formation of these memories through the creation of modular structures during the learning, inhibition controls the memory storage capacity of a network. This claim is highlighted in Fig. 3.3 (b) where the number of inhibitory neurons and more precisely **the number of pairs of Hebbian and anti-Hebbian inhibitory neurons is related to the number of different neural assemblies that may admit independent spontaneous recalls (not synchronised with the rest) and consequently be maintained**. Several studies have also revealed the role of inhibition in the storage capacity of the brain and the retention of these memories [145, 255, 336]. In par-

ticular, in our experiment, we quantified this maximal capacity for a network of N neurons as $\frac{N}{3}$ memory items for an "optimal" network composed of $\frac{2N}{3}$ inhibitory neurons. In that case, this capacity is higher than that of the Hopfield model ($\simeq 0.14N$) [186] but is similar for a more realistic network with 20% of inhibitory neurons as observed in the cortex, giving a theoretical limit fixed at $0.1N$ items. This can question on the reasons why using such E/I ratio as in biology rather than a ratio favouring a better storage capacity. We can argue that this storage capacity is more than sufficient for such a system and that it allows encoding more complex information by dedicating more excitatory neurons in each memory.

For comparison, considering an excitatory-inhibitory ratio of $80 - 20$, we have a capacity of $M = 0.2N + 1$ **memory items for** N θ**-neurons and** $M = 0.1N$ **memory items for** N **QIF neurons**. This difference can be explained by the difference of coupling (electrical versus chemical) between the two models, making more complicated to obtain synchronous regimes in the latter case. Basically, for phase models, only one type of inhibition is required, versus the need of both Hebbian and anti-Hebbian for spiking neurons. Nevertheless, as we have shown previously, in the spiking neural model, we obtain a better capacity for learning and retaining more complex information (e.g., mixed selectivity) without any direct dependency between inhibition and neurons coding different items (unlike θ-neurons). This was confirmed in Chapter 4 with sensory information, where although memory items admit several mixed-selectivity neurons, they are correctly recalled and retained thanks to the pair of inhibitory neurons associated to each of them. Ultimately, this recall process (random or deterministic with oscillators) of memories can be perceived as an active storage of information, where once the information has been sufficiently learned during stimulation, the network autonomously consolidates stored memories.

The organisation of memory

In the first part of this thesis using phase models notably in Chapter 2, we have carried out several additional experiments to validate the robustness of the results and to better understand the mechanisms at the basis of memory storage and consolidation. On the one hand, we have shown that stored memories can be retrieved with a brief stimulation recall even if the excitatory assemblies associated to the applied stimuli are no longer present in the connectivity, but are preserved only in the connections involving the inhibitory neurons, as shown in Fig. 2.8(b). In the light of these results, one could speculate that the main purpose of the connections between excitatory neurons is to store short-term memories while **the links associated with the inhibitory neurons** – which have been reinforced during the consolidation post-training phase – **correspond to long-term memory storage**. The functional connectivity initially induced by the excitatory plasticity through the input information is then maintained in the long-term by the inhibitory plasticity [78, 87, 273, 424]. Therefore, as long as inhibitory neurons preserve their connections, short-term memories could be erased to process and learn new information. This result resembles the one found in [402] for excitatory-inhibitory networks with inhibitory plasticity.

These ideas about inhibitory links and memory encoding are reaffirmed in experiments

of Fig. 3.5 in Chapter 3, where informations appear to be stored in the links associated with inhibitory units (E-I and I-E links). Indeed, in addition to being sufficient to provoke spontaneous recalls for the reasons discussed in Sec. 4.4, these links are sufficient to perform recognition tasks with the activation of the inhibitory units after the presentation of a learned pattern in same manner as with θ-neurons. However, for more complex tasks such as external recall (or reconstruction of memory) as in Fig. 3.7, it appears that memories must be encoded or reloaded into the E-E links. We could again speculate that the main purpose of these E-E connections is to store short-term memories that will be directly used for working memory tasks or encoded in **links associated with inhibitory neurons in order to store long-term memories**. Thus, as long as inhibitory neurons preserve their connections, short-term memories could be erased to process and learn new information, but they could be also reloaded as shown in Fig. 3.5.

An analogy can be made with the functional and structural connectivity of the brain. Indeed, reorganization of the structural connectivity between neurons occurs in response to plasticity, new sensory information and during memory consolidation [63, 175, 252]. Logically, this structural connectivity plays an important role in shaping the functional organization of the brain [183, 202, 295]. However, studies found that regions with strong structural connections do not necessarily have strong functional connections [102] and reciprocally regions with weak structural connections may admit strong functional connections [202]. This suggest that although structural connectivity provides a foundation for functional connectivity, the relationship between the two is not always straightforward and is region and task specific [183, 390]. In our case, we have seen that the structural modules of E-E connections can disappear and reappear without significantly affecting the dynamics at rest. In fact, as soon as some particular structural connections (E-I and I-E links) are sufficiently learned, they form the basis of the cognitive functions and can be used to reconstruct the initial information as in Fig. 3.5 (b). For example, in the multimodal application of Chapter 4, features of memory items are also encoded in E-I and I-E connections of each sensory areas (see Fig. 4.3 (d)). However, in this case, these information are also encoded in E-E connections with the hub label neurons, allowing higher-level processing across modalities.

The processing of memory

As seen, inhibitory neurons, and their connections, allow the coherent storage and encoding of learned memory for processing in different tasks. During the external recall experiments of Sec. 3.3.4 and particularly in the recognition and generation tasks of Chapter 4, we demonstrated the **possibility to effectively triggering one part of memory by stimulating an associated part**. In the case of recognition, this process corresponded to the presentation of a sensory pattern (auditory or visual) to activate the corresponding hub label neurons. In the case of generation, the reverse process consisted of stimulating the hub label neurons to obtain the associated bimodal information. This highlights the concept of associative memory, where there are functional correlations between information from the same structures [362]. A phenomenon linked to this idea is that of

latching dynamics, which can be expressed as the transitions between associated memories [326, 350, 420]. More precisely, the dynamics in the network can be seen as jumps from one attractor to another according to their physical connection and so their semantic relations [11, 326, 350, 420]. In our model, it corresponds to the external activation of a part of the network followed by the induced recall of the memory item. This mechanism can remind that present in the Hopfield's recurrent neural network, where the initial activation of particular units leads to the activation of another state and so on [185]. In general, in these models, the system converges towards a final attractor or to an infinite recursion whose the duration depends on the strength of the attractors and the size of the network [326, 381]. In our case, regardless of whether the activation of a memory attractor is obtained spontaneously or by external stimulation of associate information, the network does not remain in this state, thanks in particular to anti-Hebbian inhibition, and returns to a A-I state as already discussed in Sec. 4.4.

Thus, in line with discussions of Sec. 4.4, feedback inhibition provided by anti-Hebbian inhibitory neurons allows for temporally coherent reactivation of memory items during processing while avoiding abnormal and excessive activity blocked in a specific location in the brain (see Fig. 3.2 (c)) as in epileptic seizures [90, 142, 358]. In addition, during the integration of multimodal information at the level of hubs in the Chapter 4, **feed-forward inhibition**, provided by Hebbian inhibitory neurons, allows **better decision-making capacity**. In a first place, each sensory area processes inputs in a unimodal way, highlighting the aspect of segregation [378], where local Hebbian inhibitory neurons enable a better features selectivity by inhibiting non-congruent ones before transmitting them to the global hub neurons. Then, during the integration of features from both modalities, the hubs compete with each other, mutually inhibiting each other thanks to Hebbian inhibitors in order to remove uncertainties (notably due to mixed selective information and fluctuation of neurons) and thus make a decision. The effectiveness of our architecture in bimodal recognition and generation tasks confirms the role of inhibition in coherent multimodal integration.

Perspectives

This thesis covered a wide range of fields, from AI and machine learning to neuroscience and biology, passing by dynamical systems and physics. Consequently, in addition to the perspectives of improvements or additional experiments addressed individually in the themes of each chapter, more general questions arise in the different fields treated.

The impact of physical topology

The physical topology of the network is an aspect that has not really been addressed in this study. Indeed, throughout this thesis, we have considered an *all-to-all* connectivity where all neurons can potentially target all the other neurons and where only synaptic plasticity shapes the structures. We can justify this choice by: (1) the desire to show the role of adaptation in the formation of modular structures emphasizing the emergence of the different

sensory areas in immature brain and, (2) the fact of considering relatively small networks (a few hundred of neurons) in order to perform phenomenological studies attempting to explain general concepts of memory maintenance. However, **sensory systems such as vision and somatosensory organize information spatially** through neurons with similar response properties, forming topographic maps that represent neighbouring regions in sensory space, preserving the spatial relationships of sensory stimuli [204, 205, 277]. In addition, the brain admits sparse connectivity [123, 352] with in particular long-range connections mainly provided by excitatory pyramidal neurons [299, 426] although some GABAergic cells also project to different brain areas [74]. In our architecture (see Fig. 3.3 (a)), Hebbian inhibitory neurons play this role to some extent. To explain this phenomenon, we could imagine that the Hebbian inhibitory neurons associated with a cluster are physically deported to distant locations in the network to locally inhibit other clusters, so that long-range connections are only ensured by excitators. This would highlight again the difference between physical and functional connectivity, which are not necessary directly correlated.

One idea for taking this physical topology into account would be to represent the network as a **2D lattice** where the physical distance would impact the probability of a neuron targeting another. More precisely, the nature of the pre-synaptic neuron would determine the possibility of admitting connections of varying lengths. Thus, such rules would make it possible to admit a relatively sparse adjacency matrix with local inhibition and short- and long-range excitation. These local inhibitions would also raise the question of the distribution of inhibitory neurons in the network (randomly polluted, localized pools, etc.). This topology would also provide an opportunity to consider a parameter that has been omitted until now, namely the impact of synaptic propagation as a function of the distance between neurons. Finally, based on this particular topological map, the idea would be to reproduce our original experiment of Chapter 4 by applying the sensory inputs this time to real localized areas in order to obtain physically segregated modalities.

Perspectives in artificial intelligence and machine learning

In addition to the perspectives enunciated in Sec. 4.4 of Chapter 4, our spiking neural model could be used in other applications in the field of pure machine learning or bio-inspired AI in general. Indeed, as well as providing certain biological constraints, our architecture solves issues often encountered in classical ANNs, in particular the capacity to keep a **constant adaptation of its weights without catastrophic forgetting** of the information learned, while allowing the network to admit **spontaneous dynamics**. Thus, this adaptability could be of real interest for real-time applications, especially given the relative simplicity of the model employed (i.e. QIF neurons), the small number of resources required to perform sensory tasks (hundreds neurons in our case), the rapid convergence of weights (tens of seconds) and the reactivity and precision of spike events (of the order of milliseconds). We can imagine the example of a robot in a constantly changing environment where the system needs to be able to adapt quickly while retaining the memory of past experiences.

Also, the multimodal application of Chapter 4 has shown the possibility to effectively integrate two different modalities (auditory and visual). Therefore, we can extend to a **larger number of modalities**, enabling more complex integrations. For example, we could integrate several information from the visual modality, such as movement and colour detections. Similarly, in the auditory modality, the confrontation of the two captation sources (i.e. the two hears) could be of interest in the case of spatial localisation by hearing. These integration of sensory information can also be connected to motor control system by creating association between sensory features such as a sound to a movement such as that of the vocal cords for potential robotic applications. Associations could even be done between different motor controls, as in the gesture-speech synchrony [244, 306]. To all this could also be added **attention mechanisms** [36, 324] or other control signals making it possible to simulate more complex cognitive tasks. From a more machine learning perspective, our neural model could be used in a reservoir computing architecture [180, 392] where spiking neurons in the reservoir disentangle information and encode memories, and then a linear regression is performed to read out the activity of the neurons. Such an experiment would highlight the most beneficial integration approach (hubs or linear regression).

Applications in neuroscience and biology

Despite the multidisciplinary aspect of this thesis, the main focus has remained on computational neuroscience, with particular emphasis on the long-term memory mechanisms. This link with neuroscience could be strengthened by replicating our main experiment on the formation of modular structures via adaptation to selective stimuli by replacing digit input data with data on the activations of brain areas over time, such as functional magnetic resonance imaging (fMRI) data. Still in this idea of better matching the **experimental data**, we could implement the architecture of a complete cortical column with the different layers of the visual or auditory cortex. Based on our established model, the spiking neurons subjected to constant STDP would be distributed between the different layers, with in particular input neurons in layer 4 and inhibitory neurons in layer 5 [48]. Thus, we could reproduce our experiment on the formation and maintenance of memories on a larger scale. To go even further, this experiment can also be considered at the whole brain level using a real connectome as in [99, 164]. In this case, modelling individual cells may have its limitations given the size of the network, and an approach such as the **mean-field** theory may be of real interest [72, 77, 121]. Note that in our case, we would need to estimate three different transfer functions to approximate the three types of neuron population considered in our model [72, 77, 121].

From a different perspective, we anticipate that our model could simulate certain biological phenomena and in particular **neural disorders** by tuning or removing certain key parameters of our model. For instance, as we have already discussed in Sec. 4.4, the absence of feedback inhibition in the network gives rise to abnormal activity that can be compared to epilepsy [90, 142, 358]. Conversely, the lack of feed-forward inhibition has been at the origin of problems with decision-making and integration of information that can refer to autism spectrum disorder and attention deficit hyperactivity disorder [115, 397]. In this

way, these experiments will provide a better understanding of the key role of these mechanisms in these disorders. Finally, we foresee that some of the computational experiments performed in this thesis could be conducted *in vitro*, for example, by considering neuronal cultures with or without inhibitory neurons and by applying different patterns of localised electrical or opto-genetical stimulations to sub-groups of neurons.

Appendices

Appendix A

Supplementary results with phase models

In this appendix chapter, we describe supplementary results obtained with phase models such as Kuramoto oscillators and Theta neurons.

A.1 Impact of topologies in Kuramoto networks

This appendix is devoted to reproducing the Sec. 1.2.2 experiments with different topologies instead of an "all-to-all" connectivity. The aim is therefore to study the impact of the network topology (i.e., presence or absence of connections between oscillators) on coupled Kuramoto oscillators subjected to adaptation. To do this, we set up several types of adjacency matrix (parameter a_{ij} in Eq. (1.1)) in order to observe the possible effects on the organization of the weights as well as on the phase dynamics. Indeed, until now, we have been dealing with a "fully connected" topology where all oscillators are bidirectionally connected to each other. In Fig. A.1, we display the adjacency matrices of some of topologies tested.

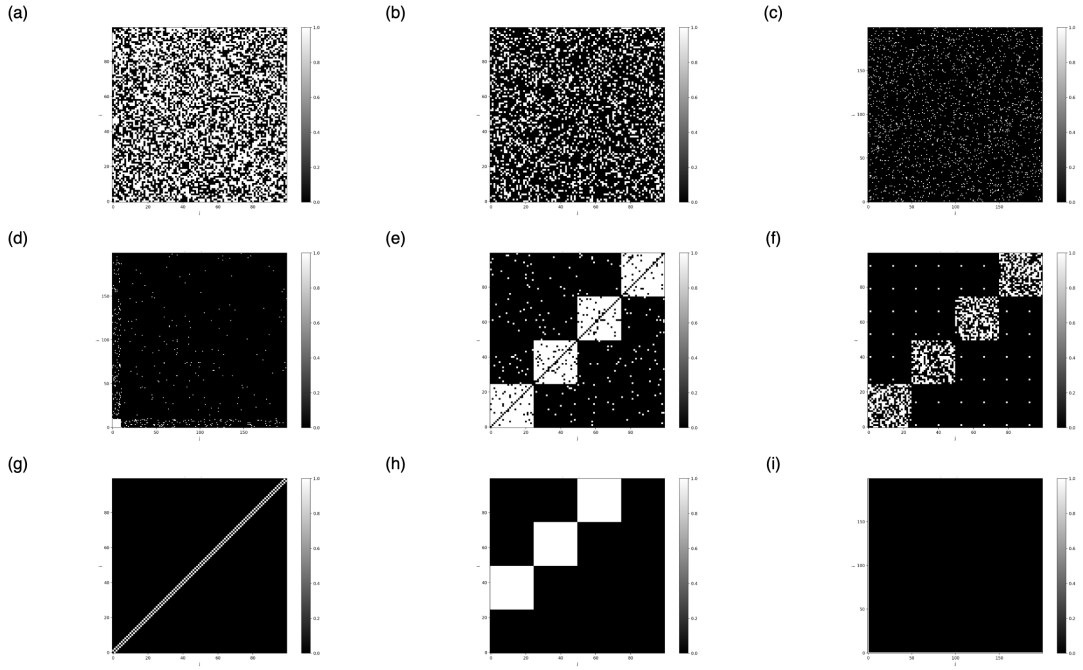


Figure A.1: Adjacency matrices of some network topologies tested. (a) Randomly connected in uniform way. (b) Randomly connected in regular way (i.e., same degree of connection for each neuron, here $\frac{N}{4}$). (c) Randomly connected with a probability of $\frac{1}{20}$. (d) Scale free topology (Barabási-Albert model [29]). (e) Small world topology [281]. (f) Modular topology with hubs [421]. (g) Line topology (with bidirectional links). (h) Multilayer topology (with 4 layers and unidirectional links). (i) "All to one" topology (with bidirectional links).

The first observation is that whatever the type of topology, the states associated with the different learning rules described above are always found after convergence. The second observation is that the sparser the topology (as in Fig. A.1 (c), (d), (g) and (i)), the longer the system takes to converge towards a particular state. This can be explained by the fact that each neuron has fewer contributions from its neighbours to synchronize with each other and that information can take longer time to arrive in some configurations. Finally, in some particular organizations (as in Fig. A.1 (g) and (h)), we observe that the absence of bidirectional links (or feedback) leads to a relatively lower capacity for synchronization of the network. This can be understood from Eq. (1.1) since with unidirectional links, only the oscillators of the incoming links will try to synchronize with their neighbours while those in the outgoing links will only follow their own dynamics.

Concerning synaptic connectivity, Fig. A.2 shows that apart from the links imposed by the topologies, we find the same organization of weights in each case as in Sec. 1.2.2. In fact, we find two alternating zones of positive and negative links for the Hebbian symmetric rule, alternative diagonal stripes of positive and negative links for the Hebbian asymmetric rule and a random organisation for the Anti-Hebbian symmetric rule. Note that we do not

display all weights matrices since they are all extremely similar.

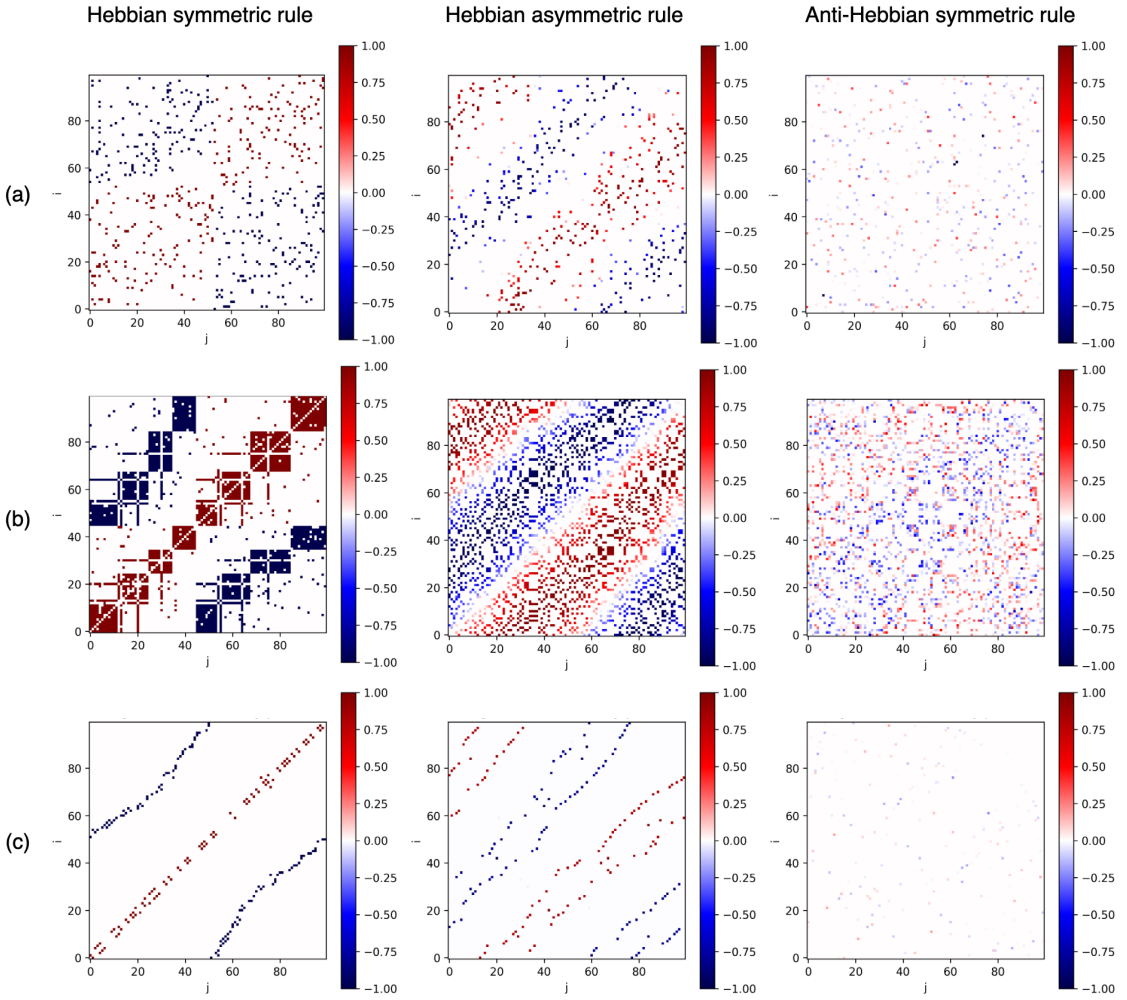


Figure A.2: Weight matrices (sorted by phases) obtained with (a) a randomly connected topology with a probability of $\frac{1}{20}$ (see Fig. A.1 (c)), (b) a small world topology (see Fig. A.1 (e)) and (c) a line topology with bidirectional links (see Fig. A.1 (g)). Each topology is evaluated with the Hebbian symmetric rule, the Hebbian asymmetric rule and the Anti-Hebbian symmetric rule.

We can conclude from this part that, although the original topology of the network has no real impact on the final results, we have seen that sparse and unidirectional connectivity is not conducive to effective network development. For these reasons, we will continue to consider a "fully connected" topology in the following.

A.2 Alternative input stimulation with Kuramoto oscillators

In this appendix, additional results are presented on the impact of different types of inputs applied to coupled Kuramoto oscillators.

A.2.0.1 Frequency-based adaptation

This subsection focuses on alternative input vectors to be learned with different plasticity rules as part of the frequency-based adaptation with coupled Kuramoto oscillators.

Two values input In Fig. A.3, we reproduce exactly the same experiment as in Fig. 1.6 of the main text with an input made of two values, but this time considering the Hebbian asymmetric plasticity rule. No major differences are observed here, apart from different dynamics within each cluster (i.e. a splay state, see Sec. 1.2.2).

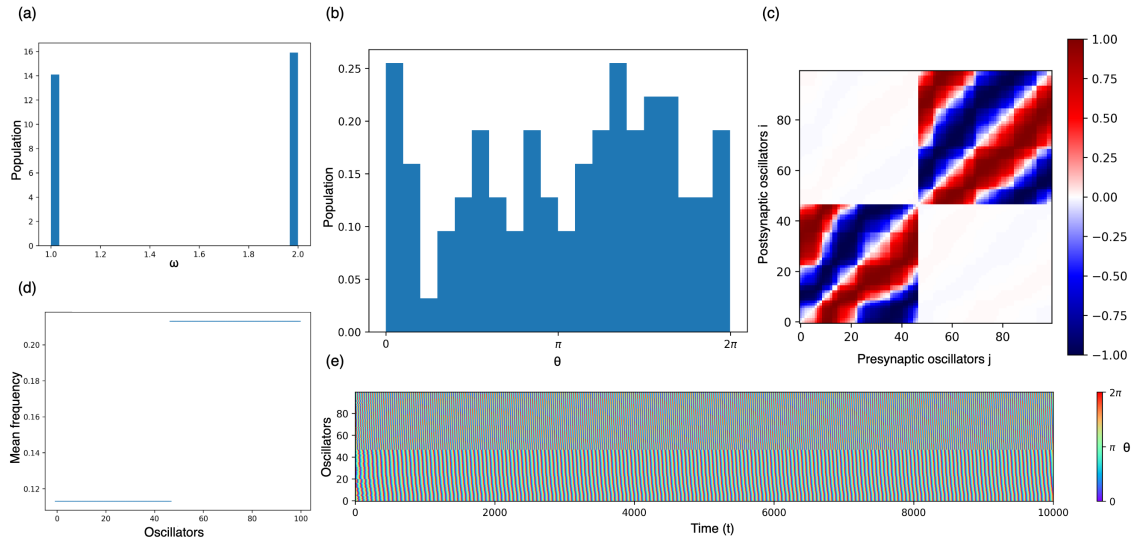


Figure A.3: Learning of a two values input using a Hebbian asymmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the distribution of neurons ($N = 100$) by phases (logarithmic scale), (c) the weighted connectivity matrix (sorted by mean frequencies and phases), (d) the mean frequency of each oscillator and (e) the phase patterns (sorted by mean frequencies and phases). Parameters: $\alpha = 0.3\pi$ and $\beta = 0$.

