A RESERVOIR OF TIMESCALES IN RANDOM NEURAL NETWORKS

by

NICOLAE ISTRATE

A DISSERTATION

Presented to the Department of Physics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: Nicolae Istrate

Title: A Reservoir of Timescales in Random Neural Networks.

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

Timothy Cohen	Chair
Luca Mazzucato	Advisor
Eric Corwin	Core Member
Jayson Paulose	Core Member
Marina Guenza	Institutional Representative

and

Krista Chronister	Vice Provost of Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2022

DISSERTATION ABSTRACT

Nicolae Istrate

Doctor of Philosophy

Department of Physics

June 2022

Title: A Reservoir of Timescales in Random Neural Networks.

The temporal activity of many biological systems, including neural circuits, exhibits fluctuations simultaneously varying over a large range of timescales. The mechanisms leading to this temporal heterogeneity are yet unknown. Here we show that random neural networks endowed with a distribution of self-couplings, representing functional neural clusters of different sizes, generate multiple timescales of activity spanning several orders of magnitude. When driven by a time-dependent broadband input, slow and fast neural clusters preferentially entrain slow and fast spectral components of the input, respectively, suggesting a potential mechanism for spectral demixing in cortical circuits.

CURRICULUM VITAE

NAME OF AUTHOR: Nicolae Istrate

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon, USA
The College of Wooster, Wooster, Ohio, USA

DEGREES AWARDED:

Doctor of Philosophy in Physics, 2022, University of Oregon
Master of Science in Physics, 2018, University of Oregon
Bachelor of Arts in Physics, 2015, The College of Wooster

AREAS OF SPECIAL INTEREST:

Theoretical and Computational Neuroscience

PROFESSIONAL EXPERIENCE:

Graduate Employee, University of Oregon, 2016-2022

PUBLICATIONS:

Stern, M., Istrate, N. and Mazzucato, L., 2021. A reservoir of timescales in random neural networks. arXiv preprint arXiv:2110.09165.

ACKNOWLEDGEMENTS

Many people who deserve to be acknowledged for their contributions and influences on my personal and professional development.

First and foremost, I would like to offer my gratitude to my wonderful big family, both from Moldova and US: Cornelia, Viorel, Bogdan-Laurențiu Istrate, Leigh, Peter, Oliver Rye, and Caitlin, Walter, Astrid, and Alice Banta. Their love, support, and encouragement have guided me my entire life. Thank you for always being there, for all your advice, and welcoming me into your family. I love you all!

A big note of gratitude goes to my teachers and academic mentors. My adviser, Dr. Luca Mazzucato, took me under his tutelage and guided me through the world of neuroscience. From him, I learned how to be a better scientist, a better programmer, and most importantly, a better human being. In addition, I would also like to acknowledge Dr. Timothy Cohen for being a remarkable committee chair and helping me through this entire process. Thank you to Drs. Eric Corwin, Jayson Paulose, and Marina Guenza for being inspirational teachers and serving on my committee.

A special thanks goes to our collaborator Merav Stern. Her previous research, along with current contributions and insights into this work, have been crucial for this project.

Next, I would like to give a shout out to our joined lab: Dr. James Murray for being a mentor and asking tough but important questions during our lab meetings; David Wyrick for being a great colleague, housemate, and friend; Elliott Abe for his inquisitive mind and insights into experimental neuroscience; Lia Papadopoulos, Christian Schmid, and Ben Lemberger for the insightful and fun conversations on a variety of topics ranging from free will to paragliding.

I would not have been able to get to this point without the support of my friends. Oliver Rye, my second brother, thank you for re-introducing me to the great outdoors and for the wonderful touring trips. Prakrit Shrestha, our friendship has lasted for a decade. I will never forget the sleepless nights in Taylor Hall trying to learn GR with Charlie. Gabe Barello and Zelda Grove, thank you for your support and guidance through some hard times; I am looking forward to more skiing and climbing trips. Ryan and Meli Hull, I really appreciate your friendship and enjoy our get-togethers for wine & cheese nights, oyster parties, DnD campaigns, or all of them combined! Erin Moody, I always look forward to our long bike rides.

I cannot go forward without sending *Un Mare Mulțumesc* to Andrei Bolocan and Nata Albot for keeping me closely connected (from far away) to my home, Moldova, through their fantastic show "Internetu' Grăiește."

Many others who deserve to be acknowledged for their help and support and I apologize if I forgot to mention them.

DEDICATION

During the last Apollo 15 moonwalk, Commander David Scott replicated the results of one of the most famous experiments: Galileo's dropping of objects from the leaning tower of Pisa. Commander David Scott simultaneously dropped a hammer and a falcon feather from the same height. Due to the lack of atmosphere on the moon, both objects touched the lunar surface simultaneously, thus showcasing one of the basic principles in classical physics - all objects fall at the same rate, regardless of their mass.

I did not get the chance to see this lunar experiment until my early twenties. However, when I was 8 years old, I got the opportunity to perform a version of this wonderful experiment with my grandfather, Liviu Racoviță. He taught me about the irrelevance of mass and the effects of air resistance on falling objects in his own way. He simultaneously dropped two sheets of paper from the same height. As expected, they landed at the same time. He crumbled one sheet, glued together two others, and asked me a pointed question. "If you think that objects with larger mass reached the ground first," - he was very particular about physics words, especially when it came to the difference between weight and mass - "do you expect the glued papers to fall faster?" This memory is imprinted in my brain forever because this

exact moment was my first introduction to the world of scientific inquiry, next to Papa Liviu (that's what I called my grandfather) as my personal physics teacher. His life story is inspirational! He was born in a loving family in Moșeni, a tiny village in northern Moldova. His parents valued education and ensured that their children (my grandfather and his sister) attended school, despite not having a school in their village. Every day, indifferent to the weather, my grandfather had to walk 5 miles there, then back, to attend classes, and then he had homework and house chores. There were no cars, buses, or electricity in the village. On top of this, at an early age, my grandfather had the misfortune of encountering soviet soldiers (they were in the middle of occupying the current territory of Moldova). These soldiers decided to separate my grandfather's family by deporting his father to the coal mines in Siberia. My grandfather's father was a school principal and thus seen as a threat in the eyes of the red army. Forced to grow up without a father, having to pick up the responsibilities left behind, and being publicly labeled as "coming from a dangerous family" by the occupiers did not break my grandfather's spirit and desire for knowledge. He followed his dream and dedicated his humble life to teaching students about the mysteries of the universe.

Papa Liviu passed away on July 12, 2021. However, he will always be alive in my heart and soul. He was a kind and loving father and grandfather, a fearful citizen always ready to defend his core values, and a soul who never stopped wondering and asking deep questions about the world that surrounded him.

I dedicate this work to my grandfather, Liviu Racoviță. May his life story inspires you in ways he inspired me!

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Summary

Experimental evidence shows that the temporal activity of many physical and biological systems exhibits fluctuations simultaneously varying over a large range of timescales. In condensed matter physics for example, spin glasses typically exhibit aging and relaxation effects whose timescales span several orders of magnitude [1]. In biological systems, metabolic networks of E. coli generate fluxes with a power-law distribution of rates [2, 3]. Gas release in yeast cultures exhibit frequency distributions spanning many orders of magnitude [4], endowing them with robust and flexible responses to the environment [5].

In the mammalian brain, a hierarchy of timescales in the activity of single neurons is observed across different cortical areas [6, 7, 8], as well as within the same area [9, 10, 11]. This heterogeneity was identified in the intrinsic activity of the brain at rest in the absence of stimulus. This suggest that multiple timescales are an intrinsic property of the dynamics of neural circuits. Leveraging computations over multiple timescales is beneficial for the brain. However, the neural mechanisms underlying the emergence of multiple timescales are not yet understood. We suggest here such mechanism.

In this work, we demonstrate a novel class of recurrent networks, capable of generating temporally heterogeneous activity whose multiple timescales span several orders of magnitude. We interpret this novel class of recurrent neural networks with multiple timescales as a regulating mechanism for the analysis of temporally complex external stimuli. We also suggest that the such networks can regulate the performance of complex behavior. Such behaviors (high-level schemas) can be broken down into simpler actions (low-level schemas). Thus, the short timescales in the network will

regulate the low-level schemas, meanwhile the longer timescales will regulate the higher-level schemas by flexibly sequencing the lower-level schemas.

We focus on random neuronal networks whose units are recurrently connected, with couplings that are chosen randomly. In our model, each network unit represents a functional cluster of cortical neurons with similar response properties. We interpret the unit's self-coupling as the neural cluster strength, reflecting the product of the cluster size and the average value of the recurrent synaptic coupling between its neurons. In the case where the self-couplings are zero or weak (order $1/\sqrt{N}$), random networks are known to undergo a phase transition from silence to chaos when the variance of the random couplings exceeds a critical value [12]. When the self-couplings are strong (order 1) and are all equal, a third phase appears featuring multiple stable fixed points accompanied by long transient activity [13]. In all these cases, all network units exhibit the same intrinsic timescale, estimated from their autocorrelation function.

We show that when the self-couplings are heterogeneous, a reservoir of multiple timescales emerges, where each unit's intrinsic timescale depends both on its own self-coupling and the network's self-coupling distribution. In particular, we find an exponential relationship between (a power of the) unit's self-coupling and its timescales. We analytically study the dynamics of a single unit in the limit of large self-coupling, revealing a new metastable regime described by colored noise-driven transitions between potential wells. We study the stimulus-response properties of our networks with heterogeneous self-couplings. In networks with zero or weak self-couplings chaotic activity is suppressed best at a single resonant frequency [14]. However, when we drive our networks with a time-dependent broadband input featuring a superposition of multiple frequencies, we find that the chaotic activity

is suppressed across multiple frequencies which depend on the units' respective self-couplings. We see that units with large and small self-couplings are preferentially entrained by the low and high frequency components of the input, respectively. This spectral specificity suggests that a reservoir of timescales may be a natural mechanism for cortical circuits to flexibly demix different spectral features of complex time-varying inputs.

CHAPTER I

THE BRAIN AS A HIERARCHICAL STRUCTURE

On a daily basis, the brain is bombarded with an ongoing flux of convoluted information, hence its overarching role is to analyze this information, structure it into comprehensible pieces, and form an appropriate response. To perform its role successfully, the brain evolved into spatial, highly specialized structures that perform particular functions. For example, the occipital lobes, one of the four main divisions of the cerebral cortex, located at the back of the brain, have an essential role in vision processing. The parietal lobes, located on top of the occipital lobes, are involved in processing sensory information. The temporal lobes are situated on each side of the brain hemisphere and are involved in object recognition, memory formation, and language; while the frontal lobes, found at the front of the brain, are responsible for making decision, planning, motor functions, and personality development [15]. This functional topographical separation is not sufficient. On top of it, the brain developed a spatial hierarchy responsible for signal processing at different stages. This is well illustrated in the visual system. The external images received through the eyes, the sensory organ, are projected onto the retina, where they are translated into electrical neural signals. From here, through the optic nerve, the neural signals are sent to the lateral geniculate nucleus (LGN), whose role is to map the visual input and regulate the flow and strength of the visual stimulus. Next, the neural signals follow a complicated path to the V1 through V8 regions in the visual cortex. All these regions perform specific functions but also are interconnected, thus offering feedback to each other. The feedback helps reduce processing errors and detect changes in the visual input. Some of the functions of these regions in the visual cortex are: spatial organization of the image through edge detection computations, control of visual

attention, detection of objects and their movement, executing eye movements, and even signaling other regions of the brain in case a non-visual response is required [16]. The structural organization of the visual system is illustrated in Fig. 1.

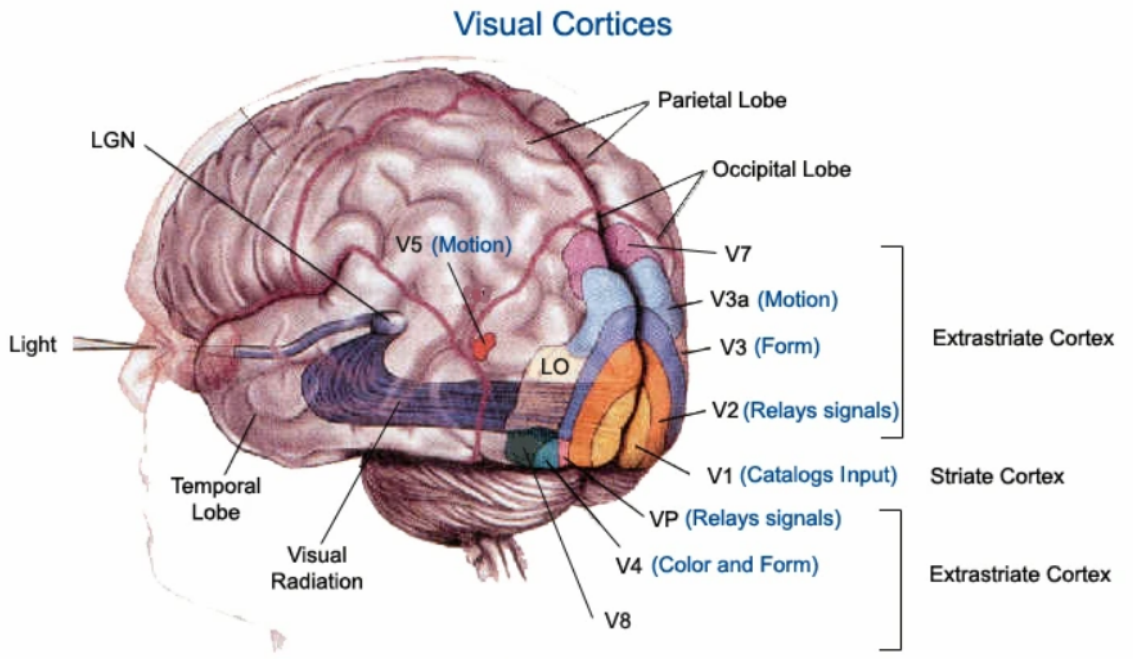

Figure 1. The structure of the visual system. The image received by the eye is transformed into electrical pulses which are sent to the LGN, and further to the occipital lobes in the V1 through V8 regions of the visual cortex. Each of these regions is specialized in performing specific tasks and provide feedback between them.

Besides the spatial hierarchical organization, the brain also manifests a hierarchical structure in its temporal dynamics. Neural circuits in the brain operate in a temporally irregular regime, where fluctuations in neural activity reveal the simultaneous presence of multiple timescales.

1.1 Bottom-Up Approach: Heterogeneous timescales

The environmental information that our brain receives is extremely complex even before it reaches the sensory organs. Before being analyzed, it has to be translated into a language that the nervous system understands (the raw information is transformed

into electrical impulses). This information is dynamic, constantly varying, and it is structured across multiple timescales. Consider the image of a tree, from a bird's-eye view perspective the tree seems static, however the raw image received through the eyes includes information about small movements of the leaves and branches, or even a squirrel cracking an acorn. On a longer timescale, the brain is able to process the static image of the tree. At the same time, on shorter timescales, it also detects rapid fluctuations due to the movement of leaves or the squirrel.

This heterogeneity of neural timescales is observed in awake animals during periods of ongoing activity, in the absence of external stimuli or behavioral tasks. This suggests that neural activity of multiple timescales may be an intrinsic property of recurrent cortical circuits. To measure a timescale, one can use the autocorrelation function of the intrinsic dynamic. The autocorrelation function is the correlation value between a signal and a delayed copy of itself for different values of delay. It informs us about the similarity between a signal and its lagged version. In case of chaotic activity, the autocorrelation decays to zero from its initial value. The timescale of the signal is dependent on how fast the autocorrelation decays. A short timescale will be characterized by a fast autocorrelation function decaying while a longer timescale corresponds to a slowly decaying autocorrelation function.

To study the temporal features in the brain, Murray et. al. [6] analyzed electrophysiological recordings in monkeys. They looked at single neuron activity in different cortical areas to characterize fluctuation in pre-stimulus neuronal activity. These fluctuations were quantified into regional intrinsic timescales through the autocorrelation function of the neuronal spike-count in each area. The results revealed a hierarchical ordering of intrinsic timescales across different brain regions, as seen in Fig.2.

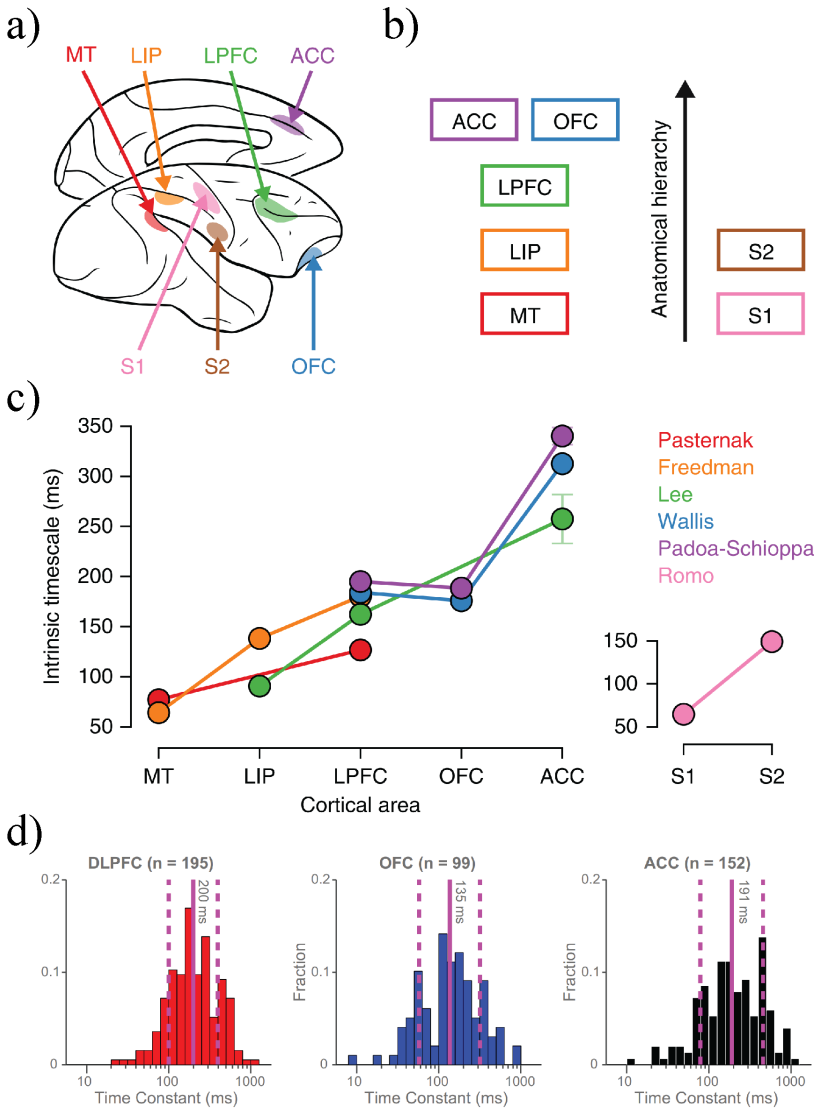

Figure 2. The analysis of single neuron spike trains revealed a consistent hierarchical ordering of intrinsic timescales across different brain regions. a) The data was collected across seven brain regions in 26 macaque monkeys: MT, LIP, LPFC, OFC, ACC, S1, and S2. b) The anatomical hierarchy of the brain regions based on long-range projections. c) The intrinsic timescales across the visual-prefrontal hierarchy (left), and the somatosensory hierarchy (right). Timescale variability within one cortical region. From left to right, we see the histogram of individual neuron timescales within the dorsal prefrontal cortex (DLPFC), orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC). The solid lines represent the mean of the timescales, $mean(\log(\tau))$, while the dashed line show the standard deviation from the mean, $mean(\log(\tau)) \pm stdev(\log(\tau))$. The parts a), b), c) of this figure were adapted from Murray et. al. [6] and part d) from Cavanagh et. al. [10].

They observed that fluctuations in the activity of sensory regions describe short timescales while fluctuations in the prefrontal cortex are slower and exhibit longer timescales. These result were consistent across multiple independently collected datasets. Another series of non-invasive experiments showed a similar heterogeneity. Spontaneous infra-slow (< 0.1 Hz) fluctuations in functional magnetic resonance imaging (fMRI) signals [17] along with fluctuations with faster frequencies 1 to 70 Hz from electroencephalography (EEG) and magnetoencephalography (MEG) experiments [18, 19] were collected from the human cortex in a resting state. These experiments showed a gradient of timescales within subcortical regions, especially in the thalamus and striatum, that mirrors the temporal hierarchy observed at the cortical level. Longer timescales were identified in higher-order transmodal regions like central-executive networks and dorsal attention networks, while significantly shorter timescales were detected in the unimodal sensory regions [20].

So far, we have seen multiple timescales in different regions of the brain, however such a heterogeneity was also observed within one cortical area. Cavanagh et. al. [10] examined the spike count autocorrelation function for individual neurons during pre-trial fixation. The neuronal activity data was collected from different regions of the prefrontal cortex in a macaque monkey. The results revealed a heterogeneity of timescales across individual neurons across regions and also within regions. Neurons showed the most variability in timescales in the dorsal prefrontal cortex (DLPFC), orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC). The distribution of timescales identified in this region are shown in Fig. 2 d).

The presence of heterogeneous timescales across different regions and also within a region of the brain is not only characteristic to only human and non-human primate

brains, but is found across different species as an intrinsic property of recurrent neural circuits.

1.2 Role of timescales

Recent studies highlighted the benefits of leveraging computations on multiple timescales when performing complex tasks in primates [21] as well as in artificial neural networks [22]. However, the role of the multiple timescales in the dynamics of the brain is not fully understood. Nevertheless, we have some insights that could explain their role. The intrinsic timescales in the brain are involved in processing and structuring the environmental external input into temporal frameworks. Hasson et.al. [23] argue that electrocorticography (ECoG), and functional imaging data suggest two main conclusions about how the information is encoded in the brain: (i) all cortical neural circuits are capable to accumulate (integrate) information over time, and (ii) the timescales over which information is accumulated manifests as a hierarchical structure: from short timescales (0 to 100 ms range) in the unimodal sensory cortical areas to long timescales (order of seconds to minutes) in higher-order transmodal areas. This hierarchical type of processing information suggests that memory is not only restricted to a specific location, but is tightly connected to how the brain processes external stimuli over multiple timescales. [23]

The heterogeneity of timescales in the brain facilitates the matching between the external stimuli and the intrinsic neuronal dynamics. They allow the brain to encode the temporal structure of external (environmental) inputs according to existing temporal structure in the brain. Shorter temporal segments from the incoming stimulus (a linguistic analogy would be the words in a story) are processed preferentially in the unimodal sensory regions with shorter timescales, meanwhile

longer intervals (paragraphs in the story) are simultaneously processed in transmodal regions with longer timescales. [20]

1.3 Top-Down Approach: Schema Theory

So far we have taken a bottom-up approach in analyzing cognitive functions. We used functional brain imaging techniques in order to correlate an observable (behavior, responses, neural dynamics) to a location in the brain. Now, let us consider a top-down approach in analyzing cognitive functions. From this perspective, we can use empirical studies of behavior, rather than brain imaging data, to develop theories about cognitive processes.

In the analysis of cognitive dynamics and behavior we make the assumption that a generally similar stimulus will trigger a generally similar response. Consider the relationship between prey and predator. A hungry frog that detects the presence of a fly (prey recognition stimulus) will attempt to catch it. Meanwhile, if the same frog detects the presence of a stork (predator detection stimulus), it will attempt to evade it. This stimulus-response relationship indicates that behavior is schematic [24]. One of these theories was first introduced by H.Head and G.Holmes and later transformed into what we currently refer to as Schema Theory by F.C.Bartlett. The basis for schema theory served the field of neuropsychology, which studies the effects of brain injuries on behavior. In 1911, while studying the effects of cerebral lesions due to disease, accidents, or surgery, in a *tour de force*, H.Head and G.Holmes were able to put forward a mechanism underlying sensation [25]. Two decades later, F.C.Bartlett [26], while studying how new memories are incorporated into the existing body of knowledge, presented an abstraction mechanism that connects stimuli to specific responses. The main idea behind the modern schema theory, is that complex behaviors (physical and cognitive) are made out of primitives (simpler behaviors)

that are sequenced together in a flexible manner. Consider the trivial behavior of drinking a cup of coffee. This behavior in itself can be considered a schema, and is made out of simpler schemas: reach out for the cup, grabbing the cup, lifting it up towards your mouth, drinking the coffee, and placing the cup on the table. This is an example of a motor schema since it involves a motor response. Other types of schemas are perceptual schemas that deal with processing of sensory input and may trigger a response in the form of a motor schema. An example of a sensory schema is the detection of prey / predator discussed at the beginning of this section. We will dive a bit deeper into schema theory in chapter 5.

Given the two approaches to neural behavior, (i) we develop a neural network mechanism that will replicate the heterogeneity of timescales observed from the data; and (ii) we propose this mechanism as a brain-inspired neural architecture able to mediate schemas at different levels. Instead of explicitly sequencing and defining the relationship on between lower- and higher-level schemas, we expect a self-organization of this functional hierarchy to emerge as a property of the network's dynamics (and learning) due to the heterogeneity of timescales. Slower timescales would represent the primitives while longer timescales their sequencing into complex schemas. Imagine a complex network architecture with a regulatory mechanism as a reservoir recurrent neural network with heterogeneous timescales trained to distinguish between inputs and provide appropriate responses.

We believe that using both, a bottom-up and a top-down approach in studying the dynamics and determining the capabilities of this neural architecture with multiple timescales will provide useful insights into complex cognitive processes which can be useful to the field of computational neuroscience and also artificial intelligence.

