Multi-Scale Spatiotemporal Neural Computation: On The Relationship Between Dynamical Attractors, Spiking Neural Networks, And Convolutional Neural Circuits

by

Eduardo Diniz

B.S., University of Brasıilia, 2015

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Eduardo Diniz

It was defended on

September 1st 2023

and approved by

Zhi-Hong Mao, Professor, Department of Electrical and Computer Engineering

Ahmed Dallal, Associate Professor, Department of Electrical and Computer Engineering

Amro El-Jaroudi, Associate Professor, Department of Electrical and Computer Engineering

Dissertation Advisors: Zhi-Hong Mao, Professor, Department of Electrical and Computer Engineering,

Howard J. Aizenstein, Professor, Department of Psychiatry

Tamer S. Ibrahim, Professor, Department of Bioengineering

Mingui Sun, Professor, Department of Neurological Surgery

Multi-Scale Spatiotemporal Neural Computation: On The Relationship Between Dynamical Attractors, Spiking Neural Networks, And Convolutional Neural Circuits

Eduardo Diniz, PhD

University of Pittsburgh, 2023

Understanding spatiotemporal neural dynamics and developing biologically-inspired artificial neural networks remain open challenges in computational neuroscience. Critical gaps persist in elucidating cortical rhythms, memory consolidation, and biological networks' remarkable spatiotemporal processing capabilities. We hypothesize that asymmetric connectivity and dedicated fast-slow processing pathways in neural systems enhance depth, robustness, and versatility in handling complex spatiotemporal patterns. Our first contribution is elucidating how neurons communicate and synchronize activity via temporally precise spikes by examining the dynamics of spike-coding networks. Developing models of cortical neural oscillators reveal the origins of spontaneous transitions between active and silent states underlying slow-wave sleep rhythms, demonstrating how the balance of excitation and inhibition orchestrates these oscillations. Our second contribution is establishing a mathematical equivalence between spike-coding and Hopfield networks by showing that fast and slow asymmetric connectivity induces equivalent cyclic attractor dynamics in both systems. The asymmetric weights in slow connections enable both models to generate complex temporal firing sequences, transitioning between quasi-attractor states representing stored memories. Simulations demonstrate the efficacy of spike-coding networks for encoding and retrieving temporal sequences while performing the n-back working memory task. Our third contribution is to harness generative adversarial networks for unpaired cross-modality translation from 3 Tesla to 7 Tesla magnetic resonance imaging. We propose a fast-slow convolutional network architecture to enhance translation performance by balancing local and global information processing. This dissertation makes significant contributions by elucidating brain mechanisms underlying rhythms and memory and unifying foundational computational frameworks while extracting principles to improve artificial neural network design.

Table of Contents

List of Figures

Preface

"Henceforth, space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."

— Hermann Minkowski, 1908

Prof. Zhi-Hong Mao has been an unwavering beacon of support, and I owe him my profound thanks for his invaluable guidance. No wonder anyone who crosses his path is instantly drawn to him; he is a paragon of humility, embodying a lifelong commitment to learning and literacy. His wisdom, patience, and unparalleled kindness truly set him apart. In an academic landscape that often champions knowledge above all, Prof. Mao reminds us of the enduring value of genuine human connection and understanding. He doesn't just teach; he is a true role model in academia and life.

To Dr. Tamer Ibrahim and Dr. Howard Aizenstein, I am deeply grateful. Dr. I, with his pioneering work in MRI and dedication to pushing the boundaries of what is possible in biomedical engineering, has been a source of motivation. On the other hand, Howard, renowned for his profound contributions to neuroimaging and work in the interface of aging and the brain, has been nothing short of a guiding star. His passion for understanding the complexities of the human brain and his patient mentorship have enriched my research experience in ways words cannot capture. Together, they co-advised and steered the MRI and deep learning aspects of my dissertation.

The "Science Without Borders" program by CNPq, the Brazilian government's funding agency, is a visionary program that has not only been instrumental in my academic journey but has also facilitated the aspirations of countless students. The initiative reflects an understanding of the universality of science and the belief that knowledge and innovation are magnified when shared across cultures and nations. Through their support during the foundational years of my academic journey, CNPq has not only invested in my future but has reaffirmed its dedication to advancing global scientific discourse. I am profoundly grateful for the opportunity they granted me, and I hold immense respect for the impact of the

"Science Without Borders" program on the broader scientific community.

To my beloved family in Brazil—to my mother, who, amidst all adversities, staunchly advocated for our education. Her tenacity and vision enabled me to earn a bachelor's degree and subsequently tread the path of academia. This journey was later mirrored by my sister, who, too, is on her path to achieving a Ph.D., standing as a testament to our mother's indefatigable spirit and belief in the transformative power of education. My sister's dedication to our family has illustrated that offering unwavering support is possible without compromising one's individuality. Her genuine behavior and steadfastness have served as a powerful lesson in preserving one's identity while being a pillar of strength for others. My father, a man of natural talent, taught himself the intricacies of electronic devices without ever stepping into a college. His ability to master complexities purely out of passion and need has always been an inspiring testament to what dedication and self-belief can accomplish. Their combined sacrifices have etched an enduring legacy for our family.

To Danielle, my pillar of strength. Through the tumultuous waves of life, your support remained, even when I was pulled under. Your love and faith, especially in the final stretch of this degree, have been the guiding light that led me through the darkest of nights. And to Tino, a source of unparalleled comfort. My love for him has revealed profound layers of care I never realized I held. Together, you have been the heart of my journey.

1.0 Introduction

Neural Computation, encompassing the intricacies of brain dynamics and artificial neural networks (ANNs), is a nexus of mathematical, computational, and biological disciplines. The multifaceted nature of this interdisciplinary domain necessitates an understanding of spatiotemporal patterns, mechanisms of neural dynamics, and practical application in computational models. By synergistically combining insights from these spheres, we aspire to bridge the existing chasm between biological and artificial intelligence (AI).

In the forthcoming sections, we delve deeper into the intricacies of biological and artificial neural systems, exploring their unique strengths, inherent challenges, and potential synergy avenues. We begin with a detailed examination of spatiotemporal patterns, fast and slow processing pathways, and asymmetric connectivity—critical constructs spanning both realms. Throughout this dissertation, we will dissect the state-of-the-art methodologies in ANNs, the complexity of biological neural dynamics, and the promising frontier where these two domains intersect.

1.1 General Background

Understanding biological neural networks' dynamics and computational capabilities remains a fundamental challenge in neuroscience. Advanced theoretical modeling and analytical techniques from dynamical systems theory and mathematical biology have provided key insights into the emergence of complex neural phenomena (Dayan and Abbott, 2005). For instance, the origins of coordinated firing patterns, macroscale brain rhythms, and high-order functions like memory consolidation have been illuminated through detailed models of spike timing neural coordination (Boerlin et al., 2013), coupled neural oscillators (Breakspear, 2017), and attractor networks (Hopfield, 1982).

In parallel, ANNs have rapidly advanced machine learning and AI systems, drawing inspiration from neuroscience (Hassabis et al., 2017). AI models like convolutional neural

networks (CNNs) (LeCun et al., 1998), recurrent neural networks (RNNs) (Lipton et al., 2015), and spiking neural networks (Maass, 1997) have achieved state-of-the-art results across diverse tasks from computer vision to natural language processing. Especially noteworthy are generative adversarial networks (GANs) (Goodfellow et al., 2014), which can learn to mimic complex high-dimensional distributions like medical images (Yi et al., 2019).

However, gaps remain between our understanding of biological computation and the capabilities of artificial models (Marblestone et al., 2016). The characterization and understanding of spatiotemporal patterns present an intriguing nexus where the realms of biological and ANNs converge and offer synergistic opportunities. The inherent complexity of these patterns, with their intricate interplay of spatial structures and temporal dynamics, is reminiscent of the challenges faced in delineating biological neural dynamics and the computational undertakings of their artificial counterparts.

1.2 Spatiotemporal Patterns in Biological and Artificial Systems

Spatiotemporal patterns are ubiquitous phenomena in the natural world, characterized by correlated changes in space and time. Think of the rhythmic dance of starlings in murmuration, the orchestrated progression of cells during embryonic development, or the rhythmic waves along a shoreline. These patterns inherently capture a sequence of events or states, evolving across space over a specific duration. They are pivotal in diverse fields, from physics and chemistry to ecology and social sciences. Such patterns offer a lens to understand complex systems' behavior, where the spatial arrangement and temporal sequence of elements are interdependent and jointly influence the observed outcome. Their study provides insights into the mechanisms underlying these phenomena and informs strategies to predict, and at times control, their future evolution.

1.2.1 Historical Context

The fascination with spatiotemporal dynamics in biology can be traced back to the early observations of coordinated behaviors in flocking birds, schooling fish, and synchronized firefly flashes. Researchers were captivated by how individual organisms, seemingly acting independently, could give rise to such large-scale coordinated behaviors. Turing's seminal work in the 1950s on morphogenesis, where he proposed reaction-diffusion systems to explain pattern formation in animal coats (Turing, 1952), laid the foundation for understanding how simple interactions at a local level could manifest into complex spatial patterns. Subsequent work in neuroscience sought to understand the brain's spatiotemporal patterns, from the synchronized firing of neurons to the propagation of waves across neural populations.

In parallel, cellular automata became an instrumental model in the 1940s and 50s of Computer Science, with pioneers like John von Neumann and later Stephen Wolfram exploring how simple rules could lead to complex spatiotemporal evolutions (Wolfram, 1983). The desire to emulate and understand biological systems led to the birth of ANNs in the late 1950s and 60s, with Rosenblatt's perceptron being one of the earliest examples (Rosenblatt, 1958). As computers became more powerful, the complexity of these systems grew, leading to the development of architectures specifically designed to handle spatiotemporal data.

The interplay between biology and computation became more pronounced as the decades passed. Biological insights drove innovations in machine learning, while computational models provided tools and frameworks to dissect and predict biological phenomena. Today, the study of spatiotemporal patterns stands at this vibrant intersection, enriched by a legacy of interdisciplinary collaboration and mutual inspiration.

1.2.2 Artificial Networks and Their Limitations

While advanced artificial networks like CNNs, RNNs, and GANs excel in recognizing and generating patterns across spatial and temporal domains, they often lack the depth of encoding, robust persistence, and the nuanced hierarchies that biological systems exhibit when processing spatiotemporal patterns (Lake et al., 2017).

CNNs, for instance, have been extensively employed in image and video processing tasks.

Their layered architecture mimics, to some extent, the hierarchies observed in the visual system, with earlier layers capturing low-level features like edges and contours and deeper layers representing high-level abstractions and semantic entities (Krizhevsky et al., 2012). However, the static nature of CNNs, with fixed feedforward connections, contrasts with the adaptability and feedback mechanisms inherent in biological systems.

RNNs, on the other hand, are designed to capture temporal dependencies in sequences, making them suitable for tasks such as natural language processing and time-series prediction (Hochreiter and Schmidhuber, 1997). The mechanism by which they store memory, though inspired by the temporal dynamics of biological systems, is oversimplified. In biological neurons, the history of activations is stored in a complex interplay of ionic currents, synaptic strengths, and dendritic computations (Magee and Johnston, 2005). In contrast, in RNNs, this memory is condensed into state vectors updated through matrix multiplications.

GANs are a recent addition to the family of neural networks, designed for generating new instances that resemble a given set of training data (Goodfellow et al., 2014). Their capability to synthesize intricate patterns is a testament to their power in representing spatiotemporal data. However, while GANs can produce novel patterns, their representation of them is more brittle than the plastic, context-aware representation in biological networks.

1.2.3 Biological Neural Systems: Hierarchy, Compositionality, and Storage

Biological neural systems naturally handle these patterns with an exceptional hierarchy and compositionality, as observed in the visual system, where neurons progressively extract higher-level features (Hubel and Wiesel, 1962). Similarly, compositionality is a crucial brain property that allows the encoding of complex spatiotemporal patterns, as observed in language processing (Fodor and Pylyshyn, 1988).

Such complex spatiotemporal patterns are encoded in synaptic weights and neuronal parameters. One prominent mechanism is spike-timing-dependent plasticity (STDP). In this process, synapses are strengthened or weakened based on the relative timing of pre- and post-synaptic spikes (Bi and Poo, 1998). This is a critical component of pattern completion, where brief input cues can reconstruct the entire stored pattern (O'Reilly and Rudy, 2001).

Moreover, neural circuits can dynamically transition and replay these patterns. Replay is particularly important in the consolidation of memory, where the brain replays sequences of neuron firing patterns observed during waking hours (Ji and Wilson, 2007). As for dynamic transitioning, it is speculated to occur via the activation of separate populations of neurons or changes in the firing pattern within a single population (Harris et al., 2003).

Reliable persistence without corruption over time is another critical aspect of these patterns. Biological systems showcase a resilience to adversarial attacks and noisy inputs, stemming from their evolutionary history, that contemporary artificial networks often struggle with (Szegedy et al., 2013). This resilience in biological systems is hypothesized to result from the intricate balance between excitation and inhibition, local and global network dynamics, and short-term and long-term plasticity mechanisms (Buzsaki and Draguhn, 2004).

Lastly, storing many such patterns and their flexible expression is paramount. This aspect is strongly linked to neuronal multistability, where a single neuron or neuronal network can exhibit various stable states (Izhikevich, 2007). This property allows the storage of multiple patterns and their flexible retrieval upon receiving specific inputs.

1.3 Asymmetric Connectivity Motifs in Biological Networks

Asymmetry in connectivity is pivotal in orchestrating the complex spatiotemporal processing capabilities observed in these networks (Sporns, 2011). In essence, asymmetric connectivity is not merely a random occurrence but is evolutionarily optimized for sophisticated information processing within the intricate structures of neural networks (Song et al., 2005).

Within cortical hierarchies, feedforward and feedback pathways introduce distinct forms of asymmetry, serving critical functions in information processing. Feedforward pathways enable the hierarchical progression of feature extraction, while feedback mechanisms provide top-down modulation, enhancing the specificity and contextuality of perceptions (Felleman and Van Essen, 1991). This organization ensures an intricate balance between sensory stimuli's detailed representation and contextual information integration, vital for a coherent perceptual experience (Rao and Ballard, 1999).

Inhibitory interneurons are quintessential for imposing asymmetric regulation on excitatory networks. By their very nature, these interneurons counter excitation, ensuring that the network's dynamics are stabilized and avoid runaway activity (Isaacson and Scanziani, 2011). This is fundamental in maintaining the homeostasis of neural circuits, ensuring the appropriate balance between excitation and inhibition. This is crucial for efficient information processing and preventing pathological activity (Vogels et al., 2011).

The hippocampal formation, critical for memory processes, features an asymmetric sequential connectivity pattern among its subregions: the dentate gyrus, CA3, CA1, and the subiculum. This supports the transformation and integration of information, with each subregion contributing uniquely to memory encoding, consolidation, and retrieval (Amaral and Witter, 1989). This organization is instrumental for temporal association and pattern separation, foundational processes for episodic memory formation (Teyler and Rudy, 2007).

Neuromodulatory systems, encompassing components like acetylcholine, dopamine, and serotonin, can diffusely project and exert asymmetric gain control over their target networks (Marder, 2012). This means that neuromodulators can dynamically adjust the sensitivity and responsiveness of neuronal populations, thereby tuning network dynamics based on the context or behavioral state. Such asymmetric modulatory influences play a crucial role in processes ranging from learning and attention to mood regulation (Dayan and Yu, 2006).

The biological neural landscape is replete with asymmetric architectural motifs that collectively define the multifaceted dynamics of information processing. Through feedforward sweeps of activity, feedback contextual modulation, pulsed inhibitory controls, and neuromodulatory influences, these asymmetries together shape the complex spatiotemporal patterns of activity. However, perturbations in these asymmetric interactions can manifest as failures in coordinating distributed processing and disruptions in temporal sequencing, often underlying neurological and psychiatric conditions (Bressler and Menon, 2010). Hence, understanding these asymmetric motifs can offer profound insights into the intricacies of biological systems, providing avenues for understanding their depth, robustness, and adaptability in spatiotemporal encoding.

1.4 Fast and Slow Pathways in Biological Systems

Biological neural networks' intricate tapestry of fast and slow processing pathways epitomizes versatility in spatiotemporal processing. Rapid encoding mechanisms ensure the preservation of transient, dynamic details. In parallel, slower pathways allow refining these encodings, embedding them within broader contexts. This integrated approach facilitates the emergence of highly sophisticated neural representations (Buzsaki and Draguhn, 2004).

Biological systems leverage the interplay between rapid feedforward sensory pathways and slower feedback mechanisms. The former are apt at capturing dynamic stimuli through immediate feature extraction (Lamme and Roelfsema, 2000). In contrast, the latter afford contextual refinement over extended timescales, integrating information across different hierarchies and bolstering representational accuracy and sophistication (Gilbert and Li, 2013).

Excitatory pyramidal neurons respond instantly within local circuits, enabling rapid signal propagation. Conversely, inhibitory interneurons, often characterized by their slower temporal profiles, meticulously fine-tune these dynamics, ensuring stability and preventing over-excitation (Isaacson and Scanziani, 2011). This duality effectively balances rapid activity and slow, stabilizing modulation.

The hippocampal formation also illustrates the fast-slow dichotomy, featuring a rapid feedforward trisynaptic loop, which interfaces with the neocortex through slower reciprocal connections (Buzsáki, 1996). This enables memory consolidation by replaying previously encoded sequences, transferring episodic memories for long-term storage (Carr et al., 2010).

Neuromodulatory systems often operate on slower timescales. Through diffuse projections, these systems induce state transitions in faster networks. Acetylcholine, in particular, is instrumental in gating plasticity, modulating cortical dynamics, and adjusting representations based on the prevailing behavioral contexts (Hasselmo and Sarter, 2011).

As currently conceptualized, ANNs often lack dedicated pathways distinguishing between fast and slow temporal scales. Many architectures, especially those in mainstream deep learning, consist of homogenous components operating on roughly similar timescales, thereby missing out on the multi-scale temporal processing in biological systems (Bengio et al., 2015).

Therefore, incorporating dedicated fast and slow processing streams into artificial neural

architectures could be transformative. Slow global networks, for instance, might modulate and gate the representations in faster local pathways, mirroring biological top-down modulation (Friston, 2010, Lee and Mumford, 2003). By embedding structured fast-slow motifs, artificial systems could attain heightened temporal fidelity (Kiebel et al., 2008), depth (Hassabis et al., 2017), and adaptability (Lake et al., 2017), moving closer to the versatility exhibited by their biological counterparts (Kriegeskorte, 2015).

1.5 Research Gaps

While advanced artificial networks have made strides in mimicking neural processes, a discernible gap remains between their capabilities and the nuanced intricacies of biological systems. Neural Computation, bridging the vast landscapes of biological and artificial realms, sits at a pivotal juncture of understanding and innovation. The exploration of spatiotemporal patterns, fundamental to both domains, elucidates the intricate dance between spatial structures and temporal dynamics. The historical backdrop has witnessed groundbreaking research and models that offer insights into this domain. However, three primary gaps remain in comprehensively grasping spatiotemporal patterns and their representations:

1. **Depth of Encoding:** Biological systems inherently process spatiotemporal patterns with remarkable depth and hierarchy. In contrast, while modern ANNs can recognize and even generate these patterns, they often lack the intricate depth of encoding observed in their biological counterparts.
2. **Robustness and Resilience:** Biological neural systems have evolved over millennia to be inherently resilient to adversarial attacks, noise, and degradation. Conversely, artificial systems, though powerful, have shown vulnerabilities, especially in noisy environments or under adversarial conditions.
3. **Natural Versatility:** Biological systems exhibit versatility in handling a multitude of spatiotemporal patterns simultaneously, seamlessly transitioning between them based on context. On the other hand, artificial systems often require separate specialized

architectures to handle different patterns or tasks, lacking the inherent multitasking dynamism observed in biological networks.

Fast and slow asymmetric motifs in biological systems epitomize the sophistication of neural computation, facilitating robust, layered, and adaptable spatiotemporal processing. Embracing these principles, particularly when designing and conceptualizing artificial neural architectures, offers a promising avenue. Drawing inspiration from such biological models, and incorporating them into machine learning paradigms, can bridge existing gaps and elevate the capabilities of artificial systems (Hassabis et al., 2017).

1.6 Dissertation Objectives

Biological systems exhibit remarkable hierarchical encoding depth, noise resilience, and versatile transitions between stored patterns. There is a critical need to capture these capabilities in artificial systems. Asymmetric fast-slow motifs in biological networks offer inspiration, but their core computational principles need further elucidation.

The overall objectives of this dissertation are to (i) elucidate spike timing coordination mechanisms enabling cortical state transitions, providing insights into the resilience gap, (ii) to establish a mathematical equivalence between Hopfield networks and spike-coding networks by introducing structured asymmetric connectivity, inducing equivalent cyclic attractor transitions that enable both models to store static memory patterns that can be replayed as temporal sequences, thereby addressing the versatility gap, (iii) to develop an unpaired image-to-image translation framework using GANs as a testbed for evaluating benefits of separate fast-slow processing pathways. This provides insights into the depth gap.

Our central hypothesis is that **asymmetric connectivity and dedicated fast and slow processing pathways in neural systems enhance depth, robustness, and versatility in handling complex spatiotemporal patterns**.

The rationale for this dissertation is that synergistically combining insights from theoretical neuroscience and models of neural computation will provide advances not attainable through either alone. Mathematical modeling will elucidate biological principles related to

the identified gaps. GAN implementations will assess the practical potential of bio-inspired architectures. This combined approach aims to elucidate mesoscale brain mechanisms while extracting lessons to shape AI.

1.7 Specific Aims

To attain the overall objectives, three specific aims will be pursued (Figure 1.1). The aims explore whether introducing separate fast-slow processing streams and asymmetric connectivity motifs can impart biological sophistication. Aim 1 models balanced excitation-inhibition, inspired by patterned biological inhibition regulating network stability. It investigates resilience mechanisms for adaptable transitions between brain states. Aim 2 introduces structured asymmetry into associative memory models to unify pattern completion and temporal sequencing, assessing versatile memory transitions lacking in ANNs. Aim 3 explores separate processing streams in a GAN framework, emulating feedforward-feedback and neuromodulatory asymmetries, to enhance translation depth and fidelity.

1.7.1 Aim 1

Our Aim 1 is **to elucidate the spike coordination mechanisms enabling synchronization and rhythmic transitions between active and silent cortical states when balancing fast excitatory slow inhibitory activity**. *We hypothesize that the spontaneous transitions between Up and Down states emerge from the coordinated balance of excitation and inhibition in cortical networks when near a codimension-two bifurcation point.* This hypothesis aligns with the resilience and versatility gaps, as balanced interactions are critical for biological systems' stability and repertoire.

We predict that:

(i) Modulating the relative excitatory-inhibitory activity will determine the frequency and duration of state transitions. This prediction provides insights into the resilience gap by revealing stabilizing mechanisms.

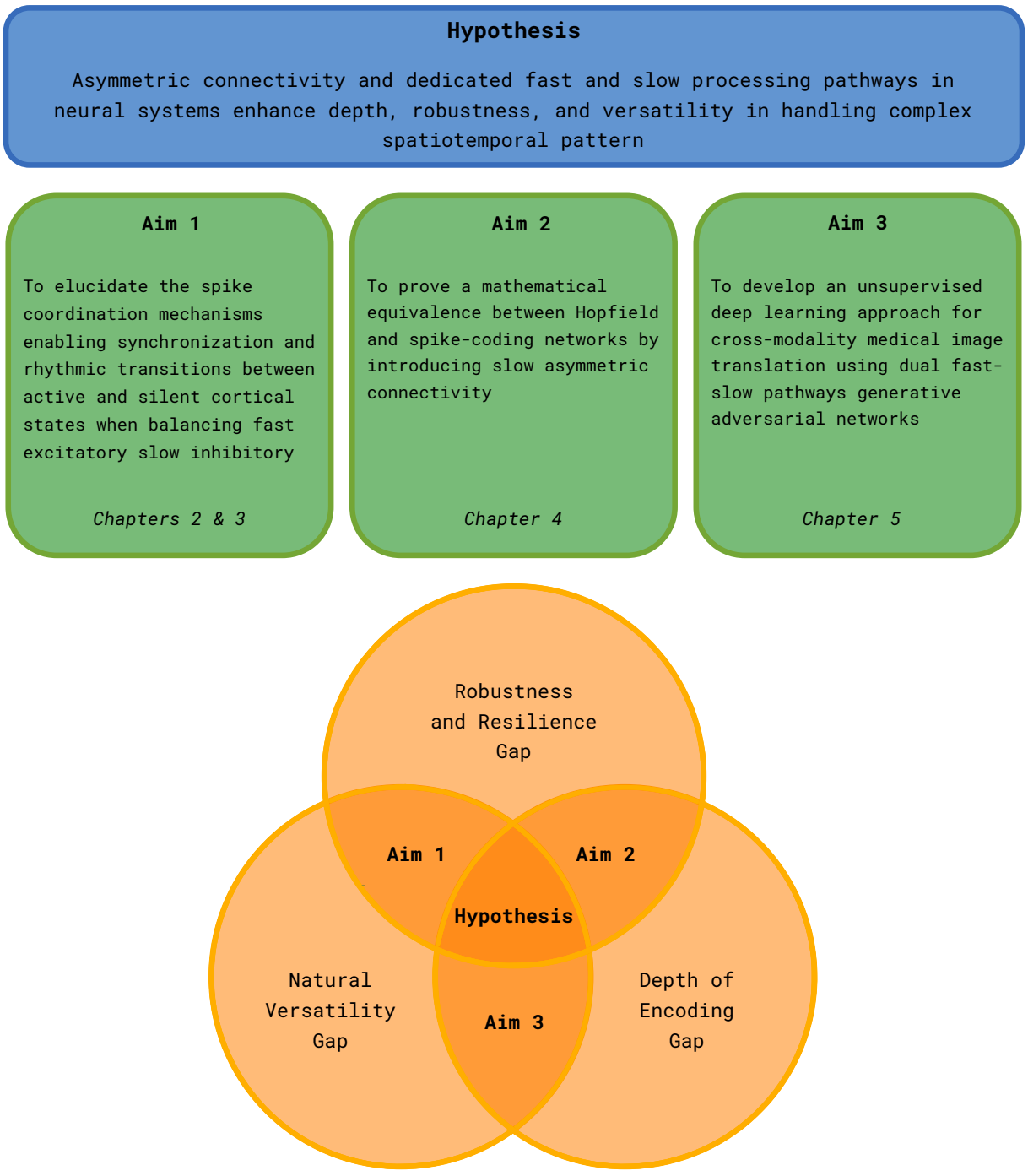


Figure 1.1: **Hypothesis and specific aims pursued in this dissertation.** Our hypothesis (blue box) is evaluated by addressing three aims (green boxes). Aims 1, 2, and 3, are pursued in Chapters 2 and 3, 4, and 5, respectively. The Venn diagram (orange circles) illustrates which research gaps each aim addresses.

(ii) Identifying the bifurcation structure underlying this excitation-inhibition balance will provide insights into the origin of cortical slow oscillations. This prediction addresses the encoding depth gap by connecting microscale dynamics to macroscale rhythms.

To test this, we will use the following approach (Chapters 2 and 3):

1. Derive membrane potential dynamics for spike-coding neuron networks.
2. Examine phase-locking for synchronizing spike times across neurons.
3. Apply population modeling to cortical neurons exhibiting Up-Down state transitions.
4. Analyze the effect of noise and coupling strength on transitions in oscillator networks.
5. Identify the bifurcation mechanism underlying cortical rhythm generation.

This theoretical modeling approach will advance our understanding of how interactions at the neuronal level coordinate to generate macroscopic brain rhythms.

1.7.2 Aim 2

Our Aim 2 is **to prove a mathematical equivalence between Hopfield and spike-coding networks by introducing slow asymmetric connectivity**. *We hypothesize that incorporating synapses with rotational symmetry will induce equivalent cyclic attractor transitions in both models.* This hypothesis directly tackles the versatility and encoding gaps regarding transitioning between stored memories in a meaningful sequence.

We predict that:

(i) Introducing asymmetry into Hopfield networks through structured perturbations of the synaptic weights will induce cyclic dynamics, allowing sequential transitions between stored memory states. This addresses the versatility gap by demonstrating memory chaining capacity.

(ii) Similarly, adding structured asymmetric synaptic connectivity in spike-coding networks will allow them to exhibit cyclic dynamics and maintain temporal order information. This prediction provides insights into the versatility gap by unifying associative completion and temporal sequencing.

(iii) The cyclic dynamics emerging from structured asymmetric synaptic weights will unify the computational capabilities of Hopfield and spike-coding networks. This prediction bridges theory to application, aligning with the depth and encoding gap.

(iv) Simulating spike-coding networks with appropriate tuning of slow synaptic weights on a working memory task will demonstrate their capability to encode and discriminate temporal sequences. This assessment tackles the resilience gap regarding handling noise and partial inputs.

To test this, we will use the following approach (Chapter 4):

1. Present the concepts of associative memory in modern Hopfield networks and their ability to store patterns through Hebbian learning.
2. Show equivalence between the spike-coding and Hopfield networks when parameterized by cyclic permutation matrices.
3. Generalize the mathematical equivalence framework through circulant matrices that enable modeling arbitrary temporal dependencies.
4. Relate the models to CNNs, which also perform a form of discrete convolution by using convolutional filters to detect spatial patterns.
5. Simulate a spiking neural network implementing the n-back working memory task to thoroughly characterize the ability of spike-coding networks to encode and retrieve meaningful temporal sequences.

By connecting these foundational neural network models through a common mathematical framework, we aim to provide an integrated understanding of the associative and dynamic aspects of neural computation.

1.7.3 Aim 3

Our Aim 3 is **to develop an unsupervised deep learning approach for cross-modality medical image translation using dual fast-slow pathways generative adversarial networks**. We will focus on the challenging task of translating 3 Tesla (3T) magnetic resonance imaging (MRI) scans to 7 Tesla (7T) MRI. Based on the predictions from Aim 2 that architectural asymmetry and separate slow-fast processing pathways can

enhance model performance, *we hypothesize that a dual pathway GAN generator will improve translation accuracy compared to a standard single pathway generator.* This hypothesis directly assesses the prediction from Aims 1 and 2 that separate asymmetric pathways confer benefits like enhanced robustness and depth, aligning with those gaps.

We predict that:

(i) The model can generate realistic synthetic 7T MR images from 3T MR inputs while preserving structural and semantic information. This tests the depth of encoding gap by evaluating multi-scale translation fidelity.

(ii) The model will demonstrate robust performance in translating tissue-specific characteristics between modalities. This prediction addresses the resilience gap by assessing consistency across datasets.

(iii) Implementing a dual pathway GAN generator will enhance translation accuracy compared to a standard single pathway generator. This directly tackles the depth and resilience gaps by validating the benefits of architectural asymmetry.

To test this, we will use the following approach:

1. Curate multi-site datasets of unpaired 3T and 7T structural MRI scans as training and validation data for the GAN models.
2. Implement standard and dual pathway CycleGAN architectures using 2D convolutional neural networks for unsupervised 3T to 7T MRI translation.
3. Optimize model hyperparameters, including network width and depth, loss functions, and training procedures to enhance translation fidelity.
4. Quantitatively evaluate model performance using segmentation metrics, including Dice coefficient and Percentual Area Distance for key tissue types.
5. Compare standard and dual pathway models to assess the predicted benefits of introducing asymmetric processing streams.

The focus on translational impact, rigorous quantification, and clinical relevance in the evaluation ensures Aim 3 meets the standards for an applied biomedical engineering contribution. This comprehensive approach encompassing robust dataset curation, model devel-

opment, and quantitative validation, aims to rigorously assess the real-world applicability of conceptual predictions on the benefits of architectural asymmetry for medical image analysis.

While Aims 1 and 2 establish core theoretical principles through detailed modeling of biological phenomena, Aim 3 demonstrates the practical potential of these findings in deep neural networks. The focus on architectural asymmetry in Aim 3, inspired by the coordinated excitation-inhibition balance in Aim 1 and structured synaptic perturbations in Aim 2, bridges theory and application. Collectively, the three Aims provide an integrative platform spanning neural computation models and their implementation in machine learning systems.

1.8 List of Contributions

Contributions in Chapter 2:

- Derived dynamical equations governing membrane potential dynamics in networks of spiking neurons.
- Demonstrated how populations of leaky integrate-and-fire neurons can encode signals into spike times through phase-locking.
- Analyzed stability of phase-locked states and impact of synaptic time constants.
- Generalized coding network to handle multiple input signals.

Contributions in Chapter 3:

- Developed Wilson-Cowan neural oscillator network exhibiting transitions between Up and Down states.
- Demonstrated the spontaneous emergence of slow oscillations from the balance between excitation and inhibition.
- Analyzed the effects of noise and coupling strength on the frequency and duration of the state transitions.
- Revealed Bogdanov-Takens bifurcation underlies rhythm genesis in slow oscillations.

Contributions in Chapter 4:

- Reviewed mathematical principles behind modern Hopfield networks and their ability to store memories through Hebbian learning.

- Demonstrated the equivalence between spike-coding and Hopfield models by representing asymmetric weights in Hopfield networks as cyclic permutation matrices.
- Generalized the mathematical equivalence framework to circulant matrices that enable modeling complex temporal dependencies.
- Related the discrete convolution implemented in these networks to CNNs.
- Simulated the n-back working memory task in the spike-coding network to demonstrate the ability of spike-coding networks to encode and retrieve temporal sequences.

Contributions in Chapter 5:

- Compiled unpaired datasets of 3T and 7T structural MR images to train CycleGAN networks.
- Developed 2D CycleGAN model to translate 3T MRI to synthetic 7T MRI in an unsupervised manner using unpaired training data.
- Implemented Fast-Slow U-Net architecture to enhance translation through separate fast and slow processing pathways.
- Evaluated quality of synthesized 7T images using segmentation metrics including Dice coefficient and percent area difference.
- Demonstrated potential of CycleGAN for spatial adaptive normalization to combine 3T and 7T MR data.

1.9 Overview of the Dissertation Structure

Aim 1 will be addressed in Chapters 2 and 3. Chapter 2 comprehensively examines spike-coding networks' mathematical foundations and dynamics. Key topics include spike time encoding via phase-locking, stability analysis of phase-locked solutions, and extensions to handling multiple encoded signals. The theoretical modeling approach will provide insights into spike coordination mechanisms for synchronizing distributed neuronal populations.

Chapter 3 builds upon the spike timing coordination principles from Chapter 2 to investigate the population dynamics underlying cortical Up-Down state transitions during slow-wave sleep. Key topics include developing a network model of coupled neural oscillators

exhibiting bistable Up-Down states, analyzing the modulation of state transitions by noise and coupling strength, and identifying a Bogdanov-Takens bifurcation structure giving rise to the spontaneous emergence of slow oscillations from the delicate balance of excitation and inhibition. This modeling approach sheds light on the mesoscopic mechanisms orchestrating macroscale brain rhythms.

Aim 2 will be covered in Chapter 4. We prove that introducing asymmetry to Hopfield and spike-coding networks elicits equivalent cyclic attractor dynamics, allowing both models to generate temporal sequences. Structured synaptic weight perturbations induce cycles in Hopfield networks, while structured asymmetric connectivity creates equivalent cycles in spike-coding networks. This unifies their ability to store static memory patterns that can be replayed as complex timing sequences. Simulations of spike-coding networks on the n-back task demonstrate this sequence generation capacity. Thus, Chapter 4 provides a unified view of associative memory models through the lens of cyclic dynamics.

Aim 3 will be explored in Chapter 5, where we present our work on developing a 2D CycleGAN model for cross-modality 3T to 7T MRI translation. We train the model on unpaired 3T and 7T datasets and implement the Fast-Slow U-Net architecture to improve multi-scale processing. The model is evaluated using tissue segmentation tasks to quantify its efficacy in generating realistic 7T MRI from 3T inputs. This approach demonstrates CycleGAN's potential to aid spatial adaptive normalization for combining historical 3T and modern 7T data, accelerating brain imaging research. The key contributions are the compilation of multi-site 3T and 7T MR datasets, developing of an unsupervised 2D CycleGAN model, implementing the Fast-Slow U-Net variant, and quantitative evaluation using segmentation metrics. The results highlight CycleGAN's capacity to bridge the technological divide between MRI field strengths while preserving compatibility with existing data.

Chapter 6 summarizes this dissertation research's key findings and contributions. It also discusses potential future directions from this theoretical and computational modeling work on elucidating brain rhythms and memory consolidation mechanisms. References cited across all chapters are collected in the bibliography.

2.0 Spike-Based Predictive Coding: A Framework for Neural Network Implementation

2.1 Chapter Summary

This chapter explores spike-based predictive coding networks' mathematical foundations and dynamics. These networks encode information using precise spike times rather than average firing rates. We derive the equations governing the membrane potential dynamics of spiking neurons and their synchronization through phase-locking. Stability analysis reveals how parameters like the synaptic time constant affect network synchronization. At extremes, specific parameters lead to complex behaviors like metastability. We also generalize the network to encode multiple signals through linear combinations of spike trains. Overall, examining spike-coding network dynamics offers insights into how neurons communicate and self-organize in the brain.

2.2 Introduction

The human brain is a complex network of approximately 86 billion neurons, which communicate via electrical signals to transmit information and mediate cognitive processes (Azevedo et al., 2009). Understanding how these neurons encode and decode information in the form of electrical impulses, often represented as spike trains, is one of the central challenges in computational neuroscience.

Spike coding refers to information encoded in the timing of action potentials, also known as spikes, within neurons. These spikes represent rapid, transient shifts in a neuron's membrane potential. This form of neural communication contrasts the rate coding approach, where the average firing rate of a neuron is considered the primary carrier of information (Gerstner et al., 2014). By focusing on the precise timing of individual spikes, spike coding allows for a more detailed and intricate understanding of neuronal information processing.

Since the seminal work of Lapicque at the beginning of the 20th century, numerous models have been proposed to describe the biophysics of neural spiking behavior, with the Hodgkin-Huxley model and the leaky integrate-and-fire model being amongst the most renowned ones (Hodgkin and Huxley, 1952, Lapicque, 1907). Over the years, the field has embraced diverse perspectives, exploring stochastic aspects, adapting firing thresholds, and integrating synaptic dynamics, among others (Brette and Gerstner, 2005, Friston et al., 2006).

In recent years, a new paradigm has emerged exploring predictive coding in neural networks, a hypothesis postulating that the brain is a prediction machine (Rao and Ballard, 1999). This framework has been adapted to the context of spiking neurons, leading to the development of spike-based predictive coding networks. These networks have demonstrated substantial potential for modeling neuronal dynamics and serving as the groundwork for more complex computational models (Eliasmith and Anderson, 2003, Boerlin et al., 2013).

We start with a comprehensive exploration of the mathematical and theoretical foundations of spike-coding networks, delving into the differential equations that govern their behavior and the extension of these concepts to encompass multiple encoded functions. Then, we derive the phase-locking conditions and examine the impact of different parameters on synchronization and stability. Phase-locking—the process in which two or more oscillating systems establish a fixed phase relationship—has significant implications for neuronal synchronization, communication, and information processing within neural circuits (Buzsaki and Draguhn, 2004). Finally, we provide crucial insights into the behavior of the networks under extreme conditions. The complex mechanisms underlying these regimes advance our understanding of neuronal dynamics and hint at the potential applications in ANNs and computational models that strive to mimic the brain's sophisticated coding strategies.

2.3 Constructing the Spike-Coding Network

2.3.1 Modeling Spiking Neurons

Suppose we want to encapsulate a signal, denoted as $f(t)$, within the activities of a neural network. Initially, we impose a constraint on this network, specifying that it comprises *spiking model neurons*. These neurons are characterized by generating a short-lived pulse or spike when their membrane voltage surpasses a predetermined threshold. Without this crossing, the neurons remain dormant and devoid of output. Hence, our task is to translate $f(t)$ into spike trains emitted by these neurons.

Decoding $f(t)$ can be achieved by filtering and summing these spike trains, effectively enabling synaptic integration. Specifically, the response of neuron i ($i = 1, 2, \ldots, N$) to an incoming current striking its membrane is defined as the spike train (Figure 2.1b:

$$\delta_i(t) = \sum_k \delta(t - t_i^k). \tag{2-1}$$

Here, $\delta(\cdot)$ represents a Dirac δ function, while t_i^k symbolizes the spike times of neuron i.

Next, we posit the normalized synaptic current, $s_i(t)$, to be the filtered form of the spike train $\delta_i(t)$ using an exponential filter (Figure 2.1c:

$$s_i(t) = \delta_i(t) * e^{-t/\tau} = \int_0^\infty e^{-t'/\tau} \delta_i(t - t') dt'. \tag{2-2}$$

In this equation, τ signifies the timescale of the filter. We assume that the neural membrane exhibits capacitive properties, resulting in the temporal filtering of its inputs.

This formulation can equivalently be represented as the dynamical equation:

$$\tau \frac{ds_i}{dt} = -s_i + \tau \delta_i. \tag{2-3}$$

We can construe this equation as a model of post-synaptic potential. Notably, with the use of this simplified synapse model, each presynaptic spike from neuron i causes an instantaneous increase of 1 in the normalized synaptic current, *i.e.*,

$$s_i \to s_i + 1. \tag{2-4}$$

This current then decays exponentially toward 0, with a time constant τ between spikes.

Finally, we define the output of the network, $z(t)$, as a weighted sum of the normalized synaptic currents:

$$z(t) = \sum_{i=1}^{N} \omega_i s_i(t). \tag{2-5}$$

Our final assumption is that *the output weights* $\boldsymbol{\omega_i}$ *are known* and they span the phase space. Figure 2.1a summarizes our configuration.

2.3.2 Decoding the Signal

Our objective is to derive the dynamical equations characterizing individual neurons and synapses, to discern the nature of the driven input signal $f_D(t)$, processed through weights $\boldsymbol{u_D}$, and to determine the pattern of connectivity (the $N \times N$ matrix of fast synapses strength $\boldsymbol{J}^{\text{fast}}$) required for the network to generate suitable spike trains at appropriate times to accurately represent the signal $f(t)$, *i.e.*, $z(t) \approx f(t)$. Furthermore, we desire to express $f_D(t)$, $\boldsymbol{u_D}$, and $\boldsymbol{J}^{\text{fast}}$ solely in terms of $\boldsymbol{\omega}$ and $s(t)$.

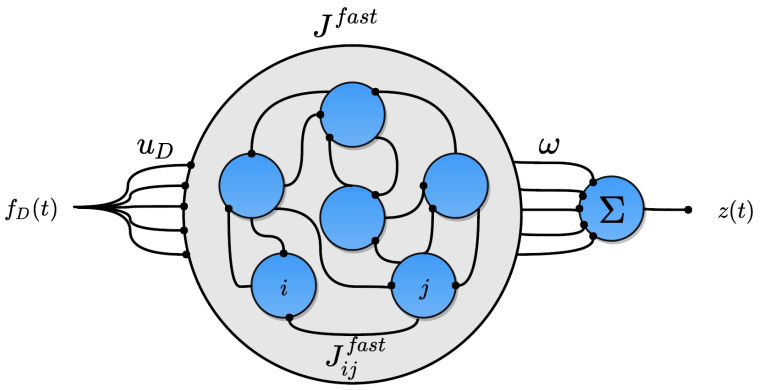
To achieve our goal, we must solve the classic credit-assignment problem in network learning, *i.e.*, we must determine how each neuron i in the network should behave such that the network's output $z(t)$ approximates the desired signal $f(t)$ as closely as possible. This can be conceptualized as an optimization problem. One approach is to define the mean squared error of the signal's decoding over time, denoted as $E(t_0)$:

$$E(t_0) = \int_{t_0}^{T} \left(f(t) - z(t)\right)^2 dt. \tag{2-6}$$

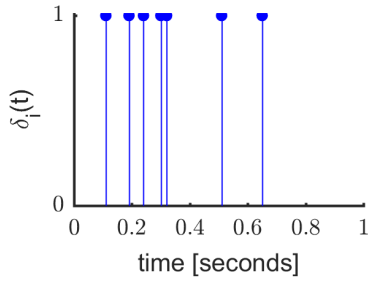
The objective then is to pinpoint the minimal set of spike trains $\delta_i(t)$, $(i = 1, 2, \ldots, N)$, during the interval $[t_0, T]$ such that $E(t_0)$ reaches a minimum. It is worth noting that our optimization strategy does not entail adjusting the *fixed* output weights, which are predetermined; instead, we optimize the timing of the spikes.

This approach deviates from rate coding networks, wherein the recurrent connections, $\boldsymbol{J}^{\text{fast}}$, are random and fixed, while the decoding weights are learned. Therefore, the optimization problem we must resolve is:

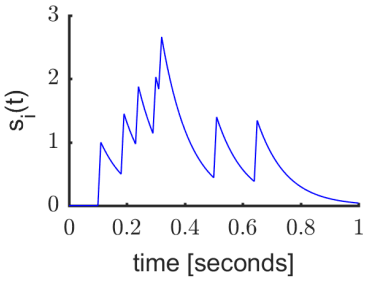
$$\min_{\delta_1, \ldots, \delta_N} E(t_0). \tag{2-7}$$



(a)



(b)



(c)

Figure 2.1: **The spike-based neural network.** (a) Structure of the network. A given input f_D is supplied to the network via weights $\boldsymbol{u_D}$. The neurons within the network are interconnected through fast synapses, the strengths of which are given by the matrix $\boldsymbol{J}^{\text{fast}}$. Specifically, the synaptic current generated in a postsynaptic neuron i by a presynaptic neuron j is dictated by the synaptic weight J_{ij}^{fast} multiplied by the normalized synaptic current $s_j(t)$ from neuron j. The network output is computed by summing, with weights $\boldsymbol{\omega}$, the synaptic currents of the neurons. (b) An example of a spike train from neuron i, $\delta_i(t)$, and (c) the corresponding normalized synaptic current, $s_i(t)$. $\tau = 100$ milliseconds.

Gradient-based techniques do not provide a feasible solution, as establishing a differentiable error measure suitable for spike trains is still a topic of ongoing research. Nevertheless,

Boerlin et al. (2013) proposes a simple heuristic, which entails utilizing a spike rule for the greedy minimization of the cost function $E(t_0)$. This rule asserts that neuron i should only generate a spike at time t if such an action would reduce the decoding error at time t.

Let us consider $E(t|\text{no spike}) = (f(t) - z(t))^2$ as the decoding error at time t, assuming no spikes occurred at that time, and $E(t|\text{neuron } i \text{ spikes}) = (f(t) - \bar{z}(t))^2$ as the decoding error at time t, assuming neuron i fired at that time.