Gaussian input In Fig. A.4, we reproduce exactly the same experiment as in Fig. 1.7 of the main text with an input made of values distributed with a Gaussian, but this time considering the Hebbian asymmetric plasticity rule. From a structural point of view, no major differences are observed apart from the organization of positive and negative weights within each cluster, which is typical of the structure obtained with the splay state (see Sec. 1.2.2). Concerning the dynamics of these clusters, the "plateau" observed in the mean frequencies in the previous case are less clear here. Indeed, it seems that even oscillators

in a same cluster do not admit exactly the same mean frequency (see panel (b)) compared with the Hebbian symmetric where they had exactly the same frequency. We can explain this firstly by the fact that, thanks to the Gaussian distribution, no oscillators have exactly the same natural frequency. In the case of the symmetric function, the rule favours strict synchronization (or anti-phase desynchronization) of the oscillators, while the asymmetric rule does not, which may explain why the oscillators are able to keep their own frequency regime. Except this small aspect, the main results seem consistent from one plasticity rule to the other.

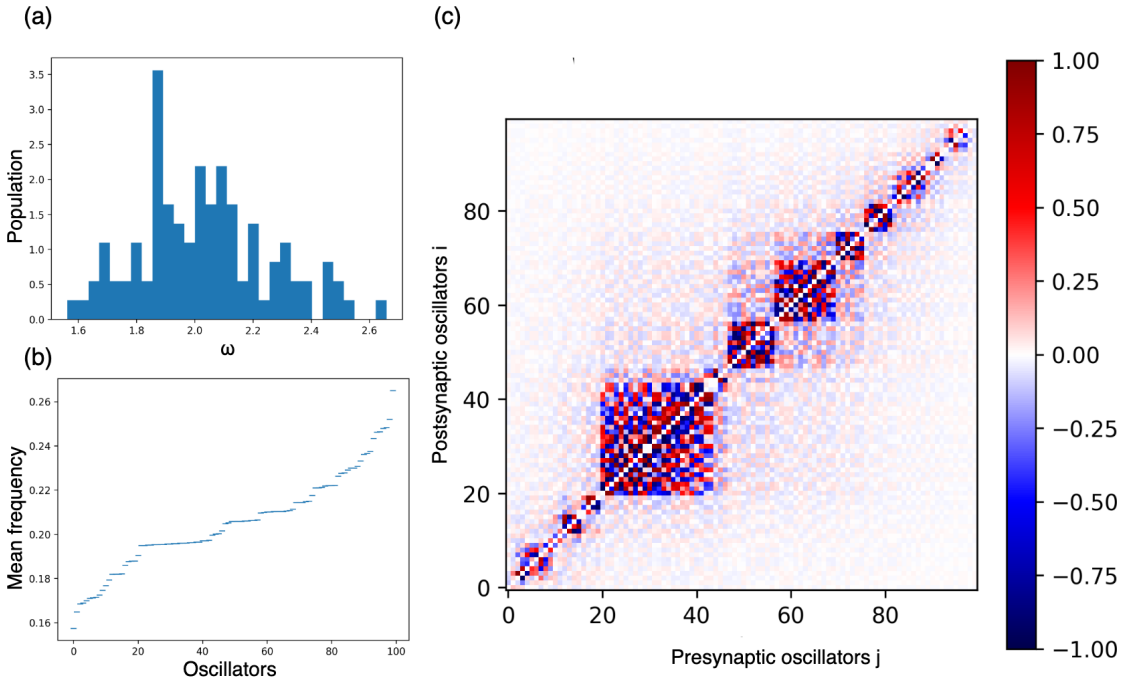


Figure A.4: Learning of a Gaussian input using a Hebbian asymmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the mean frequency of each oscillator and (c) the weighted connectivity matrix (sorted by mean frequencies and phases). Parameters: $\alpha = 0.3\pi$ and $\beta = 0$.

Uniform input As a final input vector, we consider values uniformly distributed between $0,75$ and $1,25$. The results obtained with the two plasticity rules are represented in Fig. A.5 and Fig. A.6. On the side of the Hebbian symmetric rule, we observe the formation of five modular structures in the weight connectivity (c). Although the natural frequencies in panel (a) are uniform, the oscillators associate into five clusters, each with its own mean frequency (see panel Fig. A.5 (b)). This confirms the intuition of the previous experiments showing firstly that oscillators with close natural frequencies can associate to form a cluster and decouple from the others. Secondly, once formed, these clusters maintain a unique mean frequency thanks to the characteristics of the Hebbian symmetric rule. In contrast, with the Hebbian asymmetric rule, clusters are a little harder to distinguish in the

weight connectivity (c). Although small modular structures seem to appear, the boundaries between them are not always clear. In addition, as with the Gaussian distribution, the oscillators all admit a different mean frequency (see panel Fig. A.6 (b)) in accordance with the natural frequency distribution in panel (a).

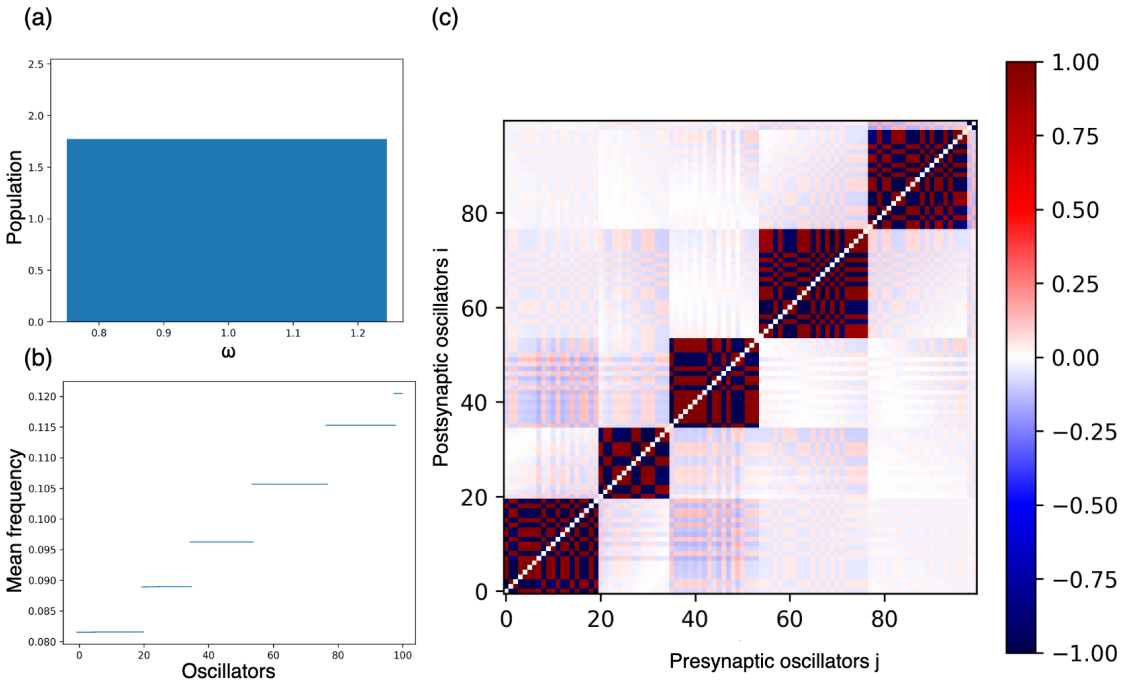


Figure A.5: Learning of a uniform input using a Hebbian symmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the mean frequency of each oscillator and (c) the weighted connectivity matrix (sorted by mean frequencies and phases). Parameters: $\alpha = 0$ and $\beta = -0.5\pi$.

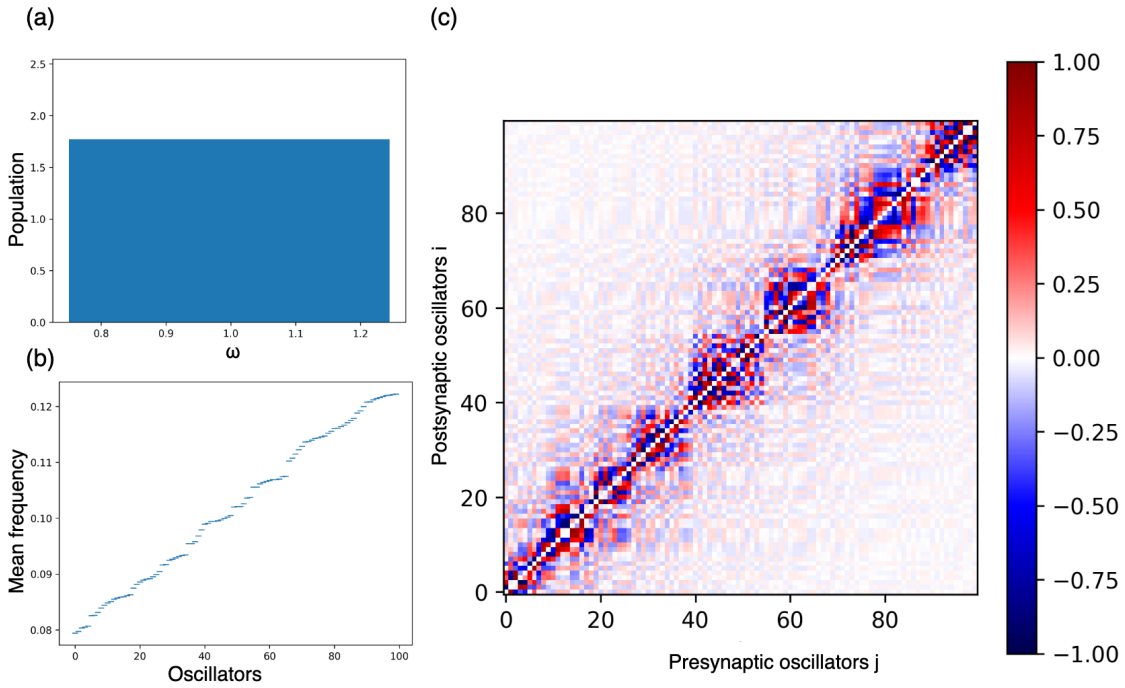


Figure A.6: Learning of a uniform input using a Hebbian asymmetric rule with: (a) the distribution of natural frequencies/inputs, (b) the mean frequency of each oscillator and (c) the weighted connectivity matrix (sorted by mean frequencies and phases). Parameters: $\alpha = 0.3\pi$ and $\beta = 0$.

A.2.0.2 Synchrony-based adaptation

This subsection focuses on alternative input vectors to be learned with different plasticity rules as part of the synchrony-based adaptation with coupled Kuramoto oscillators.

Four inputs In Fig. A.7, we reproduce exactly the same experiment as in Fig. 1.9 of the main text with four different input vectors made of binary values, but this time considering the Hebbian asymmetric plasticity rule. No major differences are observed here, apart from the different dynamics within each cluster (i.e. a splay state, see Sec. 1.2.2) and the associated structural organization.

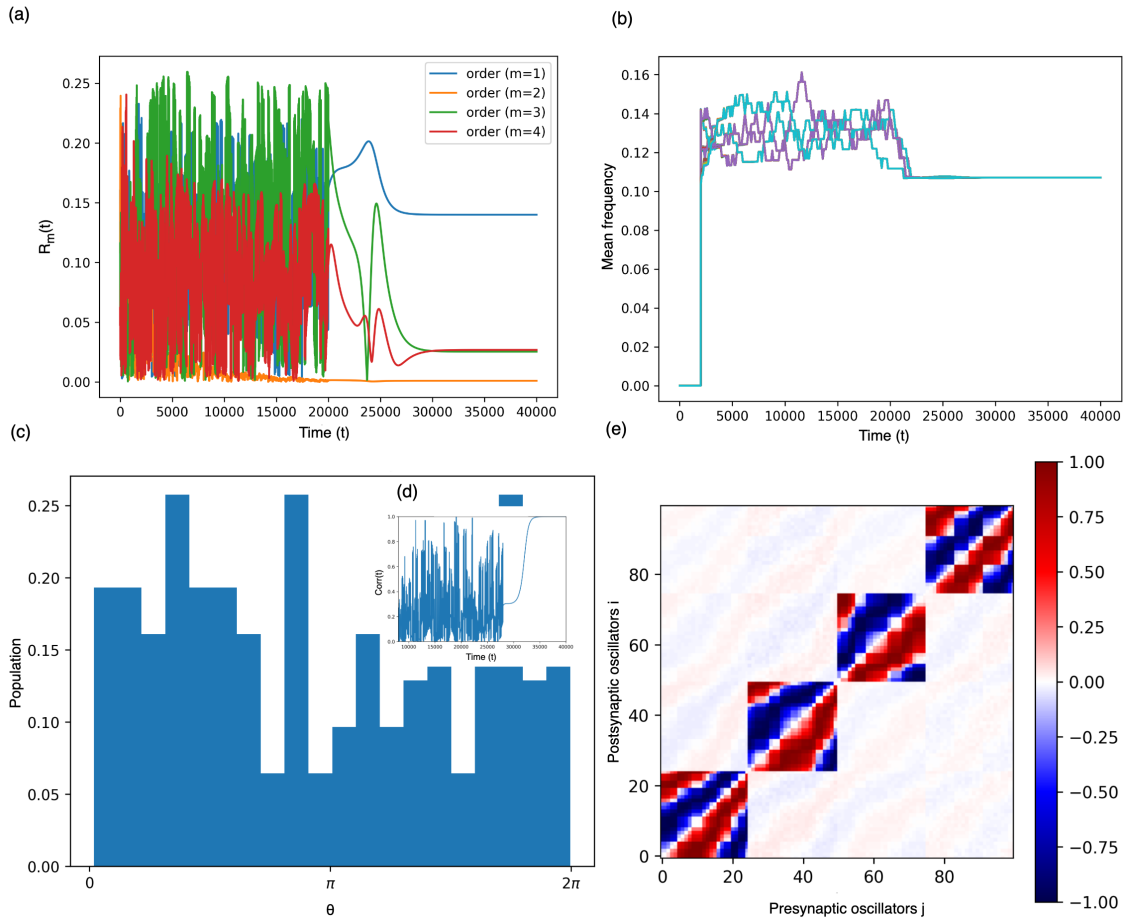


Figure A.7: Learning of four binary inputs randomly selected using a Hebbian asymmetric rule with: (a) the time development of the order parameters, (b) the mean frequencies of each oscillator through the time, (c) the distribution of neurons ($N = 100$) by phases (logarithmic scale) after learning phase, (d) the autocorrelations of the phase pattern and (e) the weighted connectivity matrix (sorted by phases in each cluster). Parameters: duration of learning = 20000 iteration steps, $\alpha = 0.3\pi$ and $\beta = 0$.

Two overlapping inputs In Fig. A.8, we reproduce exactly the same experiment as in Fig. 1.10 of the main text with two overlapping input vectors made of binary values, but this time considering the Hebbian asymmetric plasticity rule. No major differences are observed here, apart from the different dynamics within each cluster and the hubs (i.e. a splay state, see Sec. 1.2.2) and the associated structural organization.

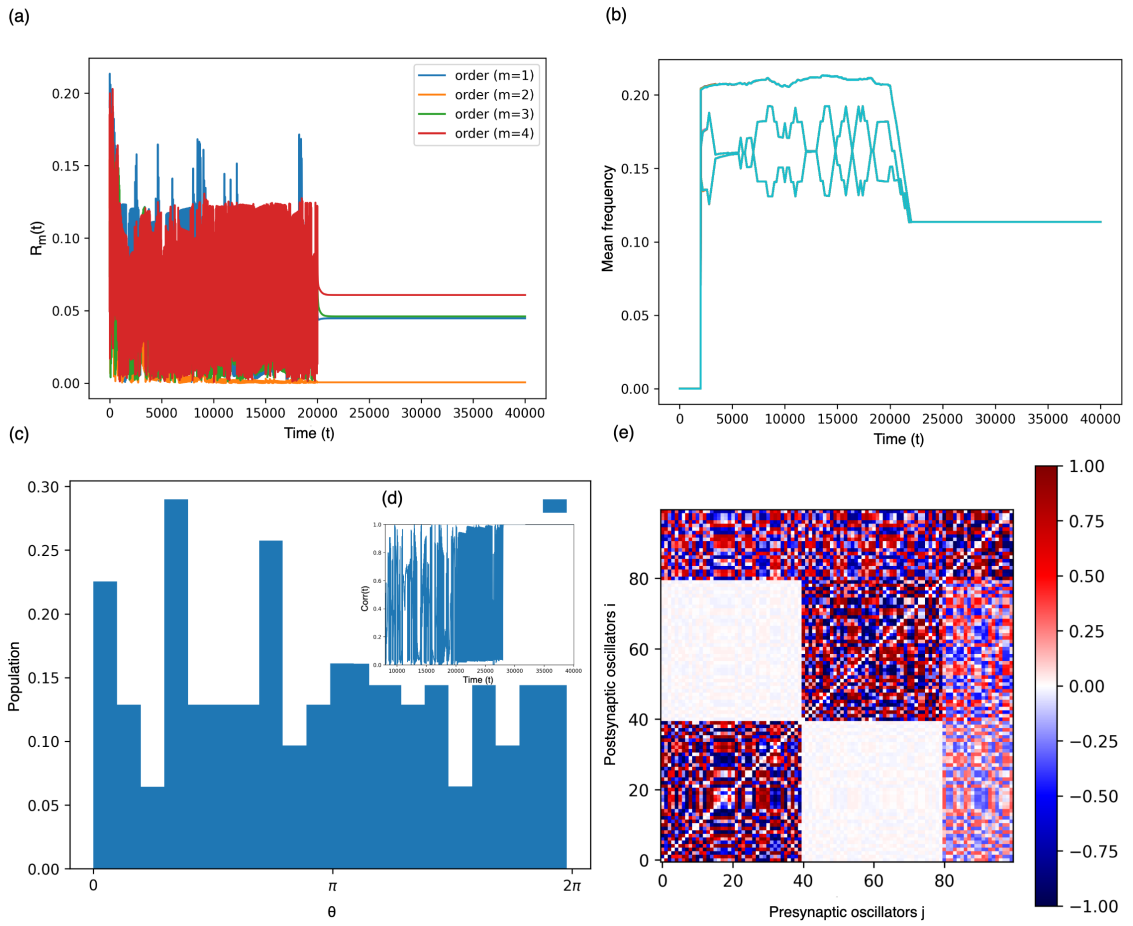


Figure A.8: Learning of two overlapping binary inputs randomly selected using a Hebbian asymmetric rule with: (a) the time development of the order parameters, (b) the mean frequencies of each oscillator through the time, (c) the distribution of neurons ($N = 100$) by phases (logarithmic scale) after learning phase, (d) the autocorrelations of the phase pattern and (e) the weighted connectivity matrix. Parameters: duration of learning = 20000 iteration steps, $\alpha = 0.3\pi$ and $\beta = 0$.

Recognition patterns This alternative experiment aims to give some premises for learning and recognition of simple patterns. To do this, we present three input vectors to a network composed of 10 Kuramoto oscillators in the same manner as in Sec. 1.3.2 of the main text. These vectors are made up of 7 binary values encoding the information and 3 binary values encoding the item label. Thus, as represented in the protocol in Fig. A.9, the aim of the learning phase is to associate a binary pattern with a label, using the temporal correlation of these inputs. In this way, during the testing phase, where only the patterns are presented, the activation of certain oscillators should imply the reactivation of a particular label oscillator and therefore the recognition of the pattern.

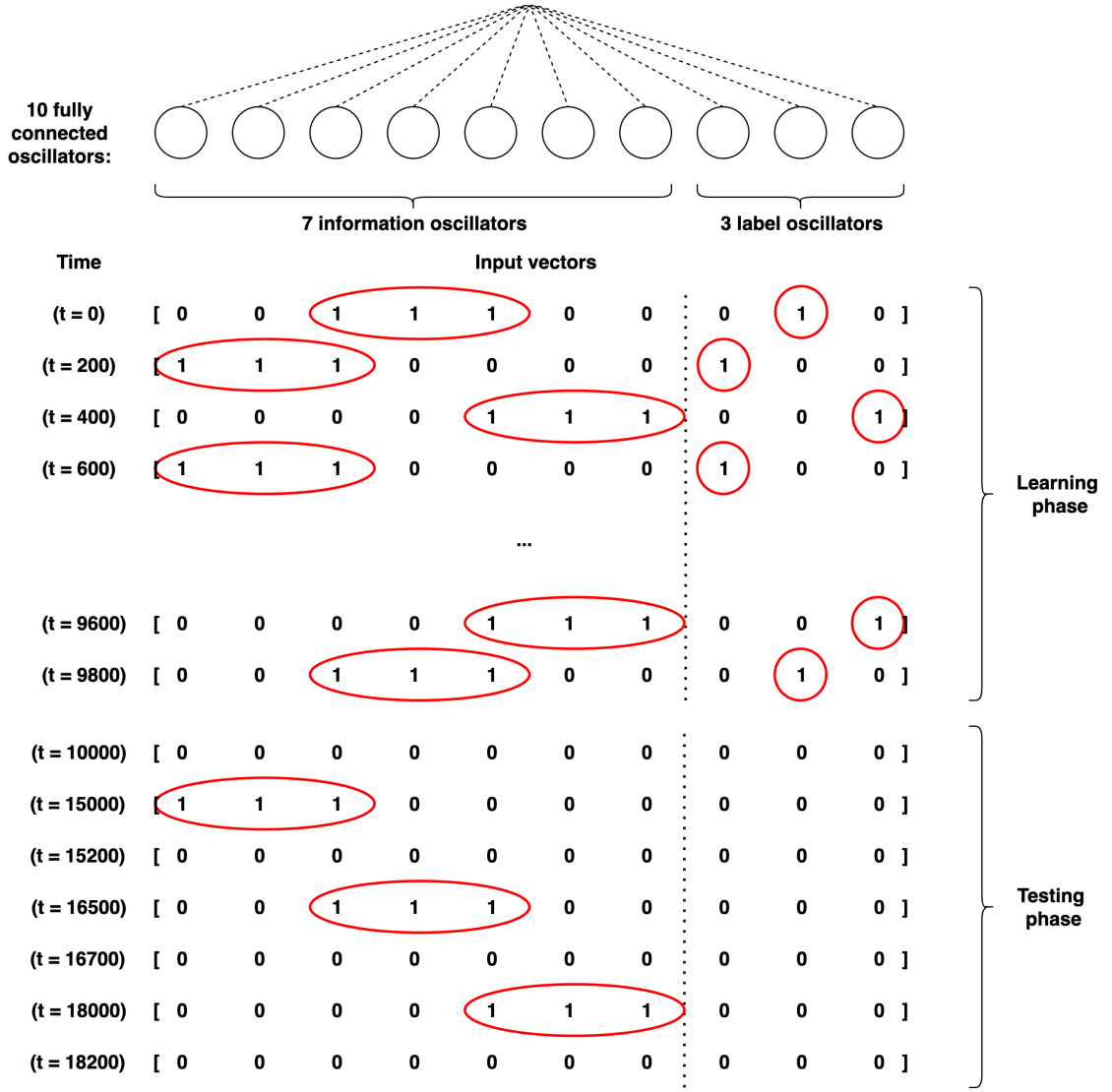


Figure A.9: Diagram showing the protocol for learning and recognising three patterns in a network of ten Kuramoto oscillators. During the learning phase, input vectors are applied alternately to the oscillators in a random manner during 200 iteration steps. In these vectors, 7 values are devoted to information and 3 values to the label associated with the pattern. During the learning phase, the labels are active playing a supervisory role, unlike during the testing phase. During the testing phase, each pattern is presented alone one after the other (with a short rest in between) in order to observe which label reacts most to a given pattern. The red circles highlight the active oscillators in the input vectors.

The results obtained with this protocol are represented in Fig. A.10. Firstly, as expected, when a pattern is presented during the testing phase, the associated oscillator label admits a mean frequency of higher amplitude, whatever the pattern and the plasticity rule employed (see panel (b)). We also observe that the labels of patterns that share oscillators

in their code show an increase in frequency. For example, the first pattern sharing one oscillator in common with the second pattern, the presentation of the first pattern implies a frequency increase of both labels. However, the frequency amplitude of the corresponding label remains higher, which confirms the possibility of the network to correctly recognize simple binary patterns. Apart from the learning phase and the presentation of a pattern, the network remains as before in the state associated with its plasticity function, as evidenced by the order parameters in panel (a). Concerning the connectivity obtained after the learning phase in panel (c), the associations between the information oscillators and labels are clearly visible with the oscillators coding for the same motif strongly connected (i.e. weights close to 1 or -1) and those uncorrelated with weak connections. Similarly, the overlaps brought by oscillators coding for different items are characterized by medium-value connections, highlighting the ambiguity of the information. This experiment is a first step in linking our results to AI applications and will subsequently be useful for learning to recognize sensory modalities. Although not developed here, we can also imagine the reverse process where we stimulate a given label to generate the associated pattern. This aspect will be developed further in chapter 4.