CHAPTER II

THE BRAIN AS A RECURRENT NEURAL NETWORK

The magnificent capabilities of the brain are evident. Despite the essential roles it plays for living organisms, the brain's anatomical and physiological complexity still baffles scientists. The brain is an organ made out of smaller cells called neurons. Traditionally, neurons are viewed as the basic computation units of the nervous system. They are specialized unidirectional cells that communicate between each other through voltage pulses called action potentials. Morphologically, the cell body of a neuron can be divided into dendrides, soma, and the axon, as shown in Fig. 3 a). Functionally, the dendrites, which are tree-like branched structures extending from the soma, receive input pulses from the surrounding neurons through the exchange of complex molecules[1], called neurotransmitters. There are multiple types of neurotransmitters like glutamate, dopamine, gama-aminobutyric acid, serotonin. Along with many other functions, they are also responsible for the closing/opening of ion channels to allow for ions to circulate in-and-out through the cell membrane. When a neuron incorporates these ions the cell membrane potential (the voltage difference between the inside and outside of the neuron) changes. The main body of the neuron, the soma, accumulates these ions and works as an integrator of the potential. Once the membrane potential reaches a threshold, a positive feedback chemical process is initiated and the neuron generates an action potential (electrical pulse) that propagates through the axon, a thin fiber extending from the soma towards other post-synaptic neurons. To study neurons and their behavior, neuroscientists use a variety of computational models. These models can account for different levels of complexity. There exist models ranging from simple binary neurons [27] to models of

[1]for alternative communication channels between neurons, see also electric synapses.

neurons described by large systems of coupled differential equations that can account for biochemical processes (e.g. Hodgkin-Huxley model [28]). Given such a large variety of models, it can be hard to chose a specific one. Hence, the choice of the model is ultimately determined by the nature of the question that the researcher is posing. For example, the Hodgkin-Huxley models of neurons are useful to understand the dynamics of the biophysical ion channels and identify potential mechanisms for firing patterns in different types of neurons. Meanwhile, firing rate models, which we focus on in this work, are a good choice when the goal is to study the dynamics of a network of neurons. Due to their relative simplicity, the firing rate models allows us to use analytical tools to study the input-output properties of neurons and can be easily expanded to account for interactions of multiple neurons.

In the next sections, we will start from the dynamics of a single neuron and slowly build up to a rate neural network that can manifest diverse chaotic activity. We will also provide analytical and numerical tools useful to analyze the temporal dynamics of such neuronal networks.

2.1 From Neurons to Rate Networks

Although action potentials are unique and slightly vary in duration, amplitude, and shape, on timescales larger than their average duration ($\approx$ 1 ms), it is useful to treat action potentials as point processes that convey information through their timing. Hence, mathematically, a sequence of n action potentials can be represented by a list of times t_i with $i = 1, 2, ..., n$ at which they occurred (Fig 3 b). The algebraic form of a spike sequence of a neuron is represented as a sum of Dirac δ functions known as the neural response function,

$$\rho(t) = \sum_{i=1}^{n} \delta(t - t_i). \tag{2.1}$$

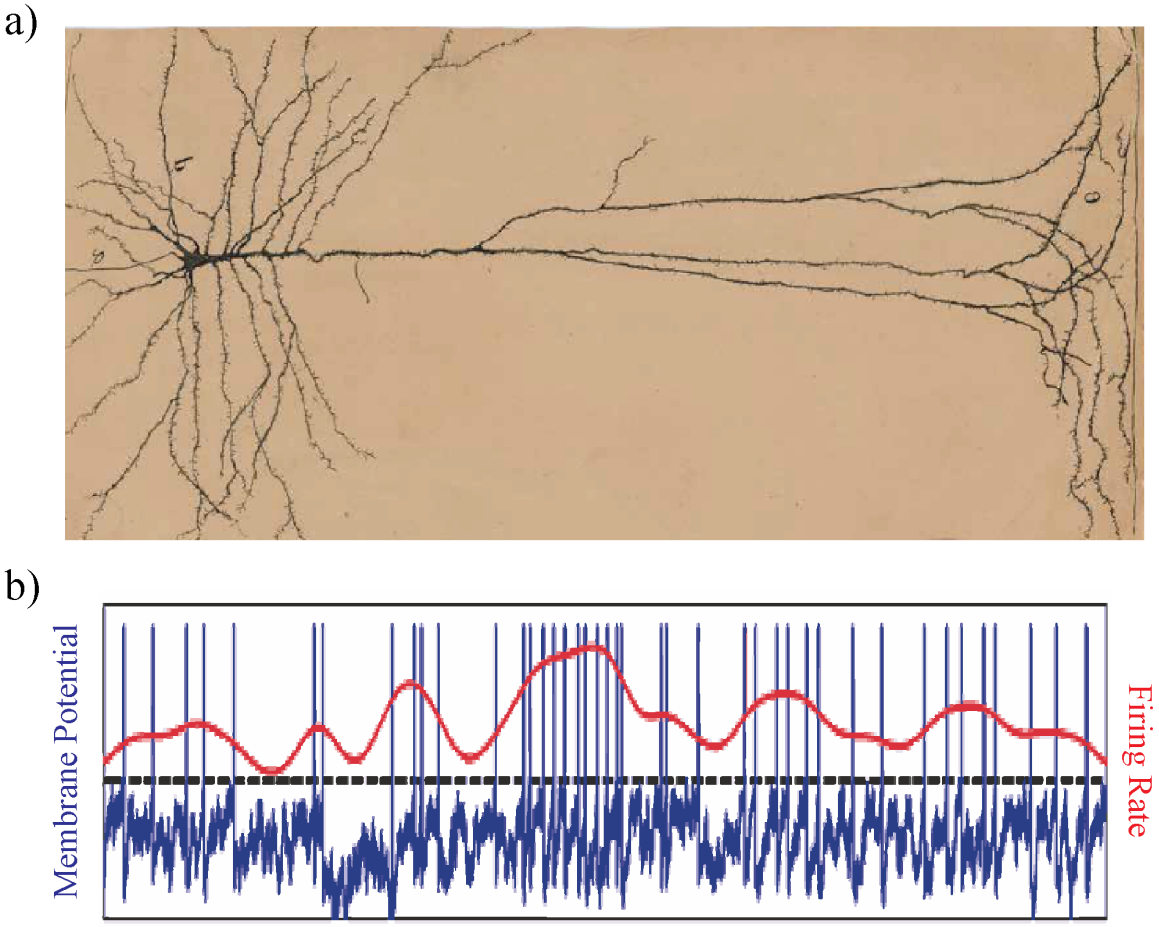

Figure 3. Structure of a neuron and its dynamic. a) A hand drawn image of a pyramidal neuron by Ramon y Cajal [29]. The soma (left) is located below the cerebral cortex. Branching on the left are the dendrites (order of mm) that receive signals from the presynamptic neurons. On the left, we see the the axon and axon terminals that communicate signals to other neurons. b) A cartoon representation of a neuron's behavior. In blue we see the membrane potential during a series of action potentials, the dashed blacked line represents the threshold value that the membrane potential needs to reach before an action potential is initiated. In red we have the firing rate of the neuron, calculated by binning action potentials.

The neural response function for a given neuron tends to vary from trial to trial and thus needs to be treated statistically or probabilistically. A more useful way to treat the dynamics of a neuron is to consider its firing rate. To move from the spike sequence to a firing rate, we can count the number of spikes over short interval of

times and divide it by the set interval. To reduce the variability across trials, we also average across multiple trials. As a result, we can move from a point process, to a continuous rate model Fig. 3 b).

Another reason to consider a rate model comes from experimental considerations. A series of experimental recording and imaging techniques used in neuroscience (e.g. EEG, MEG, fMRI, or widefield Ca^{2+}-sensitive imaging techniques) measure the average neuron population activity, rather than the activity of a single neurons. The rate model allows us to analyze such experiments, since they can be easily generalized to populations of neurons with similar response properties. Such generalized models are knows as Wilson-Cowan models [30].

The brain is made out of a large number of neurons (the human brain contains around $90 - 100$ billion neurons). Simulating the dynamics of such a system at the level of a single neuron is a tedious, if not impossible, task. To solve this problem we can turn to how related issues have been handled in physics. When studying the dynamics of gas atoms, physicists realized that solving the equation of motion for a large number (1 mol $\approx 6.022 \times 10^{23}$ units) of gas molecules is close to impossible. Instead, statistical physics was developed as a bridge between the microscale and macroscale description of the system. The key aspect of this approach is to define macroscopic observables as statistical properties of the microstates of the system. For example, the internal energy of a gas is defined as the average energy of all microstates. The analogy can be transferred to a system of neurons. Instead of studying the dynamics of networks of single neurons, we move to a macro scale by considering interactions between sub-populations of neurons, i.e. clusters. The way we perform this coarse-graining is by grouping neurons with similar properties together. These clusters serve as the building blocks of the macrosopic neural model. Such an approach

was taken by H.Wilson and J.Cowan [30]. They analyzed the collective behavior of large numbers of neurons using methods from statistical mechanics. Their model describes the dynamics of two separate populations of neurons, one excitatory and the other inhibitory. The schematic of interaction is shown in Fig. 4.

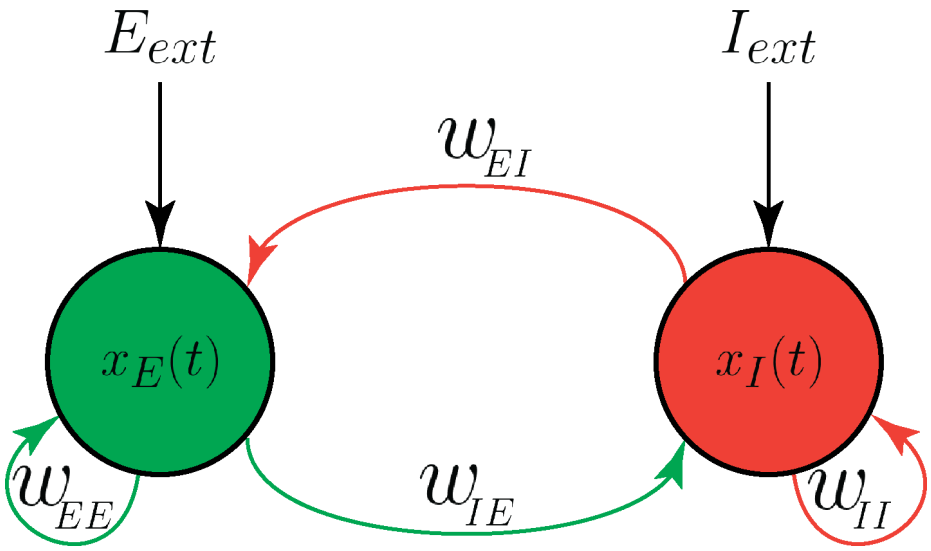

Figure 4. The Wilson Cowan Model describes the behavior of two sub-populations of neurons, one excitatory x_E and the other inhibitory x_I. Both sub-populations can interact with each other. Self-interactions are also allowed to account for the internal dynamics inside the sub-population. The model also allows for independent external input to each sub-population.

The model is represented as a set of coupled differential equations

$$\theta \frac{d}{dt} x_E(t) = -x_E(t) + \phi \left[w_{EE} x_E(t) - w_{EI} x_I(t) + E_{ext} \right]$$
$$\theta \frac{d}{dt} x_I(t) = -x_I(t) + \phi \left[w_{IE} x_E(t) - w_{II} x_I(t) + I_{ext} \right],$$

where the weights w_{EE}, w_{EI}, w_{II}, , w_{EI} are positive and describe the connection strength between the excitatory E and inhibitory I subpopulations. Notice that the inhibitory sub-population has the role of reducing the firing rate of the excitatory

sub-population. In the Wilson Cowan model, a sigmoid function was chosen as the activation function ϕ.

2.2 Cluster Model: Rate equations

In this work, we also consider a population based rate model. Our model consists of N clusters of neurons. Mathematically, each cluster can be represented as a localized continuous variable Q_i, $i = 1, 2, ..., N$ ranging between $Q_i(t) \in [-1.0, 1.0]$. Each cluster is associated with a local field $x_i(t)$ such that $Q_i(t) = \phi\left[x_i(t)\right]$, where ϕ is an activation function that defines the input-output relationship for each cluster. Biologically, $x_i(t)$ can be interpreted as the mean activity, also known as the firing rate or population activity.

Now that we defined the building blocks of our model, we can focus on the interactions and dynamics of such clusters. For start, we consider a linear recurrent neural network (RNN), a fully connected rate model with a linear activation function:

$$\frac{d}{dt}x_i(t) = -x_i(t) + \sum_j J_{ij}x_j(t), \tag{2.2}$$

where J_{ij} is an $N \times N$ matrix describing the strength of the recurrent connections between clusters. In a linear neural network the activation function $\phi[x] = x$. There are two types of dynamics present in the linear RNN in Eq. 2.2: stable and unstable. The stable patterns are described by a simple decay to zero or a non-trivial fixed point in case constant input is added to the network, meanwhile the unstable patterns represent cluster activity that blows up, as illustrated in Fig. 5.

To get a more diverse behavior we can introduce non-linearilty in the model through and activation function ϕ. The reason for introducing a non-linearity is motivated by the physiology of neuronal connections. As mentioned before, neurons communicate with each through action potentials. However, the neural signals are

$$\frac{d}{dt}x_i(t) = -x_i(t) + \sum_j J_{ij}x_j(t)$$

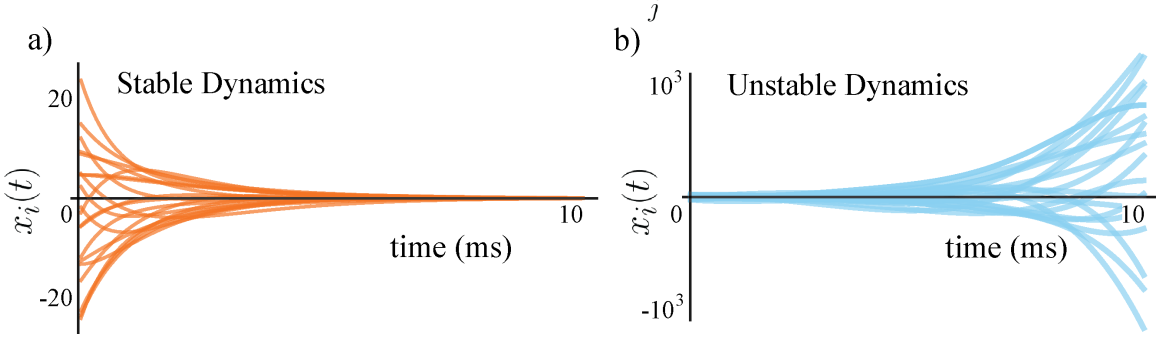

Figure 5. The network rate dynamics $x_i(t)$ in linear recurrent neural network can be stable or unstable.a) In the stable dynamics regime, the network starts at its initial conditions and quickly reaches a stable fixed point. b) In the unstable dynamics regime, the network starts at its initial conditions and as time passes, the firing rate of all units blows up to large values.

not directly shared, but depend on the amount and type of neurotranmitters released through the axon terminals of the pre-synaptic neurons and absorbed by the dendrites of the post-synaptic neuron. In a linear RNN the current in the model cluster depends on the pre-synaptic currents of other clusters weighted by the coupling strength. However, the neuron's output is not the sum of all of the inputs; it actually performs a nonlinear operation on it. For example, neurons integrate the inputs until a certain threshold is reached before an action potential happens. Neurons also are able to regulate the input, experimental evidence shows that there exist a saturating nonlinearity for large inputs [31]. Therefore, it stands to reason that the output (i.e. firing rate[2]) of a a cluster of neurons is a non-linear computation of the inputs due to all the clusters connected to it. We introduce the non-linearity in our model by assuming a sigmoid shape of the activation function ϕ; we chose $\phi[\cdot] \equiv \tanh[\cdot]$. It is worth noticing that for weak currents, the linear and non-linear RNN are equivalent,

[2]the negative firing rates are problematic and there are ways of countering that. However, we want to consider them as though they are relative to some kind of background.

while for stronger current, the nonlinearlity of the activation function prevents the network from blowing up. Other popular choices include a threshold-linear transfer function, logistic function, or even transfer functions fitted to experimental data. Given that our network is also made of clusters of neurons, we have to account for the internal activity of each cluster as well. This can be done by adding a self coupling term to the set of coupled ordinary differential equations from Eq. 2.2. Hence, the final form of our model becomes:

$$\frac{d}{dt}x_i(t) = -x_i(t) + s_i\phi\left[x_i(t)\right] + \sum_j J_{ij}\phi\left[x_j(t)\right]. \tag{2.3}$$

We can interpret Eq. 2.3 through the Kirchoff's laws perspective. If we imagine a cluster of neurons as a electrical element. The left hand side describes the average change in the potential, the first term on the right hand side describes the average current leakage through the membrane of neurons, the second terms describes the average current due to the activity of neurons in the cluster scaled by a self-coupling parameter s_i proportional to the number of neurons in the cluster. The last term on the right hand side describes the average current due to the activity of all other clusters. It is important to notice that the form of Eq. 2.3 implies that time is a dimensionless parameter. In this model, we are using a homogeneous time constant $\theta_i = 1$ ms for all clusters. The effect of non-homogeneous time constants on the temporal dynamics present in the network are described in Appendix B.

Compared to the linear RNN, the dynamical patterns in our model, rather than being trivial or exploding, become transient or ongoing. We recover the decay to a trivial fixed points seen in Fig. 5, however the network is also capable of oscillatory activity along with other rich, chaotic ongoing activities, Fig. 6.

$$\frac{d}{dt}x_i(t) = -x_i(t) + \sum_j J_{ij}\phi\left[x_j(t)\right]$$

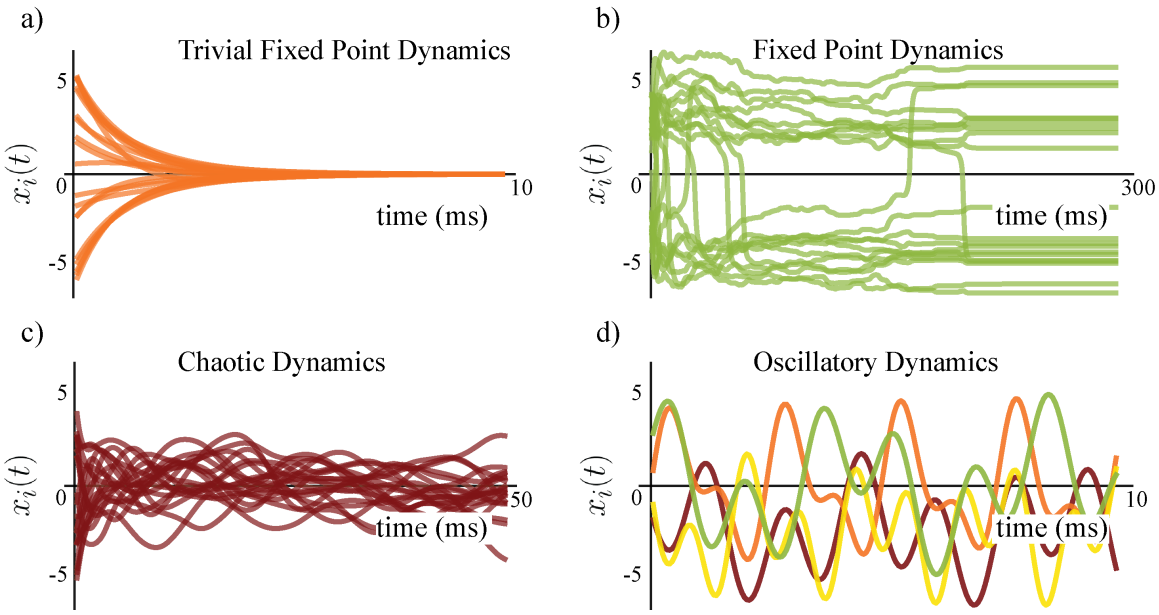

Figure 6. The network rate dynamics $x_i(t)$ in non-linear recurrent neural networks can be stable, chaotic, and also oscillatory. a) In the stable dynamics regime, the network starts at its initial conditions and quickly reaches a trivial stable fixed point. b) In the stable dynamics regime, the network starts at its initial conditions and fluctuates until it reaches a non-trivial fixed point. c) In the chaotic dynamics regime, the network firing rate fluctuates in a chaotic manner. d) In the oscillatory dynamics regime the firing rate of the units create complex, but periodic fluctuations.

The long-time behavior of the network depends on the particular realization of the **J** matrix, however in the large N thermodynamic limit we can identify a typical behavior. [12]

The case that we shall consider in this work is a non-symmetric **J** matrix. In this scenario, besides fixed points, the system can also exhibit limit cycles and even chaotic behavior.

Even in this regime, one can consider a nearly infinite set of connectivity matrices. We will focus on a fully connected RNN models where the J_{ij} and J_{ji} elements in the

connectivity matrix are uncorrelated. Furthermore, our model considers the elements of matrix **J** to be a collection of independent and identically distributed random Gaussian variables with zero mean and variance g^2/N, where g is a gain parameter that controls the recurrent coupling strength. The value of this coupling defines the type of synaptic connection: $J_{ij} > 0$ describes an excitatory connection, meanwhile $J_{ij} < 0$ describes an inhibitory one.

2.3 Dynamic Mean Field Theory

Our goal is to gain an analytical understanding of the dynamical behavior in our network. In particular, we are interested in determining the timescales present in the network along with understanding how they can be manipulated. For a large network size, N, certain quantities, the autocorrelation function being one of particular interest, are self-averaging. This means that their expectation value can be obtained by averaging over multiple network configurations [32]. Hence, we use a path integral approach to derive the Dynamic Mean Field Theory of our network. Path integral methods are well known in high energy physics and have been extended for the study of stochastic dynamics in statistical systems [33], critical phenomenon and renormalization group [34, 35, 36], and also stochastic dynamics of spin glasses [37, 38, 39]. Applying path integral methods to deterministic dynamical systems such as the one in Eq. 2.3 is not that common. However, the application of this methodology is facilitated by the presence of asynchronous chaotic states which generates deterministic fluctuations with stationary statistics. Our strategy of the path integral approach is to derive a generating functional for the relevant correlation and response function induced by the dynamics in our RNN.

Here we apply the dynamic mean field theory [40] to an RNN with self-couplings. We start by discretizing a finite dimensional time period, $t \in [t_0 = 0, T]$, into n

segments of duration dt, such that $ndt = T$ remains fixed. In discretization of Eq. 2.3, we use the Ito convention [41] to obtain:

$$x_i^{a+1} = x_i^a + \left(-x_i^a + s_i \phi \left[x_i^a \right] + \sum_j J_{ij} \phi \left[x_j^a \right] \right) dt, \tag{2.4}$$

where the notation x_i^a refers to the activity rate of cluster i at time $t = a * dt$. To recover the continuous limit, we take $n \to \infty$ and $dt \to 0$. Given the form above, we can write the probability density functional for the path $x(t)$ as a product of δ functions:

$$\begin{aligned} \mathcal{P}\left(x_i^1, x_i^2, ..., x_i^n \middle| x_i^0 \right) &= \prod_{a=0}^{n-1} \delta \left(x_i^{a+1} - y_i^{a+1}(x_i^a) \right) \\ &= \prod_{a=0}^{n-1} \delta \left(x_i^{a+1} - x_i^a - \left(-x_i^a + s_i \phi \left[x_i^a \right] + \sum_j J_{ij} \phi \left[x_j^a \right] \right) dt \right), \end{aligned} \tag{2.5}$$

where $y_i^{a+1}(x_i^a)$ was interpreted as the solution to Eq. 2.4 at timestep $a+1$ given the activity rate x_i^a at the previous timestep a. Using the Fourier representation of a Dirac Delta function

$$\delta(x) = \frac{1}{2\pi} \int e^{-ikx} dx,$$

we can write the probability density function as:

$$\begin{aligned}
\mathcal{P}\left(x_i^{1,..,n-1}\middle|x_i^0\right) &= \int \prod_a \frac{dk_i^a}{2\pi} \exp\left[-ik_i^a\left(x_i^{a+1} - x_i^a - \left(-x_i^a + s_i\phi[x_i^a] + \sum_j J_{ij}\phi[x_j^a]\right)dt\right)\right] \\
&= \int \prod_a \frac{dk_i^a}{2\pi} \exp\left[-ik_i^a\left(\frac{x_i^{a+1} - x_i^a}{dt} - \left(-x_i^a + s_i\phi[x_i^a] + \sum_j J_{ij}\phi[x_j^a]\right)\right)dt\right] \\
&= \int \prod_a \frac{d\tilde{x}_i^a}{2\pi} \exp\left[-\tilde{x}_i^a\left(\frac{x_i^{a+1} - x_i^a}{dt} - \left(-x_i^a + s_i\phi[x_i^a] + \sum_j J_{ij}\phi[x_j^a]\right)\right)dt\right],
\end{aligned} \tag{2.6}$$

where in the last line $\tilde{x}_i^a = ik_i^a$ and the appropriate integral is over the complex axis. Moving back to the continuous space by taking the appropriate limit of $dt \to 0$ and $n \to \infty$ while $T = n * dt$ remains constant. In this limit the following quanties take their continuous form as follows:

$$\begin{array}{lll}
x_i^a \to x_i(t) & \frac{x_i^{a+1}-x_i^a}{dt} \to \frac{d}{dt}x_i(t) \equiv d_t x_i(t) & \sum_a \to \int_{t_0=0}^{T} dt \\
\tilde{x}_i^a \to \tilde{x}_i(t) & & \\
\prod_{a,i} \frac{d\tilde{x}_i^a}{2\pi} \to \mathcal{D}\tilde{x}_i & & \prod_{a,i} dx_i^a \to \mathcal{D}x_i
\end{array}.$$

With these changes we can rewrite the probability density functional in terms of an action:

$$\mathcal{P}\left(x|x_0\right) \equiv \prod_i \mathcal{P}\left(x_i^{1,..,n-1}\middle|x_i^0\right) = \int \prod_i \mathcal{D}\tilde{x}_i \ \exp\{[-S\left(x_i(t), \tilde{x}_i(t)\right)]\}, \tag{2.7}$$

where the action is explicitly defined as:

$$S\left(x_i(t), \tilde{x}_i(t)\right) = \int dt\ \tilde{x}_i(t)\left(\left(d_t + 1\right)x_i(t) - s_i\phi[x_i(t)] - \sum_j J_{ij}\phi[x_j(t)]\right). \tag{2.8}$$

The action S is known as an example of an Martin-Siggia-Rose-Janssen-de Dominicis action, while the new field $\tilde{x}_i(t)$ is called an auxiliary field, or a response field. Such a

mathematical representation of the probability density function is a very useful and well studied representation in physics as well.