A critical question to address here pertains to the effect of a spike on the output; that is, *how does the new output, $\bar{z}(t)$, adjust if neuron i fires at time t?* To answer this, we write a dynamical equation for $z(t)$:

$$\begin{aligned}
\frac{dz}{dt} &= \sum_{i=1}^{N} \omega_i \frac{ds_i}{dt} && \text{(time derivative of Eq. (2-5))} \\
&= -\frac{1}{\tau} \sum_{i=1}^{N} \omega_i s_i + \sum_{i=1}^{N} \omega_i \delta_i && \text{(Using Eq. (2-3))} \\
&= -\frac{1}{\tau} z + \sum_{i=1}^{N} \omega_i \delta_i && \text{(Using Eq. (2-5))} \\
\therefore \quad \tau \frac{dz}{dt} &= -z + \sum_{i=1}^{N} \tau \omega_i \delta_i. && \text{(2-8)}
\end{aligned}$$

Note that Eq. (2-8) has a similar form to Eq. (2-3), and by analogy, we can conclude that each spike from neuron i instantaneously increases the output by ω_i:

$$z \to z + \omega_i,$$

and then the output decays exponentially toward 0 with a time constant τ between spikes. Consequently, the new output $\bar{z}(t)$, given that neuron i fired at time t is:

$$\bar{z} = z + \omega_i. \tag{2-9}$$

Thus, we arrive at the *spiking rule for neuron i*: Neuron i should fire whenever the following condition is met:

$$
\begin{aligned}
E(t|\text{neuron } i \text{ spikes}) &< E(t|\text{no spikes}) && \text{(2-10)}\\
(f-\bar{z})^2 &< (f-z)^2 \\
(f-(z+\omega_i))^2 &< (f-z)^2 && \text{(Using Eq. (2-9))}\\
-2f\omega_i + 2z\omega_i + \omega_i^2 &< 0 && \text{(Expanding the squares)}\\
\omega_i(z-f) &< -\frac{\omega_i^2}{2} \\
\omega_i(f-z) &> \frac{\omega_i^2}{2}. && \text{(2-11)}
\end{aligned}
$$

If we consider $\omega_i(f-z)$ as the membrane voltage, $v_i(t)$, and $\omega_i^2/2$ as the firing threshold, T_i of neuron i, then we can define the voltage of neuron i and the firing threshold as

$$v_i(t) := \omega_i(f-z) \tag{2-12}$$

and

$$T_i := \frac{\omega_i^2}{2}. \tag{2-13}$$

This brings us to conclude that whenever the error $(f-z)$ exceeds $\omega_i/2$, neuron i will fire a spike, as illustrated by the following firing rule:

$$v_i(t) = \omega_i(f-z) > \frac{\omega_i^2}{2} = T_i. \tag{2-14}$$

To refine our understanding, we aim to derive a model that encapsulates this behavior for individual neurons within the network. We formulate a dynamical equation for the neuron's

potential $v_i(t)$. Therefore, we take the time derivative of Eq. (2-12):

$$\begin{aligned}
\frac{dv_i}{dt} &= \omega_i \left(\dot{f} - \frac{dz}{dt} \right) && (\dot{f} := \frac{df}{dt}) \\
&= \omega_i \dot{f} + \frac{1}{\tau} \omega_i z - \sum_{j=1}^{N} \omega_i \omega_j \delta_j && \text{(Using Eq. (2-8))} \\
\therefore \quad \tau \frac{dv_i}{dt} &= \tau \omega_i \dot{f} + \omega_i z - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j \\
&= \tau \omega_i \dot{f} + [\omega_i f - \omega_i f] + \omega_i z - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j \\
&= -\omega_i (f - z) + \omega_i (\tau \dot{f} + f) - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j \\
&= -v_i + \omega_i (\tau \dot{f} + f) - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j && \text{(Using Eq. (2-12))} \\
\therefore \quad \tau \frac{dv_i}{dt} &= -v_i + \omega_i (\tau \dot{f} + f) + \sum_{j=1}^{N} (-\tau \omega_i \omega_j) \delta_j. && \text{(2-15)}
\end{aligned}$$

From Eq. (2-15), it becomes apparent that every spike originating from neuron j results in an instantaneous decrease in the voltage of neuron i by an amount of $\omega_i \omega_j$,

$$v_i \to v_i - \omega_i \omega_j. \tag{2-16}$$

Specifically, when $i = j$, it leads to the self-reset

$$v_i \to v_i - \omega_i^2, \tag{2-17}$$

which implies that neuron i fires when $v_i = T_i$, followed by a reset to $v_i = T_i - \omega_i^2 = -T_i$.

The membrane potential of neuron i, $v_i(t)$, essentially tracks the decoding error $f(t) - z(t)$. Once this decoding error surpasses $T_i = \omega_i / 2$ (refer to Eq. (2-14)), neuron i is triggered to fire a spike. This, in turn, modifies the network's output as $z \to z + \omega_i$ (refer to Eq. (2-9)). Additionally, the firing of neuron i resets its membrane voltage to $v_i \to -T_i$.

Eq. (2-15) along with the firing condition stated in Eq. (2-14) collectively describe a network of leaky integrate-and-fire neurons. The input to this network is characterized by

$$f_D(t) = \tau \dot{f} + f, \tag{2-18}$$

whereas the input weights are denoted by $\boldsymbol{u_D} = \boldsymbol{\omega}$, and the fast synaptic connections are expressed as $J_{ij}^{\text{fast}} = -\tau\omega_i\omega_j$.

The approach to understanding rate-based networks—in which neurons communicate via continuous variables contrasting with the spike-based communication in spike-coding networks—often involves a bid to compensate for the effects of nonlinearities (*e.g.*, firing rule) and synaptic connections $\boldsymbol{J}^{\text{fast}}$. This can be achieved by selecting a suitable output weight $\boldsymbol{\omega}$. A common strategy is setting $f_D(t)$ to $f(t)$, then tuning $\boldsymbol{\omega}$ such that z equals f. That way, the network could encode $f(t)$ via the neurons' firing rates and later decode it through $\boldsymbol{\omega}$ (Amit and Brunel, 1997). Despite the reasonable results this technique produces for rate-based networks, it often falls short when applied to spiking networks, whether they are rate-coding or spike-coding networks[1]. Here, the deficiency is attributable to the inability to offset the effects of the low-pass filtering operation and the resultant phase delay—occurring during the conversion from spikes to normalized synaptic currents—through the mere adjustment of $\boldsymbol{\omega}$. A viable solution is to input a high-pass filtered, phase-advanced version of the desired output, $f_D = \tau\dot{f} + f$, into the network (Abbott and Kepler, 2005).

2.4 Autonomous Encoding of Signals

Eq. (2-15) offers the dynamic model for the spiking neurons and expresses $\boldsymbol{u_D}$ and $\boldsymbol{J}^{\text{fast}}$ solely in terms of $\boldsymbol{\omega}$. However, f_D remains related to the signal f. In this state, the network essentially operates as an autoencoder (Hinton and Salakhutdinov, 2006). It accepts the desired signal $f(t)$ as an input, represents it through neuronal activity, and recovers $f(t)$ at the output by decoding the spike trains. Even though we have determined how to drive the network to achieve $f(t)$ at the output, our ultimate goal is to develop an autonomous system, *i.e.*, a network that can encode f independently without requiring it as input. This demands a three-fold approach: ensuring that the network can generate f_D (as given in

[1]Spiking networks can come in two flavors: rate-coding and spike-coding. In rate-coding networks, neurons communicate through spikes and encode the target output in the neurons' *firing rates*. In contrast, spike-coding networks communicate through spikes but encode the target output in the *spike times*. Both are spiking networks communicating through spikes but differ in how they encode the target output.

Eq. (2-18)) autonomously; identifying a set of recurrent connections that can accomplish this with satisfactory accuracy; and ensuring the stability of the solution with respect to the network dynamics (Doya et al., 1992).

For the case of a linear dynamic system that produces f as

$$\tau \dot{f} = -f + Af + c(t), \tag{2-19}$$

where $c(t)$ is a general control signal, we can identify such an input.

Specifically, we can rewrite Eq. (2-15) by integrating Eq. (2-19):

$$\begin{aligned}
\tau \frac{dv_i}{dt} &= -v_i + \omega_i A f - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j + \omega_i c && \text{(Using Eq. (2-19))} \\
&= -v_i + \omega_i A z - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j + \omega_i c && \text{(Since } f \approx z\text{)} \\
&= -v_i + \sum_{j=1}^{N} \omega_i A \omega_j s_j - \sum_{j=1}^{N} \tau \omega_i \omega_j \delta_j + \omega_i c && \text{(Using Eq. (2-5))} \\
\therefore \quad \tau \frac{dv_i}{dt} &= -v_i + \sum_{j=1}^{N} (-\tau \omega_i \omega_j) \delta_j + \sum_{j=1}^{N} (\omega_i A_{ij} \omega_j) s_j + \omega_i c(t). && \text{(2-20)}
\end{aligned}$$

Now, along with $J_{ij}^{\text{fast}} = -\tau \omega_i \omega_j$, we have

$$J_{ij}^{\text{slow}} = \omega_i A \omega_j. \tag{2-21}$$

J_{ij}^{slow} represents a set of connections that effectively implement the dynamic system given by Eq. (2-19). The synapses described by $\boldsymbol{J}^{\text{slow}}$ transform the normalized synaptic currents s_j, therefore slower recurrent connections when juxtaposed with $\boldsymbol{J}^{\text{fast}}$, which modifies the raw spikes δ_j instead, hence the adopted fast/slow terminology.

The interplay between rapid recurrent synapses ($\boldsymbol{J}^{\text{fast}}$) and slower recurrent synapses ($\boldsymbol{J}^{\text{slow}}$) serve a two-fold purpose, with the former primarily orchestrating a competitive dynamic among neurons and the latter fostering cooperation. For example, consider the interaction between neuron i and neurons k that share sign ($\omega_k \omega_i > 0$). When neuron i fires, it excites similar neurons, immediately incrementing their voltages $v_k(t)$ proportional to $\omega_k \omega_i$. Such cooperative connections provide predictions trajectory of $f(t)$. In situations where $f(t)$

remains static (setting $A = 1$ and $c(t) = 0$ in Eq. (2-19)), these connections enable the network to sustain ongoing activity, thereby preventing $z(t)$ from decaying to zero.

Conversely, when neuron i interacts with neurons j that have an opposite sign ($\omega_j\omega_i < 0$), the action of neuron i firing prompts an inhibition of dissimilar neurons, as evidenced by an immediate decrement in their voltages $v_j(t)$ proportional to $\omega_j\omega_i$.

At first glance, one might attempt to conclude that fast synapses may correspond to observed AMPA excitatory synapses in biological neural networks, and slower synapses match NMDA inhibitory synapses. Similarly, one might assume that speedy GABA_A synapses relate to fast responses and lethargic GABA_B synapses to slower responses. However, this assumption is overly simplistic. The relationship between these in biological neurons is far more intricate due to the diverse processes that modify these synapses (Koch, 2004).

Figure 2.2 offers a network representation of the autonomous network that has been constructed, with Figure 2.2a detailing the resultant output feedback structure that follows by modifying Eq. (2-18) given the assumptions Eq. (2-19) and $f \approx z$:

$$\tau\dot{f} = -f + Af + c \qquad \text{(Eq. (2-19))}$$

$$\tau\dot{f} + f = Af + c$$

$$f_D(t) = Af(t) + c(t) \qquad \text{(Using Eq. (2-18))}$$

Consequently, we must pass the network output through the linear dynamical system defined by A, append $c(t)$, and then feed this as the network input. Importantly, this negative feedback loop does not need to be explicitly implemented in the network but can be realized within the neural population (Amit, 1989). As illustrated in Figure 2.2b, the arrangement presented in Figure 2.2a is equivalent to the inclusion of a new set of synaptic connections with strengths corresponding to Eq. (2-21). Consequently, the network's connectivity is entirely contingent upon $\boldsymbol{\omega}$ and the desired dynamics A.

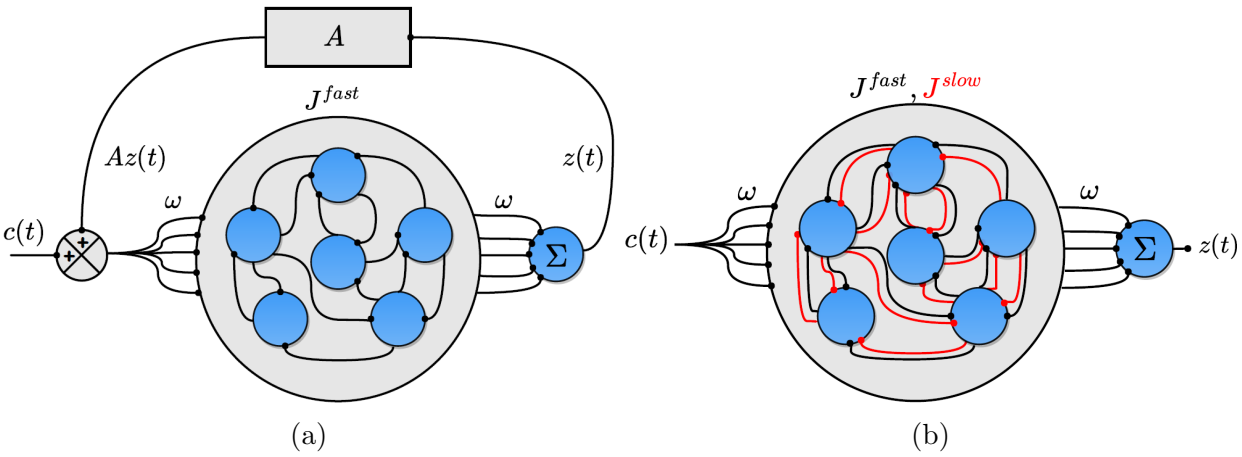


Figure 2.2: **Network representation of the autonomous spike-coding network.** (a) Schematic of the output feedback mechanism employed to generate the input in terms of the output, premised on the assumption that $f(t)$ satisfies Eq. (2-19). (b) Schematic portraying an equivalent network constructed via the inclusion of an additional set of synapses ($\boldsymbol{J}^{\text{slow}}$) with strengths as indicated by Eq. (2-21). Red connections indicate slow synaptic connections, while black connections denote fast ones. $c(t)$ represents the control signal.

2.4.1 Spike-Trigger Adaptation

Before moving forward, it is necessary to address an additional consideration: the case where ω_i is markedly small. As indicated in Eq. (2-9), if ω_i is small, neuron i will minimally influence z, yet will fire at a remarkably high rate, given the prediction of a shallow firing threshold by Eq. (2-13). This issue could be avoided through the introduction of a quadratic cost term controlled by the gain μ to encourage as few spikes as possible and to foster a uniform distribution of spikes across neurons:

$$E(t_0) = \int_{t_0}^{T} \left[\left(f(t) - z(t) \right)^2 + \mu \sum_{i=1}^{N} s_i^2(t) \right] dt. \tag{2-22}$$

From this, we can re-derive the **spiking rule for neuron i** as follows:

$$
\begin{aligned}
&E(t|\text{neuron } i \text{ spikes}) < E(t|\text{neuron } i \text{ does not spike}) \\
&(f-(z+\omega_i))^2 + \mu(s_i^2+1) < (f-z)^2 + \mu s_i^2 && \text{(Using Eq. (2-4))} \\
&-2f\omega_i + 2z\omega_i + \omega_i^2 + \mu s_i + \mu < 0 \\
&\omega_i(z-f) + \mu s_i < -\frac{1}{2}(\omega_i^2+\mu). \\
&\omega_i(f-z) - \mu s_i > \frac{1}{2}(\omega_i^2+\mu). && \text{(2-23)}
\end{aligned}
$$

Following similar logic to prior discussions, we define the voltage of neuron i as

$$v_i(t) := \omega_i(f-z) - \mu s_i \tag{2-24}$$

and the firing threshold as

$$T_i := \frac{1}{2}(\omega_i^2+\mu). \tag{2-25}$$

In this context, the dynamic equation for $v_i(t)$ becomes:

$$
\begin{aligned}
&\frac{dv_i}{dt} = \omega_i\left(\dot{f} - \frac{dz}{dt}\right) - \mu\frac{ds_i}{dt} && \text{(Derivative of Eq. (2-24))} \\
\therefore\quad &\tau\frac{dv_i}{dt} = \tau\omega_i\dot{f} + \omega_i z - \sum_{j=1}^{N}\tau\omega_i\omega_j\delta_j - \mu\left(-s_i + \tau\delta_i\right) && \text{((2-3) \& (2-8))} \\
&\quad = \tau\omega_i\dot{f} + [\omega_i f - \omega_i f] + \omega_i z + \mu s_i - \sum_{j=1}^{N}\tau\omega_i\omega_j\delta_j - \tau\mu\delta_i \\
&\quad = -[\omega_i(f-z) - \mu s_i] + \omega_i(\tau\dot{f} + f) - \sum_{j=1}^{N}\tau\omega_i\omega_j\delta_j - \tau\mu\delta_i \\
&\quad = -v_i + \omega_i(\tau\dot{f} + f) - \sum_{j=1}^{N}\tau\omega_i\omega_j\delta_j - \tau\mu\delta_i && \text{(From (2-24))} \\
\therefore\quad &\tau\frac{dv_i}{dt} = -v_i + \omega_i(\tau\dot{f} + f) + \sum_{j=1}^{N}(-\tau\omega_i\omega_j)\delta_j - \tau\mu\delta_i. && \text{(2-26)}
\end{aligned}
$$

As suggested by Eq. (2-26), each spike from neuron j causes a momentary decrease in voltage by $\omega_i\omega_j$ (a concept previously introduced; refer to Eq. (2-16)). This results in a reset:

$$v_i \to v_i - \omega_i^2 - \mu. \tag{2-27}$$

This implies that neuron i fires when $v_i = T_i$ and then resets to $v_i = T_i - \omega_i^2 - \mu = -T_i$, with T_i given by Eq. (2-25). Eq. (2-26) also implies that $J_{ij}^{\text{fast}} = -\tau(\omega_i\omega_j + \mu\delta_{ij})^2$.

The role of the quadratic cost gain μ is similar to introducing a hyperpolarizing current in the membrane potential of neurons that just spiked and relates to previous findings on spike-triggered adaptation (Liu and Wang, 2001, Benda and Herz, 2003).

2.4.2 Multiple Encoded Functions

The previous discussion has primarily been confined to instances where $f(t)$ is a scalar function, thereby limiting Eq. (2-19) to represent either a low-pass filter or an integrator within a one-dimensional context. To expand our framework to a more extensive range of system implementations, we introduce the possibility of $\boldsymbol{f}(t)$ being a vector. Variable and parameter definitions utilized throughout the following derivation can be found in Table 2.1.

Adapting the derivations of the earlier sections to support multiple encoded functions is a straightforward process. As before, signal decoding proceeds via the computation of a weighted sum of normalized synaptic currents as shown in Eq. (2-28):

$$z_k(t) = \sum_{j=1}^{N} \omega_{kj} s_j(t), \tag{2-28}$$

where ω_{kj} represents the weight of neuron j's contribution to the output $f_k(t)$.

In this generalized framework, the network dynamics evolve according to the following:

$$\tau\frac{dv_i}{dt} = -v_i + \sum_{k=1}^{K} \omega_{ik}(\tau\dot{f}_k + f_k) - \tau\sum_{k=1}^{K}\omega_{ik}\sum_{j=1}^{N}\omega_{kj}\delta_j - \tau\mu\delta_i, \tag{2-29}$$

complemented by the spiking rule:

$$v_i > T_i = \frac{1}{2}\left(\sum_{k=1}^{K}\omega_{ik}^2 + \mu\right). \tag{2-30}$$

[2]By the notation δ_{ij} it is meant the Kronecker delta function, *i.e.*, $\delta_{ij} = 0$ when $i \neq j$, and $\delta_{ij} = 1$ when $i = j$. Not to be confused with δ_i, which denotes the spike train of neuron i.

An autonomous network capable of encoding $\boldsymbol{f}(t)$ can be constructed assuming that $f_k(t)$ are generated by a linear dynamic system represented as follows:

$$\tau \dot{f}_k = -f_k + \sum_{p=1}^{K} A_{kp} f_p + c_k. \tag{2-31}$$

Suppose this assumption holds and the approximation $\boldsymbol{f} \approx \boldsymbol{z}$ is in effect. In that case, we can re-write Eq. (2-29) to yield:

$$\tau \frac{dv_i}{dt} = -v_i + \sum_{k=1}^{K} \omega_{ik} \sum_{j=1}^{N} A_{kj} \omega_{pj} s_j - \tau \sum_{k=1}^{K} \omega_{ik} \sum_{j=1}^{N} \omega_{kj} \delta_j - \tau \mu \delta_i + \sum_{k=1}^{K} \omega_{ik} c_k, \tag{2-32}$$

with the respective fast and slow synaptic currents expressed as

$$J_{ij}^{\text{fast}} = -\tau \left(\sum_{k=1}^{K} \omega_{ik} \omega_{kj} + \mu \delta_{ij} \right)$$

and

$$J_{ij}^{\text{slow}} = \sum_{k=1}^{K} \sum_{p=1}^{K} \omega_{ik} A_{kp} \omega_{pj}.$$

Interestingly, this network structure can autonomously represent any linear dynamical system by simply choosing $\boldsymbol{A}$ appropriately.

In matrix notation, Eqs. (2-28)–(2-32) are

$$\boldsymbol{z}(t) = \boldsymbol{\Omega} \boldsymbol{s}(t), \tag{2-33}$$

$$\tau \dot{\boldsymbol{v}} = -\boldsymbol{v} + \boldsymbol{\Omega}^T (\tau \dot{\boldsymbol{f}} + \boldsymbol{f}) - \tau \boldsymbol{\Omega}^T \boldsymbol{\Omega} \boldsymbol{\delta} - \tau \mu \boldsymbol{\delta} + \boldsymbol{\Omega}^T \boldsymbol{c}, \tag{2-34}$$

with spiking rule

$$\boldsymbol{v} > \boldsymbol{T} = \frac{1}{2} \text{diag} \left(\boldsymbol{\Omega}^T \boldsymbol{\Omega} + \mu \boldsymbol{I} \right). \tag{2-35}$$

Assuming $\boldsymbol{f}(t)$ can be generated by a linear dynamic system of the form

$$\tau \dot{\boldsymbol{f}} = -\boldsymbol{f} + \boldsymbol{A} \boldsymbol{f} + \boldsymbol{c}, \tag{2-36}$$

we can re-write Eq. (2-34) as

$$\tau \dot{\boldsymbol{v}} = -\boldsymbol{v} + \boldsymbol{\Omega}^{\text{slow}} \boldsymbol{s} + \boldsymbol{\Omega}^{\text{fast}} \boldsymbol{\delta} + \boldsymbol{\Omega}^T \boldsymbol{c}, \tag{2-37}$$

where $\boldsymbol{\Omega}^{\text{fast}} = -\tau \left(\boldsymbol{\Omega}^T \boldsymbol{\Omega} + \mu \boldsymbol{I} \right)$ and $\boldsymbol{\Omega}^{\text{slow}} = \boldsymbol{\Omega}^T \boldsymbol{A} \boldsymbol{\Omega}$.

Table 2.1: List of variables and parameters for the autonomous spike-coding network.

Variable/Parameter	**Description**
τ	synapse dynamics timescale (scalar)
N	number of neurons (scalar)
K	number of inputs (scalar)
$\boldsymbol{f}(t) = [f_1(t), \ldots, f_K(t)]^T$	encoded signals ($K \times 1$ vector)
$\boldsymbol{c}(t) = [c_1(t), \ldots, c_K(t)]^T$	control signals ($K \times 1$ vector)
$\boldsymbol{z}(t) = [z_1(t), \ldots, z_K(t)]^T$	decoded (or estimated) signals ($K \times 1$ vector)
$\boldsymbol{\delta}(t) = [\delta_1(t), \ldots, \delta_N(t)]^T$	spike trains ($N \times 1$ vector)
$\boldsymbol{s}(t) = [s_1(t), \ldots, s_N(t)]^T$	normalized synaptic currents ($N \times 1$ vector)
$\boldsymbol{v}(t) = [v_1(t), \ldots, v_N(t)]^T$	membrane potentials ($N \times 1$ vector)
$\boldsymbol{T} = [T_1, \ldots, T_N]^T$	firing thresholds ($N \times 1$ vector)
μ	quadratic cost term gain (scalar)
$\boldsymbol{A}$	describes the linear temporal evolution of the encoded signal $\boldsymbol{f}(t)$ ($K \times K$ matrix)
$\boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{K1} & \omega_{K2} & \cdots & \omega_{KN} \end{bmatrix}$	decoding weights ($K \times N$ matrix); ω_{ki} is the weight of contribution of neuron i to the k-th decoded signal $z_k(t)$
$\boldsymbol{\Omega}_i = [\omega_{1i}, \ldots, \omega_{Ki}]^T$	i-th column of $\boldsymbol{\Omega}$ ($K \times 1$ vector)
$\boldsymbol{J^{fast}} = -\tau \boldsymbol{\Omega}^T \boldsymbol{\Omega}$	fast synaptic connections ($N \times N$ matrix); $J_{ij}^{fast} = -\tau \boldsymbol{\Omega}_i^T \boldsymbol{\Omega}_j$
$\boldsymbol{J^{slow}} = \boldsymbol{\Omega}^T \boldsymbol{A} \boldsymbol{\Omega}$	slow synaptic connections ($N \times N$ matrix); $J_{ij}^{slow} = \boldsymbol{\Omega}_i^T \boldsymbol{A} \boldsymbol{\Omega}_j$

In summary, the autonomous network of spiking neurons encoding the linear dynamical

system in Eq. (2-36) within the spike times of neurons are given by the following set of differential equations:

$$\tau \dot{\boldsymbol{v}} = -\boldsymbol{v} + \boldsymbol{\Omega}^{\text{slow}} \boldsymbol{s} + \boldsymbol{\Omega}^{\text{fast}} \boldsymbol{\delta} + \boldsymbol{\Omega}^T \boldsymbol{c} \tag{2-38}$$

$$\tau \dot{\boldsymbol{s}} = -\boldsymbol{s} + \tau \boldsymbol{\delta}, \tag{2-39}$$

where

$$\boldsymbol{\Omega}^{\text{fast}} = -\tau \left(\boldsymbol{\Omega}^T \boldsymbol{\Omega} + \mu \boldsymbol{I} \right), \tag{2-40}$$

$$\boldsymbol{\Omega}^{\text{slow}} = \boldsymbol{\Omega}^T \boldsymbol{A} \boldsymbol{\Omega}. \tag{2-41}$$

2.5 Oscillatory Patterns in Spike-Coding Networks

Our focus in this section lies primarily in the analytical examination of spike-coding networks, specifically those operating in regimes supportive of frequency-locked solutions. This realm of operation is characterized by network units acting as oscillators, exhibiting a synchronous firing pattern that shares a common period (Hoppensteadt and Izhikevich, 1997). This behavior encapsulates all phase-locked solutions, notably the synchronous and traveling wave states, foundational to understanding the complex dynamics of neural networks (Ermentrout, 1996, Kopell and Ermentrout, 2002, Kuramoto, 1984).

To facilitate comprehension, we reiterate our experimental setup and recapitulate the dynamical equations which serve as the foundation for our subsequent mathematical analysis. The following system defines the spike-coding network of N uniform LIF neurons:

$$\tau \frac{dv_i}{dt} = -v_i - \tau\epsilon \sum_{j=1}^{N} J_{ij}^{\text{fast}} \delta_j(t) + \epsilon \sum_{j=1}^{N} J_{ij}^{\text{slow}} s_j(t) + \omega_i I_i, \tag{2-42a}$$

$$\tau \frac{ds_i}{dt} = -s_i + \tau \delta_i(t). \tag{2-42b}$$

Here, I_i denotes a persistent external bias, and the terms $J_{ij}^{\text{fast}} = \omega_i \omega_j$ and $J_{ij}^{\text{slow}} = \omega_i A_{ij} \omega_j$ designate recurrent connections from the jth neuron to the ith neuron. The function $\delta_i(t) = \sum_{n \in \mathbb{Z}} \delta(t - t_i^n)$ is the spike train of the ith neuron, with firing times t_i^n. This structure entails

neurons communicating with one another through the synaptic transmission of spike trains. The global strength of these synaptic connections is represented by the parameter ϵ.

In this context, $s_i(t)$ signifies the temporal evolution of the postsynaptic response to the spike train $\delta_i(t)$; *i.e.*, each spike is transformed into a postsynaptic current, as dictated by the kernel implicit in Eq. (2-42b):

$$s_i(t) = (\alpha * \delta_i)(t), \tag{2-43}$$

where

$$\alpha(t) = e^{-t/\tau}\Theta(t) \tag{2-44}$$

is the standardized postsynaptic response kernel, and $\Theta(t) = 1$ for $t > 0$.

Applying the method of variation of parameters, one can solve equations (2-42) and obtain:

$$v_i(t) = v_i(0)e^{-t/\tau} + \omega_i I_i \left(1 - e^{-t/\tau}\right) + \epsilon\frac{1}{\tau}\sum_{j,m} J_{ij}^{\text{slow}} A(t - t_j^m, 0) - \epsilon\sum_{j,m} J_{ij}^{\text{fast}} E(t - t_j^m, 0), \tag{2-45}$$

where

$$E(x, y) := e^{-(x+y)/\tau}\Theta(y) \tag{2-46}$$

and $A(x, y)$ denotes the integral

$$A(x, y) := e^{-x/\tau}\int_0^x e^{u/\tau}\alpha(u + y)du. \tag{2-47}$$

From the perspective of the ith neuron, it perceives a sequence of action potentials occurring at times $\{t_j^m\}$, with each spike eliciting a synaptic response conforming to (2-42b). A spike is triggered when v_i reaches the threshold ϑ_i, immediately resetting v_i to the resting potential, $-\vartheta_i$. By denoting the firing times of the ith neuron by $\{t_i^m\}$, we can conceptualize the neuron as a functional unit that translates $\{t_j^m\} \to \{t_i^m\}$.

We can construct an implicit map of the firing times for the spike-coding network under consideration. Here, ϑ_i symbolizes the firing threshold, implying $v_i(t_i^{n-}) = \vartheta_i$ and $v_i(t_i^{n+}) = -\vartheta_i$. Integrating Eq. (2-42) from t_i^n to t_i^{n+1} yields the nonlinear firing time map:

$$\vartheta_i = \omega_i I_i(1 - e^{-(t_i^{n+1} - t_i^n)/\tau}) + \epsilon\frac{1}{\tau}\sum_{j,m} J_{ij}^{\text{slow}} A(t_i^{n+1} - t_j^m) - \epsilon\sum_{j,m} J_{ij}^{\text{fast}} E(t_i^{n+1} - t_j^m), \tag{2-48}$$

where $t_i^{n+1} > t_j^m, \forall j, m$.

2.5.1 Phase-Locked States

We restrict our attention to phase-locked solutions of (2-42) in which every oscillator fires with a self-consistent period T, to be determined. For instance, consider the case where two neurons continue firing periodically when coupled. Suppose that neuron $i = 1$ fires at times $t_1^n = nT$, where n is an integer, while neuron i=2 fires at $t_2^n = (n - \phi_2)T$. Thus, both neurons are firing at the same frequency but are separated by a phase $0 < \phi_2 < 1$. In general, the spike times satisfy,

$$t_i^n = (n - \phi_i)T, \tag{2-49}$$

with constant phases ϕ_i.

Plugging the ansatz (2-49) in (2-48), we arrive at the algebraic conditions for phase-locking in the spike-coding network:

$$\vartheta_i = \omega_i I_i(1 - e^{-T}) - \epsilon \sum_j J_{ij}^{\text{fast}} E_T(\phi_j - \phi_i) + \epsilon \frac{1}{\tau} \sum_j J_{ij}^{\text{slow}} A_T(\phi_j - \phi_i), \tag{2-50}$$

where

$$A_T(\phi) := \sum_m A(T, (\phi - m)T) = e^{-T/\tau} \int_0^T e^{u/\tau} \alpha_T(u + \phi) du, \tag{2-51}$$

$$\alpha_T(\phi) := \sum_m \alpha((\phi - m)T), \tag{2-52}$$

and

$$E_T(\phi) := \sum_m E((m - \phi)T) = \sum_m e^{-(\phi - m)T}, \tag{2-53}$$

with $A_T(\phi)$, $\alpha_T(\phi)$ and $E_T(\phi)$ extended outside the range $0 < \phi < 1$, by making them periodic functions of ϕ with period T.

Choosing ϕ_1 as a reference phase, Eq. (2-50) generates a system of N equations for the unknown period T and the remaining $N - 1$ relative phases $\phi_j - \phi_1$. These equations could be evaluated for several coupling functions $\alpha(t)$, yielding the algebraic conditions for phase-locking in a spike-coding network with various topologies.

$E_T(\phi)$ represents the "instantaneous phase interaction" in that it depicts the immediate consequences of a neuron's spiking on the membrane potential of other interconnected neurons, predominantly through fast synaptic connections. Essentially, $E_T(\phi)$ is a normalized

sum of exponential terms characterizing the disparities in spike times amongst different neurons, weighted by the inverse of the exponential decay rate, $1 - e^{-T/\tau}$ (Roxin et al., 2005). It encapsulates the nature of an immediate synaptic response induced by the spiking activity of a neuron, as reflected in its phase ϕ. Intuitively, $E_T(\phi)$ depicts the extent to which a spike from a presynaptic neuron instantly impacts the postsynaptic neuron's membrane potential, culminating in an instantaneous change in its potential (Wang, 2010).

$A_T(\phi)$ is a "delayed phase interaction function" that encapsulates the lagged influence of a neuron's spiking on other neurons, predominantly through slow synaptic connections. This delay is triggered by the conversion process of a spike into a postsynaptic current, as represented by the integral in its definition. The normalization here is also conducted by $1 - e^{-T/\tau}$, which approximates the contribution from various spikes (Roxin et al., 2005). At an intuitive level, $A_T(\phi)$ represents how the postsynaptic neuron's membrane potential is altered over time, after the firing of the presynaptic neuron, factoring in the process of postsynaptic current generation and its subsequent decay (Hansel et al., 1995).

In aggregate, $E_T(\phi)$ and $A_T(\phi)$ articulate the phase interaction effects, thereby encapsulating the impact of one neuron's spiking activity on another, while taking into account the time scales of synaptic transmission (Ermentrout and Terman, 2010).

2.5.2 Stability of Phase-Locked States

Stability analysis plays a vital role in the study of dynamical systems. Here, we're interested in the stability of the phase-locked solutions of the spike-coding network (2-42). In simple terms, we want to know whether small perturbations in the phase will be damped out or amplified over time, determining whether the system returns to the phase-locked state or moves away from it, respectively.

Perturbations to the phase-locked state of the system can be formally expressed as:

$$t_i^n = (n - \phi_i)T + \delta_i^n. \tag{2-54}$$

Here, δ_i^n constitutes infinitesimal perturbations in the phases of the neurons.

Next, we linearized the firing time map (2-48) around the phase-locked state. This is achieved by taking the first-order Taylor expansion of the map with respect to the perturbations. The non-linear term involving the spike times will become a linear term involving the phase perturbations $\delta\phi_i$. The linearized map gives us a set of linear differential equations that describe how the perturbations evolve.

Specifically, we substitute Eq. (2-54) into Eq. (2-48) and unfold the resultant expression as a power series in the perturbations δ_i^n. Phase-locked equations (2-50) can be recovered at the $\mathcal{O}(1)$ order, while the $\mathcal{O}(\delta)$ order yields an infinite-order linear difference equation. This equation forms an explicit map for the perturbations:

$$\begin{aligned}
0 &= \omega_i I_i e^{-T/\tau}(\delta_i^{n+1} - \delta_i^n) \\
&+ \epsilon \sum_j J_{ij}^{\text{fast}} E_T(\phi_j - \phi_i)(\delta_i^{n+1} - \delta_j^m) - \epsilon \frac{1}{\tau} \sum_j J_{ij}^{\text{slow}} A_T(\phi_j - \phi_i)(\delta_i^{n+1} - \delta_i^n) \\
&+ \epsilon \sum_j J_{ij}^{\text{slow}} \alpha_T(\phi_j - \phi_i)(\delta_i^{n+1} - \delta_i^n) + \epsilon \sum_j J_{ij}^{\text{slow}} H_T(\phi_j - \phi_i)(\delta_i^n - \delta_j^m).
\end{aligned} \tag{2-55}$$

Here, $H_T(\phi)$ is defined by:

$$H_T(\phi) := e^{-T/\tau} \int_0^T e^{u/\tau} \alpha_T'(u + \phi) du, \tag{2-56}$$

where $\alpha_T' \equiv d\alpha_T/du$.

The function $H_T(\phi)$ is a time-dependent quantity defined in terms of the time derivative of α_T. In essence, $H_T(\phi)$ is a measure of how the membrane potential of the ith neuron responds to a small change in the phase difference between it and another neuron in the network. The intuition behind this function is that it captures how phase perturbations propagate through the network, influencing the timing of each neuron's spikes.

Eq. (2-55) represents the linearized map at phase-locked solutions, premised on the assumption that the relatively insignificant $\mathcal{O}(\delta^2)$ terms can be safely disregarded.

To facilitate understanding of this intricate expression, let us contemplate the case of an uncoupled network with constant input $I_i(t) = I_i$. In such a setting, the asymptotic value of the membrane potential for the ith neuron is $\omega_i I_i$. If $\omega_i I_i > \vartheta_i$, the neuron exhibits periodic

firing. Assuming $v_i(0) = 0$, Eq. (2-45) allows us to solve for the time of the first spike, t_i^1, leading to the following relationship:

$$\vartheta_i = \omega_i I_i \left(1 - e^{-t_i^1/\tau}\right).$$

Subsequently, we find:

$$t_i^1 = \tau \ln\left(\frac{\omega_i I_i}{\omega_i I_i - \vartheta_i}\right) \equiv T_i. \tag{2-57}$$

This equation captures T_i, the period of the isolated ith neuron, thereby shedding light on the intrinsic rhythm of the individual neuron when not subjected to any external influence or coupling with other neurons.

By utilizing the approximation $T \approx T_i$ and integrating Eq. (2-57) into Eq. (2-50), we get:

$$\epsilon \sum_j J_{ij}^{\text{fast}} E_T(\phi_j - \phi_i) - \epsilon \frac{1}{\tau} \sum_j J_{ij}^{\text{slow}} A_T(\phi_j - \phi_i) = 0.$$

Thus, Eq. (2-55) denoting the perturbation map can be condensed to the following form:

$$\left[\omega_i I_i - \vartheta_i + \epsilon \sum_j J_{ij}^{\text{slow}} \alpha_T(\phi_j - \phi_i)\right] (\delta_i^{n+1} - \delta_i^n) = \epsilon \sum_j J_{ij}^{\text{slow}} H_T(\phi_j - \phi_i)(\delta_i^n - \delta_j^m). \tag{2-58}$$

To determine the rate and direction in which the perturbations evolve, we consider solutions of the form:

$$\delta_i^n = \lambda^n \delta_i. \tag{2-59}$$

This represents a specific form of solutions to the perturbation map that describes how perturbations in the phase of firing neurons evolve. Here, δ_i^n denotes the perturbation in the firing time of the i^{th} neuron at the n^{th} spike, and λ is an eigenvalue of the Jacobian of the firing map at the phase-locked state (2-55), that characterizes the rate of evolution of the perturbations. The intuition behind this solution is that perturbations evolve geometrically over time—that is, they either grow or shrink at a constant rate as determined by λ. If $\text{Re}(|\lambda|) < 1$, then the perturbations shrink over time, suggesting the stability of the equilibrium state. On the other hand, if $\text{Re}(|\lambda|) > 1$, the perturbations grow over time, suggesting the instability of the equilibrium.

Substituting (2-59) in (2-48):

$$\left[\omega_i I_i - \vartheta_i + \epsilon \sum_j J_{ij}^{\text{slow}} \alpha_T(\phi_j - \phi_i)\right](\lambda - 1)\delta_i = \epsilon \sum_j J_{ij}^{\text{slow}} \left[\tilde{H}_T(\phi_j - \phi_i, \lambda)\delta_j - H_T(\phi_j - \phi_i)\delta_i\right]. \quad (2\text{-}60)$$

Here, $\tilde{H}_T(\phi, \lambda)$ is defined as:

$$\tilde{H}_T(\phi, \lambda) := e^{-T/\tau} \int_0^T e^{u/\tau} \sum_m \lambda^{-m} \alpha'(u + \phi) du. \quad (2\text{-}61)$$

$\tilde{H}_T(\phi, \lambda)$ is a more general form of $H_T(\phi)$ that includes a dependence on λ, the eigenvalue of the perturbation map.

Interestingly, for $\delta_i = \delta$, $\lambda = 1$ denotes a trivial solution, symbolizing perturbations concurrent with the phase-locked solutions. This implies that they are degenerate for uniform phase shifts in the firing times, indicated by $t_i^n \to t_i^n + \delta$. The stability criterion can thus be expressed in terms of nontrivial solutions to Eq. (2-60): if all remaining solutions λ fulfill $\text{Re}(\lambda) < 0$, the phase-locked state is linearly stable. This indicates that δ_i tends towards 0 geometrically fast as n approaches infinity.

2.5.3 Special Case—Stability of Synchrony

Although solutions λ are typically found through numerical integration, we can further analytically investigate the stability characteristics of the synchronous state, a specific phase-locked solution where $\phi_i = \phi$ for all i and an arbitrary ϕ. For the sake of simplicity, suppose $I_i = I$ holds for all i and a fixed I such that $\omega_i I > \vartheta_i$. The collective period of oscillation, T_c, then adheres to:

$$T_c = \tau \ln\left(\frac{\omega_i I}{\omega_i I - \vartheta_i}\right). \quad (2\text{-}62)$$

Under the condition that ω_i/ϑ_i is constant, T_c is self-consistent, signifying a synchronous state of the system.

By setting $T = T_c$ in Eq. (2-50), the existence of a synchronous solution is contingent upon the equation:

$$I_i = \frac{I}{\vartheta_i}\left[1 - \epsilon\frac{1}{\tau}\gamma_i^{\text{slow}} A_T(0) + \epsilon\gamma_i^{\text{fast}} E_T(0)\right], \quad (2\text{-}63)$$

where γ_i^{slow} and γ_i^{fast} are defined as $\sum_j J_{ij}^{\text{slow}}$ and $\sum_j J_{ij}^{\text{fast}}$ respectively.

Additionally, it is important to note that:

$$H_T(0) = A'_T(0). \tag{2-64}$$

Also, integrating by parts and using Eq. (2-62), we can demonstrate:

$$\begin{aligned}
A'_T(\phi) &= e^{-T/\tau} \int_0^T e^{u/\tau} \alpha'_T(u+\phi) du \\
&= \alpha_T(\phi+T) - e^{-T/\tau}\alpha(\phi) - \tfrac{1}{\tau} A_T(\phi) && \text{(integration by parts)} \\
&= \alpha_T(\phi)\left[1 - e^{-T/\tau}\right] - \frac{1}{\tau} A_T(\phi) && (\alpha(T+\phi) = \alpha(\phi)) \\
&= \alpha_T(\phi)\left[1 - \frac{\omega_i I - \vartheta_i}{\omega_i I}\right] - \frac{1}{\tau} A_T(\phi) && \text{(using (2-62))} \\
\therefore \quad \frac{\omega_i I}{\vartheta_i} A'_T(\phi) &= \alpha_T(\phi) - \frac{1}{\tau}\frac{\omega_i I}{\vartheta_i} A_T(\phi). && (2\text{-}65)
\end{aligned}$$

$A'_T(\phi)$ captures the rate of change of the membrane potential response to the phase difference. The intuition behind this is that it encapsulates the sensitivity of each neuron's membrane potential to changes in the phase difference between them and other neurons.

Enforcing synchrony, the condition of Eq. (2-63), and the identities (2-64) and (2-65), Eq. (2-60) can be reformulated to:

$$\left[(\lambda - 1)\left(\tfrac{\omega_i I}{\vartheta_i}\left(\epsilon\gamma_i^{\text{fast}} E_T(0) + \epsilon\gamma_i^{\text{slow}} A'_T(0) + 1\right) - \vartheta_i\right) + \epsilon\gamma_i^{\text{slow}} A'_T(0)\right]\delta_i = \epsilon\tilde{H}_T(0,\lambda)\textstyle\sum_j J_{ij}^{\text{slow}}\delta_j. \tag{2-66}$$

The left-hand side describes the dynamics of a single neuron's perturbation (δ_i) due to external stimuli and its interaction with the overall network. It includes the effect of the strength of external stimuli ($\omega_i I/\vartheta_i$), intrinsic neuronal properties such as the fast and slow timescales of neuronal dynamics (γ_i^{fast} and γ_i^{slow}), and the external fields ($E_T(0)$ and $A'_T(0)$).

The right-hand side describes the network topology's collective influence on the neuron's perturbation through the sum $\sum_j J_{ij}^{\text{slow}}\delta_j$. Here, J_{ij}^{slow} represents the slow synapse connectivity matrix and δ_j the perturbation of neuron j. The term $\tilde{H}_T(0,\lambda)$ is a scaled version of the first derivative of the external field, including the effects of the phase-locked state (λ).

Under the assumption of homogeneity in the neuronal dynamics, *i.e.*, when γ_i^{slow} and γ_i^{fast} are independent of i, we can further simplify (2-66) by diagonalizing it in terms of the eigenvalues of $\boldsymbol{J}^{\text{slow}}$ and suitably align $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$ with one of the eigenvectors. In

this case, γ^{slow} is the degenerate eigenvalue corresponding to the eigenvector $(1, \ldots, 1)$ that correspond to constant phase shift. Let σ_p, for $p = 1, \ldots, N$, denote the eigenvalues of J_{ij}^{slow}. Then, (2-66) simplifies to

$$(\lambda - 1)\left(\frac{\omega_i I}{\vartheta_i}\left(\epsilon\gamma_i^{\text{fast}} E_T(0) + \epsilon\gamma_i^{\text{slow}} A_T'(0) + 1\right) - \vartheta_i\right) + \epsilon\gamma_i^{\text{slow}} A_T'(0) - \epsilon\sigma_p \tilde{H}_T(0, \lambda) = 0. \quad (2\text{-}67)$$

This simplification allows us to express the original matrix equation in terms of eigenvalues of $\boldsymbol{J}^{\text{slow}}$ denoted by σ_p. Specifically, it gives the synchrony condition for each eigendirection p of the network connectivity matrix, associating synchrony to a particular mode of collective neuronal activity represented by the eigenvector corresponding to σ_p.

2.6 Analysis of Postsynaptic Potentials for the Exponential Synaptic Kernel

In the case of the postsynaptic response given by (2-44), $E_T(\phi)$, $A_T(\phi)$, and $\tilde{H}_T(\phi, \lambda)$ reduce to a geometric series, which can subsequently be evaluated into a closed form. Let's start with $E_T(\phi)$, representing the effective connectivity between neurons. It is calculated by summing up the total postsynaptic potentials, given by Eq. (2-44), at every firing time of a periodic neuron firing with period T and phase ϕ:

$$E_T(\phi) = \sum_{n=0}^{\infty} e^{-(nT+\phi)/\tau}. \quad (2\text{-}68)$$

This is a geometric series with a common ratio $r = e^{-T/\tau}$. The sum of this series, when $|r| < 1$, is given by $1/(1 - r)$. Applying this formula to the series, we obtain

$$E_T(\phi) = \frac{e^{-\phi/\tau}}{1 - e^{-T/\tau}}. \quad (2\text{-}69)$$

The quantity $A_T(\phi)$, on the other hand, represents the average phase advancement due to the postsynaptic potentials. It's calculated by summing up the weighted postsynaptic potentials, where the weight is the time since the last firing, over all firings. The contribution to the phase advancement from the nth firing is $(nT + t)e^{-(nT+t)/\tau}$, where t ranges from 0 to T. Therefore, $A_T(\phi)$ can be expressed as an infinite sum of integrals:

$$A_T(\phi) = \sum_{n=0}^{\infty} \int_0^T (nT + t) e^{-(nT+t+\phi)/\tau} dt. \tag{2-70}$$

Solving the integral inside the summation, we have

$$\int_0^T (nT + t) e^{-(nT+t+\phi)/\tau} dt = \left[-(nT + t) e^{-(nT+t+\phi)/\tau} - \tau e^{-(nT+t+\phi)/\tau} \right]_0^T. \tag{2-71}$$

Calculating this expression at $t = T$ and $t = 0$, and subtracting the latter from the former, we get

$$-(n+1)Te^{-((n+1)T+\phi)/\tau} - \tau e^{-((n+1)T+\phi)/\tau} + (nT)e^{-nT/\tau} + \tau e^{-(nT+\phi)/\tau}. \tag{2-72}$$

Summing this up over all n gives us the expression for $A_T(\phi)$:

$$A_T(\phi) = \frac{1}{1 - e^{-T/\tau}} \left[\phi T + \frac{T e^{-T/\tau}}{1 - e^{-T/\tau}} \right] e^{-\phi T/\tau}. \tag{2-73}$$

To derive the equations for $A'_T(\phi)$ and $\tilde{H}_T(\phi, \lambda)$, we need to differentiate $A_T(\phi)$ and evaluate a sum that involves a function of $E_T(\phi)$, respectively.