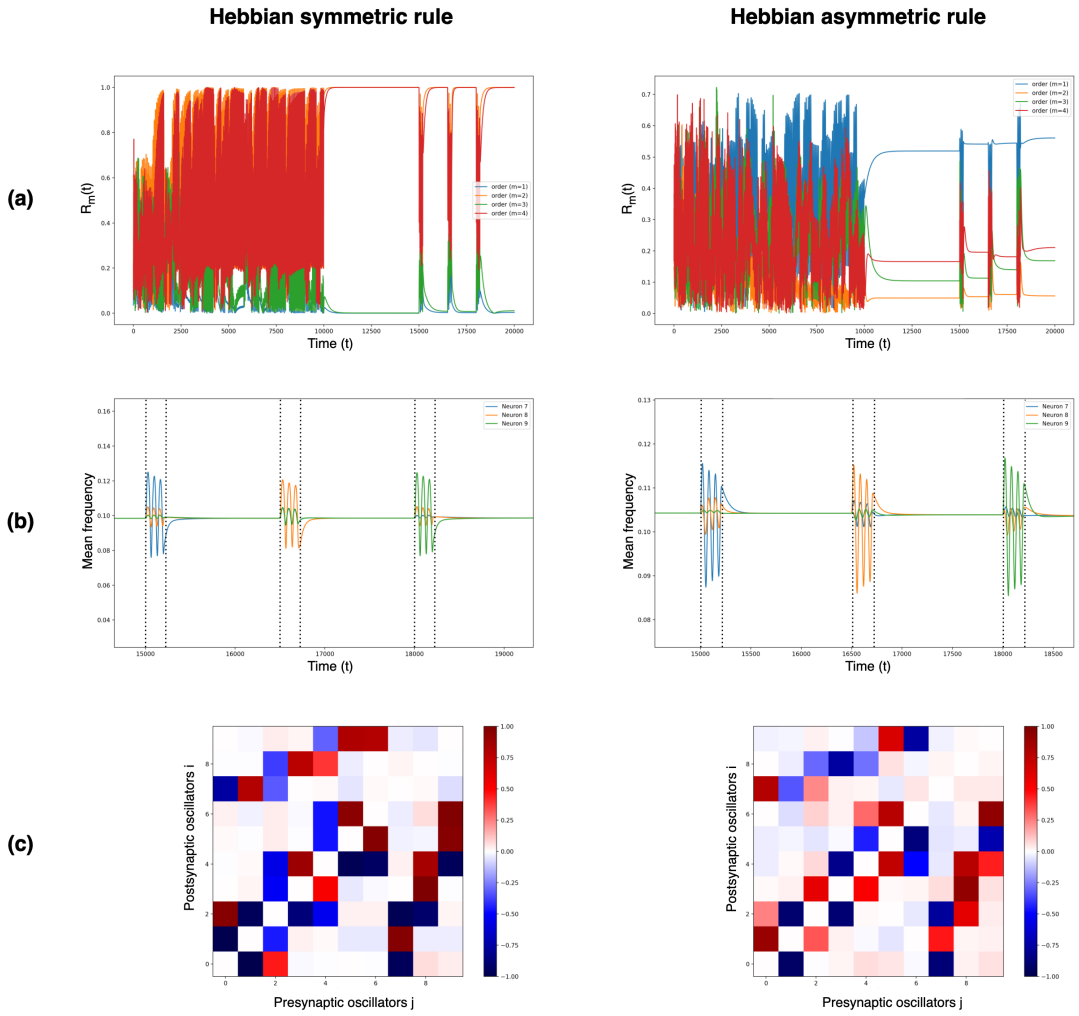


Figure A.10: Results obtained for the learning and recognition three patterns in a network of ten Kuramoto oscillators using the Hebbian symmetric rule and Hebbian asymmetric rule. (a) The time development of the order parameters during the learning phase (time 0 to 10000) and the testing phase (time 10000 to 20000). (b) The mean frequencies of the 3 label oscillators through the time (neuron 7 for pattern 1, neuron 8 for pattern 2, neuron 9 for pattern 3), the dashed lines represent the ranges where an input is presented (i.e. pattern 1, pattern 2 and then pattern 3). (c) the weighted connectivity matrix after the learning phase. Parameters: number of iteration = 20000, $\alpha = 0$, $\beta = -0.5\pi$ for the Hebbian symmetric rule and $\alpha = 0.3\pi$, $\beta = 0$ for the Hebbian asymmetric rule.

Frequency and synchrony based adaptation In this last alternative experiment, we mix the two approaches of Sec. 1.3.1 and Sec. 1.3.2 by considering input vectors containing values of different amplitudes, as represented in Fig. A.11 (a). Specifically, we consider four input vectors with values taking the form of Gaussian of amplitude 2 and standard deviation of 12.5, centred in oscillators 12, 37, 62 and 87 respectively. This consequently

implies overlaps in the stimulated oscillators as in Fig. 1.10. Except these differences in the inputs content, the learning protocol remains the same as in Sec. 1.3.2. The weight matrices obtained after learning with the two plasticity rules are represented in Fig. A.11 (b) and (c). First of all, we do not observe any significant difference between the rules. In both cases, small modular structures seem to appear at the centre and extremities of each Gaussian input where the oscillators receive inputs of similar frequencies. On the slopes of the Gaussian, where the frequency differences are larger, connections only form between oscillators that are physically close in the vector or at the opposite end (according to the centre of the Gaussian). Similarly, the overlaps induced by the inputs create connections between the different areas stimulated. These particular forms of input explain the complex structures obtained with coupling weights of intermediate values. The dynamics of the network are not described here, since they are similar to those of the previous experiments, with a dependency linked solely to the plasticity rule used.

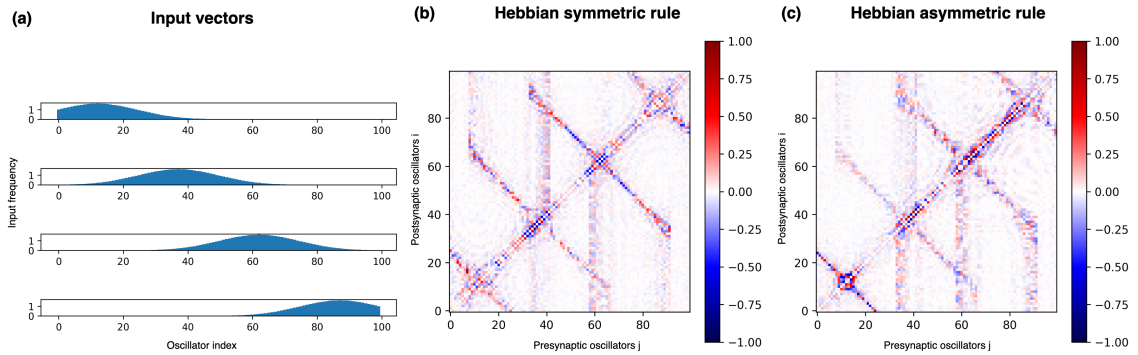


Figure A.11: Learning of four overlapping Gaussian inputs randomly selected with: (a) the four input vectors learned (amplitude = 2, standard deviation = 12.5, means = 12, 37, 62, 87), (b) the weighted connectivity matrix using the Hebbian symmetric rule and (c) the weighted connectivity matrix using the Hebbian asymmetric rule. Parameters: $\alpha = 0$, $\beta = -0.5\pi$ for the Hebbian symmetric rule and $\alpha = 0.3\pi$, $\beta = 0$ for the Hebbian asymmetric rule.

A.3 Alternative stimulation protocols with θ-neurons

In this appendix, we present results relative to additional stimulation protocols which consist of variations of the initial protocols depicted in the Sec. 2.2. The dynamics are simulated using the θ-model as in Chapter 2.

Emergence of splay states The idea of this experiment is to start from the connectivity matrix shown in panel (1) of Fig. A.12 where all the excitatory neurons are decoupled one from another, but where each of them has a single post-synaptic connection to an inhibitory neuron. The latter consequently inhibits all neurons except the one with which it is associated. As evident from the raster plot (2) displayed at time T0 of Fig. A.12, all

pairs of excitatory-inhibitory neurons form pseudo-clusters that are desynchronized from the other pairs forming a splay state. Certainly, this represents an ideal limiting case since we artificially create the structure without recurring to any learning process. Nevertheless, it represents the extremal case in which a maximum number of possible clusters can be generated, confirming that one inhibitory neuron per cluster is theoretically sufficient to maintain each pair of neurons desynchronized from the others.

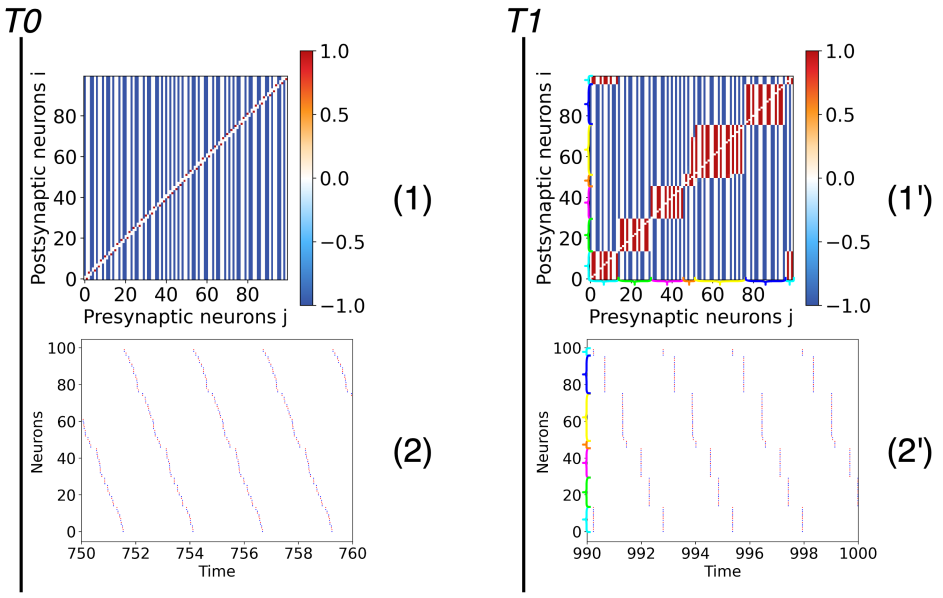


Figure A.12: Splay states in the θ-model. Results of the experiment where the connectivity matrix is initialized to exhibit $N/2$ clusters corresponding to a splay state. The time T0 corresponds to the beginning of a resting phase in absence of any stimulations, after a short transient period $t_t = 750$ has been discarded. The time T1 corresponds to the end of a long period of spontaneous activity during which synaptic weights are consolidated. Panels labeled (1) and (1') represent the weight matrices at times T0 and T1, respectively: the color denotes if the connection is excitatory (red), inhibitory (blue) or absent (white). Panels (2) and (2') are raster plots at times T0 and T1, displaying the firing times of excitatory (red dots) and inhibitory (blue dots) neurons. Note that in both figures, for clarity the neurons are sorted by phases. Also to better visualize the splay state, a homogeneous system without noise is considered in this experiment. The cyan, green, magenta, orange, yellow and dark blue brackets represent clusters 1, 2, 3, 4, 5 and 6 respectively in weight matrices and raster plots after the consolidation phase.

However, after a long period of spontaneous activity and consolidation, this structure and this particular dynamics are not maintained. Indeed as explained in Sec. 2.3.2 of the main text, the phase (time) potentiation window of the plasticity function determines the interval within which the neurons are considered to be correlated. The clusters forming the splay state spike at very close (though strictly different) times, however the plasticity forces the temporally close clusters to merge over the long term, as evident in the weighted

connectivity matrix (1') obtained at long time after consolidation (T1). Thus, the network reaches a stable state with about 5 clusters spiking at distinct times (see raster plot (2')).

Two memory patterns with randomly stimulated neurons In the previous experiments, when a group of neurons was stimulated, all the neurons of the group received the same stimulation input. In this experiment, only a certain percentage of the neurons within the group receive the stimuli. The neurons are randomly selected with a probability of 50%. The protocol and the results corresponding to this experiment are depicted in Figs. A.13(a) and A.13(b).

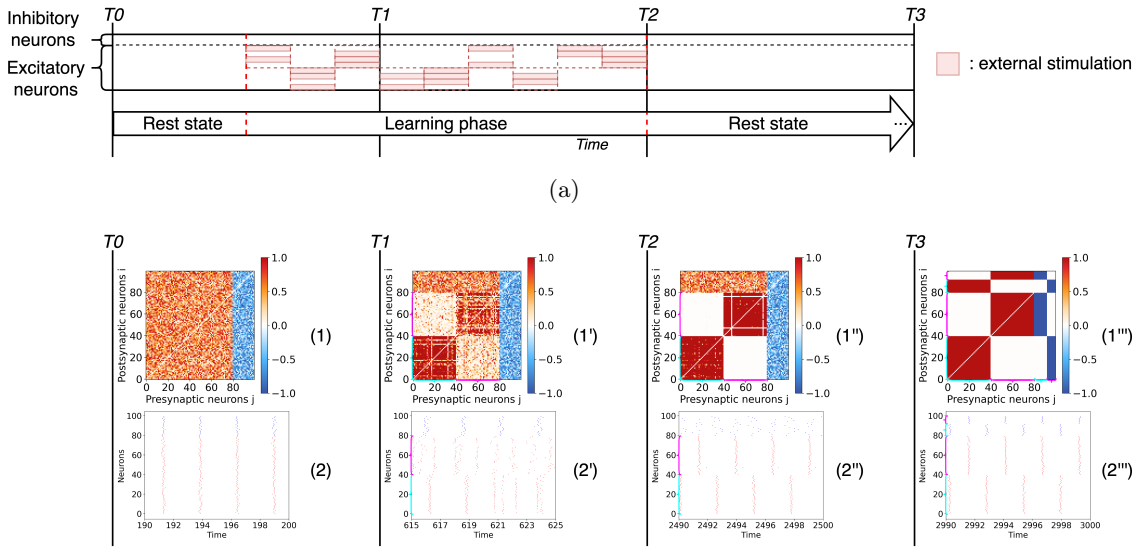
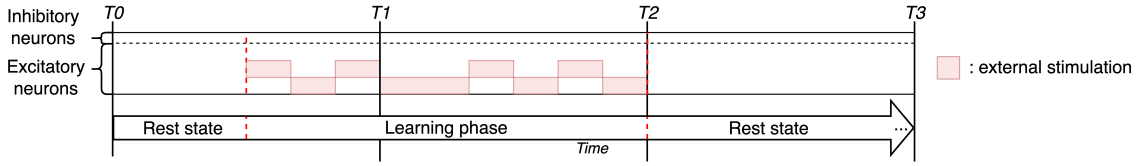


Figure A.13: Entrainment of a networks of excitatory and inhibitory θ-neurons with two stimuli applied to random subsets of neurons within each group. (a) Schema of the experiment consisting of the stimulation of randomly chosen excitatory neurons of each of the two groups every stimulation period. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. Note that the inhibitory neurons are sorted by phases at time T3 for visualization purposes. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

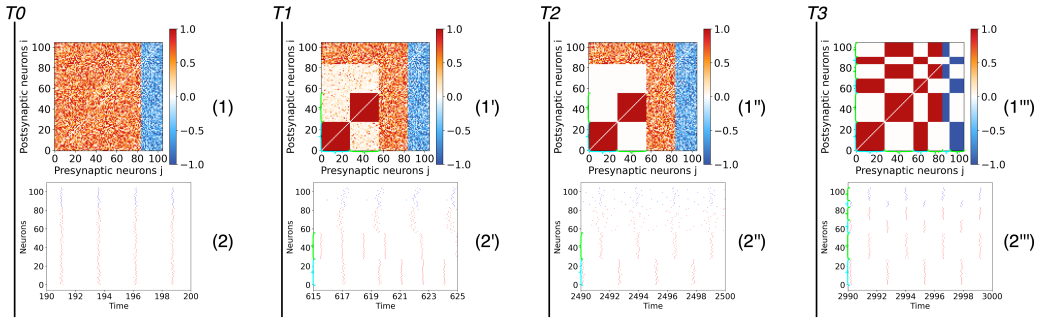
The first difference observed here is that in the connectivity matrix (1") obtained after the learning (time T2), the clusters are less pronounced due to the random nature of the activation of the neurons that constitute each of them. Despite this, the two groups can be still distinguished, firing almost together at different times (raster plots (2")). Afterwards, the reinforcement phase (time T3) allows to obtain the same connectivity and dynamical evolution as in the original protocol. The interest of this protocol is to demonstrate the possibility of creating memory patterns with more realistic input signals. Indeed in the

brain, not all the neurons in an area are stimulated at the same time.

Three groups of excitatory neurons, with only two subject to learning The idea of this experiment is to reproduce the same original protocol but this time keeping a group of excitatory neurons unstimulated and therefore not subject to a strong adaptation. See the protocol of this experiment and its results in Figs. A.14(a) and A.14(b).



(a)



(b)

Figure A.14: Entrainment of a network of excitatory and inhibitory θ-neurons with two non-overlapping stimuli and a group of free excitatory neurons. (a) Stimulation protocol with two stimulated groups of excitatory neurons and another group of unstimulated neurons. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. Note that the inhibitory neurons and the initially untrained excitatory neurons are sorted by phases at time T3 for visualization purposes. The cyan and green brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

The first observation after the learning phase (time T2) is that the two structural clusters created, connectivity matrix (1"), are also visible in the raster plots as two groups of neurons firing in anti-phase (2"). However, the group of unstimulated excitatory neurons fires quite irregularly at this time as the inhibitory neurons. This similarity with the inhibitory group is also observable after the consolidation process (time T3) where the unstimulated excitatory neurons associate with one of the two excitatory clusters and its associated inhibitory neurons. Thus, we can clearly see in matrix (1"') that both the untrained excitatory neurons and the inhibitory population are splitted into two structural clusters, each of them associated to one of the two memory patterns making them grow. Depending on the initial conditions it can happen that the two groups are not equally

populated and in extreme cases the unstimulated excitatory neurons can associate also to only one of the two memory patterns.

Note that this phenomenon is only possible when the original memory patterns involve a sufficient number of neurons with respect to the unstimulated group. When the number of the latter prevails, the two stimulated groups tend to synchronize and only one pattern emerges in the end. As discussed in the multi-cluster learning, the number of inhibitory neurons associated to each excitatory cluster has also a role in determining which is the minimal number of excitatory neurons in each memory pattern that can remain separated on the long run.

Three memory patterns learned, two maintained This experiment is analogous to the experiment reported in Fig. 2.5 in Sec. 2.3.2. In particular, three different stimuli are presented to the excitatory neurons as in the main text, but now only one inhibitory neuron is present in the network. The results are reported in Fig. A.15.

At the steps T0 and T1, we logically obtain similar results as in the main text. However at time T2 after the learning phase, although the three clusters are well formed as before (see connectivity matrix (1") in Fig. A.15(b)), two of them share a close dynamics, i.e. almost synchronized (as shown in in raster plot (2") in Fig. A.15(b)). The direct consequence of this is that in the long term (time T3), these two structural modules merge in only one (panel (1"')) and they are now completely synchronized (panel (2"')). In other words, one of the three learned items is forgotten and finally only two are memorized. This experiment confirms the previously established limit by showing that one inhibitory neuron is not sufficient to maintain three memories ($N_{inhibitory} = 1 \ngeq M - 1 = 2$ with $M = 3$ clusters). However, it also shows that a single inhibitory neuron is sufficient for maintaining two memories ($N_{inhibitory} = 1 \geq M - 1 = 1$ with $M = 2$ clusters).

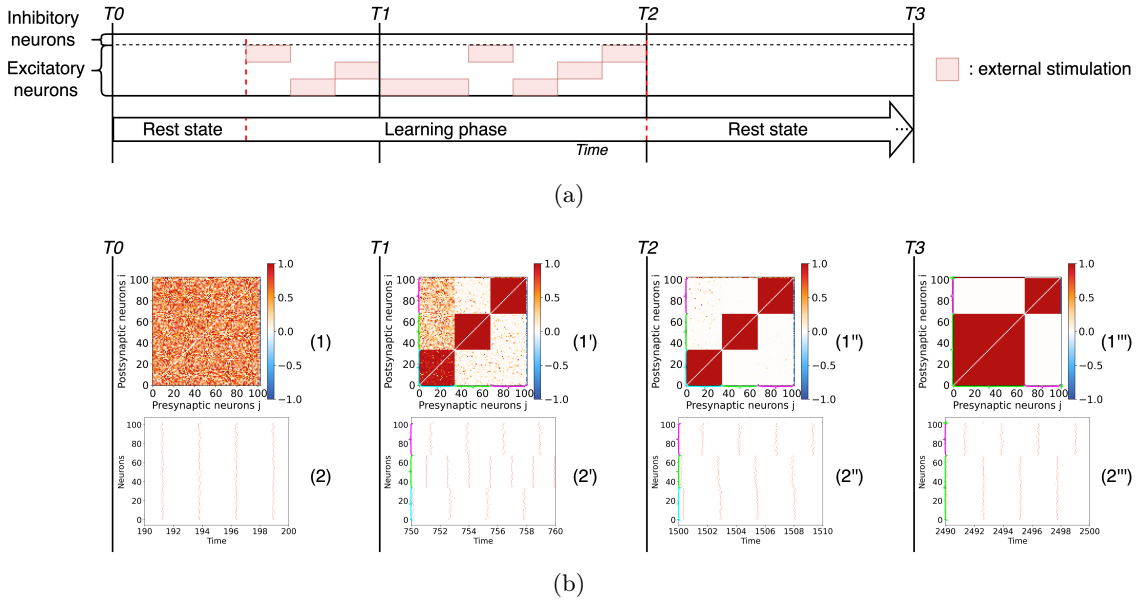


Figure A.15: Entrainment of a network of excitatory and inhibitory θ-neurons with three stimuli in presence of only one inhibitory neuron. (a) Stimulation protocol of the excitatory neurons via three different non-overlapping stimuli. (b) Entrainment results at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. The cyan, green and magenta brackets represent clusters 1, 2 and 3 respectively when they are visible in weight matrices and raster plots.

Unstable hub neurons In this last alternative stimulation protocol, we reproduce the same experiment shown in Fig. 2.7 in Sec. 2.3.3. As in the original case, we allow the two stimuli to share eight neurons among them but this time there are only sixteen inhibitory neurons in the network. The results of the experiment are reported in Fig. A.16.

During the steps T0 and T1, as expected we obtain similar results to the one reported in the main text. However at time T2 after the learning phase, although the two clusters and the "hub neurons" are well formed as in the main text (see connectivity matrix (1") in Fig. A.16(b)), now the two clusters almost synchronize (see raster plot (2") in Fig. A.16(b)). Consequently over the long post-training period (time T3), the clusters merge into one (panel (1"')) and they become completely synchronized (panel (2"')). Therefore, in addition to being unstable over the long term, these hub neurons make the modules unmaintainable. As speculated in the main text, the excitatory connections between modules that are mediated by the hubs must be compensated by an equivalent amount of inhibition. This corroborates the previously established limit by showing that sixteen inhibitory neurons are not sufficient to maintain eight hubs ($N_{inhibitory} = 16 \not\geq 2 * M + 1 = 17$ with $M = 8$ hubs).

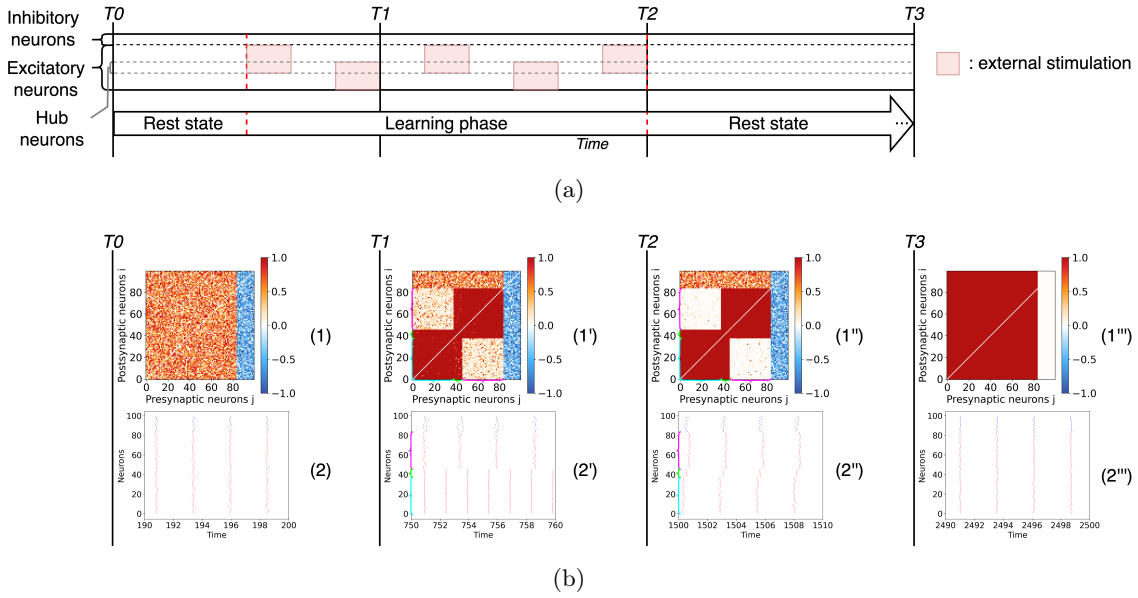


Figure A.16: Entrainment of a network of excitatory and inhibitory θ-neurons with two overlapping stimuli and 16 inhibitory neurons. (a) Schema of the experimental protocol showing that the two presented stimuli involve 8 shared neurons. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. The cyan, magenta and green brackets represent clusters 1, 2 and the hubs respectively when they are visible in weight matrices and raster plots.

A.4 Maintenance of neural assemblies with oscillators models

Throughout the chapter 2, single units have been simulated considering the θ-neuron model. We now reproduce the results of Fig. 2.3 of this chapter considering networks of phase oscillators (the Kuramoto model as in chapter 1) and of oscillators characterized by their phase and amplitude (the Stuart-Landau model). The adaptation of the synaptic weights will be governed by the same rule employed for the θ model and reported in Sec.2.2.