The next step is to write the moment generating functional for the probability density function in Eq. 2.7.

For a random variable Y with probability density function $p(y)$, one can find the n-th moment of Y as $\langle Y^n \rangle = \int y^n p(y) dy/Z$, where Z is the normalization constant. However, instead of calculating moments one by one, it is more efficient to define the moment generating funcition as

$$Z(\lambda) \equiv \int e^{\lambda y} p(y)\ dy.$$

To determine the n-th moment of Y, we just take the n-th order derivatives of $Z(\lambda)$:

$$\langle Y^n \rangle = \frac{1}{Z(0)} \frac{d^n}{d\lambda^n} Z(\lambda)\bigg|_{\lambda=0}$$

In our case the moment generating functional becomes:

$$Z(\lambda) = \int \mathcal{D}\tilde{x}\mathcal{D}x\ \exp\Bigg[\int dt \sum_i \tilde{x}_i(t)\Big((\partial_t + 1)x_i(t) - s_i\phi\left[x_i(t)\right] - \sum_j J_{ij}\phi\left[x_j(t)\right]\Big) + \lambda_i(t)x_i(t)\Bigg], \tag{2.9}$$

where $\mathcal{D}x = \prod_i \mathcal{D}x_i$.

Recall that the dynamics of the network in Eq. 2.3 depend on their statistics as a whole, in other words on the particular value of the gain parameter g. To capture these properties that are generic to the ensemble of the models, we further proceed by averaging the moment generating functional above over the normal distribution

describing connectivity matrix $\mathbf{J} \sim \mathcal{N}(0, g^2/N)$:

$$\begin{aligned}
\bar{Z} &= \langle Z(\lambda) \rangle_{\mathbf{J}} \\
&= \sqrt{\frac{N}{2\pi g^2}} \prod_{i,j} \int dJ_{ij}\ Z(\lambda) \exp\left\{ \left[-\frac{J_{ij}^2 N}{2g^2} \right] \right\} \\
&= \int \mathcal{D}\tilde{x}\mathcal{D}x \exp\left\{ \left[-\int dt \sum_i \tilde{x}_i(t) \big((d_t + 1)x_i(t) - s_i \phi\left[x_i(t)\right] \big) + \lambda_i(t) x_i(t) \right] \right\} \\
&\times \exp\left\{ \left[\frac{1}{2} \int dt dt' \left(\sum_i \tilde{x}_i(t) \tilde{x}_i(t') \right) \left(\frac{g^2}{N} \sum_j \phi\left[x_j(t)\right] \phi\left[x_j(t')\right] \right) \right] \right\}.
\end{aligned} \tag{2.10}$$

In deriving the last line we completed the square and made use of the property that the elements J_{ij} are independent and identically distributed. The coupling term between $\tilde{x}_i$ and x_j fields in the last line describes quantities dependent on four fields and we would like to decouple them into products of pair of fields. We can do that by defining the quantity

$$Q_1(t, t') \equiv \frac{g^2}{N} \sum_j \phi\left[x_j(t)\right] \phi\left[x_j(t')\right].$$

This new quantity represents a superposition of N weakly correlated contributions, hence, in the $N \to \infty$ limit, it will approach a Gaussian distribution by the Central Limit Theorem. This allows us to replace $Q_1(t, t')$ by its expectation value when performing a saddle point approximation to the lowest order. We enforce this definition by multiplying the disordered averaged moment generating functional by unity expressed as an integral over a Dirac delta function in its integral form:

$$1 = \int dQ_1 \delta[\ f(Q_1)] = \frac{1}{2\pi} \int dQ_1 dQ_2 \exp\{[Q_2\ f(Q_1)]\},$$

where the Q_2 integral is over the imaginary axis, recall $\tilde{x} = ik$. As a result, Eq. 2.10 becomes:

$$\begin{aligned}\bar{Z} = &\int \mathcal{D}\tilde{x}\mathcal{D}x \exp\left\{\left[-\int dt \sum_i \tilde{x}_i\big((d_t+1)x_i - s_i\phi\left[x_i(t)\right]\big) + \lambda_i x_i\right]\right\} \\ &\times \exp\left\{\left[\frac{1}{2}\int dtdt' \sum_i \tilde{x}_i(t)\tilde{x}_i(t')Q_1(t,t')\right]\right\} \\ &\times \int dQ_2 \exp\left\{\left[-\int dtdt' \frac{N}{g^2}Q_2(t,t')Q_1(t,t') + Q_2(t,t')\sum_j \phi\left[x_j(t)\right]\phi\left[x_j(t')\right]\right]\right\}.\end{aligned} \tag{2.11}$$

The disordered averaged moment generating functional can be written in a more elegant form. However, first we have to clarify the meaning behind the changes we are about to make. As previously mentioned, our network is made out of clusters of neurons whose sizes are quantified by the self-coupling term s_i. It is possible that every cluster is unique. However, we want to explore network dynamics generated by interactions between sub-populations of clusters of the same size (same s_i). Hence, we introduce the quantity $s_\alpha \in S$, where S denotes the set of different values of self-couplings s_α, indexed by $\alpha = 1, 2, ..., A$ and $A \leq N$. Further, we denote by N_α the number of units with the same self-coupling s_α, and accordingly by $n_\alpha = N_\alpha/N$ their fraction. Under this conditions, the disordered averaged moment generating

functional becomes:

$$\bar{Z} = \int dQ_2 \exp\Bigg\{ \Bigg[-\frac{N}{g^2} \int dt dt' Q_1(t,t') Q_2(t,t') + \sum_{\alpha=1}^{A} N_\alpha \ln[\Omega_\alpha] \Bigg] \Bigg\},$$

$$\Omega_\alpha = \int \mathcal{D}\tilde{x}_\alpha \mathcal{D}x_\alpha \exp\Bigg[-\int dt\ \tilde{x}_\alpha(t) \big((d_t + 1) x_\alpha(t) - s_\alpha \phi\left[x_\alpha(t)\right] \big) + \lambda_\alpha(t) x_\alpha(t)$$
$$+ \frac{1}{2} \int dt dt' \tilde{x}_\alpha(t) Q_1(t,t') \tilde{x}_\alpha(t') + \int dt dt' \phi\left[x_\alpha(t)\right] Q_1(t,t') \phi\left[x_\alpha(t')\right] \Bigg]. \tag{2.12}$$

In writing Eq. 2.12 we accounted for A types of clusters sharing the same value of self-couplings s_α. We also used that the auxiliary fields Q_1 and Q_2 couple to sums of the fields x_i^2 and $\phi\left[x_i\right]^2$. This allows to factorize the functional integrals of the fields into products of N_α factors of Ω_α. Next, we consider the thermodynamic limit of $N \gg 1$ by taking advantage of the saddle point approximation, i.e.

$$\lim_{M\to\infty} \int_a^b e^{Mf(x)} dx = \sqrt{\frac{2\pi}{-Mf\,''(z_0)}} e^{Mf(z_0)} \quad \text{where} \quad f\,'(z_0) = 0.$$

In the saddle point approximation we search for the stationary point of the action by requiring that:

$$0 = \frac{\partial S[Q_1, Q_2]}{\partial Q_{1,2}} = \frac{\partial}{\partial Q_{1,2}} \left[-\frac{N}{g^2} \int dt dt' Q_1(t,t') Q_2(t,t') + \sum_{\alpha=1}^{A} N_\alpha \ln[\Omega_\alpha] \right].$$

This corresponds to finding the point in the space of (Q_1, Q_2) which provides the dominant contribution to the probability mass. By applying the properties of the

moment generating functional,

$$\begin{aligned} 0 &= -\frac{N}{g^2}Q_1^*(t,t') + \sum_{\alpha=1}^{A}\frac{N_\alpha}{\Omega_\alpha}\frac{\partial\Omega_\alpha}{\partial Q_2(t,t')}\bigg|_{Q^*} \\ &\Leftrightarrow \frac{N}{g^2}Q_1^*(t,t') = \sum_{\alpha=1}^{A} n_\alpha \left\langle \phi\left[x_\alpha(t)\right]\phi\left[x_\alpha(t')\right]\right\rangle \equiv C_\phi(t,t'), \end{aligned} \tag{2.13}$$

$$\begin{aligned} 0 &= -\frac{N}{g^2}Q_2^*(t,t') + \sum_{\alpha=1}^{A}\frac{N_\alpha}{\Omega_\alpha}\frac{\partial\Omega_\alpha}{\partial Q_1(t,t')}\bigg|_{Q^*} \\ &\Leftrightarrow Q_2^*(t,t') = \frac{g^2}{2}\left\langle \tilde{x}(t)\tilde{x}(t')\right\rangle = 0, \end{aligned} \tag{2.14}$$

where $C_\phi(t,t')$ represents the average autocorrelation function of the network. Meanwhile, the second saddle point $Q_2^* = 0$ must vanish since it alters the normalization of the moment generating functional by mixing the retarded and non-retarded time derivatives which gives rise to acausal response functions [40, 12, 42].

Inserting the saddle points back into Eq. 2.12, the moment generating functional becomes

$$\begin{aligned} \bar{Z} = \sum_{\alpha=1}^{A}\int \mathcal{D}\tilde{x}_\alpha\mathcal{D}x_\alpha \exp\Bigg[&- N_\alpha\int dt\tilde{x}_\alpha\big((\partial_t+1)x_\alpha - s_\alpha\varphi_\alpha + \lambda x_\alpha\big) \\ &+ \frac{g^2}{2}\int dtdt'\tilde{x}_\alpha C_\varphi(t,t')\tilde{x}_\alpha\Bigg] \end{aligned} \tag{2.15}$$

This result indicates that in the large thermodynamic limit of $N \gg 1$ a network with $A \leq N$ self-couplings from Eq. 2.3 simplifies into A coupled differential equations sharing a common term:

$$\begin{aligned}
\frac{d}{dt}x_1(t) &= -x_1(t) + s_1\phi\left[x_1(t)\right] + \eta(t) \\
\frac{d}{dt}x_2(t) &= -x_2(t) + s_2\phi\left[x_2(t)\right] + \eta(t) \\
&\cdots \\
\frac{d}{dt}x_A(t) &= -x_A(t) + s_A\phi\left[x_A(t)\right] + \eta(t),
\end{aligned} \tag{2.16}$$

where $\eta(t)$ can be interpreted as common Gaussian noise shared by all clusters with the autocorrelation $\langle\eta(t)\eta(t')\rangle = g^2 C_\phi(t,t')$ calculated in Eq. 2.13. This result is quiet remarkable since it tells us that in our model, clusters of the same size exhibit the same temporal dynamics. This simplifies our situation a lot, instead of studying the dynamics of $N \gg 1$ individual clusters, we can replace it by a model of $A \ll N$ interacting clusters that share a common source of Gaussian noise whose statistics are self-consistent and can be easily calculated.

2.4 Mean Field and Autocorrelation Function

In the previous section we showed that an RNN with $N \gg 1$ clusters behaves statistically identical to model of $A \ll N$ distinct clusters that share a common source of Gaussian noise. In what follows we will provide an analytical analysis of the network autocorrelation function. This will help with understanding and quantifying the types of temporal dynamics we identify in our networks.

For start, let us consider the simple case of the RNN in Eq. 2.3 with no self-couplings, ($s_i = 0$),

$$\frac{d}{dt}x_i(t) = -x_i(t) + \sum_j J_{ij}\phi\left[x_j(t)\right]. \tag{2.17}$$

The dynamics of this RNN model were extensively studied by Sompolinsky, Crisanti, and Sommers[12]. In the large N limit, the effective mean-field dynamics of this

network simplifies to the dynamics of a single unit

$$(d_t + 1)x(t) = \eta(t), \tag{2.18}$$

driven by a Gausian noise with with auto-correlation $\langle \eta(t)\eta(t') \rangle = g^2 C_\phi(t, t')$ as shown in Eq. 2.13.

Our goal is to represent the self consistent autocorrelation as motion of a particle in a potential. First, let us use the knowledge of the noise statistics by first multiplying Eq. 2.18 by itself at a different time t' and then taking the expectation value of the new expression:

$$(d_t + 1)(d_{t'} + 1)C_x(t, t') = g^2 C_\phi(t, t'), \tag{2.19}$$

where $C_x(t, t') \equiv \langle x(t)x(t') \rangle$. We are interested in the stationary statistics of the system, i.e. $C_x(t, t') = C_x(\tau)$ where $t - t' = \tau$. Given that C_ϕ is time-translation invariant, $C_\phi(t + \tau, t)$ is also a function of τ. The derivatives in the equation above, become derivatives with respect to τ:

$$d_{t'} = d_{t-\tau} = -d_\tau$$

$$d_t = d_{\tau - t'} = d_\tau.$$

Hence, we obtain:

$$(-d_\tau^2 + 1)C_x(\tau) = g^2 C_\phi(\tau). \tag{2.20}$$

By solving this equation we can learn about the covariance function $C_x(\tau)$ between two points τ apart as a function of the initial condition given by the variance $C_0 \equiv C_x(0)$.

We continue with the replacement of our variable $x(t)$ by a sum of two other Gaussian variables that satisfy the same statistics:

$$\begin{aligned} x(t) &= \alpha z_1(t) + \beta z_2(t) \quad \text{such that} \quad \left\langle z_1(t)^2 \right\rangle = \left\langle z_2(t)^2 \right\rangle = 1 \\ 0 &= \langle x(t) \rangle = \alpha \langle z_1(t) \rangle + \beta \langle z_2(t) \rangle \\ &\Rightarrow \quad \langle z_1(t) \rangle = \langle z_2(t) \rangle = 0 \\ C_0 &= \langle (\alpha z_1(t) + \beta z_2(t)) (\alpha z_1(t) + \beta z_2(t)) \rangle = \alpha^2 + \beta^2. \end{aligned}$$

If we let

$$\beta = \frac{C_0}{\sqrt{C_x(\tau)}} \quad \Rightarrow \quad \alpha = \sqrt{C_0 - \frac{C_x(\tau)^2}{C_0}},$$

the initial value of our variable becomes

$$x(0) = \sqrt{c_0} z_2(t).$$

Before going further in finding an analytical expression for the autocorrelation function, we notice that given the moment generating functional of a Gaussian theory in Eq. 2.15 with $s = 0$ and $x(t)$ as a zero mean Gaussian variable, the second moment completely determines the distribution, hence $C_\phi(\tau) = g^2 f_\phi(C_x(\tau), C_0)$, where we used

$$f_u(C_x(\tau), C_0) = \int \mathcal{D}z_1 \mathcal{D}z_2 \; u\left(\sqrt{C_0 - \frac{C_x(\tau)^2}{C_0}} z_1(t) + \frac{C_x(\tau)}{\sqrt{C_0}} z_2 \right) u\left(\sqrt{C_0} z_2(t) \right), \tag{2.21}$$

where $\mathcal{D}z_i = \exp\left(\frac{-z_i^2}{2}\right) dz_i$. By substituting the result above into Eq. 2.20, we obtained a closed system for determining the autocorrelation functions for $x(t)$ and $\phi(t)$.

Given the result we have arrived to, it is worth looking back at Eq. 2.20. Its mathematical form describes the motion of particle with unit mass in a potential, however, there is a slight complication. While in the case of Newtonian mechanics, C_0 would describe the initial position of the particle, here, we have a slight complication due to the relationship between C_ϕ and C_0 in Eq. 2.21. This complication indicates that the potential in which the particle moves, depends not only on the position of the particle $C_x(\tau)$ at time τ, but also on the initial position of the particle C_0.

The potential function can be written in the following manner

$$V(C_x, C_0) = -\frac{1}{2}C_x^2 + g^2 f_\Phi(C_x, C_0) - g^2 f_\Phi(0, C_0), \tag{2.22}$$

where the last term is an arbitrary constant that ensures that $V(0, C_0) = 0$ and Φ is the primitive of ϕ. In obtaining the potential, we made use of Price's theorem [43, 40] (i.e. $d_c\ f_\Phi = f_{d_c\ \Phi} = f_\phi$) and that the dependence of $f_u(C_x, C_0)$ on τ comes only through $C_x(\tau)$. This allows us to write Eq. 2.20 in terms of the potential $V(C_x, C_0)$

$$-d_\tau^2 C_x(\tau) = -\frac{\partial}{\partial C_x} V(C_x, C_0). \tag{2.23}$$

The shape of the potential (Fig. 7) is determined by the interplay between two opposing terms: the $-1/2\ C_x^2$ downward bend and the upwards bend due to $g^2 f_\Phi(C_x, C_0)$.

A further in depth analysis hints at multiple types of behavior given the specific value of the gain parameter g. For $g < 1$ the potential (Fig. 7) takes the form of a parabola centered at 0 driving the system into the zero fixed point, $x_i(t) = 0$. For $g > 1$ the $g^2 f_\Phi$ starts to dominate the curvature of the potential close to $C_x \approx 0$.

With increasing C_0, the curvature of the potential at $C_x \approx 0$ changes from negative to positive. In the intermediate regime, the potential assumes a double well shape. [40]

The behavior of the system in the case of $g > 1$ can be studied through finding the maximum Lyapunov exponent. The maximum Lyapunov exponent describes the sensitivity of the system to initial conditions by measuring the asymptotic growth rate of infinitesimal perturbation. We consider two replicas of the same network with the same **J** matrix. Generally following the same procedures as before, we find that the maximum Lyapunov exponent is positive. This indicates that the distance between two initially close trajectories grows exponentially with time, characteristic of chaotic dynamics [12, 44].

Using the conservation of energy argument, we can numerically determine the *initial position* of the particle, C_0, through a bisectioning algorithm. The goal is to find the root of $V(C_0, C_0) = 0$.

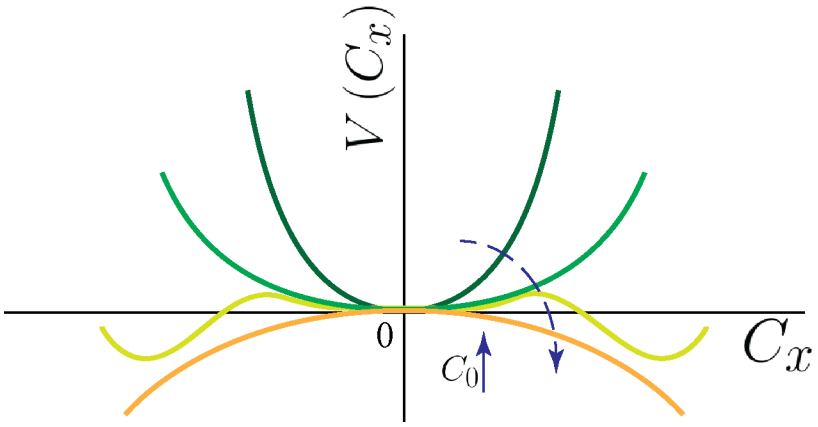

Figure 7. The shape and change of the potential $V(C_x, C_0)$ as a function of the autocorrelation function $C_x(\tau)$ for increasing values of the chaotic variance C_0. The initial value of the aurocorrelation function, C_0, is analogous to a *initial position* of the unit mass particle, its value increases as the network is more chaotic. This figure was adapted from M.Helias et.al. [40].

Finally, we can obtain the autocorrelation function by integrating Eq. 2.23. This can be done numerically by rewriting the second order differential equation as a coupled set of first order equations:

$$\begin{cases} \dfrac{d}{d\tau}\, y(\tau) = C_x(\tau) - g^2 f_\phi\Big(C_x(\tau), C_0\Big) \\ \dfrac{d}{d\tau}\, C_x(\tau) = y(\tau) \end{cases}, \tag{2.24}$$

where $\tau > 0$ and $y(0) = 0$. The agreement between this result and a numerical simulation of a network of $N = 2000$ neurons with gain $g = 2.0$ can be seen in Fig. 8 a), and the behavior of the network in each mode can be seen in Fig. 8 b) and c).

Having familiarized ourselves with the Dynamic Mean Field Theory calculation for Eq. 2.3, in the next chapter we will explore the set of possible dynamical patterns along with the temporal behavior that our RNN class is capable of.

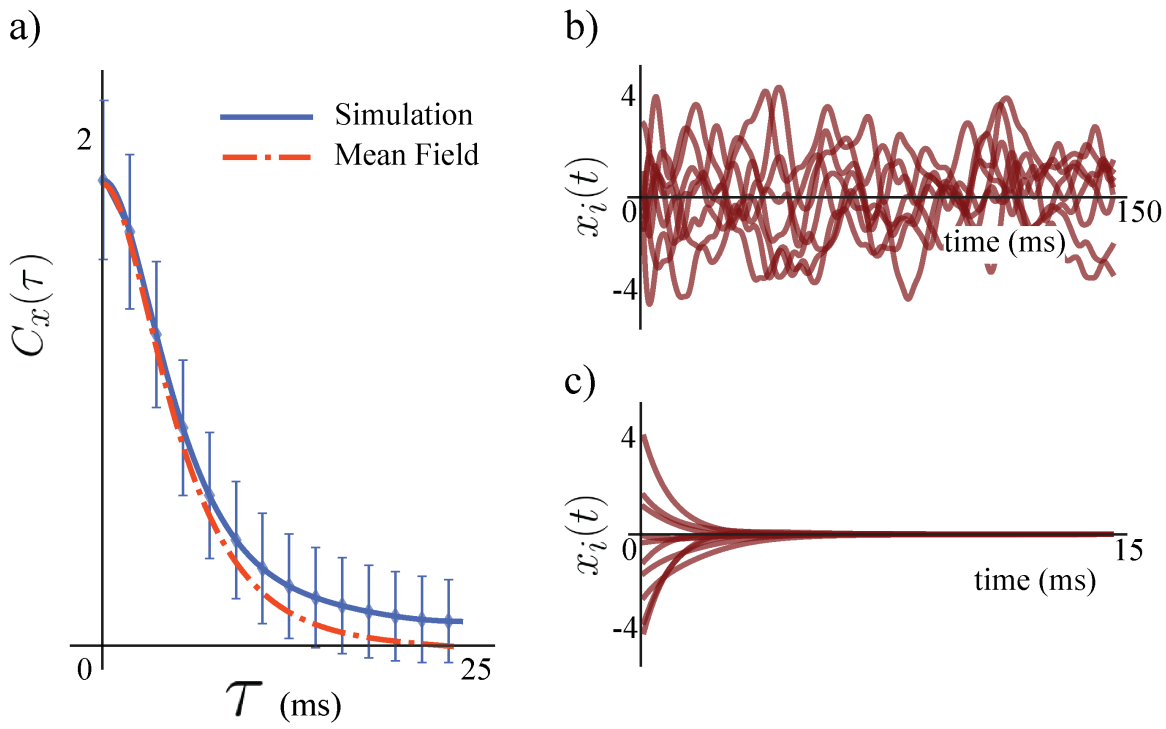

Figure 8. Behavior of a recurrent neural network in Eq. 2.17. (a) The autocorrelation function for $g = 1.8$ derived from the mean field result (red dashed line) in Eq. 2.24 compared to a network simulation (blue solid line) of $N = 2000$ clusters. (b) Illustrates the chaotic dynamics of the recurrent neural network in a). (c) Illustrates the decay to trivial fixed point dynamics in a recurrent neural network with $N = 2000$ clusters and gain $g = 0.5$.

CHAPTER III

NETWORK DYNAMICS AND TIMESCALES

Now that we have an analytical tool for studying recurrent neural networks, we can explore the temporal dynamics present in the network. Our goal is to understand the conditions necessary for the network to generate multiple timescales as well as the effect of different parameters on these timescales.

To build a bridge between our theoretical model and experiments, we have to quantify neuronal population dynamics in the temporal domain. We use a practical metric like the autocorrelation function of the firing rate to accomplish this. The value of the autocorrelation function at no delay describes the total variance of the temporal dynamics present in the network, meanwhile, the value of the autocorelation function at some time delay $\tau > 0$ gives us an insight of the temporal structure present in the network [14]. In a chaotic regime, which can describe the intrinsic spontaneous activity in the cortical and subcortical regions, the autocorrelation function decays to zero as the time delay increases. This behavior is indicative of uncorrelated firing rate dynamics for large time intervals. Thus, a slowly decaying autocorrelation function indicates temporal stability and corresponds to slow fluctuations of the neuron's firing rate, thus a large timescale. Meanwhile, a fast decaying autocorrelation function indicates variability (fast fluctuations) in the firing rate and corresponds to shorter timescales. To quantify the timescales present in the network we use the width at half amplitude of the autocorrelation functions.