The derivative of $A_T(\phi)$ with respect to ϕ gives us $A'_T(\phi)$:

$$\begin{aligned}
A'_T(\phi) &= \frac{d}{d\phi} \left[\frac{1}{1 - e^{-T/\tau}} \left[\phi T + \frac{T e^{-T/\tau}}{1 - e^{-T/\tau}} \right] e^{-\phi T/\tau} \right] \\
&= -T \frac{1}{1 - e^{-T/\tau}} \left[\phi T + \frac{T e^{-T/\tau}}{1 - e^{-T/\tau}} \right] e^{-\phi T/\tau} + \frac{T}{1 - e^{-T/\tau}} e^{-\phi T/\tau} \\
A'_T(\phi) &= -T A_T(\phi) + T e^{-\phi T/\tau}.
\end{aligned}$$

Now, $\tilde{H}_T(\phi, \lambda)$ is defined as

$$\begin{aligned}
\tilde{H}_T(\phi, \lambda) &= \sum_{n=0}^{\infty} \lambda^n E_T(n\phi) \\
&= \sum_{n=0}^{\infty} \lambda^n \frac{e^{-n\phi T/\tau}}{1 - e^{-T/\tau}} \\
\tilde{H}_T(\phi, \lambda) &= \frac{1}{1 - e^{-T/\tau}} \sum_{n=0}^{\infty} (\lambda e^{-\phi T/\tau})^n,
\end{aligned}$$

which is a geometric series with ratio $r = \lambda e^{-\phi T/\tau}$. The sum of this series is given by $1/(1-r)$ when $|r| < 1$, hence we have

$$\tilde{H}_T(\phi, \lambda) = \frac{1}{(1 - e^{-T/\tau})(1 - \lambda e^{-\phi T/\tau})}.$$

Substitute the above expressions into Eq. (2-66), we get:

$$(\lambda - 1)(\alpha_i \beta_i - \vartheta_i)\delta_i + \xi_i \delta_i = \eta \sum_j J_{ij}^{\text{slow}} \delta_j, \tag{2-74}$$

where $\alpha_i \coloneqq \frac{\omega_i I}{\vartheta_i}$, $\beta_i \coloneqq \epsilon \gamma_i^{\text{fast}} E_T(0) + \epsilon \gamma_i^{\text{slow}} A'_T(0) + 1$, $\xi_i \coloneqq \epsilon \gamma_i^{\text{slow}} A'_T(0)$, and $\eta \coloneqq \epsilon \tilde{H}_T(0, \lambda)$

Now, for Eq. (2-67), we need to introduce one more term: $\zeta_p \coloneqq \epsilon \sigma_p \tilde{H}_T(0, \lambda)$. Then, using the terms defined earlier, Eq. (2-67) can be rewritten as:

$$(\lambda - 1)(\alpha_i \beta_i - \vartheta_i) + \xi_i - \zeta_p = 0 \tag{2-75}$$

2.6.1 Exploring Extreme Regimes

In a complex system like this, further analytical simplifications are only meaningful under specific conditions or certain parameter regimes. Nevertheless, the analytical expressions can provide insights into the system's qualitative behavior and guide future modeling.

Let's consider the limits of some of the parameters of our equations and how they affect the system. When τ approaches infinity, the time constant of the postsynaptic response becomes very large. From the definitions of $E_T(0)$, $A_T(0)$, $A'_T(0)$, and $\tilde{H}_T(0, \lambda)$, we can observe that in this case, these terms become independent of the time delay T and instead become constants. This implies that the system's dynamics become time-invariant.

The term ϵ in the equations measures the coupling strength between neurons. If ϵ approaches zero, the coupling strength becomes very weak. In this case, all terms involving ϵ vanish, which includes β_i, ξ_i, η and ζ_p. Effectively, this reduces (2-74) and (2-75) to simple expressions involving only the intrinsic frequencies of the oscillators and the phase differences, leading to a set of independent oscillators.

The term ϑ_i in the equations scales the phase difference between neurons, and if this term approaches infinity, it indicates that the phase difference between neurons becomes infinitesimally small. In this limit, the first part of both Eqs. (2-74) and (2-75) vanish (due

to α_i being finite), effectively simplifying the system of equations. The system behavior then depends mostly on the slow coupling terms.

In the limit as γ^{fast} approaches infinity, the influence of $E_T(0)$ (which multiplies γ^{fast}) in both (2-74) and (2-75) will dominate all other terms. The implication is that synchronization in the system is chiefly driven by the terms with $E_T(0)$ and potentially leads to an immediate phase locking, assuming ϵ is finite. This is a scenario where the fast connections in the network would primarily control synchronization. When γ^{fast} goes to zero, the terms involving $E_T(0)$ in Eqs. (2-74) and (2-75) become negligible. In this case, synchronization depends on the terms involving $A'_T(0)$ (related to slow coupling) and the individual intrinsic properties of each oscillator (characterized by λ, ω_i, ϑ_i).

Analogous to the corresponding γ^{fast} case, when γ^{slow} goes to zero, the terms in (2-74) and (2-75) involving $A'_T(0)$ become negligible. In this case, synchronization is mainly influenced by the terms involving $E_T(0)$ (related to fast coupling) and the individual properties of each oscillator. As γ^{slow} approaches infinity, the terms involving $A'_T(0)$ in (2-74) and (2-75) will dominate. This means slow connections primarily drive the synchronization. This could result in a very slow phase locking process, as the system will need time to respond to the slow interactions. The longer transient time before reaching a synchronized state allows the system to exhibit rich and complex dynamics.

2.6.2 Implications of Dominant Slow Connections and Multistability

As γ^{slow} edges towards infinity, the phase-locking process becomes largely dictated by the network's slow connections, referring to the "slow" coupling of oscillators. This transformation underscores the critical role of frequency interplay and the oscillators' phase relationships in determining system behavior (Buzsaki and Draguhn, 2004).

In neuroscience, "slow" connections encapsulate the connections via modulatory neurotransmitter systems or comparably slower chemical synapses with longer latency periods Destexhe and Sejnowski (2003). These slow connections are pivotal in modulating the network's global state, influencing network behavior over extended timescales, and contrasting the fast synaptic connections.

A predominant γ^{slow} results in phase-locking between oscillators on significantly slower timescales, resonating with observations in large-scale brain networks. In such systems, slower frequencies commonly control faster frequencies' activities through "phase-amplitude coupling" (Voytek et al., 2010). Given their precision, the phase relationships between oscillators become pivotal. This shift can lead to "multistability," where multiple stable phase-locked conditions are possible (Rosenblum and Pikovsky, 2003). This ability to switch between different coherent activity patterns based on the oscillators' phase relationships can provide a multistable network.

The concept of "dynamical relaying" (Vicente et al., 2008) adds another layer of complexity to networks dominated by slow connections. Under such circumstances, an indirect pathway might sometimes be more expedient than a direct one, primarily due to the delay-intrinsic frequency interplay within the nodes. This interplay may precipitate phase-locking along indirect pathways, creating more complex synchronization patterns.

These insights parallel the transient dynamics highlighted by Deco and Jirsa (2012), and Litwin-Kumar and Doiron (2012), showing that the synchronization process's nuances can provide key insights into the underlying mechanisms. Deco and Jirsa (2012) discussed the notions of multistability and "ghost attractors" in the context of ongoing cortical activity, indicating the potential for the system to inhabit several states before reaching synchronization. Furthermore, Litwin-Kumar and Doiron (2012) demonstrated the implications of slow dynamics on network variability in a cortical network with clustered connections. They found that the time taken to reach a synchronized state directly influences the variability of the network, leading to high variability when slow dynamics are present.

They also resonate with Engel and Singer (2001) temporal binding concept and the structured flows in the dynamical framework proposed by Huys et al. (2014). Engel and Singer (2001) suggested that these transients could play a critical role in forming sensory awareness, as the brain's perception of external stimuli is often determined during these periods. The analysis by Huys et al. (2014) provides a dynamical framework for understanding motor behavior, emphasizing the structured flows on manifolds that may occur during the transition to synchronization.

2.7 Conclusions

This chapter focused on a network of spiking model neurons, distinguished by their ability to generate a short-lived pulse or spike when their membrane voltage surpasses a specified threshold. Encoding a signal, denoted by $f(t)$, into a series of spikes, or spike trains, forms the cornerstone of the spike-coding network (Eliasmith and Anderson, 2003, Boerlin et al., 2013, Abbott et al., 2016).

Building upon this premise, we discussed the membrane potential dynamics, v_i, for neuron i. We set up a differential equation (Eq. (2-15)) representing the evolution of v_i, considering an integration process, inhibitory feedback, and a stimulus-dependent drive, which led us to Eq. (2-26). The neuron's firing rule was defined so that whenever v_i reaches a threshold T_i, it fires and resets its potential to $v_i = T_i - \omega_i^2 - \mu = -T_i$ (Eq. (2-25)). We further unpacked that the quadratic cost gain μ contributes a hyperpolarizing current to the membrane potential, particularly influencing the recently spiked neuron, thereby suggesting a mechanism for spike-triggered adaptation (Liu and Wang, 2001, Benda and Herz, 2003).

Subsequently, we focused on extending the network's functionality from scalar to vector-encoded operations. While we initially limited our discussions to scalar function $f(t)$ and corresponding low-pass filter or integrator systems, we broadened the scope by allowing $\boldsymbol{f}(t)$ to operate as a vector function, thus expanding the range of systems the network could potentially implement (Gütig, 2014). In this generalized network, the signals $\boldsymbol{f}(t)$ were decoded through a weighted sum of normalized synaptic currents, with each neuron contributing to the output based on its assigned weight.

We then presented an autonomous network model capable of encoding $\boldsymbol{f}(t)$, with $f_k(t)$ functions generated by a linear dynamic system (Boerlin et al., 2013). We could reformulate the network equations using an approximation, concluding that any linear dynamical system could be autonomously implemented with an appropriate choice of $\boldsymbol{A}$.

We derived the phase-locked solutions within the framework of the spike-coding network. This exploration commenced with the derivation of generic conditions fostering phase-locking, culminating in a set of algebraic equations contingent upon the relative phases of the oscillators alongside the self-consistent collective period. Subsequent analysis pivoted

towards evaluating the linear stability of these phase-locked solutions through perturbations of the firing times and the consequent examination of the asymptotic behavior of these perturbations. Notably, we observed that a pair of parameters dictated the stability of the solutions—λ, and the time constant of the synaptic response, τ.

Analysis of the phase-locking dynamics of the network gave us invaluable insights into how neurons synchronize their activity, how different parameters affect this synchronization, and the stability of these phase-locked states. The analysis of extreme conditions, where certain parameters approach infinity, sheds light on the robustness of these synchronization phenomena and their susceptibility to changes in network parameters.

When parameters such as γ^{slow} grew unbounded, we observe an increase in the coupling strength between the neurons leading to intriguing behaviors, such as metastability. Studying these phenomena further underscores the delicate balance of parameters needed to maintain a stable and functional neuronal network, and highlights the necessity of local and network-level control mechanisms.

This theoretical examination exposed the complexity of spike-coding networks and provided valuable perspectives on neuronal communication. As we continue to refine our understanding of the dynamics of neuronal networks, biomimetic models like the spike-coding networks, despite their simplicity, continue to offer us profound insights into the complex mechanisms underlying neuronal behavior and inter-neuron communication.

3.0 The Interplay of Synchronization and Origination in Slow Oscillations

3.1 Chapter Summary

This chapter investigates the neural mechanisms underlying the spontaneous transitions between active (Up) and silent (Down) states in cortical networks during slow-wave sleep. A model of weakly coupled Wilson-Cowan neural oscillators is analyzed, demonstrating bistability between quiescent and active states. Stochastic fluctuations and the coupling strength between neural populations modulate transitions between these bistable states. Varying key parameters reveals a Bogdanov-Takens bifurcation underlies the genesis of slow oscillations. This provides insight into the complex coordination of excitation and inhibition required to orchestrate the rhythmic Up-Down state transitions underlying slow-wave activity.

3.2 Introduction

Among the numerous strategies to model extensive brain networks, mesoscopic-level models such as firing rate or population models have proven to be one of the most effective approaches. These models focus on the average spike rates of neuron groups rather than scrutinizing the individual activity of each neuron within the network (Ermentrout, 1998a). This approach provides two primary advantages: computational efficiency, which eases the simulation of large-scale brain networks (Markram, 2006, Izhikevich and Edelman, 2008), and suitability for comparison with experimental data, facilitating correlation with real-world phenomena (Deco et al., 2008, Bojak et al., 2010).

Standard neuroimaging techniques such as electroencephalogram (EEG), local field potential (LFP) recordings, magnetoencephalogram (MEG), and functional magnetic resonance imaging (fMRI) measure the collective activity of large neuron populations rather than individual intracellular potentials. Thus, population models may serve as more representative approximations of emergent collective behavior within cortical columns.

In the context of brain region networks, each node symbolizes a neuronal ensemble (typically ranging between $10^3 - 10^6$ neurons). Each ensemble carries its dynamic description, distinguished by its mathematical complexity and biological detail. The widely recognized model formulated by Wilson and Cowan (WC) (Wilson and Cowan, 1972, 1973) is the focus of this discussion, as it describes cortical area activity. In the WC model, each network node is an interconnected pair of excitatory and inhibitory neuronal populations.

Wilson and Cowan adopted statistical methodologies to examine the collective behavior of substantial neuron populations. Their model embodies a mean-field approximation of a coupled network of excitatory and inhibitory neurons, focusing on their mean activity level, a strategy analogous to approximating the motion of gas molecules with mean-field terms such as temperature and pressure (Kittel, 2004). This mean-field approach has widespread applications in statistical physics, especially when addressing computationally or analytically arduous problems. Its efficacy in computational neuroscience is similarly recognized (see Deco et al. (2008) and the references therein).

The current analysis emphasizes the dynamical behavior of WC oscillators within a locally coupled, homogeneous system context. Despite the model's simplicity, disregarding elaborate topological and physiological detail, it manifests various dynamical behaviors. This repertoire encompasses equilibrium states, multistability, oscillations, traveling waves, and chaotic dynamics. The model produces coordinated spontaneous transitions between two distinct membrane potentials with the appropriate setup. It successfully accounts for many of the properties of active Up and quiescent Down states observed in corticothalamic networks during slow-wave sleep (SWS).

3.3 Background

3.3.1 Slow-Wave Sleep

SWS is critical in various other physiological functions. Aside from facilitating the process of memory consolidation (Plihal and Born, 1997, 1999, Girardeau et al., 2009), SWS

contributes to the regulation of metabolic processes (Iliff et al., 2012, Sharma et al., 2010), and fostering brain development and plasticity (Wilhelm et al., 2008). The glymphatic system, a waste clearance pathway in the brain, shows increased activity during SWS, suggesting a role in metabolic waste clearance (Åkerstedt and Nilsson, 2003, Fultz et al., 2019). Furthermore, the synaptic homeostasis hypothesis posits that SWS aids in the recalibration of synaptic strength, balancing the need for plasticity with the finite metabolic and structural limits of the brain (Tononi and Cirelli, 2014).

A quintessential feature of SWS lies in its regulatory mechanisms, specifically the spatiotemporal orchestration of synchronous EEG rhythms, including delta waves, spindles, and K-complexes. These components, characterized by their specific amplitude and frequency, are integral to non-rapid eye movement (NREM) sleep, marked by reduced physiological activity and sensory awareness (Steriade, 1997).

Stages 3 and 4 of NREM sleep are particularly significant due to their association with SWS, a state delineated by slow-wave activity (SWA) surpassing 75 microvolts and occurring at less than 1 Hz. This occurrence is non-trivial, accounting for approximately 10% to 20% of a robust young adult's sleep period (Carskadon and Dement, 2017).

There is heightened intrigue in comprehending the collective dynamics of corticothalamic networks that underlie slow oscillations during SWS. The slow oscillations are a macroscopic behavioral manifestation that engages a broad range of electrophysiological categories of cortical and thalamic neurons. They result in alternating periods of activity and silence within corticothalamic networks, spanning approximately one second (Steriade et al., 1993a,b).

Two distinct states demarcate these oscillations: the Up states, characterized by depolarization and intense synaptic activity (spanning both excitatory and inhibitory), and the Down states, typified by neuronal hyperpolarization and a relative quiescence (Figure 3.1). These oscillations are posited to originate intracortically. They can be observed *in vitro* in cortical preparations, as well as *in vivo* in deafferented pieces of the cortex (Steriade et al., 1993b, McCormick and Sanchez-Vives, 2000, Destexhe and Sejnowski, 2003). This intracortical origin hypothesis is further bolstered by suppressing these oscillations in the thalamus following cortical deafferentation (Timofeev and Steriade, 1996).

Contrastingly, more contemporary studies posit that the full manifestation of slow waves

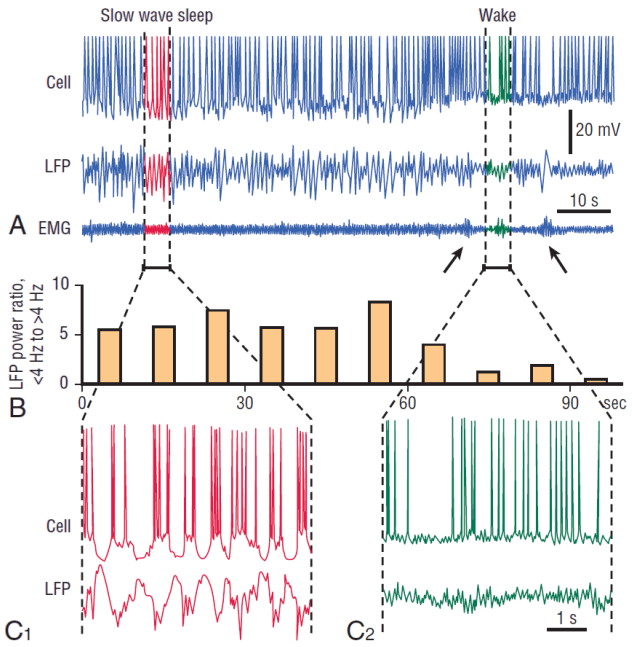


Figure 3.1: **Slow oscillations in local cortical field potentials (LFP) and the membrane potential of an adult cat's cortical neuron during non-rapid eye movement (NREM) sleep.** (A) Simultaneous intracellular, LFP, and electromyography (EMG) recording during sleep and wakefulness. The animal is in NREM sleep at the beginning of the recording, transitioning to an awake state after approximately 70 seconds (arrows indicate EMG activation). Action potentials are truncated in the intracellular recording. (B) Levels of delta power in the LFP are higher during NREM sleep than during waking. Plotted are 10-second bins of the ratio of spectral power ($<$ 4 Hz/ $>$ 4 Hz) recorded in the LFP. (C) Intracellular activity and LFP recording from (A) are shown at an expanded time scale. Note clear fluctuations of the membrane potential between Up and Down states during SWS (C_1) in association with the slow oscillation ($<$ 1 Hz) in the LFP. During wakefulness (C_2), the cell is tonically depolarized, and no sustained episodes of hyperpolarization are present. Adopted from (Mukovski et al., 2007).

during NREM sleep necessitates the involvement of both the cortex and thalamus (David et al., 2013). This assertion opposes the early findings of Steriade et al. (1993b). Despite

initial inhibition, slow wave activity within the deafferented cortical slab gradually recovers, suggesting a possible thalamic contribution (Lemieux et al., 2014).

The enduring phenomena of Up and Down states within corticothalamic networks are attributed to the intersection of inherent neuronal properties and the synaptic organization of these circuits (Steriade et al., 1991, McCormick and Sanchez-Vives, 2000, Destexhe and Sejnowski, 2003, Crunelli and Hughes, 2010). Each slow oscillation cycle can originate from any cortical site, cascading sequentially toward other cortical regions.

In adult humans, the initiation of activity can occur anywhere within the cortex but exhibits a predilection towards the frontal areas, from where it diffuses (Massimini et al., 2004). A similar propagation pattern is observable in feline subjects, where slow oscillations can originate anywhere within the cortex. However, activity preferentially appears at the frontier between areas 5 and 7, spreading anterolaterally (Volgushev et al., 2006).

An intriguing observation is the synchronous onsets of silent states, which occur even more synchronously than the onsets of activity (Volgushev et al., 2006). This suggests a large-scale synaptic mechanism capable of effecting long-range synchronization, thus facilitating the transition from activity to silence. However, the network mechanisms underlying these spontaneous state alternations remain largely enigmatic.

Pertinent questions arise from these observations. How does activity originate when the network is in a silent state? Does activity onset have a precise cortical location, or can it originate anywhere? Inversely, how and where does silence initiate? Moreover, perhaps the most intriguing question pertains to the observation that the onsets of silence appear even more synchronized and precise than the onsets of activity.

Among the manifold hypotheses endeavoring to decipher the origins of active states during SWS, three distinct propositions—namely, the propagation hypothesis, the synchronization hypothesis, and the integration hypothesis. Each conjecture embodies potential mechanistic insights, offering prismatic perspectives that shed light on the otherwise elusive nature of these states (Lüthi and McCormick, 1998, McCormick and Sanchez-Vives, 2000, Cossart et al., 2003, Timofeev et al., 2000). A meticulous examination of each hypothesis, including its conjectures and potential limitations, is required to untangle the convoluted enigma of active state genesis during SWS.

The *propagation hypothesis* postulates the existence of a specific set of neurons that pioneer the genesis of the active state, which subsequently propagates laterally with systematic site-specific onset delays. This hypothesis postulates that during silent states, h-currents, which activate when the membrane potential plunges beyond approximately -50 to -70 millivolts (Lüthi and McCormick, 1998), incrementally depolarize Layer V pyramidal neurons towards the Up state equilibrium potential, situated at approximately -30 millivolts (McCormick and Sanchez-Vives, 2000). In contradistinction, the *synchronization hypothesis* contends that a heterogeneous ensemble of neurons maintains the capacity for sporadic spontaneous firing even amidst the silent state. The harmonized orchestration of these neuronal ensembles directs the transition from silence to activity (Cossart et al., 2003).

Lastly, the *integration hypothesis* proposes the potential for the genesis of weaker activity to commence sporadically from multiple locations. This hypothesis suggests that the accumulation of spike-independent neurotransmitter release may precipitate depolarizing events. These events, in turn, could activate inward currents that depolarize specific neurons to the firing threshold. If sufficient events transpire within a postsynaptic integration period, additional neurons may become involved (Timofeev et al., 2000).

No singular hypothesis proposed thus far has proven entirely satisfactory in its explanatory power or congruity with empirical findings. While illuminating fragments of the giant puzzle, each postulation inevitably encounters contradictions or limitations when confronted with experimental results (Luczak et al., 2007, Timofeev et al., 2000, Cossart et al., 2003, McCormick and Sanchez-Vives, 2000). For instance, extracellular signal recording experiments have revealed unique spiking patterns for Layer V neurons during Up states, displaying a stereotypical sequential activation pattern. This pattern belies the supposition of randomized spontaneous firing synchronizing or integrating into activity, thereby bringing into question the tenability of the synchronization and integration hypotheses (Timofeev et al., 2000, Cossart et al., 2003). It instead bolsters the propagation hypothesis, which posits a well-localized origin with subsequent propagation (McCormick and Sanchez-Vives, 2000).

Nevertheless, the propagation hypothesis has its inconsistencies. The preferential location for the onset of activity varies across species (Massimini et al., 2004, Volgushev et al., 2006) and changes with age within species (Kurth et al., 2010), leading to a degree of incon-

gruity within the hypothesis itself. Furthermore, the propagation hypothesis hinges on the assumption of focal activity possessing sufficient intensity. However, the h-current in neocortical neurons is relatively feeble and unlikely to propagate activity to target cells effectively, thereby casting doubts on its candidacy as a primary mechanism (Timofeev et al., 2002).

To address these shortcomings and reconcile the seeming contradictions, Chauvette et al. (2010) advanced a hybrid mechanism of active state initiation, merging elements from both the propagation and integration hypotheses. This postulation acknowledges that local neocortical activity could originate in any neuron and at any depth. However, it emphasizes that the Layer V pyramidal cells, with their staggering synaptic count of 50,000-60,000—significantly surpassing any other cortical cell type (DeFelipe and Fariñas, 1992)—are more susceptible to initial activation in a stochastic setting of neurotransmitter release. The propensity for pyramidal cells in deep layers to intrinsically burst (Connors and Gutnick, 1990) enhances their potential to instigate Up states in their targets (Timofeev et al., 2000).

The complex choreography of active states during SWS involves a delicate equilibrium of excitation and inhibition, ensuring their steady maintenance (Shu et al., 2003, Haider et al., 2006, Beltramo et al., 2013, Lemieux et al., 2015). However, this dynamic ballet does not dance indefinitely, and there inevitably arises the question: *How do these active states conclude, thereby birthing silent states?* Previous computational models have implicated mechanisms such as activity-dependent synaptic depression (Holcman and Tsodyks, 2006, Kilpatrick and Bressloff, 2010) or spike adaptation through the activation of hyperpolarizing intrinsic currents (McCleney and Kilpatrick, 2016) in orchestrating the onset of silence.

However, the work of Volgushev et al. (2006) introduced a twist in the tale by demonstrating that the onset of silence manifests more synchronously than its termination. This revelation instigates a reevaluation of traditional understanding about state alternation, as no cell-level mechanism, such as firing adaptation or synaptic depression, can independently account for the synchronous onset of silence. This is because these processes would imply the onset of silence at diverse times across different neurons. Consequently, the puzzle begs for a network mechanism that could effectuate a switch from activity to silence, likely accomplished via active inhibition (Steriade et al., 1993a).

This theoretical void has been filled with two contending hypotheses for the onset of

silence from activity, and both are premised on a transient imbalance between excitation and inhibition via the disfacilitation of subpopulations of excitatory neurons (Beltramo et al., 2013, Lemieux et al., 2015). The *pacemaker hypothesis* proposes that cortical regions are subject to direct, widespread inhibitory input from extracortical areas, suggesting a global signal cues the Down state (Lemieux et al., 2014). In contrast, the *disynaptic relay hypothesis* postulates a more nuanced mechanism wherein widespread excitatory inputs synchronously recruit a subpopulation of cortical inhibitory interneurons that hyperpolarize a subset of cortical neurons below the firing threshold. The dampened activity of these hyperpolarized neurons attenuates the excitatory drive on their targets, thereby sparking a domino effect of disfacilitation that eventually silences the entire network (Lemieux et al., 2015).

Notwithstanding these hypotheses, the conundrum persists; the network mechanisms underlying the synchronized onsets of silence in corticothalamic neurons during SWS remain enigmatic. In sum, the genesis of active cortical states during sleep is predicated on spontaneous, spike-independent release, with a propitious bias towards layer V pyramidal neurons. This predilection stems from their substantial synaptic inputs and expansive divergence of outputs, engendering an environment ripe for initiating active states (Chauvette et al., 2010). On the other hand, the denouement of these states can be ascribed to network mechanisms predicated on active inhibition.

Two prominent theories illuminate this network dynamics. The first, intracortical in its machinations, posits that excitatory neurons could recruit a subset of inhibitory neurons via disynaptic pathways. This, in turn, would instigate a cascade of disfacilitation, essentially nudging the network towards a state of silence (Lemieux et al., 2015). The second theory invokes a more global mechanism, suggesting that widespread inhibitory inputs from extracortical areas could act as the proverbial "handbrake," arresting cortical activity and thus heralding the onset of silence (Lemieux et al., 2014).

Together, these observations provide a robust yet evolving framework for understanding the orchestration of active and silent states during sleep. However, much like the cortical states, our comprehension of these mechanisms remains fluid.

3.3.2 Neural Oscillator

Neural oscillators, believed to play a crucial role in generating oscillating activity in the brain, are prevalent throughout the neural system (Rakic, 1975). As the fundamental unit of our network, the neural oscillator model is defined such that the excitatory neurons' population generates action potentials that excite the inhibitory neuron population. This inhibitory population then counteracts with inhibition, forming a feedback loop between the excitatory and inhibitory units (see Figure 3.2).

$$\tau_x \frac{dx_j}{dt} + x_j = F_x(w_{ee}x_j + w_{ie}y_j + T_j + \sigma_x\gamma_j), \tag{3-1a}$$

$$\tau_y \frac{dy_j}{dt} + y_j = F_y(w_{ei}x_j + w_{ii}y_j + S_j + \sigma_y\eta_j), \tag{3-1b}$$

$$F_s(u) = \frac{1}{1 + e^{-\beta_s(u-\theta_s)}}, \qquad s \in \{x, y\}, \tag{3-1c}$$

The fractions of firing excitatory and inhibitory neurons are denoted by x_j and y_j, respectively. The variables $\tau_x, \tau_y > 0$ are their corresponding membrane time constants. The inputs T_j and S_j represent the aggregate currents from external sources (e.g., sensory organs, other brain regions, implanted electrodes) reaching the jth excitatory and inhibitory neurons, respectively. The parameters w_{ie} and w_{ei}, known as synaptic, signify inhibitory-excitatory and excitatory-inhibitory interconnections coupling the ith neural oscillator populations, while w_{ee} and w_{ii} denote feedback parameters, representing excitatory-excitatory and inhibitory-inhibitory interconnections within the jth neural oscillator. The random variables γ and η are normally distributed with zero mean, unit variance, and amplitudes σ_x and σ_y, respectively, introduced to account for the intrinsic neuronal differences and to encourage spontaneous state switching. The nonlinearity F_s is a logistic function with threshold θ_s and scaling parameter β_s, where $s \in x, y$. The function F_s can be interpreted as a firing probability rather than an actual firing rate, adopting a logistic function.

First, it is worth noting the clear distinction between the neurons in the model: they are either excitatory or inhibitory, a classification that harkens back to "Dale's principle" (Dale, 1935, Eccles, 1976), a helpful framework in mathematical neuroscience. This principle postulates that a neuron exerts the same physiological action at all its targets; that is, all synapses of a neuron are either excitatory or inhibitory.

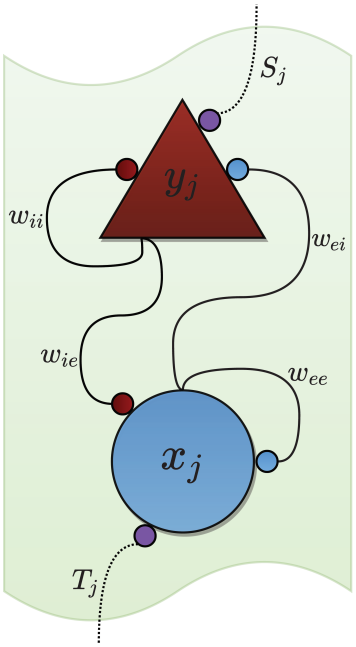


Figure 3.2: **A neural oscillator model.** It consists of two units, an excitatory x_i (blue circle) and an inhibitory y_i (red triangle), forming a feedback loop. Labels for units inputs and outputs correspond to Eqs. (3-1a) and (3-1b). Small blue circles denote excitatory synaptic connections, red triangles inhibitory synaptic connections, and purple circles represent connections that could have excitatory or inhibitory effects. Dotted curves indicate extra-cortical inputs and solid one's cortico-cortical connections.

The system (3-1) fundamentally parallels the seminal work done by Wilson and Cowan in their model (Wilson and Cowan, 1972, 1973). However, the distinction lies in the approach to incorporating refractoriness: the WC model does this by premultiplying the firing rates, denoted as F_x and F_y, by factors of $(1 - r_x x_j)$ and $(1 - r_y y_j)$ respectively, a feature absent in the system above. The introduction of these terms acknowledges the refractory dynamics encompassing both neuronal populations. They signify the proportion of neurons that retain the potential to fire, contingent on their refractory periods, represented by r_x and r_y.

Nevertheless, it is noteworthy that these multiplicative factors, despite contributing to the complexity of the model, do not significantly impact the analysis of the equations. The

effects these factors introduce can effectively be compensated for by rescaling the parameters of the nonlinearities F_x and F_y (Pinto et al., 1996, Curtu and Ermentrout, 2001). Consequently, without loss of generality, assigning values of zero to r_x and r_y is warranted.

Additionally, the firing rates F_x and F_y can be considered as modified instances of the conventional logistic function, with the specific property that $F_x(0) = F_y(0) = 0$. The terms x_j and y_j pertain to the low-pass filtered versions of the proportions of active excitatory and inhibitory cells. This modification has the aggregate effect of eliminating oscillatory components from neural responses that exceed 100 Hz.

Despite its seemingly minimalist representation, the (3-1) model can emulate numerous phenomena observed in actual neural systems. This includes firing rates, bistability, and varied oscillatory behaviors. Notably, these myriad phenomena are captured despite the model being an extreme abstraction of the rich neurodynamics inherent to cortical regions.

3.3.3 Phase Portraits

The seminal work of Beer (1995) undertook a comprehensive analysis of two-population networks, which was subsequently expanded by Ermentrout (1998b) to accommodate gaps in Beer's assessment. These inquiries illuminate that the entire conduct of two-cell neural networks can be systematically classified and predicted. Specifically, if the function F_s, where s is a member of the set x, y, exhibits a monotonic nature (such as squashing functions akin to F_s) and the weight of the connections adhere to either $w_{ei}w_{ie} > 0$ or $w_{ee}, w_{ii} < 0$, there are no limit cycles (oscillations) and merely fixed points in the WC model (Eq. (3-1)). As a result, comprehensive qualitative dynamical analysis can be performed by exploring nullcline intersections; these are curves on the x–y plane characterized by $dx/dt = dy/dt = 0$. It is noteworthy that the intersections of nullclines correspond to fixed points.

Our interest primarily lies in the dynamics of cortical slow oscillations, leading us to consider neural oscillators that can present limit cycles. This condition is met when both self-excitation ($w_{ee} \geq 0$ and $w_{ii} \leq 0$) and a negative feedback loop between the two populations ($w_{ie} \geq 0$ and $w_{ei} \leq 0$) exist. Figure 3.3 graphically portrays a typical limit cycle on the x–y plane. In this diagram, time progresses counter-clockwise along the limit cycle. The phase

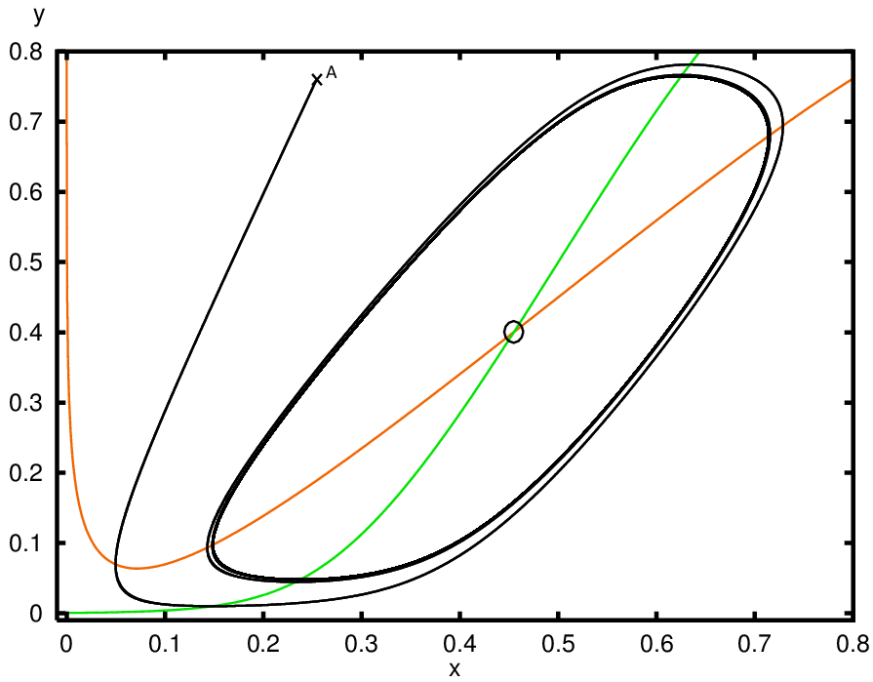


Figure 3.3: **Stable limit cycle in the single neural oscillator model** (3-1)**.** The orange curve represents the x-nullcline ($dx/dt = 0$) and the green one represents the y-nullcline ($dy/dt = 0$). There is a single fixed point, labeled (◯), where the nullclines intersect: an unstable focus. The oscillator starts at the upper-left corner, marked by (□), and quickly converges to the limit cycle solution. The parameters for this simulation are: $\tau_x = 1$, $\tau_y = 0.7$, $w_{ee} = 15$, $w_{ie} = -10$, $w_{ei} = 20$, $w_{ii} = -5$, $T = -3$, $S = -7.5$, $\sigma_x = \sigma_y = 0$, $\theta_x = \theta_y = 0$ and $\beta_x = \beta_y = 1$.

plane is divided into four regions by the x– and y–nullclines (illustrated as orange and green curves), wherein dx/dt and dy/dt exhibit different signs. The oscillator initiates in the upper-left region, where x and y decrease ($dx/dt < 0$ and $dy/dt < 0$) until the trajectory intersects the x–nullcline. Post-intersection, x begins to increase ($dx/dt > 0$), while y continues to decrease ($dy/dt < 0$). This is a consequence of the inhibitory population's activity reducing to such an extent that it ceases to suppress the firing of the excitatory neurons.

As the trajectory enters the lower-right region ($dx/dt > 0$ and $dy/dt > 0$), x and y increase. At this stage, a sufficient number of excitatory neurons are active and commence

the recruitment of the inhibitory population. When the trajectory reaches the upper-right region, sufficient inhibitory neurons are active to inhibit the excitatory population ($dx/dt < 0$ and $dy/dt > 0$). Finally, when the balance between excitation and inhibition is disrupted once more, y decreases, and the cycle begins anew, akin to the predator-prey relationship described in the Lotka-Volterra equations (Hoppensteadt, 2006).

The intersection of the x- and y-nullclines, denoted by a small circle at the center of Figure 3.3, represents an equilibrium corresponding to an unstable focus. Inferring the number and location of these equilibria from the analysis of Eq. (3-1) can be challenging. However, once the nullclines are identified, the task becomes geometrically straightforward. We used the Euler method to numerically solve the equations in subsequent simulations, with a time step of $\Delta t = 0.05$.

Eq. (3-1) supports asymptotic behaviors apart from stable oscillations. By merely varying the external inputs T_j and S_j, we can generate a phase portrait that is qualitatively distinct from the one illustrated in Figure 3.3. Figure 3.4 depicts a qualitatively different portrait of two stable equilibria at different membrane potentials and a saddle point. Saddle points play a crucial role as their stable manifolds form separatrices, which divide the plane into the domains of attraction for multiple stable fixed points.

But the neural oscillator supports yet another qualitatively different dynamic: the coexistence of Up and Down as stable states, *i.e.*, bistability. The phase portrait of this scenario is illustrated in Figure 3.5a. In a striking departure from the scenario portrayed in Figure 3.4, where two stable resting equilibria existed at the subthreshold membrane potential for firing, in the present case, the upper-right node transformed into a stable focus. This focus, near the point $(0.35, 0.25)$ as seen in Figure 3.5a, is characterized by an attenuated oscillation encircling it. Therefore, introducing a slight noise factor can trigger spontaneous transitions between Up and Down states, a phenomenon illustrated in Figure 3.5b[1].

From a dynamic systems perspective, these qualitative changes in the phase portrait correspond to bifurcations of neural oscillator dynamics. FitzHugh (1961) pioneered local bifurcation analysis in neuroscience, followed by Ermentrout and Cowan (1979) in the analy-

[1]The execution of this and the simulations in the next section relied on the XPP software package, a robust tool often used in such computational neuroscience contexts (Ermentrout, 2002).

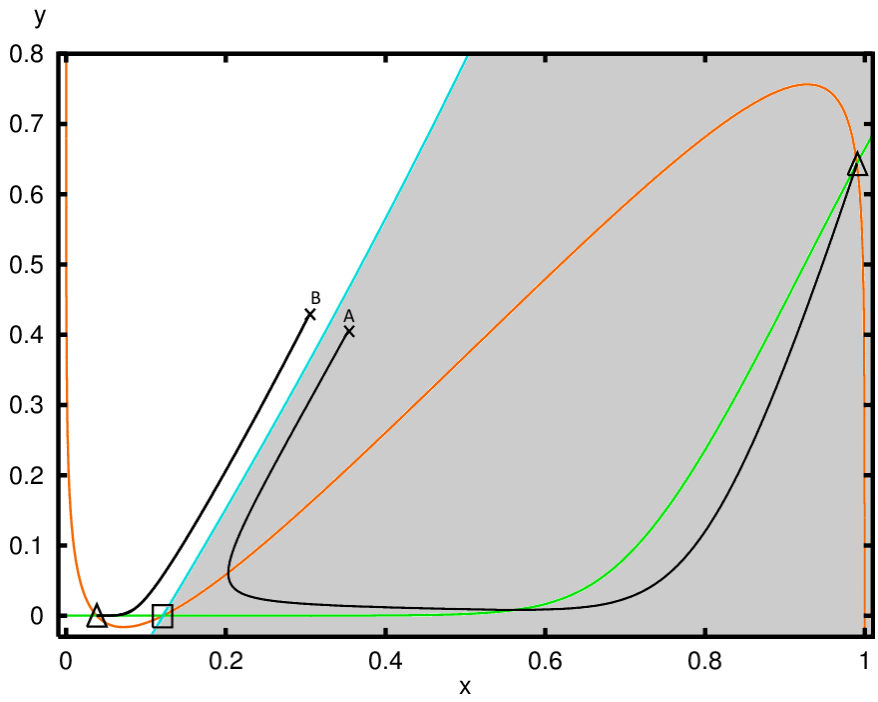
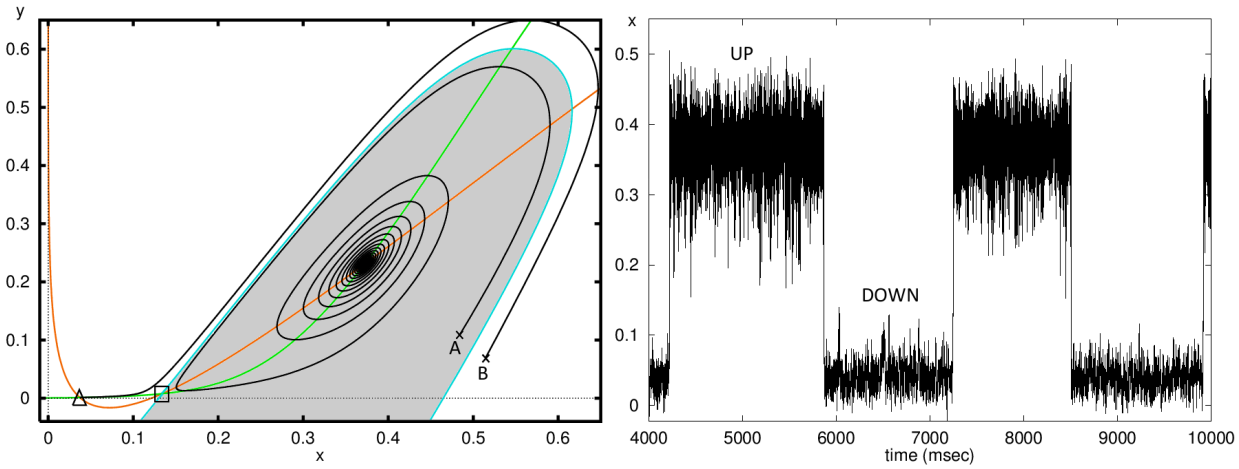


Figure 3.4: **Multiple dynamical attractors in the single neural oscillator model** (3-1)**.** The fixed points labeled (△) are stable nodes and the one labeled (□) is a saddle. As before, the orange curve represents the x–nullcline, and the green one represents the y–nullcline. The blue curve is the stable manifold of the saddle point; it divides the phase plane into two regions: the gray region is the attraction domain of the stable node in the upper-left corner, and the complementary region is the attraction domain of the stable node in the lower-left corner. Hence, trajectories to the left of the stable manifold will converge to the lower-left stable node (*e.g.*, trajectory starting in point A), and trajectories to the right of the stable manifold will converge to the upper-right stable node (*e.g.*, trajectory starting in point B). Parameters are the same as in Figure 3.3, except $T = -3.8$ and $S = -16$.

sis of the WC equations. For a comprehensive bifurcation analysis of the WC equations, the reader is referred to Borisyuk and Kirillov (1992) and Hoppensteadt and Izhikevich (1997).



(a) (b)

Figure 3.5: **Bistability and oscillation in the single neural oscillator model** (3-1)**.** (a) The fixed point labeled (△) is a stable node corresponding to the Down state, the fixed point labeled (□) is a saddle, and the fixed point located near the point $(0.35, 0.25)$ is an unstable focus corresponding to the Up state. As before, the orange curve represents the x-nullcline, the green one represents the y-nullcline, and the blue curve is the stable manifold of the saddle point. The gray region is the attraction domain of the Up state, and oscillators starting there (*e.g.*, the trajectory beginning in A) converge to the limit cycle. Conversely, oscillators starting in the complementary region (*e.g.*, the trajectory beginning in B) converge to the resting state. The parameters for this simulation are: $\tau_x = 1$, $\tau_y = 0.7$, $w_{ee} = 15$, $w_{ie} = -10$, $w_{ei} = 20$, $w_{ii} = -5$, $T = -3.8$, $S = -7.5$, $\sigma_x = \sigma_y = 0$, $\theta_x = \theta_y = 0$ and $\beta_x = \beta_y = 1$. (b) A simulation of (3-2) when additive white noise is added to the inputs, showing spontaneous state switching. Parameters as in (a) with $\sigma_x = \sigma_y = 0.015$.

