

COLLECTIVE BEHAVIOR IN COMPLEX NETWORKS: APPLICATIONS TO BIOLOGY

By

Shuyue Xue

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Physics — Doctor of Philosophy  
Computational Mathematics, Science and Engineering — Dual Major

2025

## ABSTRACT

This dissertation investigates the principles of collective behavior in complex networks, focusing on how modular architectures and non-reciprocal interactions shape system dynamics and adaptive responses. The research integrates two complementary approaches: a data-driven analysis of biological networks (Track I) and a theoretical exploration of a non-reciprocal Hopfield model (Track II).

Track I employs Differential Network (DN) analysis to examine longitudinal RNA-sequencing data from two biological systems: human saliva following PPSV23 vaccination and primary B-cells subjected to Rituximab treatment. This methodology uncovers stimulus-specific modular reorganization within these networks. Key findings include the identification of temporally ordered activation patterns among gene communities: in saliva, fifteen gene communities show a well-ordered activation cascade; in B cells, fourteen communities cluster into three temporal response classes. Functional enrichment confirms that each module specializes in coherent pathways, while hub-gene analysis highlights IL4R (saliva) and PELI1 (B cells) as putative drivers of the observed immune and drug responses. These results suggest that inter-module couplings are pivotal in orchestrating complex biological processes.

Track II develops a theoretical framework using a non-reciprocal Hopfield network, featuring two interacting subnetworks (termed similarity and differential), to elucidate the dynamic mechanisms that could drive such modular behaviors. This investigation utilizes a combination of mean-field theory, stochastic Langevin equations, Master Equation formalism, and large-scale Glauber Monte Carlo simulations. The model exhibits a rich phase diagram with distinct paramagnetic, memory retrieval, and limit-cycle dynamical regimes. These phases are separated by Hopf and fold bifurcation lines. Critical dynamics near these lines are characterized by scaling exponents  $\zeta = \frac{1}{2}$  (Hopf) and  $\zeta = \frac{1}{3}$  (fold), distinct response-time

laws ( $|F|^{-2/3}$  vs.  $|F|^{-1/2}$ ), and a limit-cycle coherence time  $T \sim N$  that quantifies finite-size effects. These analytical predictions are numerically validated.

By bridging empirical observations of adaptive modular responses in biological systems with a mechanistic understanding derived from a tractable theoretical model, this dissertation offers significant insights into how interactions both within and between network modules collectively govern global network behavior and facilitate functional adaptation in complex systems.

Dedicated to my mom and dad, grandparents, mentors, and friends.

## ACKNOWLEDGMENTS

I am deeply grateful to the many wonderful human beings who have surrounded and supported me throughout my course of graduate school.

Foremost, I feel fortunate to have loving parents and grandmother, whom I've missed dearly. Their daily video calls and messages have been a source of comfort and motivation. They have provided me with a stable and nurturing safety net wherever I go, while instilling values that guide every decision I make. My cousins have been like siblings to me, and my uncle and aunt have offered wisdom whenever I faced difficult choices. Their sacrifices and belief in my potential have been a constant source of strength.

My gratitude goes to my advisor, Professor Carlo Piermarocchi, and my co-advisor, Professor George Mias, for their mentoring throughout this journey. They supported my development through every stage of the research process, from defining problems and interpreting results to writing papers, selecting journals, and carefully navigating the peer review process. This comprehensive experience refined my technical skills while fostering my growth toward becoming an independent researcher.

With his depth of experience and scientific intuition, Carlo often offered perspectives that were valuable in interpreting and validating my results. Much of my work in theoretical condensed matter took shape under his guidance, where I benefited from his expertise in the field. I treasure the memories of traveling to conferences with him, from local events and regional group trips to navigating the APS March Meeting together. He consistently pointed me toward relevant talks and discussions, which helped me stay focused and engaged. His passion for science brought lasting energy, even through the longest days.

George is the one who first brought me into the fields of genetics and bioinformatics.

While navigating this new area often required me to find my own path through trial and error, I appreciated his continued engagement and efforts in supporting me by sharing resources, monitoring my progress, and encouraging me to connect within the field. His guidance in opening these interdisciplinary doors has been meaningful in shaping my future research direction.

I extend my appreciation to the remaining members of my committee: Professors Morten Hjorth-Jensen, Kirsten Tollefson, Michael Murillo, and Jianrong Wang, for their thoughtful guidance, steady support, and the mentorship. Each of them has influenced me in unique ways that have shaped who I am today—from my research interests and scientific insight, to my skills in computation and writing, and even to how I build meaningful relationships or approach leisure.

From Morten, I have adopted his habit of connecting dots between analytical derivations, computational simulations, and physical intuitions, though I continue to aspire toward his level of mastery. His elegant problem-solving style and generous mentorship—in both scientific thinking and career development—have had a lasting and meaningful impact on me.

I'm deeply grateful to Kirsten for her insight, inclusive leadership, and steadfast, effective guidance during challenging moments. Her strategic instinct in finding me paths that truly served my best interests was something I found invaluable, and her guidance remains treasured in my memories.

Michael has been my go-to mentor in data science. His sharp technical insight, fluency in cutting-edge algorithms, and ever-curious outlook make every chat with him both enlightening and entertaining. Whether showing me the latest cool techniques or sharing thought-provoking writing strategies, I could always count on learning something new.

I also learned a great deal from Jianrong in computational medicine. His course of-

ferred a structured view into this rapidly evolving field, and his early guidance in shaping research directions and identifying promising project ideas proved valuable in my academic development.

Beyond my committee, I would also like to thank Professor Mohammad Maghrebi for his collaboration on my Hopfield network project. His expertise in phase transitions contributed new perspectives to our analytical approach. Meetings with him and Carlo turned out to be thought-provoking conversations that enriched the development of my research.

I'm fortunate to have met Professor Remco Zegers, whose mentorship extended far beyond academic matters. Whether sharing career insights or offering practical wisdom about time management and balancing hobbies with research, he consistently opened new perspectives for me on sustainable academic life. I'll remember my chats with him fondly

My thanks are also due to the Department of Physics and Astronomy for many years of Teaching Assistantship funding that supported my graduate work. A warm thank you to Kim Crosslan, who has been like a wonderful department mom to all of us, and to Susie Harwood-Brown for handling my conference reimbursements. Kim, I'll especially miss your warm hugs!

In addition, I'm thankful to my former lab mates (who have now become friends), Eren Veziroglu and Elliot Lu. Their enthusiasm and dedication to science are a source of my motivation. Thank you for those uplifting phone calls and chats!