Such a definition of timescales requires us to consider only regions of the parameter space in which the network temporal dynamics is continuous and avoids fixed points, meaning that the intrinsic dynamics of the network are chaotic. As we will see in this chapter, chaotic phases are common in large networks with random

connectivity. Chaotic dynamics have been proposed as a mechanism for generating cortical variability [45], meanwhile, in the field of machine learning, the rich network dynamics present at the edge of the chaotic phase have been useful in learning complex temporal patterns [46].

In the rest of this chapter we will explore the network dynamics and identify the timescales present in recurrent neural networks with homogeneous and heterogeneous self-couplings. Also, we will examine the emergence of multiple timescales within one network, and study their dependence on network parameters such as gain g, self coupling s_i, and cluster sub-population size N_i.

3.1 Timescale Dynamics: Homogeneous Self-coupling

The case of an RNN with homogeneous self-couplings, s, i.e. clusters having the same number of neurons, has been extensively studied by Stern et. al. [13]. Based on the values of the parameters g and s, three distinct types of network dynamics were identified: the trivial fixed point regime, the chaotic regime, and the transient chaos with fixed points regime. The parameter space of phase transformation along with examples of network dynamics in each phase can be seen in Fig. 5. Generally speaking, having the gain parameter g and the self-coupling s, has an effect of producing rich dynamics in our network. By changing their values we can alternate between different types of dynamics in the RNN.

To determine the phase boundaries of the chaotic region, we can explore the fixed-point solutions present in the network. We will follow the work done by Stern et.al. [13]. For illustrating the general procedure, we examine the simple case of the network in the trivial fixed point region. To analyze stability of the $x_i(t) = 0$ solution,

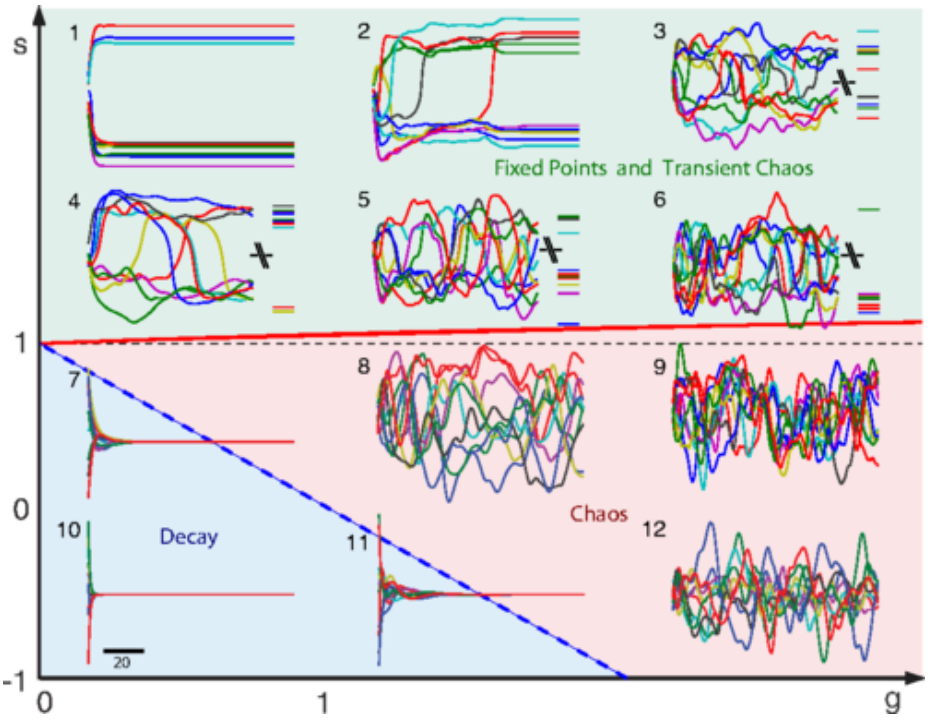

Figure 9. The phase space and dynamic of the network as a function of the gain g and self-coupling s parameters. Below the long dashed blue line any initial activity in the network decays to zero. Above the solid red curve, the network exhibits transient irregular activity that eventually settles into one out of a number of possible nonzero stable fixed points. In the region between these two curves, the network activity is chaotic. This figure was created by and borrowed from M. Stern et al. [13]

we compute the stability matrix of Eq. 2.3, with one self-coupling in the region:

$$0 = -x_i(t) + s\phi\left[x_i(t)\right] + \sum_j J_{ij}\phi\left[x_j(t)\right].$$

This is achieved through a Taylor expansion of $\phi\left[x_i(t)\right]$ around the trivial fixed point $x_i(t) = 0$. Neglecting higher order terms, we get

$$\begin{aligned} 0 &= (-1+s)x_i(t) + \sum_j J_{ij}x_j(t) \\ &= \sum_j \left[(-1+s)\delta_{ij} + J_{ij}\right] x_j(t), \end{aligned}$$

where the Kronecker delta δ_{ij} is used to match the appropriate indices. Hence, the stability matrix around the trivial fixed point is:

$$M_{ij} = (-1 + s)\delta_{ij} + J_{ij}.$$

Given the statistical properties of the connectivity matrix $\mathbf{J}$, the elements of whom are independent and identically distributed random variables extracted from a Gaussian distribution of zero mean and variance g^2/N, we can make use of the circlular law from random matrix theory [47, 48] to gain information about the phase boundary separating the trivial fixed-point and chaotic region. The eigenvalues of the stability matrix $\mathbf{M}$ belong in the complex plane and lie in a circle of radius g centered at $(-1 + s)$ on the real axis. The stability of the zero fixed-point requires the real part of the eigenvalues to be negative, which is satisfied by the simple inequality $(-1 + s) + g < 0$. Meanwhile, the phase boundary is determined by the condition $s = 1 - g$, which corresponds to the dashed blue line in Fig. 9. For $s > 1 - g$, the system is chaotic, while for $s < 1 - g$ the system reaches fixed points. For the case of the non-trivial fixed points, the procedure is similar, we expand around a non-zero fixed point, determine the stability matrix, and then examine its eigenvalues. For the case of nontrivial fixed points, the elements of the stability matrix are described by

$$M_{ij} = \left[-1 + s(1 - \tanh[x_j]^2)\right] \delta_{ij} + J_{ij}(1 - \tanh[x_j]^2).$$

As before, the stability of the fixed points requires that the real part of the eigenevalues to be negative. Unfortunately, the eigenvalue distribution of the stability matrix is not as easy to compute as in the trivial fixed-point case. It requires a careful derivation with the use of random matrix theory due to the non-normal nature of the

stability matrix. A careful derivation strategy is described by Ahmadian et.al. [49] meanwhile the results have been studied by Stern et.al. [13]. The boundary between the chaotic and nontrivial fixed points corresponds to the solid red line in Fig. 9.

As explained in the previous chapter, in the limit of a large number of homogeneous clusters, the dynamics of the network resembles the dynamics a single cluster driven by noise (Eq. 2.16):

$$\begin{aligned}
\frac{d}{dt}x_i(t) &= -x_i(t) + s\phi\left[x_i(t)\right] + \sum_{j=1}^{N} J_{ij}\phi\left[x_j(t)\right] \\
&\Downarrow \quad N \gg 1 \\
\frac{d}{dt}x(t) &= -x(t) + s\phi\left[x_i(t)\right] + \eta(t).
\end{aligned}$$

Differently put, in the large N limit, the network temporal dynamics are described by one autocorrelation function, hence only one timescale is present in the network. The magnitude of this timescale can be modified by changing the values of the gain g and the self coupling s parameter while the network remains in the chaotic regime. As we get closer to the zero fixed point transition line, the dynamics of the network slows down and the timescale increases. A larger timescale can be also achieved by increasing the value of the self-coupling term. The timescale dependence on the two network parameters can be seen in Fig. 10

3.2 Multiple Timescales: Heterogeneous Self-coupling

While recurrent neural networks with homogeneous self-coupling can create timescales of diverse magnitudes depending on the values of the gain and self coupling parameters, all clusters share the same timescale. A simple way of generating heterogeneous timescales in our network model is to consider a network with clusters of different sizes. For start, we consider the network from Eq. 2.3 with equal number of

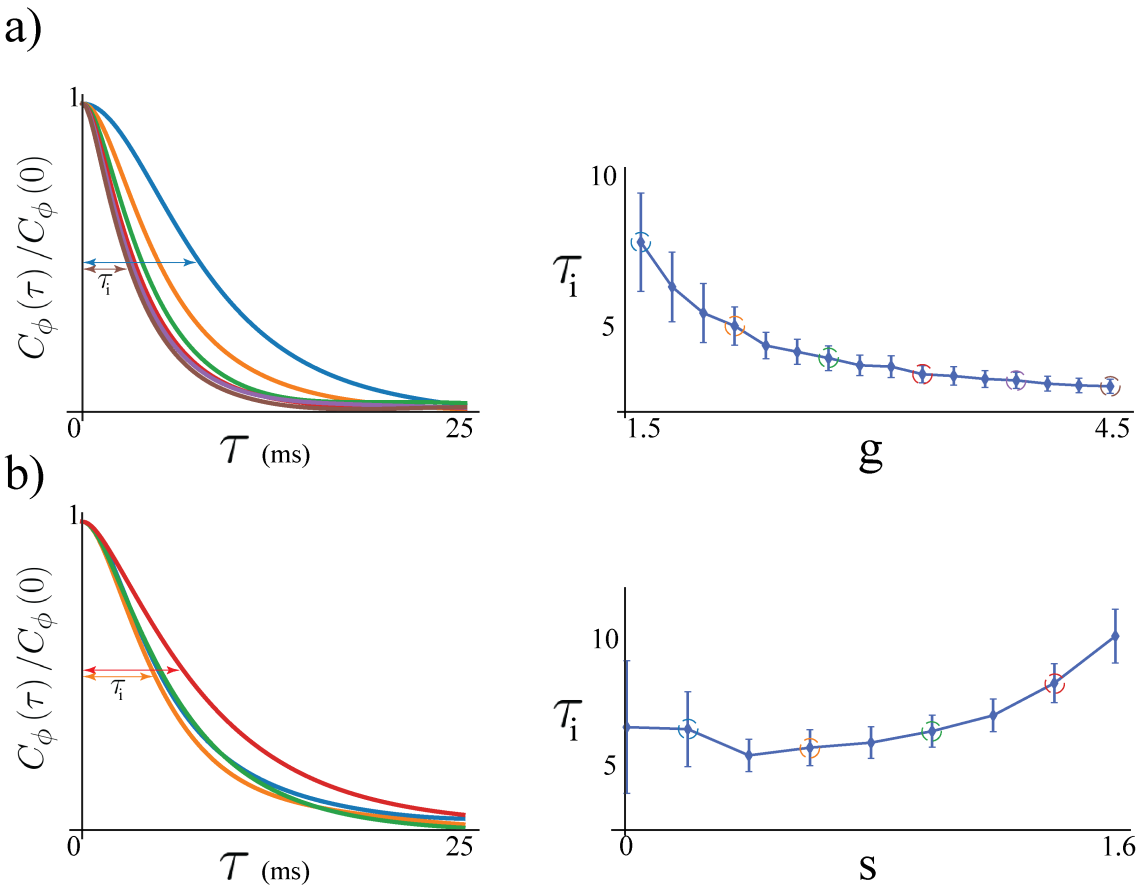

Figure 10. Autocorrelation and timescale dependence on network parameters (Eq. 2.3.a) Normalized autocorrelation function (left) and timescale dependence on the gain g parameter (right) for a network with $N = 3000$ and $s = 0$. b) Normalized autocorrelation function (left) and timescale dependence on the self coupling s parameter in a network with $N = 3000$ and $g = 2.0$.

interacting neuron clusters ($N_1 = N_2$) of two sizes, s_1 and s_2. In a large network, the dynamic mean field theory shows that the dynamics of such a network is equivalent to the interaction of only two clusters that share a common source of Gaussian noise, $\eta(t)$. The statistics of the noise are determined by the mean field as described in Eq. 2.13 and 2.16. In these network models, we recover the same three types of network dynamics as in the homogeneous self-coupling case. We can determine the phase

transition boundaries through analyzing the stability of the fixed points, through the methods referenced in the previous section.

Networks with heterogeneous self-couplings exhibit a complex landscape of fixed points x^*_α, obtained as the self-consistent solutions to the static version of Eq. 2.16 and Eq. 2.13, namely

$$x_\alpha - s_\alpha \tanh(x_\alpha) = \eta \ , \tag{3.1}$$

where the mean field η has zero mean and its variance is given by

$$\begin{aligned} \langle \eta^2 \rangle &= g^2 C_\phi \\ C_\phi(\tau) &= \sum_{\alpha \in A} n_\alpha \langle \phi[x_\alpha]^2 \rangle \ . \end{aligned} \tag{3.2}$$

The solution for each unit depends on its respective s_α (Fig. 11). If $s_\alpha < 1$ a single interval around zero is available. For $s_\alpha > 1$, for a range of values of η, x^*_α can take values in one of three possible intervals. However, the available solutions in the latter case are further restricted by stability conditions.

We can derive the stability condition by expanding the dynamical system in Eq. 2.3 around the fixed point and requiring that all eigenvalues of the corresponding stability matrix are negative. To determine the onset of instability we look for conditions such that at least one eigenvalue develops a positive real part. An eigenvalue of the stability matrix exists at a point z in the complex plane if [13, 49]

$$g^2 \sum_{\alpha \in A} n_\alpha \left\langle \frac{\left[1 - \tanh^2(x_\alpha)\right]^2}{\left[z + 1 - s_\alpha \left(1 - \tanh^2(x_\alpha)\right)\right]^2} \right\rangle > 1. \tag{3.3}$$

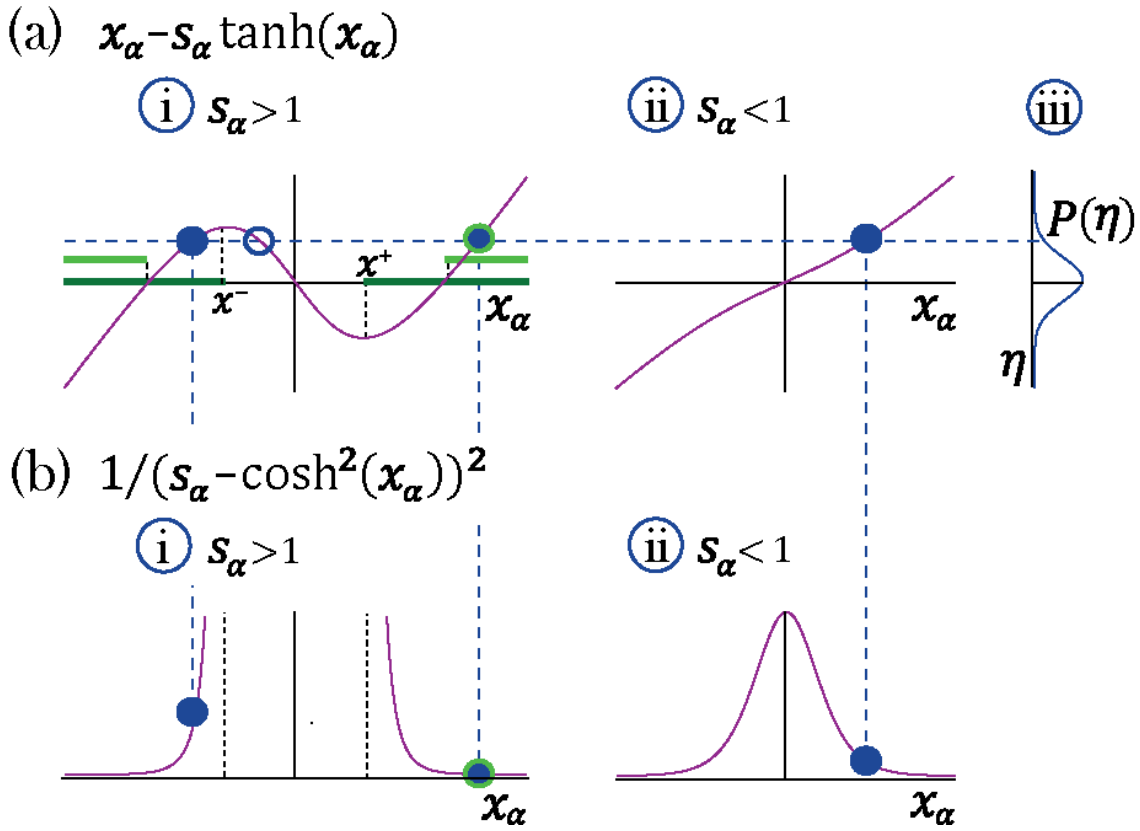

Figure 11. Transition to chaos with multiple self-couplings: Fixed point solutions and stability. a-i) The fixed point curve $x_\alpha - s_\alpha \tanh x_\alpha$, from Eq. 3.1, for $s_\alpha > 1$. Stable solutions are allowed within the dark green region. b-i) The shape of a unit's contribution to stability $q^{-1} = (s_\alpha - \cosh {x_\alpha}^2)^{-2}$, from Eq. 3.4. Stable solutions of $x_\alpha - s_\alpha \tanh x_\alpha = \eta$, filled blue circles in (a-i), with different $|x|$ values contribute differently to stability. At the edge of chaos only a fixed point configuration with all units contributing most to stability (minimal q^{-1}) is stable, light green region in (a-i). a-ii) The curve $x_\alpha - s_\alpha \tanh x_\alpha$ for $s_\alpha < 1$. a-iii) A possible distribution of the Gaussian mean-field η. A representative fixed point solution is illustrated by the dashed blue line: for $s_\alpha < 1$ a single solution exists for all values of η, (filled blue circle in a-ii);For $s_\alpha > 1$ multiple solutions exist (a-i) for some values of η; some of them lead to instability (empty blue circle in a-i). The other two solutions may lead to stability (filled blue circles in a-ii), although only one of them will remain stable at the edge of chaos (encircled with green line in a-i).

Since the denominator of the expression above is z plus the slope of the curve in Fig. 11a-i, a solution whose value x^*_α gives a negative slope (available when $s_\alpha > 1$) leads to a vanishing value of the denominator at some positive z and hence to a positive eigenvalue and instability. Hence, the n_α fraction of units with $s_\alpha > 1$ at a stable fixed point are restricted to have support on two disjoint intervals $[x^*_\alpha(s_\alpha) < x^-_\alpha(s_\alpha)] \cup [x^*_\alpha(s_\alpha) > x^+_\alpha(s_\alpha)]$. We refer to this regime as multi-modal, a direct generalization of the stable fixed points regime found in [13] for a single self-coupling $s > 1$, characterized by transient dynamics leading to an exponentially large number of stable fixed points. For the n_α portion of units with $s_\alpha < 1$, the stable fixed point is supported by a single interval around zero.

A fixed point solution becomes unstable as soon as an eigenvalue occurs at $z = 0$, obtaining from Eq. (3.3) the stability condition

$$g^2 \sum_{\alpha \in A} n_\alpha \langle q_\alpha^{-1} \rangle \leq 1 \ , \tag{3.4}$$

where $q_\alpha = \left[s_\alpha - \cosh^2(x_\alpha) \right]^2$. For $s_\alpha > 1$ the two possible consistent solutions to Eq. 3.1 that may result in a stable fixed point (from the two disjoint intervals in Fig. 11a-i), contribute differently to q_α. Larger $|x^*_\alpha|$ decreases q_α^{-1} (Fig. 11b-i), thus improving stability. Choices for distributions of x^*_α along the two intervals become more restricted as g increases or s_α decreases, since both render higher values for the stability condition, Eq. 3.4, forcing more solutions of x_i to decrease q_α^{-1}. This restricts a larger fraction of x^*_α at the fixed points to the one solution with higher absolute value. At the transition to chaos, a single last and most stable solution exists with all x_i values chosen with their higher absolute value x^*_α (Fig. 11a-i, light green segments). For those with $s_\alpha < 1$ only one solution is available, obtained by the distribution of

η through consistency (Eq. 3.1) at the fixed point. In this configuration, the most stable solution is exactly transitioning from stability to instability where Eq. 3.4 reaches unity. Hence the transition from stable fixed points to chaos occurs for a choice of g and $P(s)$ such that solving consistently Eq. 3.1 and 3.2 leads to saturate the stability condition (Eq. 3.4) at one.

We illustrate the discussion above in the case of a network with two sub-populations with n_1 and $n_2 = 1 - n_1$ portions of the units with self-couplings s_1 and s_2, respectively. In the (s_1, s_2) plane, this model gives rise to a phase portrait with a single chaotic region separating four disconnected stable fixed-point regions (Fig. 12a). A unit's activity is determined by its own self-coupling, the network's distribution of self-couplings and g. We will first discuss the stable fixed points, which present qualitatively different structures depending on the values of the self-couplings. When both self-couplings $s_1, s_2 < 1$, the only possibility for a stable fixed point is the trivial solution, with all $x_i = 0$ (Fig. 12a), where the network activity quickly decays to zero. When at least one self-coupling is greater than one, there are three stable fixed point regions (Fig. 12a-i, a-ii, a-iii); in these three regions, the network activity starting from random initial conditions unfolds via a long-lived transient periods, then it eventually settles into a stable fixed point. This transient activity with late fixed points is a generalization of the network phase found by M.Stern et.al. [13] When both self-couplings are greater than one ($s_1, s_2 > 1$) the fixed point distribution in each sub-population is bi-modal (Fig. 12a-ii,iii). When $s_1 > 1$ and $s_2 < 1$, the solutions for the respective sub-populations are localized around bi-modal fixed points and around zero, respectively (Fig. 12a-i). In the case of a Gaussian distribution of self-couplings in the stable fixed point regime, a complex landscape of stable fixed points emerges. The unit values at the stable fixed points continuously interpolates between the zero

(for units with $s_i < 1$) and the multiple values bi-modal cases (for units with $s_i > 1$) within the same network, Fig. 14a).

In the chaotic phase, we can estimate the intrinsic timescales τ_i of a cluster unit $x_i(t)$ from its autocorrelation function $C_\phi(\tau) = \langle \phi[x_i(t)]\phi[x_i(t+\tau)] \rangle$ as the width at half amplitude. For networks with two self-couplings, s_1 and s_2, we identified a parametric separation of timescales between the sub-populations. This separation quantified as the ratio τ_1/τ_2, is dependent on the network parameters g, s_1, and s_2, and also on the sub-population ratio of the clusters sharing the same self coupling n_1/n_2 ($n_\alpha = N_\alpha/N$). A visual summary of the parametric separation of timescales in such a network can be seen in Fig. 13. The observed general behavior is that in a network with similar n_1 and n_2 sub-populations, the timescales τ_1 and τ_2 increase concomitantly for large values of the self-couplings, meanwhile their ratio τ_2/τ_1 increases fast at first and then starts to plateau with larger values of the self-couplings. The story changes when we consider a small sub-population n_2 with a large self-coupling s_2 interacting with a large sub-population n_1 with a small self-coupling s_1. In this case the timescale τ_1, for the large sub-population, remains relatively small, while the timescale τ_2, of the small sub-population, increases supra-linearly with the increase of the self-coupling s_2. We examine the reason of switch in the behavior of timescales the next section.