3.3.4 Weakly Interconnected Oscillatory Networks

Intriguingly, the weak interconnectivity among neural oscillators does not impede their oscillatory patterns (Hoppensteadt and Izhikevich, 1997). We repurpose the model (3-1) into an array encompassing N neural oscillators to investigate further the properties inherent to

a network of such weakly linked oscillators. This array has a unique configuration, featuring nearest-neighbor coupling with periodic boundary conditions (refer to Figure 3.6).

$$\tau_x \frac{dx_j}{dt} + x_j = F_x(w_{ee}\overline{x}_j - w_{ie}\overline{y}_j + T_j + \sigma_x\gamma_j), \tag{3-2a}$$

$$\tau_y \frac{dy_j}{dt} + y_j = F_y(w_{ei}\overline{x}_j - w_{ii}\overline{y}_j + S_j + \sigma_y\eta_j), \tag{3-2b}$$

$$F_s(u) = \frac{1}{1 + e^{-\beta_s(u-\theta_s)}}, \qquad s \in x, y, \tag{3-2c}$$

The equations are parameterized as per the following conditions:

$$\overline{x}_j = \begin{cases} (1-c_x)x_j + (c_x/2)[x_{j+1} + x_N] & \text{if } j = 1, \\ (1-c_x)x_j + (c_x/2)[x_{j+1} + x_{j-1}] & \text{if } 1 < j < N, \\ (1-c_x)x_j + (c_x/2)[x_1 + x_{j-1}] & \text{if } j = N, \end{cases} \tag{3-3}$$

Correspondingly,

$$\overline{y}_j = \begin{cases} (1-c_y)y_j + (c_y/2)[y_{j+1} + y_N] & \text{if } j = 1, \\ (1-c_y)y_j + (c_y/2)[y_{j+1} + y_{j-1}] & \text{if } 1 < j < N, \\ (1-c_y)y_j + (c_y/2)[y_1 + y_{j-1}] & \text{if } j = N. \end{cases} \tag{3-4}$$

The parameters c_x and c_y, nonnegative by definition, indicate the network's coupling strength between neighboring excitatory and inhibitory populations.

3.4 Network Mechanisms behind Slow-Wave Oscillations

We will use the network defined by (3-2)–(3-4) to explore the genesis of slow oscillations. The network includes $N = 50$ WC neural oscillators with the parameters set to $\tau_x = 1$, $\tau_y = 0.7$, $w_{ee} = 15$, $w_{ie} = -15$, $w_{ei} = 10$, $w_{ii} = -5$, $T = -3.75$, $S = -5.4$, $\sigma_x = \sigma_y = 0.02$, $\theta_x = \theta_y = 0$, $\beta_x = \beta_y = 1$, $c_x = 0.3$, and $c_y = 0.11$. Thus, we assume that each neural oscillator operates in a setting where Down and Up states coexist, as illustrated in Figure 3.5a. This, however, does not imply that isolated oscillators present bistability.

Probing for indications of Down to Up and Up to Down state transitions propagating through the network is particularly enlightening. Our point of departure was a network devoid of activity, with activity variables $x_j = y_j = 0$ for $1 \le j \le 50$. After that, the system could evolve per Eq. (3-2) for up to 100 seconds. The period was divided into bins of size $\Delta t = 0.05$ ms for a detailed examination.

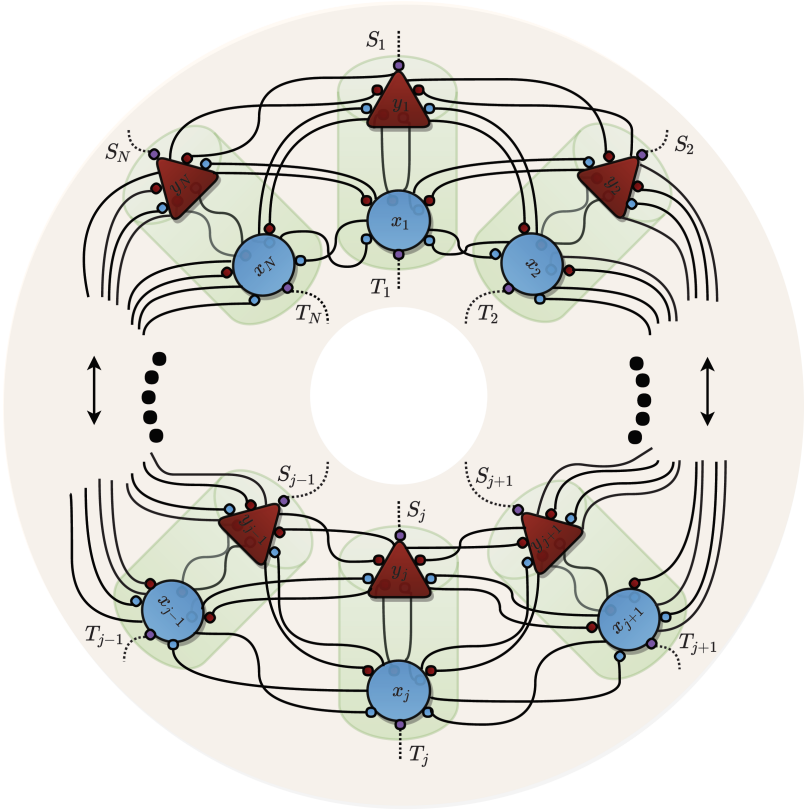


Figure 3.6: **Periodic array of neural oscillators with nearest-neighbor coupling.** An array of N neural oscillators, as in Figure 3.2, each featuring nearest-neighbor coupling in a ring configuration - that is, x_0 is identical to x_N, x_{N+1} corresponds to x_1, and y_0 and y_{N+1} align with y_N and y_1, respectively. Excitatory synaptic connections are represented by small blue circles, inhibitory connections by red circles, and purple circles symbolize connections that could be either excitatory or inhibitory. Dotted lines indicate extra-cortical inputs, while solid lines demonstrate cortico-cortical connections. While intra-oscillator connections are only partially displayed for simplicity, they should be considered existing, as per Figure 3.2. Similarly, connection weights are not depicted due to brevity. Lastly, the wiring configuration reflects the coupling and boundary conditions stated in Eqs. (3-2–3-4).

3.4.1 Emergence of Slow Oscillations

Macroscopically, the network generates spontaneous oscillations (Figure 3.7a). The phenomenon is quite robust, and the sharp state transitions imply that the population of neurons transitions between Up and Down states synchronously (Figure 3.7c). This resembles the observations by Volgushev et al. (2006) in the cat neocortex during SWS; likewise, we also see a robust long-range correlation of oscillators during each activity regime and, more significantly, during state transitions.

Interestingly, the irregular pattern of activity suggests the network's inherent heterogeneity; not all Up and Down states are identical, and the duration in each state varies. This irregularity indicates an intrinsic dynamism at each state transition. This dynamic behavior may contribute to the network's robustness and adaptability, allowing it to respond to varying inputs and conditions, a salient feature crucial to any physiological system.

Figure 3.7c further unravels the critical role played by the network's inherent nonlinearities F_s and population activity in these state transitions. The silent phase (bluish regions) reflects a lack of activity across the network, supported by the individual oscillators' intrinsic bistability rather than active inhibition. However, the onset of widespread firing (reddish regions) is triggered by an increase in the activity of a subset of excitatory populations due to background noise, leading to a cascading recruitment of neighboring oscillators.

Subsequently, as inhibitory populations' activity escalates, the balance between excitation and inhibition tilts, propelling the excitatory populations towards their stable manifolds and consequently returning the network to its silent state. This termination process intriguingly sidesteps neuronal adaptation, with slow oscillations engendered solely through a strategic interplay between excitation and inhibition. The oscillation frequency and state duration are largely dictated by the distribution of active populations and the overall degree of coupling, a finding that further underscores the network's complex multi-scale coordination.

3.4.2 Noise Modulation

To investigate how stochastic fluctuations modulate the behavior of the network, specifically how they influence the active population distribution during the Up and Down states,

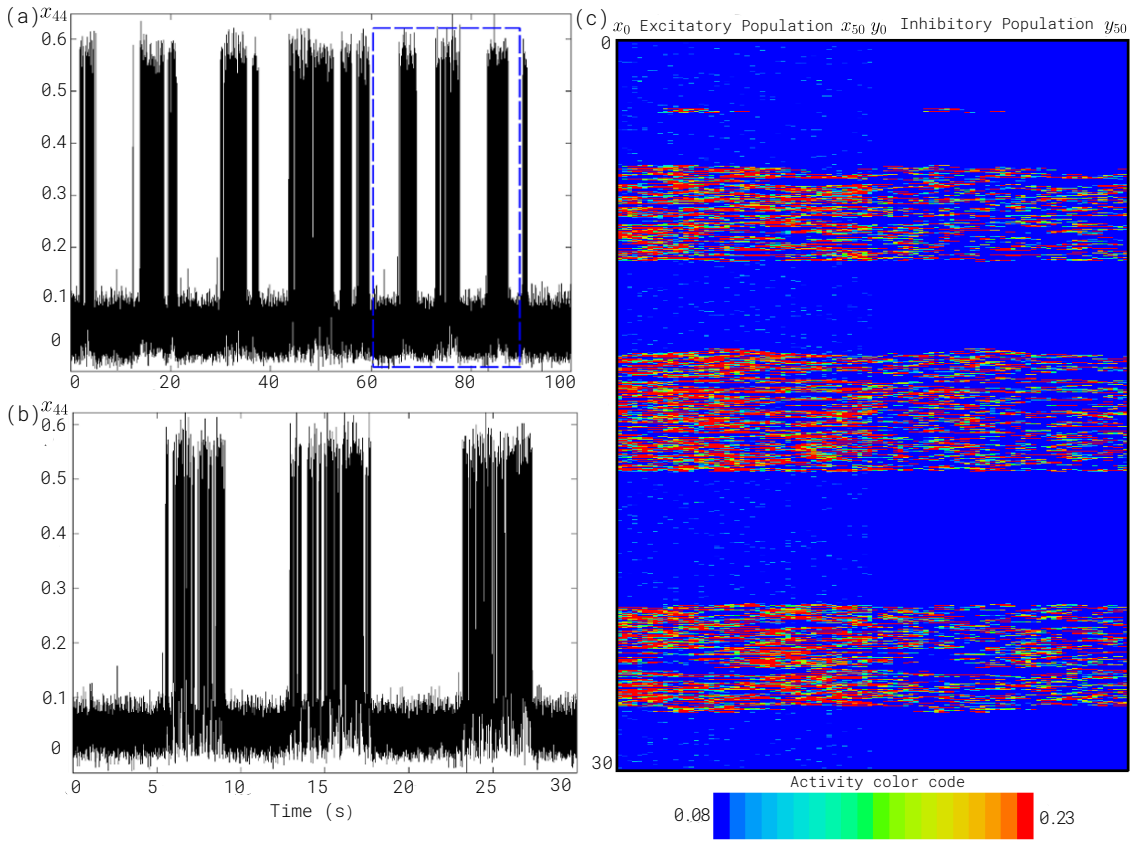


Figure 3.7: **Emergence of slow-wave oscillations in the neural oscillator network.** (a) Activity of the x_{44} excitatory population over 100 seconds. (b) Expanded x_{44} population activity corresponding to the highlighted blue box in (a) over the interval from $t = 60$ to $t = 90$ seconds. (c) The time-space plot of the activity of the entire network of oscillators over the interval from $t = 60$ to $t = 90$ seconds. Time "runs downwards vertically," and oscillators are aligned on the horizontal axis. Excitatory (left region) and inhibitory populations (right region) were separated for easy visualization. The transition from one state to the other was defined to occur at $(0.23, 0.08)$, according to the location of the saddle point in Figure 3.5a and accounting for the effects of noise. Blue activity denotes silence, and red sustained firing. Noise amplitude: $\sigma_x = \sigma_y = 0.02$; degree of coupling: $c_x = 0.3$, and $c_y = 0.11$.

we methodically manipulated the σ_y parameter while keeping σ_x invariant. The network model delineated in Figure 3.7 served as our foundational benchmark for these explorations.

Figure 3.8 shows a network where σ_y was curtailed by 20%. The upper trace represents the activity of the x_{44} excitatory population over 100 seconds, while the rest of the figure provides a detailed account of the network behavior within the initial 30 seconds. The frequency of state transitions dwindled after reducing the inhibitory noise component. For instance, a single, approximately 5-second-long transition to the Up state can be observed within the first 30 seconds. However, this state was observed three times in the original scenario (Figure 3.7), each time of a different, and generally longer, duration.

On the contrary, elevating σ_y by 20% generally truncated the durations of both Up and Down states, as demonstrated in Figure 3.9. For example, a significantly greater number of state transitions can be observed within the first 30 seconds (Figure 3.9b), compared to the corresponding spans in Figures 3.7 and 3.8. Several transitions occur even before the entire network has the chance to engage in the activation of all populations, particularly evident in the initial 10 seconds of Figure 3.9c.

Intuitively, mitigating the influence of noise results in fewer transitions and a reduction in the duration of states due to the diminished probability of threshold crossings. Conversely, an increase in noise augments the frequency of state transitions and generates short-lived states due to the accentuated number of threshold crossings. When the noise amplitude becomes excessively large, as depicted in Figure 3.10b, the network transits states faster than the time requisite to engage all oscillators, leading to epochs of meta-activity, where only isolated subsets of populations are active.

3.4.3 Coupling Modulation

It is naturally anticipated that the intensity of the coupling will determine the frequency of state changes and the duration that the network persists in each state. To scrutinize the modulatory effect of coupling on the network's behavior, we manipulated the coupling degree among the excitatory populations, denoted as c_x, while keeping c_y constant. Initially, based on our fundamental network model, we studied the consequences of reducing the coupling.

Figure 3.10a shows a time-space plot of a network in which c_x was decreased by 50%. The widespread silence observed throughout 100 seconds substantiates our preliminary ob-

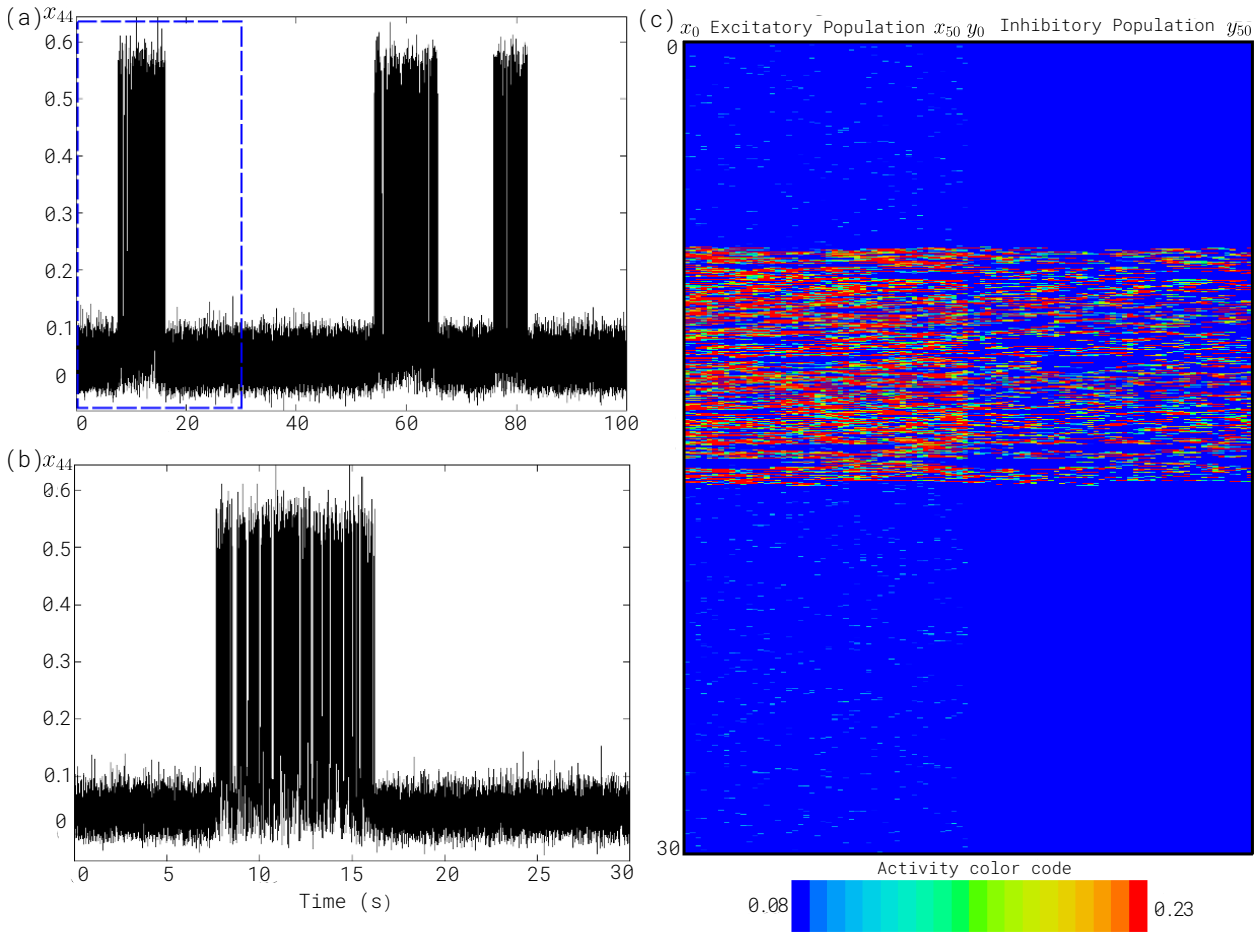


Figure 3.8: **Effects of decreasing the inhibitory noise in the neural oscillator network.** (a) Activity of the x_{44} excitatory population over 100 seconds. (b) Expanded x_{44} population activity corresponding to the highlighted blue box in (a) over the time interval from $t = 0$ to $t = 30$ seconds. (c) Time-space plot of the activity of the entire network of oscillators over the time interval from $t = 0$ to $t = 30$ seconds. Noise amplitude: $\sigma_x = 0.02$, $\sigma_y = 0.016$; degree of coupling: $c_x = 0.3$, $c_y = 0.11$.

servation that the Down states are sustained due to the absence of activity rather than due to active inhibition. Here, c_x was diminished to below what was necessary to couple the remaining active excitatory populations (as depicted by sparsely scattered red dots throughout the image), thereby obstructing activity propagation across the entire network. Essentially,

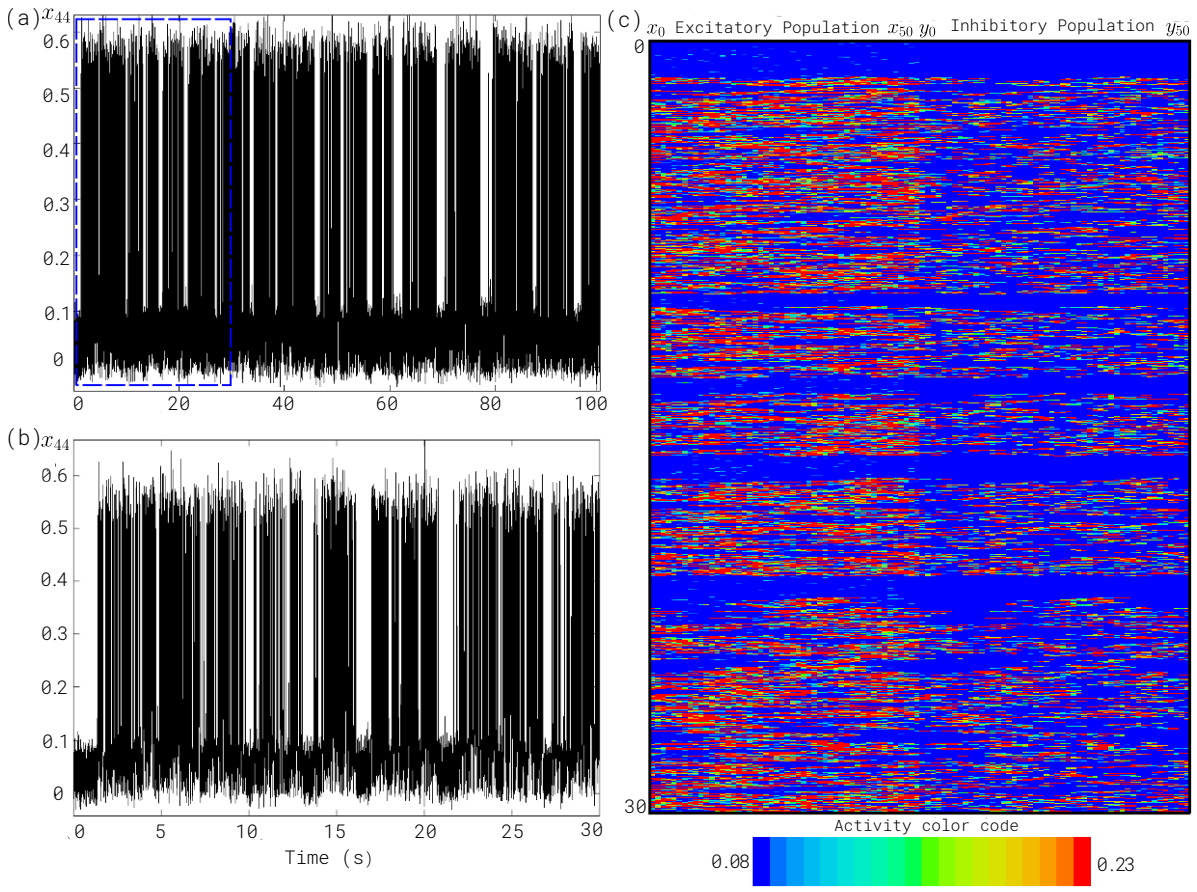


Figure 3.9: **Effects of increasing the inhibitory noise in the neural oscillator network.** (a) Activity of the x_{44} excitatory population over 100 seconds. (b) Expanded x_{44} population activity corresponding to the highlighted blue box in (a) over the time interval from $t = 0$ to $t = 30$ seconds. (c) The time-space plot of the activity of the entire network of oscillators over the time interval from $t = 0$ to $t = 30$ seconds. Noise amplitude: $\sigma_x = 0.02$, $\sigma_y = 0.024$; degree of coupling: $c_x = 0.3$, $c_y = 0.11$.

we rendered the Up state unstable by lessening the coupling.

We then proceeded to increase c_x by 50%. As evident from Figure 3.10b, after roughly 10 seconds, the network transitions from the Down to the Up state and persists for the remaining 90 seconds. In this instance, the imbalance between excitation and inhibition becomes apparent. However, the uncontrolled excitation among the inhibitory populations

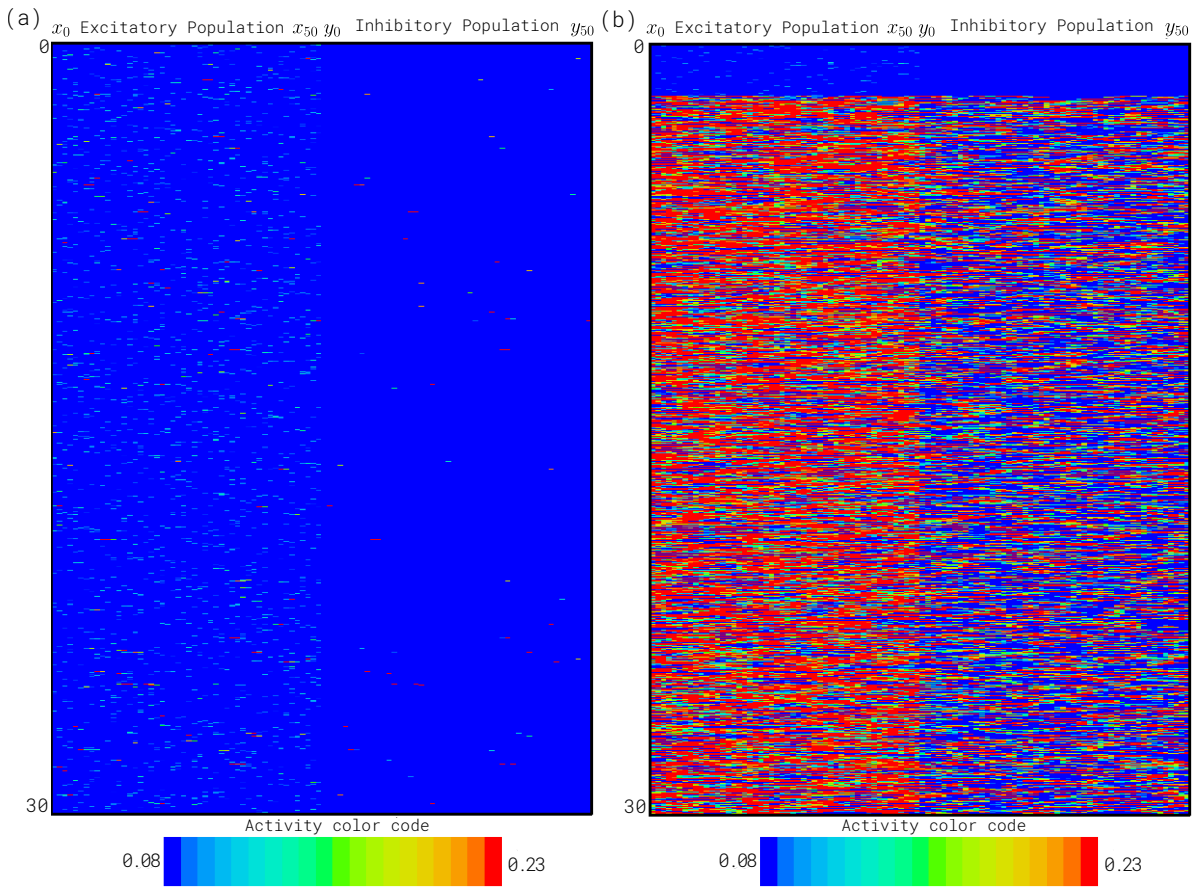


Figure 3.10: **Effects of varying the degree of coupling in the neural oscillator network.** (a) Time-space plot showcasing the activity of the entire network of oscillators over 100 seconds. Noise amplitude: $\sigma_x = \sigma_y = 0.02$; degree of coupling: $c_x = 0.15$, and $c_y = 0.11$. The network remains silent for the entire duration. (b) Time-space plot of the activity over 100 seconds showing sustained firing post Down to Up transition. Noise amplitude: $\sigma_x = \sigma_y = 0.02$; degree of coupling: $c_x = 0.45$, and $c_y = 0.11$.

drives the excitatory ones towards their stable manifolds; the heightened degree of coupling fortifies the Up state sufficiently to withstand inhibitory onslaughts. In this context, the Down state is destabilized by the escalating coupling among the excitatory populations.

3.5 Bifurcation Analysis

Identical patterns of silence and sustained firing can be achieved by increasing or decreasing the degree of coupling among the inhibitory populations. Consequently, the origin of slow oscillations may be comprehended as a combination of excitation and inhibition alone, with the relative imbalance of activity among the populations causing the destabilization of either the Up or the Down state. In this respect, c_x and c_y can be perceived as bifurcation parameters and, as such, are accountable for the qualitative shifts in behavior as depicted in Figures 3.7c, 3.10a, and 3.10b.

Stable Down and Up states coexistence can be understood as the system operating near a saddle-node bifurcation point. A saddle-node bifurcation corresponds to creating or annihilating a pair of fixed points, an essential process in our model for generating the bistability between Up and Down states. The spontaneous transitions between them under stochastic noise can be seen as stochastic perturbations driving the system across the bifurcation point.

However, stochastic driving cannot account for the observed rhythmic transition between high and low activity when there is a balance between excitation and inhibition. In this scenario, a Hopf bifurcation, creating or annihilating a limit cycle surrounding a fixed point, can account for spontaneously generated rhythmic transition between Up and Down states.

In the context of our model, the Bogdanov-Takens bifurcation, a type of codimension-two bifurcation, characterized by the coincidence of a saddle-node and a Hopf bifurcation, provides an elegant mathematical framework for describing the mechanism of the Up and Down state transitions. Specifically, the model parameters c_x and c_y control the proximity to the bifurcation point, modulating the bistability and the frequency of state transitions.

3.6 Conclusions

Our investigation has been focused on the influence of stochastic perturbations and the strength of coupling within a straightforward population model, designed to emulate the slow oscillations witnessed in cortical networks during SWS. This network comprised WC

neural oscillators, with neighboring oscillators coupled and with periodic boundary conditions applied. The oscillators functioned in a bistable region, that is, a region where stable quiescent and excited states coexist. State transitions were induced by infusing independent white noise into the neuron subpopulations.

Unlike previous modeling studies (Bazhenov et al., 2002, Holcman and Tsodyks, 2006, Parga and Abbott, 2007, Duc et al., 2015, McCleney and Kilpatrick, 2016), our model does not resort to any underlying systematic slow process like firing-rate adaptation or synaptic depression to mechanistically explain the stochastic origin of active states and slow oscillations. Instead, the onsets of silence are mediated by a network mechanism per the disynaptic relay hypothesis, where inhibition following excitation induces a chain reaction of disfacilitation terminating activity (Lemieux et al., 2015)

The collective dynamics of the network exhibited spontaneous and synchronized state transitions. Depending on the degree of coupling between the inhibitory and excitatory populations, we observed complete silence, sustained widespread activity, or rhythmic oscillations. Additionally, the amplitude of the noise was found to modulate the frequency and duration of states during slow oscillations. The interplay between the saddle-node and Hopf bifurcations in the Bogdanov-Takens bifurcation, as we vary the coupling parameters c_x and c_y, parsimoniously accounts for the rich tapestry of dynamics we observed. Overall, These modeling results demonstrate how the complex coordination of excitation and inhibition can spontaneously generate rhythmic transitions between Up and Down states.

4.0 Associative Memory and Temporal Sequence Learning

4.1 Chapter Summary

This chapter establishes an equivalence between two seminal neural network models—Hopfield networks and spike-coding networks—that have largely been studied independently. Hopfield networks rely on fixed point attractors to store memories, while spike-coding networks emphasize precise spike timing for temporal processing. We prove these models can be unified under the same mathematical framework using circulant matrices to represent asymmetric weights in Hopfield networks. This introduces dynamics that enable complex spike sequences generated from the stored attractor states. Our analysis combines two central, yet disconnected models in neural computation, providing a unified lens for understanding the interlinked principles of associative memory and temporal sequence learning in recurrent neural networks. Specifically, we demonstrate that: (1) Stable states in Hopfield networks provide a substrate for generating spike sequences in spiking networks; (2) Asymmetric weights implemented via circulant matrices induce dynamics between stored attractor states; (3) The two models are mathematically equivalent under this formulation. Overall, this work elucidates how neural networks can learn spatial and temporal patterns, underscoring the importance of jointly considering associative and dynamic aspects in modeling biological neural computation.

4.2 Background

4.2.1 Classical Hopfield Networks

John Hopfield pioneered formulating a novel paradigm in the landscape of associative memories by introducing Hopfield Networks in 1982 (Hopfield, 1982). The core concept encapsulates a particular type of associative memory wherein the recurrent synaptic con-

nections between neurons offer a scaffold for a Lyapunov function, synonymously termed an energy function, to preside over the temporal dynamics of the neuron states. Such a framework enables the activity of the neurons—the state variables, to undergo a monotonically non-increasing evolution towards a local minimum of this Lyapunov function, these terminal points characterizing the fixed-point attractors of the neural network dynamics.

The mechanics of the Hopfield networks, or their computation functionality, are orchestrated by the temporal evolution of their state. A binary state architecture was proposed, evolving in discrete time intervals. This dynamical process can be elegantly encapsulated through the following mathematical expression:

$$v_i[t+1] = \begin{cases} 1, & \text{if} \quad \sum_{j=1}^{N} J_{ij} v_j[t] + I_j > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4-1}$$

In this model, $v_i[t]$ signifies the activity or state of the i-th neuron, with the total number of neurons in the network denoted by N. I_j represents the bias input impinging upon neuron j, and J_{ij} encapsulates the synaptic weight associated with the connection from neuron j to neuron i. At each time step, a neuron is selected in a stochastic manner, and the ensuing state is determined by the Eq. (4-1) (Hopfield, 1982).

The synaptic weight matrix $\boldsymbol{J}$ must exhibit specific characteristics for the network dynamics to be governed by a Lyapunov function. The diagonal entries of this matrix must be null, implying no self-connections, or $J_{ii} = 0, \forall i$, and the matrix must be symmetric, *i.e.*, $J_{ij} = J_{ji}, \forall i, j$. Under these premises, the energy function $E[t]$ can be defined as follows (Goles-Chacc et al., 1985, Bruck, 1990):

$$E[t] = -\frac{1}{2} v^T \boldsymbol{J} v - v^T I = -\frac{1}{2} \sum_{i,j} J_{ij} v_j[t] v_i[t] - \sum_j I_j v_j[t]. \tag{4-2}$$

This energy function $E[t]$ thus acts as a Lyapunov function for the dynamical system (4-1).

The aforementioned Lyapunov function is the nexus binding the principles of associative memory and a Hopfield network. An associative memory paradigm is distinctly identified by its capacity for cued pattern recall; when given a fragment of a stored pattern, the system strives to output the complete, nearest matching stored memory pattern.

Let us consider a reservoir of M patterns denoted as $\{\xi^\mu\}_{\mu=1}^M$. If the synaptic weight matrix $\boldsymbol{J}$ is configured such that each pattern ξ^μ is a local minimum of the Lyapunov function (4-2), then any initial state (or query) within the basin of said attractor will inevitably converge to the state corresponding to the pattern ξ^μ. This propensity is a testament to the network's associative recall capacity (Hopfield, 1982).

The construction of a synaptic weight matrix, $\boldsymbol{J}$, that projects the M patterns, $\boldsymbol{\xi}^\mu$, to the attractors of the Lyapunov function (4-2) is an essential step in building Hopfield networks. In its most rudimentary form, this matrix can be composed by the summation of outer products of the given patterns:

$$\boldsymbol{J} = \sum_{\mu=1}^{M} \boldsymbol{\xi}^\mu \cdot \boldsymbol{\xi}^{\mu T}. \tag{4-3}$$

This particular learning rule, characterizing the formation of the synaptic weight matrix, is labeled as a Hebbian rule. This nomenclature stems from synaptic weights being proportional to the covariance of the pre- and postsynaptic states, aligning with Donald Hebb's postulate of learning and synaptic plasticity (Hebb, 1949).

So far, we have primarily discussed Hopfield networks characterized by binary state variables evolving in discrete time intervals. Hopfield, in 1984, proposed an extension of this model to a domain where both the state variables and time manifest as continuous quantities.

In this continuous setting, let us denote v_i as the internal potential of a neuron i and let g be a bounded, real, and monotonically increasing function, typically a sigmoid function. This function g transmutes a neuron's potential into a firing rate, signified by $s_i = g(v_i)$. Suppose the internal potential of a neuron i reacts instantaneously to the input from neuron j via synaptic weight J_{ij}. It then follows an exponential decay with a constant τ.

The temporal evolution of the neurons' internal potential can be comprehensively described by a system of N ordinary differential equations:

$$\tau \frac{dv_i}{dt} = -v_i + \sum_{j=1}^{N} \tau J_{ij} g(v_j) + I_i, \tag{4-4}$$

where I_i represents the bias input to neuron i (Hopfield, 1984).

Analogous to the discrete counterpart, if the synaptic weight matrix, $\boldsymbol{J}$, exhibits symmetry, we can define a Lyapunov function for the continuous system as follows:

$$E = -\frac{1}{2} v^T \boldsymbol{J} v - v^T I - \frac{1}{\tau} \sum_{j=1}^{N} \int^{g(v_j)} g^{-1}(z) dz. \tag{4-5}$$

This energy function is a Lyapunov function for the dynamical system (4-4) (Hopfield, 1984).

4.2.2 Modern Hopfield Networks

The original formulations of Hopfield networks, both discrete and continuous versions, unfortunately, suffer from a fundamental limitation–their storage capacity scales linearly with the number of inputs (Hopfield, 1982). Furthermore, this retrieval system requires the memories to be entirely uncorrelated for flawless pattern retrieval. This condition often needs to be revised in complex real-world applications where correlation among memories is the norm rather than an exception.

The introduction of strong nonlinearities into the activation function g has been instrumental in addressing the storage capacity limitations inherent in classical Hopfield networks. This adaptation has allowed for a super-linear scaling relationship between the number of inputs and the network's storage capacity, as elucidated in studies by Krotov and Hopfield, Demircigil et al. (2016, 2017). This advanced model, christened as "Modern Hopfield Networks" or "Dense Associative Memories," marks a significant departure from the classical formulations. These modern networks' discrete and continuous versions have been studied extensively (Krotov and Hopfield, 2016, Ramsauer et al., 2020).

The binary neuron model in modern Hopfield networks has been adequately discussed in the literature (Krotov and Hopfield, 2016). However, for continuity and comprehensive understanding, we will focus on the case of continuous variables and time from now on.

A general system of differential equations in the form of (4-4), where $s_j = g(v_j)$, for a non-linear function g, can exhibit complex dynamics. The choice of nonlinearities and initial conditions can heavily influence the nature of these dynamics.

Modern Hopfield networks, though similar in structure, tend to exhibit more tractable dynamics due to their intrinsic design. They are engineered to have an underlying energy

function $E(t)$ that is guaranteed to decrease over time, *i.e.*, $dE(t)/dt \leq 0$. This characteristic ensures the convergence of the system dynamics to fixed-point attractor states, simplifying the analysis and comprehension of their behavior.

Let us consider a fully connected network of N neurons characterized by a symmetric matrix of weights. Each synaptic connection weight from neuron j to neuron i is represented by $J_{ij}, 1 \leq i \leq N, 1 \leq j \leq N$. We represent the activity of each neuron i, v_i, using its activation function g_i and timescale τ_i. The evolution of v_i over time can be modeled by the firing rate, Eq. (4-6), is:

$$\tau_i \frac{dv_i}{dt} = -v_i + \sum_{j=1}^{N} J_{ij} g_j(v_j). \tag{4-6}$$

A Lagrangian function $L\left(\{v_i\}_{i=1}^{N}\right)$ is assumed to encapsulate the dependency on the state of (4-6). Consequently, we can express each activation function g_i as a partial derivative of the Lagrangian with respect to v_i:

$$g_i := \frac{\partial L}{\partial v_i}. \tag{4-7}$$

The introduction of the Lagrangian function allows for the derivation of a global energy function for (4-6), which is defined as (4-8):

$$E = \boldsymbol{v}^T \boldsymbol{g} - L - \frac{1}{2} \boldsymbol{g}^T \boldsymbol{J} \boldsymbol{g}. \tag{4-8}$$

The first term on the right-hand side, $\boldsymbol{v}^T \boldsymbol{g} - L$, represents the Legendre transform of the Lagrangian L with respect to the vector $\boldsymbol{v}$ (Krotov, 2021).

The temporal evolution of the energy function $E(t)$, represented by its derivative with respect to time, is defined as:

$$\frac{E(t)}{dt} = -(\nabla \boldsymbol{v})^T \boldsymbol{H}_v \nabla \boldsymbol{v}, \tag{4-9}$$

where $\boldsymbol{H}_{vij} = \tau_i \partial^2 v / \partial v_i \partial v_j$ represents the Hessian of $\boldsymbol{v}$.

The energy function is guaranteed to decrease over time ($dE(t)/dt \leq 0$) if the Hessian $\boldsymbol{H}_v$ is positive semi-definite (Krotov, 2021). Furthermore, if the energy function (4-8) is bounded from below, $\boldsymbol{v}$ will converge to a fixed-point attractor (Krotov, 2021).

The nature of the Lagrangian function significantly influences the behavior of the activation functions g_i. Each g_i is a non-linear function of its corresponding state variable v_i for

additive Lagrangian functions. However, for non-additive Lagrangian functions, each g_i can be a function of a subset of the state vector $\boldsymbol{v}$.

The cases of non-additivity become particularly intriguing when the activation function g_i embodies canonical neural computations. Specifically, we have instances where g_i enables divisive normalization, a mechanism widely referenced in the literature (Reynolds and Heeger, 2009, Ni et al., 2012, Ohshiro et al., 2011, Carandini and Heeger, 2012). Alternatively, g_i can facilitate contrast normalization by implementing softmax functions (Heeger, 1992, Geisler and Albrecht, 1992, Bonin et al., 2005). These instances highlight the fascinating breadth of computational capabilities within the non-additive domain.

We now shift our attention to a two-tier network consisting of both fast and slow neurons, denoted respectively by $v_i^f, 1 \leq i \leq N^f$ and $v_j^s, 1 \leq j \leq N^s$. This system is differentiated by their kinetic time constants, with fast neurons exhibiting a significantly smaller constant than their slower counterparts. This disparity can be mathematically represented as $\tau_i = \tau^f, \forall 1 \leq i \leq N^f$, $\tau_j = \tau^s, \forall 1 \leq j \leq N^s$, where $\tau^f \ll \tau^s$.

Moreover, the output functions of fast and slow neurons (h_i^f and g_j^s, respectively) are contingent on the activities within their corresponding neuronal populations. That is, $h_i^f = h\left(v_i^f i = 1^{N_f}\right)$ and $g_j^s = g\left(v_j^s j = 1^{N_s}\right)$.

We then model the temporal evolution of the neurons' state through the following system of differential equations:

$$\tau^f \frac{dv_i^f}{dt} = \sum_{j=1}^{N^s} J_{ij} g_j^s - v_i^f + I_i, \tag{4-10}$$

$$\tau^s \frac{dv_j^s}{dt} = \sum_{i=1}^{N^f} J_{ji} h_i^f - v_j^s + I_j. \tag{4-11}$$

These equations contain activation functions h^f and g^s, which are derivatives of their respective Lagrangian functions: $h_i^f = \partial L^f / \partial v_i^f$ and $g_j^s = \partial L^s / \partial v_j^s$. The Lagrangian functions are denoted as follows:

$$L^f = \log\left(\sum_{i=1}^{N^f} e^{v_i^f}\right), \tag{4-12}$$

$$L^s = \frac{1}{2}\sum_{j=1}^{N^s} (v_j^s)^2. \tag{4-13}$$

For such defined Lagrangians, we can leverage the averaging theory when $\tau^f \ll \tau^s$ to express Eq. (4-11) in terms of the steady-state solutions of Eq. (4-10).

The resulting averaged equations for the slowly-varying neurons under these conditions are given by (Ramsauer et al., 2020):

$$\tau^s \frac{dv_j^s}{dt} = \sum_{i=1}^{N^f} J_{ij} \text{softmax}\left(\sum_{j=1}^{N^s} J_{ij} v_j^s\right) - v_j^s + I_j, \tag{4-14}$$

This equation, fascinatingly, has an underlying energy function:

$$E = \frac{1}{2}\sum_{j=1}^{N^s}(v_j^s)^2 - \log\left(\sum_{i=1}^{N^f}\exp\left(\sum_{j=1}^{N^s} J_{ij} v_j^s\right)\right). \tag{4-15}$$

This remarkable framework of Eqs. (4-14) and (4-15) has a distinct resemblance to the self-attention mechanism (Vaswani et al., 2017) brought to prominence by the transformative deep learning architecture (Ramsauer et al., 2020, Krotov and Hopfield, 2020).

4.2.3 Storage and Retrieval of Temporal Sequences in Hopfield Networks

A fundamental form of short-term memory can retain the order of a sequence of stimuli (Fuster and Alexander, 1971, Howard and Kahana, 2002); for instance, recognizing whether a given stimulus has been presented before in a sequence. Implementing such a system within an ANN requires the capability to store and retrieve temporal sequences, which can be thought of as ordered chains of patterns. A robust short-term memory system of this type must satisfy three primary criteria:

1. The production of temporal sequences should be autonomous, not dependent on downstream feedback or upstream modulation (Hochreiter and Schmidhuber, 1997);
2. The rhythmic output must manifest as a network-level phenomenon rather than as an intrinsic property of individual neurons (Buzsaki and Draguhn, 2004);
3. The network should possess a rich set of rhythmic outputs modulated by afferent inputs (Fries, 2005, Rajan et al., 2016). We can interpret the output as a cyclic transition of stored patterns, with the network capable of chunking and arranging these patterns, producing many temporal sequences (Buzsáki, 2010).

We aim to construct a theoretical framework encompassing these requirements by relying on fundamental neural elements, that is, modeling the desired behavior in terms of the most straightforward possible neurons and synaptic forms. As discussed in Section 2.2.1, a symmetric synaptic weight matrix, akin to Eq. (4-3), leads to pattern storage using fixed-point attractors. The natural progression from this point is to explore the impact of introducing non-symmetric synapses on the dynamic landscape of the network.

The implications of injecting noise into the ideal symmetric "Hebbian" synapses, thereby inducing asymmetry, have been previously examined analytically (Sompolinsky and Kanter, 1986). Consider a network composed of N neurons with synaptic weights $\bar{J}_{ij} = J_{ij} + \eta_{ij}$, where η_{ij} follows a normal distribution $\mathcal{N}(0, , \sigma^2/N)$, and J_{ij} constitutes the Hebbian term characterizing a well-functioning network. Sompolinsky and Kanter (1986) demonstrated that the network could withstand substantial levels of σ, that is, high degrees of asymmetry, without a significant decline in performance.

By demonstrating the network's resilience to asymmetry, Sompolinsky *et al.*'s findings provide a basis for understanding how asymmetric synapses could contribute to network dynamics without undermining the base Hebbian storage capacity. It also paves the way for further investigations into how structured asymmetry might engender novel dynamics, such as the emergence of rhythmic output patterns that could underlie the system's capacity for storing and retrieving temporal sequences.

Structured asymmetry within synaptic weights has been the subject of rigorous study, with systematic investigations performed by Hopfield that provide an insightful perspective on the matter (Hopfield, 1982, 1984). A simple yet effective approach considered by Hopfield involved supplementing the existing synaptic matrix with an additional set of synapses, which can be represented as:

$$\boldsymbol{H} = \lambda \sum_{\mu=1}^{M} \boldsymbol{\xi}^{\mu+1} \cdot \boldsymbol{\xi}^{\mu T}. \tag{4-16}$$

In this equation, λ characterizes the relative strength of the new set of synapses compared to the original set represented by $\boldsymbol{J}$ in Eq. (4-3), and M represents the total number of stored patterns. For ease of analysis and clarity, we treat the original and additional synapses as two distinct sets; however, it is not essential for the practical operation of the network, as

a single synapse connecting every pair of neurons can potentially comprise both symmetric and asymmetric components.

The fundamental idea underpinning this arrangement is to establish a correlation between sequentially adjacent patterns, *i.e.*, $[\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^\mu, \boldsymbol{\xi}^{\mu+1}, \ldots, \boldsymbol{\xi}^M]$. This intention is reflected in the outer products $\boldsymbol{\xi}^{\mu+1} \cdot \boldsymbol{\xi}^{\mu T}$ in Eq. (4-16). Moreover, depending upon the final pattern $\boldsymbol{\xi}^{M+1}$, we either obtain a finite sequence concluding at pattern $\mu = M$ (if $\boldsymbol{\xi}^{M+1} = \boldsymbol{\xi}^M$), or a periodic sequence transitioning from states $\boldsymbol{\xi}^r, \forall 1 \leq r < M$, to $\boldsymbol{\xi}^M$, if $\boldsymbol{\xi}^{M+1} = \boldsymbol{\xi}^r$.

Taking into account the newly introduced synaptic connections, $\bar{\boldsymbol{J}} = \boldsymbol{J} + \boldsymbol{H}$, the resultant effect on a particular neuron's activity (for instance, neuron i) can be quantified as:

$$\sum_{j=1}^{N} J_{ij} g_j(v_j) + \lambda \sum_{j=1}^{N} H_{ij} g_j(v_j) = \sum_{\mu=1}^{M} \xi_i^\mu \sum_{j=1}^{N} \xi_j^\mu g_j(v_j) + \lambda \sum_{\mu=1}^{M} \xi_i^{\mu+1} \sum_{j=1}^{N} \xi_j^\mu g_j(v_j), \tag{4-17}$$

In this equation, the inner sums of the terms on the right-hand side represent the overlaps between pattern $\boldsymbol{\xi}^\mu$ and the current network state. A comparison of Eq. (4-17) with the second term on the right-hand side of (4-6) can provide enlightening insights into the effects of introducing asymmetry.