Finally, a massive thank you to all my friends I've had along the way. To Meiqi Liu, Kang Yu, Annabelle Miller, Swapna Mishra, Boyao Zhu, Zhite Yu, Xiaoyi Sun, Jane Kim, Julie Butler, Erin White, Jeanetta Mohlke-Hill, and Stephen Hemmerle: you have all enriched my journey in countless ways. I cherish the moments we shared with laughs over coffee, dinners, parties, and local travels, and look forward to the adventures that lie ahead.

## PREFACE

The original scientific content presented in this dissertation is composed in part by one peer-reviewed publication and one preprint which has been submitted for review. The following outlines how the chapters of this dissertation relate to these works:

**Chapter 2–3:** chapters has been adapted from the following publication:

Xue S, Rogers LRK, Zheng M, He J, Piermarocchi C, Mias GI. (2022). Applying differential network analysis to longitudinal gene expression in response to perturbations. *Frontiers in Genetics*, **13**, 1026487.

The code supporting the research in this chapter is available at: <https://github.com/gmiaslab/DifferentialNetworks>

**Chapters 4–7:** chapters have been adapted from the following manuscript under review:

Xue, S., Maghrebi, M., Mias, G. I., & Piermarocchi, C. (2025). Critical Dynamics of Non-Reciprocal Hopfield Networks. *arXiv preprint arXiv:2501.00983*. (Submitted for peer review).

The code supporting the research in these chapters is available at: <https://github.com/shuyue13/non-reciprocal-Hopfield>

**Additional Contributions:** I contributed to related research that provided experience for the work presented in Chapter 3:

Mias, G.I., Singh, V.V., Rogers, L.R.K., Xue, S., Zheng, M., Domanskyi, S., Kanada, M., Piermarocchi, C., & He, J. (2021). Longitudinal saliva omics responses to immune perturbation: a case study. *Scientific Reports*, **11**, 2046. <https://doi.org/10.1038/s41598-020-80605-6>

All computational code provided in the repositories listed in this preface and utilized for the analyses within this dissertation was authored by Shuyue Xue. For inquiries regarding the code, please contact [xueshuy1@msu.edu](mailto:xueshuy1@msu.edu) or [shuyue.xue413@gmail.com](mailto:shuyue.xue413@gmail.com).

## TABLE OF CONTENTS

<b>Preface</b> . . . . .	<b>viii</b>
<b>LIST OF TABLES</b> . . . . .	<b>xi</b>
<b>LIST OF FIGURES</b> . . . . .	<b>xii</b>
<b>Chapter 1. Introduction</b> . . . . .	<b>1</b>
1.1 Motivations . . . . .	1
1.1.1 Adaptive Responses in Complex Systems . . . . .	1
1.1.2 Networks and Modularity as an Organizing Principle . . . . .	2
1.1.3 Differential Modular Responses in Biological Networks . . . . .	3
1.1.4 From Observation to Mechanism . . . . .	4
1.1.5 A Non-Reciprocal Hopfield Network for Modular Dynamics . . . . .	5
1.1.6 Dissertation Overview and Integrative Perspective . . . . .	6
1.2 Dissertation Outline . . . . .	7
<b>Chapter 2. Foundations of Differential Network Analysis</b> . . . . .	<b>10</b>
2.1 Introduction to Network Science in Biological Systems . . . . .	10
2.2 Fundamentals of Network Theory . . . . .	13
2.2.1 Basics Definitions and Representations . . . . .	13
2.2.2 Network Centrality: Identifying Important Nodes . . . . .	15
2.2.3 Community Structure . . . . .	16
2.2.4 Community Detection Algorithms . . . . .	17
2.2.4.1 Modularity . . . . .	19
2.2.5 The Louvain Algorithm . . . . .	22
2.2.5.1 Computational Limitations . . . . .	25
2.3 Differential Network (DN) Analysis . . . . .	26
2.3.1 Concept and Formulation . . . . .	26
2.3.2 Topological Analysis of Differential Networks . . . . .	28
<b>Chapter 3. Application of Differential Network Analysis to Salivary and B-Cell Transcriptomic Data</b> . . . . .	<b>30</b>
3.1 Data Acquisition . . . . .	31
3.1.1 Data Availability . . . . .	32
3.2 Application Workflow . . . . .	33
3.2.1 Data Preprocessing . . . . .	34
3.2.2 Gene Selection . . . . .	36
3.2.3 Co-Expression Networks Construction . . . . .	36
3.2.4 Differential Networks Construction . . . . .	37
3.2.5 Community Specific Time-Resolved Analysis . . . . .	38
3.2.6 Pathway Enrichment . . . . .	39
3.2.7 Community Hub Identification . . . . .	39

3.2.8	Results formatting and visualization	39
3.3	Results	40
3.3.1	Saliva DN	40
3.3.2	Saliva Communities Temporal Visualization	42
3.3.3	Pathway Enrichment of Saliva Communities	43
3.3.4	B Cell DN	46
3.3.5	Pathway Enrichment of B Cell Communities	47
3.3.6	B Cell Communities Temporal Visualization	51
3.3.7	Community Hubs	52
3.3.8	Biological Considerations	52
3.4	Discussion	56
3.4.1	Perturbation Induces Temporal Responses	57
<b>Chapter 4. Mean Field Theory of a Non-Reciprocal Hopfield Network</b>		<b>59</b>
4.1	Introduction	59
4.2	Cyclic Hopfield networks	62
4.2.1	Mean field solution	63
4.2.2	Near cusp dynamics	67
<b>Chapter 5. Critical Cyclic Behavior</b>		<b>70</b>
5.1	Critical Dynamics	70
5.1.1	Near the cusp with $\beta\lambda_- = 0$ and $\beta\lambda_+ \approx 1$	70
5.1.2	Hopf bifurcation line: $\beta\lambda_- \neq 0$ while $\beta\lambda_+ = 1$	72
5.1.3	Limit cycle phase near the Hopf line	73
5.1.4	Fold line	74
5.1.5	Near the fold line	76
5.1.6	External drive	79
<b>Chapter 6. Master Equation and Spectral Analysis</b>		<b>82</b>
6.1	Master Equation	82
6.2	Numerical diagonalization of Liouvillian	85
6.2.1	Exact expectations and correlation functions	86
<b>Chapter 7. Monte Carlo Validation with Glauber Dynamics</b>		<b>89</b>
7.1	Glauber Dynamics	89
7.1.1	Numerical tests of critical behavior	93
7.2	Conclusions	96
7.2.1	Code Availability	98
<b>Chapter 8. Conclusions</b>		<b>99</b>
8.1	Track I: key findings and contributions	99
8.2	Track II: key findings and contributions	101
8.3	Future Directions	102
<b>BIBLIOGRAPHY</b>		<b>105</b>