In this section we explored models of recurrent neural networks with heterogeneous self-couplings. The dynamics of such networks describes a complex parameters space of phase transformations. Limiting our analysis to the chaotic regime, we have shown that the mechanism of creating multiple timescales in a network requires the interaction between sub-populations of different size clusters, i.e. heterogeneous self couplings. The separation of timescales in such networks

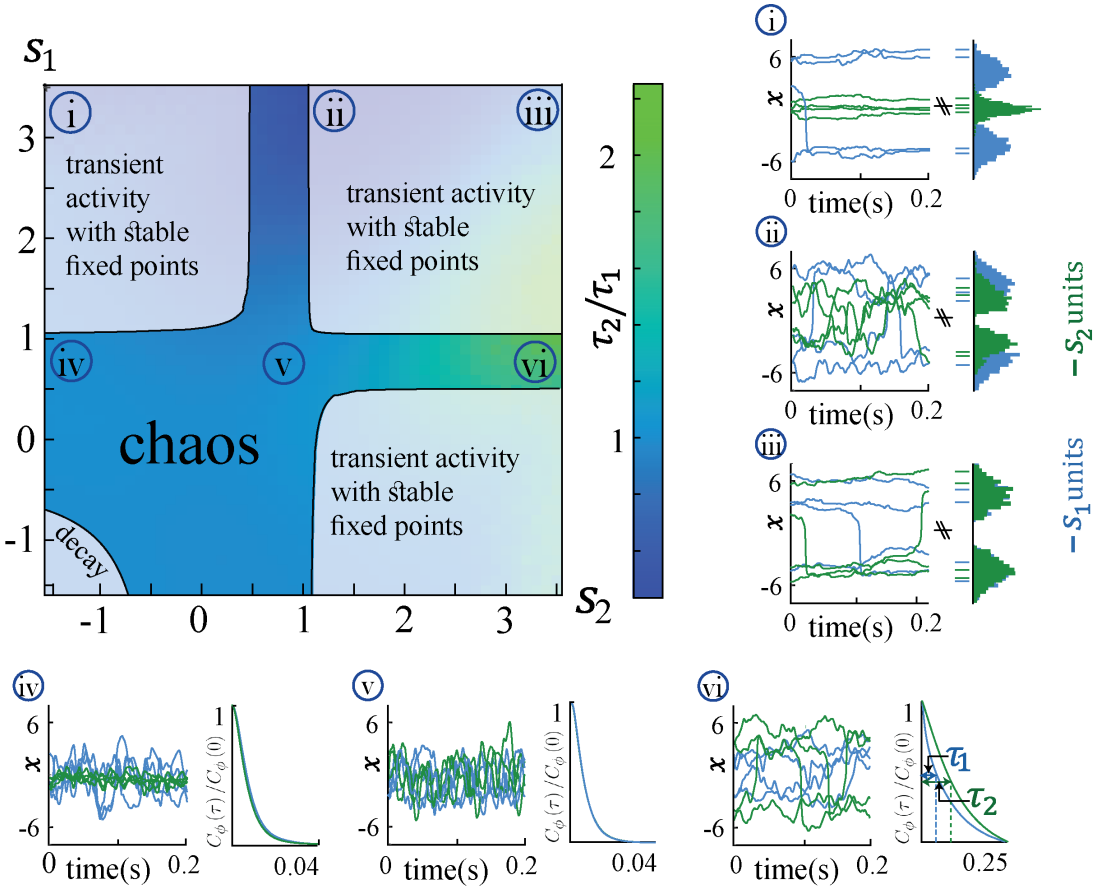

Figure 12. Ratio of autocorrelation timescales τ_2/τ_1 of units with self-couplings s_2 and s_1, respectively (τ_i is estimated as the half width at half max of a unit's autocorrelation function, see panels iv) - vi), in a network with $n_1 = n_2 = 0.5$ and $g = 2$ and varying s_1, s_2. A central chaotic phase separates four different stable fixed point regions with or without transient activity. Black curves represent the transition from chaotic to stable fixed point regimes. i)-iii) Activity across time during the initial transient epoch (left) and distributions of unit values at their stable fixed points (right), for networks with $N = 1000$ and (i) $s_1 = 3.2, s_2 = -1.5$, (ii) $s_1 = 3.2, s_2 = 1.2$, (iii) $s_1 = 3.2, s_2 = 3.2$. iv) - vi) Activity across time (left) and normalized autocorrelation functions $C(\tau)/C(0)$, (right) of units with (iv) $s_1 = 0.8, s_2 = -1.5$, (v) $s_1 = 0.8, s_2 = 0.8$, (vi) $s_1 = 0.8, s_2 = 3.2$.

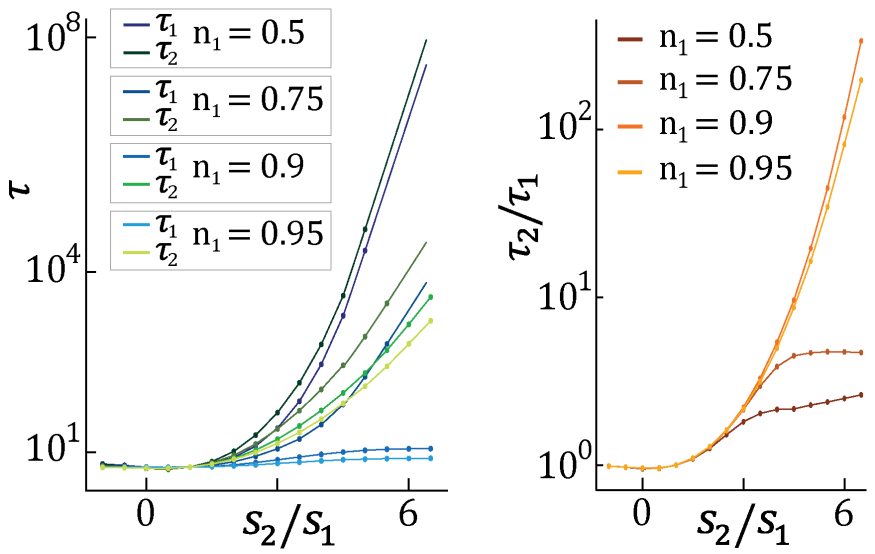

Figure 13. Timescales τ_2, τ_1 (left) and their ratio τ_2/τ_1 (right) for fixed $s_1 = 1$ and varying s_2, as a function of the relative size of the two populations $n_1 = N_1/N, n_2 = N_2/N$ (at $g = 2$, $N = 2000$; average over 20 network realizations.)

is dependent on the network parameters and, with the right tuning, can span over several orders of magnitude.

3.3 A Reservoir of Timescales

In the chaotic phase we can estimate the intrinsic timescale τ_i of a unit x_i from its autocorrelation function $C_\phi(\tau) = \langle \phi[x_i(t)]\phi[x_i(t+\tau)]\rangle_t$ as the half width at its autocorrelation half maximum (Fig. 12a-vi, τ_1 and τ_2). The chaotic phase in the network is characterized by a large range of timescales that can be simultaneously realized across the units with different self-couplings.

In a network with two self-couplings s_1 and s_2 in the chaotic regime, we found that the ratio of the timescales τ_2/τ_1 increases as we increase the self-couplings ratio s_2/s_1 (Fig. 12b). The separation of timescales depends on the relative fraction n_2/n_1 of the slow and fast populations. When this fraction drops below 10%, the log of the

timescale ratio exhibits a supralinear dependence on the self-coupling ratio, leading to a vast separation of timescales.

In a case of a lognormal distribution of self-couplings, in the chaotic regime the network generates a reservoir of multiple timescale τ_i of chaotic activity across network units, spanning across several orders of magnitude (Fig. 14b). For long tailed distributions such as the lognormal, mean field theory can generate predictions for rare units with large self-couplings from the tail end of the distribution by solving Eq. 2.16 and the continuous version of Eq. 2.13 highlighting the exponential relation between a unit's self-coupling and its autocorrelation decay time.

We do not have to limit our analysis to the case of a network with discrete self couplings and can also consider the case when s_i are extracted from a continuous distribution. To see the temporal capabilities of our network, we considered the case of a long tailed distribution of self-couplings. In the chaotic regime, such a network is able to generate multiple timescales τ_i spanning across multiple orders of magnitude as a function of the clusters' self-couplings s_i. The span of timescales present in a chaotic RNN with self-couplings extracted from a lognormal distribution are shown in Fig. 14 b).

3.4 Timescale Separation in the Bistable Chaotic Regime

As previously mentioned, the ratio and magnitudes of the timescales present in a network with two self couplings depends on network parameters like the gain g, the magnitude and ratio of self couplings s_1 and s_2, and also on the fraction of the numbers of clusters in each subpopulation, $n_1 = N_1/N$ and $n_2 = N_2/N$. In this section we will focus on the latter with the goal of gaining an analytical understanding of parametric separation of timescales seen in Fig. 13 b). For this analysis, we consider the special case of a network from Eq. 2.3 where a large subpopulation of clusters, $N_1 = N - 1$,

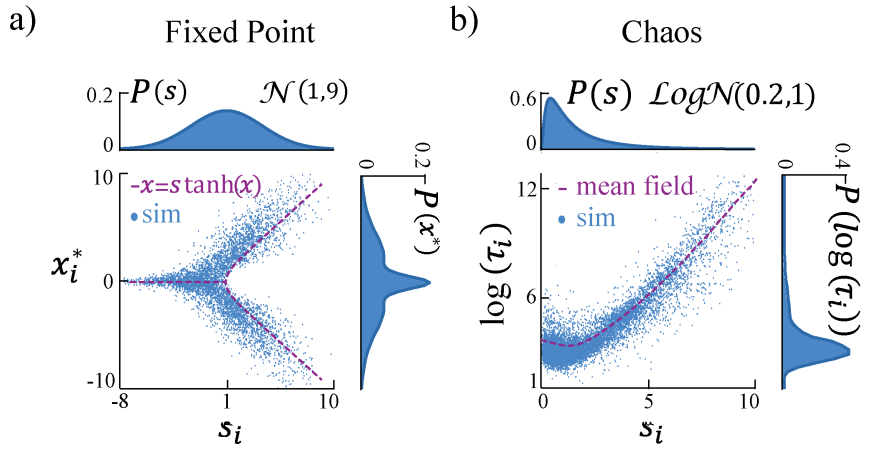

Figure 14. a) In a network with a Gaussian distribution of self-couplings (mean $\mu = 1$ and variance $\sigma^2 = 9$), and $g = 2.5$, the stable fixed point regime exhibits a distribution of fixed point values interpolating between around the zero fixed point (for units with $s_i \leq 1$) and the multi-modal case (for units with $s_i > 1$). The purple curve represents solutions to $x = s\tanh(x)$. b) In a network with a lognormal distribution of self-couplings (parameters $\mu = 0.2$ and $\sigma^2 = 1$), and $g = 2.5$, autocorrelation timescales τ_i in the chaotic phase span several orders of magnitude as functions of the units' self-couplings s_i (purple curve shows the dynamic mean-field predictions for τ_i).

with $s_1 = 1.0$ containing all but one slow cluster unit, x_2, described by a large self coupling $s_2 \gg s_1$. A illustration of this network is depicted in Fig. 15 a).

In the large N limit the behavior of the probe unit is described by the dynamical equation

$$\frac{d}{dt}x_2(t) = -x_2(t) + s_2\phi\left[x_2(t)\right] + \eta(t), \tag{3.5}$$

where $\eta(t)$ is interpreted as a source of external colored noise with autocorrelation relation

$$\begin{aligned}\langle \eta(t)\eta(t+\tau)\rangle &= g^2 C(\tau) \\ &\approx g^2 \left\langle \phi\left[x_1(t)\right]\phi\left[x_1(t+\tau)\right]\right\rangle .\end{aligned}$$

In the large N limit, the contribution to the mean field term from the probe unit is negligible, hence not included in the expression for the autocorrelation function. The

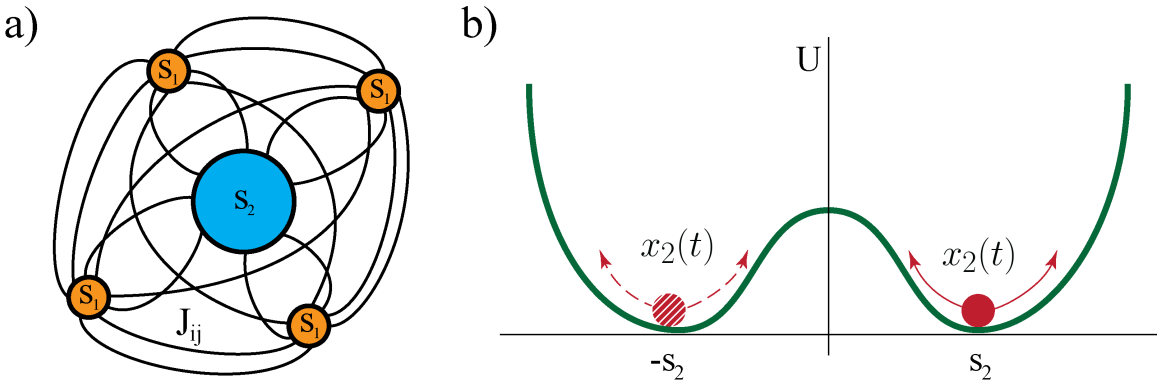

Figure 15. a) A graphical representation of an RNN with a single slow cluster s_2 (probe unit) interacting with multiple fast clusters (s_1 , $N_1 = N-1$). b) A graphical representation of the dynamics of the probe unit in a double well potential governed by external noise.

noise term η is described by its strength $D = \int_0^\infty d\tau C(\tau)$ and a timescale τ_1, which is also refereed to as the color.

In the case of a large cluster, $s_2 \gg s_1$, the dynamics of the probe units $x_2(t)$ is described by a bi-stable chaotic phase where its temporal activity is localized around two critical points. An analogy to this situation would be an object located in a double-potential well whose movements are governed by external noise, as illustrated in Fig. 15 b). The critical points are located at $x_2 = x^{\pm} \approx \pm s_2$. The probe unit switches between the critical points at random times. The interval of time the probe unit spends in on of the arms of the potential well is known as the dwell time, T. The behavior of the probe cluster and the distribution of dwell times T can be seen in Fig. 16 a-i,iii). For comparison, we also show the results for probe unit driven by external white noise as well b-i,iii).

Given the large timescale $\tau_1 > 1$ ($\tau_1 \approx 7.9$ for $s_1 = 1$ and $s_2 = 5$), the behavior of the probe unit falls under the strong colored noise regime. The stationary probability distribution $p(x)$, Fig. 16 a-ii,b), satisfies the the condition for using the colored noise

approximation to the Fokker Plank equation [50, 51]. We provide a refresher on the Fokker Plank formalism in Appendix A.

$$p(x) = Z^{-1}|h(x)| \exp\left[-U_{eff}(x)/D\right] \ , \tag{3.6}$$

where Z is a normalization constant, $h(x) \equiv 1 - \tau_1 f'(x)$, and the effective potential $U_{eff}(x) = -\int^x f(y)h(y)dy$ is given by

$$U_{eff} = \frac{x^2}{2} - s_2 \log\cosh(x) + \frac{\tau_1}{2} f(x)^2 - U_{min} \ . \tag{3.7}$$

The distribution $p(x)$ has support in the region $h(x) > 0$ comprising two disjoint intervals $|x| > x_c$ where $\tanh(x_c)^2 = 1 - \frac{1+\tau_1}{\tau_1 s_2}$ (Fig. 16b). $p(x)$ is concentrated around the two minima $x^{\pm} \simeq \pm s_2$ of U_{eff}. The main effect of the strong color $\tau_1 \gg 1$ is to sharply decrease the variance of the distribution around the minima $x^{\pm}$. This is evident from comparing the colored noise with a white noise, when the latter is driving the same bi-stable probe $dx_2/dt = -x_2 + s_2\phi(x_2) + \xi(t)$, where $\xi(t)$ is a white noise with an equivalent strength to the colored noise, Fig. 16a-iv,v,vi. The naive potential for the white noise case $U = x^2/2 - s_2 \log\cosh(x)$ is obtained from Eq. 3.6 by sending $\tau_1 \to 0$ in the prefactor h and in potential U_{eff}. It results in wider activity distribution compared to our network generated colored noise, in agreement with the simulations, Fig. 16a,b.

In our network generated colored noise the probe unit's temporal dynamics is captured by the mean first passage time $\langle T \rangle$ for the escape out of the potential well:

$$\begin{aligned} \langle T \rangle &= \int_{-s_2}^{-x_c} \frac{dx}{D} \frac{h(x)^2}{p(x)} \int_{-\infty}^{x} p(y)dy \\ &\simeq 2\pi \sqrt{U''_{eff}(x_-)\rho''(x_f)} \exp\left(\frac{\Delta}{D}\right) \ , \end{aligned} \tag{3.8}$$

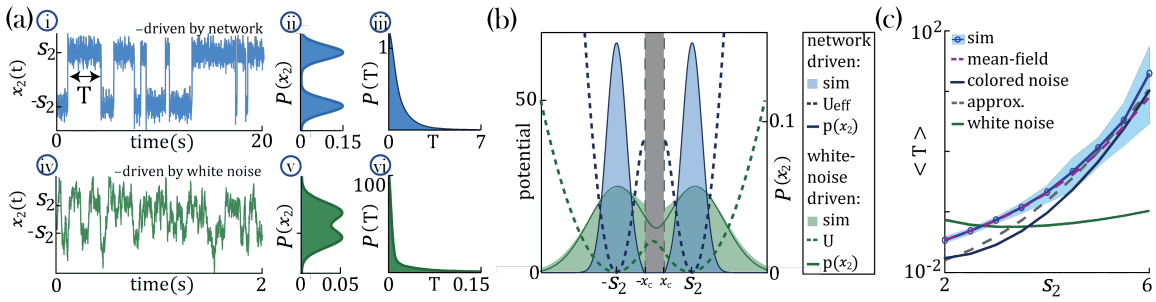

Figure 16. Separation of timescales and metastable regime. (a) Examples of bistable activity. i,iv - time courses; ii,v - histograms of unit's value across time; iii,vi - histograms of dwell times. (a-i,ii,iii) An example of a probe unit x_2 with $s_2 = 5$, embedded in a neural network with $N = 1000$ units, $N_1 = N - 1$ units with $s_1 = 1$ and $g = 1.5$. (a-iv,v,vi) An example of a probe unit driven by white noise. (b) The unified colored noise approximation stationary probability distribution $p(x_2)$ (dark blue curve, Eq. 3.6, its support excludes the shaded gray area) from the effective potential U_{eff} (dashed blue curve) captures well the activity histogram (same as (a-ii)); whereas the white noise distribution $p(x_2)$ (dark green curve, obtained from the naive potential U, dashed green curve) captures the probe unit's activity (same as (a-v)) when driven by white noise, and deviates significantly from the activity distribution when the probe is embedded in our network. (c) Average dwell times,$\langle T \rangle$, in the bistable states. Simulation results, mean and 95% CI (blue curve and light blue background, respectively; An example for the full distribution of T is given in (a-iii)). Mean-field prediction (purple curve). The mean first passage time from the unified colored noise approximation (Eq. 3.8, black curve) captures well the simulated dwell times. An approximation for the unified colored noise (Eq. 3.9, gray dashed line) achieves good results as well. the white noise average dwell times are significantly different (green curve).

where $\Delta = \rho(x_f) - U_{eff}(x_-)$ and $\rho = U_{eff} + D \log h$ (Appendix C). We evaluated the integrals by steepest descent around x^- and $-x_f$, where $\tanh(x_f)^2 \simeq 1 - 1/2s_2$. The agreement between Eq, 3.8 and simulation results improves with increasing s_2, as expected on theoretical ground [52, 53]. The asymptotic scaling for large s_2 is

$$\log\left[\langle T \rangle\right] \sim \frac{\tau_1 + 1}{2D}\left(s_2^2 - s_2 \log[s_2]\right). \tag{3.9}$$

The agreement between the analytical result in Eq. 3.9 and simulation are presented in Fig. 16 c).

Given a large cluster ($s_2 \gg s_1$), the dynamics in the network generate a parametric separation of timescales between the sub-populations x_1 with intrinsic timescale τ_1 and the probe cluster x_2 whose dynamics is characterized by two separate timescales, a large one, T, which is the average dwell time in one of the arms of the potential, and a small one τ_1 which describes the fast fluctuations around the metastable states $x^{\pm}$. This analytical method can be generalized to a network with $N - p$ clusters with self-coupling s_1 and $p \ll N$ probe clusters with self-coupling $s_2 \gg s_1$. This analysis explains the change in the individual timescales behavior seen in Fig. 13 in the case of larger values of n_1. The slow dynamics of each probe unit x_α is captured by its own bistable switching time T_α in Eq. 3.8 and all slow units are driven by a shared external colored noise $\eta(t)$ with timescale τ_1. In summary, in our model multiple timescales can be robustly generated with specific values, varying over several orders of magnitude. [54].

Is the relationship between the unit's self-coupling and its timescale relying on single-unit properties, or does it rely on network effects? To answer this question, we compare the dynamics of a unit when driven by a white noise input vs. the self-consistent input generated by the rest of the recurrent network (i.e., the mean field). If the neural mechanism underlying the timescale separation was a property of the single-cell itself, we would observe the same effect regardless of the details of the input noise. We found that the increase in the unit's timescale as a function of s_2 is absent when driving the unit with white noise, and it only emerges when the unit is driven by the self-consistent mean field. We thus concluded that this neural mechanism is not an intrinsic property of a single unit but requires the unit to be part of a recurrently connected network.

In summary, we have presented a category of recurrent neural network models with self-coupling terms capable of chaotic activity. The network is made out of clusters units, where each cluster represents a conglomeration of neurons. The size of the cluster is quantified through the self-coupling term s_i. A network made out of homogeneous clusters, $s_i = s$, manifests a single timescales. The sufficient condition to generate multiple timescales within the network is to consider heterogeneous clusters, i.e. clusters of different sizes s_i. In the case of an RNN with heterogeneous clusters, the separation of timescales generated by the network depends on the magnitude of the self-coupling terms and also on the ratio of same size clusters, n_i, present in the network. Given the right set of parameters, our network is able to generate multiple timescales ranging over multiple orders of magnitude, as observed in experimental literature.

CHAPTER IV

RNN NETWORK WITH INPUT

A different approach to study neural systems is considering their responses to different external inputs. The input-output relationship can reveal important characteristics of the neural system. Such an approach is very common across disciplines like information theory, physics, engineering, and neuroscience.

In this section we will study the effect of external input on the ongoing intrinsic dynamics of recurrent neural network. A useful approach in studying network responses to oscillatory input has been presented by K.Rajan et. al. [14]. We will explicitly present this approach here, and then apply it to our case.

4.1 RNN with Sinusoidal Input

We introduce this analysis method by first considering external input $H_i(t)$ to individual clusters in a recurrent neural network (Eq. 2.3) with no self-couplings, $s_i = 0$. The external input to each cluster is a sinusoidal function with shared amplitude I and frequency f across all clusters, and individual phase γ_i for each cluster.

$$\begin{aligned} \frac{d}{dt}x_i(t) &= -x_i(t) + \sum_j J_{ij}\phi\left[x_j(t)\right] + H_i(t) \\ H_i(t) &= I\cos\left[2\pi f t + \gamma_i\right] \end{aligned} \tag{4.1}$$

The phase γ_i is extracted randomly from a uniform distribution in the $[0, 2\pi]$ range. The individual phase for each cluster assures that the sinusoidal input does not introduce global temporal coherence. This corresponds to a situation where the population of clusters favors a wide range of spatiotemporal phases [14]. In the

original work, the authors use a slightly different transfer function:

$$\phi\left[x\right] = \begin{cases} r_0 \tanh\left[\frac{x}{r_0}\right] & \text{, for } x \leq 0 \\ \\ (2 - r_0)\tanh\left[\frac{x}{2-r_0}\right] & \text{, for } x > 0 \end{cases} .$$

This choice of transfer function is more biologically plausible compared to the more common choice of $\tanh[\cdot]$. Notice that when the parameter $r_0 = 1$ we recover our transfer function $\phi[x] = \tanh[x]$. We would like to remind the reader that while negative firing rates are problematic, we can avoid this issue by consider the firing rate from our transfer function as relative to some background r_b:

$$\frac{d}{dt}x_i(t) = -x_i(t) + \sum_j J_{ij}\Big(\phi\left[x_j(t)\right] + r_b\Big).$$

Hence, the different choices for the transfer function in this work does not affect the analysis method nor the main results.

In order to make a biological equivalence to the spontaneous and ongoing intrinsic activity recorded in cortical and subcortical areas of the brain, we restrict our analysis to the the regime of chaotic activity in the recurrent neural network (i.e. $g > 1$). To understand the effects of the sinusoidal input on the chaotic activity we can analyze the average autocorrelation function of firing rate across all clusters:

$$C_\phi(\tau) = \frac{1}{N}\sum_{i=1}^{N}\left\langle \phi\left[x_i(t)\right]\phi\left[x_i(t+\tau)\right]\right\rangle .$$

As discussed in chapter 2, we can gain a thorough understating of the network dynamics by applying the dynamic mean field theory and considering the large N limit of clusters in the network. In this limit, the total recurrent input from all

the other clusters, mathematically represented as $\sum_j J_{ij}\phi\left[x_j(t)\right]$, is approximated as Gaussian noise $\eta(t)$. The self-consistent conditions for this approximation are determined by matching the first two moments of the mean-field variable $\eta(t)$ and the network firing rate $\phi\left[x_i(t)\right]$:

$$\langle \eta(t) \rangle = \langle \phi\left[x_i(t)\right] \rangle = 0$$
$$\langle \eta(t)\eta(t+\tau) \rangle = g^2 C_\phi(\tau)$$

An analytical expression for the average auto-correlation function in the large N limit can be determined by treating the intrinsic chaotic dynamics and the oscillatory dynamics (due to the input stimulus) independently. We write that $x_i(t) = x_i^{chaos}(t) + x_i^{osc}(t)$, where $x_i^{chaos}(t)$ satisfies the set of coupled differential equations describing the intrinsic chaotic dynamics without input:

$$\frac{d}{dt}x_i^{chaos}(t) = -x_i^{chaos}(t) + \sum_j J_{ij}\phi\left[x_j^{chaos}(t) + x_j^{osc}(t)\right],$$

while $x_i^{osc}(t)$ is a steady state solution that takes account of the sinusoidal input into the network:

$$\frac{d}{dt}x_i^{osc}(t) = -x_i^{osc}(t) + I\cos\left[2\pi f t + \gamma_i\right].$$

Such an approach allows to explicitly express the average autocorrelation function as a contribution from the intrinsic chaotic component and the oscillatory stimulus

component:

$$C_\phi(\tau) = \int_0^{2\pi} \frac{d\gamma}{2\pi} \int \mathcal{D}z_3 \int Dz_1 \phi \Bigg[\sqrt{C_x(0) - |C_x(\tau)|}\, z_1 + \operatorname{sgn}\left(C_x(\tau)\right) \sqrt{|C_x(\tau)|}\, z_3 \frac{I}{\sqrt{1+(2\pi f)^2}} \cos[\gamma] \Bigg] \times \int \mathcal{D}z_2 \phi \Bigg[\sqrt{C_x(0) - |C_x(\tau)|}\, z_2 + \sqrt{|C_x(\tau)|}\, z_3 + \frac{I}{\sqrt{1+(2\pi f)^2}} \cos\left[2\pi f t + \gamma\right] \Bigg],$$

where z_1, z_2, and z_3 are Gaussian random variables with mean zero and unit variance, while $\mathcal{D}z_i = dz_i \exp[-z_i^2/2]/\sqrt{2\pi}$. For more details see K.Rajan et. al. [14]

The interplay between the intrinsic chaotic activity and the oscillatory external input seen in the analytical expression for the average autocorrelation function is also observed in the network dynamics. Depending on the strength of the external input, the network dynamics undergoes a transition from chaos to complete entrainment.