To simplify, let us analyze the scenario where the M patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^M$ are random, normalized, and uncorrelated. Assuming that the current state of the network aligns with a given pattern p, denoted $\boldsymbol{\xi}^p$, all the overlaps in Eq. (4-17) will be null for $\mu \neq p$ and equal to one for $\mu = p$. Therefore, Eq. (4-17) simplifies to:

$$\xi_i^p + \lambda \xi_i^{p+1}. \tag{4-18}$$

It is observable from this equation that the parameter λ essentially determines the pattern (either $\boldsymbol{\xi}^p$ or $\boldsymbol{\xi}^{p+1}$) towards which the state of neuron i will evolve in the next instance.

Assuming a condition where all neurons synchronously update, and λ is appropriately adjusted, a network expressed by (4-6) with a synaptic weight matrix as given by (4-16) will reproduce the specified sequence of patterns. Each subsequent pattern follows the prior one with each update of the state. However, one must note that the network requires precise tuning of parameters, such as λ, and is vulnerable to noise. Even minor correlations between patterns can lead to errors in the sequence of output patterns (Hopfield, 1982).

Another noteworthy concern is the assumption of synchronous updates, which seems untenable in biological and artificial neural network scenarios. Under asynchronous updates, where different neurons are updated at different times, the output at any given time step would be a superposition of patterns. This occurs because the update of a particular neuron depends on the current state of all other neurons, which includes the new states of neurons updated earlier. Consequently, the effective current at the neuron being updated at a given time step would not be a superposition of current and next state, as suggested by the idealized scenario of Eq. (4-18), but a superposition of multiple patterns.

Furthermore, addressing the transient nature of the patterns produced in the current model is critical. In this scheme, the network spends only a single step at each pattern in the sequence. Though the network successfully traverses the sequence of patterns, these patterns do not persist in the output long enough to be "cognitively" perceptually distinct (Hopfield, 1984). This discrepancy questions the model's utility in simulating cognitive processes where information must be retained over more extended, meaningful periods.

In summary, while this method introduces an exciting perspective on modeling the sequential dynamics of neural networks, several challenges arise from the required precision in parameter tuning, the biological and practical infeasibility of synchronous updates, and the transient nature of the network states.

To effectively capture the robust and temporally stable behavior of a pattern sequence within the neural network, our model must evolve to accommodate quasi-attractors, a concept widely recognized within the domain of dynamical systems (Conley, 1988). Quasi-attractors embody the idealized features of both stability and robustness of fixed-point attractors while integrating an additional property: the destabilization of the quasi-attractor as it persists for prolonged periods, thereby enabling the network state's transition to the succeeding quasi-attractor. This is a pivotal characteristic as we desire network states to stabilize for substantial periods, allowing for "cognitive" recognition.

Our model further imposes the constraint that synaptic weights can only be modified on a timescale significantly slower than the network's primary timescale. Echoing the principles outlined in Section 2.2.1, we will maintain a solution where the fast synaptic weights are kept static. Modifying synapse strengths will be reserved for the more complex task of learning.

The additional set of synapses, which previously enabled the network to generate a sequence of patterns, is paradoxically the origin of instability, as it promotes the departure from the current pattern as soon as the network state begins to align. The solution to this issue, as suggested independently by Sompolinsky and Kanter (1986) and Kleinfeld (1986), introduces a *synaptic mechanism* that effectively delays the synapses responsible for pattern transitions (as indicated in Eq. (4-16)) in comparison to the synapses responsible for robust pattern retrieval (as denoted in Eq. (4-3)). This "synaptic mechanism" is an umbrella term encompassing various biological mechanisms capable of implementing such delays, including a simple delay in direct synaptic contact between two neurons or introducing relays through other neurons or complete networks.

4.2.3.1 Delay Mechanisms

First, let us consider the scenario where a simple delay is implemented through the slow synapses. We denote $s_i(t)$ as the neuron i activity, delayed by τ_d seconds. Formally, $s_i(t) = v_i(t - \tau_d)$. The contributions to the state of neuron i from the network consequently become:

$$\sum_{\mu=1}^{M} \xi_i^\mu \sum_{j=1}^{N} \xi_j^\mu g_j(v_j(t)) + \lambda \sum_{\mu=1}^{M} \xi_i^{\mu+1} \sum_{j=1}^{N} \xi_j^\mu g_j(v_j(t - \tau_d)). \tag{4-19}$$

Following similar lines of analysis as previously established, if the network state aligns with pattern p at $t = 0$, then Eq. (4-19) simplifies to ξ_i^p for $0 < t < \tau_d$. At $t = \tau_d$, the second term in Eq. (4-19) becomes nonzero, resulting in $\xi_i^p + \lambda \xi_i^{p+1}$, mirroring Eq. (4-18), but with a critical difference.

In the following τ_d seconds ($\tau_d < t < 2\tau_d$), there will be no extraneous terms inducing transitions towards the subsequent patterns in the sequence, as the information about state transition will not propagate through the transition synapses due to the implemented time delay. This modification allows all neurons to update their states towards the following pattern in the sequence, even in the asynchronous update scenario. The network lingers in each pattern for an interval τ_d, where τ_d can be fine-tuned so that each pattern remains stable long enough to facilitate "cognitive" recognition.

The scope of our model is not restricted to simple delays. The delaying mechanism can

be further generalized by substituting the simple delay with a filter for postsynaptic activity transmitted over $\boldsymbol{J}^{\text{slow}}$. Consider the following formulation:

$$s_i(t) = v_i * w(t) = \int_0^\infty w(-t')v_i(t-t')dt', \tag{4-20}$$

In this instance, $w(t)$ denotes the filter applying a weighted time-average over the signal arriving via $\boldsymbol{J}^{\text{slow}}$, with $w(t)$ being non-negative and normalized, *i.e.*,

$$\int_0^\infty w(t)dt = 1. \tag{4-21}$$

By definition, the average time interval τ_d is given by:

$$\tau_d = \int_0^\infty tw(t)dt.$$

Additionally, we assume that $w(t)$ does not overweight contributions from more than two patterns simultaneously, *i.e.*, if the network spends a period T in each pattern, then $\tau_d < 2T$.

Notably, if $w(t) = \delta(t - \tau_d)$, then Eq. (4-20) reverts to the simple time delay case. Moreover, for the specific scenario of an exponential decay filter, defined as:

$$w(t) = \begin{cases} \frac{1}{\tau_d} e^{-t/\tau_d}, & \text{if} \quad t > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4-22}$$

The network's contribution to a given neuron's activity via the slow synapses equates to the normalized synaptic current (refer to Eq. (2-2)) in the spiking network introduced in Section 2.2.1. This correlation underscores the intricate relationships between varying network architectures and synaptic mechanisms.

The normalization condition given by Eq. (4-21) has noteworthy implications for the period T spent in each pattern for the case of Eq. (4-22). Specifically, T is given by

$$T = \tau_d \ln\left(\frac{2\lambda}{\lambda - 1}\right), \tag{4-23}$$

with the stipulation that $T < 2\tau_d$ if the condition $e^2/(e^2 - 2) < \lambda < e/(e - 2)$ is satisfied.

By the same reasoning as that developed in Section 2.2.2, a meaningful energy function can be defined as a mean field approximation for the slow synapses, despite the lack of a landscape of stable fixed-point attractors. The mean field contribution from the additional

slow, asymmetric synapses introduces an extra term to the energy function that underlies the fast, symmetric synapses. Qualitatively, this additional term will deepen the valley representing the fixed-point attractor for the network's current state pattern $\boldsymbol{\xi}^p$ for the duration of the averaging period T, after which this valley is gradually filled in. In contrast, the valley representing the subsequent pattern $\boldsymbol{\xi}^{p+1}$ is deepened.

4.2.3.2 Correlated Patterns

Addressing the restrictions imposed by uncorrelated patterns is vital, as this might seem somewhat limiting, especially considering the high correlation of stimuli in their natural or direct mapped form. However, numerous strategies have been proposed to address this concern, and several warrant further discussion.

The first of these strategies adopts an agnostic stance, positing that raw stimuli are transformed into uncorrelated patterns before storage. At the time of retrieval, these patterns are subsequently re-mapped into their representative physical or sensory forms. An alternative strategy suggests subtracting the mean correlation between the directly mapped stimuli from the synaptic matrices, thereby preserving local neuronal interactions and connectivity.

The most comprehensive approach, and the one we adopt here, involves projecting the synaptic matrices into a more conducive space. If we denote $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^M$ as the set of M patterns to be stored, and define $\boldsymbol{J}^{\text{fast}}$ according to Eq. (4-25), then the new set of patterns can be expressed as:

$$\boldsymbol{\zeta}^\mu = \sum_{p=1}^{M} \boldsymbol{J}^{\text{fast}^{-1}} \boldsymbol{\xi}^p, \tag{4-24}$$

where $\boldsymbol{J}^{\text{fast}^{-1}}$ represents the inverse of the matrix $\boldsymbol{J}^{\text{fast}}$, the new patterns $\boldsymbol{\zeta}^\mu$ will be orthogonal to the set $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^M$ and will therefore be uncorrelated.

4.3 Connecting Hopfield and Spike-Coding Networks

In this section, we derive the important result connecting the predictive coding framework discussed in Chapter 2 and the concept of associative memories through Hopfield networks.

To ensure terminological consistency and lucidity, let us rename the stabilizing synapses matrix (Eq. (4-3)) as $\boldsymbol{J}^{\text{fast}}$ and the transition synapses matrix (Eq. (4-16)) as $\boldsymbol{J}^{\text{slow}}$ to reflect their delayed behavior:

$$\boldsymbol{J}^{\text{fast}} = \sum_{\mu=1}^{M} \boldsymbol{\xi}^{\mu} \cdot \boldsymbol{\xi}^{\mu T}, \tag{4-25}$$

$$\boldsymbol{J}^{\text{slow}} = \lambda \sum_{\mu=1}^{M} \boldsymbol{\xi}^{\mu+1} \cdot \boldsymbol{\xi}^{\mu T}. \tag{4-26}$$

It is now clear we are drawing a parallel between the synaptic matrices $\boldsymbol{J}^{\text{fast}}$ and $\boldsymbol{J}^{\text{slow}}$ and $\boldsymbol{\Omega}^{\text{fast}}$ (Eq. (2-40)) and $\boldsymbol{\Omega}^{\text{slow}}$ (Eq. (2-41)) from the spike-coding network. We are ready to show that *the Hopfield Network with asymmetric synaptic weights is equivalent to the autonomous spike-coding network in* (2-20)*.*

The dynamics of a continuous Hopfield network including both $\boldsymbol{J}^{\text{fast}}$ and $\boldsymbol{J}^{\text{slow}}$ are:

$$\begin{aligned}
\frac{dv_i}{dt} &= -v_i + \sum_{\mu=1}^{M} \xi_i^{\mu} \sum_{j=1}^{N} \xi_j^{\mu} g_j(v_j) + \lambda \sum_{\mu=1}^{M} \xi_i^{\mu+1} \sum_{j=1}^{N} \xi_j^{\mu} g_j(v_j(t-\tau_d)) + I_i \\
&= -v_j + \sum_{j=1}^{N} \sum_{\mu=1}^{M} \xi_i^{\mu} \xi_j^{\mu} g_j(v_j) + \sum_{j=1}^{N} \sum_{\mu=1}^{M} \lambda \xi_i^{\mu+1} \xi_j^{\mu} g_j(v_j(t-\tau_d)) + I_j \\
\therefore \quad \frac{dv_i}{dt} &= -v_i + \sum_{j=1}^{N} J_{ij}^{\text{fast}} g_j(v_j) + \sum_{j=1}^{N} J_{ij}^{\text{slow}} g_j(v_j(t-\tau_d)) + I_i,
\end{aligned} \tag{4-27}$$

where i, j index the neurons, μ indexes the patterns, N is the number of neurons, M is the number of patterns, v_i is the potential of neuron i, I_i is a bias input current to neuron i, ξ_i^{μ} represents the ith component of the μth stored pattern, g_j is an activation function, $v_j(t-\tau_d)$ is the delayed input stimulus, and λ weights the asymmetric synapses $\boldsymbol{H}$.

From a term to term comparison with Eq. (2-32), reproduced below for clarity,

$$\begin{aligned}
\tau \frac{dv_i}{dt} &= -v_i - \tau \sum_{k=1}^{K} \omega_{ik} \sum_{j=1}^{N} \omega_{kj} \delta_j + \sum_{k=1}^{K} \omega_{ik} \sum_{j=1}^{N} A_{kj} \omega_{pj} s_j + \sum_{k=1}^{K} \omega_{ik} c_k \\
\therefore \quad \tau \frac{dv_i}{dt} &= -v_i + \sum_{j=1}^{N} \Omega_{ij}^{\text{fast}} \delta_j + \sum_{j=1}^{N} \Omega_{ij}^{\text{slow}} s_j + \Omega_{ij} c_j,
\end{aligned} \tag{4-28}$$

we can see the following:

- The $\Omega_{ij}^{\text{fast}}$ weights in (4-28) are akin to the symmetric J_{ij}^{fast} weights in (4-27). Both represent correlations within a stored pattern.
- The leak/decay v_j term is present in both models.
- The I_j bias input is akin to $\omega_i c$ in (4-28).
- The neuron time constant τ in (4-28) is equivalent to $1/\lambda$ in (4-28); the inverse of the asymmetric synapses $\boldsymbol{H}$ relative strength.
- The delayed asymmetric J_{ij}^{slow} weights in (4-27) are analogous to the slow recurrent synapses $\Omega_{ij}^{\text{slow}}$in (4-28). For a carefully chosen $\boldsymbol{A}$, we can show they are equivalent.

If $\boldsymbol{A}$ is defined such that $A_{ij} = 1$ if $j = (i+1) \mod N$ and 0 otherwise, that is, $\boldsymbol{A}$ is a $N \times N$ matrix with the following form:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{4-29}$$

Then multiplying $\boldsymbol{\xi}^\mu$ by A will result in $\boldsymbol{\xi}^{\mu+1}$. That is,

$$A\boldsymbol{\xi}^\mu = \boldsymbol{\xi}^{\mu+1}. \tag{4-30}$$

Applying the ansatz (4-30), Eq. (4-27) becomes:

$$\frac{dv_i}{dt} = -v_j + \sum_{j=1}^{N}\sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu g_j(v_j) + \sum_{j=1}^{N}\sum_{\mu=1}^{M} \lambda (A\xi^\mu)_i \xi_j^\mu g_j(v_j(t-\tau_d)) + I_j$$

$$\frac{dv_i}{dt} = -v_j + \sum_{j=1}^{N}\sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu g_j(v_j) + \sum_{j=1}^{N}\sum_{\mu=1}^{M} \lambda \xi_i^\mu A_{ij} \xi_j^\mu g_j(v_j(t-\tau_d)) + I_j.$$

This matches Eq. (4-28) by identifying:

$$g_j(v_j) \equiv s_j$$

$$\boldsymbol{\xi}^\mu \equiv \boldsymbol{\Omega}_i,$$

where $\boldsymbol{\Omega}_i = [\omega_{1i}, \ldots, \omega_{Ki}]^T$ is the i-th column of $\boldsymbol{\Omega}$ and represent the decoding weights of neuron i. Each pattern $\boldsymbol{\xi}^\mu$ corresponds to signal z_k in the spike-coding framework.

Therefore, with this particular permutation matrix $\boldsymbol{A}$*,* (2-32) *is equivalent to* (4-27) *and we establish an equivalence between the predictive coding framework of spike-coding networks and the associative memory of continuous Hopfield networks.*

4.3.1 Cyclic Permutation Matrices

The matrix under consideration, $\boldsymbol{A}$, is characterized as a permutation matrix, specifically defined by a square binary matrix structure with precisely one entry of 1 for each row and column, and 0s dispersed otherwise (Brualdi et al., 1991). The particular $\boldsymbol{A}$ in (4-29) is known as a Sylvester's cyclic permutation matrix (Aitken, 2017).

Sylvester's permutation matrices carry distinct properties and significant implications for a system's dynamics governed by such matrices. First, their orthogonality implies that the system is invariant under the transformation represented by the matrix (Strang, 2022). Furthermore, orthogonal matrices preserve the inner product, meaning that the angles and lengths are preserved (Blank et al., 2008).

For Hopfield networks, the permutation matrix's transpose equivalent to its inverse creates a new system configuration from the original connectivity matrix $\boldsymbol{J}^{\text{slow}}$. Such invariance could prompt symmetry or uniformity in the network's dynamic behavior, affecting its pattern recognition and memory recall capacity (Rao et al., 2002).

In the context of spike-coding networks, the transformation denoted by $\boldsymbol{\Omega}^T\boldsymbol{A}\boldsymbol{\Omega}$ yields a new weight matrix that conserves the geometric attributes of the original weight matrix $\boldsymbol{\Omega}$. Since weight matrices denote the synaptic connections' strength, this transformation might culminate in symmetrical or uniform synaptic connections in the transformed network. This uniformity can stimulate synchronous neuron firing, creating stable patterns or oscillations within the network (Dayan and Abbott, 2005).

Sylvester's permutation matrices have a determinant of -1, suggesting that the system undergoes a reflection or a parity transformation (Horn and Johnson, 2012). This has implications for physical systems described by such matrices, which include changes in direction but maintain the system's volume. For example, in physics, this is related to time-reversal symmetry, where a negative determinant implies that the system behaves the same way if time is reversed (Sakurai and Commins, 1995).

In the framework of the Hopfield network, this could potentially lead to a reflection or inversion in the network's operational behavior. For instance, the transformation $\boldsymbol{J}^{\text{slow}}$ will negate the connectivity matrix given $\boldsymbol{H}$ is an identity matrix. In the context of network

dynamics, this could give rise to flipping in the stored or retrieved patterns, as commonly observed in the stability of anti-patterns (Roelfsema and Ooyen, 2005).

For spike-coding networks, this could mean that the network's dynamics exhibit a balance between excitation and inhibition or represent some form of reset mechanism in the neuronal dynamics, where a certain state leads to a reversal in the neurons' behavior (Izhikevich, 2007).

Eigenvalues of these matrices are the nth roots of unity, implying the system's dynamics are cyclical; after n sequential transformations, the system returns to its initial state, highlighting the periodic and, thus, predictable behavior of the system.

In a Hopfield network scenario, this could evoke periodic or rhythmic patterns in the network's dynamics, such as cyclic shifts in memory recall patterns or oscillations between stable states based on input and initial conditions (Chen and Amari, 2001).

In the context of a spike-coding network, this could signify oscillatory activity. These cycles might correspond to various rhythmic activities observed in the brain, such as theta, alpha, beta, and gamma rhythms (Buzsaki and Draguhn, 2004). Additionally, it might denote cyclic dynamics in a neuron chain or a ring network, where activity cyclically propagates through the network (Ermentrout and Terman, 2010).

In conclusion, a system governed by a permutation matrix exhibits a periodic behavior. It is also invariant under the transformation represented by the matrix (owing to orthogonality), and time-reversal symmetry (as the determinant is -1). For a spike-coding network, these properties could result in symmetrical synaptic connections, excitatory-inhibitory balance, and the emergence of oscillations. For Hopfield networks, these properties potentially introduce symmetry, stability of "anti-patterns," and cyclic shifts in memory recall patterns or oscillations between various stable states.

4.3.2 Generalization to Circulant Matrices

In the presented framework, two important theoretical constructs converge—the predictive coding framework and the associative memory dynamics of Hopfield networks. Specifically, the Hopfield network with asymmetric synaptic weights shows equivalency with the autonomous spike-coding network for a cyclic permutation matrix. Now, we generalize to

the case of arbitrary circulant matrices.

A circulant matrix has the property that its columns (or rows) are cyclic shifts of each other. In other words, a circulant matrix has the form $\boldsymbol{C_c}$ (Gray et al., 2006):

$$\boldsymbol{C_c} = \begin{bmatrix} c_0 & c_{N-1} & c_{N-2} & \cdots & c_1 \\ c_1 & c_0 & c_{N-1} & \cdots & c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N-1} & c_{N-2} & \cdots & c_1 & c_0 \end{bmatrix}. \tag{4-31}$$

For any circulant matrix $\boldsymbol{C_c}$ of size $N \times N$, there exists an associated polynomial $p(\boldsymbol{c})$ in the cyclic permutation matrix $\boldsymbol{A}$ (Eq. (4-29)), such that:

$$\boldsymbol{C_c} = c_0\boldsymbol{I} + c_1\boldsymbol{A} + c_2\boldsymbol{A}^2 + ... + c_{N-1}\boldsymbol{A}^{N-1}, \tag{4-32}$$

where $\boldsymbol{I}$ is the identity matrix, c_k, $\forall 0 \leq k \leq N-1$ are the polynomial coefficients, and each power of $\boldsymbol{A}$ corresponds to a cyclic shift of the matrix (Pan, 2001).

As a result, we can express any circulant matrix as a linear combination of the cyclic shift matrices $\boldsymbol{A}^n$, where n is the shift count. By introducing the circulant matrix $\boldsymbol{C_c} = \sum_{n=0}^{N-1} c_n\boldsymbol{A}^n$ into the definition of $\boldsymbol{J}^{\text{slow}}$, allows us to capture more complex interaction dynamics in the network. The new definition of $\boldsymbol{J}^{\text{slow}}$ in terms of $\boldsymbol{C_c}$ is give by:

$$\boldsymbol{J}^{\text{slow}} = \lambda \sum_{\mu=1}^{M} \sum_{n=0}^{N-1} c_n \boldsymbol{\xi}^{\mu+n} \cdot \boldsymbol{\xi}^{\mu T}. \tag{4-33}$$

This matrix captures more general interaction dynamics in the network, by weighting all the different shifted patterns in determining the influence of neuron j on neuron i.

With the given generalized $\boldsymbol{J}^{\text{slow}}$, (4-27) becomes:

$$\begin{aligned} \frac{dv_i}{dt} &= -v_i + \sum_{j=1}^{N} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu g_j(v_j) + \lambda \sum_{j=1}^{N} \sum_{\mu=1}^{M} \sum_{n=0}^{N-1} c_n \xi_i^{n+\mu} \xi_j^\mu g_j(v_j(t-\tau_d)) + I_i. \\ &= -v_i + \sum_{j=1}^{N} J_{ij}^{\text{fast}} g_j(v_j) + \lambda \sum_{j=1}^{N} \sum_{n=0}^{N-1} c_n \xi_i^\mu A_{ij}^n \xi_j^\mu g_j(v_j(t-\tau_d)) + I_i \\ \therefore \quad \frac{dv_i}{dt} &= -v_i + \sum_{j=1}^{N} J_{ij}^{\text{fast}} g_j(v_j) + \sum_{j=1}^{N} J_{ij}^{\text{slow}} g_j(v_j(t-\tau_d)) + I_i. \end{aligned} \tag{4-34}$$

Given the new form of $\mathbf{\Omega}^{\text{slow}}$, the correspoding $\mathbf{\Omega}^{\text{slow}}$ in (4-34) is:

$$\mathbf{\Omega}^{\text{slow}} = \sum_{n=0}^{N-1} c_n \mathbf{\Omega}^T \boldsymbol{A}^n \mathbf{\Omega}, \tag{4-35}$$

which is obtained by replacing $\boldsymbol{A}$ with the circulant matrix $\boldsymbol{C_c}$.

The polynomial $\boldsymbol{C_c}$ defines the neural interaction topology. Specifically, the coefficients c_n in $\boldsymbol{C_c}$ determine how much of the nth shifted pattern will affect the current neuron state. Consequently, $\boldsymbol{C_c}$ defines how information propagates through the network and how the network's state evolves. When $n = 1$ and $c_0 = 0$, this reduces to the 1-step cyclic permutation case. Higher order terms like $c_2\boldsymbol{A}^2$ describe longer range correlations.

It is worth noting that the generalization of the networks to incorporate circulant matrices does not dilute the models' original purpose. In contrast, it fortifies them by accounting for more complex scenarios and introducing additional parameters that may enhance the models' predictive power and fidelity to real neural dynamics (Marullo and Agliari, 2020). Varying the coefficients c_n allows flexible control over the temporal window and structure of the correlations (Cheng et al., 2015). For example, high-order terms of the form $c_i\boldsymbol{A}^i$, $i \geq 2$, can model longer-range dependencies, like slowly oscillating brain rhythms.

Therefore, expressing $\boldsymbol{C_c}$ as a polynomial in $\boldsymbol{A}$ provides a general way to systematically map cyclic, convolving dynamics between Hopfield and spike-coding networks across multiple scales. The quadratic form $\mathbf{\Omega}^T\boldsymbol{C_c}\mathbf{\Omega}$ could be seen as representing the aggregate effect of these synaptic weights at multiple timescales on the neural activity, implying a *multi-scale model of neural computation and learning*. The term with the identity matrix, $c_0\mathbf{\Omega}^T\boldsymbol{I}\mathbf{\Omega}$, would represent the current influence of the synaptic weights on the neural activity. The term $c_1\mathbf{\Omega}^T\boldsymbol{A}\mathbf{\Omega}$ would then capture the influence of synaptic weights with a one-step delay, $c_2\mathbf{\Omega}^T\boldsymbol{A}^2\mathbf{\Omega}$ would capture the influence with a two-step delay, and so on.

The proposed interpretation mirrors the notion of multi-timescale learning within reinforcement learning theory. This theory acknowledges that different learning facets can, and more importantly, should advance at disparate rates (Sutton and Barto, 2018).

In parallel, this interpretation aligns with empirical evidence in neuroscience, specifically with the observed plasticity changes in the brain's synapses. Plasticity changes can span several timescales, from short-term alterations persisting in seconds to long-term modifications lasting days or even longer (Fusi et al., 2005).

Furthermore, this model could also be useful in capturing and studying phenomena like STDP, where the change in a synapse's strength depends on the precise timing of pre- and postsynaptic spikes (Markram et al., 2011). Likewise, it could be useful in modeling synaptic consolidation, where synaptic changes induced during periods of wakefulness are solidified during sleep (Diekelmann and Born, 2010), a central mechanism for memory storage and consolidation, balancing the need for learning new information and maintaining the stability of neuronal networks (Abel and Lattal, 2001).

4.3.3 Circular Convolutions

Circulant matrices have been extensively studied for their unique properties and diverse applications in a range of fields, from image processing (Gonzales and Wintz, 1987), coding theory (McEliece, 2002), to communication systems (Proakis, 2007).

Circulant matrices offer a unique implementation of circular convolutions due to their specific structure, which allows for rotations and shifts in their entries. Let's begin with the definition of circular convolution to appreciate the connection between circulant matrices and circular convolutions. For two vectors of length N, denoted as $\boldsymbol{a} = [a_0, a_1, ..., a_{N-1}]$ and $\boldsymbol{b} = [b_0, b_1, ..., b_{N-1}]$, the circular convolution of $\boldsymbol{a}$ and $\boldsymbol{b}$ is given by:

$$(\boldsymbol{a} \circledast \boldsymbol{b})n = \sum_{k=0}^{N-1} a_k b_{(n-k) \bmod N}, \tag{4-36}$$

where n ranges from 0 to $N-1$, the operation $(n-k) \bmod N$ denotes the modulus after division, ensuring the indices are cyclic (Gray et al., 2006).

Now, let's consider a circulant matrix $\boldsymbol{C_b}$, whose first column is defined by the vector $\boldsymbol{b} = [b_0, b_1, ..., b_{N-1}]$. Each subsequent column of $\boldsymbol{C_b}$ is a cyclic shift of the previous column. When $\boldsymbol{C_b}$ multiplies $\boldsymbol{a}$, the resulting vector is exactly the circular convolution of $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$\boldsymbol{C_b a} = \boldsymbol{a} \circledast \boldsymbol{b}. \tag{4-37}$$

One of the most significant properties of circulant matrices is that they are diagonalizable by the discrete Fourier transform (DFT) matrix (Davis, 1979). This property is critical because it allows using the fast Fourier transform (FFT), an efficient algorithm for computing DFTs, in matrix-vector multiplication involving circulant matrices (Lu, 1989).

To demonstrate this property, let $\boldsymbol{F}_N$ be the $N \times N$ DFT matrix, whose elements are given by $\boldsymbol{F}_N[k, n] = \omega_N^{kn}$, where $\omega_N = e^{2\pi i/N}$ is the Nth root of unity. It can be shown that

$$\boldsymbol{C_b} = \boldsymbol{F}_N^{-1} \boldsymbol{D} \boldsymbol{F}_N, \tag{4-38}$$

where $\boldsymbol{D}$ is a diagonal matrix with elements equal to the DFT coefficients of the vector $\boldsymbol{b}$.

This result implies that the circular convolution of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, which can be computed as $\boldsymbol{C_b a}$ if $\boldsymbol{C_b}$ is a circulant matrix with first column equal to $\boldsymbol{b}$, can also be computed in the Fourier domain using the FFT as follows:

$$\boldsymbol{C_b a} = \boldsymbol{F}_N^{-1} \boldsymbol{D} \boldsymbol{F}_N \boldsymbol{a} = \boldsymbol{F}_N^{-1} (\boldsymbol{F}_N \boldsymbol{y}) \odot (\boldsymbol{F}_N \boldsymbol{x}), \tag{4-39}$$

where $\odot$ denotes the Hadamard (element-wise) product.

This result shows that circular convolution in the time domain corresponds to point-wise multiplication in the Fourier domain (Bracewell, 1986), a fundamental property used in many digital signal processing applications (Oppenheim, 1999).

This opens the door to a more in-depth investigation of the system in the frequency domain since its dynamics becomes decouple, allowing for easier analysis, *i.e.*, the solution to each frequency component of the system can be examined independently. This yields valuable insights into the oscillatory behavior of the system, such as the frequency content of the system's response and the effects of delays on this response.

Furthermore, stability can be evaluated by examining the roots of the system's characteristic equation in the Fourier domain. Specifically, we can analyze the spectral radius of the DFT of the matrices, which represents the maximum magnitude of their eigenvalues. If the spectral radius is less than or equal to one, the system is stable; otherwise, it is unstable. Importantly, this stability criterion is significantly easier to check than directly analyzing the potentially high-dimensional dynamical system.

Therefore, the application and implications of circulant matrix $\boldsymbol{C}$ and its induced circular convolution have far-reaching impacts on both Hopfield Networks and spike-coding networks. They introduce robust computational mechanisms that allow for modeling cyclic temporal dynamics and correlations, thus playing a pivotal role in memory and sequencing.

For Hopfield Networks, the circulant matrix $\boldsymbol{C}$ serves to implement a time-delayed autocorrelation function between stored patterns $\boldsymbol{\xi}^\mu$. This function enables the creation of cyclic transitions between stored memories, thereby executing a shift register as associations cycle through time. The introduction of asymmetric synapses, which encode spatial correlations using circulant matrices, highlights the suitability of these structures for convolution. Through these cyclic dynamics, robust transitioning between stored states is enabled, a factor that reduces corruption and improves overall network performance.

In the context of spike-coding networks, employing the special $\boldsymbol{C}$ matrix leads to the implementation of circular convolution through slow synapses $\boldsymbol{\Omega}^T\boldsymbol{C}\boldsymbol{\Omega}$, hence facilitating correlations of states across time. This process allows the network to dynamically reconstruct temporal sequences from the stored patterns, introducing a new dimension of flexibility and responsiveness in memory function. The cyclic permutation of states essentially implements a form of delay line memory, a technique that further underscores the role of cyclic structures in network memory and sequencing. Importantly, this cyclic structure provides a stable way to introduce delays without interference between associations, thus ensuring smooth and efficient network operations.

4.3.4 Universality with Circulant Structure

Previously, we showed that spike-coding networks can implement arbitrary linear dynamical systems by appropriately choosing the matrix $\boldsymbol{A}$ (Chapter 2). We now show that this result holds even if $\boldsymbol{A}$ is replaced in the original formulation by the circulant matrix $\boldsymbol{C}_c$.

Let's consider the following linear time-invariant (LTI) system, represented as

$$\dot{\boldsymbol{x}} = -\boldsymbol{x} + \boldsymbol{A}\boldsymbol{x} + \boldsymbol{c},$$

where $\boldsymbol{A}$ is the system matrix and $\boldsymbol{c}$ denotes the external input.

The equation's homogeneous solution is often expressed in the form of

$$\boldsymbol{x}_h(t) = e^{(\boldsymbol{A}-\boldsymbol{I})t}\boldsymbol{x}(0),$$

where $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{x}(0)$ designates the initial conditions. This characterizes the system's natural dynamics, *i.e.*, that is independent of the external input.

The additional term $\boldsymbol{c}$ yields a particular solution that must be determined based on the nature of $\boldsymbol{c}(t)$. Techniques such as the method of undetermined coefficients, variation of parameters, or Laplace transform can be used to find this particular solution, $\boldsymbol{x}_p(t)$. The overall solution is $\boldsymbol{x}(t) = \boldsymbol{x}_h(t) + \boldsymbol{x}_p(t)$.

An essential concept within this context is the system's impulse response, $\boldsymbol{h}(t)$, defined as the system's output when the input $\boldsymbol{c}(t)$ is an impulse function, denoted as $\boldsymbol{\delta}(t)$. This response carries comprehensive information about the system's characteristics; it can be used to compute the system's output for any given input through the convolution operation, *i.e.*, $\boldsymbol{x}(t) = \boldsymbol{h}(t) \star \boldsymbol{c}(t)$ (Chen, 1984).

Let us delve into a scenario where a vector $\boldsymbol{h} = [h_0, h_1, h_2, ..., h_{N-1}]^T$ whose elements are the impulse response of an LTI system is the kernel for a circulant matrix $\boldsymbol{C_h}$. In this scenario, the following association holds:

1. Circulant matrix $\boldsymbol{C}_h$ is comprised of columns representing cyclic shifts of $\boldsymbol{h}$.
2. Multiplying $\boldsymbol{C}_h$ by a vector $\boldsymbol{x}$ corresponds to the circular convolution of $\boldsymbol{h}$ and $\boldsymbol{x}$.
3. Given proper zero-padding of $\boldsymbol{x}$, the circular convolution with $\boldsymbol{h}$ approximates a linear convolution with $\boldsymbol{h}$.

Therefore, it is inferred that this matrix, defined by the impulse response of a filter, implements the corresponding filtering operation, with the cyclic structure capturing the temporal evolution of the filter's response, *i.e.*, $\boldsymbol{C_h x}$ mirrors the operation of the LTI system on input $\boldsymbol{x}$. Hence, constraining $\boldsymbol{A}$ as a circulant matrix does not limit the class of linear systems the spike-coding network implements. The circulant structure preserves the network's universal linear function approximation property.

4.3.5 Convolutional Neural Circuits

The circulant matrix acts as a discrete LTI filter, effectively implementing a discrete convolution with the network state. This mechanism creates a sliding filter effect, unearthing correlations across time and space. Notably, this operation resembles the convolutional layers in convolutional neural networks (CNNs).

In the context of Hopfield networks, asymmetric synapses implement convolution with a circulant matrix, correlating the stored patterns. This process mirrors CNN filters' strategy to convolve their weights across space. In parallel, slow synapses in spike-coding networks perform a function akin to computing a convolutional filter that transitions between temporal states. This resembles the hierarchical processing inherent in CNNs, assembled by stacking multiple convolutional layers.

The associative memory properties and dynamical features of Hopfield and spike-coding networks could inspire enhancements to CNNs. A hallmark of these models is their ability to handle spatio-temporal data naturally. In contrast, standard CNNs often struggle with temporal information. Thus, embedding spatio-temporal processing capabilities into CNNs could boost their effectiveness in video and sequence data tasks.

For instance, introducing external memory modules, motivated by Hopfield's associative memory, could empower CNNs to store and retrieve valuable prior activations relevant to a task, potentially augmenting their performance (Weston et al., 2014). Furthermore, segmenting features into fast and slow timescales, mirroring the fast and slow synapses of spike-coding networks, could equip CNNs with a more intricate mechanism for temporal processing. Here, fast features detect immediate cues, while slow features integrate context over time (Tavanaei et al., 2019). Chapter 5 introduces a CNN inspired by this dual "fast-slow" pathway to tackle an image-to-image translation problem.

4.4 Simulations

4.4.1 Methods

This section describes the simulation methods, including the spiking neural network architecture, training procedures, and parameters used to implement the n-back working memory task. The network consists of interconnected excitatory and inhibitory spiking neurons with synaptic plasticity. Training involves presenting sequences of stimuli and modulating the network's synaptic weights to produce the desired output spike patterns. Key parameters include the number of excitatory and inhibitory cells, synaptic strengths, neuronal dynamics, and training algorithms.

4.4.1.1 The n-back Task

The n-back task, originating from Kirchner's study in 1958, has proven to be an indispensable tool in cognitive psychology and neuroscience, particularly for assessing working memory capabilities (Kirchner, 1958, Baddeley, 2003). The task operates by presenting subjects with a series of stimuli and requiring them to identify when a current stimulus matches the one from n trials prior. This fundamental setup permits extensive flexibility in research designs by allowing adjustment in key variables such as the memory span n, the set of stimuli N, and the interstimulus intervals t_i (Jaeggi et al., 2010a).

The theoretical construct of the n-back task is grounded in the working memory model, where the task is seen to place a significant demand on the central executive component—the center of attentional control. In the context of the n-back task, this component is engaged in ongoing information management and updating, which is scaled according to the value of n, indicating the memory span needed (Conway et al., 2005).

Adjustments in the task parameters can effectively modulate its complexity and difficulty. A higher n value elevates the task's difficulty due to increased information load and potential for interference, challenging the participant's cognitive resource management (Jaeggi et al., 2010b). Similarly, a larger stimulus set N enhances discrimination difficulty within the task by necessitating differentiation among an increased number of stimuli, thereby heightening

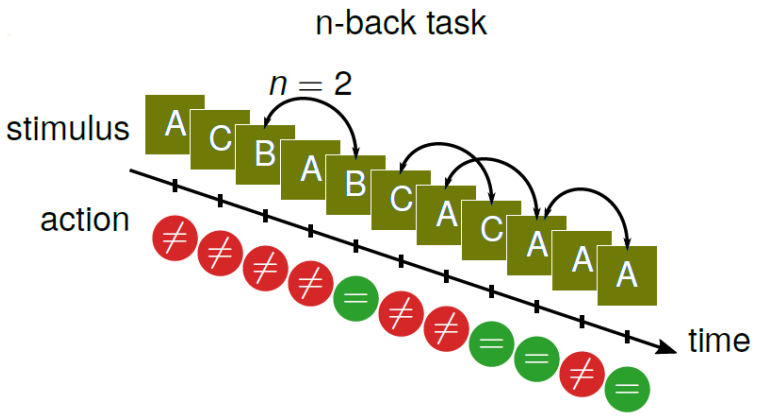


Figure 4.1: **The *n*-back task.** In each trial, the subject has to memorize a sequence of stimuli, drawn from the stimuli set (here {A, B, C}) and perform an action whenever the current stimulus is equal to the stimulus presented n steps before (here $n = 2$).

cognitive load (Oberauer, 2005). Interstimulus intervals t_i, the time lapses between consecutive stimuli, also contribute to the task's modulation. Shorter intervals necessitate rapid cognitive processing and response, whereas longer intervals can amplify memory load as the retention of information over a longer duration is required (Zakay, 1989).

Performance on the n-back task is conventionally quantified by the ratio of correctly executed actions, a direct measure of working memory accuracy. This metric is often supplemented with reaction times to evaluate cognitive efficiency (Jaeggi et al., 2010a). Effective and efficient working memory is characterized by high accuracy and minimal reaction times.

In essence, the n-back task is a robust and adaptable instrument for exploring working memory and cognitive functioning. The manipulation of key task variables enables the generation of detailed cognitive profiles and contributes to our understanding of the nature and boundaries of human working memory.

4.4.1.2 Network Architecture

To demonstrate the capabilities of the unified Hopfield and spike-coding network framework, we trained a network consisting of 200 interconnected excitatory and inhibitory spiking

neurons, as described in Section 2.2.1, to perform the 5-back memory task. The first 100 neurons were modeled as excitatory units and the last 100 as inhibitory.

The synaptic connectivity includes fast synapses $\mathbf{\Omega}_{\text{fast}}$ for pattern storage (Section 4.2.1) and slow synapses $\mathbf{\Omega}_{\text{slow}}$ for sequencing (Section 4.3.2). The recurrent synaptic weights $\mathbf{\Omega}$ were randomly initialized based on uniform and binomial distributions to simulate biologically realistic randomness and variation in the network connectivity (Sjostrom et al., 2008).

For excitatory neurons (indices 1 to 100), $\Omega_{ij} \sim B(1, 0.3) \cdot U(0.06, 0.1)$, where $B(1, 0.3)$ refers to a binomial distribution with $n = 1$ trial and probability $p = 0.3$. This results in a binary sample of either 0 or 1, with a 70% chance of being 1. $U(0.06, 0.1)$ refers to a continuous uniform distribution ranging from 0.06 to 0.1. By multiplying the binomial sample by the uniform sample, we obtain synaptic weights of 0 (with 70% probability) or take on continuous values between 0.06 and 0.1 (with 30% probability). This models the patchy connectivity observed in cortical networks, where each neuron only connects to a subset of other neurons (binomially distributed), with the strengths of those present connections varying continuously (uniformly distributed).

Likewise, for inhibitory neurons (indices 101 to 200), $\Omega_{ij} \sim B(1, 0.7) \cdot U(-0.1, -0.06)$, *i.e.*, synaptic weights that are either 0 (with 30% probability) or take on continuous values between -0.06 and -0.1. The negatives represent an inhibitory reversal potential (Ferrante et al., 2013). Overall, this initialization scheme produces a largely inhibitory-dominated network with sparse, patchy recurrent projections reflecting biological constraints (Perin et al., 2011, Song et al., 2005, Yoshimura et al., 2005).

The recurrent feedback weights $\mathbf{\Omega}_{\text{slow}}$ are scaled to $0.1\mathbf{\Omega}_{\text{fast}}$ to keep the overall network excitation below the threshold, maintaining stability and preventing runaway dynamics (Murphy and Miller, 2003). The spiking cost $\mu = 0.1$ added in $\mathbf{\Omega}^T\mathbf{\Omega} + \mu\boldsymbol{I}$ (see Eq. (2-37)) is used to implement a form of sparse and distributed coding across the neural network, regulating the propensity of individual neurons to fire (Boerlin et al., 2013), as explained in Section 2.4.1.

To evaluate the network's robustness, we introduced different types of noise and perturbations during simulations:

Gaussian voltage noise with a standard deviation of ($\sigma_V = 0.1$) was introduced into each

neuron to encourage irregular spiking patterns, replicating the diverse spiking activities exhibited by cortical neurons (Mainen and Sejnowski, 1995). This variability in spiking behavior is rooted in various biophysical noise sources, including channel noise due to the random opening and closing of ion channels, and synaptic noise due to the stochasticity in neurotransmitter release and reception (Faisal et al., 2008).

To emulate the inherent variability in sensory inputs, Gaussian noise ($\sigma_{\text{stim}} = 0.1$) was superimposed onto the input pixels at each step (Brunton et al., 2013). Real sensory inputs are subject to random fluctuations due to multiple factors, including transduction noise in sensory receptors and external environmental noise.

A drifting Ornstein-Uhlenbeck (OU) process ($\sigma_{\text{OU}} = 0.05$, $\lambda_{\text{OU}} = 10$ was utilized to model slow input fluctuations over time. The OU process is a continuous-time stochastic process that exhibits two key features: (1) it is Gaussian, signifying that any linear combination of the process at different points in time follows a Normal distribution (Gardiner et al., 1985), and (2) it is Markovian, implying that the future state of the process is solely determined by the current state, independent of its history (Van Kampen, 1992). The OU process is a fundamental model for mean-reverting processes, effectively representing fluctuations around a long-term mean value. The given parameters induce a rapid return to the mean within 10 seconds, with the process increments having a standard deviation of 0.05. This is intended to mimic effects such as the drift in attentional focus or slow sensory variations (Lundqvist et al., 2010).

A synaptic transmission delay τ_{delay} of 2 milliseconds was incorporated into the model to reflect biological limitations on the propagation of spikes between neurons. Delays in signal transmission can arise due to several factors, including synaptic delay, axonal conduction delay, and dendritic processing delay (Swadlow, 1994, Buzsáki et al., 2012)

To simulate neuronal death, 40% of neurons were silenced midway through trials. In biological systems, neurons can die for various reasons, from normal aging to pathological conditions, testing the resilience and adaptability of neural networks (Haspel et al., 2021).

Finally, the network's spiking dynamics were simulated via the forward Euler numerical integration method with a fine-grained time step of 0.1 milliseconds. Unlike their continuous counterparts, spiking neural networks represent and transmit information through discrete

events. Thus, accurately capturing these dynamics requires numerical methods that can handle such discontinuities and nonlinearity. The forward Euler method offers a simple yet effective solution, as it is an explicit method capable of tracking the abrupt changes associated with spike events (Press, 2007). Furthermore, the small time step size of 0.1 milliseconds is chosen to ensure numerical stability and accuracy. This fine temporal resolution allows the simulation to capture the millisecond-scale precision inherent in neural computation faithfully, given the rapid timescale of spiking activity (Lytton and Hines, 2005).

4.4.1.3 Network Training

The spike-coding network has fixed recurrent weights $\boldsymbol{\Omega}$ for pattern storage, as described in Chapter 2. Given the weight initialization procedure from the previous section, the patterns stored in $\boldsymbol{\Omega}$ are not directly "cognitively" interpretable as the input targets $\boldsymbol{\zeta}$. Therefore, using the procedure in Section 4.2.3.2, we train a projection matrix $\boldsymbol{W}$ between the patterns stored in $\boldsymbol{\Omega}$ and the desired patterns $\boldsymbol{\zeta}^{\mu}$:

$$\boldsymbol{\zeta}^{\mu} = \boldsymbol{W}\boldsymbol{\xi}^{\mu}.$$

$\boldsymbol{W}$ is initialized randomly and trained using the following Hebbian-like learning rule:

$$\Delta \boldsymbol{W} \propto \eta \left(\boldsymbol{\zeta} - \boldsymbol{W}\boldsymbol{\xi}\right) \boldsymbol{\xi}^{T}$$

This learning rule tunes the mapping $\boldsymbol{W}$ to associate the fixed random patterns in $\boldsymbol{\Omega}$ with the actual patterns $\boldsymbol{\zeta}$ we want to store.

The slow weights $\boldsymbol{\Omega}_{\text{slow}}$ implement temporal sequencing and are structured as follows:

$$\boldsymbol{\Omega}_{\text{slow}} = \sum_{n} c_n \boldsymbol{A}^n,$$

where $\boldsymbol{A}$ is a cyclic permutation matrix and c_n are the trainable coefficients.

The $\boldsymbol{c} = [c_1, \ldots, c_n]$ coefficients are updated according to the unsupervised learning rule:

$$\Delta \boldsymbol{c} \propto -\alpha \left(\boldsymbol{A}^n \boldsymbol{s} - \boldsymbol{\Omega}_{\text{slow}} \boldsymbol{s}\right).$$

This learning rule modifies the c_n to tune the structure of $\mathbf{\Omega}_{\text{slow}}$ and embed the 5-back computation. Therefore, after training, the network leverages the fast weights $\mathbf{\Omega}_{\text{fast}}$ for memory storage and the trained slow weights $\mathbf{\Omega}_{\text{slow}}$ to generate complex spike sequences for retaining and comparing stimuli over time.