The Kuramoto model So we consider the Kuramoto model [225, 226, 228] for coupled phase oscillators with adaptation defined in chapter 1. Therefore, it is important to compare our findings for a classical neuronal model with the paradigmatic Kuramoto model, where the evolution of the phase θ_i of an oscillator is now described by:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{g}{N}\Big(\sum_{j=1}^{N} \kappa_{ij} \sin(\theta_j - \theta_i)\Big) + I_i(t) + \xi_i(t) \quad . \tag{A.1}$$

The natural frequency of oscillator i is denoted by ω_i, g represents the global coupling strength, while κ_{ij} is the relative directed coupling from the oscillator j towards the oscil-

lator i. In this context, $I_i(t)$ represents an external driving term and $\xi_i(t)$ represents an additive Gaussian noise.

The Stuart-Landau network model One of the limitations of the Kuramoto model is the unrealistic nature of its dynamics compared to biological oscillators, in particular the fact that the evolution of the amplitudes is neglected. The normal form for a non-linear oscillator in the proximity of a Hopf bifurcation is represented by the Stuart-Landau model which describes not only the phase variations but also the amplitude variations of the oscillator as enunciated in chapter 1. The dynamics of this oscillator is given in terms of a complex variable $z_i = \rho_i e^{\hat{\imath}\theta_i} = x_i + \hat{\imath} y_i$ where ρ_i represents the amplitude of the ith unit, $\theta_i \in [-\pi, \pi[$ its phase and $\hat{\imath}^2 = -1$. The evolution of the i-th oscillator in a network is described by Deco et al. [110]:

$$\frac{dz_i}{dt} = z_i\Big[\alpha_i + \hat{\imath}(\omega_i + I_i(t)) - |z_i^2|\Big] + \frac{g}{N}\Big(\sum_{j=1}^{N} \kappa_{ij} z_j\Big) + \xi_i(t). \tag{A.2}$$

Most of the quantities appearing in Eq. (A.2) have been already defined in Eq. (A.1). The only new element is the real parameter α_i which governs the type of dynamics of the system: $\alpha_i < 0$ leads to a stable fixed point and $\alpha_i > 0$ to a stable limit cycle oscillation [110]. The Gaussian noise $\xi_i(t)$ is here considered to be complex.

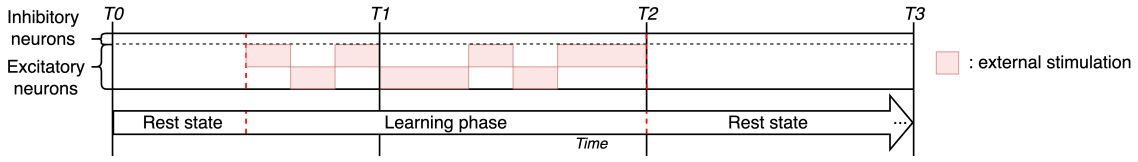
In order to guarantee a similar firing rate at rest (where the spike is defined like in the θ-neuron model) and a stimulation of the same order, we have adapted the specific parameters of these two models as indicated in table A.1. Parameters not indicated in the table are considered identical to the ones employed for the θ-neuron model.

Table A.1: Parameters for the Kuramoto and Stuart-Landau oscillator networks

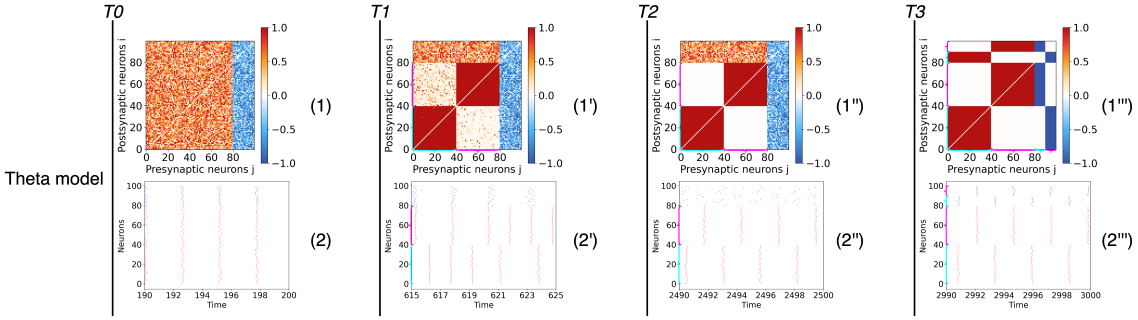
Parameters	Values
ω_i	$\mathcal{N}(2.45,\ 0.02)$
α	$\mathcal{N}(0.0,\ 0.2)$
g	2
$I(t)$	$\{0, 3.46\}$
$\xi(t)$	$\mathcal{N}(0.0,\ 0.2)$

Experiment Following the same stimulation protocol as in Fig. 2.2 of the main text (see Fig. A.17(a)), the corresponding results for the θ-model, the Kuramoto and the Stuart-Landau models are reported in Figs. A.17(b) – A.17(d). Essentially, the same results are obtained for the three models as seen in the emerging weight matrices after learning (1") and reinforcement (1"'), and in the corresponding spiking behaviour (raster plot (2"')). The Kuramoto and the Stuart-Landau oscillators seem just a little noisier than the θ-neurons. We have obtained with the Kuramoto and Stuart-Landau network models quite similar results also for all the other protocols and numerical experiments performed in chapter 2

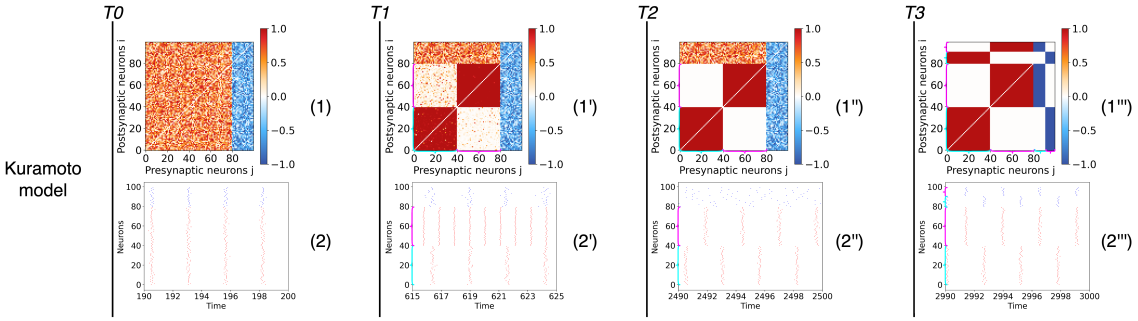
for the θ-neuron. These results prove our findings to be quite general and not limited to a specific model of oscillator.



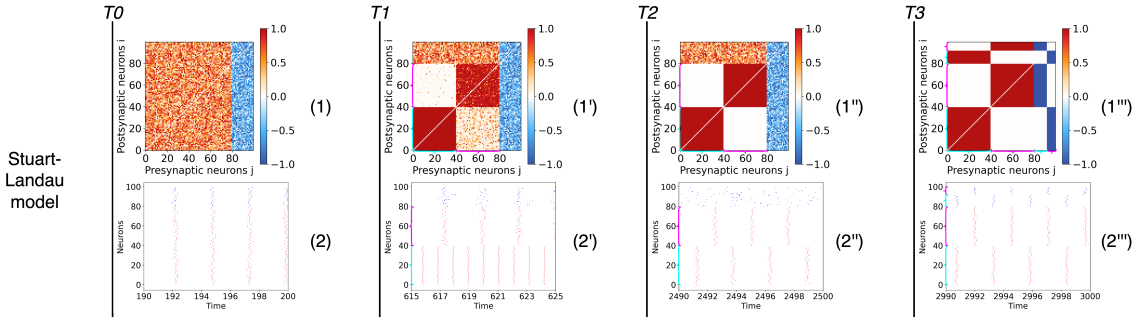
(a)



(b)



(c)



(d)

Figure A.17: Entrainment of networks of 80% excitatory and 20% inhibitory neurons for three different oscillatory models. (a) Schema of the experiment leading to the emergence of two clusters due to the stimulation of two non-overlapping excitatory neuronal populations. (b) Results for a network of θ-neurons, (c) for a network of Kuramoto phase oscillators and (d) for a network of Stuart-Landau oscillators at different instants of the simulation. The time labels and the graphs have the same significance and content as in Fig. 2.3. Note that the inhibitory neurons are sorted by phases at time T3 for visualization purposes. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

Appendix B

Supplementary results with spiking neurons

In this appendix chapter, we describe supplementary materials with network made of spiking neurons.

B.1 Alternative stimulation protocols with QIF neurons

In this appendix, we present results relative to additional stimulation protocols which consist of variations of the initial protocols depicted in Fig. 3.2. The simulations use QIF neurons subject to STDP as in Chapter 3.

B.1.1 Learn 3 clusters

In this alternative protocol, we reproduce the experiment of Fig. 3.2 (e) of the main text but this time considering three stimuli to be learned instead of two. The results obtained are described in Fig. B.1. We obtain similar results with this time the formation of three modular structures in the weighted connectivity (1) at time T2 involving spontaneous recalls of the three different memories in the dynamics of raster plot (2).

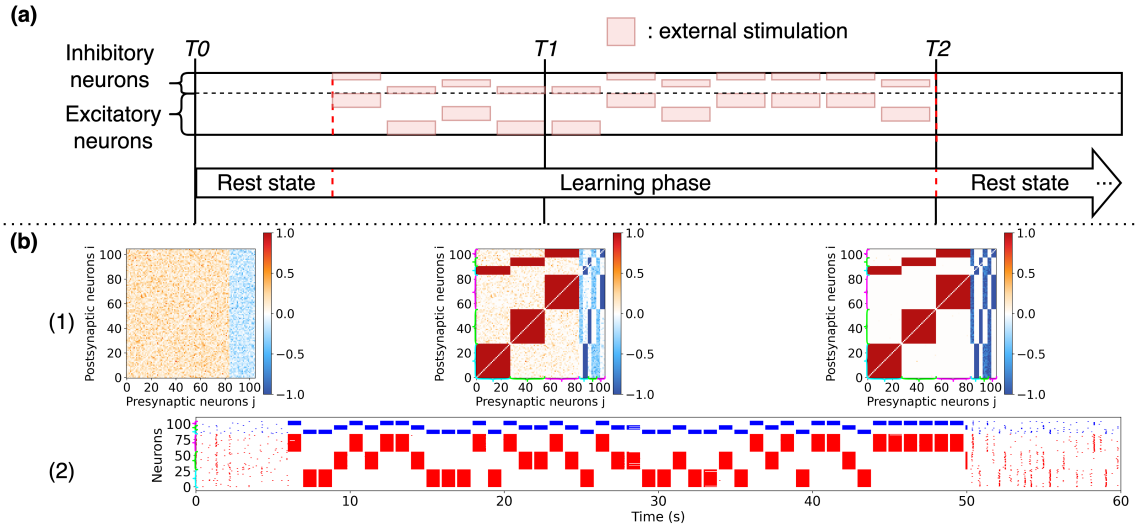


Figure B.1: Learning of 3 stimuli in a network of QIF neurons. (a) Experimental protocol consisting of the stimulation of three non-overlapping neuronal populations of QIF neurons with plastic synapses. Stimuli are presented in temporal alternation. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan, green and magenta brackets represent clusters 1, 2 and 3 respectively when they are visible in weight matrices and raster plots.

B.1.2 Learn 4 clusters

In this alternative protocol, we reproduce the experiment of Fig. 3.2 (e) of the main text but this time considering four stimuli to be learned instead of two. The results obtained are described in Fig. B.2. We obtain similar results with this time the formation of four modular structures in the weighted connectivity (1) at time T2 involving spontaneous recalls of the four different memories in the dynamics of raster plot (2).

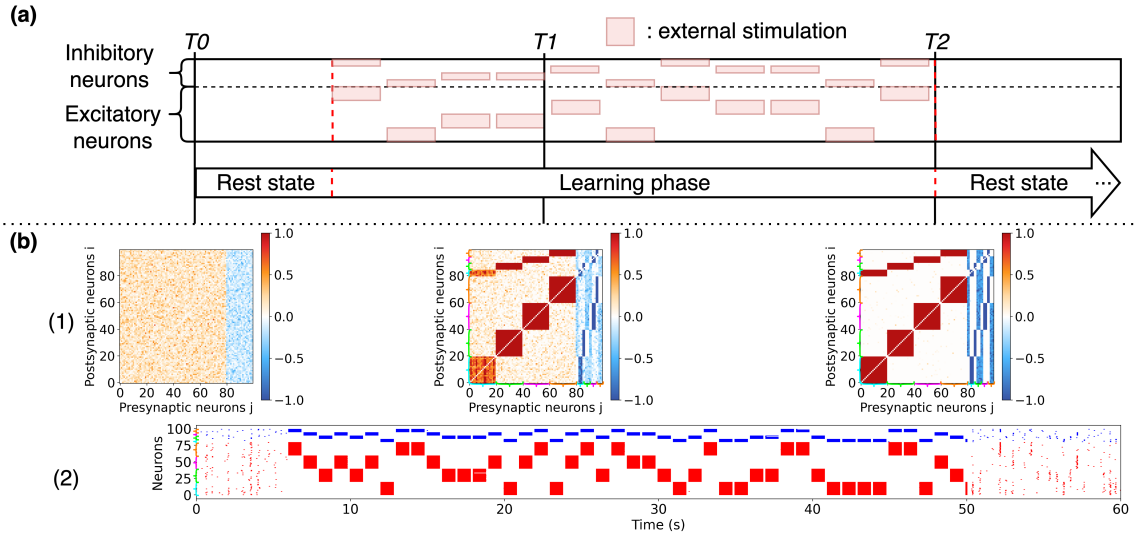


Figure B.2: Learning of 4 stimuli in a network of QIF neurons. (a) Experimental protocol consisting of the stimulation of four non-overlapping neuronal populations of QIF neurons with plastic synapses. Stimuli are presented in temporal alternation. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan, green, magenta and orange brackets represent clusters 1, 2, 3 and 4 respectively when they are visible in weight matrices and raster plots.

B.1.3 Learn 4 clusters sharing hubs

In this alternative protocol, we reproduce the experiment of Fig. 3.6 of the main text but this time considering four stimuli sharing 8 neurons to be learned instead of two stimuli. The results obtained are described in Fig. B.3. We obtain similar results as in Fig. B.2 with the formation of four modular structures in the weighted connectivity (1) at time T2, accompanied by the formation of hubs, as in the main experience, which are here connected (incoming and outgoing connections) with the four clusters. Regarding the dynamics in raster plot (2), in analogy with the main experiment, we obtain different types of spontaneous recalls with recalls of one of the four clusters alone, the recalls of one of them accompanied by the hubs or the recall of hubs alone. This experiment again highlights the richness of the different dynamics that the network can display and maintain for a certain period of time.

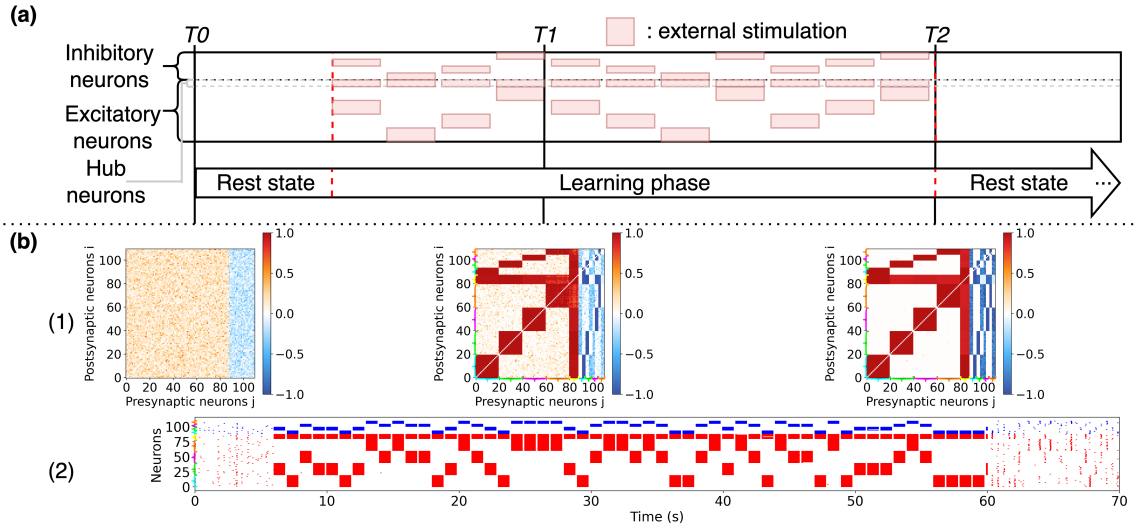


Figure B.3: Learning of 4 overlapping stimuli in a network of QIF neurons. (a) Schema of the experiment protocol showing that the four presented stimuli involve 8 shared neurons. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan, green, magenta, orange and yellow brackets represent clusters 1, 2, 3, 4 and the hubs respectively when they are visible in weight matrices and raster plots.

B.1.4 Untrained group of neurons

This alternative protocol is analogous to that of experiment of Fig. 3.2 (e) of the main text. The only difference lies in the fact that a group of excitatory neurons is never stimulated and so untrained. The results obtained are described in Fig. B.4. On the one hand, we obtain analogous results with the formation of two modular structures in the weighted connectivity (1) at time T2 involving spontaneous recalls of the two different memories in the dynamics of raster plot (2). On the other hand, neurons of the untrained group are weakly connected to each other in accordance with the absence of stimulation and are decoupled from the other clusters while receiving Hebbian inhibition connections from them. As a result, these neurons spike in a totally asynchronous and irregular way, without impacting the dynamics of the rest of the network.

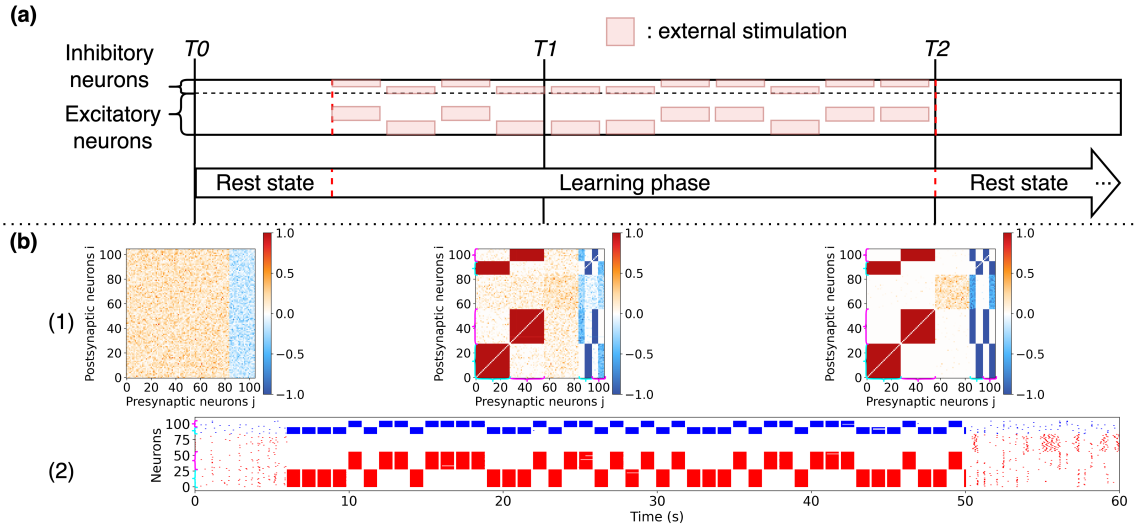


Figure B.4: Three groups of excitatory neurons, with only two subject to learning in a network of QIF neurons. (a) Experimental protocol consisting of the stimulation of two non-overlapping neuronal populations and a group of free excitatory neurons of QIF neurons with plastic synapses. Stimuli are presented in temporal alternation. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

B.1.5 Randomly stimulated neurons

This alternative protocol is analogous to that of experiment of Fig. 3.2 (e) of the main text. The only difference lies in the fact that when a cluster is selected during learning, a random number of neurons (with a probability of 0.5) is stimulated. The results obtained are described in Fig. B.5. The direct consequence is that the two modular structures in the weighted connectivity (1) at time T2 are less well formed compared to the original experiment. Nevertheless, the clusters remain decoupled with the same feed-forward and feedback inhibition described in the main text. Only the weights within the clusters appear to be more random. Therefore, although the spontaneous recalls of the two memory items are present in the dynamics of raster plot (2), they appears to be somewhat sparser and less synchronized. We assume that this is largely due to the fact that the connections to and from the inhibitory neurons are incomplete. Indeed, in Fig. 3.5 of the main text, the randomness of the E-E connections does not impact the spontaneous recalls. This highlights the need to reach a convergence of the weights linked to inhibition to correctly memorize the items. Nevertheless, this experiment also shows that even by partially learning memory items, averaged over time, the entire original item is somehow retrieved and learned.

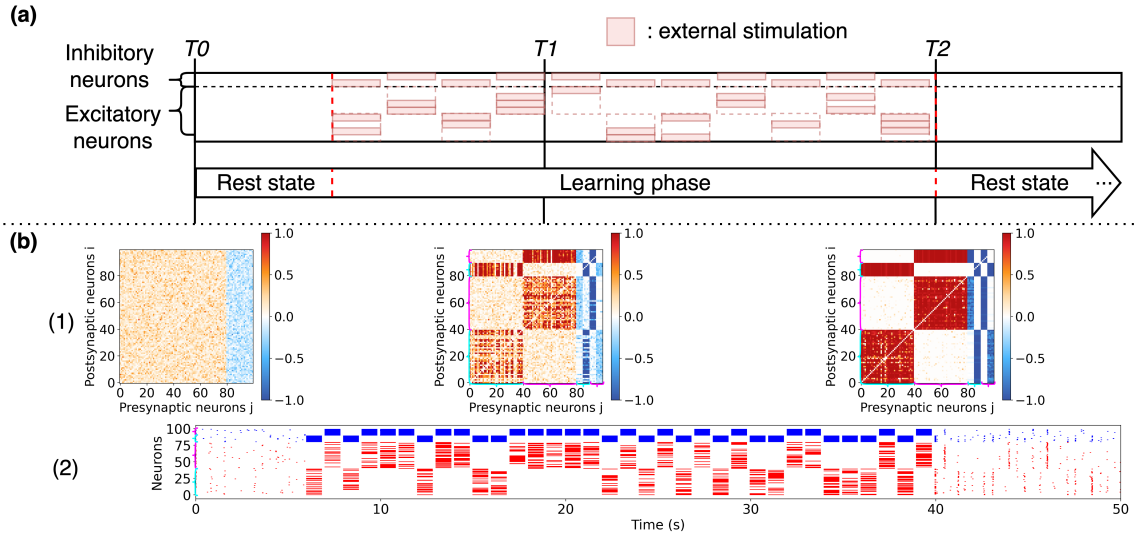


Figure B.5: Learning of 2 memory patterns with randomly stimulated QIF neurons. (a) Experimental protocol consisting of the random stimulation of neurons of two non-overlapping neuronal populations with plastic synapses. Stimuli are presented in temporal alternation. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

B.1.6 Random stimulation values

This last alternative protocol is again analogous to that of experiment of Fig. 3.2 (e) of the main text. The only difference lies in the fact that when a cluster is stimulated (excitatory and inhibitory neurons), the neurons within it receive inputs of random values (i.e. involving frequencies between 50 and 100 Hz). The results obtained are described in Fig. B.6. We observe very similar results to those in Fig. B.5. However in this case, the weights connections of (1) at time T2 seem stronger than in the previous experiment. As a result, the spontaneous recalls appear to be more visible in the dynamics of raster plot (2). In addition to the conclusions made in the previous study, this experience allows us to say that the spatio-temporal correlations of the applied inputs are more impactful on the formation of the modular structures than the intensities of these same inputs. Indeed, while the frequency of the inputs applied to the neurons necessarily has an impact on the encoding of information and the construction of memory, as can be seen here, compared with the previous experiment where the neurons received the same inputs but at times that were not necessarily correlated, the structure is much better learned in this case.