## LIST OF TABLES

Table 3.1: Reactome pathway enrichment analysis. Statistically significant pathways (FDR < 0.05) are summarized for saliva DN communities C0 and C1. In the full analysis, we omitted small communities with fewer than 8 genes [1], and 12 communities (C0 to C11) qualified for the pathway analysis. . . . .	44
Table 3.2: Reactome pathway enrichment analysis. Statistically significant pathways are summarized for primary B cell DN community C2 and C4. In the full analysis, we omitted small communities with fewer than 8 genes [1], and 14 communities (C0 to C13) qualified for the pathway analysis. . . . .	49

## LIST OF FIGURES

Figure 2.1: Schematic of the Louvain Algorithm . . . . .	24
Figure 2.2: Schematic of Differential Network Construction . . . . .	28
Figure 3.1: <b>Workflow Overview.</b> Our methodology starts with time course experimental data, followed by network construction, differential network determination, community detection, sequential ordering by activation pattern, pathway analyses of individual communities, community hub gene interpretation, and final results including analyses and temporal trend visualizations. . . . .	35
Figure 3.2: <b>Differential network analysis for the saliva experiment.</b> a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C0 (top) and C1 (bottom) communities with red highlights indicating genes found in statistically significant Reactome pathways (FDR $\leq$ 0.05), and their corresponding edges in the network. c) Heatmaps of C0 (top) and C1 (bottom) over 24 hours. Columns represent time points while rows denote gene identifiers. The row data show the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalizing using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (also known as the Farthest Point Algorithm) [2, 3]. . . . .	41
Figure 3.3: <b>Differential network analysis for the B cell experiment.</b> a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C2 (top) and C4 (bottom) communities with red highlights indicating genes found in statistically significant Reactome pathways (FDR $\leq$ 0.05), and their corresponding edges in the network. c) Heatmaps of C2 (top) and C4 (bottom) over 15 hours (6 time points). Columns represent time points while rows denote genes. These row data demonstrate the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalizing using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (Farthest Point Algorithm) [2, 3]. . . . .	48

Figure 4.1:	Phase portraits (left) and the phase diagram (right) for the two-memory cyclic Hopfield model. Left: The panel shows the dynamical behavior for different values of $\lambda_+$ and $\lambda_-$ . The six phase portraits ( <b>a</b> to <b>f</b> ) show the trajectory dynamics, with green arrows indicating the vector fields of the derivatives. In <b>e</b> and <b>f</b> , empty circles represent saddle points, while solid circles denote stable points (sinks). Right: The phase diagram is divided into three regions of different dynamical behavior: Limit Cycles (diagonal lines with a purple background), Memory Retrieval (dotted dark green background), and Paramagnetic (vertical stripes with a light green background). These phases are bounded by bifurcation lines: fold bifurcation lines (dark green lines) and the Hopf bifurcation (purple vertical line). The positions of the six trajectory plots ( <b>a</b> - <b>f</b> ) are indicated by corresponding labels on the phase diagram. . . . .	65
Figure 5.1:	Representative trajectories at the fold transition as a function of $t$ with $N = 1300$ and $D = 1$ . One can notice several features that are absent (deep) in the limit cycle phase: First, there is a larger variation between different trajectories. This highlights the role of noise in inducing phase slips. Second, the interval between jumps is rather long, and scales as $T \sim N^{1/3}$ , in this case roughly of the order $T \sim 10$ compared to deep in the limit cycle phase where it is of order 1. Again this is due to noise as the period should diverge when $N \rightarrow \infty$ . . . . .	75
Figure 5.2:	Effective potential $V(\theta)$ for different values of $\epsilon$ . Depending on its sign, $\epsilon$ alters the steepness in $V(\theta)$ , thereby affecting system's stability. A negative $\epsilon$ initiates an uphill start in $V$ , which poses a potential hill for $\theta$ as a metastable state till it overcomes the hill. For $\epsilon = 0$ , $V$ starts flat, then slips down at a faster rate than in the negative $\epsilon$ scenario. A positive $\epsilon$ triggers an immediate downhill movement in $V$ , swiftly driving the system into the oscillatory phase at an even faster rate. The two horizontal dashed lines mark the 1 <sup>st</sup> local minima of the $\epsilon = -0.5$ and $\epsilon = 0$ cases. . . . .	79
Figure 6.1:	Real and Imaginary part of the second smallest eigenvalue of the Liouvillian matrix, $\Lambda_2$ , as a function of $\beta\lambda_+$ for a fixed value of $\beta\lambda_- = 0.17$ and $N_S = N_D = N/2$ . . . . .	85
Figure 6.2:	Exact $\langle m_2(t) \rangle$ solved from the master equation for different system sizes $N$ with $\beta\lambda_+ = 1.3$ and $\beta\lambda_- = 0.17$ . As $N$ increases, the oscillations become slower and more pronounced. . . . .	88
Figure 6.3:	Two-time correlation function $C_{2,2}(t, \tau)$ for $M_2$ at $t = 15$ , $\beta\lambda_+ = 1.3$ and $\beta\lambda_- = 0.17$ , for various $N$ . As $N$ increases, $C_{2,2}(t, \tau)$ exhibits defined oscillations. . . . .	88