We are already familiar with the chaotic state for the network with no input ($I = 0$, illustrated in Fig 17 a)). The firing rate fluctuations are irregular and the autocorrelation function decays from its initial value $C_\phi(0)$ to zero. This interplay manifests in the average power spectrum density as well, which in the chaotic regime, exponentially decays as a function of increasing frequency and does not show any spikes. The suppression of power density for large frequency fluctuations is a characteristic of a chaotic behavior [14].

In the case of a sufficiently strong input (large $I \geq I_c$, illustrated in Fig 17 c)), the firing rate of the clusters oscillates at the frequency of the external input. These oscillations are also present in the autocorrelation function. In this regime, the amplitudes of the oscillations in the autocorrelation function are equal and match its

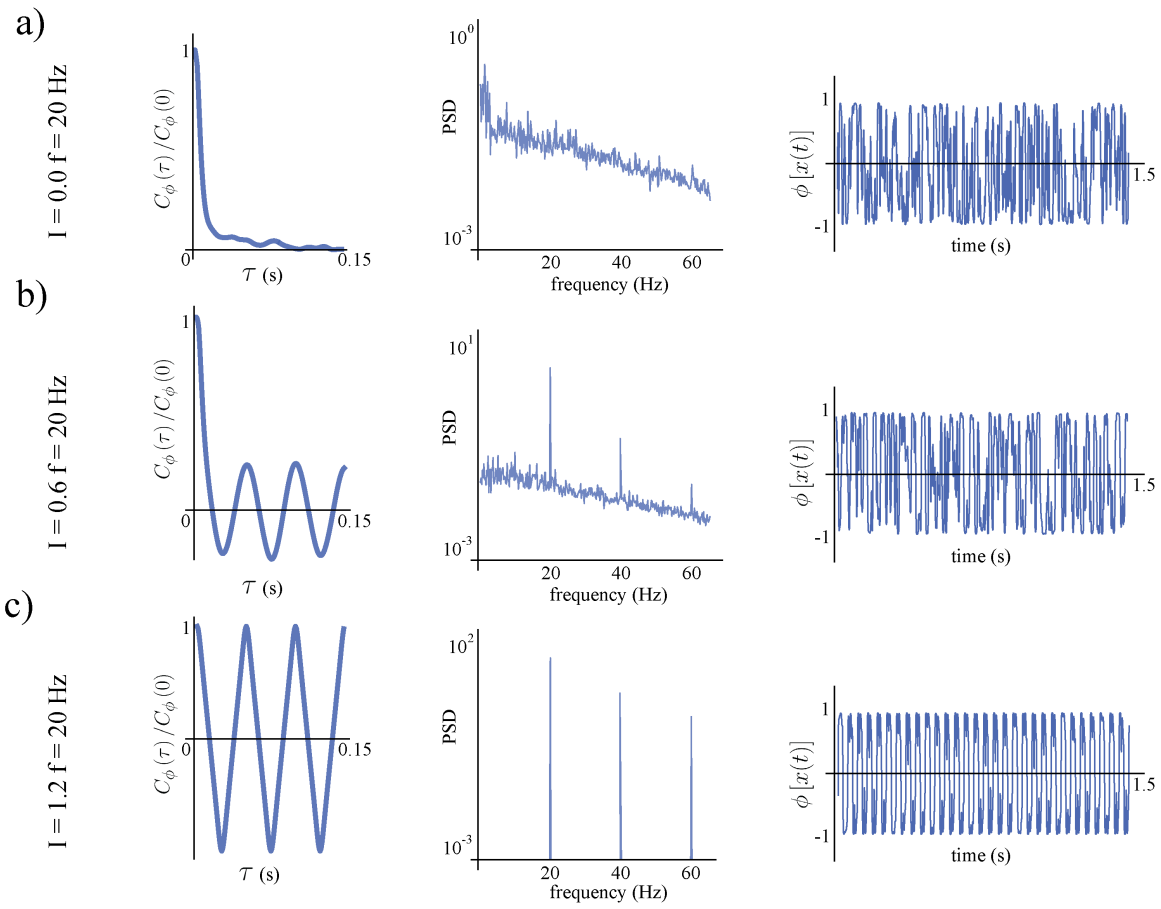

Figure 17. Network dynamics and transition from chaos to driven activity as a function of external signal amplitude I. a) Chaotic RNN with no input ($I = 0$). b) Interplay between chaos and signal in an RNN with sinusoidal input ($I = 0.6$, $f = 20$ Hz). c) Driven(no chaos) RNN with sinusoidal input ($I = 1.2$, $f = 20$ Hz). In all cases (a,b,c), from left to right: the normalized average autocorrelation function, power spectrum density, and the firing rate dynamics.

initial value $C_\phi(0)$. This indicates that the external stimulus dominates the network by fully suppressing the intrinsic chaos. This new phase in the network is seen in the power spectrum density of the firing rate dynamics. In the driven regime, the network power is focused only at the frequency f of the input signal and its following harmonics. The broad decaying power spectrum seen in the chaotic regime is completely gone.

When the network is driven by a weaker signal ($0 < I < I_c$, illustrated in Fig 17 b)), we observe an convoluted dance between the two phases, an interplay between chaos and driven activity. The firing rate shows periodic fluctuations on top of a chaotic background. This also manifests in the autocorrelation function, which oscillates at the frequency of the input signal at large time-lags τ, however the amplitude of these oscillations is smaller compared to the initial value of the autocorrelation function $C_\phi(0)$. This driven-chaos interplay is also present in the power spectrum, where we can identify the broad decay for large frequencies characteristic to chaos and also peaks at the input frequency f and its harmonics.

A useful way to quantify this transition is through the autocorrelation function. In the chaotic regime (with no external input), the initial amplitude of the averaged autocorrelation function $C_\phi(0)$ describes the total variance of the firing rates of all clusters in the network and the contribution from the external signal is null. In the driven regime (no chaos), the initial amplitude of the averaged autocorrelation function $C_\phi(0)$, along with the amplitude of all the subsequent oscillations, quantify the external sinusoidal signal introduced in the network. The contribution from the intrinsic chaos is null in this regime. In the intermediate regime, the initial amplitude of the averaged autocorrelation function $C_\phi(0)$ is a combination of the chaos present in the network and the external stimulus. Meanwhile at larger values

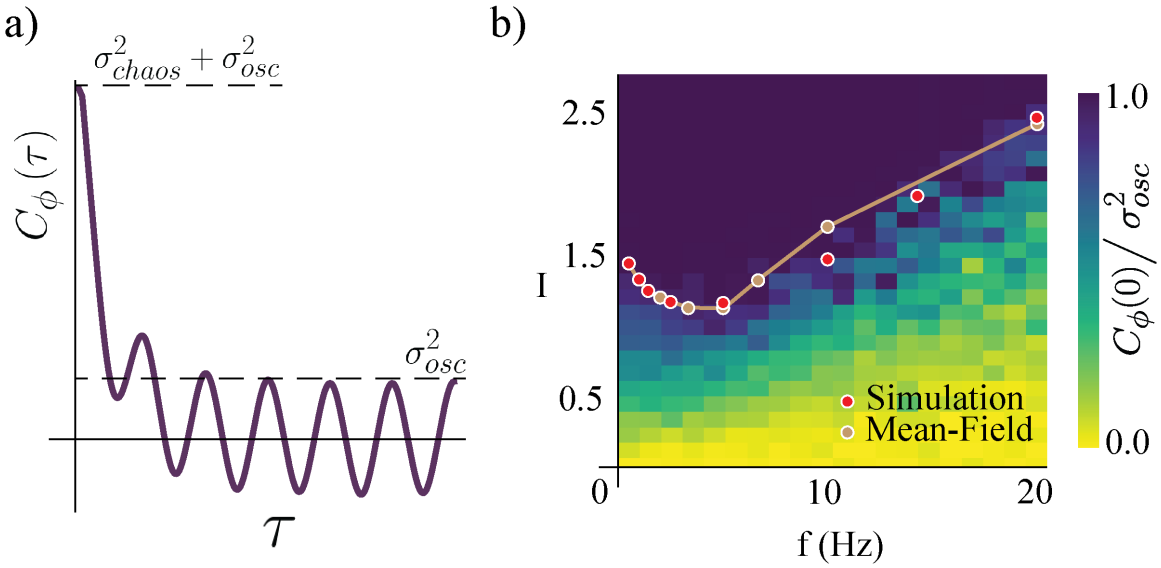

Figure 18. a) Average autocorrelation function of an RNN with sinusoidal input. The shape of the autocorrelation function is quantified as an interplay of the variance due to the intrinsic chaos σ^2_{chaos} and the variance due to the sinusoidal external input σ^2_{osc}. b) Phase-transition curves showing the critical input amplitude I_c that separates the chaotic regime from the driven regime as a function of frequency, for the following network parameters $g = 2.0$ and $N = 5000$. This figure was adapted from K.Rajan et.al. [14]

of the delay $C_\phi(\tau > 0)$, the contribution from the intrinsic chaos drops to zero leaving only the contributions from the external signal. Hence, we divide the network response into a chaotic and a oscillatory component. We write the initial amplitude of the averaged autocorrelation function as a sum of the variance due to the intrinsic chaos and the variance due to the sinusoidal input: $C_\phi(0) = \sigma^2_{chaos} + \sigma^2_{osc}$. The variance due to external signal σ^2_{osc} also equals to the amplitude of the oscillations at large delay $C_\phi(\tau > 0)$, as illustrated in Fig. 18 a). Under this representation, the complete entrancement of chaos happens when $C_\phi(0) = \sigma^2_{osc}$. This condition allows us to determine the minimal values of the network parameters I and f required to completely suppress the chaotic fluctuation in a network. The transition lines for a network with gain $g = 2.0$ is shown in Fig. 18 b).

4.2 RNN: Heterogeneous Self-Couplings with Sinusoidal Input

We apply the analysis method described in the previous section to a recurrent neural network with heterogeneous timescales. We consider a network of N clusters with two self couplings s_1 and s_2 and equal ratios of sub-populations $n_1 = n_2$. Each cluster receives the same external sinusoidal input with individual phase $H_i(t) = I\cos[2\pi ft + \gamma_i]$. As before, the phase γ_i for each cluster is random and extracted from a uniform distribution in the $[0, 2\pi]$ range. In the limit of a large number of clusters $N \gg 1$, the dynamics of the recurrent neural network can be described by a set of two coupled differential equations:

$$\frac{d}{dt}x_1(t) = -x_1(t) + s_1\phi[x_1(t)] + \eta(t) + I\cos[2\pi ft + \gamma_1]$$
$$\frac{d}{dt}x_2(t) = -x_2(t) + s_2\phi[x_2(t)] + \eta(t) + I\cos[2\pi ft + \gamma_2],$$

where $\eta(t)$ can be interpreted as Gaussian noise whose moments are matched in the mean field approximation, $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t+\tau)\rangle = g^2C_\phi(\tau)$. This indicates that each sub-population responds differently to the same external input, therefore it should have its own phase-transition curve. As before, we determine the critical amplitude and frequency pair for the transition in each sub-population, by comparing the peaks of the average autocorrelation function for large delays to its initial value. The critical pair (I_c, f_c) is quantified as the minimum amplitude and frequency at which the ratio between the peaks $C_\phi(0)/\sigma^2_{osc} = 1$. The phase-transition curve for larger values of s_i will occur at higher signal amplitudes. The phase transition curve along with the network dynamics, average autocorrelation function, and power spectrum density per sub-population are shown in Fig. 19.

By studying the phase-transition curve in the network with zero self-couplings, Fig. 18, we observe the presence of a resonant frequency at which the intrinsic chaotic

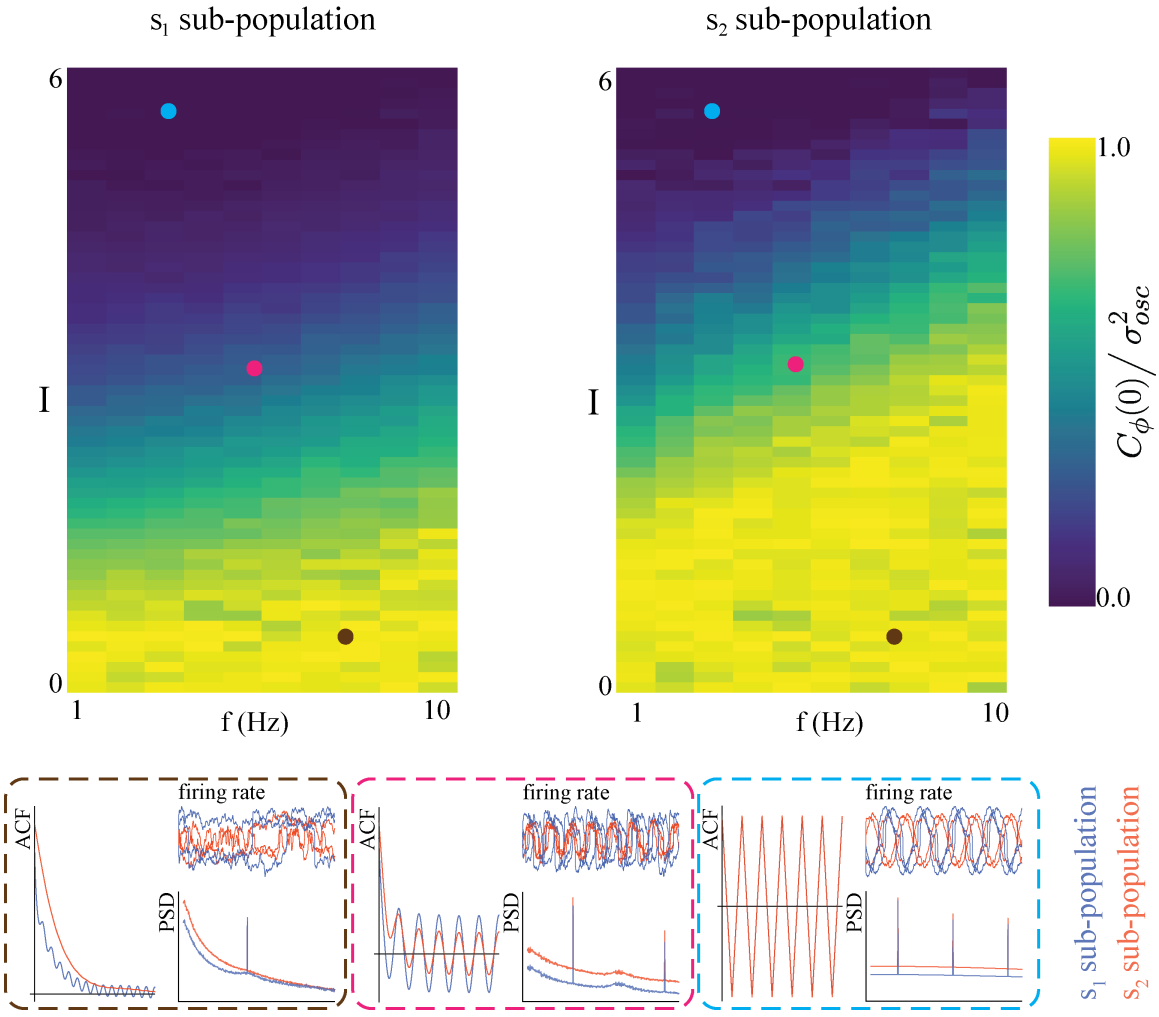

Figure 19. Phase transition from chaos to driven RNN with sinusoidal input per sub-population $s_1 = 1.0$, $s_2 = 4.0$. Below, from left to right we see the average autocorrelation functions, power spectrum density and firing rate for each sub-population (blue for s_1, and red for $_s2$). In the brown we have the RNN with 2 self-couplings driven by a sinusoidal signal with $I = 0.5$ and $f = 7$ Hz; pink box: sinusoidal signal with $I = 3.0$ and $f = 5$ Hz; cyan box: sinusoidal signal with $I = 5.5$ and $f = 3$ Hz.

activity is entrained by an external signal with the smallest amplitude. We could not detect such a resonant frequency in a network with two timesclaes.

4.3 Time-Dependent Broadband Input

We further investigated the response of our network with two different self-couplings $s_1 < s_2$ to a more complex external input. We drove the network with a broadband stimulus $H_i(t) = I \sum_{l=1}^{L} \cos[2\pi f_l t + \gamma_i]$, consisting of a superposition of L sinusoidal functions at different frequencies f_l in the $[0, 200]$ Hz range, random phases γ_i, and shared amplitude I.

We observed that each sub-population responds differently to certain ranges of frequencies. The sub-population with a slow intrinsic timescale, clusters with self-coupling s_2, prefer to concentrate more power around smaller frequencies of the broadband stimulus. Meanwhile, the sub-population with a fast intrinsic timescale, clusters with self-coupling s_1, prefer to concentrate more power around larger frequencies of the broadband stimulus. This behavior is observed by comparing the peaks in each sub-population at the stimulus frequencies of the power spectrum density in Fig 20 a).

We quantify this effect through a spectral modulation index $m(f)$. This modulation index is computed at every frequency of the broadband stimulus and represents the ratio between the difference of the power spectrum peak amplitude of each sub-population and their sum:

$$m(f) = \frac{P_2(f) - P_1(f)}{P_2(f) + P_1(f)}, \tag{4.2}$$

where $P_i(f)$ is the power spectrum amplitude at frequency f of the sub-population with self-coupling s_i. A positive value for $m(f)$ indicates a stronger entrainment at frequency f in the sub-population with self-coupling s_2 compared to the s_1 sub-

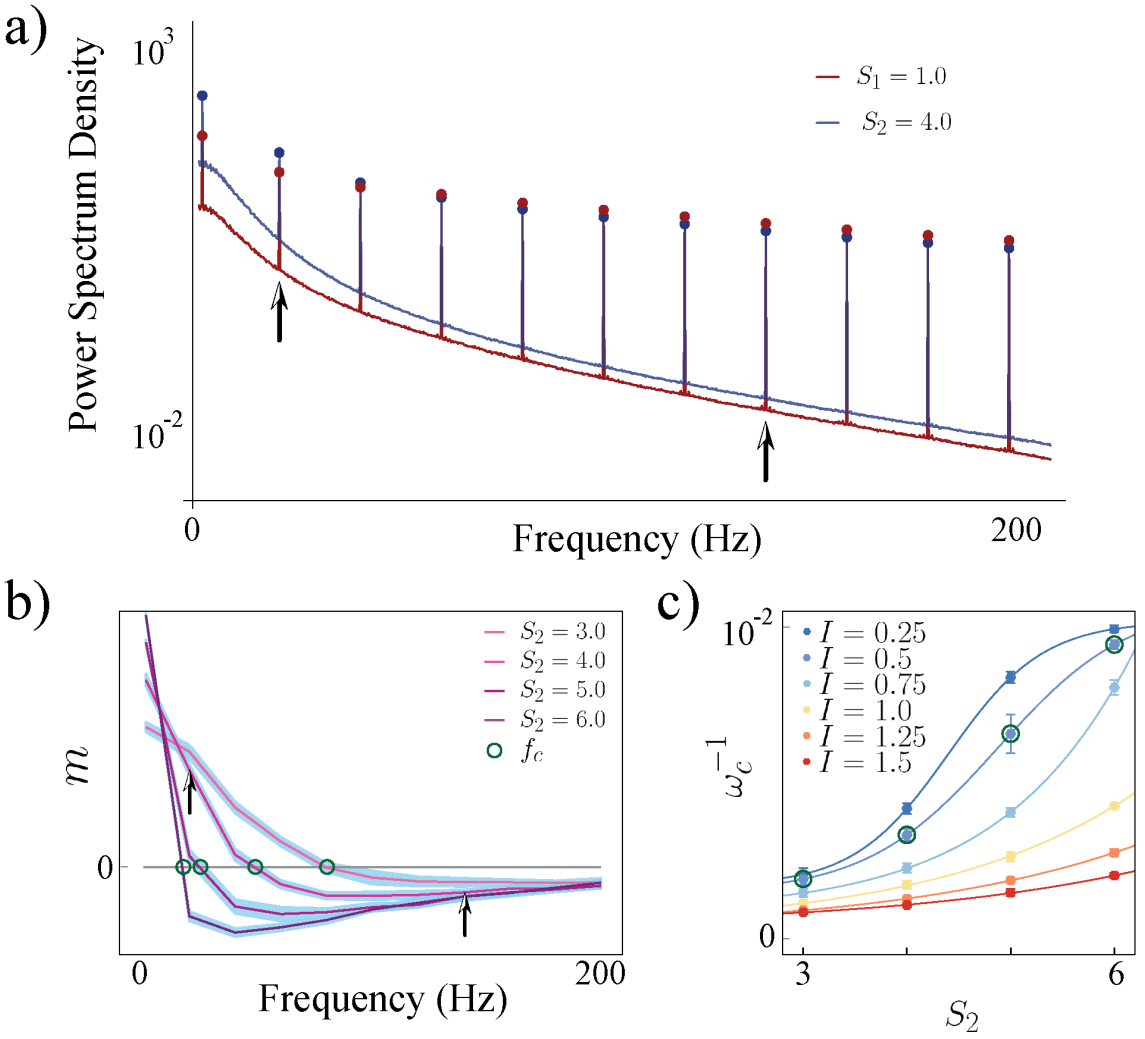

Figure 20. Network response to broadband input. (a) Power spectrum density of a network driven by time-dependent input comprising a superposition of 11 sinusoidal frequencies (see main text for details). Maroon and navy curves represent average power spectrum density in s_1 and s_2 sub-populations, respectively; circles indicate the peak in the power spectrum density amplitudes at each frequency; amplitude $I = 0.5$; $N_1 = N_2 = 1000$, $g = 3.0$, $s_1 = 1$, and $s_2 = 4$. (b) Modulation index, Eq. 4.2, of the power spectrum density amplitudes as a function of frequency in networks with $s_1 = 1$ and various s_2. The green circles mark the cutoff frequency f_c where the modulation index changes sign. (c) Cutoff period, $2\pi\omega_c^{-1}$, as a function of self coupling s_2 for different input amplitudes. An inversely proportional relation between the cut off period and the amplitude of the broadband signal is present.

population. A negative value of $m(f)$, indicates the opposite, a stronger entrainment at frequency f in the sub-population with self-coupling s_1 compared to the s_2 sub-population. The spectral modulation index for the frequency range of the input signal is shown in Fig 20 b).

We observe a consistent behavior of the spectral modulation units across the frequency range of the broadband input for different values of s_2. The crossover of $m(t)$ from positive to negative values, as the frequency increases, indicates that the low frequency components of the input entrained the slow sub-population s_2, while the fast components, large frequencies, of the input predominantly entrained the fast sub-population s_1. The frequency at which this crossover happens is called the crossover frequency f_c, from which we can determine the crossover period as $\omega_c^{-1} = 1/(2\pi f_c)$.

The dependence of the crossover period ω_c^{-1} on the self-coupling of the slow sub-population s_2 (the self-coupling of the fast sub-population was fixed to $s_1 = 1.0$) is strong at for external stimuli with low amplitude and it begins to weaken as the amplitude increases. This dependence relationship illustrated in Fig 20 c), is consistent with the suppression of chaos (the transition from a a chaotic to a driven network regime) for signals with larger amplitude as seen in the previous section.

CHAPTER V

SCHEMA THEORY AND MULTIPLE TIMESCALES

This chapter presents a potential path for the future of this project. Here, we introduce the framework behind Schema Theory in more details and also discuss how to computationally integrate this theoretical framework with our recurrent neural network with heterogeneous timescales.

5.1 Schema Theory

A lot of behaviors of living organisms can be structurally decomposed into a series of simple actions known as primitives. Primitives can be thought of as generalized actions segments that can be reused in different situations. Consider the simple action of grabbing an object, i.e. the grabbing primitive. This primitive can be used in different scenarios, independent of the object, or the sequence of actions that preceded or followed it. For example the grabbing primitive can be used when drinking some coffee (grabbing the cup), or for climbing a latter (grabbing the latter for support).

Schema theory is the concept that complex behaviors can be broken down into primitives. A schema is what is learned (or innately given) about some aspect of the world; it combines knowledge with the process of applying it [24]. Generally, schemas can be of two types: perceptual schemas and motor schemas. Perceptual schemas are structured internal representations of objects or images acquired through the sensory organs (ex: the coffee detection schema). Motor schemas, on the other hand, are easier to comprehend. They can be thought of as memory representations of adaptable parameters that define specific motor actions (ex: grabbing an object schema). Perceptual and motor schemas are not necessarily independent and can

affect each other, for example the coffee detection perceptual schema can activate grabbing an object motor schema.

It is important to make the distinction between the concept of a schema, which is an abstraction over a set of similar circumstances, and the concept of activating a schema, which is a neural computation that determines the similarities between the current situation and pre-existing concepts of similar circumstances and can also generate an appropriate response [24].