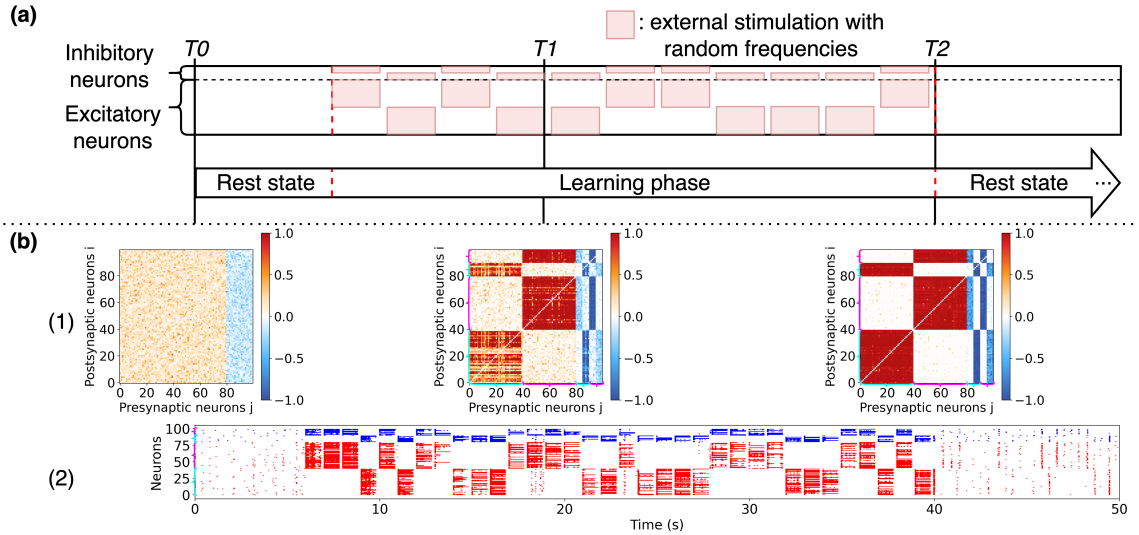


Figure B.6: Learning of 2 memory patterns with random stimulation values in a network of QIF neurons. (a) Experimental protocol consisting of the stimulation of two non-overlapping neuronal populations of QIF neurons with plastic synapses with inputs of random frequencies. Stimuli are presented in temporal alternation. (b) The results are given at different instants of the simulation. The time labels and the graphs have the same significance and content as in the main text. The cyan and magenta brackets represent clusters 1 and 2 respectively when they are visible in weight matrices and raster plots.

B.2 Supplementary materials multimodal application

In this appendix, we present supplementary materials from the multimodal application of Chapter 4.

B.2.1 Audio-visual database

The database learned during the multimodal application is represented in Fig. B.7 with the original data in (a) and (d), the extracted features in (b) and (e), and the actual learned information in (c) and (f).

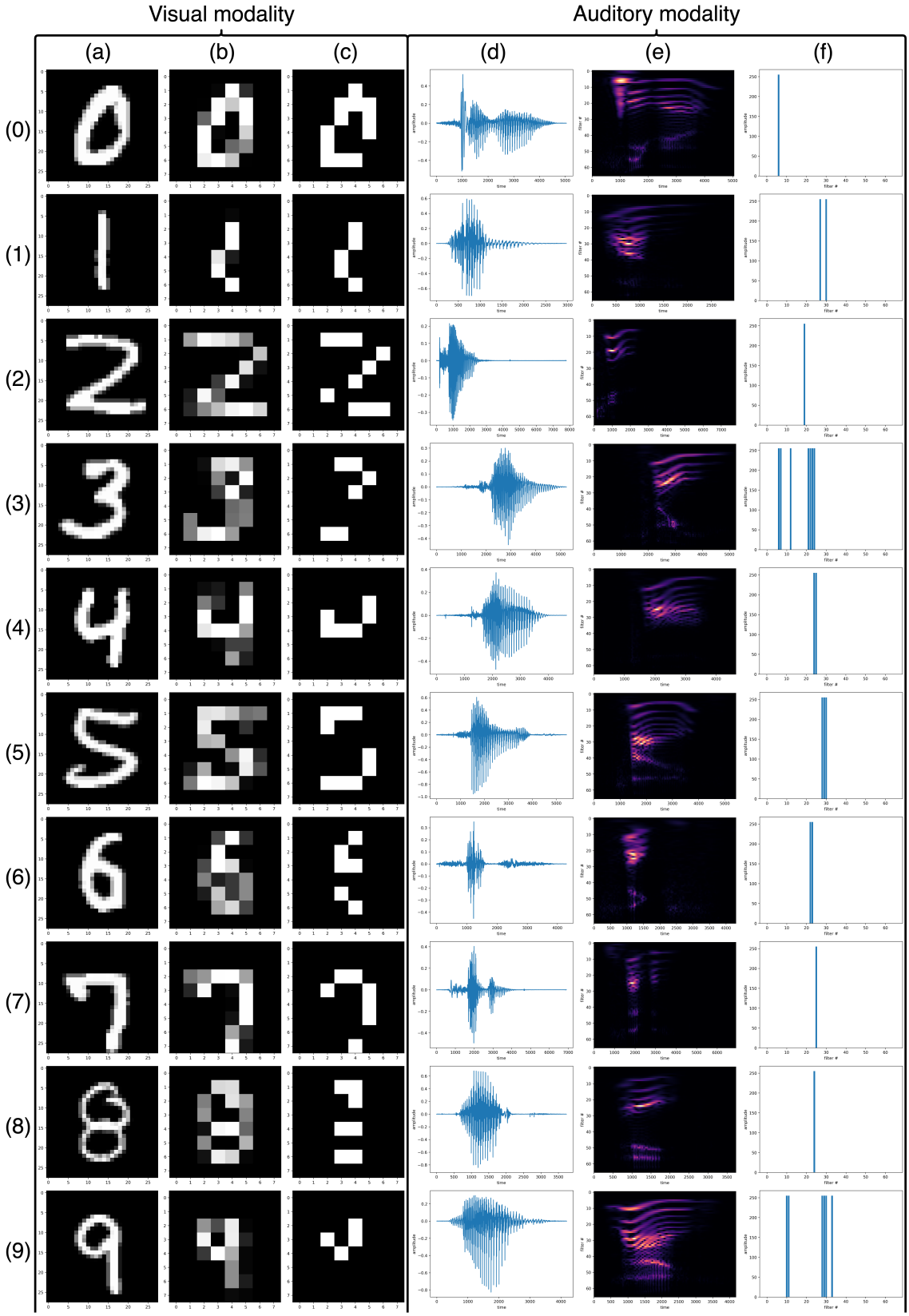


Figure B.7: The audio-visual database composed of ten digits (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) for visual and auditory modalities. (a) The original 28x28 images from the database. (b) The compressed 8x8 images. (c) The thresholded and binarized images. (d) The audio waveforms of the original sounds from the database. (e) The cochleagrams of the sounds using 66 filters. (f) The cochleagrams averaged over time, thresholded and binarized.

B.2.2 Bimodal learning

This appendix reproduces the learning experiment of Sec. 4.3.1 but this time considering bimodal learning. Indeed, unlike previously, during the learning phase, the sensory inputs from the visual and auditory modalities are presented simultaneously, as depicted in Fig. B.8 (a) and (b). In the same way, the corresponding local inhibitory labels are stimulated at the same time for both modalities.

Despite this difference, we obtained very similar results to those obtained previously, on the structure obtained after learning (panel (d)) and on its long-term maintenance (panels (e) and (f)). The major difference lies in the presence of connections between sensory areas. In fact, since features neurons and inhibitory neurons of the two modalities coding the same digit are stimulated at the same moment, their connections are strengthened, creating this direct "bridge" between modalities. This consequently compromises the segregation of modalities stated notably in the IIT and respected in previous learning approach.

For this reason, we will not present other results concerning this architecture in detail here. However, some preliminary experiments have been carried out to quantify the recognition and generation capacity of such an architecture. The performances appeared not to be better (or sometimes slightly worse). Furthermore, unimodal recognition is this case indirectly impossible since the presentation of sensory input of one modality directly activates the sensory input of the other modality (due to the direct connections between them) and makes the global labels reacting to both modalities. This might be an advantage given the direct generation from one modality to the other, but again it can question on the efficiency of these direct relations in comparison with a clear segregation and integration at a higher level via hubs.

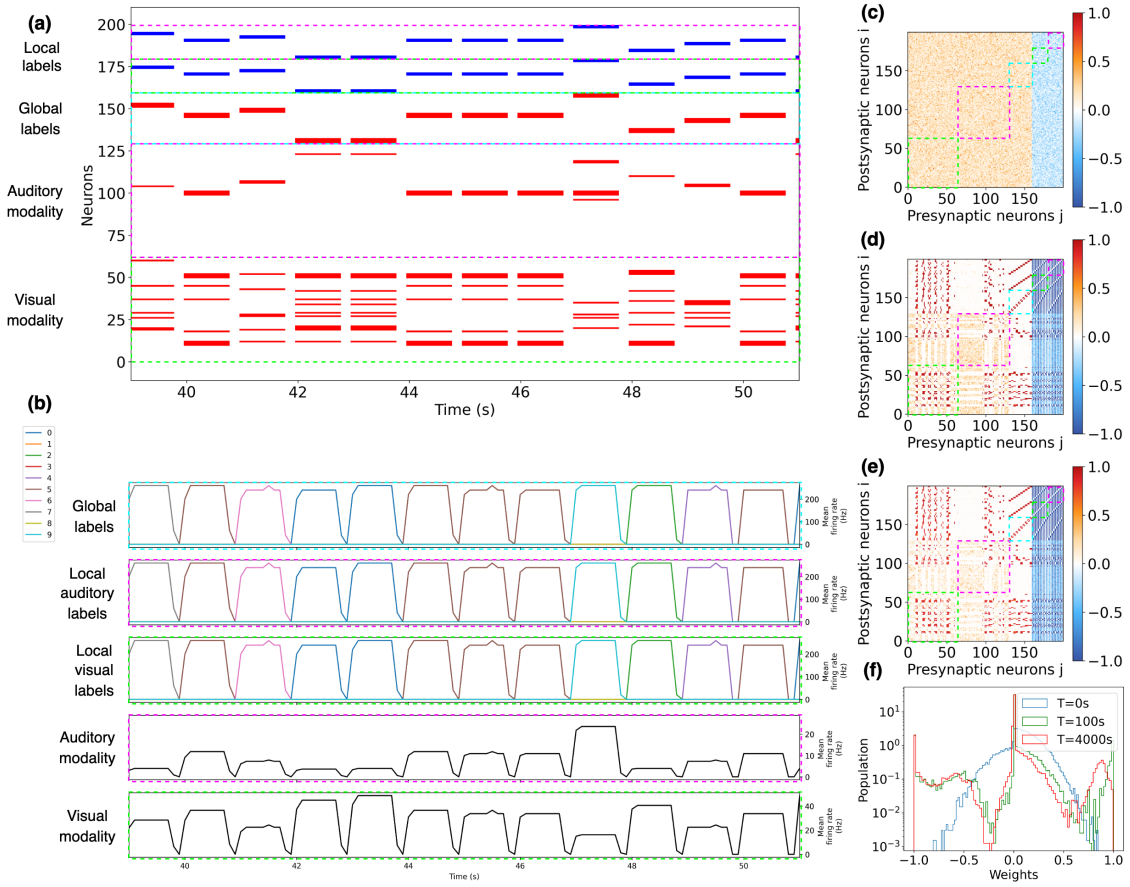


Figure B.8: Simulation of a neural network of 200 QIF neurons simultaneously learning bimodal digits. Neurons 0-63 code visual information, neurons 64-129 code auditory information, neurons 130-159 code global labels (3 consecutive neurons per label), neurons 160-179 code local labels for visual modality (2 consecutive neurons per label) and neurons 180-199 code local labels for auditory modality (2 consecutive neurons per label). The green dashed rectangles highlight the visual areas, the magenta dashed rectangles highlight the auditory areas and the cyan dashed rectangles highlight the global label area. (a) The spike raster plot during the learning phase with excitatory neurons in red, inhibitory in blue. (b) Mean firing rates of the different areas of the network during the learning phase. The colours in the label areas represent the groups of neurons associated to each digit label. (c) Weights matrix at the beginning of the simulation, before learning. (d) Weights matrix after learning (T=100s). (e) Weights matrix at the end of the simulation (T=4000s). (f) Distributions (in log scale) of the weights in the network at the different instant of the simulation (before learning t=0s (in blue), after learning T=100s (in green), at the end of the simulation T=4000s (in red)).

B.2.3 Mismatch modality recognition

This last appendix presents the results obtained in the mismatch modality recognition experiment. It corresponds to the case where the input stimuli of the two given modalities (visual and auditory) are not congruent. Each confusion matrix displays the activities of the global labels neurons in reaction to the varying auditory stimuli, while the visual stimuli are fixed to a particular value in each matrix, as shown in Fig. B.9. In this way, all combinations of bimodal stimuli are represented. The first thing we observe is that when the two modalities are congruent, as in Fig. 4.4 of the main text, we logically obtain correct recognition again (the associated global labels respond to the stimuli). In the case of non-congruent information, the decision is much more complex.

Indeed, in some cases, the visual modality seems to take precedence over the auditory modality (higher activity for the global label associated with the visual stimuli) and reversely in some cases. This is characterised visually by greater activity in general on the diagonals of the matrices (for auditory modality) and on the lines of the corresponding visual stimuli in the matrices (for visual modality). We speculate that this depends on the recognition degree of the individual modalities on each digit and on the contributions of each feature in this process. There is therefore a kind of competition between them, where too much uncertainty in both modalities leads to no response from the labels neurons or to completely wrong decisions. On the contrary, too much certainty in both modalities leads to the incapacity to make a decision with the same activation of two global labels. Ultimately, this mismatch recognition task makes the system generally inefficient, as humans are in face of incoherent information, creating a kind of cognitive dissonance. Thus, without another top-down mechanism such as attention, allowing to inhibit parasite information, such an architecture remains only subject to external cues with no sustainable decision making capacity.

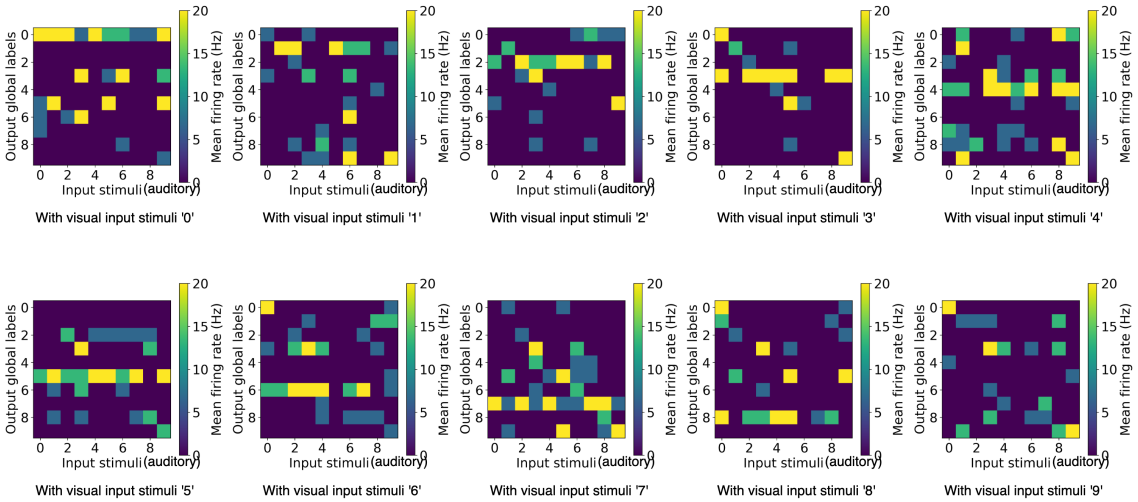


Figure B.9: Mismatch modality recognition. Each confusion matrix displays the activities of the global labels neurons in response to different auditory stimuli (1, 2, 3, 4, 5, 6, 7 , 8, 9). The visual stimulus is fixed to a particular digit in each matrix (digit 0 for the first matrix, digit 1 for the second matrix...).

Bibliography

[1] Petilla terminology: nomenclature of features of gabaergic interneurons of the cerebral cortex. *Nature Reviews Neuroscience*, 9(7):557–568, 2008. (Cited on page 56.)

[2] Moshe Abeles. *Corticonics: Neural circuits of the cerebral cortex*. Cambridge University Press, 1991. (Cited on page 54.)

[3] Priyanka A Abhang, Bharti Gawali, and Suresh C Mehrotra. *Introduction to EEG- and speech-based emotion recognition*. Academic Press, 2016. (Cited on page 57.)

[4] Juan A Acebrón, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of modern physics*, 77(1):137, 2005. (Cited on pages 11 and 14.)

[5] Robert Adler. A study of locking phenomena in oscillators. *Proceedings of the IRE*, 34(6):351–357, 1946. (Cited on page 11.)

[6] Google AI. Bard, 2023. (Cited on page 1.)

[7] Alan Eric Akil, Robert Rosenbaum, and Krešimir Josić. Balanced networks under spike-time dependent plasticity. *PLoS computational biology*, 17(5):e1008958, 2021. (Cited on pages 32 and 54.)

[8] Daniel J Amit and Daniel J Amit. *Modeling brain function: The world of attractor neural networks*. Cambridge university press, 1989. (Cited on page 12.)

[9] Daniel J Amit and Nicolas Brunel. Global spontaneous activity and local structured (learned) delay period activity in cortex. *Cerebral Cortex*, 7:237–252, 1997. (Cited on page 54.)

[10] Nancy C Andreasen, Daniel S O'Leary, Sergio Paradiso, Ted Cizadlo, Stephan Arndt, G Leonard Watkins, Laura L Boles Ponto, and Richard D Hichwa. The cerebellum plays a role in conscious episodic memory retrieval. *Human brain mapping*, 8(4):226–234, 1999. (Cited on page 2.)

[11] Louis Annnabi, Alexandre Pitti, and Mathias Quoy. A predictive coding account for chaotic itinerancy. In *Artificial Neural Networks and Machine Learning–ICANN 2021: 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 14–17, 2021, Proceedings, Part I 30*, pages 581–592. Springer, 2021. (Cited on page 109.)

[12] Takaaki Aoki and Toshio Aoyagi. Co-evolution of phases and connection strengths in a network of phase oscillators. *Physical review letters*, 102(3):034101, 2009. (Cited on pages 11, 14, 16, 28, 32, 34 and 44.)

[13] Takaaki Aoki and Toshio Aoyagi. Self-organized network of phase oscillators coupled by activity-dependent interactions. *Physical Review E*, 84(6):066109, 2011. (Cited on pages 11, 14, 16, 28, 32, 34 and 44.)

[14] Salva Ardid, Xiao-Jing Wang, and Albert Compte. An integrated microcircuit model of attentional processing in the neocortex. *Journal of Neuroscience*, 27(32):8486–8495, 2007. (Cited on page 82.)

[15] Alex Arenas and Albert Diaz-Guilera. Synchronization and modularity in complex networks. *The European Physical Journal Special Topics*, 143(1):19–25, 2007. (Cited on page 12.)

[16] Alex Arenas, Albert Díaz-Guilera, Jurgen Kurths, Yamir Moreno, and Changsong Zhou. Synchronization in complex networks. *Physics reports*, 469(3):93–153, 2008. (Cited on page 12.)

[17] Alex Arenas, Albert Diaz-Guilera, and Conrad J Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Physical review letters*, 96(11):114102, 2006. (Cited on page 12.)

[18] Alex Arenas, Alberto Fernández, and Sergio Gómez. A complex network approach to the determination of functional groups in the neural system of c. elegans. In *Workshop on Bio-Inspired Design of Networks*, pages 9–18. Springer, 2007. (Cited on page 12.)

[19] Jorge L Armony and Joseph E LeDoux. How the brain processes emotional information. 1997. (Cited on page 82.)

[20] Bassam V Atallah and Massimo Scanziani. Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron*, 62(4):566–577, 2009. (Cited on page 101.)

[21] Bernard J Baars. In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. *Journal of consciousness Studies*, 4(4):292–309, 1997. (Cited on page 2.)

[22] Bernard J Baars. The conscious access hypothesis: origins and recent evidence. *Trends in cognitive sciences*, 6(1):47–52, 2002. (Cited on page 2.)

[23] Bernard J Baars, Stan Franklin, and Thomas Zoega Ramsoy. Global workspace dynamics: cortical "binding and propagation" enables conscious contents. *Frontiers in psychology*, 4:200, 2013. (Cited on page 2.)

[24] Claudio Babiloni, Claudio Del Percio, Roberta Lizio, Giuseppe Noce, Susanna Lopez, Andrea Soricelli, Raffaele Ferri, Maria Teresa Pascarelli, Valentina Catania, Flavio Nobili, et al. Abnormalities of resting state cortical eeg rhythms in subjects with mild cognitive impairment due to alzheimer's and lewy body diseases. *Journal of Alzheimer's Disease*, 62(1):247–268, 2018. (Cited on page 10.)

[25] Stephen A Baccus and Markus Meister. Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5):909–919, 2002. (Cited on page 35.)

[26] Markus Bär, Eckehard Schöll, and Alessandro Torcini. Synchronization and complex dynamics of oscillators with delayed pulse coupling. *Angewandte Chemie-International Edition*, 51(38):9489, 2012. (Cited on page 104.)

[27] Moshe Bar. A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of cognitive neuroscience*, 15(4):600–609, 2003. (Cited on page 82.)

[28] Moshe Bar. The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1235–1243, 2009. (Cited on page 54.)

[29] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: statistical mechanics and its applications*, 281(1-4):69–77, 2000. (Cited on page 116.)

[30] Dennis L Barbour and Edward M Callaway. Excitatory local connections of superficial neurons in rat auditory cortex. *Journal of Neuroscience*, 28(44):11174–11185, 2008. (Cited on page 36.)

[31] Robert J Barry, Adam R Clarke, and Stuart J Johnstone. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. qualitative and quantitative electroencephalography. *Clinical neurophysiology*, 114(2):171–183, 2003. (Cited on page 10.)

[32] Andre M Bastos, W Martin Usrey, Rick A Adams, George R Mangun, Pascal Fries, and Karl J Friston. Canonical microcircuits for predictive coding. *Neuron*, 76(4):695–711, 2012. (Cited on pages 1 and 97.)

[33] Andre Moraes Bastos, Julien Vezoli, Conrado Arturo Bosman, Jan-Mathijs Schoffelen, Robert Oostenveld, Jarrod Robert Dowdall, Peter De Weerd, Henry Kennedy, and Pascal Fries. Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, 85(2):390–401, 2015. (Cited on page 12.)

[34] Yehezkel Ben-Ari. Excitatory actions of gaba during development: the nature of the nurture. *Nature Reviews Neuroscience*, 3(9):728–739, 2002. (Cited on page 31.)

[35] Gabriel Béna and Dan FM Goodman. Extreme sparsity gives rise to functional specialization. *arXiv preprint arXiv:2106.02626*, 2021. (Cited on page 82.)

[36] Raphaël Bergoin, Sofiane Boucenna, Alexandre Pitti, and David Cohen. A developmental model of audio-visual attention (mava): A premise for bimodal language learning in infants and robots. (Cited on pages 94 and 111.)

[37] Raphaël Bergoin, Alessandro Torcini, Gustavo Deco, Mathias Quoy, and Gorka Zamora-Lopez. Emergence and consolidation of modular neuronal networks (assemblies) driven by learning to selective stimuli and spontaneous recalls. (Cited on page 7.)

[38] Raphaël Bergoin, Alessandro Torcini, Gustavo Deco, Mathias Quoy, and Gorka Zamora-Lopez. Inhibitory neurons control the consolidation of neural assemblies via adaptation to selective stimuli. *Scientific Reports*, 13(1):6949, 2023. (Cited on pages 7, 73 and 75.)

[39] Rico Berner, Jan Fialkowski, Dmitry Kasatkin, Vladimir Nekorkin, Serhiy Yanchuk, and Eckehard Schöll. Hierarchical frequency clusters in adaptive networks of phase oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(10):103134, 2019. (Cited on pages 14, 32, 34 and 51.)

[40] Rico Berner, Eckehard Scholl, and Serhiy Yanchuk. Multiclusters in networks of adaptively coupled phase oscillators. *SIAM Journal on Applied Dynamical Systems*, 18(4):2227–2266, 2019. (Cited on pages 14, 32, 34 and 44.)

[41] Rico Berner and Alessandro Torcini. *Patterns of synchrony in complex networks of adaptively coupled oscillators*. Technische Universitaet Berlin (Germany), 2020. (Cited on pages 11, 14, 15 and 19.)

[42] Maxwell A Bertolero, BT Thomas Yeo, and Mark D'Esposito. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, 112(49):E6798–E6807, 2015. (Cited on pages 3, 4 and 51.)

[43] Ruggero G Bettinardi, Gustavo Deco, Vasileios M Karlaftis, Tim J Van Hartevelt, Henrique M Fernandes, Zoe Kourtzi, Morten L Kringelbach, and Gorka Zamora-López. How structure sculpts function: unveiling the contribution of anatomical connectivity to the brain's spontaneous correlation structure. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(4):047409, 2017. (Cited on page 12.)

[44] Richard F Betzel, John D Medaglia, Lia Papadopoulos, Graham L Baum, Ruben C Gur, Raquel E Gur, David Roalf, Theodore D Satterthwaite, and Danielle S Bassett. The modular organization of human anatomical brain networks: Accounting for the cost of wiring. *Network Neuroscience*, 1(1):42–68, 2017. (Cited on page 51.)

[45] Prateek Bhansali and Jaijeet Roychowdhury. Gen-adler: The generalized adler's equation for injection locking analysis in oscillators. In *2009 Asia and South Pacific Design Automation Conference*, pages 522–527. IEEE, 2009. (Cited on page 11.)

[46] Guo-qiang Bi and Mu-ming Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of neuroscience*, 18(24):10464–10472, 1998. (Cited on pages 1, 15, 34 and 59.)

[47] Hongjie Bi, Matteo Di Volo, and Alessandro Torcini. Asynchronous and coherent dynamics in balanced excitatory-inhibitory spiking networks. *Frontiers in Systems Neuroscience*, page 135, 2021. (Cited on page 33.)

[48] Yazan N Billeh, Binghuang Cai, Sergey L Gratiy, Kael Dai, Ramakrishnan Iyer, Nathan W Gouwens, Reza Abbasi-Asl, Xiaoxuan Jia, Joshua H Siegle, Shawn R Olsen, et al. Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron*, 106(3):388–403, 2020. (Cited on page 111.)