The training involved presenting many 4 seconds long trials of input-target pairs $(\boldsymbol{\zeta}, \tilde{\boldsymbol{y}})$ where $\boldsymbol{\zeta}$ was a random binary vector with 6 dimensions (pixels), where each pixel was either "On" (1) or "Off" (0), and $\tilde{\boldsymbol{y}}$ the desired 5-back match/non-match readout indicating a stimulus match or non-match. Learning was initiated after 1 second to avoid adapting to input transients irrelevant to the task. The learning rates η and α decayed exponentially across 100 training epochs to embed robust encoding.

4.4.2 Results

The resulting spiking network could accurately perform the 5-back memory task, dynamically retaining and comparing stimuli over time, as illustrated in Figure 4.2. The network's spiking activity depicted sparse and irregular firing patterns, closely mirroring those observed in biological cortical networks (Yuste, 2015).

Throughout the simulations, small amounts of Gaussian noise and several perturbations were introduced that were not present during training. Despite these disruptions, the network continued to perform the 5-back task accurately. The distributed encoding provided resilience against perturbations, a key feature of biological neural systems.

Notably, the network generalized well to new input sequences, indicating the learning of robust computational rules rather than simply memorizing specific associations. The capacity to handle novel sequences highlights the potential of these biologically inspired learning principles for implementing flexible working memory systems.

Importantly, our simulations show that integrating principles from the spike-coding network framework with Hopfield associative memory networks can reproduce key computational motifs of short-term memory and temporal processing in the brain. Notably, this includes robustness to neuronal death, a natural process in biological neural networks. When a portion of the neurons was eliminated, the network's overall performance did not substan-

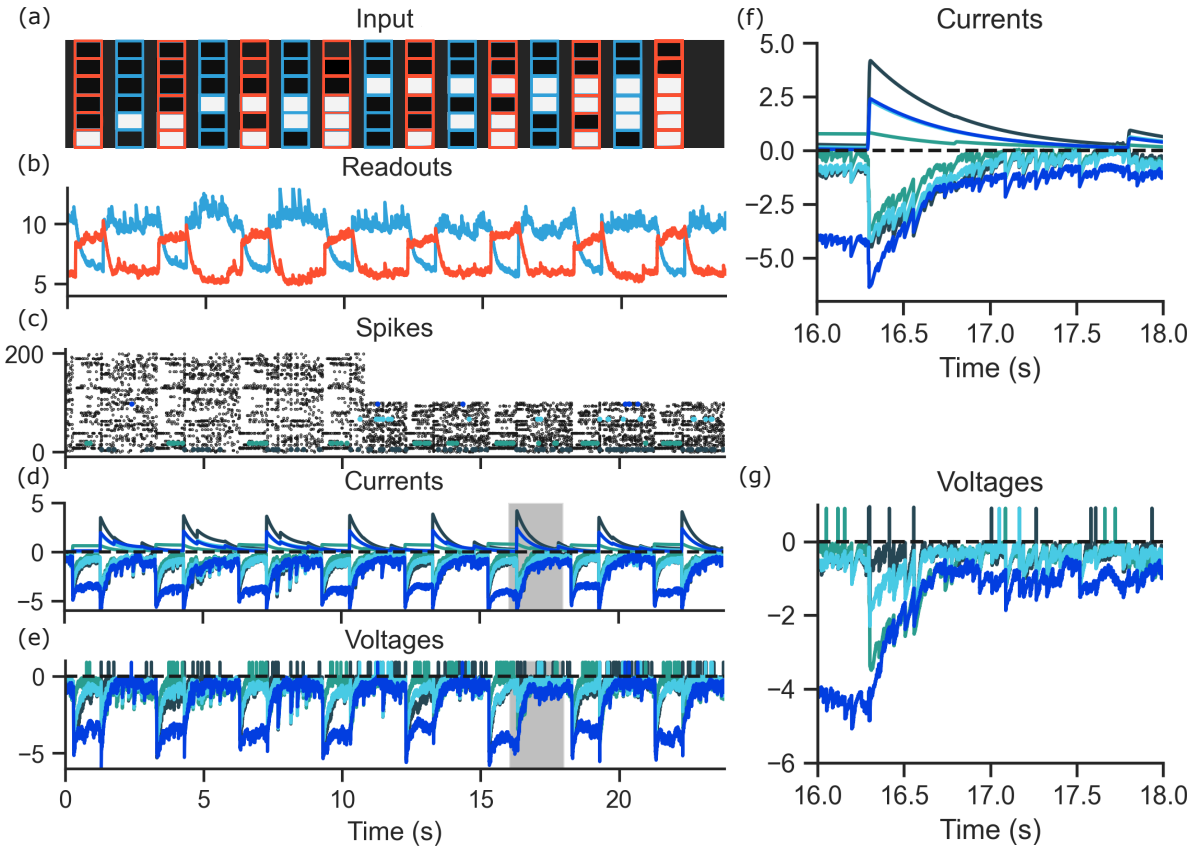


Figure 4.2: **Simulation results for a spiking neural network trained to perform the 5-back memory task.** (a) 6-pixel binary input vectors. (b) Network readout indicating match/non-match. (c) Underlying irregular spiking activity across neurons. (d) Example voltages exhibiting fluctuations around the threshold. (e) Input currents to example neurons. (f-g) Zoomed voltages and currents highlight balanced excitation/inhibition. The network exhibits complex dynamics, distributed encoding, and resilience to perturbations, including 40% unit deletions. Network parameters: $N = 200$, $\lambda = 2$, $\mathbf{\Omega}_{\text{slow}} = 0.1\mathbf{\Omega}_{\text{fast}}$, $\mu = 0.1$, $\tau_{\text{delay}} = 2$. Learning parameters: $\eta = 0.01$, $\alpha = 0.001$, decaying exponentially over 100 epochs. Noise parameters: $\sigma_V = 0.1$, $\sigma_{\text{stim}} = 0.1$, $\sigma_{\text{OU}} = 0.05$, $\lambda_{\text{OU}} = 10$.

tially change, indicating a high level of redundancy (Figure 4.2c) (Barrett et al., 2016). The remaining neurons compensated for this loss by increasing their firing rates, showcasing the adaptability of the network. Conversely, adding neurons decreases population firing rates,

and each neuron's activity can be arbitrarily sparse.

The network operates in a regime of balanced excitation and inhibition (Figures 4.2f and g). Under the influence of a given input, a group of neurons will oscillate between active patterns, maintaining a threshold-adjacent state with balanced underlying excitation and inhibition (Isaacson and Scanziani, 2011). Other neurons, detached from the current readout, experienced relatively more inhibition. This behavior aligns with research demonstrating that balanced networks can perform nonlinear transformations when subnetworks receive balanced inputs, while the rest receive more inhibition (Baker et al., 2020).

Finally, the network exhibited robustness against transmission delays. Spurious spikes induced by delays did not initiate cascades of subsequent spikes, hereby preventing the erroneous propagation of signals across the network. This resilience against delays could be linked to the network's ability to store sequences in slow synapses, a feature facilitated using a circulant matrix in our model. This functionality of the circulant matrix can be thought of as a form of "temporal error correction," where the precise time-stamp of neural firing is not essential, as long as the sequence of firing, or the relative timing, remains intact. This aspect contributes to the network's ability to handle and adapt to delays in neural transmission, further enhancing the network's robustness and adaptability (Borst and Theunissen, 1999).

4.5 Conclusion

This chapter presented a unifying framework that connects two seminal neural network models—Hopfield networks and spike-coding networks. Although differing in their emphasis on fixed point attractors versus spike timing, we proved they offer two complementary perspectives on the same underlying phenomenon—encoding associative memories and temporal sequences in recurrent networks.

The key insight underpinning this unification was recognizing the inherent linkage between these models; the stable attractor states in Hopfield networks provide a substrate for generating complex temporal spike sequences, which are dynamically replayed in spike-coding networks. In essence, Hopfield networks supply the associative memories, while spike-coding

networks provide the sequencing between these memories. Our analysis established a mathematical equivalence between these classic models by using circulant matrices to represent asymmetric weights in Hopfield networks. This introduction of asymmetry induced dynamics that enabled smooth transitions between stored attractor states.

Combining these two central yet formerly disconnected models in neural computation provides a unified lens for understanding the interlinked principles of associative memory and temporal sequence learning in recurrent networks. The proposed framework synergistically integrates complementary computational elements from each model into an integrated system. Hopfield networks contribute content-addressable associative memories implemented via attractor states. Spike-coding networks lend efficient temporal processing and sequencing enabled by precise spike patterns. Our unified perspective based on circulating dynamics between attractor states sheds light on the intricate mechanisms underlying sophisticated functions like working memory, thereby advancing theoretical models of neural computation.

Furthermore, introducing circulant matrices establishes an intriguing connection to CNNs. This link highlights potential cross-pollination between these biologically-inspired models and deep learning architectures. Insights from associative memories and spike-based dynamics could inspire enhancements in convolutional networks, such as embedding external memory modules or multi-timescale feature processing. This synthesis of foundational neural network models provides theoretical and practical value. It offers a principled integration of key concepts spanning decades of research in mathematical neuroscience. Additionally, it points towards further unification between biological and artificial neural systems by elucidating their shared algorithmic principles.

5.0 Cross-Modality Image Translation of 3 Tesla Magnetic Resonance Imaging to 7 Tesla using Generative Adversarial Networks

5.1 Chapter Summary

The rapid advancements in Magnetic Resonance (MR) Imaging technology have precipitated a new paradigm, wherein cross-modality data translation across diverse imaging platforms, field strengths, and disparate locales is increasingly challenging. This issue particularly accentuates when transitioning from 3 Tesla (3T) to 7 Tesla (7T) MR imaging systems. This study proposes a novel solution to these challenges by harnessing the power of Generative Adversarial Networks (GANs)—specifically, the CycleGAN architecture—to create synthetic 7T images from 3T data. Employing 1112 and 490 unpaired 3T and 7T MR images datasets, we trained CycleGAN models to translate 3T MR brain images into corresponding 7T data for spatial adaptive normalization across field strength. Alongside the traditional U-Net generator, we implement the Fast-Slow U-Net variant with separate local and global information processing pathways. Rigorous independent testing on 25 distinct subjects affirmed the model's proficiency in accurately predicting various tissue types, encompassing cerebral spinal fluid, gray matter, and white matter. Our approach provides a reliable and efficient methodology for synthesizing 7T images, thereby significantly aiding in the spatial adaptive normalization of longitudinal data. The Fast-Slow U-Net model showed particular promise in enhancing visual quality, with our results suggesting it results in more reliable and accurate segmentation. Furthermore, our results delineate the potential of GANs in amplifying the Contrast-to-Noise Ratio (CNR) from 3T, enhancing the diagnostic capability of the images. While acknowledging the potential risk of model overfitting, our research underscores a promising progression towards harnessing the benefits of 7T MR imaging systems in research investigations while preserving compatibility with extant 3T MR data. Thus, the research propounded herein is poised to contribute significantly to neuroimaging and the broader medical imaging domain.

5.2 Introduction

MR Imaging (MRI) technology has offered unprecedented, high-definition, and non-invasive insights into the intricacies of the human brain structure and function (Hagmann et al., 2006). This advancement has considerably enhanced our understanding of the brain's morphological and functional changes across the lifespan and in various neurodegenerative conditions (Ashburner and Friston, 2000). In addition, it has spearheaded the discovery of potential biomarkers for numerous neurological disorders (Ross et al., 2012), thereby opening the doors to early detection and therapeutic intervention.

MRI has shifted from conventional/clinical 3T MRI systems towards high-field strength 7T MRI systems (Uğurbil, 2014). The benefits 7T MRI include augmented signal-to-noise ratio (SNR), enhanced visualization of intricate brain structures due to reduced voxel size, and an improved CNR resulting from faster tissue relaxation times (Uğurbil, 2014). Despite these advantages, the increased costs, restricted availability, increased susceptibility artifacts, a higher specific absorption rate (SAR), and radiofrequency (RF) inhomogeneity may complicate image acquisition and interpretation (van Osch and Webb, 2014).

The field has seen significant advancements in RF coil design and application that are instrumental in overcoming the challenges of 7T MRI, chiefly ensuring high-quality and high-resolution acquisition (Santini et al., 2021b). Yet, transitioning longitudinal patient studies from the widely-used 3T systems to the more advanced 7T systems presents substantial hurdles. The primary reason for this is the innate differences in the images generated by the two systems, especially the contrast variations among soft tissues. These may affect direct image comparisons and complicate longitudinal analyses (Obusez et al., 2018).

Other factors, such as the temporal variability in image quality and systematic differences between varying field strengths, can complicate statistical analyses of aggregate data (Fortin et al., 2018). These difficulties underscore the need for efficient spatial adaptive data normalization methodologies capable of handling the variances across different field strengths; they are paramount in strengthening our comprehension of the neuroanatomical shifts and progressions associated with normal aging and pathological processes.

AI and machine learning technologies have significantly shaped numerous sectors, includ-

ing healthcare (Bohr and Memarzadeh, 2020). Notably, medical imaging is being improved by the potential of deep learning methods, extending their reach into the complex realm of neuroimaging (Shen et al., 2017). Specifically, applying CNNs in image classification tasks has improved diagnostic accuracy and efficiency (Litjens et al., 2017). Building on this momentum, GANs and their variants have emerged as powerful tools for medical image analysis, including tasks like image synthesis, segmentation, and data augmentation (Nie et al., 2018).

Introduced by Goodfellow et al. (2014), GANs comprise a unique model structure employing two deep neural networks: the generator and the discriminator. They engage in an adversarial process, aiming to generate synthetic data resembling the original distribution. These networks have demonstrated proficiency in modeling complex data distributions (Arjovsky et al., 2017, Salimans et al., 2016), proving beneficial for various image tasks.

CycleGANs, an extension of GANs, provide a solution for unpaired image-to-image translation tasks (Zhu et al., 2017), including photo enhancement (Chen et al., 2017), style transfer (Gatys et al., 2016), and image synthesis (Wang et al., 2018). Within the realm of medical imaging, its applications have been transformative. For example, they have shown potential in tasks like lesion synthesis (Guerrero et al., 2018). These networks can enrich rare disease datasets by generating synthetic lesion images, facilitating better disease detection and diagnosis. Additionally, they have been used for organ segmentation by transforming the images into simpler representations that ease the segmentation process (Cai et al., 2019).

One significant application pertains to cross-modality image translation, such as converting Computed Tomography (CT) images to MR images and vice versa (Wolterink et al., 2017). In practical scenarios, patients may not undergo CT and MR scans due to cost, radiation exposure, or other considerations. By training on unpaired CT and MR datasets, CycleGAN has demonstrated the capability to generate synthetic but anatomically accurate MR images from CT scans and, conversely, to produce CT-like images from MR scans.

Cross-modality image translation, defined by the machine learning and imaging community, consists of converting images from one modality to another, retaining the crucial content while modifying the style to resemble the target modality (Zhu et al., 2017). 3T and 7T, despite employing the identical principle of nuclear MR, exhibit divergent operational mechanics due to distinct field strengths. This differentiation manifests in significant

contrasts in image quality, resolution, CNR, susceptibility effects, and spectral separation (Vaughan et al., 2001). Thus, the transition from 3T to 7T MRI goes beyond reducing systematic variations within the same data type; it signifies a paradigm shift accentuated by profound alterations in image properties. In this regard, the transformation from 3T to 7T surpasses data harmonization's scope, warranting 3T MRI to be considered one modality and 7T MRI another. The conversion process from 3T to 7T via GANs fulfills this criterion, as it retains the intrinsic information (content) from the 3T images and translates the image properties (style) to align with that of 7T MRI.

This study harnesses a dataset of unpaired 3T and 7T MR images to train CycleGAN-based models capable of translating 3T MR data into synthetic 7T data. This innovative approach is directed towards addressing the issue of spatial adaptive MR data normalization, allowing the scientific community to extract the benefits offered by the 7T systems while still utilizing previous longitudinal data acquired at 3T.

Recent advancements in CNN design have led to models that separate the processing of high and low-frequency features through multiple pathways (Simonyan and Zisserman, 2014, Feichtenhofer et al., 2016, Tran et al., 2015). Based on the results of Chapter 4, we propose a fast-slow architectural motif containing distinct fast and slow pathways for local and global information, respectively. The fast pathway focuses on extracting fine details using smaller convolutional kernels, while the slow pathway captures broader semantic context through dilated convolutions. The Fast-Slow U-Net variant of CycleGAN is a potential architecture option to enhance cross-modality 3T to 7T MRI translation.

5.3 Materials and Methods

The landscape of machine learning has been significantly shaped by the emergence of GANs, as conceptualized by Goodfellow et al. (2014). This novel model structure, built upon game-theoretic principles, employs two deep neural networks in strategic interaction: the generator and the discriminator. The generator aims to generate synthetic data replicating the actual data distribution, while the discriminator strives to discern the synthesized data

from the real data. This adversarial dance evolves over iterative training sessions, yielding refined performance from both networks and improved proficiency in modeling complex data distributions (Arjovsky et al., 2017, Salimans et al., 2016).

Taking a step further, CycleGANs provides a robust solution to unpaired image-to-image translation, a challenging problem in computer vision. Zhu et al. (2017) introduced CycleGAN's architecture, comprising two GANs containing a generator and a discriminator. The generator creates compelling images of the target domain, deceiving the discriminator, which strives to differentiate real images from the translated ones within the target domain. The design of CycleGAN is such that each generator maps an input image from a source domain to a target domain and vice versa to ensure cyclical consistency (Zhu et al., 2017).

A feature that distinctly characterizes CycleGAN from traditional GANs is its cycle consistency loss. This loss function aims to ensure that the cycle of translating an image from one domain to another and back again reconstructs the original image. Let $\mathcal{X}$ denote the domain of images, and $\mathcal{Y}$ denote another domain of images. We can represent an image from $\mathcal{X}$ as $x \in \mathcal{X}$, and an image from $\mathcal{Y}$ as $y \in \mathcal{Y}$. Let's construct image datasets from each domain: $\mathcal{D}_{\mathcal{X}} = \{x_1, x_2, \ldots, x_m\}$ consisting of M images from $\mathcal{X}$, and $\mathcal{D}_{\mathcal{Y}} = \{y_1, y_2, \ldots, y_n\}$ consisting of N images from $\mathcal{Y}$. Now, let the mappings $G : \mathcal{X} \to \mathcal{Y}$ and $F : \mathcal{Y} \to \mathcal{X}$ be the generators between the two domains. Then, the cycle consistency loss is:

$$L_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim \mathcal{D}_{\mathcal{X}}} \left[\|F(G(x)) - x\|_1 \right] + \mathbb{E}_{y \sim \mathcal{D}_{\mathcal{Y}}} \left[\|G(F(y)) - y\|_1 \right]. \tag{5-1}$$

$\mathbb{E}_{x \sim \mathcal{D}_{\mathcal{X}}}$ and $\mathbb{E}_{y \sim \mathcal{D}_{\mathcal{Y}}}$ are the expectations taken over the datasets $\mathcal{D}_{\mathcal{X}}$ and $\mathcal{D}_{\mathcal{Y}}$, respectively. The terms $\|F(G(x)) - x\|_1$ and $\|G(F(y)) - y\|_1$ measure the absolute difference between the original image and the image that has been translated to the other domain and back, *i.e.*, $F(G(x))$ means G translates an image $x \in \mathcal{D}_{\mathcal{X}}$ to domain $\mathcal{Y}$ and then translated to $\mathcal{X}$ by F. Similarly, $G(F(y))$ takes an image $y \in \mathcal{D}_{\mathcal{Y}}$, translates it to domain $\mathcal{X}$ by F and then translates it back to $\mathcal{Y}$ by G. The cycle consistency loss should be small if the model performs well, indicating that the original and the cycled-back image are nearly identical.

Our study uses the CycleGAN framework to deal with unpaired 3T and 7T MR data. Figure 5.1 shows a representation of the architecture and functioning of the overall model. We implement two variations of the CycleGAN; the original integrates two generator networks

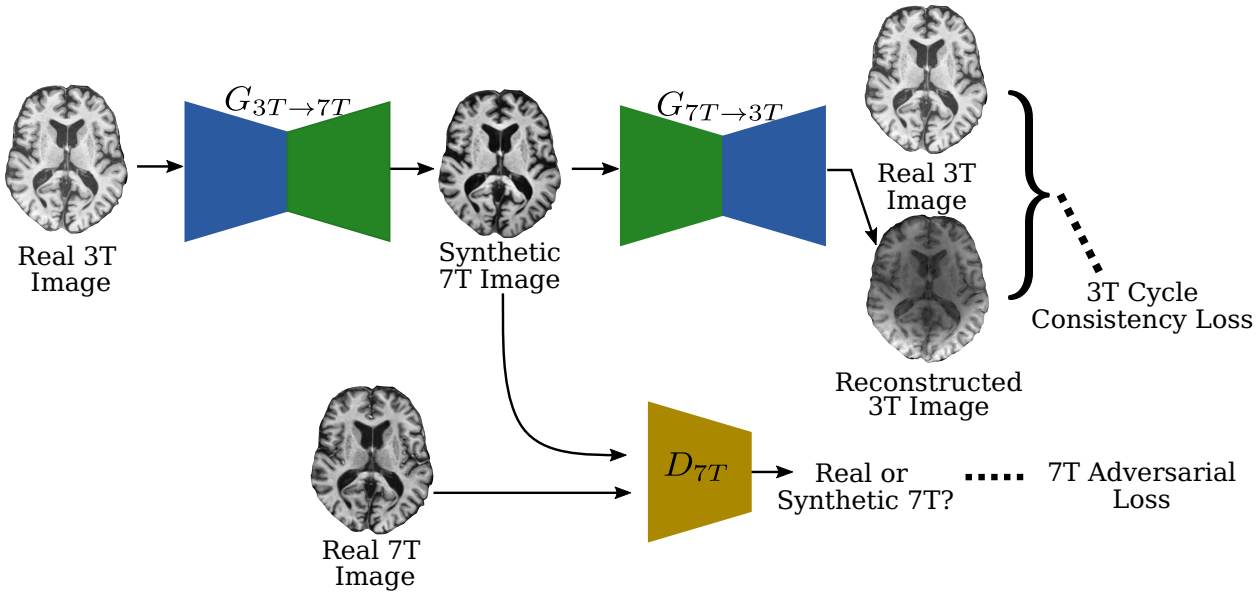


Figure 5.1: The cyclical consistency of the CycleGAN. This figure visually demonstrates the 3T cycle consistency. The 7T loop is conceptually symmetric. The generators are marked as performing the mappings $G_{3T\to 7T} : 3T \to 7T$ and $G_{7T\to 3T} : 7T \to 3T$. The real 3T image is fed to the $G_{3T\to 7T}$ (top left), yielding the synthetic 7T image (top middle). The D_{7T} discriminator differentiates between the real and synthetic 7T images and strives to minimize the 7T adversarial loss (bottom left). The synthetic 7T image is fed into the $G_{7T\to 3T}$ generator, yielding the reconstructed 3T Image. The 3T cycle consistency loss ensures the 3T image can be reconstructed from the synthetic 7T image (top right).

based on the U-Net architecture (Ronneberger et al., 2015), with inserted residual blocks between the encoding and decoding stages (He et al., 2016). The variant implements a dual pathway U-Net inspired by the connections between fast and slow processing in spike-coding, Hopfield networks, and convolutional neural circuits discussed in Section 4.3.5.

U-Net is a CNN architecture renowned for its utility in biomedical image segmentation. It is named after its U-shaped structure, consisting of a contracting path for capturing context and a symmetric expanding path for precise localization. The expanding pathway uses transposed convolutions, with skip connections transferring feature maps from the contract-

ing path to recover spatial information.

Residual blocks, proposed in the ResNet architecture by He et al. (2016) resolve training difficulties encountered with deep neural networks. These blocks consist of several convolutions, after which the input is added to the output, creating a residual connection. This configuration allows the network to learn residual functions concerning the layer inputs, facilitating the training of deeper models.

Both models use instance normalization and stride-2 convolutions. Instance normalization is often used in style transfer tasks (Ulyanov et al., 2016). It computes the mean and variance for each sample separately, leading to the relative scaling of activations. This technique is beneficial for content and style disentanglement for image generation tasks. Stride-2 convolutions help reduce the spatial dimensions of feature maps and allow the model to learn more abstract representations in deeper layers (Goodfellow et al., 2016).

Finally, our CycleGAN models use 70×70 PatchGANs for the discriminator networks. PatchGANs, as described by Isola et al. (2017), classify whether each patch in an image is real or fake instead of assessing the picture as a whole. This patch-level classification enables the generation of sharper and more contextually accurate images.

5.3.1 Fast-Slow U-Net

The key to implementing fast and slow separation in a CNN is to process different information frequencies separately, combine them, and ensure that learning is effective across the entire network. This requires a purposeful reconfiguration of the network's architecture, and we draw inspiration from the hierarchical nature of biological networks. The "fast" and "slow" features can be interpreted as separate components of the CNN architecture, each responsible for different types of information processing. Let us consider a two-pathway architecture, also known as a dual-stream architecture (Simonyan and Zisserman, 2014).

The "fast" pathway can be designed to respond to high-frequency changes in the input, capturing immediate, rapid changes in the scene. It can be characterized by small, densely connected convolutional layers that capture fine-grained detail. One possible approach is to create a shallow architecture with smaller receptive fields to handle high-frequency, spatially

detailed features (Canziani et al., 2016).

On the other hand, the "slow" pathway should be architected to detect and encode lower-frequency, more abstract features. It provides a more contextual picture of the input, capturing long-term dependencies and integrating information over time. This can be implemented with deeper convolutional layers with larger receptive fields. Another critical aspect of the slow pathway is temporal integration. In video processing, this could be achieved by 3D convolutions or recurrent layers, which can integrate information across frames (Tran et al., 2015). An equivalent mechanism incorporating a larger spatial context could be used for non-temporal applications like cross-modality image translation. One approach is to use dilated convolutions, which allow the network to have a larger receptive field without increasing the number of parameters or the amount of computation (Yu and Koltun, 2015).

The idea behind dilated convolution is straightforward. Instead of sliding the convolutional kernel through the input in contiguous steps, the operation is performed with gaps or "dilations" between the input units. This dilation process is controlled by a parameter known as the dilation rate, d. When $d = 1$, the operation is equivalent to standard convolution. For $d > 1$, the kernel is spread over a larger input area, effectively increasing the receptive field without adding more parameters (Yu and Koltun, 2015).

The primary advantage of dilated convolution is its ability to exponentially expand the receptive field without loss of resolution or coverage overlap and a significant increase in computational load. This is crucial in applications like semantic segmentation, where capturing long-range dependencies at a high resolution is essential (Wang et al., 2018). It has been used in several influential models. In the DeepLab-v3 semantic segmentation architecture, it is used to extract multi-scale features effectively (Chen et al., 2017). It has also found use in audio generation and text-to-speech synthesis models, such as Google's WaveNet, allowing the model to capture information across different time scales (Oord et al., 2016).

To combine the outputs from the fast and slow pathways, we can fuse them at each corresponding layer before proceeding to the next. The exact fusion method (*e.g.*,, concatenation, weighted sum) depends on the specific task and the nature of the data (Feichtenhofer et al., 2016). We can create a multi-level, multi-receptive field feature representation by concatenating the feature maps produced by each path along the channel dimension.

Here we propose the Fast-Slow U-Net, based on the traditional U-Net structure (Ronneberger et al., 2015) but with a significant modification involving two distinct fast and slow pathways. This modification allows for processing and exploiting the input data at different scales and contexts, allowing the model to balance efficiency and performance. The fast pathway captures fine-grained details, and the slow pathway accounts for the global context. Figure 5.3.1.2 shows a schematic representation of the modified U-Net architecture. The following sections describe each pathway and its building blocks in detail.

5.3.1.1 Fast Pathway

The fast pathway is a standard feed-forward CNN that extracts low-level details from the image. In the specific model described, the input to the fast pathway is a grayscale image of size $224 \times 224 \times 1$. The first layer of the fast pathway includes two sequential 3×3 padded convolutions with 64 filters each, followed by instance normalization and ReLU activations (Glorot et al., 2011). After this stage, a 2×2 max pooling operation (Scherer et al., 2010) reduces the spatial dimension of the feature maps by half ($112 \times 112 \times 64$).

Subsequently, similar convolutional, ReLU, and max-pooling layers are applied, but this time with 128 filters (resulting in feature maps of size $56 \times 56 \times 128$), followed by other similar layers with 256 filters (resulting in $28 \times 28 \times 256$ feature maps), and 512 filters (resulting in $14 \times 14 \times 256$ feature maps). Finally, the bottleneck comprises two sequential 3×3 padded convolutions with 1024 filters each, followed by instance normalization and ReLU activations ($14 \times 14 \times 1024$ features). The model can rapidly capture the finer details in the images through the fast pathway, providing a granular understanding of the input image.

5.3.1.2 Slow Pathway

The slow pathway operates concurrently with the fast pathway but employs dilated convolutions to capture more global and contextual information from the image. In the first layer of the slow pathway, two 3×3 dilated convolutions with a dilation rate of $d = 2$ are used, each having 64 filters and followed by ReLU activations. A max pooling operation then halves the spatial dimension, similar to the fast pathway.

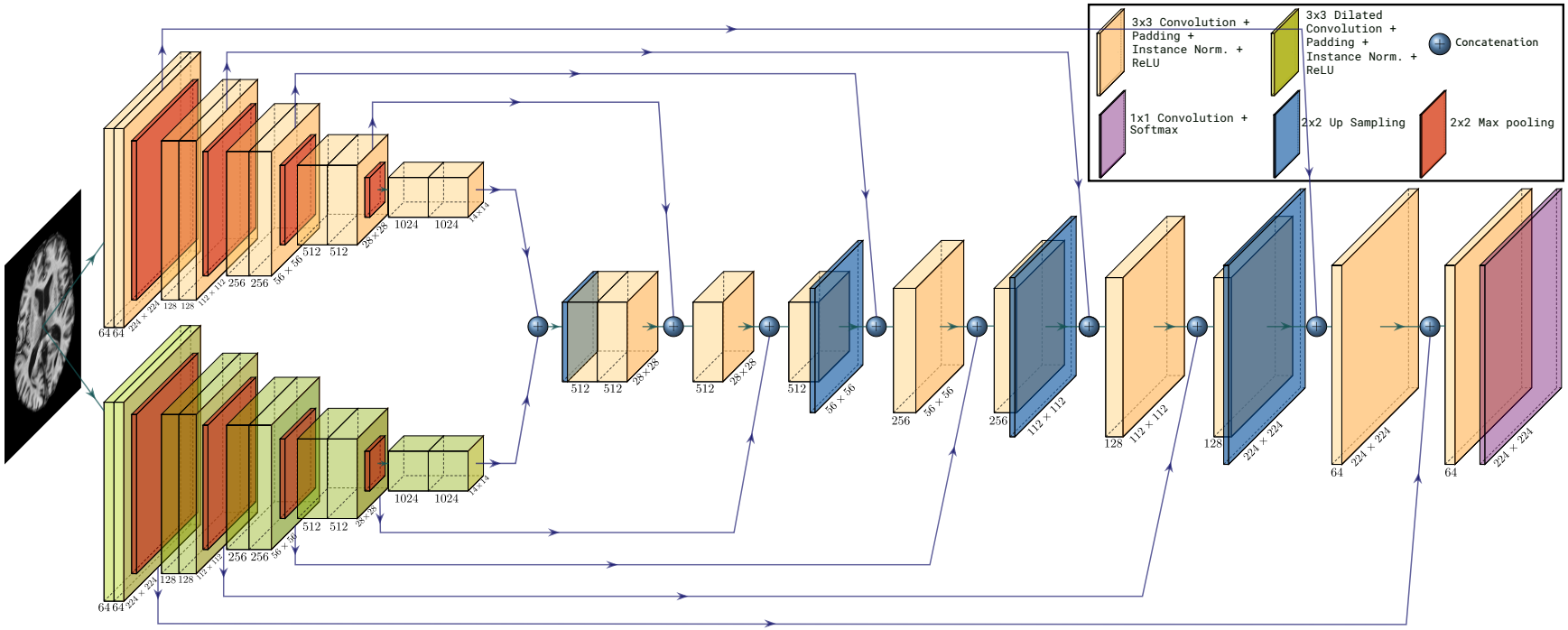


Figure 5.2: A schematic representation of the Fast-Slow U-Net architecture. The dual-pathway model, consisting of a fast pathway (left) and a slow pathway (right), processes high-frequency and low-frequency details separately. The fast pathway rapidly captures fine-grained details using standard convolutions, while the slow pathway aggregates broader contextual information through dilated convolutions. Features from both pathways are fused at different levels, ensuring effective multi-scale feature learning. The architecture culminates in a 1×1 convolution, producing a single-channel output representation.

The subsequent layers follow the same design but with dilation rates of $d = 4$, $d = 8$, $d = 16$, and $d = 16$, respectively, and an increasing number of filters (128, 256, 512, and 1024), resulting in final feature maps of size $14 \times 14 \times 1024$. The outputs of each dilated convolution were padded to match the corresponding layers in the fast pathway. This approach enables the slow pathway to aggregate more considerable contextual information, essential in many segmentation tasks, such as object identification and boundary localization.

5.3.1.3 Merging and Upsampling

Upon reaching the U-Net bottleneck, the merging and upsampling process begins. First, the bottlenecks of the slow and fast pathways are concatenated and upsampled to twice their spatial dimensions and then passed through two 3×3 convolutional layers with instance normalization and ReLU activation. Then, the resulting feature maps are concatenated with the corresponding features from the fast pathway, creating a combined feature map. The concatenated features are passed through a 3×3 convolutional layer with instance normalization and ReLU activation. After, they are concatenated with the corresponding slow pathway features, followed by a 3×3 convolutional layer with instance normalization and ReLU activation. This process is repeated for feature maps $56 \times 56 \times 256$, $112 \times 112 \times 128$, and $224 \times 224 \times 64$. This step leverages the strengths of both pathways by fusing fine-grained, local details from the fast pathway with broader, contextual details from the slow pathway.

5.3.1.4 Final Layer

In the Fast-Slow U-Net architecture, the final layer employs a 1×1 convolutional operation with a single filter. The 1×1 convolution operation, also known as network-in-network operation, has a pivotal role in transforming the depth dimension of the feature maps (Springenberg et al., 2014). It effectively functions as a channel-wise fully connected layer, relating features at each spatial location while maintaining the spatial structure of the image.

A single feature is extracted from the output of the previous layer. This single feature map encodes the transformed content of the input image in the same style, with the spatial dimensions preserved. The final output is, therefore, an image of size $224 \times 224 \times 1$.

5.3.2 Dataset

We leveraged an extensive collection of unpaired and paired 3T and 7T MR images from diverse sources to ensure a rich and comprehensive dataset. The 3T MR data were primarily sourced from the Human Connectome Project (HCP), a globally recognized brain imaging dataset that contains high-definition, multimodal brain imaging data from healthy adults (Sotiropoulos et al., 2013). Specifically, structural 3T images were sourced from a subset of 1112 scans in the HCP's 1200 subjects data release. Images were acquired using a Siemens 3T Connectome Skyra and a standard 32-channel Siemens receive head coil at Washington University in St. Louis. The scans were conducted using a T1-weighted (T1w) Magnetization Prepared RApid Gradient Echo (MPRAGE) protocol with a repetition time (TR) of $2400\,ms$, echo time (TE) of $2.14\,ms$, flip angle (FA) of $8°$, field of view (FOV) of $224\,mm \times 224\,mm$, and $0.7\,mm$ isotropic resolution, lasting approximately 8 minutes.

The inclusion criteria for the 3T HCP data required participants to be aged between 22 and 35 years with no significant history of psychiatric disorder, substance abuse, neurological or cardiovascular disease, and the ability to provide valid informed consent. Participants were required to have a Mini Mental Status Exam score above 28.

Exclusion criteria included multiple non-provoked seizures or a diagnosis of epilepsy, genetic disorders such as cystic fibrosis, use of prescription medications for migraines in the past 12 months, and conditions such as multiple sclerosis, cerebral palsy, brain tumor, stroke, sickle cell disease, thyroid hormone treatment in the past 12 months, current treatment for diabetes, head injury, and premature birth. Other exclusion criteria were the presence of unsafe metal or devices in the body (*e.g.*, cardiac pacemakers, cochlear implants, aneurysm clips), current or historical use of chemotherapy or immunomodulatory agents that could affect the brain, pregnancy, and moderate to severe claustrophobia.

The 7T dataset was sourced from the 7TBRP dataset, acquired at the RF Research Facility at the University of Pittsburgh, a specialized research center for high-field imaging. The dataset included 490 subjects' structural 7T MR images. The scans were acquired using a Siemens Magnetom 7T whole-body MR scanner and an in-house developed RF 7T Tic-Tac-Toe head coil system (Santini et al., 2018, 2021b, Krishnamurthy et al., 2019, Santini et al.,

2021a). The scans were conducted using a T1w MPRAGE protocol with a TR of 3000 ms, TE of 2.17 ms, bandwidth of 391 Hz/Px, GRAPPA reconstruction with acceleration factor $R = 2$, and 0.75 mm isotropic resolution, lasting approximately 5 minutes.

Exclusion criteria for the 7TBRP dataset included pregnancy or lactation, acute medical problems that could result in neurocognitive or brain dysfunction, including diabetes mellitus, coronary artery disease, and causes cerebral vasculities, such as peripheral vascular disease. Other exclusion criteria were contraindications to MR scanning such as electronic implants, magnetically-activated implants, tattoos above the shoulders, or brain implants. Data acquisition across all studies was conducted following the protocol approved by local institutional review boards. All participants were older than 18 and able to provide written informed consent before participation.

A set of paired 3T-7T data from the 7TBRP dataset were used to compare 3T and 7T data. The data set consisted of MR images from 25 subjects. The study-specific inclusion and exclusion criteria for the paired data were similar to those for the unpaired 7T datasets.

5.3.3 Data Preprocessing

A fundamental step in the MR data pipeline, data preprocessing is aimed at standardizing the data originating from varied sources and subjects. This step is a mitigation measure for inter-subject and inter-scanner variability (Nyul and Udupa, 1999, Tustison et al., 2010). The inherent variations in MR data necessitate the application of several preprocessing stages to ensure the data's reliability and consistency.

The first stage of preprocessing, bias field correction, was conducted using FMRIB's Automated Segmentation Tool (FAST) part of FSL (Zhang et al., 2001). This technique mitigates the intensity inhomogeneity observed within MR images due to magnetic field variations or patient positioning differences (Vovk et al., 2007). If not correctly trained, deep learning models can learn from these inhomogeneities and become biased, leading to sub-optimal performance. Bias field correction can rectify this, enabling the model to concentrate on the salient features of the image (Arnold et al., 2001). This process results in enhanced feature extraction, improved segmentation, and registration outcomes, bolstering

the effectiveness of subsequent processing and training phases (Tustison et al., 2010).

Following bias field correction, we performed spatial normalization using the Statistical Parametric Mapping (SPM) toolbox's coregister function, co-registering the images via normalized mutual information to the Montreal Neurological Institute 152 standard (MNI152) space (Ashburner, 2007, Penny et al., 2011). Normalization facilitates the comparison and integration of data across subjects, as each voxel corresponds to the same anatomical structure across all images (Evans et al., 2012). This alignment minimizes the need for the neural network to learn these invariances, potentially improving the performance (Cao et al., 2017).

In the next stage, we transformed the image intensities to a standard normal range of $[-1, 1]$ using SPM's image calculator function. This transformation minimizes inter-subject variability, fostering improved image data consistency (Nyul and Udupa, 1999). Additionally, intensity normalization allows the model to focus more on structural or semantic differences across images than absolute pixel intensities (Goodfellow et al., 2016). This is critical during the early stages of training as it prevents the gradients from either vanishing or exploding, resulting in a more stable optimization process and quicker convergence (LeCun et al., 2002). Furthermore, intensity normalization can also make the learning process less dependent on the specific units used. This is especially important when combining data from multiple sources with different measurement units or scales (Bishop and Nasrabadi, 2006).

Next, we skull-stripped the images using the Brain Extraction Tool (BET) within FSL, removing non-brain tissue from the MR images and thereby reducing dimensionality, which can enhance the efficiency of subsequent model training (Iglesias and Sabuncu, 2015, Smith, 2002). For computational efficiency, we transformed 3D images into 2D by extracting axial slices using FSL's fslslice function, which is less computationally intensive while preserving critical anatomical information.

5.3.4 Model Training

Training a deep learning model requires orchestrating various strategies and methods to optimize performance. In the context of our CycleGAN models implemented within the TensorFlow framework (Abadi et al., 2016), we had to consider various elements such as com-

putation hardware, optimization algorithms, weight initialization, regularization techniques, and learning rate schedules.

We used NVIDIA TitanX GPUs for training and adopted an iterative training strategy; the discriminators were updated four times for each update of the generators within the training loop. This prevented the generators from overpowering the discriminators, which is critical to maintaining the dynamics of the adversarial process (Goodfellow et al., 2014).

We used early stopping as a regularization technique to control the model's capacity and prevent overfitting. We halted the training when no significant improvement in the validation set's performance was observed, *i.e.*, the validation loss did not decrease by 0.1% over five consecutive epochs (Prechelt, 2002). Furthermore, L1/L2 regularization was employed to encourage the model to learn distributed and sparse representations, which can increase its generalization ability (Ng, 2004).

We used the Adam optimizer (Kingma and Ba, 2014) for the optimization algorithm. Its adaptive learning rate and efficient computation makes it a superior choice for training deep learning models, especially in achieving faster and more effective convergence. The Xavier weight initialization method was used for all the convolutional layers (Glorot and Bengio, 2010). Additionally, all the bias terms were initialized to zero. Proper weight initialization can significantly improve the convergence rate and prevent problems like vanishing or exploding gradients, making the optimization process more stable.

Finally, we used learning rate annealing to fine-tune the speed at which the model learns. The learning rate was set at 10^{-4} for the first 100 epochs and was linearly decreased to zero afterward. By gradually reducing the learning rate, the model can make significant updates to learn the gross structure of the data in the early stages and then fine-tune its weights in later stages, allowing for smoother and more effective convergence (Bengio, 2012).

5.3.5 Model Evaluation

The traditional approach to assessing GANs, primarily focusing on discriminator loss, needs an explicit measure of sample quality. This work thus evaluates the quality of training by leveraging a pre-trained deep neural network to embed the samples, inspired by Heusel

et al. (2017) methodology. We adopted a ResNet-50 architecture pre-trained on ImageNet (He et al., 2016), given its effectiveness in similarity searching in the embedding space for various forms of imaging data, including MRI, even without fine-tuning (Yosinski et al., 2014). The quality of samples, in this context, is evaluated based on their realism and "closeness" to the target domain. We calculated the Fréchet ResNet Distance (FRD) and the Cosine ResNet Distance (CRD) between each sample in the ResNet-50 embedding space to facilitate a comparison between 3T and 7T MRI features.

Let $\mathcal{D}_{\mathcal{X}'} = G(\mathcal{D}_{\mathcal{Y}})$ be the dataset of synthetic samples and ResNet($\mathcal{D}_{\mathcal{X}}$) and ResNet($\mathcal{D}_{\mathcal{X}'}$) denote the ResNet-50 embeddings of real and synthetic samples, respectively. The FRD, computed as the Fréchet distance (Dowson and Landau, 1982) or as the Wasserstein-2 distance (Vallender, 1974) between the ResNet($\mathcal{D}_{\mathcal{X}}$) and ResNet($\mathcal{D}_{\mathcal{X}'}$) embeddings is:

$$\mathrm{FRD}\left(\mathcal{N}(\,\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}}\,), \mathcal{N}\left(\mu_{\mathcal{X}'}, \Sigma_{\mathcal{X}'}\right)\right)^2 = \|\mu_{\mathcal{X}} - \mu_{\mathcal{X}'}\|_2^2 + \mathrm{Tr}(M), \tag{5-2}$$

where

$$M = \left(\Sigma_{\mathcal{X}} + \Sigma_{\mathcal{X}'} - 2\left(\Sigma_{\mathcal{X}}^{\frac{1}{2}} \cdot \Sigma_{\mathcal{X}'} \cdot \Sigma_{\mathcal{X}}^{\frac{1}{2}}\right)^{\frac{1}{2}}\right),$$

$\mathcal{N}\left(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}}\right)$ and $\mathcal{N}\left(\mu_{\mathcal{X}'}, \Sigma_{\mathcal{X}'}\right)$ are the multivariate Gaussian distributions over ResNet($\mathcal{D}_{\mathcal{X}}$) and ResNet($\mathcal{D}_{\mathcal{X}'}$), respectively, $\|\cdot\|_2$ represents the Euclidean norm, and Tr the trace.

Likewise, the CRD is the average cosine distance between each real $z \in \mathrm{ResNet}(\mathcal{D}_{\mathcal{X}})$ and synthetic $z' \in \mathrm{ResNet}(\mathcal{D}_{\mathcal{X}'})$:

$$\mathrm{CRD} = 1 - \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{z_i \cdot z_j'}{\|z_i\|_2 \cdot \|z_j'\|_2}, \tag{5-3}$$

where z_i represents the i-th feature embedding in ResNet($\mathcal{D}_{\mathcal{X}}$) and z_j' represents the j-th feature in ResNet($\mathcal{D}_{\mathcal{X}'}$). M and N are the total real and synthetic samples, respectively. This expression calculates the cosine similarity between each pair of feature embeddings, normalizes it by their Euclidean norms, and then averages the distances over all pairs.

Finally, we measured the sample quality over the synthetically generated 7T MR images. The accuracy of the generated synthetic 7T MR images was thoroughly scrutinized by cross-referencing these images with their respective real 7T ground truth counterparts. Images were segmented using FSL's FAST segmentation tool. Specifically, we used two key and

widely used evaluation metrics in medical imaging: the Dice Coefficient (Dice) and the Percentual Area Difference (PAD). These metrics were indispensable for evaluating the model's performance and provided imperative insights for future model enhancement.

The Dice, introduced by Dice (1945), measures the overlap between the predicted and the ground truth segmentation. Mathematically, for an expected pixel set $\mathcal{A}$ and the ground truth pixel set $\mathcal{B}$, the Dice is:

$$\text{Dice} = 2 \cdot \frac{|\mathcal{A} \cap \mathcal{B}|}{(|\mathcal{A}| + |\mathcal{B}|)}, \tag{5-4}$$

where $|\cdot|$ represents the cardinality of a set. The range of Dice spans from 0 (indicating no overlap) to 1 (indicating a perfect overlap), rendering it an excellent indicator of the congruence between predicted and real segmentation (Zou et al., 2004).

Conversely, the PAD measures the discrepancy in the size between the predicted and the ground truth segmentation. Specifically, it evaluates the model's accuracy in estimating the size of the segmented objects. For an expected area A and the ground truth segmentation area B, the PAD is computed as:

$$\text{PAD} = 100\% \cdot \frac{|A - B|}{B}, \tag{5-5}$$

where $|\cdot|$ is the absolute value. The PAD is inversely proportional to the performance of the segmentation, with smaller values indicating superior performance (Huang et al., 2019).