We have talked about schemas as an abstraction phenomena, and now we want to build a bridge from a theory to an implementation framework. This can be done by studying the relationship between schemas. In his studies of memory recollections, F.C.Bartlett [26] determined that in order to encode and identify a specific event within a schema-based structure it is necessary to find the similarities between a specific event and the pre-existing general schema, and also to identify how this specific event differs from the generalized version. Another point to consider is that schemas can be nested structures, meaning that more complex schemas are made out of simpler schemas, and so on, until we reach the simplest structures, primitives. Computationally, we can interpret a schema as a knowledge structure described by a set of relevant parameters. The values of these parameters determine the relationship between schemas, thus allowing for the inhibiton or activation of others schemas. To illustrate these concepts, consider the prey-predator detection schema. At the lowest level, there is an elementary features detection schema that expresses parameters like orientation, color, texture, margins. The value of these parameters can activate more higher-level schemas that can parametrically express more complex structures like body parts or movement. These new schemas can activate or inhibit even higher-level schemas. For example the identification of pray-like parameters activates the catch-

and-feed schema and inhibits a run-away response. Alternatively, the identification of predator-like parameters inhibits the catch-and-feed and activates the run-away response.

Such a relationship creates some computational requirements. First, schemas can work together by reinforcing each other, and secondly, schemas can be competitive as well. A predator can have different responses to a prey-detection schema depending on the type of prey. This last point requires a mechanism to ensure that at any time, the behavior is determined by only one schema [24].

Each schema, within the sequence of hierarchical schemas is associated with an activation function that can vary between zero and one. The value of the activation function at value $t+1$ depends on the inhibitory or excitatory inputs to the schema during the interval of time between timesteps t and $t+1$. A constant excitatory input will drive the activation value to one, meanwhile sustained inhibitory input will drive it to zero. Mathematically, we can determine the activation value A_α of a particular schema α at time $t+1$ as

$$A_\alpha(t+1) = \sigma\left[p\sigma^{-1}\left[A_\alpha(t)\right] + I_\alpha(t)\right],$$

where $\sigma[x] = 1/(1+\exp[-(x+b)])$ is the usual sigmoid function with bias b. The equation above can be generalized as a recurrence relationship:

$$A_\alpha(t+1) = \sigma\left[\sum_{i=0}^{t} p^i I_\alpha(t-i)\right].$$

The parameter $p \in [0,1]$ is known as the persistence and it describes the sensibility of the activation value to the the input $I_\alpha(t)$ relative to prior inputs. Low values of the persistence parameter favors the dominance of the input in determining $A_\alpha(t)$,

which manifests as fast fluctuations of the schema activation value. A more physical interpretation of p is as the parameter describing the temporal granularity of the system (i.e. the time interval in ms between timestep t and $t+1$). When p is close to 1, the time interval is small and and the system is continuous, while a p close to 0 describes a large time interval causing the activation value to be updated less often thus risking overshooting or undershooting the activation value at each timestep [55].

The term $I_\alpha(t)$ is the total input to the schema. It contains contributions from the surrounding schemas, but also can include more complicated contributions as top-down, bottom-up excitations, lateral inhibitions, self-activation, and noise. [56, 57]

While we can build a good understanding of the relationships between schemas, the underlying neural mechanism that governs the way in which the external stimuli are broken down into primitives and how these primitives are flexibly sequenced into higher-level schemas is currently unknown.

We propose a brain-inspired neural architecture mediated by temporal functional hierarchies (multiple timescales) able to flexible sequence primitives. In previous studies, a hierarchical structure was hardcoded into the modular structure of the computational model. The primitives represented lower modules, while the sequencing of tasks was implemented via higher modules within an ad-hoc architecture [58, 59]. An advantage of hardcoding a hierarchical structure during learning is that one module will not affect other modules. However, such an approach creates a conflict between the generalization of similar sequences and the segmentation of primitives. This generalization-segmentation conflict becomes evident when considering a set of complex behaviors that have an overlap of primitives. Consider the schema of making a cup of instant coffee with milk and sugar illustrated in Fig. 21. If we were to hardcode the sequence of lower-level schemas, then we would

get multiple, redundant, versions of the same schema: in one version one might first pour the milk, then the coffee, then the sugar; in another version one might pour the coffee, then the sugar, and then the milk. Hardcoding schemas doesn't offer a flexibility of sequencing primitives, is not a generalize method and thus requires lots of memory space for storing every possible primitive.

As an alternative to hardcoding the schema architecture and solving the generalization-segmentation conflict, we propose the use of the recurrent neural network with mutiple timescales presented in the previous chapters. The smaller clusters with short timescales will engage the primitives and low-level schemas meanwhile the larger clusters with longer timescales will have the role of sequencing primitives into more complex tasks, or higher-order schemas. We have illustrated this concept for the preparing a coffee schema in Fig. 21. Ours is a novel approach since we do not hardcode this functional hierarchy into the network. Instead, it is an emergent property of the network's dynamics and learning.

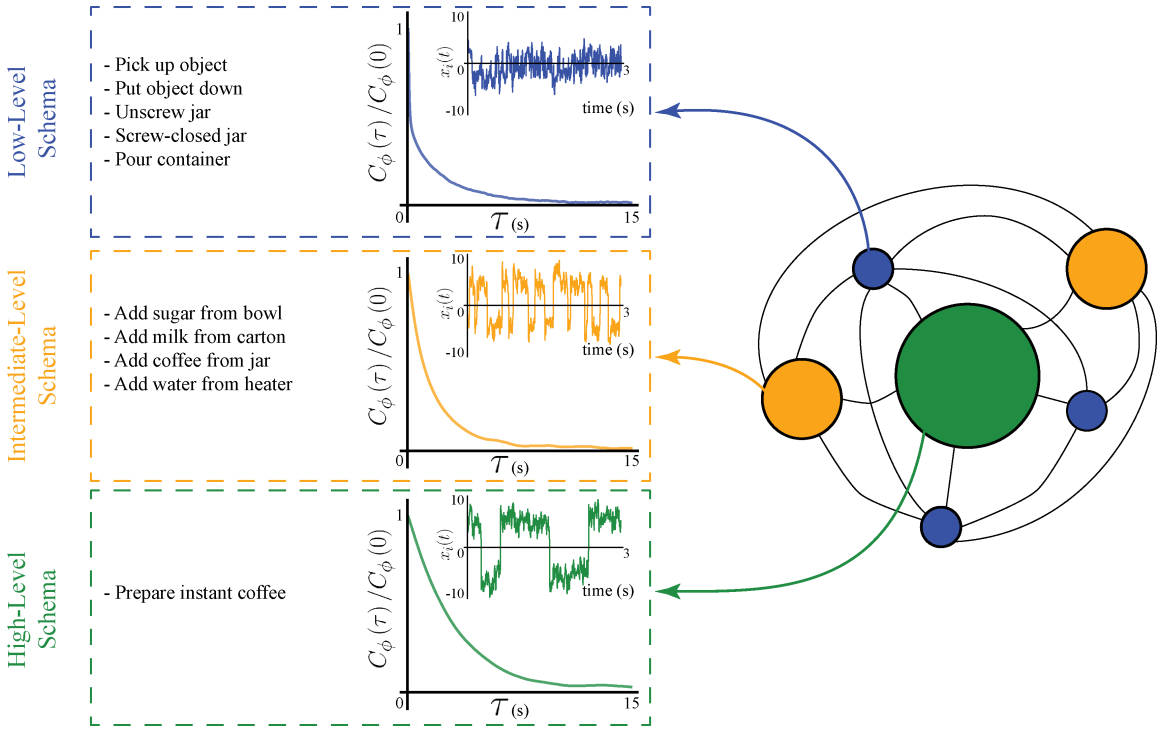

Figure 21. The complex behavior of preparing a cup of instannt coffee with milk and sugar can be broken into serveral schemas, ranging from low- to high-level schemas. The low-level schemas are governed by the fast units (short timescale clusters s_1) in our RNN. The intermediate-level schemas are governed by the slower units (intermediate timescale clusters s_2) in our RNN. Meanwhile the slowest units (longest timescale clusters s_3) keeps track of the highest-level schema and its completion.

CHAPTER VI

CONCLUSION

We demonstrated a new robust and biologically plausible network mechanism whereby multiple timescales emerge across units with heterogeneous self-couplings. In our model, units are interpreted as neural clusters, or functional assemblies, consistent with experimental evidence from cortical circuits [60, 61, 62, 63, 64] and theoretical modeling [65, 66]. We found that the neural mechanism underlying the large range of timescales is the heterogeneity in the distribution of self-couplings (representing neural cluster strengths).

A heterogeneous distribution of cluster sizes, in turn, generates a reservoir of timescales.

Several experimental studies uncovered heterogeneity of timescales of neural activity across brain areas and species. Comparison of the population-averaged autocorrelations across cortical areas revealed a hierarchical structure, varying from 50 ms to 350 ms along the occipital-to-frontal axis [6]. Neurons within the same area exhibit a wide distribution of timescales as well. A heterogeneous distribution of timescales (from 0.5 s to 50 s) was found across neurons in the oculomotor system of the fish [11] and primate brainstem [67], suggesting that timescale heterogeneity is conserved across phylogeny. During periods of ongoing activity, the distribution of single-cell autocorrelation timescales in primates was found to be right-skewed and approximately lognormal, ranging from 10 ms to 10 s [10]. Single neuron activity was found to encode long reward memory traces in primate frontal areas over a wide range of timescales up to 10 consecutive trials [9]. In these studies, autocorrelation timescales where estimated using parametric fits, which may be affected by statistical biases, although a new Bayesian generative approach might overcome this issue [68].

In this study, we estimated timescales nonparametrically as the half-width at half-maximum of the autocorrelation function. In our biologically plausible model based on a spiking network with cell-type specific connectivity, the distribution of timescales was in the range between 20 ms and 100 s, similar to the range of timescales observed in experiments [11, 67, 10]. Moreover, the distribution of cluster sizes in our model is within the $50 - 100$ neurons range, consistent with the size of functional assemblies experimentally observed in cortical circuits [60, 64]. A fundamental new prediction of our model, to be tested in future experiments, is the direct relationship between cluster strength and its timescale.

Previous neural mechanisms for generating multiple timescales of neural activity relied on single cell bio-physical properties, such as membrane or synaptic time constants [69]. In feedforward networks, developmental changes in single-cell conductance can modulate the timescale of information transmission, explaining the transition from slow waves to rapid fluctuations observed in the developing cortex [70]. However, the extent to which this single-cell mechanism might persist in presence of strong recurrent dynamics was not assessed. To elucidate this issue, we examined whether a heterogeneous distribution of single-unit integration time constants could lead to a separation of timescales in a random neural network (see Appendix B for details). In this model, half of the units were endowed with a fast time constant which we held fixed, and the other half with a slow time constant, whose value we varied across networks. We found that, although the average network timescale increased proportionally to the value of the slower time constants, the difference in autocorrelation time between the two populations remained negligible. These results suggest that, although the heterogeneity in single-cell time constants may affect the dynamics of single neurons in isolation or within feedforward circuits [70], the presence

of strong recurrent dynamics fundamentally alter these single-cell properties in a counterintuitive way. Our results suggest that a heterogeneity in single cell time constants may not lead to a diversity of timescales in presence of recurrent dynamics.

Our results further clarified that the relationship between a cluster's self-coupling and its timescale relies on the strong recurrent dynamics. This relationship is absent when driving an isolated cluster with white noise external input (Fig. 16). Indeed, the mechanism linking the self-coupling to the timescale only emerged when driving the unit with a mean field whose color was self-consistently obtained from an underlying recurrent network of self-coupled units.

Previous models showed that a heterogeneity of timescales may emerge from circuit dynamics through a combination of structural heterogeneities and heterogeneous long-range connections arranged along a spatial feedforward gradient [71, 72]. These networks can reproduce the population-averaged hierarchy of timescales observed across cortex in the range of $50-350$ ms [6, 72]. A similar network architecture can also reproduce the heterogeneous relaxation time after a saccade, found in the brainstem oculomotor circuit [11, 67], in a range between $10-50$ s [73, 74]. This class of models can explain a timescale separation within a factor of 10, but it is not known whether they can be extended to several orders of magnitude, as observed between neurons in the same cortical area [10]. Moreover, while the feedforward spatial structure underlying these two models is a known feature of the cortical hierarchy and of the brainstem circuit, respectively, it is not known whether such a feedforward structure is present within a local cortical circuit. Our model, on the other hand, relies on strong recurrent connectivity and local functional assemblies, two hallmarks of the architecture of local cortical circuits [60, 61, 62, 63, 64]. Other network models generating multiple timescales of activity fluctuations were proposed

based on self-tuned criticality with anti-hebbian plasticity [75], or multiple block-structured connectivity [76].

In our model, the dynamics of units with large self-couplings, exhibiting slow switching between bistable states, can be captured analytically using the universal colored noise approximation to the Fokker-Planck equation [52, 53]. This is a classic tool from the theory of stochastic processes, which we successfully applied to investigate neural network dynamics for the first time. This slow switching regime may underlie the emergence of metastable activity, ubiquitously observed in the population spiking activity of behaving mammals [77, 78, 79, 80, 74, 81, 82].

What is the functional relevance of neural circuits exhibiting a reservoir of multiple timescales? The presence of long timescales deeply in the chaotic regime is a new feature of our model which may be beneficial for memory capacity away from the edge of chaos [83]. Moreover, we found that, in our model, time-dependent broadband inputs suppress chaos in a population-specific way, whereby slow (fast) subpopulations preferentially entrain slow (fast) spectral components of the input. This mechanism may thus endow recurrent networks with a natural and robust tool to spatially demix complex temporal inputs [22] as observed in visual cortex [84]. Third, the presence of multiple timescales may be beneficial for performing flexible computations involving simultaneously fast and slow timescales, such as in role-switching tasks [21]; or as observed in time cells in the hippocampus [85, 86]. A promising direction for future investigation is the exploration of the computational properties of our model in the context of reservoir computing [87] or recurrent networks trained to perform complex motor and cognitive tasks [88].

APPENDIX A

FOKKER PLANK FORMALISM

The Fokker Plank formalism is used to describe a stochastic process. This formalism has been found useful in physics to study rotational and translational motion of a structured Brownian particles [89], in chemistry to describe complex chemical reaction through a protein folding model [90] and in neuroscience to describe the coupled behaviors of ion transport and channel conformation under an applied membrane potential [91]. In this work we used the Fokker Plank formalism to describe the statistical properties of clusters of neurons in a double-well potential under the influence of colored noise.

In this abstract we will provide a review of the general Fokker Plank formalism. In a stochastic process where the variable $x(t)$ describes the system, the probability distribution $p(x,t)$ obeys the Fokker plank equations:

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\left(F(x)p(x,t)\right) + D\frac{\partial^2}{\partial x^2}p(x,t), \tag{A.1}$$

where $F(x)$ is the force driving the system, and D is the diffusion coefficient. One way to solve the the Fokker Plank equation is to consider it as a Schrodinger type equation. This allows us to use operator techniques from quantum mechanics to find the $p(x,t)$, which furthers allow to obtain averages of macroscopic values through integration [92].

The solution to this equation can be obtained through the separation of variables. The probability distribution can be written as a product of a time dependent function $T(t) = \exp\{(-t|\Lambda_n|)\}$ and a position dependent function. The position dependent

function is the solution to the Schrodinger type equation:

$$-\frac{1}{2}\left[\frac{F(x)^2}{2D}+\frac{\partial}{\partial x}F(x)\right]\Psi_n(x)+D\frac{\partial^2}{\partial x^2}\Psi_n(x)=\Lambda_n\psi_n(x), \tag{A.2}$$

where the eigenfunctions $\Psi_n(x)$ form an orthonormal basis, while Λ_n are eigenvalues. Notice that the left hand side of the equation above resembles the Hamiltonian operator with the terms in square brackets playing the role of an effective potential U_{eff}

$$U_{eff}(x)=\frac{1}{2}\left[\frac{F(x)^2}{2D}+\frac{\partial}{\partial x}F(x)\right] \tag{A.3}$$

Thus, the probability distribution function can be written in terms of eigenfunctions $\Psi_n(x)$ and eigenvalues Λ_n

$$p(x,t)=\psi_0(x)\sum_n a_n\Psi_n(x)\exp\{[-t|\Lambda_n|]\}, \tag{A.4}$$

where $a_n=\Psi_n(x_0)/\Psi_0(x_0)$. The initial condition $x_0\equiv x(t=0)$ is a delta function located at x_0, and $\Psi_0(x_0)$ is the equivalent of the ground state to Eq. A.2.

A.1 Symmetric Double Well Potential

For clarity, we provide the example for finding the probability distribution function $p(x,t)$ for a system that can be related to a symmetric double well effective potential: []

$$U_{eff}(x)=\frac{1}{2}\left(|x|-\alpha\right)^2+\sqrt{\frac{D}{2}}. \tag{A.5}$$

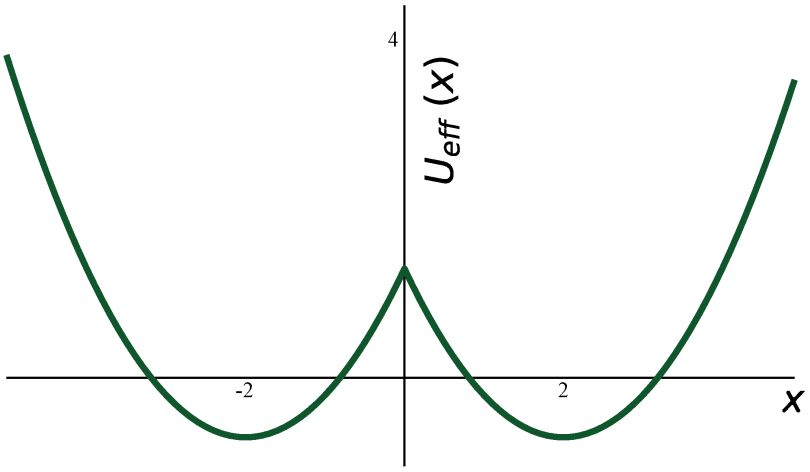

Figure A.1. An illustration of the effective potential in Eq. A.5. The two potential wells are centered at $\pm\alpha$, where $\alpha = 2.0$ and the diffusion parameter $D = 0.5$.

The shape of this potential for $\alpha = 2$ and $D = 1$ is illustrated in Fig. A.1. To find the driving force $F(x)$ we solve Eq. A.3 for the given potential $U_eff(x)$:

$$F(x) = \begin{cases} -\sqrt{2D}(x-\alpha), & \text{for } x > 0 \\ -\sqrt{2D}(x+\alpha), & \text{for } x < 0 \end{cases}.$$

Notice that for $x = \pm\alpha$, the force term $F(x)$ has two minima. This shows us the location of the minima in the potential function.

To find the spatial component of the probability distribution $p(x,t)$, we use rewrite Eq. A.2 as

$$-\frac{1}{2}\left[(|x|-\alpha)^2 + \sqrt{2D}\right]\Psi_n(x) + D\frac{\partial^2}{\partial x^2}\Psi_n(x) = \Lambda_n\psi_n(x). \tag{A.6}$$

For simplicity we chose the diffusion parameter $D = 0.5$. The equation above is analogous to the Schrodinger equation for a quantum double oscillator whose solutions are known: [93]

$$\Psi_n(x) = \frac{1}{Z} \begin{cases} \mathfrak{D}_n\left(\sqrt{2}(x-\alpha)\right), & \text{for } x \geq 0 \\ \pm\mathfrak{D}_n\left(-\sqrt{2}(x+\alpha)\right), & \text{for } x < 0 \end{cases}, \tag{A.7}$$

where $\mathfrak{D}_n$ is the parabolic cylinder function, and Z is a normalization constant. The eigenvalues of the corresponding $\Psi_n(x)$ are $\Lambda_n = -(n+\epsilon)$, where a infinitesimal shift ϵ in the potential is necessary in order ensure that the ground state eigenvalue is $\Lambda_0 = 0$).

The probability density function $p(x,t)$ becomes:

$$p(x,t) = \Psi_0(x)^2 + \frac{\Psi_0(x)}{\Psi_0(x_0)} \sum_{n>0} \Psi_n(x_0)\Psi_n(x) \exp\{[-t|\Lambda_n|]\}.$$

APPENDIX B

NEURON TIME CONSTANT

An assumption that we make for the model class of rate recurrent neural network from Eq. 2.3, is that all clusters of neurons share the same time constant, $\theta = 1$ ms.

The time constant measures the rate of change of neuron's membrane potential. Mathematically, it is represented as $\theta = R_m C_m$, where R_m is resistance and C_m is the capacitance across the membrane. The resistance is determined by the density of ion channels, while the capacitance depends on the lipid layer in the neuron's membrane. The time constant is the time it takes the membrane's potential to decay to approximately 37 % of its initial value.[94] Therefore, one can hypothesize that multiple timescales can arise in a network of neurons with heterogeneous time constants.

To examine such a possibility, we consider the RNN model from Eq.2.3 with no self-coupling term, $s_i = 0$, where each neuron can have its own time constant θ_i:

$$\theta_i \frac{d}{dt} x_i(t) = -x_i(t) + g \sum_j J_{ij} \phi \left[x_j(t) \right]. \tag{B.1}$$

Following the same strategy as in the case with self-couplings, we consider the scenario when our network contains two equal populations of neurons ($N_1 = N_2$) with time constants θ_1 and θ_2. The connectivity matrix $\mathbf{J}$ is extracted from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1/N$. The timescales, τ_i, present in the system are quantified as the width of the autocorrelation function at midpoint. Given a homogeneous time constant $\theta_1 = \theta_2$ we recover the same dynamics as seen in Eq.2.3, where only one timescale is present in the system, Fig. B.1a(i). When considering distinct time constant values $\theta_1 \neq \theta_2$, we observe a new timescale emerging in the network, Fig. B.1a(i-v). While the overall timescales, τ_i, increase with larger

time constants θ_i, the timescale separation between each population, τ_2/τ_1, remains minimal even for large time constant ratios θ_2/θ_1, Fig. B.1b.

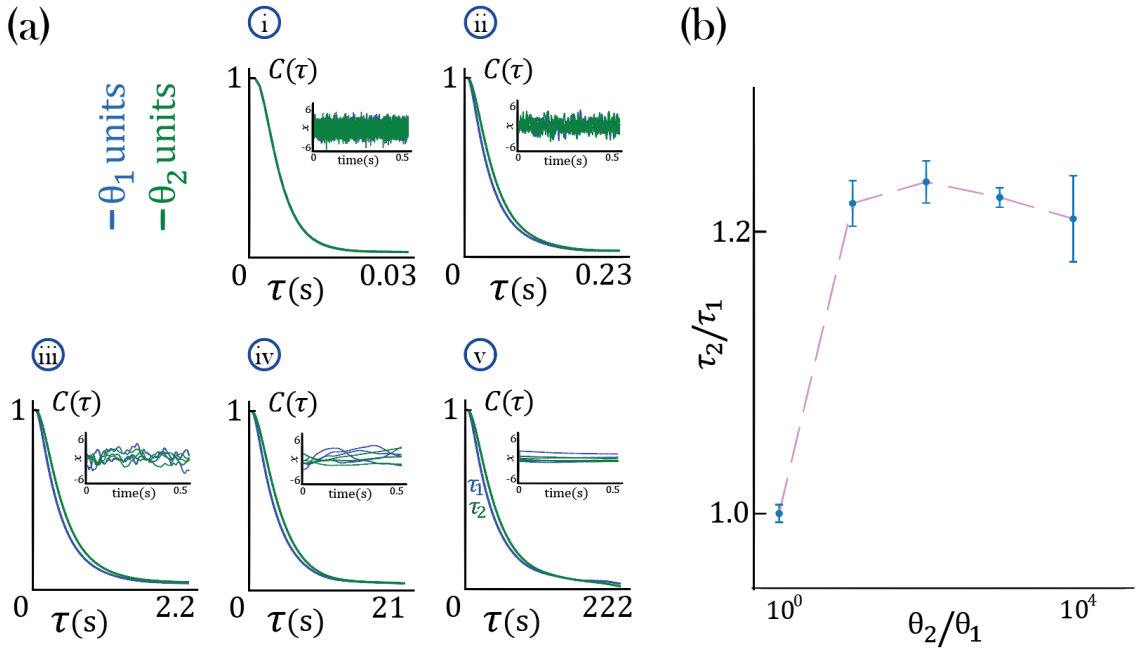

Figure B.1. Timescale analysis for an RNN with two time constants θ_i, Eq. B.1, governing equal populations of neurons ($N_1 = N_2 = 1000$) and gain $g = 2.5$. a) Average autocorrelation function for each population. The insert shows the dynamics of individual neurons from each population: blue for neurons with timeconstant θ_1 and green for neurons with timeconstant θ_2. In the networks considered here, $\theta_1 = 0.1$ ms is kept constant while: $\theta_2 = 0.1$ ms (i), $\theta_2 = 1.0$ ms (ii), $\theta_2 = 10.0$ ms (iii), $\theta_2 = 100.0$ ms (iv), $\theta_2 = 1000.0$ ms (v). b) Population timescale ratio τ_2/τ_1 for fixed timeconstant $\theta_1 = 0.1$ ms and varying θ_2.

Hence, we conclude that heterogeneous timeconstants is not an efficient mechanism to generate discrete timescales that range over multiple orders of magnitudes.