[49] Tim VP Bliss and Graham L Collingridge. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, 361(6407):31–39, 1993. (Cited on page 54.)

[50] Stefano Boccaletti, Jürgen Kurths, Grigory Osipov, DL Valladares, and CS Zhou. The synchronization of chaotic systems. *Physics reports*, 366(1-2):1–101, 2002. (Cited on page 12.)

[51] Bruno Bontempi, Catherine Laurent-Demir, Claude Destrade, and Robert Jaffard. Time-dependent reorganization of brain circuitry underlying long-term memory storage. *Nature*, 400(6745):671–675, 1999. (Cited on page 53.)

[52] Christoph Börgers and Nancy Kopell. Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. *Neural computation*, 15(3):509–538, 2003. (Cited on pages 102 and 104.)

[53] Alexander Borst and Frédéric E Theunissen. Information theory and neural coding. *Nature neuroscience*, 2(11):947–957, 1999. (Cited on page 83.)

[54] Sofiane Boucenna and Raphaël Bergoin. Bimodal recognition of vowel facilitated by an attention mechanism. (Cited on page 82.)

[55] Gordon H Bower and Paul R Cohen. Emotional influences in memory and thinking: Data and theory. *Affect and cognition*, 1, 1982. (Cited on page 54.)

[56] Michael Breakspear, Stewart Heitmann, and Andreas Daffertshofer. Generative models of cortical oscillations: neurobiological implications of the kuramoto model. *Frontiers in human neuroscience*, 4:190, 2010. (Cited on page 13.)

[57] Naama Brenner, Steven P Strong, Roland Koberle, William Bialek, and R van Steveninck. Symbols and synergy in a neural code. *arXiv preprint physics/9902067*, 1999. (Cited on page 83.)

[58] Carlos D Brody, Adrián Hernández, Antonio Zainos, and Ranulfo Romo. Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral cortex*, 13(11):1196–1207, 2003. (Cited on page 105.)

[59] William E Brownell, Charles R Bader, Daniel Bertrand, and Yves De Ribaupierre. Evoked mechanical responses of isolated cochlear outer hair cells. *Science*, 227(4683):194–196, 1985. (Cited on page 82.)

[60] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8:183–208, 2000. (Cited on page 54.)

[61] Nicolas Brunel and Peter E Latham. Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural computation*, 15(10):2281–2306, 2003. (Cited on page 55.)

[62] Nicolas Brunel and Xiao-Jing Wang. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of computational neuroscience*, 11:63–85, 2001. (Cited on page 54.)

[63] C Buchel, JT Coull, and Karl J Friston. The predictive value of changes in effective connectivity for human learning. *Science*, 283(5407):1538–1541, 1999. (Cited on page 108.)

[64] Randy L Buckner, Jorge Sepulcre, Tanveer Talukdar, Fenna M Krienen, Hesheng Liu, Trey Hedden, Jessica R Andrews-Hanna, Reisa A Sperling, and Keith A Johnson. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to alzheimer's disease. *Journal of neuroscience*, 29(6):1860–1873, 2009. (Cited on page 4.)

[65] Michael A Buice and Jack D Cowan. Field-theoretic approach to fluctuation effects in neural networks. *Physical Review E*, 75(5):051919, 2007. (Cited on page 106.)

[66] Daniel A Butts, Chong Weng, Jianzhong Jin, Chun-I Yeh, Nicholas A Lesica, Jose-Manuel Alonso, and Garrett B Stanley. Temporal precision in the neural code and the timescales of natural vision. *Nature*, 449(7158):92–95, 2007. (Cited on page 83.)

[67] György Buzsáki. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–385, 2010. (Cited on page 102.)

[68] Gyorgy Buzsaki and Andreas Draguhn. Neuronal oscillations in cortical networks. *science*, 304(5679):1926–1929, 2004. (Cited on pages 9, 10 and 54.)

[69] György Buzsáki and Xiao-Jing Wang. Mechanisms of gamma oscillations. *Annual review of neuroscience*, 35:203–225, 2012. (Cited on pages 10, 12 and 102.)

[70] Ryan T Canolty, Erik Edwards, Sarang S Dalal, Maryam Soltani, Srikantan S Nagarajan, Heidi E Kirsch, Mitchel S Berger, Nicholas M Barbaro, and Robert T Knight. High gamma power is phase-locked to theta oscillations in human neocortex. *science*, 313(5793):1626–1628, 2006. (Cited on page 12.)

[71] Ryan T Canolty and Robert T Knight. The functional role of cross-frequency coupling. *Trends in cognitive sciences*, 14(11):506–515, 2010. (Cited on pages 9 and 12.)

[72] Cristiano Capone, Matteo Di Volo, Alberto Romagnoni, Maurizio Mattia, and Alain Destexhe. State-dependent mean-field formalism to model different activity states in conductance-based networks of spiking neurons. *Physical Review E*, 100(6):062413, 2019. (Cited on page 111.)

[73] Cristiano Capone, Elena Pastorelli, Bruno Golosio, and Pier Stanislao Paolucci. Sleep-like slow oscillations improve visual classification through synaptic homeostasis and memory association in a thalamo-cortical model. *Scientific reports*, 9(1):8990, 2019. (Cited on page 82.)

[74] Antonio Caputi, Sarah Melzer, Magdalena Michael, and Hannah Monyer. The long and short of gabaergic neurons. *Current opinion in neurobiology*, 23(2):179–186, 2013. (Cited on page 110.)

[75] Jessica A Cardin, Marie Carlén, Konstantinos Meletis, Ulf Knoblich, Feng Zhang, Karl Deisseroth, Li-Huei Tsai, and Christopher I Moore. Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, 459(7247):663–667, 2009. (Cited on page 102.)

[76] Kristofor D Carlson, Micah Richert, Nikil Dutt, and Jeffrey L Krichmar. Biologically plausible models of homeostasis and stdp: stability and learning in spiking neural networks. In *The 2013 international joint conference on neural networks (IJCNN)*, pages 1–8. IEEE, 2013. (Cited on page 59.)

[77] Mallory Carlu, Omar Chehab, Leonardo Dalla Porta, Damien Depannemaecker, Charlotte Héricé, Maciej Jedynak, E Köksal Ersöz, Paolo Muratore, Selma Souihel, Cristiano Capone, et al. A mean-field approach to the dynamics of networks of complex neurons, from nonlinear integrate-and-fire to hodgkin–huxley models. *Journal of neurophysiology*, 123(3):1042–1051, 2020. (Cited on page 111.)

[78] Luis Carrillo-Reid. Neuronal ensembles in memory processes. In *Seminars in Cell & Developmental Biology*. Elsevier, 2021. (Cited on page 107.)

[79] Matthew Chalk, Boris Gutkin, and Sophie Deneve. Neural oscillations as a signature of efficient coding in the presence of synaptic delays. *Elife*, 5:e13824, 2016. (Cited on page 83.)

[80] Jean-Pierre Changeux, Alexandros Goulas, and Claus C Hilgetag. A connectomic hypothesis for the hominization of the brain. *Cerebral cortex*, 31(5):2425–2449, 2021. (Cited on page 3.)

[81] MK Chawla, JF Guzowski, V Ramirez-Amaya, P Lipa, KL Hoffman, LK Marriott, PF Worley, BL McNaughton, and Carol A Barnes. Sparse, environmentally selective expression of arc rna in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus*, 15(5):579–586, 2005. (Cited on page 102.)

[82] Tsuhan Chen and Ram R Rao. Audio-visual integration in multimodal communication. *Proceedings of the IEEE*, 86(5):837–852, 1998. (Cited on page 82.)

[83] Chiayu Q Chiu, Gyorgy Lur, Thomas M Morse, Nicholas T Carnevale, Graham CR Ellis-Davies, and Michael J Higley. Compartmentalization of gabaergic inhibition by dendritic spines. *Science*, 340(6133):759–762, 2013. (Cited on page 87.)

[84] RY Cho, RO Konecky, and Cameron S Carter. Impairments in frontal cortical γ synchrony and cognitive control in schizophrenia. *Proceedings of the National Academy of Sciences*, 103(52):19878–19883, 2006. (Cited on page 10.)

[85] Dominique Chu and Huy Le Nguyen. Constraints on hebbian and stdp learned weights of a spiking neuron. *Neural Networks*, 135:192–200, 2021. (Cited on page 35.)

[86] Chiara Cirelli, Christina M Gutierrez, and Giulio Tononi. Extensive and divergent effects of sleep and wakefulness on brain gene expression. *Neuron*, 41(1):35–43, 2004. (Cited on page 105.)

[87] Claudia Clopath, Lars Büsing, Eleni Vasilaki, and Wulfram Gerstner. Connectivity reflects coding: a model of voltage-based stdp with homeostasis. *Nature neuroscience*, 13(3):344–352, 2010. (Cited on pages 28 and 107.)

[88] Jeff Clune, Jean-Baptiste Mouret, and Hod Lipson. The evolutionary origins of modularity. *Proceedings of the Royal Society b: Biological sciences*, 280(1755):20122863, 2013. (Cited on page 5.)

[89] Susan K Coltman, Joshua GA Cashaback, and Paul L Gribble. Both fast and slow learning processes contribute to savings following sensorimotor adaptation. *Journal of neurophysiology*, 121(4):1575–1583, 2019. (Cited on pages 35, 36 and 52.)

[90] Juliette Courson, Mathias Quoy, Yulia Timofeeva, and Thanos Manos. The role of onset location and edge resection in epileptic seizure propagation in mice brain: a computational study. In *NeuroFrance 2023*, 2023. (Cited on pages 109 and 111.)

[91] Fergus IM Craik and Robert S Lockhart. Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, 11(6):671–684, 1972. (Cited on page 54.)

[92] Francis Crick and Christof Koch. A framework for consciousness. *Nature neuroscience*, 6(2):119–126, 2003. (Cited on page 2.)

[93] Sharon M Crook, G Bard Ermentrout, and James M Bower. Spike frequency adaptation affects the synchronization properties of networks of cortical oscillators. *Neural Computation*, 10(4):837–854, 1998. (Cited on page 13.)

[94] Nicolas A Crossley, Andrea Mechelli, Jessica Scott, Francesco Carletti, Peter T Fox, Philip McGuire, and Edward T Bullmore. The hubs of the human connectome are

generally implicated in the anatomy of brain disorders. *Brain*, 137(8):2382–2395, 2014. (Cited on page 4.)

[95] Hiroaki Daido. Order function and macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators. *Progress of theoretical physics*, 88(6):1213–1218, 1992. (Cited on page 15.)

[96] Hiroaki Daido. Generic scaling at the onset of macroscopic mutual entrainment in limit-cycle oscillators with uniform all-to-all coupling. *Physical review letters*, 73(5):760, 1994. (Cited on page 15.)

[97] H Dale. Pharmacology and nerve endings. *Br Med J*, 2:1161–1163, 1934. (Cited on pages 29 and 32.)

[98] Peter Dallos. The active cochlea. *The journal of neuroscience*, 12(12):4575, 1992. (Cited on page 82.)

[99] Fabrizio Damicelli, Claus C Hilgetag, and Alexandros Goulas. Brain connectivity meets reservoir computing. *PLoS Computational Biology*, 18(11):e1010639, 2022. (Cited on page 111.)

[100] Fabrizio Damicelli, Claus C Hilgetag, Marc-Thorsten Hütt, and Arnaud Messé. Modular topology emerges from plasticity in a minimalistic excitable network model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(4):047406, 2017. (Cited on pages 3, 5, 51 and 104.)

[101] Fabrizio Damicelli, Claus C Hilgetag, Marc-Thorsten Hütt, and Arnaud Messé. Topological reinforcement as a principle of modularity emergence in brain networks. *Network Neuroscience*, 3(2):589–605, 2019. (Cited on page 5.)

[102] Jessica S Damoiseaux, SARB Rombouts, Frederik Barkhof, Philip Scheltens, Cornelis J Stam, Stephen M Smith, and Christian F Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, 103(37):13848–13853, 2006. (Cited on page 108.)

[103] James A D'amour and Robert C Froemke. Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex. *Neuron*, 86(2):514–528, 2015. (Cited on page 31.)

[104] Luigi De Gennaro, Cristina Marzano, Domenica Veniero, Fabio Moroni, Fabiana Fratello, Giuseppe Curcio, Michele Ferrara, Fabio Ferlazzo, Luana Novelli, Maria Concetta Pellicciari, et al. Neurophysiological correlates of sleepiness: a combined tms and eeg study. *Neuroimage*, 36(4):1277–1287, 2007. (Cited on page 57.)

[105] Isabel Dean, Nicol S Harper, and David McAlpine. Neural population coding of sound level adapts to stimulus statistics. *Nature neuroscience*, 8(12):1684–1689, 2005. (Cited on page 83.)

[106] Dominique Debanne, Beat H Gähwiler, and Scott M Thompson. Long-term synaptic plasticity between pairs of individual ca3 pyramidal cells in rat hippocampal slice cultures. *The Journal of physiology*, 507(Pt 1):237, 1998. (Cited on page 34.)

[107] Gustavo Deco and Viktor K Jirsa. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *Journal of Neuroscience*, 32(10):3366–3375, 2012. (Cited on page 70.)

[108] Gustavo Deco, Viktor K Jirsa, and Anthony R McIntosh. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12(1):43–56, 2011. (Cited on page 12.)

[109] Gustavo Deco, Viktor K Jirsa, Peter A Robinson, Michael Breakspear, and Karl Friston. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology*, 4(8):e1000092, 2008. (Cited on page 106.)

[110] Gustavo Deco, Morten L Kringelbach, Viktor K Jirsa, and Petra Ritter. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Scientific reports*, 7(1):1–14, 2017. (Cited on pages 10, 70 and 134.)

[111] Gustavo Deco, Adrián Ponce-Alvarez, Dante Mantini, Gian Luca Romani, Patric Hagmann, and Maurizio Corbetta. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33(27):11239–11252, 2013. (Cited on page 70.)

[112] Stanislas Dehaene, Jean-Pierre Changeux, Lionel Naccache, Jérôme Sackur, and Claire Sergent. Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences*, 10(5):204–211, 2006. (Cited on page 2.)

[113] Stanislas Dehaene and Lionel Naccache. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2):1–37, 2001. (Cited on page 2.)

[114] Athena Demertzi, Enzo Tagliazucchi, Stanislas Dehaene, Gustavo Deco, Pablo Barttfeld, Federico Raimondo, Charlotte Martial, Davinia Fernández-Espejo, Benjamin Rohaut, HU Voss, et al. Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Science advances*, 5(2):eaat7603, 2019. (Cited on page 2.)

[115] Eleni A Demetriou, Amit Lampit, Daniel S Quintana, Sharon L Naismith, Yun JC Song, Julia E Pye, Ian Hickie, and Adam J Guastella. Autism spectrum disorders: a meta-analysis of executive function. *Molecular psychiatry*, 23(5):1198–1204, 2018. (Cited on page 111.)

[116] Sophie Denève and Christian K Machens. Efficient codes and balanced networks. *Nature neuroscience*, 19(3):375–382, 2016. (Cited on page 83.)

[117] Li Deng. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine*, 29(6):141–142, 2012. (Cited on page 84.)

[118] Nicolas Deperrois and Michael Graupner. Short-term depression and long-term plasticity together tune sensitive range of synaptic plasticity. *PLoS computational biology*, 16(9):e1008265, 2020. (Cited on pages 35 and 52.)

[119] Alain Destexhe. Self-sustained asynchronous irregular states and up–down states in thalamic, cortical and thalamocortical networks of nonlinear integrate-and-fire neurons. *Journal of computational neuroscience*, 27:493–506, 2009. (Cited on page 54.)

[120] Federico Devalle, Alex Roxin, and Ernest Montbrió. Firing rate equations require a spike synchrony mechanism to correctly describe fast oscillations in inhibitory networks. *PLoS computational biology*, 13(12):e1005881, 2017. (Cited on pages 12 and 13.)

[121] Matteo Di Volo, Alberto Romagnoni, Cristiano Capone, and Alain Destexhe. Biologically realistic mean-field models of conductance-based networks of spiking neurons with adaptation. *Neural computation*, 31(4):653–680, 2019. (Cited on page 111.)

[122] Jeremy S Dittman, Anatol C Kreitzer, and Wade G Regehr. Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *Journal of Neuroscience*, 20(4):1374–1385, 2000. (Cited on pages 35 and 52.)

[123] Rodney J Douglas and Kevan AC Martin. Mapping the matrix: the ways of neocortex. *Neuron*, 56(2):226–238, 2007. (Cited on page 110.)

[124] Yadin Dudai. The restless engram: consolidations never end. *Annual review of neuroscience*, 35:227–247, 2012. (Cited on pages 54, 63, 87 and 102.)

[125] Gerald M Edelman. Neural darwinism: selection and reentrant signaling in higher brain function. *Neuron*, 10(2):115–125, 1993. (Cited on page 82.)

[126] Jean-Baptiste Eichenlaub, Alain Nicolas, Jérôme Daltrozzo, Jérôme Redouté, Nicolas Costes, and Perrine Ruby. Resting brain activity varies with dream recall frequency between subjects. *Neuropsychopharmacology*, 39(7):1594–1602, 2014. (Cited on page 105.)

[127] Russo Eleonora, Alessandro Treves, et al. The phase space of lateral thought. In *Advances in Cognitive Neurodynamics (III)*, pages 483–489. Springer-Verlag, 2013. (Cited on page 104.)

[128] Andreas K Engel and Pascal Fries. Beta-band oscillations—signalling the status quo? *Current opinion in neurobiology*, 20(2):156–165, 2010. (Cited on page 10.)

[129] Bard Ermentrout. An adaptive model for synchrony in the firefly pteroptyx malaccae. *Journal of Mathematical Biology*, 29(6):571–585, 1991. (Cited on page 11.)

[130] Bard Ermentrout. Type i membranes, phase resetting curves, and synchrony. *Neural computation*, 8(5):979–1001, 1996. (Cited on pages 11 and 33.)

[131] Bard Ermentrout, Matthew Pascal, and Boris Gutkin. The effects of spike frequency adaptation and negative feedback on the synchronization of neural oscillators. *Neural computation*, 13(6):1285–1310, 2001. (Cited on pages 11, 13 and 32.)

[132] G Bard Ermentrout and Nancy Kopell. Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM journal on applied mathematics*, 46(2):233–253, 1986. (Cited on pages 32 and 33.)

[133] G Bard Ermentrout and Nancy Kopell. Multiple pulse interactions and averaging in systems of coupled neural oscillators. *Journal of Mathematical Biology*, 29(3):195–217, 1991. (Cited on page 13.)

[134] Oxana Eschenko, Cesare Magri, Stefano Panzeri, and Susan J Sara. Noradrenergic neurons of the locus coeruleus are phase locked to cortical up-down states during sleep. *Cerebral Cortex*, 22(2):426–435, 2012. (Cited on page 105.)

[135] Christian W Eurich, Klaus Pawelzik, Udo Ernst, Andreas Thiel, Jack D Cowan, and John G Milton. Delay adaptation in the nervous system. *Neurocomputing*, 32:741–748, 2000. (Cited on page 34.)

[136] Daniel E Feldman. Timing-based ltp and ltd at vertical inputs to layer ii/iii pyramidal cells in rat barrel cortex. *Neuron*, 27(1):45–56, 2000. (Cited on page 34.)

[137] Daniel E Feldman. Synaptic mechanisms for plasticity in neocortex. *Annual review of neuroscience*, 32:33–55, 2009. (Cited on page 4.)

[138] Daniel J Felleman and David C Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 1(1):1–47, 1991. (Cited on page 82.)

[139] Robert Fettiplace and Anthony J Ricci. Adaptation in auditory hair cells. *Current opinion in neurobiology*, 13(4):446–451, 2003. (Cited on page 35.)

[140] LA Finelli, H Baumann, AA Borbély, and P Achermann. Dual electroencephalogram markers of human sleep homeostasis: correlation between theta activity in waking and slow-wave activity in sleep. *Neuroscience*, 101(3):523–529, 2000. (Cited on page 57.)

[141] Luca A Finelli, Alexander A Borbély, and Peter Achermann. Functional topography of the human nonrem sleep electroencephalogram. *European Journal of Neuroscience*, 13(12):2282–2290, 2001. (Cited on page 105.)

[142] Robert S Fisher, Carlos Acevedo, Alexis Arzimanoglou, Alicia Bogacz, J Helen Cross, Christian E Elger, Jerome Engel Jr, Lars Forsgren, Jacqueline A French, Mike

Glynn, et al. Ilae official report: a practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482, 2014. (Cited on pages 109 and 111.)

[143] Peter Földiak. Forming sparse representations by local anti-hebbian learning. *Biological cybernetics*, 64(2):165–170, 1990. (Cited on page 59.)

[144] Nicolas Fourcaud-Trocmé. Integrate and fire models, deterministic. In *Encyclopedia of Computational Neuroscience*, pages 1683–1689. Springer, 2022. (Cited on page 56.)

[145] Paul W Frankland and Bruno Bontempi. The organization of recent and remote memories. *Nature reviews neuroscience*, 6(2):119–130, 2005. (Cited on pages 53 and 106.)

[146] Pascal Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480, 2005. (Cited on pages 9, 12 and 54.)

[147] Pascal Fries. Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual review of neuroscience*, 32:209–224, 2009. (Cited on page 13.)

[148] Pascal Fries. Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235, 2015. (Cited on page 10.)

[149] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138, 2010. (Cited on page 70.)

[150] Robert C Froemke, Mu-ming Poo, and Yang Dan. Spike-timing-dependent synaptic plasticity depends on dendritic location. *Nature*, 434(7030):221–225, 2005. (Cited on page 35.)

[151] Christos Galanis and Andreas Vlachos. Hebbian and homeostatic synaptic plasticity—do alterations of one reflect enhancement of the other? *Frontiers in Cellular Neuroscience*, 14:50, 2020. (Cited on page 35.)

[152] Jason R Gerstner, John N Koberstein, Adam J Watson, Nikolai Zapero, Davide Risso, Terence P Speed, Marcos G Frank, and Lucia Peixoto. Removal of unwanted variation reveals novel patterns of gene expression linked to sleep homeostasis in murine cortex. *BMC genomics*, 17:377–387, 2016. (Cited on page 105.)

[153] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014. (Cited on page 33.)

[154] Marcus Ghosh, Gabriel Bena, Volker Bormuth, and Dan FM Goodman. Multimodal units fuse-then-accumulate evidence across channels. *bioRxiv*, pages 2023–07, 2023. (Cited on page 82.)

[155] Charles D Gilbert. Microcircuitry of the visual cortex. *Annual review of neuroscience*, 6(1):217–247, 1983. (Cited on page 81.)

[156] Charles D Gilbert and Wu Li. Adult visual cortical plasticity. *Neuron*, 75(2):250–264, 2012. (Cited on page 82.)

[157] Matthieu Gilson, Gustavo Deco, Karl J Friston, Patric Hagmann, Dante Mantini, Viviana Betti, Gian Luca Romani, and Maurizio Corbetta. Effective connectivity inferred from fmri transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. *Neuroimage*, 180:534–546, 2018. (Cited on page 3.)

[158] Matthieu Gilson and Tomoki Fukai. Stability versus neuronal specialization for stdp: long-tail weight distributions solve the dilemma. *PloS one*, 6(10):e25339, 2011. (Cited on page 60.)

[159] Matthieu Gilson, Nikos E Kouvaris, Gustavo Deco, Jean-François Mangin, Cyril Poupon, Sandrine Lefranc, Denis Rivière, and Gorka Zamora-López. Network analysis of whole-brain fmri dynamics: A new framework based on dynamic communicability. *NeuroImage*, 201:116007, 2019. (Cited on page 4.)

[160] Matthieu Gilson, Timothée Masquelier, and Etienne Hugues. Stdp allows fast rate-modulated coding with poisson-like spike trains. *PLoS computational biology*, 7(10):e1002231, 2011. (Cited on page 59.)

[161] Matthieu Gilson, Ruben Moreno-Bote, Adrián Ponce-Alvarez, Petra Ritter, and Gustavo Deco. Estimation of directed effective connectivity from fmri functional connectivity hints at asymmetries of cortical connectome. *PLoS computational biology*, 12(3):e1004762, 2016. (Cited on page 4.)

[162] Corinna Giorgi and Silvia Marinelli. Roles and transcriptional responses of inhibitory neurons in learning and memory. *Frontiers in Molecular Neuroscience*, 14:113, 2021. (Cited on pages 32 and 106.)

[163] Bruno Golosio, Chiara De Luca, Cristiano Capone, Elena Pastorelli, Giovanni Stegel, Gianmarco Tiddia, Giulia De Bonis, and Pier Stanislao Paolucci. Thalamo-cortical spiking model of incremental learning combining perception, context and nrem-sleep. *PLoS Computational Biology*, 17(6):e1009045, 2021. (Cited on page 82.)

[164] Alexandros Goulas, Fabrizio Damicelli, and Claus C Hilgetag. Bio-instantiated recurrent neural networks: Integrating neurobiology-based network topology in artificial networks. *Neural Networks*, 142:608–618, 2021. (Cited on page 111.)

[165] Kalanit Grill-Spector and Rafael Malach. The human visual cortex. *Annu. Rev. Neurosci.*, 27:649–677, 2004. (Cited on page 81.)

[166] Yifan Gu and Pulin Gong. The dynamics of memory retrieval in hierarchical networks. *Journal of computational neuroscience*, 40(3):247–268, 2016. (Cited on pages 3, 5 and 54.)