5.4 Results

The performance of the model was assessed using multiple parameters, including generators, discriminators, and cycle consistency losses in the validation dataset; FRD and CRD embedding quality measures; and a detailed analysis of the prediction for various brain tissue types and Dice and PAD sample quality metrics from independent test data.

5.4.1 Training Performance

We assessed the models' performance by first analyzing the generator and discriminator losses in the validation dataset. The generator and discriminator essentially participate in a minimax game, where the generator seeks to create images that the discriminator cannot distinguish from the real 7T images. In contrast, the discriminator aims to accurately classify real images from synthetic ones (Goodfellow et al., 2014).

Figures 5.3a and 5.3b show that the tug-of-war between generators and discriminators (for both CycleGAN models) resulted in the generators' losses gradually decreasing, demonstrating the generators' increasing proficiency in creating images that the discriminators struggle to differentiate from real MR images. In tandem, the discriminators' losses also decreased, indicative of their improving ability to distinguish between real and generated images. At minima, the convergence of the losses substantiates the model's efficacy in producing synthetic MR-like images from real data.

However, this is evidence of improvement only during the early training stages. Beyond that, since generator and discriminator losses plateaued after the 14th epoch, the cycle-consistency losses suggest that the output kept improving. Hence, to quantitatively evaluate the progression of the CycleGAN models, we apply the FRD and CRD metrics described in Methods 5.3.5 to the data generated at each training epoch. This strategy allows us to scrutinize the training dynamics of the CycleGANs. As shown in Figures 5.3c and 5.3d, FRD and CRD though slightly noisy, stop decreasing beyond the 90th and 140th epochs, for the Fast-Slow and Traditional CycleGAN, respectively. Overall, Figure 5.3 demonstrates the embedded MR images capture more relevant features, with measures of transformation quality showing a consistent and smooth decrease throughout the training process.

5.4.2 Visual Representation of Results

Further, an independent test dataset evaluated the models' generalization ability, where we used the models to translate 3T images from new subjects, unseen during training, into synthetic 7T images. Figure 5.4 exhibits examples of real and synthetic 7T MR image pairs generated by the Traditional CycleGAN model. Immediately, the synthetic images manifest

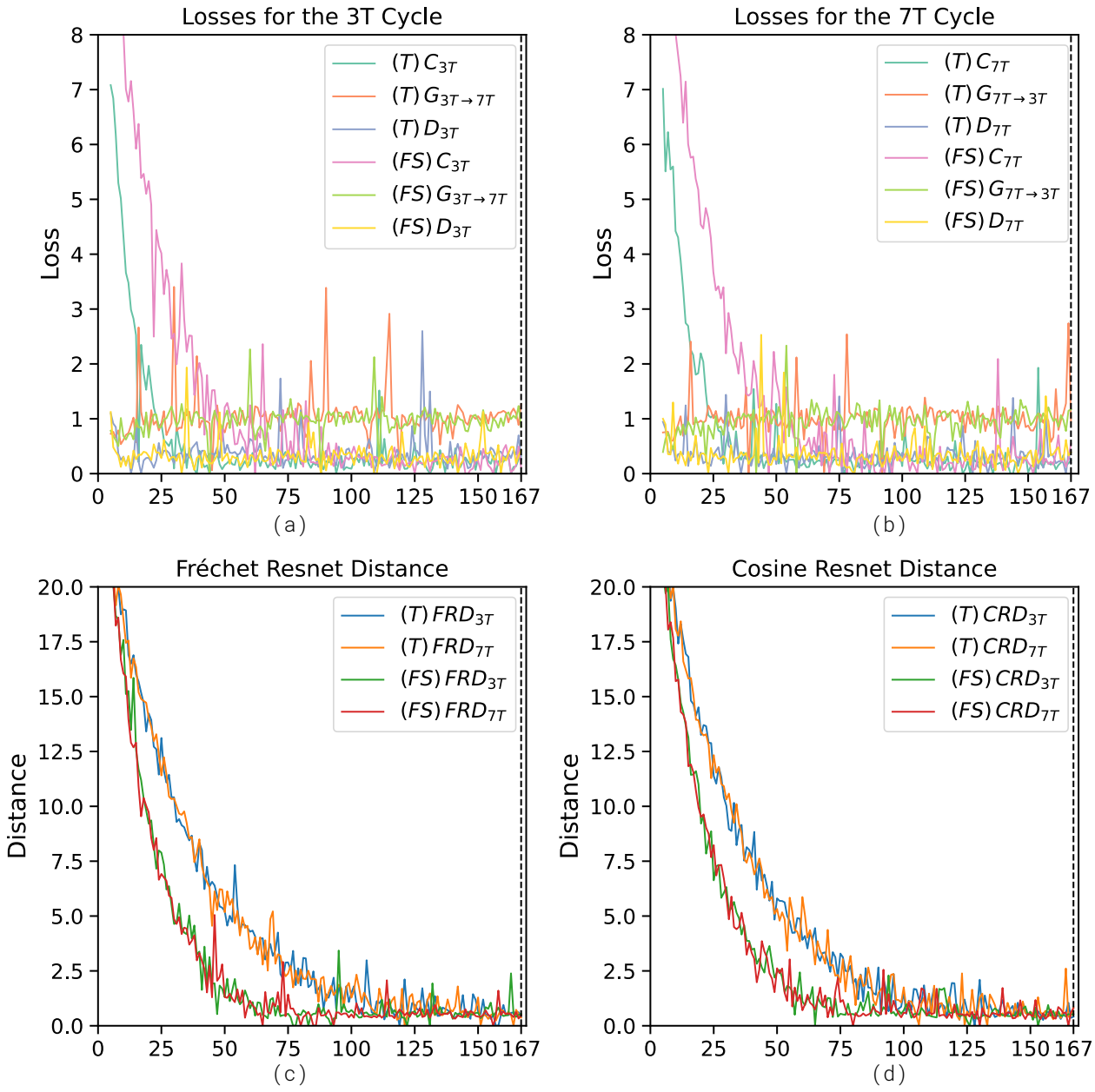


Figure 5.3: Progression of CycleGAN models training dynamics and quality metrics for 3T to 7T MRI translation. (a) Generator, Discriminator, and Cycle Consistency Losses for the 3T loop for the Traditional (T) and the Fast-Slow (FS) CycleGAN, (b) Generator, Discriminator, and Cycle Consistency Losses for the 7T loop (c) Fréchet ResNet Distance (FRD), (d) Cosine Resnet Distance (CRD). The vertical black dotted line mark epoch 167, where we halted training for both models.

convincing resolution and contrast properties that mimic the attributes of real 7T images, demonstrating the model's success in generating high-fidelity 7T-like images from 3T data. Moreover, the Fast-Slow model generated synthetic 7T with more accurate, sharper tissue boundaries and finer anatomical details. This indicates the dual pathway design enhanced the model's ability to preserve local textures and global tissue contrasts.

5.4.3 Tissue Type Specific Prediction

Diving deeper, we analyzed the models' performance in generating specific brain tissue types—the cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM)—by computing the Dice and the PAD. The results, visualized using box plots (Figure 5.5a), revealed that the Traditional (T) CycleGAN achieved a median Dice of 82.97% for CSF, 80.92% for GM, and 89.68% for WM. The Fast-Slow (FS) CycleGAN showed superior performance with a median Dice of 87.82% for CSF, 85.31% for GM, and 90.45% for WM.

Figure 5.5b has the PAD scores. The median PAD for the T model was 6.98% for CSF, 7.67% for GM, and 4.72% for WM. The FS model had similar performance with a median PAD of 5.64% for CSF, 2.28% for GM, and 4.72% for WM. These median scores demonstrate consistent performance across all three tissue types, with mean Dice significantly high and PAD reasonably low, corroborating both models' robustness in translating detailed tissue-specific characteristics from 3T to 7T.

We utilized robust statistical methods to objectively assess the differences in performance between the FS and T models. Normality testing is an essential preliminary step before performing parametric tests. The Shapiro-Wilk test is a widely accepted method for assessing the normality of data distributions (Shapiro and Wilk, 1965). Paired t-tests are commonly used to compare two related groups, assessing the differences in their means (Student, 1908). However, multiple comparisons increase the risk of false positives (Type I errors). The Holm-Bonferroni correction method effectively adjusts the p-values to control the familywise error rate (Holm, 1979). Effect sizes complement significance tests, measuring the magnitude of the differences. Cohen's d is a widely used effect size measure, where values of 0.2, 0.5, and 0.8 represent small, medium, and large effects, respectively (Cohen, 1992).

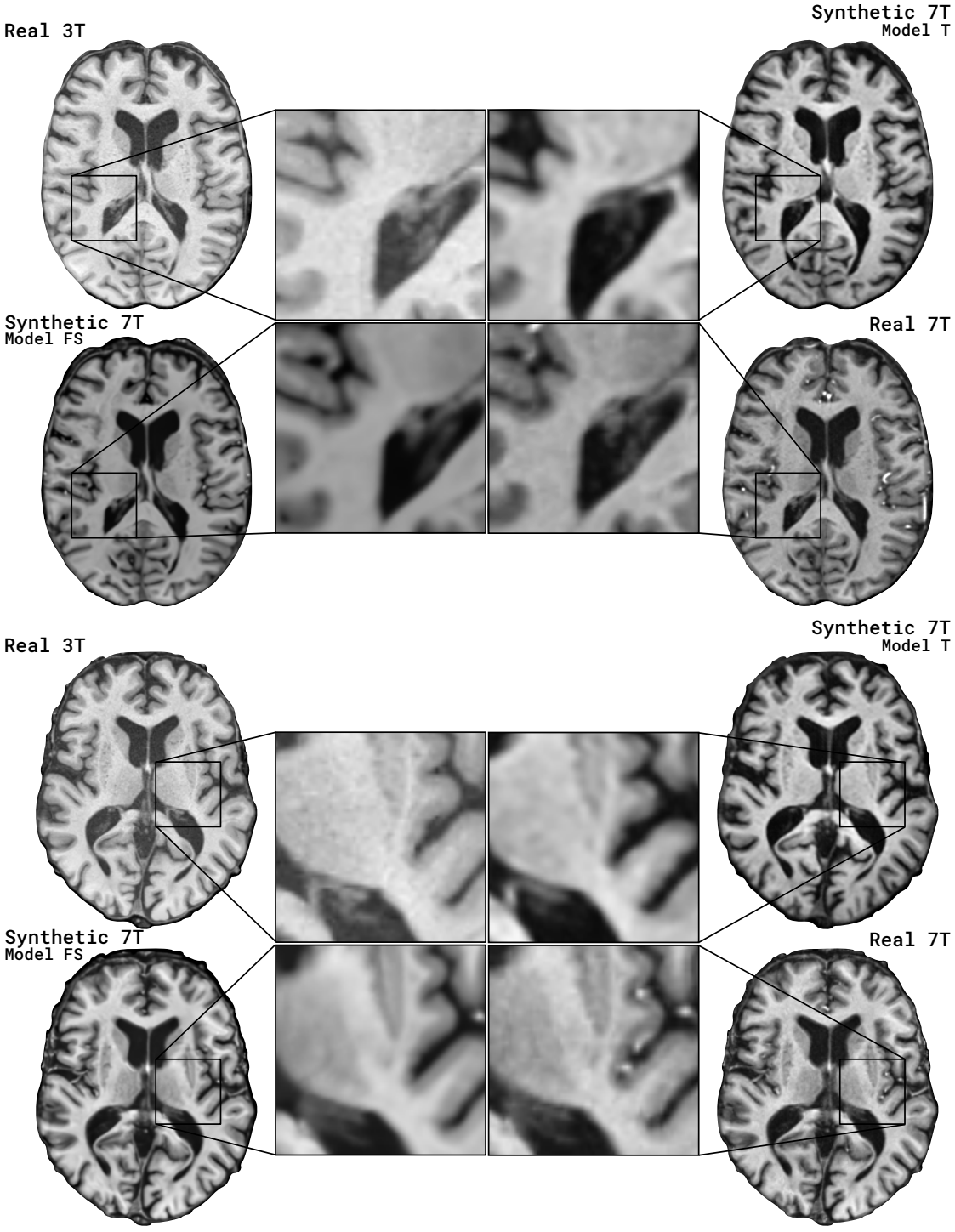


Figure 5.4: Examples of paired, real 3T and 7T MR images and the corresponding synthetic 7T images generated by the Traditional (T) and Fast-Slow (FS) CycleGAN model.

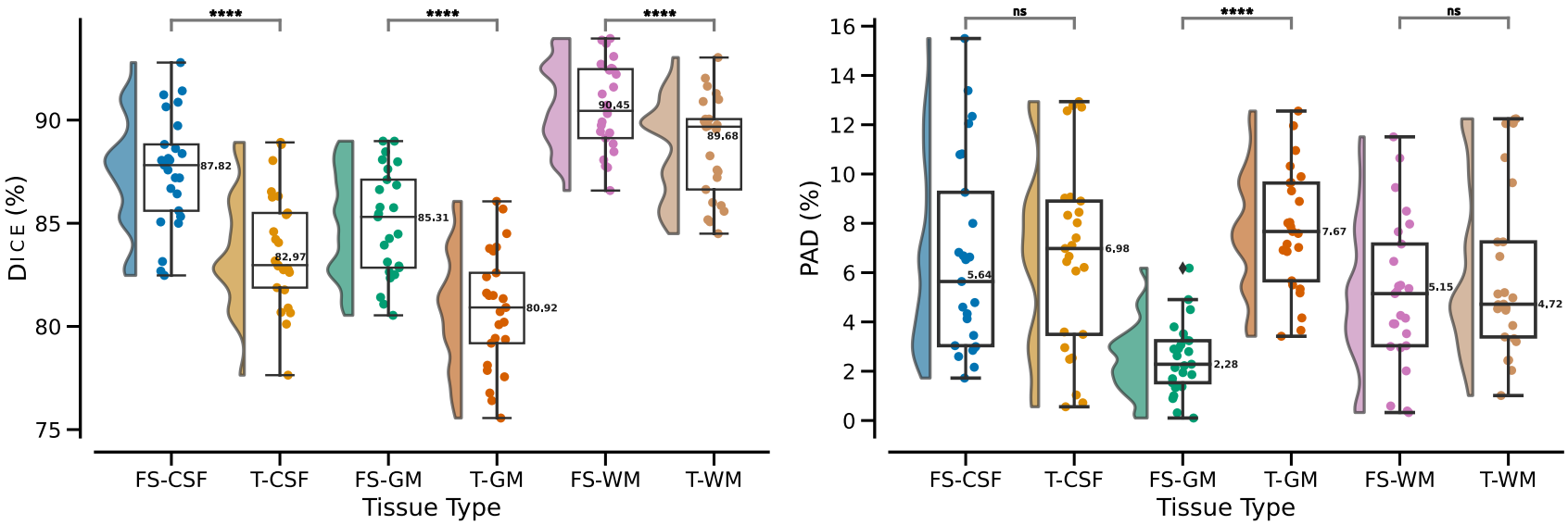


Figure 5.5: **Group-level segmentation performance for the validation (25 subjects) dataset across different tissue types—Cerebral Spinal Fluid (CSF), Gray Matter (GM), and White Matter (WM)—for the Traditional (T) and Fast-Slow (FS) CycleGAN models.**. We plot single-patient data overlaid on group-level whisker-box plots (center, median; box, 25th to 75th percentiles; whiskers, 1.5 interquartile range) and the smoothed data distribution. The median values are shown. The stars above the plots denote statistical significance, with $p \leq 0.0001$ represented by four asterisks (****). ns denotes no statistical significance.

The Shapiro-Wilk test results indicated that the data distributions for both the FS and T models were generally normal, as shown by the p-values greater than 0.05 for most tissue types. In the FS model, the p-values were 0.749 for Dice-CSF, 0.257 for Dice-GM, 0.414 for Dice-WM, 0.027 for PAD-CSF, 0.710 for PAD-GM, and 0.586 for PAD-WM. In the T model, the p-values were 0.718 for Dice-CSF, 0.899 for Dice-GM, 0.188 for Dice-WM, 0.152 for PAD-CSF, 0.870 for PAD-GM, and 0.007 for PAD-WM.

The paired t-tests revealed statistically significant differences between the FS and T models regarding the Dice metric across all tissue types, with the FS model outperforming the T model. The p-values for Dice were $\ll 0.0001$ for all tissue types (Dice-CSF: $p = 6.38 \times 10^{-13}$, Dice-GM: $p = 2.38 \times 10^{-11}$ Dice-WM: $p = 5.74 \times 10^{-8}$). However, there were no significant differences in the PAD metric for any tissue type except GM (PAD-CSF: $p = 0.932$, PAD-GM: $p = 5.74 \times 10^{-9}$, PAD-WM: $p = 0.389$).

Cohen's d effect sizes further illustrated the magnitude of the differences between the two models. Large effect sizes were observed for the Dice in the CSF (1.40) and GM (1.53), and a moderate effect size for the WM tissue type (0.84). Conversely, the effect sizes for the PAD metric were negligible for the CSF (-0.03) and WM (-0.27) tissue types and large for the GM tissue type (-2.60), indicating substantial differences between the two models.

In conclusion, the FS model significantly outperformed the T model regarding the Dice metric across all tissue types. Still, no significant differences existed in the PAD metric for any tissue type, except GM. The effect sizes indicated these differences were substantial, especially for the CSF and GM tissues.

5.5 Conclusions

This study sought to contribute with a novel approach to spatial adaptive MR data normalization between the 3T and 7T MR modalities, a persistent issue in medical imaging. Utilizing CycleGAN, an unsupervised generative adversarial network, our model demonstrated promising performance in generating clear 7T-like MR images from 3T inputs, evidenced by the high Dice and PAD (Taha and Hanbury, 2015, Zou et al., 2004). These scores affirm

both models' competence in maintaining the original morphological features of the brain and preserving the clinical relevance of the images, which has a significant implication in furthering neurological research (Moeskops et al., 2016, Ronneberger et al., 2015).

Both models' successful performance, especially in tasks requiring global contrast properties, supports the hypothesis that our approach enables a seamless transition to 7T MR systems without jeopardizing the quality of previously obtained 3T data (Ashburner and Friston, 2000, Johansen-Berg and Behrens, 2006). This holds promise in facilitating consistent and robust data analysis, mitigating potential bias and loss of statistical power associated with missing data in longitudinal studies (Schaer et al., 2008, Little and Rubin, 2019).

Previous research efforts have probed into creating 7T images using advanced deep-learning methodologies (Hou et al., 2016, Qin et al., 2019). Nevertheless, our model stands out due to its unsupervised learning aspect and capacity to function without resorting to frequency domain information (Klosowski and Frahm, 2017, Huang et al., 2019). The CycleGAN models' effectiveness in minimizing image variance underlines its potential to tackle the formidable challenge of cross-modality image translation between different MR field strengths (Adriany et al., 2008, Van der Velden et al., 2015).

Our model offers a promising approach to cross-modality MR image translation, allowing the utilization of 3T and 7T MRI technologies in longitudinal studies examining brain health (Uğurbil, 2014, Keuken and Forstmann, 2015). Our study constitutes a significant leap towards exploiting the capabilities of cutting-edge neuroimaging technologies without devaluing the rich trove of existing imaging data. This spatial adaptive normalization tactic helps bridge the technological gap, potentially accelerating the identification and validation of imaging biomarkers for neurological conditions (Bourgeat et al., 2015, Zhang et al., 2015).

The Fast-Slow U-Net aimed to enhance CycleGAN performance by separately processing local, high-frequency input features and global, low-frequency context. The motivation was handling multi-scale information would aid in complex 3T to 7T MRI translation, inspired by our conclusions from Section 4.3.5. Overall, the Fast-Slow U-Net CycleGAN experiment formed an important step in an ongoing endeavor to determine optimal network architectures for cross-modality 3T to 7T MRI translation using CycleGANs.

5.5.1 Limitations

Our research has made a significant step forward in advancing the application of CycleGAN to transform 3T to 7T MR data. The success of the translation task does, however, depend heavily on various factors, including the quality and nature of the input data, model architecture, and training procedure. Despite our progress, it is essential to consider these factors as potential limitations that may impact the results.

One of the primary limitations concerns the 2D nature of the implemented model. Adopting a 2D CycleGAN model was primarily due to computational considerations; however, the inherent lack of three-dimensional context by independently processing each slice results in slice-to-slice inconsistencies (Çiçek et al., 2016, Milletari et al., 2016). Additionally, patch-based image processing might result in losing local spatial information, which is critical in medical image segmentation (Tajbakhsh et al., 2020). These inconsistencies may impact the quality and fidelity of the synthesized 7T images. As Yushkevich et al. (2006) have demonstrated, 3D analysis can significantly enhance image interpretation and extract critical information from complex structures. Although implementing 3D CycleGAN architecture would alleviate this concern, the extensive computational demand often hinders its deployment.

The second potential pitfall is associated with an intrinsic property of CycleGANs, namely the non-bijective or many-to-many mapping between domains (Zhu et al., 2017). This can lead to struggles with controlling the mode of output (Ghosh et al., 2018) and non-functional transformations, which introduce potentially unrealistic information, or "hallucinations," into the synthetic images (Almahairi et al., 2018). These limitations and the difficulty of maintaining the semantic consistency of anatomical structures in synthesized images (Chartsias et al., 2018) hinder their diagnostic applicability (Wang et al., 2020). Consequently, future research should focus on devising strategies to rectify this challenge, such as regularizing the model with a perceptual loss (Johnson et al., 2016) or improving training procedures to facilitate a better generalization (Roth et al., 2020).

Variations in image quality across MR scanners and imaging protocols can also influence our model's performance (Jovicich et al., 2006, Kruggel et al., 2010). This issue is frequently encountered in multi-center studies and can generate images that do not accurately represent

real-world variance, reducing its practical use. Therefore, future research is warranted to assess the model's resilience against such variations and verify its performance on diverse datasets, including those featuring pathological changes (van Opbroek et al., 2015).

The question of data dependence and the potential for error in translation must also be mentioned. The translation's success relies heavily on the quality of the input data, meaning that errors during preprocessing, like inaccurate skull-stripping or normalization, could adversely impact the output (Iglesias and Sabuncu, 2015). Similarly, the translation process could inadvertently introduce or exaggerate noise or artifacts not present in authentic 7T images, leading to misleading interpretations (Chen et al., 2017).

The performance evaluation of segmentation algorithms plays a significant role in medical imaging. It determines the efficacy and reliability of these algorithms in clinical practice. We employed the Dice and PAD scores. Despite its widespread usage, the Dice score has a few limitations. One major drawback is that it needs to account for the geometric correspondence between the predicted and ground truth segments (Taha and Hanbury, 2015). It does not consider how well the predicted segment's shape, location, and orientation match the ground truth. In some scenarios, this information is critical, mainly when the size and position of the segment are clinically significant (*e.g.*, tumor segmentation) (Crum et al., 2006).

Conversely, the PAD evaluates the size estimation accuracy of segmentation results. While PAD effectively measures size discrepancies, it does not provide information on the spatial overlap between the predicted and the ground truth segmentation. Thus, two segmentations with the same area but in different locations would yield a zero PAD, failing to account for the spatial mismatch (Isensee et al., 2017). This limitation is critical in situations where the accurate localization of the object of interest is essential.

5.5.2 Future Directions

In conclusion, our research underscores the transformative potential of blending advanced machine learning techniques with clinical applications, particularly within neuroimaging (Obermeyer and Emanuel, 2016, Esteva et al., 2019). Our results provide a powerful testimony to the potential of CycleGAN in handling unpaired data, thus offering a compelling

solution to MR translation across different field strengths.

However, despite its merits, the CycleGAN models we investigated have limitations, and these challenges delineate promising avenues for future enhancement. Our study used generators based on the U-Net architecture (Ronneberger et al., 2015), which have delivered excellent results in medical image segmentation tasks. However, U-Net-based models primarily exploit local dependencies in the data, potentially neglecting long-range relationships between pixels. In contrast, with the advent of Vision Transformers (ViTs), there is potential to enhance our model's performance further (Dosovitskiy et al., 2020). ViTs can capture these intricate, long-range dependencies, which could be highly advantageous for tasks involving complex morphological structures such as brain MR images (Wang et al., 2018). This, in turn, could improve the segmentation performance, particularly for brain images that exhibit non-local relationships between anatomical structures (Petit et al., 2021).

Overall, the Fast-Slow U-Net experiment formed an important step in an ongoing endeavor to determine optimal network architectures for cross-modality 3T to 7T MRI translation. The results indicate that the Fast-Slow model significantly outperforms the Traditional model, especially for the CSF and GM tissue types. Further investigations are needed to understand the contributing factors to these differences and validate the models' performance in larger datasets and across diverse populations. Nevertheless, the Fast-Slow U-Net remains a promising avenue for future exploration.

6.0 Discussion

6.1 Summary of Key Findings and Contributions

6.1.1 Spike-Coding Dynamics and Neural Synchronization (Chapter 2)

Chapter 2 explored the mathematical foundations and dynamics of spike-coding networks. These networks encode information using precise spike times rather than average firing rates to encode information. The equations governing the membrane potential dynamics of spiking neurons were derived. It was shown how populations of leaky integrate-and-fire neurons can synchronize their spike times through a process known as phase-locking. Stability analysis revealed how key parameters like the synaptic time constant impact network synchronization. Varying these parameters revealed regimes exhibiting complex synchronization behaviors like metastability. The spike-coding network formulation was also generalized to handle encoding multiple input signals through linear combinations of spike trains. Overall, examining spike-coding network dynamics provided insights into how neurons communicate and self-organize to generate coordinated spiking activity and how synchronization depends on intrinsic neuronal and synaptic properties. Deriving the dynamical equations and exploring synchronization phenomena advances our theoretical understanding of the mechanisms enabling diverse temporal coding strategies employed by neural circuits in the brain.

6.1.2 Origins of Slow Oscillations in Cortical Networks (Chapter 3)

Chapter 3 investigated the neural mechanisms underlying the spontaneous transitions between active Up and silent Down states in cortical networks during slow-wave sleep. A model of weakly coupled Wilson-Cowan neural oscillators was analyzed, demonstrating bistability between quiescent and active states. Stochastic fluctuations and the coupling strength between neural populations were found to modulate transitions between these bistable states. Varying key parameters revealed that a Bogdanov-Takens bifurcation underlies the genesis of slow oscillations, providing insight into the complex coordination of excitation and inhibition

required to orchestrate the rhythmic Up-Down state transitions. This modeling elucidated how the interplay between noise, coupling topology, and nonlinear population dynamics gives rise to the spontaneous emergence of slow oscillations from cortical circuits. Revealing the bifurcation structure governing state transitions advances our mechanistic understanding of how sleep rhythms originate and are controlled.

6.1.3 Unified View of Associative Memory Models (Chapter 4)

Chapter 4 established a unifying framework connecting two foundational neural network models—Hopfield and spike-coding networks. Although differing in their emphasis on fixed point attractors versus spike timing, these classic models offer complementary perspectives on the same phenomenon—encoding associative memories and temporal sequences in recurrent networks. The key insight enabling this unification was recognizing the inherent linkage between the models; the stable attractor states in Hopfield networks provide a substrate for generating complex temporal spike sequences, which are dynamically replayed in spike-coding networks. In essence, Hopfield networks supply the associative memories, while spike-coding networks provide the sequencing between these memories. This analysis established a mathematical equivalence between the models using circulant matrices to represent asymmetric weights in Hopfield networks. This introduction of carefully structured asymmetry induced dynamics that enabled smooth transitions between stored attractor states. The circulant structure was further generalized to capture more complex temporal dependencies and correlations critical for tasks like the n-back working memory task. Overall, this work provided a unified lens for understanding the interlinked principles of associative memory and temporal sequence learning in neural networks. By synergistically integrating complementary computational elements from each model, it elucidated a shared algorithmic foundation spanning decades of research in mathematical neuroscience.

6.1.4 Cross-Modality MRI Translation with CycleGAN (Chapter 5)

Chapter 5 presented a novel deep learning approach for cross-modality translation of 3T MRI to 7T MRI using GANs. The model aimed to synthesize 7T-like images from

widely available 3T data, overcoming the challenges in transitioning longitudinal studies to more advanced 7T systems. The approach employed unpaired 3T and 7T MR datasets to train CycleGAN models for unsupervised 3T to 7T MRI translation. A specialized Fast-Slow U-Net variant was implemented alongside standard U-Net generators to enhance multi-scale image processing. Rigorous quantitative evaluation affirmed the model's efficacy in predicting key tissue types from independent test data. This research underscores GANs' potential for spatial adaptive normalization of MR data across field strengths. Generating synthetic 7T images from 3T inputs provides an efficient solution to bridge the technological divide and preserve compatibility with existing data. The results highlight deep learning's transformative capabilities in medical imaging and the need for innovative specialized architectures adapted to the problem structure. The work elucidates a promising cross-disciplinary approach blending neuroscience, medical imaging, and AI.

6.2 Synthesis and Significance of Research Directions

6.2.1 Elucidating Mechanisms of Neural Synchronization and Sequencing

Chapters 2 and 4 provided complementary modeling perspectives that elucidate neural networks' mechanisms to synchronize distributed neurons and sequence spatiotemporal patterns. Chapter 2 analyzed spike-coding networks and showed how populations of neurons can precisely coordinate their spike timing through phase-locking. Chapter 4 examined asymmetric Hopfield networks and revealed how structured synaptic weight perturbations induce cyclic attractor dynamics, allowing the storage and replay of temporal sequences. Despite their differing emphases, analysis of these foundational models exposed unifying principles governing how neural circuits orchestrate flexible spike-based coding. The synchronization phenomena in spike-coding networks and the sequential replay of memories in asymmetric Hopfield networks rely on shared mathematical principles and circuit mechanisms. These insights advance our theoretical grasp of the algorithms leveraged by biological neural computation to coordinate precise spike timing and represent temporal sequences and memories.

This knowledge helps establish guiding principles for designing more sophisticated neural network models capable of dynamic synchronization and sequencing.

6.2.2 Multi-Scale Modeling of Brain Rhythms and Memory Consolidation

A key theme connecting Chapters 3 and 4 is the multi-scale coordination underlying complex neural dynamics and computation. Chapter 3 demonstrated how slow oscillations originate from the delicate balance of excitation and inhibition within cortical neural populations. Mathematical analysis revealed this mesoscopic population-level rhythm genesis depends on a Bogdanov-Takens bifurcation, linking microscopic neuron properties to global brain states. Chapter 4 provided a multi-timescale circulant matrix framework to unify microscopic spike-coding networks and macroscopic Hopfield associative memories. This generalization using circulant matrices could capture neural interactions across diverse timescales, from fast synaptic currents to slow neuromodulation. Together, these examples showcase the importance of multi-scale modeling, from biophysical neurons to emergent rhythms and memories, to unravel cross-level mechanisms in the brain. They provide theoretical foundations to bridge cellular, population, and systems neuroscience approaches.

6.2.3 Bridging Biological and Artificial Neural Systems

Challenging applications like medical image analysis underscore key gaps between state-of-the-art AI systems and biological intelligence. As evidenced in Chapter 5, while Deep Learning models like CycleGAN can synthesize compelling 7T MR data from 3T inputs, significant improvement remains compared to human-level complex visual interpretation and understanding. Fundamentally, ANNs lack the levels of specialized domain knowledge and architectural adaptation that allow biological neural networks to handle complexity and generalize robustly. Implementing the Fast-Slow U-Net variant took inspiration from the spike-coding networks and asymmetric Hopfield networks analyzed in Chapters 2 and 4 to improve task performance. Those models demonstrated the utility of separate pathways for local and global information processing. However, much scope exists to integrate additional brain-inspired motifs inspired by those analyses, like hierarchical processing, top-down feedback,

and neuromodulation. Such bio-inspired enhancements tailored to particular tasks could significantly advance AI by instilling some of the flexibility and efficiency of biological neural networks. More broadly, this cross-disciplinary approach combining neuroscience, medical imaging, and AI offers a promising direction for developing more human-like intelligence—interpretable, data-efficient, and trustworthy systems that learn from nature.

6.3 Limitations and Future Research Directions

6.3.1 Extending Theoretical Models with Physiological Details

The conceptual models explored in this dissertation, including the spike-coding networks in Chapter 2 and the Wilson-Cowan neural oscillators in Chapter 3, provide valuable theoretical insights into the computational principles and dynamical mechanisms underlying neural processing. However, these simplified models necessarily omit many physiological details of real neurons. As Gerstner et al. (2014) suggested, incorporating additional biophysical realism could improve the models' accuracy and experimental predictive power.

For instance, prior work has shown the precise shape of action potentials can influence synchronization in spiking neural networks (Izhikevich, 2004). Neurons exhibit a diversity of spike shapes depending on their molecular composition and morphology (Bean, 2007). Adding spike shapes beyond simple delta functions could reveal new mechanisms by which postsynaptic waveform characteristics modulate coding and communication in local circuits.

Furthermore, axonal propagation delays can significantly impact synchronization phenomena (Vicente et al., 2008). Large-scale wiring diagrams have provided new data on the distribution of axonal conduction delays in cortical networks (Markram et al., 2015). Incorporating biologically constrained models of axonal delays could enhance the predictions of synchronization principles derived in Chapter 2.

Additionally, dendrites' active properties and complex geometry are crucial in neuronal integration and plasticity (Gidon and Segev, 2012). Recent advances in neuron morphology imaging could enable the integration of detailed dendritic morphologies into neural network

models (Gouwens et al., 2019). This could uncover new computational principles stemming from interactions between dendritic nonlinearities and network connectivity patterns.

In summary, while conceptual models provide important theoretical insights, adding physiological details like empirically-based spike shapes, axonal delays, and dendritic morphologies could lead to more biologically realistic models with enhanced experimental predictive power. This dissertation developed foundations to incorporate such details through multi-scale modeling spanning subcellular to systems levels.

6.3.2 Large-Scale Simulations and In Vivo Experiments

The models developed in this dissertation are limited in scale, typically comprising hundreds of model neurons. An important future direction is scaling up the models to simulate larger neural systems and capture emergent phenomena arising in whole-brain networks.

Recent advances in supercomputing infrastructure and simulation software enable modeling spiking activity across billions of neurons (Kunkel et al., 2014). Large-scale simulations of full cortical networks could examine how the computational principles identified in the simpler models manifest at larger scales, where interactions across areas and feedback loops come into play. Such large-scale models could better capture complex spatiotemporal patterns like those observed empirically in multi-electrode recordings (Timme et al., 2016).

Additionally, *in vivo* experiments are essential to validate the predictions from conceptual and simulation models (Kopell et al., 2011). Techniques like optogenetics, electrophysiology, and imaging enable precise perturbation and observation of neural activity in living systems (Grosenick et al., 2015). Testing theoretical predictions in animal models could refine and extend the modeling frameworks developed in this dissertation.

For instance, experimentally driving specific cell types predicted to control transitions between synchronous Up and Down states could causally probe their computational roles *in vivo* (Lewis et al., 2012). Comparing the statistics of neuronal recordings against simulation data can also help validate and constrain theoretical models (Gerstner and Naud, 2009).

An integrated approach combining multi-scale modeling with empirical measurements is necessary to align theory and experiments. The modeling approaches in this disserta-

tion provide foundations to bridge conceptual models, large-scale simulations, and biological experiments through mutually informative multi-scale exploration of neural systems.

6.3.3 Advancing AI through Bio-Inspiration

This dissertation developed conceptual models to advance our theoretical comprehension of biological neural computation. Looking forward, the insights gained could help inspire new directions to improve AI systems.

Fundamental principles in neuroscience like sparse coding, hierarchical processing, and neuromodulatory control could provide blueprints for enhancing Deep Learning architectures (Bengio et al., 2015, Kriegeskorte, 2015). For example, introducing winner-take-all dynamics or divisive normalization motivated by neuroscience models has improved network efficiency and generalization (Carandini and Heeger, 2012, Riesenhuber and Poggio, 1999). Neuromodulated learning, where different plasticity rules operate in distinct network layers, could enable more flexible training (Miconi et al., 2018).

Additionally, specialized architectures adapted to the structure of particular problems have shown promise in tasks like medical imaging and time series analysis (Bai et al., 2018). The Fast-Slow U-Net model developed in Chapter 5 demonstrates how dual-pathway designs separating local and global processing could benefit image analysis applications. Brain-inspired architectures incorporating recurrent processing for working memory could enhance sequential learning (Frank et al., 2001, O'Reilly and Frank, 2006).

Finally, developing more interpretable AI systems is critical for trust and transparency in medical and scientific applications (Samek et al., 2017). Incorporating neuroscience-based representations and objectives could yield more explainable models than end-to-end Deep Learning (Kietzmann et al., 2017). Integrating conceptual theories with practical machine learning could catalyze interpretable AI that learns from nature.

This research aims to complement data-driven AI by distilling essential computational principles from neuroscience. While not directly prescriptive, these theoretical insights may offer inspiration for developing the next generation of intelligent systems that integrate synergistically with human capabilities.

6.4 Conclusion

This dissertation advanced our understanding of the algorithms and mechanisms underlying neural computation in the brain through integrated modeling across multiple scales.

The research elucidated principles of spike timing coordination that enable distributed neurons to synchronize or sequence activity patterns flexibly. It revealed how global brain rhythms can spontaneously emerge from the nonlinear dynamics of interacting populations. A unifying mathematical framework that links conceptual models of associative memory and temporal processing in recurrent networks was introduced.

Together, these findings provide unified mechanistic explanations spanning neural coding, system-level rhythms, and memory consolidation. The models synergistically incorporate biophysical spiking neurons, mesoscopic population interactions, and abstract associative memories—bridging cellular to systems neuroscience. Analyses of network synchronization, nonlinear dynamics, bifurcations, and circulant weight matrices revealed common motifs underlying diverse computational capabilities.

This work elucidates a path toward interpretable AI by distilling essential algorithms from neuroscience. While not directly prescriptive, these conceptual insights may guide the development of brain-inspired computing. Separating fast and slow representations could enhance multi-scale processing in deep networks. Circulant connectivity could introduce robust sequential processing. Rich dynamics modeled through bifurcations may enable context-dependent computation. Integrating such motifs could yield more efficient, generalizable, and trustworthy AI systems that productively cooperate with human capabilities.

The convergence of neuroscience and AI is a promising cross-disciplinary approach to unraveling the complexities of natural and artificial systems. This research aims to accelerate progress at the intersection of these fields by elucidating unifying principles spanning levels of analysis, from biophysical mechanisms to collective computations. By revealing shared mathematical foundations, it contributes towards a unified understanding of biological and AI while highlighting enhancements to emerge from their synergistic integration.

Bibliography

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, and Matthieu Devin. Tensorflow: Large–scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016. arXiv:1603.04467.

Larry F Abbott and Thomas B Kepler. Model neurons: from hodgkin-huxley to hopfield. In *Statistical Mechanics of Neural Networks: Proceedings of the Xlth Sitges Conference Sitges, Barcelona, Spain, 3–7 June 1990*, pages 5–18. Springer, 2005.

Larry F Abbott, Brian DePasquale, and Raoul-Martin Memmesheimer. Building functional networks of spiking model neurons. *Nature neuroscience*, 19(3):350–355, 2016.

Ted Abel and K Matthew Lattal. Molecular mechanisms of memory acquisition, consolidation and retrieval. *Current opinion in neurobiology*, 11(2):180–187, 2001.

Gregor Adriany, Pierre-Francois Van de Moortele, Johannes Ritter, Steen Moeller, Edward J Auerbach, Can Akgün, Carl J Snyder, Thomas Vaughan, and Kâmil Uğurbil. A geometrically adjustable 16-channel transmit/receive transmission line array for improved rf efficiency and parallel imaging performance at 7 tesla. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 59(3): 590–597, 2008. ISSN 0740–3194.

Alexander Craig Aitken. *Determinants and matrices*. Read Books Ltd, 2017.

Torbjörn Åkerstedt and Peter M Nilsson. Sleep as restitution: an introduction. *Journal of internal medicine*, 254(1):6–12, 2003.

Amjad Almahairi, Sai Rajeshwar, Alessandro Sordoni, Philip Bachman, and Aaron Courville. Augmented cyclegan: Learning many–to–many mappings from unpaired data. In *International conference on machine learning*, pages 195–204. PMLR, 2018. ISBN 2640–3498.

David G Amaral and Menno P Witter. The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience*, 31(3):571–591, 1989.

Daniel J Amit. *Modeling brain function: The world of attractor neural networks*. Cambridge university press, 1989.

Daniel J Amit and Nicolas Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 7(3):237–252, 1997.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017. ISBN 2640–3498.

J. B. Arnold, J. S. Liow, K. A. Schaper, J. J. Stern, J. G. Sled, D. W. Shattuck, A. J. Worth, M. S. Cohen, R. M. Leahy, J. C. Mazziotta, and D. A. Rottenberg. Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *Neuroimage*, 13(5):931–43, 2001. ISSN 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1006/nimg.2001.0756. URL `https://www.ncbi.nlm.nih.gov/pubmed/11304088`.

J. Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007. ISSN 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2007.07.007. URL `https://www.ncbi.nlm.nih.gov/pubmed/17761438`.

J. Ashburner and K. J. Friston. Voxel–based morphometry–the methods. *Neuroimage*, 11 (6 Pt 1):805–21, 2000. ISSN 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1006/nimg.2000.0582. URL `https://www.ncbi.nlm.nih.gov/pubmed/10860804`.

Frederico AC Azevedo, Ludmila RB Carvalho, Lea T Grinberg, José Marcelo Farfel, Renata EL Ferretti, Renata EP Leite, Wilson Jacob Filho, Roberto Lent, and Suzana Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5): 532–541, 2009.

Alan Baddeley. Working memory: looking back and looking forward. *Nature reviews neuroscience*, 4(10):829–839, 2003.

Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

Cody Baker, Vicky Zhu, and Robert Rosenbaum. Nonlinear stimulus representations in neural circuits with approximate excitatory-inhibitory balance. *PLoS computational biology*, 16(9):e1008192, 2020.

David GT Barrett, Sophie Deneve, and Christian K Machens. Optimal compensation for neuron loss. *Elife*, 5:e12454, 2016.

Maxim Bazhenov, Igor Timofeev, Mircea Steriade, and Terrence J Sejnowski. Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *The Journal of Neuroscience*, 22(19):8691–8704, 2002.

Bruce P Bean. The action potential in mammalian central neurons. *Nature Reviews Neuroscience*, 8(6):451–465, 2007.

Randall D Beer. On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4):469–509, 1995.

Riccardo Beltramo, Giulia D'Urso, Marco Dal Maschio, Pasqualina Farisello, Serena Bovetti, Yoanne Clovis, Glenda Lassi, Valter Tucci, Davide De Pietri Tonelli, and Tommaso Fellin. Layer-specific excitatory circuits differentially control recurrent network dynamics in the neocortex. *Nature neuroscience*, 16(2):227–234, 2013.

Jan Benda and Andreas VM Herz. A universal model for spike-frequency adaptation. *Neural computation*, 15(11):2523–2564, 2003.

Yoshua Bengio. *Practical recommendations for gradient–based training of deep architectures*, pages 437–478. Springer, 2012.

Yoshua Bengio, Dong-Hyun Lee, Jorg Bornschein, Thomas Mesnard, and Zhouhan Lin. Towards biologically plausible deep learning. *arXiv preprint arXiv:1502.04156*, 2015.

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.

Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

Jiri Blank, Pavel Exner, and Miloslav Havlicek. *Hilbert space operators in quantum physics*. Springer Science & Business Media, 2008.

Martin Boerlin, Christian K Machens, and Sophie Denève. Predictive coding of dynamical variables in balanced spiking networks. *PLoS computational biology*, 9(11):e1003258, 2013.

Adam Bohr and Kaveh Memarzadeh. *Artificial intelligence in healthcare*. Academic Press, 2020. ISBN 0128184396.

Ingo Bojak, Thom F Oostendorp, Andrew T Reid, and Rolf Kötter. Connecting mean field models of neural activity to eeg and fmri data. *Brain topography*, 23(2):139–149, 2010.

Vincent Bonin, Valerio Mante, and Matteo Carandini. The suppressive field of neurons in lateral geniculate nucleus. *Journal of Neuroscience*, 25(47):10844–10856, 2005.

Roman M Borisyuk and Alexandr B Kirillov. Bifurcation analysis of a neural network model. *Biological Cybernetics*, 66(4):319–325, 1992.

Alexander Borst and Frédéric E Theunissen. Information theory and neural coding. *Nature neuroscience*, 2(11):947–957, 1999.

P. Bourgeat, V. L. Villemagne, V. Dore, B. Brown, S. L. Macaulay, R. Martins, C. L. Masters, D. Ames, K. Ellis, C. C. Rowe, O. Salvado, J. Fripp, and Aibl Research Group. Comparison of mr–less pib suvr quantification methods. *Neurobiol Aging*, 36 Suppl 1:S159–66, 2015. ISSN 1558–1497 (Electronic) 0197–4580 (Linking). doi: 10.1016/j.neurobiolaging.2014.04.033. URL `https://www.ncbi.nlm.nih.gov/pubmed/25257985`.

Ronald Newbold Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 1986.

Michael Breakspear. Dynamic models of large-scale brain activity. *Nature neuroscience*, 20 (3):340–352, 2017.

Steven L Bressler and Vinod Menon. Large-scale brain networks in cognition: emerging methods and principles. *Trends in cognitive sciences*, 14(6):277–290, 2010.

Romain Brette and Wulfram Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of neurophysiology*, 94(5):3637–3642, 2005.

Richard A Brualdi, Herbert John Ryser, et al. *Combinatorial matrix theory*, volume 39. Springer, 1991.

Jehoshua Bruck. On the convergence properties of the hopfield model. *Proceedings of the IEEE*, 78(10):1579–1585, 1990.

Bingni W Brunton, Matthew M Botvinick, and Carlos D Brody. Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340(6128):95–98, 2013.

György Buzsáki. The hippocampo-neocortical dialogue. *Cerebral cortex*, 6(2):81–92, 1996.

György Buzsáki. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68(3): 362–385, 2010.

Gyorgy Buzsaki and Andreas Draguhn. Neuronal oscillations in cortical networks. *science*, 304(5679):1926–1929, 2004.

György Buzsáki, Costas A Anastassiou, and Christof Koch. The origin of extracellular fields and currents–eeg, ecog, lfp and spikes. *Nature reviews neuroscience*, 13(6):407–420, 2012.

J. Cai, Z. Zhang, L. Cui, Y. Zheng, and L. Yang. Towards cross–modal organ translation and segmentation: A cycle– and shape–consistent generative adversarial network. *Med Image Anal*, 52:174–184, 2019. ISSN 1361–8423 (Electronic) 1361–8415 (Linking). doi: 10.1016/j.media.2018.12.002. URL `https://www.ncbi.nlm.nih.gov/pubmed/30594770`.

Alfredo Canziani, Adam Paszke, and Eugenio Culurciello. An analysis of deep neural network models for practical applications. *arXiv preprint arXiv:1605.07678*, 2016.

Xiaohuan Cao, Jianhua Yang, Jun Zhang, Dong Nie, Minjeong Kim, Qian Wang, and Dinggang Shen. Deformable image registration based on similarity–steered cnn regression. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11–13, 2017, Proceedings, Part I 20*, pages 300–308. Springer, 2017. ISBN 3319661817.

Matteo Carandini and David J Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012.