Figure 7.1: Magnetization $m_2(t)$ in Glauber dynamics as it approaches the mean field solution. Each dashed line with markers represents a single realization for system sizes $N = 100$ (blue circles), 1,000 (orange triangles), 10,000 (red crosses), and 100,000 (green diamonds) at $\beta\lambda_+ = 1.3$ and $\beta\lambda_- = 0.17$ , compared to the mean field solution (solid black line). As $N$ increases, the simulations match the mean field predictions, with larger systems nearly overlapping with the mean field curve. . . . .	90
Figure 7.2: Average magnetization, $m_2(t)$ in Glauber dynamics over 1,000 realizations. The $N = 100$ system shows a significant decay within approximately two periods. At $N = 1,000$ , the oscillation initially matches the mean field period but begins to shorten around $t/\tau_0 = 400$ while damping out. A larger system exhibits prolonged oscillation persistence, yet still with a noticeable damping. The largest $N$ (the green curve) approximates an infinite system and more closely recovers the oscillations of the mean field solution. . . . .	91
Figure 7.3: Equivalence of the master equation solution and averages of Glauber dynamics. System parameters and initial conditions are same of those in Fig. 7.1 and 7.2: $\beta\lambda_+ = 1.3$ , $\beta\lambda_- = 0.17$ , $m_1(0) = 1$ , and $m_2(0) = 0$ . The system has 100 spins in each subnet $S$ and $D$ , totaling $N = 200$ . Colored dashed curves represent the averages from Monte Carlo simulation runs. As the sampling size increases, the average trajectory of all stochastic paths converges to the Liouvillian solution. . . . .	92
Figure 7.4: Autocorrelation on the fold line at $\beta\lambda_+ = 1.25$ and $\beta\lambda_- = 0.1025$ . For different values of $N$ . The time axis is scaled according to Eq. 5.21 to show collapsing into a single function. The absolute value of the autocorrelation decays purely exponentially, while its real and imaginary components show underdamped oscillations. . . . .	93
Figure 7.5: $\theta$ rotations following the activation of a constant $F$ for a system of $N = 10^6$ on the fold line with $\beta\lambda_+ = 1.25$ and $\beta\lambda_- = 0.1025$ . Time and angles are rescaled according to Eq. 5.32, which is valid for $\theta \ll 1$ . . . . .	94
Figure 7.6: Hopf critical exponent: $\beta\lambda_+ = 1.0$ , $\beta\lambda_- = 1.7$ . The time axis and autocorrelation are scaled according to Eq. 5.13 to show collapsing into a single function. . . . .	95

Figure 7.7: Characteristic decay times of autocorrelation near the Hopf line and above the Fold line grow linearly with  $N$ , following predictions of Eqs. 5.15 and 5.26. Below the Fold line, the decay time grows exponentially with  $N$ , following Eq. 5.28. Errors in the data are too small to be visible and have been omitted for clarity. The fits for all three regimes indicate strong statistical agreement between the data and the respective fitted models, with  $R^2$  values of 0.969 (near Hopf), 0.980 (above Fold), and 0.993 (below Fold).

# Chapter 1. Introduction

## 1.1 Motivations

### 1.1.1 Adaptive Responses in Complex Systems

Complex systems, particularly biological organisms, face the fundamental challenge of navigating and responding effectively to dynamic and often unpredictable environments. Successful adaptation in such contexts necessitates more than mere reactivity; it demands the capacity for **adaptive responses**. This aligns with cybernetic principles, such as Ashby's Law of Requisite Variety, which posits that a system's internal repertoire of responses must be at least as diverse as the environmental challenges it encounters to maintain stability and function [4].

Living organisms offer tangible illustrations: a cell, for instance, typically maintains homeostasis by correcting minor perturbations, yet it can also undergo profound state changes when presented with specific developmental or environmental signals. This dual capability (i.e., robustness in the face of noise and transformative flexibility in response to specific cues) exemplifies the essential balance underlying adaptive responses, often orchestrated by sophisticated networks of interactions that dynamically adjust to different challenges. Without mechanisms to produce condition-specific actions, a system would either remain rigidly unresponsive to external stresses or become erratically unstable, failing to preserve critical functions. The imperative, therefore, is clear: adaptivity and specificity in responses are fundamental requirements for complex systems, from engineered networks to living cells.

## 1.1.2 Networks and Modularity as an Organizing Principle

Whether the basic units are genes, proteins, or neurons, the organizational backbone of these complex adaptive systems is typically a network of interacting components. A central tenet in the study of such systems is that a network’s architecture, its structure, shapes its dynamics and functional capabilities. One architectural principle observed across diverse complex systems—from biological networks and engineered systems like AI neural networks [5] to social and technological infrastructures—is **modularity**. In a broad sense, modularity refers to a pattern of organization where elements are grouped into subsets that are relatively autonomous yet interacting; these connections can be physical (e.g., protein-protein interactions), dynamical (e.g., in gene regulatory networks), or statistical (e.g., correlations among phenotypic traits leading to variational modules) [6]. The concept of modularity is contextually defined. For example, a biological network may involve variational, functional, or developmental modules. The specific type depends on the system and the processes being considered [6]. This organization echoes Herbert Simon’s insights on nearly-decomposable systems, which, due to the semi-independence of their modules, can adapt more rapidly and efficiently to changing environments [7].

Modularity is widely believed to confer significant functional advantages. It facilitates **functional specialization**, so different parts of the network can perform particular tasks. It enhances **robustness**, as perturbations within one module are less likely to catastrophically affect the entire system because their impact tends to be localized. And it promotes **evolvability**—the capacity to generate adaptive phenotypic variation—by allowing modules to be refined, recombined, or repurposed with reduced interference to other functions [8, 9, 10, 5]. Crucially for this dissertation, modularity provides a organizational basis for **differential**

**collective behaviors**, where different subunits (modules) of the network exhibit unique responses yet together they form a coordinated outcome contributing to an overarching system-wide dynamic.

### 1.1.3 Differential Modular Responses in Biological Networks

In biological systems, particularly at the molecular level of gene regulatory networks, observing and understanding these differential modular responses is paramount for deciphering cellular logic and disease mechanisms. Gene networks must orchestrate precise reactions to a vast array of signals. Deciphering how these intricate networks achieve such specificity from high-throughput omics data presents a significant challenge in systems biology [11]. This dissertation first addresses this challenge through the application of **Differential Network (DN) Analysis** (see Chapters 2 and 3). Unlike traditional differential gene expression analysis, which focuses on individual genes changing in isolation, DN analysis shifts the focus to the level of interactions, seeking to identify stimulus-specific alterations in connectivity [12, 13]. Importantly, this approach allows for the identification of functional modules or **communities that exhibit differential activity patterns** in response to particular perturbations, such as specific activation, repression, or distinct temporal profiles. As detailed in Chapters 2 and 3, this methodology reveals how biological systems, when faced with stimuli, mobilize distinct, interacting sets of genes that form functional communities. These communities, jointly, manifest patterns of sequential activation, thereby empirically substantiating the strategy of using a modular architecture to generate differentiated responses.