[167] Bilal Haider, Alvaro Duque, Andrea R Hasenstaub, and David A McCormick. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *Journal of Neuroscience*, 26(17):4535–4545, 2006. (Cited on pages 101 and 102.)

[168] John E Hall and Michael E Hall. *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences, 2020. (Cited on page 82.)

[169] Seung Kee Han, Christian Kurrer, and Yoshiki Kuramoto. Dephasing and bursting in coupled neural oscillators. *Physical Review Letters*, 75(17):3190, 1995. (Cited on page 13.)

[170] D Hansel and G Mato. Existence and stability of persistent states in large neuronal networks. *Physical Review Letters*, 86(18):4175, 2001. (Cited on page 55.)

[171] David Hansel, German Mato, and Claude Meunier. Phase dynamics for weakly coupled hodgkin-huxley neurons. *Europhysics Letters*, 23(5):367, 1993. (Cited on page 103.)

[172] Oliver Hardt, Karim Nader, and Lynn Nadel. Decay happens: the role of active forgetting in memory. *Trends in cognitive sciences*, 17(3):111–120, 2013. (Cited on page 59.)

[173] Christopher D Harvey, Philip Coen, and David W Tank. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*, 484(7392):62–68, 2012. (Cited on page 105.)

[174] Donald O Hebb. The first stage of perception: growth of the assembly. *The Organization of Behavior*, 4:60–78, 1949. (Cited on pages 1, 20, 51 and 103.)

[175] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Science editions, 1949. (Cited on page 108.)

[176] Takao K Hensch. Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience*, 6(11):877–888, 2005. (Cited on page 4.)

[177] Christoph S Herrmann. Human eeg responses to 1–100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137:346–353, 2001. (Cited on page 57.)

[178] CS Herrmann and Tamer Demiralp. Human eeg gamma oscillations in neuropsychiatric disorders. *Clinical neurophysiology*, 116(12):2719–2733, 2005. (Cited on page 12.)

[179] Claus-C Hilgetag, Gully APC Burns, Marc A O'Neill, Jack W Scannell, and Malcolm P Young. Anatomical connectivity defines the organization of clusters of cortical areas in the macaque and the cat. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1393):91–110, 2000. (Cited on pages 3, 51 and 103.)

[180] Xavier Hinaut and Peter Ford Dominey. Real-time parallel processing of grammatical structure in the fronto-striatal system: A recurrent network simulation study using reservoir computing. *PloS one*, 8(2):e52946, 2013. (Cited on page 111.)

[181] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952. (Cited on page 55.)

[182] Sonja B Hofer, Ho Ko, Bruno Pichler, Joshua Vogelstein, Hana Ros, Hongkui Zeng, Ed Lein, Nicholas A Lesica, and Thomas D Mrsic-Flogel. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature neuroscience*, 14(8):1045–1052, 2011. (Cited on page 31.)

[183] Christopher J Honey, Olaf Sporns, Leila Cammoun, Xavier Gigandet, Jean-Philippe Thiran, Reto Meuli, and Patric Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040, 2009. (Cited on page 108.)

[184] Hyunsuk Hong and Steven H Strogatz. Conformists and contrarians in a kuramoto model with identical natural frequencies. *Physical Review E*, 84(4):046202, 2011. (Cited on page 13.)

[185] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982. (Cited on pages 32, 106 and 109.)

[186] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982. (Cited on page 107.)

[187] David H Hubel. *Eye, brain, and vision.* Scientific American Library/Scientific American Books, 1995. (Cited on page 82.)

[188] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962. (Cited on page 102.)

[189] David H Hubel and Torsten N Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1):215–243, 1968. (Cited on page 83.)

[190] David H Hubel and Torsten N Wiesel. Brain mechanisms of vision. *Scientific American*, 241(3):150–163, 1979. (Cited on page 81.)

[191] JM Hupé, AC James, BR Payne, SG Lomber, P Girard, and J Bullier. Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394(6695):784–787, 1998. (Cited on page 82.)

[192] Bruce Hutcheon and Yosef Yarom. Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends in neurosciences*, 23(5):216–222, 2000. (Cited on page 10.)

[193] Jeffry S Isaacson and Massimo Scanziani. How inhibition shapes cortical activity. *Neuron*, 72(2):231–243, 2011. (Cited on page 102.)

[194] Eugene M Izhikevich. *Dynamical systems in neuroscience*. MIT press, 2007. (Cited on page 12.)

[195] Zohar Jackson, César Souza, Jason Flaks, Yuxin Pan, Hereman Nicolas, and Adhish Thite. Jakobovski/free-spoken-digit-dataset: v1. 0.8. *Zenodo, August*, 2018. (Cited on page 84.)

[196] Ben H Jansen and Vincent G Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological cybernetics*, 73(4):357–366, 1995. (Cited on page 11.)

[197] Ole Jensen and Ali Mazaheri. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in human neuroscience*, 4:186, 2010. (Cited on page 10.)

[198] Ole Jensen and Claudia D Tesche. Frontal theta activity in humans increases with memory load in a working memory task. *European journal of Neuroscience*, 15(8):1395–1399, 2002. (Cited on page 9.)

[199] Xiaoxuan Jia and Adam Kohn. Gamma rhythms in the brain. *PLoS biology*, 9(4):e1001045, 2011. (Cited on page 10.)

[200] Yingyezhe Jin and Peng Li. Ap-stdp: A novel self-organizing mechanism for efficient reservoir computing. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1158–1165. IEEE, 2016. (Cited on page 60.)

[201] Marian Joëls, Zhenwei Pu, Olof Wiegert, Melly S Oitzl, and Harm J Krugers. Learning under stress: how does it work? *Trends in cognitive sciences*, 10(4):152–158, 2006. (Cited on page 54.)

[202] Heidi Johansen-Berg, TEJ Behrens, MD Robson, I Drobnjak, MFS Rushworth, JM Brady, SM Smith, DJ Higham, and PM Matthews. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences*, 101(36):13335–13340, 2004. (Cited on page 108.)

[203] Marcia K Johnson and Carol L Raye. Reality monitoring. *Psychological review*, 88(1):67, 1981. (Cited on page 54.)

[204] Jon H Kaas. Topographic maps are fundamental to sensory processing. *Brain research bulletin*, 44(2):107–112, 1997. (Cited on page 110.)

[205] Jon H Kaas and Christine E Collins. The organization of sensory cortex. *Current opinion in neurobiology*, 11(4):498–504, 2001. (Cited on page 110.)

[206] Jon H Kaas and Troy A Hackett. Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, 97(22):11793–11799, 2000. (Cited on page 82.)

[207] Eric R Kandel. The molecular biology of memory storage: a dialogue between genes and synapses. *Science*, 294(5544):1030–1038, 2001. (Cited on page 54.)

[208] Takeshi Kaneko. Local connections of excitatory neurons in motor-associated cortical areas of the rat. *Frontiers in neural circuits*, 7:75, 2013. (Cited on page 36.)

[209] Hiroyuki K Kato, Shea N Gillet, Andrew J Peters, Jeffry S Isaacson, and Takaki Komiyama. Parvalbumin-expressing interneurons linearly control olfactory bulb output. *Neuron*, 80(5):1218–1231, 2013. (Cited on page 87.)

[210] Larry C Katz and Carla J Shatz. Synaptic activity and the construction of cortical circuits. *Science*, 274(5290):1133–1138, 1996. (Cited on page 4.)

[211] Steffen Katzner, Ian Nauhaus, Andrea Benucci, Vincent Bonin, Dario L Ringach, and Matteo Carandini. Local origin of field potentials in visual cortex. *Neuron*, 61(1):35–41, 2009. (Cited on page 102.)

[212] Richard Kempter, Wulfram Gerstner, and J Leo Van Hemmen. Hebbian learning and spiking neurons. *Physical Review E*, 59(4):4498, 1999. (Cited on pages 28 and 59.)

[213] Richard Kempter, Wulfram Gerstner, and J Leo Van Hemmen. Spike-based compared to rate-based hebbian learning. *Advances in neural information processing systems*, 11:125–131, 1999. (Cited on pages 28, 32 and 54.)

[214] James M Kilner, Karl J Friston, and Chris D Frith. Predictive coding: an account of the mirror neuron system. *Cognitive processing*, 8:159–166, 2007. (Cited on page 97.)

[215] Florence I Kleberg, Tomoki Fukai, and Matthieu Gilson. Excitatory and inhibitory stdp jointly tune feedforward neural circuits to selectively propagate correlated spiking activity. *Frontiers in computational neuroscience*, 8:53, 2014. (Cited on pages 32, 37, 54, 56 and 59.)

[216] Wolfgang Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2-3):169–195, 1999. (Cited on pages 9 and 105.)

[217] Florian Klimm, Javier Borge-Holthoefer, Niels Wessel, Jürgen Kurths, and Gorka Zamora-López. Individual node's contribution to the mesoscale of complex networks. *New Journal of Physics*, 16(12):125006, 2014. (Cited on pages 4 and 51.)

[218] Jens G Klinzing, Niels Niethard, and Jan Born. Mechanisms of systems memory consolidation during sleep. *Nature neuroscience*, 22(10):1598–1610, 2019. (Cited on page 79.)

[219] Eric I Knudsen. Sensitive periods in the development of the brain and behavior. *Journal of cognitive neuroscience*, 16(8):1412–1425, 2004. (Cited on page 4.)

[220] Giacomo Koch, Viviana Ponzo, Francesco Di Lorenzo, Carlo Caltagirone, and Domenica Veniero. Hebbian and anti-hebbian spike-timing-dependent plasticity of human cortico-cortical connections. *Journal of Neuroscience*, 33(23):9725–9733, 2013. (Cited on page 59.)

[221] Helga Kolb, Eduardo Fernandez, and Ralph Nelson. The organization of the retina and visual system. *Webvision-the organization of the retina and visual system*, 2005. (Cited on page 81.)

[222] LD Kolibius, Frédéric Roux, George Parish, M Ter Wal, M Van Der Plas, Ramesh Chelvarajah, Vijay Sawlani, DT Rollings, Johannes Lang, Stephanie Gollwitzer, et al. Hippocampal neurons code individual episodic memories in humans. *bioRxiv*, pages 2021–06, 2021. (Cited on page 53.)

[223] Stephen M Kosslyn, William L Thompson, and Nathaniel M Alpert. Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage*, 6(4):320–334, 1997. (Cited on page 82.)

[224] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012. (Cited on page 1.)

[225] Yoshiki Kuramoto. International symposium on mathematical problems in theoretical physics. *Lecture notes in Physics*, 30:420, 1975. (Cited on pages 10, 14 and 133.)

[226] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975. (Cited on pages 10, 14, 33 and 133.)

[227] Yoshiki Kuramoto. Collective synchronization of pulse-coupled oscillators and excitable units. *Physica D: Nonlinear Phenomena*, 50(1):15–30, 1991. (Cited on page 13.)

[228] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence*. Courier Corporation, 2003. (Cited on pages 10, 14, 15 and 133.)

[229] Carlo R Laing. The dynamics of networks of identical theta neurons. *The Journal of Mathematical Neuroscience*, 8(1):1–24, 2018. (Cited on page 33.)

[230] Victor AF Lamme. Why visual attention and awareness are different. *Trends in cognitive sciences*, 7(1):12–18, 2003. (Cited on pages 2 and 82.)

[231] Karri Lamsa, Joost H Heeroma, and Dimitri M Kullmann. Hebbian ltp in feed-forward inhibitory interneurons and the temporal fidelity of input discrimination. *Nature neuroscience*, 8(7):916–924, 2005. (Cited on pages 56 and 59.)

[232] Karri P Lamsa, Joost H Heeroma, Peter Somogyi, Dmitri A Rusakov, and Dimitri M Kullmann. Anti-hebbian long-term potentiation in the hippocampal feedback inhibitory circuit. *Science*, 315(5816):1262–1266, 2007. (Cited on pages 56 and 59.)

[233] Lev D Landau. On the problem of turbulence. In *Dokl. Akad. Nauk USSR*, volume 44, page 311, 1944. (Cited on page 11.)

[234] Matthew Larkum. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends in neurosciences*, 36(3):141–151, 2013. (Cited on page 82.)

[235] Peter E Latham and Sheila Nirenberg. Synergy, redundancy, and independence in population codes, revisited. *Journal of Neuroscience*, 25(21):5195–5206, 2005. (Cited on page 83.)

[236] Peter E Latham, BJ Richmond, PG Nelson, and S Nirenberg. Intrinsic dynamics in neuronal networks. i. theory. *Journal of neurophysiology*, 83(2):808–827, 2000. (Cited on page 55.)

[237] Steven Laureys. The neural correlate of (un) awareness: lessons from the vegetative state. *Trends in cognitive sciences*, 9(12):556–559, 2005. (Cited on page 2.)

[238] Steven Laureys, Adrian M Owen, and Nicholas D Schiff. Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology*, 3(9):537–546, 2004. (Cited on page 2.)

[239] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015. (Cited on page 1.)

[240] Yann LeCun, Corinna Cortes, Chris Burges, et al. Mnist handwritten digit database, 2010. (Cited on page 84.)

[241] Seung-Hee Lee, Alex C Kwan, Siyu Zhang, Victoria Phoumthipphavong, John G Flannery, Sotiris C Masmanidis, Hiroki Taniguchi, Z Josh Huang, Feng Zhang, Edward S Boyden, et al. Activation of specific interneurons improves v1 feature selectivity and visual perception. *Nature*, 488(7411):379–383, 2012. (Cited on page 102.)

[242] Pascal Leimer, Michael Herzog, and Walter Senn. Synaptic weight decay with selective consolidation enables fast learning without catastrophic forgetting. *BioRxiv*, page 613265, 2019. (Cited on pages 36 and 104.)

[243] Huiyan Li, Xiaojuan Sun, and Jinghua Xiao. Degree of synchronization modulated by inhibitory neurons in clustered excitatory-inhibitory recurrent networks. *EPL (Europhysics Letters)*, 121(1):10003, 2018. (Cited on pages 13, 32, 37, 56 and 103.)

[244] Juliette Lozano-Goupil, Stéphane Raffard, Delphine Capdevielle, Emilie Aigoin, and Ludovic Marin. Gesture-speech synchrony in schizophrenia: A pilot study using a kinematic-acoustic analysis. *Neuropsychologia*, 174:108347, 2022. (Cited on page 111.)

[245] Leonhard Lücken, Oleksandr V Popovych, Peter A Tass, and Serhiy Yanchuk. Noise-enhanced coupling between two oscillators with long-term plasticity. *Physical Review E*, 93(3):032210, 2016. (Cited on page 34.)

[246] Tanushree B Luke, Ernest Barreto, and Paul So. Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons. *Neural computation*, 25(12):3207–3234, 2013. (Cited on page 33.)

[247] Yotam Luz and Maoz Shamir. Balancing feed-forward excitation and inhibition via hebbian inhibitory synaptic plasticity. *PLoS computational biology*, 8(1):e1002334, 2012. (Cited on page 59.)

[248] Kaiser M. Brain architecture: a design for natural computation. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 365(1861):3033–3045, 2007. (Cited on pages 3 and 51.)

[249] Wolfgang Maass. Networks of spiking neurons: the third generation of neural network models. *Neural networks*, 10(9):1659–1671, 1997. (Cited on pages 55 and 83.)

[250] Mojtaba Madadi Asl, Alireza Valizadeh, and Peter A Tass. Delay-induced multistability and loop formation in neuronal networks with spike-timing-dependent plasticity. *Scientific reports*, 8(1):1–15, 2018. (Cited on page 34.)

[251] Mojtaba Madadi Asl, Alireza Valizadeh, and Peter A Tass. Propagation delays determine neuronal activity and synaptic connectivity patterns emerging in plastic neuronal networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(10):106308, 2018. (Cited on page 34.)

[252] Nicolas Maingret, Gabrielle Girardeau, Ralitsa Todorova, Marie Goutierre, and Michaël Zugaro. Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nature neuroscience*, 19(7):959–964, 2016. (Cited on page 108.)

[253] Yuri L Maistrenko, Borys Lysyansky, Christian Hauptmann, Oleksandr Burylko, and Peter A Tass. Multistability in the kuramoto model with synaptic plasticity. *Physical Review E*, 75(6):066207, 2007. (Cited on pages 11 and 32.)

[254] Robert C Malenka and Mark F Bear. Ltp and ltd: an embarrassment of riches. *Neuron*, 44(1):5–21, 2004. (Cited on page 54.)

[255] Eve Marder and Jean-Marc Goaillard. Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*, 7(7):563–574, 2006. (Cited on page 106.)

[256] Nikola T Markov, Maria Ercsey-Ravasz, Camille Lamy, Ana Rita Ribeiro Gomes, Loïc Magrou, Pierre Misery, Pascale Giroud, Pascal Barone, Colette Dehay, Zoltán Toroczkai, et al. The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences*, 110(13):5187–5192, 2013. (Cited on page 4.)

[257] Henry Markram, Wulfram Gerstner, and Per Jesper Sjöström. Spike-timing-dependent plasticity: a comprehensive overview. *Frontiers in synaptic neuroscience*, 4:2, 2012. (Cited on page 34.)

[258] Cristina Marzano, Michele Ferrara, Federica Mauro, Fabio Moroni, Maurizio Gorgoni, Daniela Tempesta, Carlo Cipolli, and Luigi De Gennaro. Recalling and forgetting dreams: theta and alpha oscillations during sleep predict subsequent dream recall. *Journal of Neuroscience*, 31(18):6674–6683, 2011. (Cited on page 105.)

[259] Richard H Masland. The fundamental plan of the retina. *Nature neuroscience*, 4(9):877–886, 2001. (Cited on page 81.)

[260] Reece Mazade and Jose Manuel Alonso. Synergy in cortical networks. *Neuron*, 104(2):184–185, 2019. (Cited on page 83.)

[261] Mark A McDaniel and Gilles O Einstein. The importance of cue familiarity and cue distinctiveness in prospective memory. *Memory*, 1(1):23–41, 1993. (Cited on page 54.)

[262] James L McGaugh. Memory–a century of consolidation. *Science*, 287(5451):248–251, 2000. (Cited on pages 53 and 54.)

[263] James L McGaugh. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.*, 27:1–28, 2004. (Cited on page 54.)

[264] Michael M Merzenich and Koichi Sameshima. Cortical plasticity and memory. *Current opinion in neurobiology*, 3(2):187–196, 1993. (Cited on page 82.)

[265] M-Marsel Mesulam. From sensation to cognition. *Brain: a journal of neurology*, 121(6):1013–1052, 1998. (Cited on page 82.)

[266] David Meunier, Renaud Lambiotte, and Edward T Bullmore. Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, 4:200, 2010. (Cited on pages 3 and 51.)

[267] Midjourney. Midjourney. https://www.midjourney.com/, 2022. (Cited on page 1.)

[268] Christoph Miehl, Sebastian Onasch, Dylan Festa, and Julijana Gjorgjieva. Formation and computational implications of assemblies in neural circuits. *The Journal of Physiology*, 2022. (Cited on page 5.)

[269] Kaare Mikkelsen, Alberto Imparato, and Alessandro Torcini. Sisyphus effect in pulse-coupled excitatory neural networks with spike-timing-dependent plasticity. *Physical Review E*, 89(6):062701, 2014. (Cited on page 32.)

[270] Renato E Mirollo and Steven H Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990. (Cited on pages 11 and 12.)

[271] Bratislav Mišić and Olaf Sporns. From regions to connections and networks: new bridges between brain and behavior. *Current opinion in neurobiology*, 40:1–7, 2016. (Cited on page 4.)

[272] Gianluigi Mongillo, Omri Barak, and Misha Tsodyks. Synaptic theory of working memory. *Science*, 319(5869):1543–1546, 2008. (Cited on page 54.)

[273] Gianluigi Mongillo, Simon Rumpel, and Yonatan Loewenstein. Inhibitory connectivity defines the realm of excitatory plasticity. *Nature neuroscience*, 21(10):1463–1470, 2018. (Cited on pages 32, 36, 37, 49, 56 and 107.)

[274] Cyril Monier, Julien Fournier, and Yves Frégnac. In vitro and in vivo measures of evoked excitatory and inhibitory conductance dynamics in sensory cortices. *Journal of neuroscience methods*, 169(2):323–365, 2008. (Cited on page 36.)

[275] Sharlen Moore and Kishore V Kuchibhotla. Slow or sudden: re-interpreting the learning curve for modern systems neuroscience. *IBRO Neuroscience Reports*, 13:9–14, 2022. (Cited on pages 35 and 52.)

[276] Yamir Moreno and Amalio F Pacheco. Synchronization of kuramoto oscillators in scale-free networks. *Europhysics Letters*, 68(4):603, 2004. (Cited on page 12.)

[277] Vernon B Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, 20(4):408–434, 1957. (Cited on page 110.)

[278] Mark Müller-Linow, Claus C Hilgetag, and Marc-Thorsten Hütt. Organization of excitable dynamics in hierarchical biological networks. *PLoS computational biology*, 4(9):e1000190, 2008. (Cited on pages 3, 4, 51 and 103.)

[279] James M Murray and G Sean Escola. Remembrance of things practiced with fast and slow learning in cortical and subcortical pathways. *Nature Communications*, 11(1):6441, 2020. (Cited on pages 35 and 52.)

[280] John D Murray, Alberto Bernacchia, David J Freedman, Ranulfo Romo, Jonathan D Wallis, Xinying Cai, Camillo Padoa-Schioppa, Tatiana Pasternak, Hyojung Seo, Daeyeol Lee, et al. A hierarchy of intrinsic timescales across primate cortex. *Nature neuroscience*, 17(12):1661–1663, 2014. (Cited on page 106.)

[281] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003. (Cited on page 116.)

[282] Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y Ng. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 689–696, 2011. (Cited on page 82.)

[283] Kuniaki Noda, Yuki Yamaguchi, Kazuhiro Nakadai, Hiroshi G Okuno, and Tetsuya Ogata. Audio-visual speech recognition using deep learning. *Applied intelligence*, 42:722–737, 2015. (Cited on page 82.)

[284] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013. (Cited on page 34.)

[285] Michael Okun and Ilan Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature neuroscience*, 11(5):535–537, 2008. (Cited on pages 101 and 102.)

[286] Dennis DM O'Leary. Do cortical areas emerge from a protocortex? *Trends in neurosciences*, 12(10):400–406, 1989. (Cited on page 4.)

[287] Dennis DM O'Leary and Yasushi Nakagawa. Patterning centers, regulatory genes and extrinsic mechanisms controlling arealization of the neocortex. *Current opinion in neurobiology*, 12(1):14–25, 2002. (Cited on page 4.)

[288] OpenAI. Dall·e 2 pre-training mitigations. https://openai.com/research/dall-e-2-pre-training-mitigations, 2022. (Cited on page 1.)

[289] OpenAI. Openai: Introducing chatgpt. https://openai.com/blog/chatgpt, 2022. (Cited on page 1.)

[290] Jill X O'Reilly, Saad Jbabdi, Matthew FS Rushworth, and Timothy EJ Behrens. Brain systems for probabilistic and dynamic prediction: computational specificity and integration. *PLoS biology*, 11(9):e1001662, 2013. (Cited on page 105.)

[291] Delphine Oudiette, James W Antony, Jessica D Creery, and Ken A Paller. The role of memory reactivation during wakefulness and sleep in determining which memories endure. *Journal of Neuroscience*, 33(15):6672–6678, 2013. (Cited on page 79.)

[292] Dennis DM O'Leary and Setsuko Sahara. Genetic regulation of arealization of the neocortex. *Current opinion in neurobiology*, 18(1):90–100, 2008. (Cited on page 4.)

[293] LJ Paciorek. Injection locking of oscillators. *Proceedings of the IEEE*, 53(11):1723–1727, 1965. (Cited on page 11.)

[294] Cesare V Parise. Crossmodal correspondences: Standing issues and experimental guidelines. *Multisensory research*, 29(1-3):7–28, 2016. (Cited on page 82.)

[295] Hae-Jeong Park, Karl J Friston, Chongwon Pae, Bumhee Park, and Adeel Razi. Dynamic effective connectivity in resting state fmri. *NeuroImage*, 180:594–608, 2018. (Cited on page 108.)

[296] Aishwarya Parthasarathy, Roger Herikstad, Jit Hon Bong, Felipe Salvador Medina, Camilo Libedinsky, and Shih-Cheng Yen. Mixed selectivity morphs population codes in prefrontal cortex. *Nature neuroscience*, 20(12):1770–1779, 2017. (Cited on page 47.)

[297] Joel Pearson and Stephen M Kosslyn. The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the national academy of sciences*, 112(33):10089–10092, 2015. (Cited on page 82.)

[298] Yaël Perez, France Morin, and Jean-Claude Lacaille. A hebbian form of long-term potentiation dependent on mglur1a in hippocampal inhibitory interneurons. *Proceedings of the National Academy of Sciences*, 98(16):9401–9406, 2001. (Cited on page 59.)

[299] Leopoldo Petreanu, Tianyi Mao, Scott M Sternson, and Karel Svoboda. The subcellular organization of neocortical excitatory connections. *Nature*, 457(7233):1142–1145, 2009. (Cited on page 110.)

[300] James Pickles. An introduction to the physiology of hearing. In *An Introduction to the Physiology of Hearing*. Brill, 2013. (Cited on page 82.)

[301] Charlotte Piette, Jonathan Touboul, and Laurent Venance. Engrams of fast learning. *Frontiers in cellular neuroscience*, 14:575915, 2020. (Cited on pages 35 and 52.)

[302] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. Synchronization: a universal concept in nonlinear science, 2002. (Cited on pages 11 and 12.)