Valerie A Carr, Indre V Viskontas, Stephen A Engel, and Barbara J Knowlton. Neural activity in the hippocampus and perirhinal cortex during encoding is associated with the durability of episodic memory. *Journal of Cognitive Neuroscience*, 22(11):2652–2662, 2010.

Mary A. Carskadon and William C. Dement. Normal human sleep: An overview. In Meir H. Kryger, T. Roth, and William C. Dement, editors, *Principles and Practice of Sleep Medicine*, chapter 2, pages 15–24. Elsevier, 6th edition, 2017.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u–net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer–Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19*, pages 424–432. Springer, 2016. ISBN 3319467220.

A. Chartsias, T. Joyce, M. V. Giuffrida, and S. A. Tsaftaris. Multimodal mr synthesis via modality–invariant latent representation. *IEEE Trans Med Imaging*, 37(3):803–814, 2018. ISSN 1558–254X (Electronic) 0278–0062 (Print) 0278–0062 (Linking). doi: 10.1109/TMI.2017.2764326. URL `https://www.ncbi.nlm.nih.gov/pubmed/29053447`.

Sylvain Chauvette, Maxim Volgushev, and Igor Timofeev. Origin of active states in local neocortical networks during slow sleep oscillation. *Cerebral cortex*, 20(11):2660–2674, 2010.

Chi-Tsong Chen. *Linear system theory and design*. Saunders college publishing, 1984.

Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, pages 1511–1520, 2017.

Tianping Chen and Shun Ichi Amari. Stability of asymmetric hopfield networks. *IEEE Transactions on Neural Networks*, 12(1):159–163, 2001.

Yu Cheng, Felix X Yu, Rogerio S Feris, Sanjiv Kumar, Alok Choudhary, and Shi-Fu Chang. An exploration of parameter redundancy in deep networks with circulant projections. In *Proceedings of the IEEE international conference on computer vision*, pages 2857–2865, 2015.

J Cohen. A power primer. *Psychological bulletin*, 112(1):155–159, 1992.

Charles Conley. The gradient structure of a flow: I. *Ergodic Theory Dynam. Systems*, 8 (Charles Conley Memorial Issue), 1988.

Barry W Connors and Michael J Gutnick. Intrinsic firing patterns of diverse neocortical neurons. *Trends in neurosciences*, 13(3):99–104, 1990.

Andrew RA Conway, Michael J Kane, Michael F Bunting, D Zach Hambrick, Oliver Wilhelm, and Randall W Engle. Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, 12:769–786, 2005.

Rosa Cossart, Rafael Yuste, and Dmitriy Aronov. Attractor dynamics of network up states in the neocortex. *Nature*, 423(6937):283–288, 2003.

W. R. Crum, O. Camara, and D. L. Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Trans Med Imaging*, 25(11):1451–61, 2006. ISSN 0278–0062 (Print) 0278–0062 (Linking). doi: 10.1109/TMI.2006.880587. URL `https://www.ncbi.nlm.nih.gov/pubmed/17117774`.

Vincenzo Crunelli and Stuart W. Hughes. The slow (¡1 hz) rhythm of non-rem sleep: a dialogue between three cardinal oscillators. *Nature Neuroscience*, 13(1):9–17, 2010.

Rodica Curtu and Bard Ermentrout. Oscillations in a refractory neural net. *Journal of mathematical biology*, 43(1):81–100, 2001.

Henry Dale. Pharmacology and nerve-endings. *Journal of the Royal Society of Medicine*, 28 (3):319–332, 1935.

François David, Joscha T. Schmiedt, Hannah L. Taylor, Gergely Orban, Giuseppe Di Giovanni, Victor N. Uebele, John J. Renger, Régis C. Lambert, Nathalie Leresche, and Vincenzo Crunelli. Essential thalamic contribution to slow waves of natural sleep. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 33(50):19599–19610, 2013.

Philip J Davis. *Circulant matrices*, volume 2. Wiley New York, 1979.

Peter Dayan and Laurence F Abbott. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT press, 2005.

Peter Dayan and Angela J Yu. Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, 17(4):335–350, 2006.

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

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 Comput Biol*, 4(8):e1000092, 2008.

Javier DeFelipe and Isabel Fariñas. The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Progress in neurobiology*, 39 (6):563–607, 1992.

Mete Demircigil, Judith Heusel, Matthias Löwe, Sven Upgang, and Franck Vermet. On a model of associative memory with huge storage capacity. *Journal of Statistical Physics*, 168(2):288–299, 2017.

A. Destexhe and T. J. Sejnowski. Interactions between membrane conductances underlying thalamocortical slow-wave oscillations. *Physiological Reviews*, 83(4):1401–1453, 2003.

Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3): 297–302, 1945. ISSN 0012–9658.

Susanne Diekelmann and Jan Born. The memory function of sleep. *Nature Reviews Neuroscience*, 11(2):114–126, 2010.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

DC Dowson and BV666017 Landau. The fréchet distance between multivariate normal distributions. *Journal of multivariate analysis*, 12(3):450–455, 1982. ISSN 0047–259X.

Kenji Doya et al. Bifurcations in the learning of recurrent neural networks 3. *learning (RTRL)*, 3:17, 1992.

Khanh Dao Duc, Pierre Parutto, Xiaowei Chen, Jérôme Epsztein, Arthur Konnerth, and David Holcman. Synaptic dynamics and neuronal network connectivity are reflected in the distribution of times in up states. *Frontiers in computational neuroscience*, 9, 2015.

John Eccles. From electrical to chemical transmission in the central nervous system. *Notes and records of the Royal Society of London*, 30(2):219–230, 1976.

Chris Eliasmith and Charles H Anderson. *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT press, 2003.

Andreas K Engel and Wolf Singer. Temporal binding and the neural correlates of sensory awareness. *Trends in cognitive sciences*, 5(1):16–25, 2001.

B. Ermentrout. Phase-plane analysis of neural activity. In Michael A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*, pages 732–738. MIT Press, 1998a.

B. Ermentrout. Neural networks as spatio-temporal pattern-forming systems. *Reports on progress in physics*, 61(4):353, 1998b.

B. Ermentrout. *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*, volume 14. Siam, 2002.

Bard Ermentrout. Type i membranes, phase resetting curves, and synchrony. *Neural computation*, 8(5):979–1001, 1996.

Bard Ermentrout and David H Terman. *Mathematical foundations of neuroscience*, volume 35. Springer, 2010.

G Bard Ermentrout and Jack D Cowan. A mathematical theory of visual hallucination patterns. *Biological cybernetics*, 34(3):137–150, 1979.

A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean. A guide to deep learning in healthcare. *Nat Med*, 25(1):24–29, 2019. ISSN 1546–170X (Electronic) 1078–8956 (Linking). doi: 10.1038/s41591--018--0316--z. URL `https://www.ncbi.nlm.nih.gov/pubmed/30617335`.

A. C. Evans, A. L. Janke, D. L. Collins, and S. Baillet. Brain templates and atlases. *Neuroimage*, 62(2):911–22, 2012. ISSN 1095–9572 (Electronic) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2012.01.024. URL `https://www.ncbi.nlm.nih.gov/pubmed/22248580`.

A Aldo Faisal, Luc PJ Selen, and Daniel M Wolpert. Noise in the nervous system. *Nature reviews neuroscience*, 9(4):292–303, 2008.

Christoph Feichtenhofer, Axel Pinz, and Andrew Zisserman. Convolutional two-stream network fusion for video action recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1933–1941, 2016.

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.

Michele Ferrante, Michele Migliore, and Giorgio A Ascoli. Functional impact of dendritic branch-point morphology. *Journal of Neuroscience*, 33(5):2156–2165, 2013.

Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445, 1961.

Jerry A Fodor and Zenon W Pylyshyn. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2):3–71, 1988.

J. P. Fortin, N. Cullen, Y. I. Sheline, W. D. Taylor, I. Aselcioglu, P. A. Cook, P. Adams, C. Cooper, M. Fava, P. J. McGrath, M. McInnis, M. L. Phillips, M. H. Trivedi, M. M. Weissman, and R. T. Shinohara. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*, 167:104–120, 2018. ISSN 1095–9572 (Electronic) 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2017.11.024. URL `https://www.ncbi.nlm.nih.gov/pubmed/29155184`.

Michael J Frank, Bryan Loughry, and Randall C O'Reilly. Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience*, 1:137–160, 2001.

Pascal Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480, 2005.

Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138, 2010.

Karl Friston, James Kilner, and Lee Harrison. A free energy principle for the brain. *Journal of physiology-Paris*, 100(1-3):70–87, 2006.

Nina E Fultz, Giorgio Bonmassar, Kawin Setsompop, Robert A Stickgold, Bruce R Rosen, Jonathan R Polimeni, and Laura D Lewis. Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. *Science*, 366(6465):628–631, 2019.

Stefano Fusi, Patrick J Drew, and Larry F Abbott. Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611, 2005.

Joaquin M Fuster and Garrett E Alexander. Neuron activity related to short-term memory. *Science*, 173(3997):652–654, 1971.

Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 3. springer Berlin, 1985.

Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2414–2423, 2016.

Wilson S Geisler and Duane G Albrecht. Cortical neurons: isolation of contrast gain control. *Vision research*, 32(8):1409–1410, 1992.

Wulfram Gerstner and Richard Naud. How good are neuron models? *Science*, 326(5951): 379–380, 2009.

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.

Arnab Ghosh, Viveka Kulharia, Vinay P Namboodiri, Philip HS Torr, and Puneet K Dokania. Multi–agent diverse generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8513–8521, 2018.

Albert Gidon and Idan Segev. Principles governing the operation of synaptic inhibition in dendrites. *Neuron*, 75(2):330–341, 2012.

Charles D Gilbert and Wu Li. Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14(5):350–363, 2013.

Gabrielle Girardeau, Karim Benchenane, Sidney I Wiener, György Buzsáki, and Michaël B Zugaro. Selective suppression of hippocampal ripples impairs spatial memory. *Nature neuroscience*, 12(10):1222–1223, 2009.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial*

intelligence and statistics, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.

Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings, 2011.

Eric Goles-Chacc, Françoise Fogelman-Soulié, and Didier Pellegrin. Decreasing energy functions as a tool for studying threshold networks. *Discrete Applied Mathematics*, 12(3): 261–277, 1985.

Rafael C Gonzales and Paul Wintz. *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc., 1987.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016. ISBN 0262337371.

Nathan W Gouwens, Staci A Sorensen, Jim Berg, Changkyu Lee, Tim Jarsky, Jonathan Ting, Susan M Sunkin, David Feng, Costas A Anastassiou, Eliza Barkan, et al. Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nature neuroscience*, 22(7):1182–1195, 2019.

Robert M Gray et al. Toeplitz and circulant matrices: A review. *Foundations and Trends® in Communications and Information Theory*, 2(3):155–239, 2006.

Logan Grosenick, James H Marshel, and Karl Deisseroth. Closed-loop and activity-guided optogenetic control. *Neuron*, 86(1):106–139, 2015.

R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, R. Wolz, M. C. Valdes-Hernandez, D. A. Dickie, J. Wardlaw, and D. Rueckert. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *Neuroimage Clin*, 17:918–934, 2018. ISSN 2213–1582 (Electronic) 2213–1582 (Linking). doi: 10.1016/j.nicl.2017.12.022. URL `https://www.ncbi.nlm.nih.gov/pubmed/29527496`.

Robert Gütig. To spike, or when to spike? *Current opinion in neurobiology*, 25:134–139, 2014.

P. Hagmann, L. Jonasson, P. Maeder, J. P. Thiran, V. J. Wedeen, and R. Meuli. Understanding diffusion mr imaging techniques: from scalar diffusion–weighted imaging to diffusion tensor imaging and beyond. *Radiographics*, 26(Suppl 1):S205–23, 2006. ISSN 1527–1323 (Electronic) 0271–5333 (Linking). doi: 10.1148/rg.26si065510. URL `https://www.ncbi.nlm.nih.gov/pubmed/17050517`.

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. *The Journal of neuroscience*, 26(17):4535–4545, 2006.

David Hansel, Germán Mato, and Claude Meunier. Synchrony in excitatory neural networks. *Neural computation*, 7(2):307–337, 1995.

Kenneth D Harris, Jozsef Csicsvari, Hajime Hirase, George Dragoi, and György Buzsáki. Organization of cell assemblies in the hippocampus. *Nature*, 424(6948):552–556, 2003.

Gal Haspel, Kristen E Severi, Lisa J Fauci, Netta Cohen, Eric D Tytell, and Jennifer R Morgan. Resilience of neural networks for locomotion. *The Journal of Physiology*, 599(16): 3825–3840, 2021.

Demis Hassabis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick. Neuroscience-inspired artificial intelligence. *Neuron*, 95(2):245–258, 2017.

Michael E Hasselmo and Martin Sarter. Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology*, 36(1):52–73, 2011.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

Donald O Hebb. The first stage of perception: growth of the assembly. *The Organization of Behavior*, 4(60):78–60, 1949.

David J Heeger. Normalization of cell responses in cat striate cortex. *Visual neuroscience*, 9(2):181–197, 1992.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time–scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

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.

David Holcman and Misha Tsodyks. The emergence of up and down states in cortical networks. *PLoS Comput Biol*, 2(3):e23, 2006.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pages 65–70, 1979.

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79:2554–2558, 1982.

J. J. Hopfield. Collective processing and neural states. *Modeling and analysis in biomedicine. Singapore: World Scientific*, 1984.

F. Hoppensteadt. Predator-prey model. *Scholarpedia*, 1(10):1563, 2006.

Frank C. Hoppensteadt and Eugene M. Izhikevich. *Weakly connected neural networks*, volume 126. Springer, 1997.

Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

Le Hou, Dimitris Samaras, Tahsin M Kurc, Yi Gao, James E Davis, and Joel H Saltz. Patch–based convolutional neural network for whole slide tissue image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2424–2433, 2016.

Marc W Howard and Michael J Kahana. A distributed representation of temporal context. *Journal of mathematical psychology*, 46(3):269–299, 2002.

X. Huang, W. Sun, T. B. Tseng, C. Li, and W. Qian. Fast and fully–automated detection and segmentation of pulmonary nodules in thoracic ct scans using deep convolutional neural networks. *Comput Med Imaging Graph*, 74:25–36, 2019. ISSN 1879–0771 (Electronic) 0895–6111 (Linking). doi: 10.1016/j.compmedimag.2019.02.003. URL `https://www.ncbi.nlm.nih.gov/pubmed/30954678`.

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.

Raoul Huys, Dionysios Perdikis, and Viktor K Jirsa. Functional architectures and structured flows on manifolds: a dynamical framework for motor behavior. *Psychological review*, 121 (3):302, 2014.

J. E. Iglesias and M. R. Sabuncu. Multi–atlas segmentation of biomedical images: A survey. *Med Image Anal*, 24(1):205–219, 2015. ISSN 1361–8423 (Electronic) 1361–8415 (Print) 1361–8415 (Linking). doi: 10.1016/j.media.2015.06.012. URL `https://www.ncbi.nlm.nih.gov/pubmed/26201875`.

Jeffrey J Iliff, Minghuan Wang, Yonghong Liao, Benjamin A Plogg, Weiguo Peng, Georg A Gundersen, Helene Benveniste, G Edward Vates, Rashid Deane, Steven A Goldman, et al. A paravascular pathway facilitates csf flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid β. *Science translational medicine*, 4(147): 147ra111–147ra111, 2012.

Jeffry S Isaacson and Massimo Scanziani. How inhibition shapes cortical activity. *Neuron*, 72(2):231–243, 2011.

Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H Maier-Hein. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 287–297. Springer, 2017. ISBN 3319752375. Revised Selected Papers 3.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image–to–image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

Eugene M Izhikevich. Which model to use for cortical spiking neurons? *IEEE transactions on neural networks*, 15(5):1063–1070, 2004.

Eugene M Izhikevich. *Dynamical systems in neuroscience*. MIT press, 2007.

Eugene M Izhikevich and Gerald M Edelman. Large-scale model of mammalian thalamocortical systems. *Proceedings of the national academy of sciences*, 105(9):3593–3598, 2008.

Susanne M Jaeggi, Martin Buschkuehl, Walter J Perrig, and Beat Meier. The concurrent validity of the n-back task as a working memory measure. *Memory*, 18(4):394–412, 2010a.

Susanne M Jaeggi, Barbara Studer-Luethi, Martin Buschkuehl, Yi-Fen Su, John Jonides, and Walter J Perrig. The relationship between n-back performance and matrix reasoning–implications for training and transfer. *Intelligence*, 38(6):625–635, 2010b.

Daoyun Ji and Matthew A Wilson. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature neuroscience*, 10(1):100–107, 2007.

Heidi Johansen-Berg and Timothy EJ Behrens. Just pretty pictures? what diffusion tractography can add in clinical neuroscience. *Current opinion in neurology*, 19(4):379, 2006.

Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real–time style transfer and super–resolution. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part II 14*, pages 694–711. Springer, 2016. ISBN 3319464744.

J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. Macfall, B. Fischl, and A. Dale. Reliability in multi–site structural mri studies: effects of gradient non–linearity correction on phantom and human data. *Neuroimage*, 30(2):436–43, 2006. ISSN 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2005.09.046. URL `https://www.ncbi.nlm.nih.gov/pubmed/16300968`.

M. C. Keuken and B. U. Forstmann. A probabilistic atlas of the basal ganglia using 7 t mri. *Data Brief*, 4:577–82, 2015. ISSN 2352–3409 (Print) 2352–3409 (Electronic) 2352–3409 (Linking). doi: 10.1016/j.dib.2015.07.028. URL `https://www.ncbi.nlm.nih.gov/pubmed/26322322`.

Stefan J Kiebel, Jean Daunizeau, and Karl J Friston. A hierarchy of time-scales and the brain. *PLoS computational biology*, 4(11):e1000209, 2008.

Tim C Kietzmann, Patrick McClure, and Nikolaus Kriegeskorte. Deep neural networks in computational neuroscience. *BioRxiv*, page 133504, 2017.

Zachary P Kilpatrick and Paul C Bressloff. Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. *Physica D: Nonlinear Phenomena*, 239(9):547–560, 2010.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint, 2014. arXiv:1412.6980.

Wayne K Kirchner. Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, 55(4):352, 1958.

Charles Kittel. *Elementary statistical physics*. Courier Corporation, 2004.

D Kleinfeld. Sequential state generation by model neural networks. *Proceedings of the National Academy of Sciences*, 83(24):9469–9473, 1986.

J. Klosowski and J. Frahm. Image denoising for real–time mri. *Magn Reson Med*, 77(3):1340–1352, 2017. ISSN 1522–2594 (Electronic) 0740–3194 (Linking). doi: 10.1002/mrm.26205. URL `https://www.ncbi.nlm.nih.gov/pubmed/27079944`.

Christof Koch. *Biophysics of computation: information processing in single neurons*. Oxford university press, 2004.

Nanacy Kopell, Miles A Whittington, and Mark A Kramer. Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proceedings of the National Academy of Sciences*, 108(9):3779–3784, 2011.

Nancy Kopell and G Bard Ermentrout. Mechanisms of phase-locking and frequency control in pairs of coupled neural oscillators. *Handbook of dynamical systems*, 2:3–54, 2002.

Nikolaus Kriegeskorte. Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annual review of vision science*, 1:417–446, 2015.

N. Krishnamurthy, T. Santini, S. Wood, J. Kim, T. Zhao, H. J. Aizenstein, and T. S. Ibrahim. Computational and experimental evaluation of the tic–tac–toe rf coil for 7 tesla mri. *PLoS One*, 14(1):e0209663, 2019. ISSN 1932–6203 (Electronic) 1932–6203 (Linking). doi: 10.1371/journal.pone.0209663. URL `https://www.ncbi.nlm.nih.gov/pubmed/30629618`.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

Dmitry Krotov. Hierarchical associative memory. *arXiv preprint arXiv:2107.06446*, 2021.

Dmitry Krotov and John Hopfield. Large associative memory problem in neurobiology and machine learning. *arXiv preprint arXiv:2008.06996*, 2020.

Dmitry Krotov and John J Hopfield. Dense associative memory for pattern recognition. *Advances in neural information processing systems*, 29, 2016.

F. Kruggel, J. Turner, L. T. Muftuler, and Initiative Alzheimer's Disease Neuroimaging. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the adni cohort. *Neuroimage*, 49(3):2123–33, 2010. ISSN 1095–9572 (Electronic) 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2009.11.006. URL `https://www.ncbi.nlm.nih.gov/pubmed/19913626`.

Susanne Kunkel, Maximilian Schmidt, Jochen M Eppler, Hans E Plesser, Gen Masumoto, Jun Igarashi, Shin Ishii, Tomoki Fukai, Abigail Morrison, Markus Diesmann, et al. Spiking network simulation code for petascale computers. *Frontiers in neuroinformatics*, 8:78, 2014.

Yoshiki Kuramoto. *Chemical turbulence*. Springer, 1984.

Salomé Kurth, Maya Ringli, Anja Geiger, Monique LeBourgeois, Oskar G Jenni, and Reto Huber. Mapping of cortical activity in the first two decades of life: a high-density sleep electroencephalogram study. *The Journal of Neuroscience*, 30(40):13211–13219, 2010.

Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and brain sciences*, 40:e253, 2017.

Victor AF Lamme and Pieter R Roelfsema. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, 23(11):571–579, 2000.

LM Lapicque. Recherches quantitatives sur l'excitation electrique des nerfs. *J Physiol Paris*, 9:620–635, 1907.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Yann LeCun, Léon Bottou, Genevieve B Orr, and Klaus-Robert Müller. *Efficient backprop*, pages 9–50. Springer, 2002.

Tai Sing Lee and David Mumford. Hierarchical bayesian inference in the visual cortex. *Josa A*, 20(7):1434–1448, 2003.

Maxime Lemieux, Jen-Yung Chen, Peter Lonjers, Maxim Bazhenov, and Igor Timofeev. The impact of cortical deafferentation on the neocortical slow oscillation. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 34(16):5689–5703, 2014.

Maxime Lemieux, Sylvain Chauvette, and Igor Timofeev. Neocortical inhibitory activities and long-range afferents contribute to the synchronous onset of silent states of the neocortical slow oscillation. *Journal of neurophysiology*, 113(3):768–779, 2015.

David A Lewis, Allison A Curley, Jill R Glausier, and David W Volk. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends in neurosciences*, 35(1): 57–67, 2012.

Zachary C Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.

G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, Jawm van der Laak, B. van Ginneken, and C. I. Sanchez. A survey on deep learning in medical image analysis. *Med Image Anal*, 42:60–88, 2017. ISSN 1361–8423 (Electronic) 1361–8415 (Linking). doi: 10.1016/j.media.2017.07.005. URL `https://www.ncbi.nlm.nih.gov/pubmed/28778026`.

Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*, volume 793. John Wiley & Sons, 2019. ISBN 0470526793.

Ashok Litwin-Kumar and Brent Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature neuroscience*, 15(11):1498–1505, 2012.

Ying-Hui Liu and Xiao-Jing Wang. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *Journal of computational neuroscience*, 10:25–45, 2001.

RTMAC Lu. *Algorithms for discrete Fourier transform and convolution*. Springer, 1989.

Artur Luczak, Peter Barthó, Stephan L. Marguet, Gy´orgy Buzsáki, and Kenneth D. Harris. Sequential structure of neocortical spontaneous activity in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1):347–352, 2007.

Mikael Lundqvist, Albert Compte, and Anders Lansner. Bistable, irregular firing and population oscillations in a modular attractor memory network. *PLoS Computational Biology*, 6(6):e1000803, 2010.

Anita Lüthi and David A McCormick. H-current: properties of a neuronal and network pacemaker. *Neuron*, 21(1):9–12, 1998.

William W Lytton and Michael L Hines. Independent variable time-step integration of individual neurons for network simulations. *Neural computation*, 17(4):903–921, 2005.

Wolfgang Maass. Networks of spiking neurons: the third generation of neural network models. *Neural networks*, 10(9):1659–1671, 1997.

Jeffrey C Magee and Daniel Johnston. Plasticity of dendritic function. *Current opinion in neurobiology*, 15(3):334–342, 2005.

Zachary F Mainen and Terrence J Sejnowski. Reliability of spike timing in neocortical neurons. *Science*, 268(5216):1503–1506, 1995.

Adam H Marblestone, Greg Wayne, and Konrad P Kording. Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, 10:94, 2016.

Eve Marder. Neuromodulation of neuronal circuits: back to the future. *Neuron*, 76(1):1–11, 2012.

Henry Markram. The blue brain project. *Nature Reviews Neuroscience*, 7(2):153–160, 2006.

Henry Markram, Wulfram Gerstner, and Per Jesper Sjöström. A history of spike-timing-dependent plasticity. *Frontiers in synaptic neuroscience*, 3:4, 2011.

Henry Markram, Eilif Muller, Srikanth Ramaswamy, Michael W Reimann, Marwan Abdellah, Carlos Aguado Sanchez, Anastasia Ailamaki, Lidia Alonso-Nanclares, Nicolas Antille, Selim Arsever, et al. Reconstruction and simulation of neocortical microcircuitry. *Cell*, 163(2):456–492, 2015.

Chiara Marullo and Elena Agliari. Boltzmann machines as generalized hopfield networks: a review of recent results and outlooks. *Entropy*, 23(1):34, 2020.

Marcello Massimini, Reto Huber, Fabio Ferrarelli, Sean Hill, and Giulio Tononi. The sleep slow oscillation as a traveling wave. *Journal of Neuroscience*, 24(31):6862–6870, 2004.

Zachary T McCleney and Zachary P Kilpatrick. Entrainment in up and down states of neural populations: non-smooth and stochastic models. *Journal of mathematical biology*, pages 1–30, 2016.

David A. McCormick and Maria V. Sanchez-Vives. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nature Neuroscience*, 3(10):1027–1034, 2000.

Robert McEliece. *The theory of information and coding*. Cambridge University Press, 2002.

Thomas Miconi, Kenneth Stanley, and Jeff Clune. Differentiable plasticity: training plastic neural networks with backpropagation. In *International Conference on Machine Learning*, pages 3559–3568. Pmlr, 2018.

Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V–net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. Ieee, 2016. ISBN 1509054073.

P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. Benders, and I. Isgum. Automatic segmentation of mr brain images with a convolutional neural network.

IEEE Trans Med Imaging, 35(5):1252–1261, 2016. ISSN 1558–254X (Electronic) 0278–0062 (Linking). doi: 10.1109/TMI.2016.2548501. URL https://www.ncbi.nlm.nih.gov/pubmed/27046893.

Mikhail Mukovski, Sylvain Chauvette, Igor Timofeev, and Maxim Volgushev. Detection of active and silent states in neocortical neurons from the field potential signal during slow-wave sleep. *Cerebral Cortex*, 17(2):400–414, 2007.

Brendan K Murphy and Kenneth D Miller. Multiplicative gain changes are induced by excitation or inhibition alone. *Journal of neuroscience*, 23(31):10040–10051, 2003.

Andrew Y Ng. Feature selection, l 1 vs. l 2 regularization, and rotational invariance. In *Proceedings of the twenty–first international conference on Machine learning*, page 78, 2004.

Amy M Ni, Supratim Ray, and John HR Maunsell. Tuned normalization explains the size of attention modulations. *Neuron*, 73(4):803–813, 2012.

D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, and D. Shen. Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans Biomed Eng*, 65(12):2720–2730, 2018. ISSN 1558–2531 (Electronic) 0018–9294 (Print) 0018–9294 (Linking). doi: 10.1109/TBME.2018.2814538. URL https://www.ncbi.nlm.nih.gov/pubmed/29993445.

L. G. Nyul and J. K. Udupa. On standardizing the mr image intensity scale. *Magn Reson Med*, 42(6):1072–81, 1999. ISSN 0740–3194 (Print) 0740–3194 (Linking). doi: 10.1002/(sici)1522--2594(199912)42:6⟨1072::aid--mrm11⟩3.0.co;2--m. URL https://www.ncbi.nlm.nih.gov/pubmed/10571928.

Klaus Oberauer. Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of experimental psychology: General*, 134(3):368, 2005.

Z. Obermeyer and E. J. Emanuel. Predicting the future – big data, machine learning, and clinical medicine. *N Engl J Med*, 375(13):1216–9, 2016. ISSN 1533–4406 (Electronic) 0028–4793 (Print) 0028–4793 (Linking). doi: 10.1056/NEJMp1606181. URL https://www.ncbi.nlm.nih.gov/pubmed/27682033.

Emmanuel C Obusez, Mark Lowe, Se-Hong Oh, Irene Wang, Jennifer Bullen, Paul Ruggieri, Virginia Hill, Daniel Lockwood, Todd Emch, Doksu Moon, et al. 7t mr of intracranial pathology: preliminary observations and comparisons to 3t and 1.5 t. *Neuroimage*, 168: 459–476, 2018.

Tomokazu Ohshiro, Dora E Angelaki, and Gregory C DeAngelis. A normalization model of multisensory integration. *Nature neuroscience*, 14(6):775–782, 2011.

Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016.

Alan V Oppenheim. *Discrete-time signal processing*. Pearson Education India, 1999.

Randall C O'Reilly and Michael J Frank. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural computation*, 18(2): 283–328, 2006.

Randall C O'Reilly and Jerry W Rudy. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychological review*, 108(2):311, 2001.

Victor Pan. *Structured matrices and polynomials: unified superfast algorithms*. Springer Science & Business Media, 2001.

Nestor Parga and Larry F Abbott. Network model of spontaneous activity exhibiting synchronous transitions between up and down states. *Frontiers in neuroscience*, 1:4, 2007.

William D Penny, Karl J Friston, John T Ashburner, Stefan J Kiebel, and Thomas E Nichols. *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011. ISBN 0080466508.

Rodrigo Perin, Thomas K Berger, and Henry Markram. A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, 108(13): 5419–5424, 2011.

Olivier Petit, Nicolas Thome, Clement Rambour, Loic Themyr, Toby Collins, and Luc Soler. U–net transformer: Self and cross attention for medical image segmentation. In *Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 12*, pages 267–276. Springer, 2021. ISBN 3030875881.

David J Pinto, Joshua C Brumberg, Daniel J Simons, G Bard Ermentrout, and Roger Traub. A quantitative population model of whisker barrels: re-examining the wilson-cowan equations. *Journal of computational neuroscience*, 3(3):247–264, 1996.

Werner Plihal and Jan Born. Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, 9(4):534–547, 1997.

Werner Plihal and Jan Born. Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology*, 36(5):571–582, 1999.

Lutz Prechelt. *Early stopping–but when?*, pages 55–69. Springer, 2002.

William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

John G Proakis. *Digital signal processing: principles, algorithms, and applications, 4/E*. Pearson Education India, 2007.

C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE Trans Med Imaging*, 38(1):280–290, 2019. ISSN 1558–254X (Electronic) 0278–0062 (Linking). doi: 10.1109/TMI.2018.2863670. URL `https://www.ncbi.nlm.nih.gov/pubmed/30080145`.

Kanaka Rajan, Christopher D Harvey, and David W Tank. Recurrent network models of sequence generation and memory. *Neuron*, 90(1):128–142, 2016.

Pasko Rakic. Local circuit neurons. *Neurosciences Research Program Bulletin*, 13(3):295–416, 1975.

Hubert Ramsauer, Bernhard Schäfl, Johannes Lehner, Philipp Seidl, Michael Widrich, Thomas Adler, Lukas Gruber, Markus Holzleitner, Milena Pavlović, Geir Kjetil Sandve, et al. Hopfield networks is all you need. *arXiv preprint arXiv:2008.02217*, 2020.

Rajesh PN Rao and Dana H Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1):79–87, 1999.

Rajesh PN Rao, Bruno A Olshausen, and Michael S Lewicki. *Probabilistic models of the brain: Perception and neural function*. MIT press, 2002.

John H Reynolds and David J Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

Maximilian Riesenhuber and Tomaso Poggio. Hierarchical models of object recognition in cortex. *Nature neuroscience*, 2(11):1019–1025, 1999.

Pieter R Roelfsema and Arjen van Ooyen. Attention-gated reinforcement learning of internal representations for classification. *Neural computation*, 17(10):2176–2214, 2005.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U–net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer–Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015. ISBN 3319245732.

Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

Michael Rosenblum and Arkady Pikovsky. Synchronization: from pendulum clocks to chaotic lasers and chemical oscillators. *Contemporary Physics*, 44(5):401–416, 2003.

D. E. Ross, A. L. Ochs, J. M. Seabaugh, M. F. Demark, C. R. Shrader, J. H. Marwitz, and M. D. Havranek. Progressive brain atrophy in patients with chronic neuropsychiatric

symptoms after mild traumatic brain injury: a preliminary study. *Brain Inj*, 26(12):1500–9, 2012. ISSN 1362–301X (Electronic) 0269–9052 (Linking). doi: 10.3109/02699052.2012.694570. URL https://www.ncbi.nlm.nih.gov/pubmed/22721509.

Karsten Roth, Timo Milbich, Samarth Sinha, Prateek Gupta, Bjorn Ommer, and Joseph Paul Cohen. Revisiting training strategies and generalization performance in deep metric learning. In *International Conference on Machine Learning*, pages 8242–8252. PMLR, 2020. ISBN 2640–3498.

Alex Roxin, Nicolas Brunel, and David Hansel. Role of delays in shaping spatiotemporal dynamics of neuronal activity in large networks. *Physical review letters*, 94(23):238103, 2005.

Jun John Sakurai and Eugene D Commins. Modern quantum mechanics, revised edition, 1995.

Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

Wojciech Samek, Thomas Wiegand, and Klaus-Robert Müller. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*, 2017.

T. Santini, Y. Zhao, S. Wood, N. Krishnamurthy, J. Kim, N. Farhat, S. Alkhateeb, T. Martins, M. Koo, T. Zhao, H. J. Aizenstein, and T. S. Ibrahim. In–vivo and numerical analysis of the eigenmodes produced by a multi–level tic–tac–toe head transmit array for 7 tesla mri. *PLoS One*, 13(11):e0206127, 2018. ISSN 1932–6203 (Electronic) 1932–6203 (Linking). doi: 10.1371/journal.pone.0206127. URL https://www.ncbi.nlm.nih.gov/pubmed/30481187.

T. Santini, M. Koo, N. Farhat, V. P. Campos, S. Alkhateeb, M. A. C. Vieira, M. A. Butters, C. Rosano, H. J. Aizenstein, J. Mettenburg, E. M. Novelli, and T. S. Ibrahim. Analysis of hippocampal subfields in sickle cell disease using ultrahigh field mri. *Neuroimage Clin*, 30:102655, 2021a. ISSN 2213–1582 (Electronic) 2213–1582 (Linking). doi: 10.1016/j.nicl.2021.102655. URL https://www.ncbi.nlm.nih.gov/pubmed/34215139.

T. Santini, S. Wood, N. Krishnamurthy, T. Martins, H. J. Aizenstein, and T. S. Ibrahim. Improved 7 tesla transmit field homogeneity with reduced electromagnetic power deposition using coupled tic tac toe antennas. *Sci Rep*, 11(1):3370, 2021b. ISSN 2045–2322 (Electronic) 2045–2322 (Linking). doi: 10.1038/s41598--020--79807--9. URL https://www.ncbi.nlm.nih.gov/pubmed/33564013.

M. Schaer, M. B. Cuadra, L. Tamarit, F. Lazeyras, S. Eliez, and J. P. Thiran. A surface–based approach to quantify local cortical gyrification. *IEEE Trans Med Imaging*, 27(2):161–70, 2008. ISSN 0278–0062 (Print) 0278–0062 (Linking). doi: 10.1109/TMI.2007.903576. URL https://www.ncbi.nlm.nih.gov/pubmed/18334438.

Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*, pages 92–101. Springer, 2010.

Samuel Sanford Shapiro and Martin B Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.

Sunil Sharma, Mani Kavuru, et al. Sleep and metabolism: an overview. *International journal of endocrinology*, 2010, 2010.

D. Shen, G. Wu, and H. I. Suk. Deep learning in medical image analysis. *Annu Rev Biomed Eng*, 19:221–248, 2017. ISSN 1545–4274 (Electronic) 1523–9829 (Print) 1523–9829 (Linking). doi: 10.1146/annurev--bioeng--071516--044442. URL `https://www.ncbi.nlm.nih.gov/pubmed/28301734`.

Y. Shu, A. Hasenstaub, and D.A. McCormick. Turning on and off recurrent balanced cortical activity. *Nature*, 423(6937):288–293, 2003.

Karen Simonyan and Andrew Zisserman. Two-stream convolutional networks for action recognition in videos. *Advances in neural information processing systems*, 27, 2014.

P Jesper Sjostrom, Ede A Rancz, Arnd Roth, and Michael Hausser. Dendritic excitability and synaptic plasticity. *Physiological reviews*, 88(2):769–840, 2008.

S. M. Smith. Fast robust automated brain extraction. *Hum Brain Mapp*, 17(3):143–55, 2002. ISSN 1065–9471 (Print) 1097–0193 (Electronic) 1065–9471 (Linking). doi: 10.1002/hbm.10062. URL `https://www.ncbi.nlm.nih.gov/pubmed/12391568`.

H Sompolinsky and I Kanter. Temporal association in asymmetric neural networks. *Physical review letters*, 57(22):2861, 1986.

Sen Song, Per Jesper Sjöström, Markus Reigl, Sacha Nelson, and Dmitri B Chklovskii. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3): e68, 2005.

S. N. Sotiropoulos, S. Jbabdi, J. Xu, J. L. Andersson, S. Moeller, E. J. Auerbach, M. F. Glasser, M. Hernandez, G. Sapiro, M. Jenkinson, D. A. Feinberg, E. Yacoub, C. Lenglet, D. C. Van Essen, K. Ugurbil, T. E. Behrens, and W. U-Minn HCP Consortium. Advances in diffusion mri acquisition and processing in the human connectome project. *Neuroimage*, 80:125–43, 2013. ISSN 1095–9572 (Electronic) 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2013.05.057. URL `https://www.ncbi.nlm.nih.gov/pubmed/23702418`.

Olaf Sporns. The human connectome: a complex network. *Annals of the new York Academy of Sciences*, 1224(1):109–125, 2011.

Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*, 2014.

M. Steriade. Synchronized activities of coupled oscillators in the cerebral cortex and thalamus at different levels of vigilance. *Cerebral cortex*, 7(6):583–604, 1997.

M. Steriade, RC Dossi, and A. Nunez. Network modulation of a slow intrinsic oscillation of cat thalamocortical neurons implicated in sleep delta waves: cortically induced synchronization and brainstem cholinergic suppression. *Journal of Neuroscience*, 11(10): 3200–3217, 1991.

M. Steriade, A. Nunez, and F. Amzica. A novel slow (¡ 1 hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *Journal of Neuroscience*, 13(8):3252, 1993a.

M. Steriade, A. Nunez, and F. 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, 1993b.

Gilbert Strang. *Introduction to linear algebra*. SIAM, 2022.

Student. The probable error of a mean. *Biometrika*, 6(1):1–25, 1908.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

Harvey A Swadlow. Efferent neurons and suspected interneurons in motor cortex of the awake rabbit: axonal properties, sensory receptive fields, and subthreshold synaptic inputs. *Journal of neurophysiology*, 71(2):437–453, 1994.

Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*, 2013.

A. A. Taha and A. Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*, 15(1):29, 2015. ISSN 1471–2342 (Electronic) 1471–2342 (Linking). doi: 10.1186/s12880--015--0068--x. URL `https://www.ncbi.nlm.nih.gov/pubmed/26263899`.

N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Med Image Anal*, 63:101693, 2020. ISSN 1361–8423 (Electronic) 1361–8415 (Linking). doi: 10.1016/j.media.2020.101693. URL `https://www.ncbi.nlm.nih.gov/pubmed/32289663`.

Amirhossein Tavanaei, Masoud Ghodrati, Saeed Reza Kheradpisheh, Timothée Masquelier, and Anthony Maida. Deep learning in spiking neural networks. *Neural networks*, 111: 47–63, 2019.

Timothy J Teyler and Jerry W Rudy. The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus*, 17(12):1158–1169, 2007.

Nicholas M Timme, Shinya Ito, Maxym Myroshnychenko, Sunny Nigam, Masanori Shimono, Fang-Chin Yeh, Pawel Hottowy, Alan M Litke, and John M Beggs. High-degree neurons feed cortical computations. *PLoS computational biology*, 12(5):e1004858, 2016.

I. Timofeev and M. Steriade. Low-frequency rhythms in the thalamus of intact-cortex and decorticated cats. *Journal of Neurophysiology*, 76(6):4152, 1996.

I. Timofeev, F. Grenier, M. Bazhenov, T. J. Sejnowski, and M. Steriade. Origin of slow cortical oscillations in deafferented cortical slabs. *Cerebral cortex (New York, N.Y. : 1991)*, 10(12):1185–1199, 2000.

Igor Timofeev, Maxim Bazhenov, Terrence Sejnowski, and Mircea Steriade. Cortical hyperpolarization-activated depolarizing current takes part in the generation of focal paroxysmal activities. *Proceedings of the National Academy of Sciences*, 99(14):9533–9537, 2002.

Giulio Tononi and Chiara Cirelli. Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*, 81(1):12–34, 2014.

Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 4489–4497, 2015.

AM Turing. The molecular basis of morphogenesis. *Phil. Trans. R. Soc.*, pages 37–72, 1952.

N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4itk: improved n3 bias correction. *IEEE Trans Med Imaging*, 29(6):1310–20, 2010. ISSN 1558–254X (Electronic) 0278–0062 (Print) 0278–0062 (Linking). doi: 10.1109/TMI.2010.2046908. URL `https://www.ncbi.nlm.nih.gov/pubmed/20378467`.

Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. arXiv preprint, 2016. arXiv:1607.08022.

Kamil Uğurbil. Magnetic resonance imaging at ultrahigh fields. *IEEE transactions on biomedical engineering*, 61(5):1364–1379, 2014. ISSN 0018–9294.

SS Vallender. Calculation of the wasserstein distance between probability distributions on the line. *Theory of Probability & Its Applications*, 18(4):784–786, 1974. ISSN 0040–585X.

Tijl A Van der Velden, Michel Italiaander, Wybe JM van der Kemp, Alexander JE Raaijmakers, AM Th Schmitz, Peter R Luijten, Vincent O Boer, and Dennis WJ Klomp. Radiofrequency configuration to facilitate bilateral breast 31p mr spectroscopic imaging and high-resolution mri at 7 tesla. *Magnetic resonance in medicine*, 74(6):1803–1810, 2015. ISSN 0740–3194.

Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

A. van Opbroek, M. A. Ikram, M. W. Vernooij, and M. de Bruijne. Transfer learning improves supervised image segmentation across imaging protocols. *IEEE Trans Med Imaging*, 34 (5):1018–30, 2015. ISSN 1558–254X (Electronic) 0278–0062 (Linking). doi: 10.1109/TMI.2014.2366792. URL https://www.ncbi.nlm.nih.gov/pubmed/25376036.

Matthias JP van Osch and Andrew G Webb. Safety of ultra–high field mri: What are the specific risks? *Current Radiology Reports*, 2:1–8, 2014.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

J. T. Vaughan, M. Garwood, C. M. Collins, W. Liu, L. DelaBarre, G. Adriany, P. Andersen, H. Merkle, R. Goebel, M. B. Smith, and K. Ugurbil. 7t vs. 4t: Rf power, homogeneity, and signal–to–noise comparison in head images. *Magn Reson Med*, 46(1):24–30, 2001. ISSN 0740–3194 (Print) 0740–3194 (Linking). doi: 10.1002/mrm.1156. URL https://www.ncbi.nlm.nih.gov/pubmed/11443707.

Raul Vicente, Leonardo L Gollo, Claudio R Mirasso, Ingo Fischer, and Gordon Pipa. Dynamical relaying can yield zero time lag neuronal synchrony despite long conduction delays. *Proceedings of the National Academy of Sciences*, 105(44):17157–17162, 2008.

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.

Maxim Volgushev, Sylvain Chauvette, Mikhail Mukovski, and Igor Timofeev. Precise long-range synchronization of activity and silence in neocortical neurons during slow-wave sleep. *The Journal of Neuroscience*, 26(21):5665–5672, 2006.

U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in mri. *IEEE Trans Med Imaging*, 26(3):405–21, 2007. ISSN 0278–0062 (Print) 0278–0062 (Linking). doi: 10.1109/TMI.2006.891486. URL https://www.ncbi.nlm.nih.gov/pubmed/17354645.

Bradley Voytek, Ryan T Canolty, Avgusta Shestyuk, Nathan E Crone, Josef Parvizi, and Robert T Knight. Shifts in gamma phase–amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Frontiers in human neuroscience*, 4:191, 2010.

G. Wang, E. Gong, S. Banerjee, D. Martin, E. Tong, J. Choi, H. Chen, M. Wintermark, J. M. Pauly, and G. Zaharchuk. Synthesize high–quality multi–contrast magnetic resonance imaging from multi–echo acquisition using multi–task deep generative model. *IEEE Trans Med Imaging*, 39(10):3089–3099, 2020. ISSN 1558–254X (Electronic) 0278–0062 (Linking). doi: 10.1109/TMI.2020.2987026. URL https://www.ncbi.nlm.nih.gov/pubmed/32286966.

Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 1451–1460. Ieee, 2018.

Xiao-Jing Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3):1195–1268, 2010.

Jason Weston, Sumit Chopra, and Antoine Bordes. Memory networks. *arXiv preprint arXiv:1410.3916*, 2014.

Ines Wilhelm, Susanne Diekelmann, and Jan Born. Sleep in children improves memory performance on declarative but not procedural tasks. *Learning & memory*, 15(5):373–377, 2008.

Hugh R Wilson and Jack D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1, 1972.

Hugh R Wilson and Jack D Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80, 1973.

Stephen Wolfram. Statistical mechanics of cellular automata. *Reviews of modern physics*, 55(3):601, 1983.

J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Isgum. Generative adversarial networks for noise reduction in low–dose ct. *IEEE Trans Med Imaging*, 36(12):2536–2545, 2017. ISSN 1558–254X (Electronic) 0278–0062 (Linking). doi: 10.1109/TMI.2017.2708987. URL `https://www.ncbi.nlm.nih.gov/pubmed/28574346`.

Xin Yi, Ekta Walia, and Paul Babyn. Generative adversarial network in medical imaging: A review. *Medical image analysis*, 58:101552, 2019.

Yumiko Yoshimura, Jami LM Dantzker, and Edward M Callaway. Excitatory cortical neurons form fine-scale functional networks. *Nature*, 433(7028):868–873, 2005.

Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? *Advances in neural information processing systems*, 27, 2014.

Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.

Paul A Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C Gee, and Guido Gerig. User–guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006. ISSN 1053–8119.

Rafael Yuste. From the neuron doctrine to neural networks. *Nature reviews neuroscience*, 16(8):487–497, 2015.

Dan Zakay. Subjective time and attentional resource allocation: An integrated model of time estimation. In *Advances in psychology*, volume 59, pages 365–397. Elsevier, 1989.

W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, and D. Shen. Deep convolutional neural networks for multi–modality isointense infant brain image segmentation. *Neuroimage*, 108:214–24, 2015. ISSN 1095–9572 (Electronic) 1053–8119 (Print) 1053–8119 (Linking). doi: 10.1016/j.neuroimage.2014.12.061. URL `https://www.ncbi.nlm.nih.gov/pubmed/25562829`.

Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation–maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57, 2001. ISSN 0278–0062.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image–to–image translation using cycle–consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004. ISSN 1076–6332.