### 1.1.4 From Observation to Mechanism

The data-driven approaches such as DN analysis are powerful for documenting *what* changes within a network and revealing *that* these systems deploy differential modular responses. Nevertheless, they often leave open the question of *how and why* those changes arise from the underlying system dynamics. There remains a critical gap in mechanistic understanding: we may observe that certain modules respond differentially, but we do not fully understand the governing principles or dynamic mechanisms that drive those modular responses or enable their selective activation. For instance, how do distinct functional modules (communities of genes) coordinate their activities or influence one another to produce an observed systemic response to stress? What interacting dynamics rules govern the flow of the localized responses from one module to another, eventually cascading an effect across the entire network? These questions, which probe the causal link between modular organization, inter-module interactions, and the emergence of specific macroscopic adaptive responses, cannot be answered by observational network analysis alone. They necessitate integrative approaches that connect data-driven findings with theoretical models capable of exploring the factors that determine network behavior and enable such differentiated responses [14]. To understand these factors, it is insightful to study models with internal organization. These models either inherently possess or can develop functionally distinct component groups. This feature allows them to jointly distinguish macroscopic states or selectively process inputs. In doing so, they move beyond classical models that often assume homogeneity.

### 1.1.5 A Non-Reciprocal Hopfield Network for Modular Dynamics

To this end, the second part of this dissertation (see Chapters 4 through 7) investigates a **non-reciprocal Hopfield network** as a theoretical paradigm for studying differential dynamics and state distinction. Hopfield networks are canonical models of associative memory, inherently designed to perform *differential tasks* by distinguishing between, and retrieving, stored patterns [15, 16, 17]. The focus on a *non-reciprocal* (asymmetric) variant is instrumental because **asymmetric interactions, where the influence of component A on B is not equal to the influence of B on A, are highly prevalent and functionally significant in real-world biological systems**, such as group dynamics in active matter [18, 19], gene regulatory circuits and neural networks [20, 21], often dictating the flow of information and control.

Non-reciprocity can drive systems out of equilibrium and lead to a rich repertoire of dynamical behaviors. For instance, the limit cycles and sustained oscillations explored in this work are not typically found in their symmetric counterparts[22, 23, 24]. Within our proposed model, internal differentiation is explicitly defined through **similarity (S) and differential (D) subnetworks**, which are based on the structure of the encoded memory patterns. These subnetworks, and their distinct contributions to the system’s dynamics, are fundamental to the model’s ability to retrieve specific memories or settle into different dynamical attractors. Building on the earlier discussion about modularity, the S and D subnetworks function as effective modules with strong internal coupling and a relatively weaker, non-reciprocal interaction between them. This non-reciprocity is the key ingredient driving sequential changes in activity patterns, resembling those observed in the empirical DN analyses. Our theoretical investigation explores how parameters such as noise, interaction

strengths, external drives, and the degree of non-reciprocity govern the model’s capacity for these diverse and distinct dynamical behaviors. These parameters shape the model’s dynamical evolution and possible transitions between distinct states.

While our Hopfield model is fully connected and does not possess an explicit, pre-defined modular hierarchy in the same vein as some biological networks, understanding how even such fundamental internal differentiations (S/D subnetworks) impact dynamics provides a crucial theoretical baseline. Moreover, these insights are foundational for understanding systems with more explicit modular architectures, as studies have demonstrated that the dynamics of Hopfield-like attractor networks are indeed significantly influenced when implemented on modular graphs, akin to those observed in biological systems ([17], Ch 6.1).

### **1.1.6 Dissertation Overview and Integrative Perspective**

Therefore, this dissertation explores modular interactions in complex systems from two complementary perspectives, biological and theoretical. Track I employs a data-driven graph analysis to RNA-seq data from PPSV23 vaccination and Rituximab treatment. It detects functional modules and traces their sequential activation, suggesting that inter-module couplings drive the biological response. Track II explores these couplings in a controlled setting: a non-reciprocal Hopfield network with two predefined subsystems is examined under varied parameter settings. The comparison links empirical observations to mechanism, showing how intra- and inter-module interactions tune global network behavior.

## 1.2 Dissertation Outline

The dissertation is organized along two research tracks that can each stand alone in terms of methodology, results, and interpretation. Each track — Differential Networks (Chapters 2–3) and Hopfield spin models (Chapters 4–7) — can be read independently.

**chapter 2, Foundations of Differential Network Analysis**, first lays out the fundamentals of network science, including graph representations, centrality measures, and the concept of community structure. This chapter then briefly surveys community detection algorithms. Particular focus is given to the Louvain heuristic — its efficiency and scalability were notable at the time of the research project done in this dissertation, although contemporary algorithms like Leiden are now often preferred. The chapter proceeds to define the differential network (DN) formalism and explain its algebraic construction by contrasting condition-specific graphs. Finally, this chapter demonstrates how topological descriptors, such as degree centrality and community structure, can be reapplied to these DNs to reveal stimulus-specific rewiring patterns.

**chapter 3, Application of Differential Network Analysis to Salivary and B-Cell Transcriptomic Data**, deploys the DN pipeline on two longitudinal RNA-seq datasets: (i) human saliva sampled hourly for 24 h before and after PPSV23 vaccination, and (ii) primary B cells tracked for 15 h with and without Rituximab. The analysis reveals temporally ordered gene communities, pathway activations consistent with vaccine and drug mechanisms, and hub genes (e.g. *IL4R*, *PELI*) linked to immune modulation.

**chapter 4, Mean-Field Theory of a Non-Reciprocal Hopfield Network**, proposes a Hopfield spin model with asymmetric couplings that generate cyclic retrieval. Mean-field analysis reveals three dynamical regimes — paramagnetic, memory-retrieval, and limit-cycle —

delimited by Hopf and fold bifurcations. It also provides criteria for fixed-point stability and for the amplitude of the emergent cycle.

**chapter 5, Critical Cyclic Behaviour**, exploits center-manifold analysis and stochastic Langevin equations to obtain scaling exponents for autocorrelation and response functions on the Hopf and fold critical lines. Finite-size effects (exponent  $\zeta = 1/2$  on Hopf;  $\zeta = 1/3$  on fold) are derived analytically as well. Furthermore, the chapter investigates the system's response to external drives, revealing distinct behaviors at each critical line: enhanced resonant sensitivity and slower response scaling ( $|F|^{-2/3}$ ) on the Hopf line, versus controlled state-switching with faster response ( $|F|^{-1/2}$ ) on the fold line.

**chapter 6, Master Equation and Spectral Analysis**, recasts the model in an operator (Liouvillian) framework. Exact diagonalisation yields system-size-dependent relaxation spectra and closed-form two-time correlation functions, thereby linking microscopic spin-flip kinetics to macroscopic observables. The analysis of the Liouvillian spectrum provides insights into the system's relaxation dynamics across different phases and near critical boundaries, complementing the mean-field predictions of critical slowing down at bifurcations like the Hopf.