[303] Lucas Pinto and Yang Dan. Cell-type-specific activity in prefrontal cortex during goal-directed behavior. *Neuron*, 87(2):437–450, 2015. (Cited on pages 102 and 103.)

[304] Mark D Plumbley. Efficient information transfer and anti-hebbian neural networks. *Neural Networks*, 6(6):823–833, 1993. (Cited on page 59.)

[305] Alexandre Pouget, Peter Dayan, and Richard Zemel. Information processing with population codes. *Nature Reviews Neuroscience*, 1(2):125–132, 2000. (Cited on page 83.)

[306] Wim Pouw, Steven J Harrison, and James A Dixon. Gesture–speech physics: The biomechanical basis for the emergence of gesture–speech synchrony. *Journal of Experimental Psychology: General*, 149(2):391, 2020. (Cited on page 111.)

[307] Dale Purves, George J Augustine, David Fitzpatrick, William Hall, Anthony-Samuel LaMantia, and Leonard White. *Neurosciences*. De Boeck Supérieur, 2019. (Cited on page 82.)

[308] R Quian Quiroga, Gabriel Kreiman, Christof Koch, and Itzhak Fried. Sparse but not 'grandmother-cell'coding in the medial temporal lobe. *Trends in cognitive sciences*, 12(3):87–91, 2008. (Cited on page 83.)

[309] Pasko Rakic. Specification of cerebral cortical areas. *Science*, 241(4862):170, 1988. (Cited on page 4.)

[310] Pasko Rakic. Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews Neuroscience*, 10(10):724–735, 2009. (Cited on page 4.)

[311] Irmantas Ratas, Kestutis Pyragas, and Peter A Tass. Multistability in a star network of kuramoto-type oscillators with synaptic plasticity. *Scientific reports*, 11(1):1–15, 2021. (Cited on pages 11 and 12.)

[312] JP Rauschecker and RV Shannon. Sending sound to the brain. *Science*, 295(5557):1025–1029, 2002. (Cited on page 82.)

[313] Alfonso Renart, Jaime De La Rocha, Peter Bartho, Liad Hollender, Néstor Parga, Alex Reyes, and Kenneth D Harris. The asynchronous state in cortical circuits. *science*, 327(5965):587–590, 2010. (Cited on page 54.)

[314] Sidarta Ribeiro, Vikas Goyal, Claudio V Mello, and Constantine Pavlides. Brain gene expression during rem sleep depends on prior waking experience. *Learning & memory*, 6(5):500–508, 1999. (Cited on page 105.)

[315] Scott Rich, Michal Zochowski, and Victoria Booth. Dichotomous dynamics in ei networks with strongly and weakly intra-connected inhibitory neurons. *Frontiers in neural circuits*, 11:104, 2017. (Cited on page 104.)

[316] Fred Rieke, David Warland, Rob de Ruyter Van Steveninck, and William Bialek. *Spikes: exploring the neural code*. MIT press, 1999. (Cited on page 83.)

[317] Mattia Rigotti, Omri Barak, Melissa R Warden, Xiao-Jing Wang, Nathaniel D Daw, Earl K Miller, and Stefano Fusi. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590, 2013. (Cited on pages 47 and 106.)

[318] Mattia Rigotti, Omri Barak, Melissa R Warden, Xiao-Jing Wang, Nathaniel D Daw, Earl K Miller, and Stefano Fusi. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590, 2013. (Cited on page 105.)

[319] Henry L Roediger III and Melissa J Guynn. Retrieval processes. *Memory*, pages 197–236, 1996. (Cited on page 54.)

[320] Jamie D Roitman and Michael N Shadlen. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of neuroscience*, 22(21):9475–9489, 2002. (Cited on page 31.)

[321] Edmund T Rolls and Gustavo Deco. The noisy brain: stochastic dynamics as a principle of brain function. 2010. (Cited on page 106.)

[322] Gemma Rosell-Tarragó and Albert Díaz-Guilera. Functionability in complex networks: Leading nodes for the transition from structural to functional networks through remote asynchronization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(1):013105, 2020. (Cited on page 13.)

[323] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958. (Cited on page 55.)

[324] Nicolas P Rougier and Julien Vitay. Emergence of attention within a neural population. *Neural Networks*, 19(5):573–581, 2006. (Cited on page 111.)

[325] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986. (Cited on page 55.)

[326] Eleonora Russo and Alessandro Treves. Cortical free-association dynamics: Distinct phases of a latching network. *Physical Review E*, 85(5):051920, 2012. (Cited on pages 3 and 109.)

[327] Eleonora Russo and Alessandro Treves. The phase space of lateral thought. In *Advances in Cognitive Neurodynamics (III)*, pages 483–489. Springer, 2013. (Cited on page 3.)

[328] Paul Sauseng, Wolfgang Klimesch, Kirstin F Heise, Walter R Gruber, Elisa Holz, Ahmed A Karim, Mark Glennon, Christian Gerloff, Niels Birbaumer, and Friedhelm C Hummel. Brain oscillatory substrates of visual short-term memory capacity. *Current biology*, 19(21):1846–1852, 2009. (Cited on page 10.)

[329] Silvia Scarpetta. Critical behavior and memory function in a model of spiking neurons with a reservoir of spatio-temporal patterns. *The Functional Role of Critical Dynamics in Neural Systems*, pages 179–197, 2019. (Cited on page 59.)

[330] Daniel L Schacter and Donna Rose Addis. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481):773–786, 2007. (Cited on page 54.)

[331] Elad Schneidman, William Bialek, and Michael J Berry. Synergy, redundancy, and independence in population codes. *Journal of Neuroscience*, 23(37):11539–11553, 2003. (Cited on page 83.)

[332] Elad Schneidman, Jason L Puchalla, Ronen Segev, Robert A Harris, William Bialek, and Michael J Berry. Synergy from silence in a combinatorial neural code. *Journal of Neuroscience*, 31(44):15732–15741, 2011. (Cited on page 83.)

[333] Per B Sederberg, Michael J Kahana, Marc W Howard, Elizabeth J Donner, and Joseph R Madsen. Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, 23(34):10809–10814, 2003. (Cited on page 105.)

[334] Mario Senden, Gustavo Deco, Marcel A De Reus, Rainer Goebel, and Martijn P Van Den Heuvel. Rich club organization supports a diverse set of functional network configurations. *Neuroimage*, 96:174–182, 2014. (Cited on page 1.)

[335] Mario Senden, Niels Reuter, Martijn P van den Heuvel, Rainer Goebel, and Gustavo Deco. Cortical rich club regions can organize state-dependent functional network formation by engaging in oscillatory behavior. *NeuroImage*, 146:561–574, 2017. (Cited on page 4.)

[336] Hyojung Seo and Daeyeol Lee. Neural basis of learning and preference during social decision-making. *Current opinion in neurobiology*, 22(6):990–995, 2012. (Cited on page 106.)

[337] Michael N Shadlen and William T Newsome. Noise, neural codes and cortical organization. *Current opinion in neurobiology*, 4(4):569–579, 1994. (Cited on page 12.)

[338] Jafar Shamsi, María José Avedillo, Bernabe Linares-Barranco, and Teresa Serrano-Gotarredona. Oscillatory hebbian rule (ohr): An adaption of the hebbian rule to oscillatory neural networks. In *2020 XXXV Conference on Design of Circuits and Integrated Systems (DCIS)*, pages 1–6. IEEE, 2020. (Cited on page 34.)

[339] Murray Shanahan. A spiking neuron model of cortical broadcast and competition. *Consciousness and Cognition*, 17(1):288–303, 2008. (Cited on pages 2 and 3.)

[340] Murray Shanahan. *Embodiment and the inner life: Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press, USA, 2010. (Cited on pages 2, 3, 4, 105 and 106.)

[341] Tracey J Shors and Louis D Matzel. Long-term potentiation: what's learning got to do with it? *Behavioral and Brain Sciences*, 20(4):597–614, 1997. (Cited on page 82.)

[342] Yousheng Shu, Andrea Hasenstaub, and David A McCormick. Turning on and off recurrent balanced cortical activity. *Nature*, 423(6937):288–293, 2003. (Cited on page 102.)

[343] Markus Siegel, Tobias H Donner, Robert Oostenveld, Pascal Fries, and Andreas K Engel. Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron*, 60(4):709–719, 2008. (Cited on page 9.)

[344] Adam M Sillito and Helen E Jones. Corticothalamic interactions in the transfer of visual information. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1428):1739–1752, 2002. (Cited on page 82.)

[345] Per Jesper Sjöström, Gina G Turrigiano, and Sacha B Nelson. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–1164, 2001. (Cited on pages 34, 35 and 52.)

[346] J. Sjöström and W. Gerstner. Spike-timing dependent plasticity. *Scholarpedia*, 5(2):1362, 2010. revision #184913. (Cited on pages 36 and 60.)

[347] Per Sebastian Skardal, Dane Taylor, and Juan G Restrepo. Complex macroscopic behavior in systems of phase oscillators with adaptive coupling. *Physica D: Nonlinear Phenomena*, 267:27–35, 2014. (Cited on pages 11, 13 and 32.)

[348] Paul So, Tanushree B Luke, and Ernest Barreto. Networks of theta neurons with time-varying excitability: Macroscopic chaos, multistability, and final-state uncertainty. *Physica D: Nonlinear Phenomena*, 267:16–26, 2014. (Cited on page 33.)

[349] William R Softky and Christof Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of neuroscience*, 13(1):334–350, 1993. (Cited on page 83.)

[350] Sanming Song, Hongxun Yao, and Alessandro Treves. A modular latching chain. *Cognitive neurodynamics*, 8(1):37–46, 2014. (Cited on pages 51 and 109.)

[351] Andreas Spiegler, Thomas R Knösche, Karin Schwab, Jens Haueisen, and Fatihcan M Atay. Modeling brain resonance phenomena using a neural mass model. *PLoS computational biology*, 7(12):e1002298, 2011. (Cited on page 11.)

[352] Olaf Sporns. The human connectome: a complex network. *Annals of the new York Academy of Sciences*, 1224(1):109–125, 2011. (Cited on page 110.)

[353] Olaf Sporns. Network attributes for segregation and integration in the human brain. *Current opinion in neurobiology*, 23(2):162–171, 2013. (Cited on page 1.)

[354] Olaf Sporns and Rolf Kötter. Motifs in brain networks. *PLoS biology*, 2(11):e369, 2004. (Cited on page 13.)

[355] Michael W Spratling. A review of predictive coding algorithms. *Brain and cognition*, 112:92–97, 2017. (Cited on page 97.)

[356] Larry R Squire and Pablo Alvarez. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Current opinion in neurobiology*, 5(2):169–177, 1995. (Cited on pages 54, 63, 87 and 102.)

[357] Larry R Squire and Stuart Zola-Morgan. The medial temporal lobe memory system. *Science*, 253(5026):1380–1386, 1991. (Cited on pages 53 and 54.)

[358] Kevin J Staley and F Edward Dudek. Interictal spikes and epileptogenesis. *Epilepsy currents*, 6(6):199–202, 2006. (Cited on pages 109 and 111.)

[359] CJ Stam, ECW Van Straaten, E Van Dellen, P Tewarie, G Gong, A Hillebrand, J Meier, and P Van Mieghem. The relation between structural and functional connectivity patterns in complex brain networks. *International Journal of Psychophysiology*, 103:149–160, 2016. (Cited on page 4.)

[360] Cornelis J Stam and Jaap C Reijneveld. Graph theoretical analysis of complex networks in the brain. *Nonlinear biomedical physics*, 1:1–19, 2007. (Cited on page 4.)

[361] Roxana A Stefanescu and Viktor K Jirsa. A low dimensional description of globally coupled heterogeneous neural networks of excitatory and inhibitory neurons. *PLoS computational biology*, 4(11):e1000219, 2008. (Cited on pages 32, 55 and 103.)

[362] Federico Stella, Erika Cerasti, Bailu Si, Karel Jezek, and Alessandro Treves. Self-organization of multiple spatial and context memories in the hippocampus. *Neuroscience & Biobehavioral Reviews*, 36(7):1609–1625, 2012. (Cited on pages 3, 104 and 108.)

[363] Mircea Steriade. Impact of network activities on neuronal properties in corticothalamic systems. *Journal of neurophysiology*, 86(1):1–39, 2001. (Cited on page 54.)

[364] Mircea Steriade, Angel Nunez, and Florin Amzica. Intracellular analysis of relations between the slow ($<$ 1 hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *Journal of Neuroscience*, 13(8):3266–3283, 1993. (Cited on page 10.)

[365] Robert Stickgold. Sleep-dependent memory consolidation. *Nature*, 437(7063):1272–1278, 2005. (Cited on pages 54, 79 and 104.)

[366] Robert Stickgold and Matthew P Walker. Sleep-dependent memory consolidation and reconsolidation. *Sleep medicine*, 8(4):331–343, 2007. (Cited on page 79.)

[367] Steven H Strogatz. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1-4):1–20, 2000. (Cited on page 12.)

[368] John Trevor Stuart. On the non-linear mechanics of hydrodynamic stability. *Journal of Fluid Mechanics*, 4(1):1–21, 1958. (Cited on page 11.)

[369] John Trevor Stuart. On the non-linear mechanics of wave disturbances in stable and unstable parallel flows part 1. the basic behaviour in plane poiseuille flow. *Journal of Fluid Mechanics*, 9(3):353–370, 1960. (Cited on page 11.)

[370] Halgurd Taher, Alessandro Torcini, and Simona Olmi. Exact neural mass model for synaptic-based working memory. *PLOS Computational Biology*, 16(12):e1008533, 2020. (Cited on page 57.)

[371] Catherine Tallon-Baudry, Olivier Bertrand, Claude Delpuech, and Jacques Pernier. Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *Journal of Neuroscience*, 16(13):4240–4249, 1996. (Cited on page 57.)

[372] Christian Tetzlaff, Christoph Kolodziejski, Marc Timme, Misha Tsodyks, and Florentin Wörgötter. Synaptic scaling enables dynamically distinct short-and long-term memory formation. *PLoS Computational Biology*, 9(10):e1003307, 2013. (Cited on pages 35 and 52.)

[373] Panagiota Theodoni, Bernat Rovira, Yingxue Wang, and Alex Roxin. Theta-modulation drives the emergence of connectivity patterns underlying replay in a network model of place cells. *Elife*, 7:e37388, 2018. (Cited on pages 3, 5 and 54.)

[374] Paul Tiesinga and Terrence J Sejnowski. Cortical enlightenment: are attentional gamma oscillations driven by ing or ping? *Neuron*, 63(6):727–732, 2009. (Cited on page 13.)

[375] Emmanuelle Tognoli and JA Scott Kelso. The metastable brain. *Neuron*, 81(1):35–48, 2014. (Cited on page 70.)

[376] Giulio Tononi. An information integration theory of consciousness. *BMC neuroscience*, 5:1–22, 2004. (Cited on page 3.)

[377] Giulio Tononi. Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, 215(3):216–242, 2008. (Cited on page 3.)

[378] Giulio Tononi, Melanie Boly, Marcello Massimini, and Christof Koch. Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7):450–461, 2016. (Cited on pages 3 and 109.)

[379] Melody Torao-Angosto, Arnau Manasanch, Maurizio Mattia, and Maria V Sanchez-Vives. Up and down states during slow oscillations in slow-wave sleep and different levels of anesthesia. *Frontiers in systems neuroscience*, 15:609645, 2021. (Cited on page 105.)

[380] Jonathan Touboul, Fabrice Wendling, Patrick Chauvel, and Olivier Faugeras. Neural mass activity, bifurcations, and epilepsy. *Neural computation*, 23(12):3232–3286, 2011. (Cited on page 11.)

[381] Alessandro Treves. Frontal latching networks: a possible neural basis for infinite recursion. *Cognitive neuropsychology*, 22(3-4):276–291, 2005. (Cited on page 109.)

[382] Julia Trommershauser, Konrad Kording, and Michael S Landy. *Sensory cue integration*. Oxford University Press, 2011. (Cited on page 82.)

[383] Misha V Tsodyks and Terrance Sejnowski. Rapid state switching in balanced cortical network models. *Network: Computation in Neural Systems*, 6(2):111, 1995. (Cited on page 102.)

[384] Endel Tulving and Donald M Thomson. Encoding specificity and retrieval processes in episodic memory. *Psychological review*, 80(5):352, 1973. (Cited on page 54.)

[385] Gina G Turrigiano and Sacha B Nelson. Homeostatic plasticity in the developing nervous system. *Nature reviews neuroscience*, 5(2):97–107, 2004. (Cited on page 101.)

[386] Peter J Uhlhaas and Wolf Singer. Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *neuron*, 52(1):155–168, 2006. (Cited on page 13.)

[387] Peter J Uhlhaas and Wolf Singer. Abnormal neural oscillations and synchrony in schizophrenia. *Nature reviews neuroscience*, 11(2):100–113, 2010. (Cited on page 10.)

[388] Leslie G Ungerleider. Two cortical visual systems. *Analysis of visual behavior*, 549:chapter–18, 1982. (Cited on page 82.)

[389] Eugenio Urdapilleta, Bailu Si, and Alessandro Treves. S elforganization of modular activity of grid cells. *Hippocampus*, 27(11):1204–1213, 2017. (Cited on pages 5 and 51.)

[390] Martijn P Van Den Heuvel and Olaf Sporns. Rich-club organization of the human connectome. *Journal of Neuroscience*, 31(44):15775–15786, 2011. (Cited on pages 4, 105 and 108.)

[391] Martijn P van den Heuvel and Olaf Sporns. Rich-club organization of the human connectome. *The journal of neuroscience*, 31(44):15775–15786, 2011. (Cited on page 106.)

[392] Guy Van der Sande, Daniel Brunner, and Miguel C Soriano. Advances in photonic reservoir computing. *Nanophotonics*, 6(3):561–576, 2017. (Cited on page 111.)

[393] Eelco V Van Dongen, Atsuko Takashima, Markus Barth, Jascha Zapp, Lothar R Schad, Ken A Paller, and Guillén Fernández. Memory stabilization with targeted reactivation during human slow-wave sleep. *Proceedings of the National Academy of Sciences*, 109(26):10575–10580, 2012. (Cited on page 79.)

[394] Mark CW Van Rossum, Maria Shippi, and Adam B Barrett. Soft-bound synaptic plasticity increases storage capacity. *PLoS computational biology*, 8(12):e1002836, 2012. (Cited on page 60.)

[395] Carl Van Vreeswijk and Haim Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726, 1996. (Cited on page 54.)

[396] Lav R Varshney, Beth L Chen, Eric Paniagua, David H Hall, and Dmitri B Chklovskii. Structural properties of the caenorhabditis elegans neuronal network. *PLoS computational biology*, 7(2):e1001066, 2011. (Cited on page 51.)

[397] Janne C Visser, Nanda NJ Rommelse, Corina U Greven, and Jan K Buitelaar. Autism spectrum disorder and attention-deficit/hyperactivity disorder in early childhood: A review of unique and shared characteristics and developmental antecedents. *Neuroscience & Biobehavioral Reviews*, 65:229–263, 2016. (Cited on page 111.)

[398] Nathalia Vitureira and Yukiko Goda. The interplay between hebbian and homeostatic synaptic plasticity. *Journal of Cell Biology*, 203(2):175–186, 2013. (Cited on page 35.)

[399] Simon Vock, Rico Berner, Serhiy Yanchuk, and Eckehard Schöll. Effect of diluted connectivities on cluster synchronization of adaptively coupled oscillator networks. *arXiv preprint arXiv:2101.05601*, 2021. (Cited on page 12.)

[400] Tim P Vogels and Larry F Abbott. Signal propagation and logic gating in networks of integrate-and-fire neurons. *Journal of neuroscience*, 25(46):10786–10795, 2005. (Cited on page 54.)

[401] Tim P Vogels, Henning Sprekeler, Friedemann Zenke, Claudia Clopath, and Wulfram Gerstner. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*, 334(6062):1569–1573, 2011. (Cited on pages 32 and 59.)

[402] Tim P Vogels, Henning Sprekeler, Friedemann Zenke, Claudia Clopath, and Wulfram Gerstner. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*, 334(6062):1569–1573, 2011. (Cited on page 107.)

[403] Jack W, Blakemore C, and Malcolm P Young. Analysis of connectivity in the cat cerebral cortex. *The journal of neuroscience*, 15(2):1463–1483, 1995. (Cited on pages 3, 51 and 103.)

[404] Jack W and Malcolm P Young. The connectional organization of neural systems in the cat cerebral cortex. *Current biology*, 3(4):191–200, 1993. (Cited on page 3.)

[405] Brian A Wandell. *Foundations of vision.* sinauer Associates, 1995. (Cited on page 81.)

[406] Jing-Yi Wang, Katharina L Heck, Jan Born, Hong-Viet V Ngo, and Susanne Diekelmann. No difference between slow oscillation up-and down-state cueing for memory consolidation during sleep. *Journal of Sleep Research*, 31(6):e13562, 2022. (Cited on page 105.)

[407] Xiao-Jing Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3):1195–1268, 2010. (Cited on pages 10, 12 and 103.)

[408] Xiao-Jing Wang and György Buzsáki. Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of neuroscience*, 16(20):6402–6413, 1996. (Cited on page 13.)

[409] Fabrice Wendling, Jean-Jacques Bellanger, Fabrice Bartolomei, and Patrick Chauvel. Relevance of nonlinear lumped-parameter models in the analysis of depth-eeg epileptic signals. *Biological cybernetics*, 83(4):367–378, 2000. (Cited on page 11.)

[410] Miles A Whittington, Roger D Traub, and John GR Jefferys. Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature*, 373(6515):612–615, 1995. (Cited on pages 10 and 102.)

[411] Miles A Whittington, Roger D Traub, Nancy Kopell, Bard Ermentrout, and Eberhard H Buhl. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *International journal of psychophysiology*, 38(3):315–336, 2000. (Cited on page 31.)

[412] Arthur T Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, 16(1):15–42, 1967. (Cited on page 11.)

[413] John T Wixted. The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.*, 55:235–269, 2004. (Cited on page 59.)

[414] Hui Wu and Mukesh Dhamala. Dynamics of kuramoto oscillators with time-delayed positive and negative couplings. *Physical Review E*, 98(3):032221, 2018. (Cited on page 13.)

[415] Yuan-Jie Xiao, Lidan Wang, Yu-Zhang Liu, Jiayu Chen, Haoyu Zhang, Yan Gao, Hua He, Zheng Zhao, and Zhiru Wang. Excitatory crossmodal input to a widespread population of primary sensory cortical neurons. *Neuroscience Bulletin*, pages 1–14, 2022. (Cited on page 36.)

[416] Mingshan Xue, Bassam V Atallah, and Massimo Scanziani. Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature*, 511(7511):596–600, 2014. (Cited on page 101.)

[417] Toshiyuki Yamane, Yasunao Katayama, Ryosho Nakane, Gouhei Tanaka, and Daiju Nakano. Wave-based reservoir computing by synchronization of coupled oscillators. In *Neural Information Processing: 22nd International Conference, ICONIP 2015, Istanbul, Turkey, November 9-12, 2015, Proceedings Part III 22*, pages 198–205. Springer, 2015. (Cited on page 12.)

[418] Ben P Yuhas, Moise H Goldstein, Terence J Sejnowski, and Robert E Jenkins. Neural network models of sensory integration for improved vowel recognition. *Proceedings of the IEEE*, 78(10):1658–1668, 1990. (Cited on page 82.)

[419] Gorka Zamora-López, Yuhan Chen, Gustavo Deco, Morten L Kringelbach, and Changsong Zhou. Functional complexity emerging from anatomical constraints in the brain: the significance of network modularity and rich-clubs. *Scientific reports*, 6(1):1–18, 2016. (Cited on pages 1, 12 and 82.)

[420] Gorka Zamora-López, Eleonora Russo, Pablo M Gleiser, Changsong Zhou, and Jürgen Kurths. Characterizing the complexity of brain and mind networks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1952):3730–3747, 2011. (Cited on pages 3 and 109.)

[421] Gorka Zamora-López, Changsong Zhou, and Jürgen Kurths. Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Frontiers in neuroinformatics*, 4:1, 2010. (Cited on pages 3, 4, 89, 103, 105, 106 and 116.)

[422] Gorka Zamora-López, Changsong Zhou, and Jürgen Kurths. Exploring brain function from anatomical connectivity. *Frontiers in neuroscience*, 5:83, 2011. (Cited on pages 1, 3, 4, 51, 103, 105 and 106.)

[423] Robert J Zatorre, Pascal Belin, and Virginia B Penhune. Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences*, 6(1):37–46, 2002. (Cited on page 82.)

[424] Friedemann Zenke, Everton J Agnes, and Wulfram Gerstner. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature communications*, 6(1):1–13, 2015. (Cited on pages 35 and 107.)

[425] Lei Zhang, Shengyuan Zhou, Tian Zhi, Zidong Du, and Yunji Chen. Tdsnn: From deep neural networks to deep spike neural networks with temporal-coding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1319–1326, 2019. (Cited on pages 55 and 83.)

[426] Siyu Zhang, Min Xu, Tsukasa Kamigaki, Johnny Phong Hoang Do, Wei-Cheng Chang, Sean Jenvay, Kazunari Miyamichi, Liqun Luo, and Yang Dan. Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345(6197):660–665, 2014. (Cited on page 110.)

[427] Rüdiger Zillmer, Roberto Livi, Antonio Politi, and Alessandro Torcini. Stability of the splay state in pulse-coupled networks. *Physical Review E*, 76(4):046102, 2007. (Cited on page 47.)