APPENDIX C

ESCAPE RATE: KRAMER'S FORMULA

The problem that we want to solve is determining the rate at which a Brownian particle escapes from a potential well. For that, we consider a collection of Brownian particles located in a well of the potential $U(x)$ (Fig. C.1). Our expectation is that the particles will reach a close to equilibrium state, however some of the particles will cross the potential barrier and slowly leak out. To determine the escape rate we start from the Langevin equation for the particles in the overdamped regime ($m\frac{d}{dt}v(t) = 0$):

$$\gamma \frac{d}{dt}x(t) = -\frac{\partial}{\partial x}U(x) + \eta(t). \tag{C.1}$$

We can describe the probability density $p(x,t)$ of the particles in the potential $U(x)$ through the Fokker Plank equation (Appendix A)

$$\begin{aligned} \frac{\partial}{\partial x}p(x,t) &= \frac{\partial}{\partial x}\left[\frac{1}{\gamma}\left(\frac{\partial}{\partial x}U(x)\right)p(x,t) + D\frac{\partial}{\partial x}p(x,t)\right] \\ &= -\frac{\partial}{\partial x}J(x,t), \end{aligned} \tag{C.2}$$

where the current term $J(x,t)$ can be rewritten as:

$$\begin{aligned} J(x,t) &= -\frac{1}{\gamma}\left(\frac{\partial}{\partial x}U(x)\right)p(x,t) - D\frac{\partial}{\partial x}p(x,t) \\ &= -D\exp\left[-\frac{U(x)}{D\gamma}\right]\frac{\partial}{\partial x}\left(\exp\left[\frac{U(x)}{D\gamma}p(x,t)\right]\right). \end{aligned} \tag{C.3}$$

At equilibrium, the current term will be zero and probability density function is time independent. We can further rewrite the equation above as:

$$\frac{\partial}{\partial x}\left(\exp\left[\frac{U(x)}{\gamma D}\right]p(x,t)\right) = -\frac{J}{D}\exp\left[\frac{U(x)}{\gamma D}\right].$$

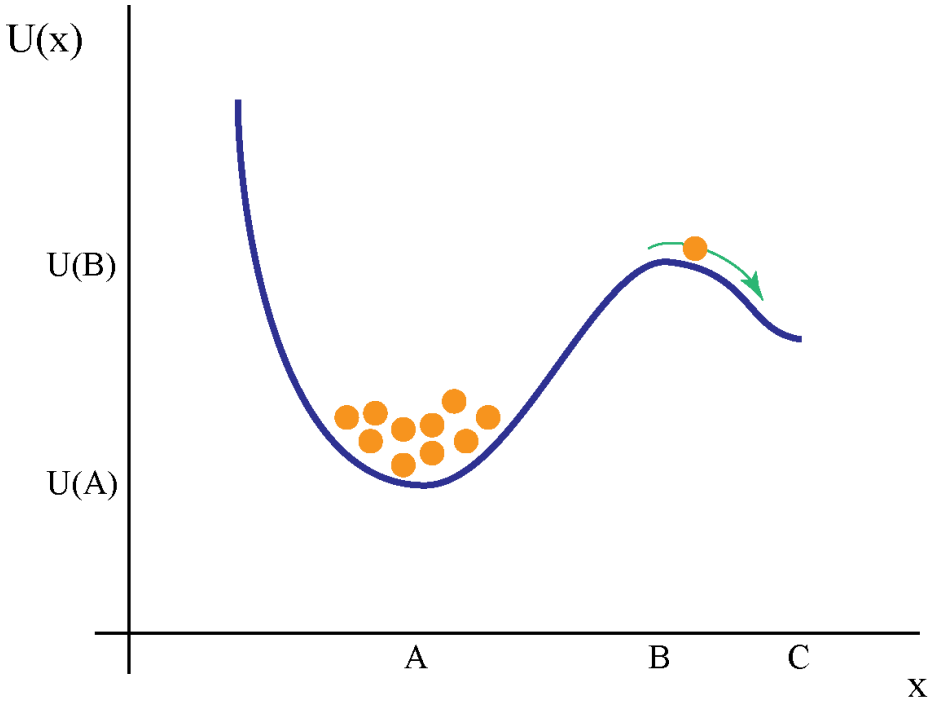

Figure C.1. Potential energy as a function of reaction coordinate. The potential function has a stable minimum at $x = A$, and a barrier at $x = B$. In order to escape, a particle must overcome the potential barrier and move from point $x = A$ to $x = C$.

We proceed by integrating this expression from $x = A$ to $x = C$, while assuming that the particles are approximately at equilibrium, at the bottom of the well A in $U(x)$. In this regime, there will be a small current across the potential barrier, thus $p(x = C)$ is very small and the probability density function can be approximated as being time independent. The small current becomes

$$J = \frac{D \exp\left[\frac{U(x)}{\gamma D}\right] p(A)}{\int_A^C \exp\left[\frac{U(x')}{\gamma D}\right] dx'},$$

while the probability distribution function

$$p(x) = p(A) \exp\left[-\frac{U(x) - U(A)}{\gamma D}\right] p(x, t).$$

If we consider the escape rate r of a particle as the conditional probability of escape per unit time given that the particle is located in the well $x = A$, and let P be the probability of the particle being inside the well, the current can be written as the product between these quantities $J = rP$. The probability of being inside the potential well can be determined by integrating around the well:

$$\begin{aligned} P &= \int_{a-\Delta}^{a+\Delta} p(x)dx \\ &= p(A)\exp\left[\frac{U(A)}{\gamma D}\right]\int_{a-\Delta}^{a+\Delta}\exp\left[-\frac{U(x')}{\gamma D}\right]dx' \\ &\approx p(A)\left(\frac{2\pi D\gamma}{|U''(A)|}\right)^{1/2}. \end{aligned}$$

To obtain the last line, we Taylor expanded around the point $x = A$ since the integrand is peaks at this point. Similarly, we can expand around the point $x = B$ to approximate the denominator in the formula for the current J.

$$\int_A^C \exp\left[\frac{U(x')}{\gamma D}\right]dx' \approx \left(\frac{2\pi D\gamma}{|U''(B)|}\right)^{1/2}\exp\left[\frac{U(B)}{\gamma D}\right].$$

Using the approximations above, we get an expression for the $r = J/P$

$$r = \frac{1}{2\pi\gamma}\left(U''(A)U''(B)\right)^{1/2}\exp\left[-\frac{U(B)-U(A)}{\gamma D}\right].$$

This formula is known as Kramer's escape rate formula. The quantity $U(B) - U(A)$ describes the height of the barrier that the particle needs to overcome in order to escape. The escape rate falls exponentially with the barrier height.

REFERENCES CITED

[1] J.-P. Bouchaud, "Weak ergodicity breaking and aging in disordered systems," *Journal de Physique I* **2** (1992) no. 9, 1705–1713.

[2] E. Almaas, B. Kovacs, T. Vicsek, Z. N. Oltvai, and A.-L. Barabási, "Global organization of metabolic fluxes in the bacterium escherichia coli," *Nature* **427** (2004) no. 6977, 839–843.

[3] M. Emmerling, M. Dauner, A. Ponti, J. Fiaux, M. Hochuli, T. Szyperski, K. Wüthrich, J. Bailey, and U. Sauer, "Metabolic flux responses to pyruvate kinase knockout in escherichia coli," *Journal of bacteriology* **184** (2002) no. 1, 152–164.

[4] M. R. Roussel and D. Lloyd, "Observation of a chaotic multioscillatory metabolic attractor by real-time monitoring of a yeast continuous culture," *The FEBS Journal* **274** (2007) no. 4, 1011–1018.

[5] M. A. Aon, M. R. Roussel, S. Cortassa, B. O'Rourke, D. B. Murray, M. Beckmann, and D. Lloyd, "The scale-free dynamics of eukaryotic cells," *PLoS One* **3** (2008) no. 11, e3624.

[6] J. D. Murray, A. Bernacchia, D. J. Freedman, R. Romo, J. D. Wallis, X. Cai, C. Padoa-Schioppa, T. Pasternak, H. Seo, D. Lee, *et al.*, "A hierarchy of intrinsic timescales across primate cortex," *Nature neuroscience* **17** (2014) no. 12, 1661–1663.

[7] J. H. Siegle, X. Jia, S. Durand, S. Gale, C. Bennett, N. Graddis, G. Heller, T. K. Ramirez, H. Choi, J. A. Luviano, *et al.*, "A survey of spiking activity reveals a functional hierarchy of mouse corticothalamic visual areas," *Biorxiv* (2019) 805010.

[8] R. Gao, R. L. van den Brink, T. Pfeffer, and B. Voytek, "Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture," *Elife* **9** (2020) e61277.

[9] A. Bernacchia, H. Seo, D. Lee, and X.-J. Wang, "A reservoir of time constants for memory traces in cortical neurons," *Nature neuroscience* **14** (2011) no. 3, 366–372.

[10] S. E. Cavanagh, J. D. Wallis, S. W. Kennerley, and L. T. Hunt, "Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice," *Elife* **5** (2016) e18937.

[11] A. Miri, K. Daie, A. B. Arrenberg, H. Baier, E. Aksay, and D. W. Tank, "Spatial gradients and multidimensional dynamics in a neural integrator circuit," *Nature neuroscience* **14** (2011) no. 9, 1150–1159.

[12] H. Sompolinsky, A. Crisanti, and H. J. Sommers, "Chaos in random neural networks," *Phys. Rev. Lett.* **61** (Jul, 1988) 259–262. https://link.aps.org/doi/10.1103/PhysRevLett.61.259.

[13] M. Stern, H. Sompolinsky, and L. F. Abbott, "Dynamics of random neural networks with bistable units," *Phys. Rev. E* **90** (Dec, 2014) 062710. https://link.aps.org/doi/10.1103/PhysRevE.90.062710.

[14] K. Rajan, L. F. Abbott, and H. Sompolinsky, "Stimulus-dependent suppression of chaos in recurrent neural networks," *Phys. Rev. E* **82** (Jul, 2010) 011903. https://link.aps.org/doi/10.1103/PhysRevE.82.011903.

[15] A. Seth, *30-second brain: The 50 most mind-blowing ideas in neuroscience, each explained in half a minute*. Icon Books Ltd, 2014.

[16] J. Werner and L. Chalupa, *The New Visual Neurosciences*. MIT Press, 2013. https://books.google.com/books?id=sjoiEAAAQBAJ.

[17] R. V. Raut, A. Mitra, S. Marek, M. Ortega, A. Z. Snyder, A. Tanenbaum, T. O. Laumann, N. U. F. Dosenbach, and M. E. Raichle, "Organization of Propagated Intrinsic Brain Activity in Individual Humans," *Cerebral Cortex* **30** (09, 2019) 1716–1734, https://academic.oup.com/cercor/article-pdf/30/3/1716/33011727/bhz198.pdf. https://doi.org/10.1093/cercor/bhz198.

[18] M. Demirtaş, J. B. Burt, M. Helmer, J. L. Ji, B. D. Adkinson, M. F. Glasser, D. C. Van Essen, S. N. Sotiropoulos, A. Anticevic, and J. D. Murray, "Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics," *Neuron* **101** (2019) no. 6, 1181–1194.e13. https://www.sciencedirect.com/science/article/pii/S0896627319300443.

[19] M. Golesorkhi, J. Gomez-Pilar, S. Tumati, M. Fraser, and G. Northoff, "Temporal hierarchy of intrinsic neural timescales converges with spatial core-periphery organization," *Communications Biology* **4** (03, 2021) 1234567890.

[20] M. Golesorkhi, J. Gomez-Pilar, F. Zilio, N. Berberian, A. Wolff, M. Yagoub, and G. Northoff, "The brain and its time: intrinsic neural timescales are key for input processing," *Communications Biology* **4** (12, 2021) .

[21] K. Iigaya, Y. Ahmadian, L. P. Sugrue, G. S. Corrado, Y. Loewenstein, W. T. Newsome, and S. Fusi, "Deviation from the matching law reflects an optimal strategy involving learning over multiple timescales," *Nature communications* **10** (2019) no. 1, 1–14.

[22] N. Perez-Nieves, V. C. Leung, P. L. Dragotti, and D. F. Goodman, "Neural heterogeneity promotes robust learning," *bioRxiv* (2021) 2020–12.

[23] U. Hasson, J. Chen, and C. J. Honey, "Hierarchical process memory: memory as an integral component of information processing," *Trends in Cognitive Sciences* **19** (2015) no. 6, 304–313. https://www.sciencedirect.com/science/article/pii/S1364661315000923.

[24] M. A. Arbib and J. J. Bonaiuto, *From neuron to cognition via computational neuroscience*. MIT Press, 2016.

[25] H. Head and G. Holmes, "Sensory disturbances from cerebral lesions," *Brain* **34** (11, 1911) 102–254, https://academic.oup.com/brain/article-pdf/34/2-3/102/933215/34-2-3-102.pdf. https://doi.org/10.1093/brain/34.2-3.102.

[26] F. Bartlett and W. Kintsch, *Remembering: A Study in Experimental and Social Psychology*. Camb. Psychol. Libr. Cambridge University Press, 1995. https://books.google.com/books?id=WG5ZcHGTrm4C.

[27] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics* **5** (1943) no. 4, 115–133.

[28] A. L. Hodgkin and A. F. Huxley, "Action potentials recorded from inside a nerve fibre," *Nature* **144** (1939) no. 3651, 710–711.

[29] P. Garcia-Lopez, V. Garcia-Marin, and M. Freire, "The histological slides and drawings of cajal," *Frontiers in neuroanatomy* **4** (2010) 9.

[30] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons.," *Biophysical journal* **12 1** (1972) 1–24.

[31] B. Li, B. N. Routh, D. Johnston, E. Seidemann, and N. J. Priebe, "Voltage-gated intrinsic conductances shape the input-output relationship of cortical neurons in behaving primate v1," *Neuron* **107** (2020) no. 1, 185–196.

[32] K. H. Fischer and J. A. Hertz, *Spin Glasses*. 1993.

[33] P. C. Martin, E. D. Siggia, and H. A. Rose, "Statistical dynamics of classical systems," *Phys. Rev. A* **8** (Jul, 1973) 423–437. https://link.aps.org/doi/10.1103/PhysRevA.8.423.

[34] C. d. Dominicis, "Techniques of field renormalization and dynamics of critical phenomena," in *J. Phys.(Paris), Colloq*, pp. C1–247. 1976.

[35] H.-K. Janssen, "On a lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties," *Zeitschrift für Physik B Condensed Matter* **23** (1976) no. 4, 377–380.

[36] C. De Dominicis and L. Peliti, "Field-theory renormalization and critical dynamics above t c: Helium, antiferromagnets, and liquid-gas systems," *Physical Review B* **18** (1978) no. 1, 353.

[37] C. De Dominicis, "Dynamics as a substitute for replicas in systems with quenched random impurities," *Physical Review B* **18** (1978) no. 9, 4913.

[38] H. Sompolinsky and A. Zippelius, "Relaxational dynamics of the edwards-anderson model and the mean-field theory of spin-glasses," *Physical Review B* **25** (1982) no. 11, 6860.

[39] A. Crisanti and H. Sompolinsky, "Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model," *Physical Review A* **36** (1987) no. 10, 4922.

[40] M. Helias and D. Dahmen, *Statistical field theory for neural networks*. Springer, 2020.

[41] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer., 2009.

[42] H. Sompolinsky and A. Zippelius, "Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses," *Phys. Rev. B* **25** (June, 1982) 6860–6875.

[43] A. Papoulis, S. Pillai, and S. Pillai, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill electrical and electronic engineering series. McGraw-Hill, 2002. `https://books.google.com/books?id=YYwQAQAAIAAJ`.

[44] A. Crisanti and H. Sompolinsky, "Path integral approach to random neural networks," *Phys. Rev. E* **98** (Dec, 2018) 062120. `https://link.aps.org/doi/10.1103/PhysRevE.98.062120`.

[45] D. Martí, N. Brunel, and S. Ostojic, "Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks," *Phys. Rev. E* **97** (Jun, 2018) 062314. `https://link.aps.org/doi/10.1103/PhysRevE.97.062314`.

[46] D. Sussillo and L. Abbott, "Generating coherent patterns of activity from chaotic neural networks," *Neuron* **63** (2009) no. 4, 544–557. `https://www.sciencedirect.com/science/article/pii/S0896627309005479`.

[47] J. Ginibre, "Statistical ensembles of complex, quaternion, and real matrices," *Journal of Mathematical Physics* **6** (1965) no. 3, 440–449, https://doi.org/10.1063/1.1704292. https://doi.org/10.1063/1.1704292.

[48] H. J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, "Spectrum of large random asymmetric matrices," *Phys. Rev. Lett.* **60** (May, 1988) 1895–1898. https://link.aps.org/doi/10.1103/PhysRevLett.60.1895.

[49] Y. Ahmadian, F. Fumarola, and K. D. Miller, "Properties of networks with partially structured and partially random connectivity,". https://arxiv.org/abs/1311.4672.

[50] P. Jung and P. Hänggi, "Dynamical systems: A unified colored-noise approximation," *Phys. Rev. A* **35** (May, 1987) 4464–4466. https://link.aps.org/doi/10.1103/PhysRevA.35.4464.

[51] P. Häunggi and P. Jung, "Colored noise in dynamical systems," *Advances in Chemical Physics* **89** (2007) 239–326.

[52] P. Hänggi and P. Jung, "Colored noise in dynamical systems," *Advances in chemical physics* **89** (1995) 239–326.

[53] P. Jung and P. Hänggi, "Dynamical systems: a unified colored-noise approximation," *Physical review A* **35** (1987) no. 10, 4464.

[54] M. Stern, N. Istrate, and L. Mazzucato, "A reservoir of timescales in random neural networks," *bioRxiv* (2021) , https://www.biorxiv.org/content/early/2021/10/12/2021.10.11.463861.full.pdf. https: //www.biorxiv.org/content/early/2021/10/12/2021.10.11.463861.

[55] R. Cooper, N. Ruh, and D. Mareschal, "The goal circuit model: A hierarchical multi-route model of the acquisition and control of routine sequential action in humans," *Cognitive Science* **38** (08, 2013) .

[56] R. Cooper and T. Shallice, "Contention scheduling and the control of routine activities," *Cognitive Neuropsychology* **17** (06, 2000) 297–338.

[57] R. Cooper, T. Shallice, J. Farringdon, *et al.*, "Symbolic and continuous processes in the automatic selection of actions," *Hybrid problems, hybrid solutions* **2** (1995) 27–37.

[58] G. R. Yang, I. Ganichev, X.-J. Wang, J. Shlens, and D. Sussillo, "A dataset and architecture for visual reasoning with a working memory," in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, eds., pp. 729–745. Springer International Publishing, Cham, 2018.

[59] M. Haruno, D. M. Wolpert, and M. Kawato, "Mosaic model for sensorimotor learning and control," *Neural computation* **13** (2001) no. 10, 2201–2220.

[60] R. Perin, T. K. Berger, and H. Markram, "A synaptic organizing principle for cortical neuronal groups," *Proc. Natl. Acad. Sci. U. S. A.* **108** (Mar., 2011) 5419–5424.

[61] W.-C. A. Lee, V. Bonin, M. Reed, B. J. Graham, G. Hood, K. Glattfelder, and R. Clay Reid, "Anatomy and function of an excitatory network in the visual cortex," *Nature* **532** (2016) no. 7599, 370–374.

[62] R. Kiani, C. J. Cueva, J. B. Reppas, D. Peixoto, S. I. Ryu, and W. T. Newsome, "Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex," *Neuron* **85** (2015) no. 6, 1359–1373.

[63] J.-e. K. Miller, I. Ayzenshtat, L. Carrillo-Reid, and R. Yuste, "Visual stimuli recruit intrinsically generated cortical ensembles," *Proceedings of the National Academy of Sciences* **111** (2014) no. 38, E4053–E4061.

[64] J. H. Marshel, Y. S. Kim, T. A. Machado, S. Quirin, B. Benson, J. Kadmon, C. Raja, A. Chibukhchyan, C. Ramakrishnan, M. Inoue, *et al.*, "Cortical layer–specific critical dynamics triggering perception," *Science* **365** (2019) no. 6453, .

[65] A. Litwin-Kumar and B. Doiron, "Slow dynamics and high variability in balanced cortical networks with clustered connections," *Nat Neurosci* **15** (2012) no. 11, 1498–505. `http://www.ncbi.nlm.nih.gov/pubmed/23001062`.

[66] D. Wyrick and L. Mazzucato, "State-dependent regulation of cortical processing speed via gain modulation," *Journal of Neuroscience* **41** (2021) no. 18, 3988–4005.

[67] M. Joshua, J. F. Medina, and S. G. Lisberger, "Diversity of neural responses in the brainstem during smooth pursuit eye movements constrains the circuit mechanisms of neural integration," *Journal of Neuroscience* **33** (2013) no. 15, 6633–6647.

[68] R. Zeraati, T. A. Engel, and A. Levina, "Estimation of autocorrelation timescales with approximate bayesian computations," *bioRxiv* (2020) .

[69] J. Gjorgjieva, G. Drion, and E. Marder, "Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance," *Current opinion in neurobiology* **37** (2016) 44–52.

[70] J. Gjorgjieva, R. A. Mease, W. J. Moody, and A. L. Fairhall, "Intrinsic neuronal properties switch the mode of information transmission in networks," *PLoS computational biology* **10** (2014) no. 12, e1003962.

[71] R. Chaudhuri, A. Bernacchia, and X.-J. Wang, "A diversity of localized timescales in network activity," *elife* **3** (2014) e01239.

[72] R. Chaudhuri, K. Knoblauch, M.-A. Gariel, H. Kennedy, and X.-J. Wang, "A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex," *Neuron* **88** (2015) no. 2, 419–431.

[73] H. K. Inagaki, L. Fontolan, S. Romani, and K. Svoboda, "Discrete attractor dynamics underlies persistent activity in the frontal cortex," *Nature* **566** (2019) no. 7743, 212–217.

[74] S. Recanatesi, U. Pereira-Obilinovic, M. Murakami, Z. Mainen, and L. Mazzucato, "Metastable attractors explain the variable timing of stable behavioral action sequences," *Neuron* (2022) .

[75] M. O. Magnasco, O. Piro, and G. A. Cecchi, "Self-tuned critical anti-hebbian networks," *Physical review letters* **102** (2009) no. 25, 258102.

[76] J. Aljadeff, M. Stern, and T. Sharpee, "Transition to chaos in random networks with cell-type-specific connectivity," *Physical review letters* **114** (2015) no. 8, 088101.

[77] M. Abeles, H. Bergman, I. Gat, I. Meilijson, E. Seidemann, N. Tishby, and E. Vaadia, "Cortical activity flips among quasi-stationary states," *Proc Natl Acad Sci USA* **92** (1995) 8616–8620.

[78] L. M. Jones, A. Fontanini, B. F. Sadacca, P. Miller, and D. B. Katz, "Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles," *Proc Natl Acad Sci U S A* **104** (2007) no. 47, 18772–7.

[79] L. Mazzucato, A. Fontanini, and G. La Camera, "Dynamics of multistable states during ongoing and evoked cortical activity," *The Journal of Neuroscience* **35** (2015) no. 21, 8214–8231.

[80] L. Mazzucato, G. La Camera, and A. Fontanini, "Expectation-induced modulation of metastable activity underlies faster coding of sensory stimuli," *Nature neuroscience* **22** (2019) no. 5, 787–796.

[81] T. A. Engel, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, and K. Boahen, "Selective modulation of cortical state during spatial attention," *Science* **354** (2016) no. 6316, 1140–1144.

[82] N. Kadmon Harpaz, D. Ungarish, N. G. Hatsopoulos, and T. Flash, "Movement decomposition in the primary motor cortex," *Cerebral Cortex* **29** (2019) no. 4, 1619–1633.

[83] T. Toyoizumi and L. Abbott, "Beyond the edge of chaos: amplification and temporal integration by recurrent networks in the chaotic regime," *Physical Review E* **84** (2011) no. 5, 051908.

[84] A. Mazzoni, S. Panzeri, N. K. Logothetis, and N. Brunel, "Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons," *PLoS Comput Biol* **4** (2008) no. 12, e1000239.

[85] B. J. Kraus, R. J. Robinson II, J. A. White, H. Eichenbaum, and M. E. Hasselmo, "Hippocampal "time cells": time versus path integration," *Neuron* **78** (2013) no. 6, 1090–1101.

[86] M. W. Howard, C. J. MacDonald, Z. Tiganj, K. H. Shankar, Q. Du, M. E. Hasselmo, and H. Eichenbaum, "A unified mathematical framework for coding time, space, and sequences in the hippocampal region," *Journal of Neuroscience* **34** (2014) no. 13, 4692–4707.

[87] D. Sussillo and L. F. Abbott, "Generating coherent patterns of activity from chaotic neural networks," *Neuron* **63** (2009) no. 4, 544–557.

[88] G. R. Yang, M. R. Joglekar, H. F. Song, W. T. Newsome, and X.-J. Wang, "Task representations in neural networks trained to perform many cognitive tasks," *Nature neuroscience* **22** (2019) no. 2, 297–306.

[89] M. H. Peters, "Fokker-planck equation and the grand molecular friction tensor for coupled rotational and translational motions of structured brownian particles near structured surfaces," *The Journal of chemical physics* **110** (1999) no. 1, 528–538.

[90] J. D. Bryngelson and P. G. Wolynes, "Intermediates and barrier crossing in a random energy model (with applications to protein folding)," *The Journal of Physical Chemistry* **93** (1989) no. 19, 6902–6915.

[91] K. Lee and W. Sung, "Ion transport and channel transition in biomembranes," *Physica A: Statistical Mechanics and its Applications* **315** (2002) no. 1-2, 79–97.

[92] H. Risken and T. Frank, *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 1996. `https://books.google.com/books?id=MG2V9vTgSgEC`.

[93] J. Brickmann, "Properties of the quantum double oscillator," *Zeitschrift für Naturforschung A* **30** (1975) no. 12, 1730–1741.

[94] W. R. Holmes, *Time Constant, τ, in Neuronal Signaling*, pp. 1–2. Springer New York, New York, NY, 2013. `https://doi.org/10.1007/978-1-4614-7320-6_37-1`.