**chapter 7, Monte-Carlo Validation with Glauber Dynamics**, reports large-scale simulations that confirm the analytic laws, map the damping of limit-cycle coherence at finite  $N$ , and quantify noise-induced phase-slip statistics. These numerical validations are crucial for establishing the model's relevance to understanding noise and finite-size effects in systems exhibiting complex oscillatory dynamics, such as those found in biology. All code is openly available at <https://github.com/shuyue13/non-reciprocal-Hopfield>.

**chapter 8, Conclusions**, summarizes the dissertations' contributions, discusses their implications for systems biology and theoretical neuroscience, and outlines future research

directions, including the prospect of a unifying framework linking differential networks and non-reciprocal Hopfield dynamics.

# Chapter 2. Foundations of Differential Network Analysis

## 2.1 Introduction to Network Science in Biological Systems

Born from foundational questions in graph theory, the study of networks has evolved into a formal and dynamic field for understanding complexity in modern science. The oldest conceptual root can be traced to Euler's 1736 solution to the Königsberg bridge problem, which often considered the first result in graph theory and thereby laid the ground for representing pairwise relations. While originated from graph theory, the modern network science was crystallized in the late 1990s and early 2000s [25]. In its formative stage, statistical physics pioneered the field and shaped the landscape of network theory [26, 27, 28, 29, 30, 31, 32]. The major progress during this period centers on universal structural properties in complex networks, such as small-world topology [33], scale-free degree distributions [34], modular organization [32], heavy-tailed clustering distributions, and assortative or disassortative mixing patterns [35, 36], etc.

As the field matured, the focus expanded to how topological structures influence dynamic processes unfolding on networks. This structural-dynamic coupling [37, 38] has illuminated phenomena such as percolation transitions, critical behavior in the Ising model, diffusion and synchronization dynamics, epidemic spreading, and control theory in complex systems [39, 40]. In parallel, the principle of maximum entropy, formulated by Boltzmann and Gibbs, provided a rigorous statistical framework for inferring probabilistic ensembles of networks, contributing

to both null models and generative models grounded in thermodynamic reasoning [41, 42].

More recently, the proliferation of massive data in biology, social science, and information technology has drawn researchers from a wide range of disciplines. This convergence has transformed network theory into a multidisciplinary field with sustained contributions from physics, biology, computer science, engineering, and the social sciences [25]. Today, network science offers a flexible formalism for modeling complex systems, where entities are abstracted as nodes (vertices) and their relationships as edges (links) [43]. This abstraction supports a wide range of mathematical and computational tools for probing structural organization, relational architecture, and collective behavior in systems composed of many interacting elements.

In biological systems, network representations manifest across multiple scales of organization [16]. At the molecular level, intricate webs of interactions govern cellular function: genes exert regulatory control over one another through complex transcriptional regulatory circuits [44]; proteins assemble into functional complexes to execute molecular processes such as signal transduction and enzymatic catalysis; and metabolites participate in elaborate biochemical pathways [17, 45]. Ascending to cellular and tissue levels, signaling cascades orchestrate the propagation of information, whereas intercellular communication networks coordinate collective responses to environmental cues or developmental signals [46]. Abstracting these multifaceted interactions into network structures enables the application of a unified analytical toolkit across remarkably diverse biological contexts. This fosters a deeper insights into systemic behaviors underlying biological complexity [43].

Once a complex system is framed as a network, one can examine its structural and dynamical details through the lens of mathematical and topological properties established in the graph theory. For example, centrality, clustering, modularity, connectivity, and network

flow - topological measures widely used in network quantification - are commonly applied to the constructed network representation to identify critical element, detect organizational structure, characterize robustness, or map interaction pathways in the underlying system. In biological systems, these quantities have revealed key regulatory hubs by centrality analysis, disease-associated modules through clustering and modularity detection, structural vulnerabilities from connectivity assessments, and transcriptional or metabolic routing via flow-based analyses [43, 17, 47, 44].

With the advent and widespread accessibility of high-throughput technologies, enormous sophisticated network-based computational methods have been developed in the past few decades, to extract and predict biological functions from large-scale omics data. These newer approaches extend beyond classical topological descriptors and are often tailored to probe specific biological aspects. Among this diverse landscape of modeling strategies, where each tuned to answer different biological questions, one major topic focuses on detecting alterations in network connectivity driven by external perturbations. Particularly valuable in this context is Differential Network (DN) analysis.

Rather than relying on a single static network, DN constructs a new network by contrasting two state-specific networks, typically a control versus perturbed condition. Specifically, by comparing the networks, it aims to filter out baseline or constitutive interactions and isolate interaction changes specifically induced by the stimulus. Thereby, the resultant network can highlight the system's adaptive response, e.g., signaling or regulatory pathways triggered by drug treatments or environmental stressors. This derived network can be further characterized using conventional graph-theoretic measures [13, 12, 48].

The DN analysis has gained wide adoption within systems biology and genomics research over the past decade [49, 50, 51]. Compared to earlier approaches such as differential expression

analysis identify individual genes with altered abundance, i.e., differentially expressed genes, the DN shifts focus toward quantifying changes in the relationships between those genes [49, 52]. This transition enabled the discovery of condition-specific regulatory switches, compensatory signaling, and module restructuring that are often inaccessible via single-gene analyses [52, 53].

When applied to time-resolved data, DN analysis acquires an additional temporal dimension, allowing one to track dynamic reorganization of networks as biological systems respond and adapt over time [54]. This temporal extension is essential for identifying coordinated gene modules exhibiting synchronized activation, and for inferring sequential logic of molecular events following perturbations [13]. Familiarity with these concepts and their application equips one with an understanding of how the DN framework can be leveraged to investigate system-wide dynamics, as further explored in the subsequent chapters.

## 2.2 Fundamentals of Network Theory

### 2.2.1 Basics Definitions and Representations

A network (or graph)  $G = (V, E)$  consists of a set of nodes  $V$  and a set of edges  $E$  connecting pairs of nodes. Networks can be directed or undirected, weighted or unweighted. Directed networks possess edges with assigned directions, often used to represent regulatory or signaling pathways. Weighted networks attribute a numerical value (weight) to each edge, capturing interaction strength or correlation magnitude.

Networks are commonly represented using adjacency matrices. For a graph with  $n$  nodes, the adjacency matrix  $A \in \mathbb{R}^{n \times n}$  is defined such that each entry  $A_{ij}$  reflects the presence and,

if applicable, the weight of the edge from node  $i$  to node  $j$ . For an unweighted, undirected graph:

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between node } i \text{ and } j \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

For undirected networks,  $A$  is symmetric ( $A_{ij} = A_{ji}$ ), while for directed networks, asymmetry is allowed ( $A_{ij}$  represents an edge from  $i$  to  $j$ ).

In weighted networks, the adjacency matrix entry  $A_{ij} = w_{ij}$  directly quantifies the strength, intensity, or confidence of the interaction between nodes  $i$  and  $j$ , regardless of whether a thresholded existence condition is applied. In such cases,  $w_{ij} \in \mathbb{R}$  and may assume values based on the domain-specific interpretation of connection weights. For instance, in social or communication networks,  $w_{ij}$  could represent message frequency or mutual affinity; in biological networks, it often reflects measures like gene co-expression, mutual information, or interaction likelihood [55, 56].

*Weights* can encode various types of pairwise relationships, depending on the modeling context:

- **Similarity:** Such as Pearson correlation between gene expression profiles.
- **Influence or Flow:** For example, regulatory impact scores in directed biological networks.
- **Statistical Dependence:** Including mutual information or partial correlation [?].

A specific and biologically meaningful subclass of weighted networks is the gene co-expression network. In such networks, each node represents a gene, and the edge weight encodes the correlation between expression profiles of gene pairs across a population of

samples. The entries of the adjacency matrix for a correlation-based network are thus given by:

$$A_{ij} = \rho_{ij}, \tag{2.2}$$

where  $\rho_{ij} \in [-1, 1]$  is the Pearson or Spearman correlation coefficient between the expression levels of gene  $i$  and gene  $j$  [57]. In practice, a threshold is often applied to  $|\rho_{ij}|$  to exclude weak or spurious associations, resulting in a sparse weighted network.

Gene co-expression networks provide a functional mapping of the transcriptional landscape. Rather than encoding physical binding or direct regulation, these networks capture statistical dependencies that often reflect shared pathways, co-regulation, or coordinated cellular roles. This abstraction enables the identification of gene modules, inference of regulatory programs, and characterization of system-wide perturbation responses.

### 2.2.2 Network Centrality: Identifying Important Nodes

Network centrality quantifies the relative importance or influence of a node within a network's topology. Among the most commonly used centrality measures is **degree centrality**, defined for a node  $v$  as

$$C_D(v) = \text{deg}(v) = \sum_j A_{vj}, \tag{2.3}$$

which simply counts the number of edges incident to  $v$ . In biological networks, high-degree nodes often represent regulatory hubs or multifunctional proteins.

**Betweenness centrality** captures the extent to which a node lies on the shortest paths

connecting other pairs of nodes in the network. Formally:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}, \quad (2.4)$$

where  $\sigma_{st}$  denotes the number of shortest paths between nodes  $s$  and  $t$ , and  $\sigma_{st}(v)$  counts how many of those paths pass through node  $v$ . A high betweenness centrality indicates that a node plays a crucial role as an intermediary or bottleneck for information flow.

While there are other centrality metrics, such as closeness centrality, which measures the average distance from a node to all others, or eigenvector centrality, which reflects influence through connectedness, in this work we focus primarily on degree-based hub detection, as it proves particularly informative for identifying hub genes in response to biological stimuli.

### 2.2.3 Community Structure

In network science, a community (or module) refers to a subset of nodes that are more densely connected to each other than to the rest of the network. The presence of communities is a topological property commonly observed in complex networks [58, 59, 32, 60]. These clusters suggest a degree of internal cohesion or functional affinity among their members, reflecting underlying organization and function of complex systems. In biological systems, nodes in the same community typically share functional relationships [46, 61, 62, 63, 64].

While the qualitative notion of communities is intuitive, in practice, uncovering these mesoscopic structures within complex networks requires formalization. This necessitates both quantitative metrics to assess the strength of community structure and the development of algorithms to partition networks based on various principles.

## 2.2.4 Community Detection Algorithms

Several algorithms have been developed for community detection, each with distinct advantages:

- **Hierarchical Clustering:** This approach builds nested communities by either progressively merging smaller communities (agglomerative) or dividing larger ones (divisive). The method developed by Girvan and Newman [58] iteratively removes edges with high betweenness centrality to reveal natural community divisions.
- **Spectral Methods:** These techniques use the eigenvectors of matrices derived from the network (such as the Laplacian matrix) to partition nodes into communities. The eigenvectors provide a continuous relaxation of the discrete clustering problem.
- **Statistical Inference:** These approaches frame community detection as a statistical inference problem, seeking the most likely partition given a generative model for the network, such as the stochastic block model.

The identification of community structure represents a fundamental task in network analysis [45, 32] - partition the network into clusters that maximize internal connectivity while minimizing external connections. The challenge lies in the fact that there is no single, universally accepted definition of a community, leading to the development of a diverse array of detection algorithms rooted in different principles and heuristics [32]. Several major approaches have emerged:

- **Hierarchical Clustering:** These methods construct a hierarchy of partitions, often represented as a dendrogram. Agglomerative approaches start with individual nodes

as communities and iteratively merge the most similar pairs or groups, while divisive approaches start with the entire network and recursively split it [65]. Similarity or distance metrics guide the merging/splitting process.

- **Spectral Methods:** Leveraging the spectral properties of graph matrices, these algorithms utilize the eigenvectors and eigenvalues of matrices such as the graph Laplacian or the adjacency matrix to partition nodes into distinct groups [66]. Often, the eigenvectors provide a low-dimensional embedding of the nodes where clusters become more apparent, relating the problem to finding informative 'cuts' in the network.
- **Statistical Inference:** This category frames community detection as an inference problem, seeking the partition that best fits a generative model of network formation. The most prominent example is the Stochastic Block Model (SBM), which assumes nodes belong to hidden blocks (communities) and the probability of an edge depends only on the block memberships of the involved nodes [67, 68]. Algorithms then aim to find the most likely assignment of nodes to blocks given the observed network structure, often via maximum likelihood or Bayesian methods [42].
- **Optimization of Quality Function:** Another major class of algorithms starts by defining a scalar quality function that measures how good a given partition is, and then searching the space of possible partitions for one that optimizes this function. The most widely adopted quality function is *modularity*, typically denoted as  $Q$  [45]. Modularity quantifies the fraction of edges within communities compared to the expected fraction if edges were placed randomly, preserving node degrees [66]. Positive and large values of  $Q$  are indicative of strong community structure. The modularity  $Q$  is discussed in Section 2.2.4.1,

Each of these fundamental principles - hierarchical clustering, spectral analysis, statistical inference, and quality function optimization - has given rise to a multitude of specific algorithms, each with its strengths and computational considerations. A comprehensive review of all categories extends beyond the scope of this chapter [32]. Given its relevance to subsequent analyses in this thesis, we will focus primarily on methods derived from the principle of modularity maximization.

A variety of algorithmic approaches have been developed to yield high values of  $Q$ . Notable examples include the spectral partitioning method that specifically uses the eigenvectors of the *modularity matrix* (as distinct from the Laplacian) [66], the widely used fast, greedy agglomerative algorithm by Clauset, Newman, and Moore (CNM) that iteratively merges communities providing the largest increase in  $Q$  [65], and other heuristic approaches such as extremal optimization techniques [69]. These algorithms offer differing trade-offs between accuracy, resolution, and computational complexity, but all share the underlying objective of optimizing modularity to reveal network communities. The Louvain method, another popular greedy modularity maximization algorithm used later in this work, also falls into this category [70].

#### 2.2.4.1 Modularity

Modularity, often denoted as  $Q$ , is a scalar metric to quantify the quality of a particular division of a network into communities. It assesses whether the number of edges falling within groups of nodes (communities) is significantly higher than what one would expect under a random network model that preserves the node degrees. In essence, modularity evaluates the deviation of the observed edge distribution from that of a null configuration model.

Formulated by Newman and Girvan [45], modularity is defined as:

$$Q = \frac{1}{2m} \sum_{i,j} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \quad (2.5)$$

where:

- $A_{ij}$  is the (weighted) adjacency matrix,
- $k_i = \sum_j A_{ij}$  is the strength (or degree, for unweighted graphs) of node  $i$ ,
- $m = \frac{1}{2} \sum_{i,j} A_{ij}$  is the total edge weight in the network,
- $c_i$  represents the community assignment of node  $i$ ,
- $\delta(c_i, c_j)$  is the Kronecker delta function, equal to 1 if nodes  $i$  and  $j$  are in the same community, and 0 otherwise.

The term  $\frac{k_i k_j}{2m}$  corresponds to the expected weight of the edge between  $i$  and  $j$  under the null model. Therefore, a high value of modularity ( $Q$  approaching 1) indicates that the observed network has significantly more intra-community edges than would be expected at random, thus revealing a pronounced modular structure. However, values of  $Q$  typically remain well below 1 in real networks; values above 0.3 or 0.4 already signify nontrivial modular organization [65].

Modularity maximization forms the basis of a broad class of community detection algorithms. These methods seek the partitioning that yields the highest modularity value, when the network is divided into regions with dense internal connectivity and sparse external ties. As an objective function, modularity is both visually inspectable and computationally tractable, and enables comparison between different partition configurations. Among the

most prominent algorithms leveraging modularity maximization are the Louvain method[70], which applies a greedy hierarchical approach, and spectral optimization techniques[59] that utilize eigenvectors of the modularity matrix.

The significance of modularity extends far beyond theoretical appeal. In biological systems, nodes in the same community typically share functional relationships [46, 61, 62, 63, 64]. As noted by Hartwell *et al.* [46], biological networks exhibit inherent modularity, with gene modules, protein complexes, and signaling pathways forming the fundamental units of cellular operations [58]. This structural organization enables both robustness and evolvability: modules can maintain internal stability while adapting semi-independently to external pressures[44].

Such modularity is particularly advantageous in fluctuating environments, where localized rewiring or adaptation must occur without globally compromising system integrity [71]. It also enhances the capacity for noise buffering and specificity in regulatory processes [17]. In genomic applications, detecting these modules provides a pathway to interpret gene co-expression patterns, identify coordinated regulatory feedback loops, and elucidate the molecular basis of disease states or treatment responses. The alignment between topological modules and annotated biological functions (e.g., Gene Ontology terms or KEGG pathways) further corroborates the biological relevance of modular organization [57].

The identification and quantification of community structure through the lens of modularity are one of central algorithms to systems-level analysis in biology. They offer a principled method to distill complex networks into interpretable, functionally coherent substructures.

## 2.2.5 The Louvain Algorithm

The Louvain algorithm, introduced in 2008 by Blondel et al. [70], became a highly influential and computationally efficient method for community detection, particularly suited for large-scale networks. Motivated by the need to overcome the scalability limitations and suboptimal convergence of prior modularity maximization techniques, Louvain employs a hierarchical greedy strategy capable of revealing community structures across multiple resolutions. The algorithm proceeds iteratively through two core phases:

1. **Local Modularity Optimization:** Initially, every node constitutes its own community.

Subsequently, for each node  $i$ , the algorithm evaluates the modularity gain ( $\Delta Q$ ) resulting from relocating  $i$  to the community of each of its neighbors. This gain,  $\Delta Q$ , represents the change in overall network modularity resulting from moving node  $i$  from its current community  $C_{old}$  to a neighbor's community  $D$ . Blondel et al. provide an efficient way to calculate this by comparing the change in internal community edge weights relative to the expected weights based on node degrees [70]. A positive  $\Delta Q$  indicates the move improves the community structure according to the modularity measure. More formally, the change in modularity  $\Delta Q$  when moving an isolated node  $i$  into a community  $D$  can be calculated as [70]:

$$\Delta Q = \left[ \frac{\Sigma_{in,D} + 2k_{i,in,D}}{2m} - \left( \frac{\Sigma_{tot,D} + k_i}{2m} \right)^2 \right] - \left[ \frac{\Sigma_{in,D}}{2m} - \left( \frac{\Sigma_{tot,D}}{2m} \right)^2 - \left( \frac{k_i}{2m} \right)^2 \right]$$

where the terms are defined as follows:

- $\Sigma_{in,D}$ : Sum of the weights of the links inside community  $D$  (before the potential move of  $i$ ).

- $k_{i,in,D}$ : Sum of the weights of the links from  $i$  to nodes currently in  $D$ .
- $\Sigma_{tot,D}$ : Sum of the weights of the links incident to nodes in  $D$  (before the potential move of  $i$ ).
- $k_i$ : Sum of the weights of the links incident to node  $i$ .
- $2m$ : Sum of all link weights in the network.

In essence, this calculation assesses whether the increased connection strength of node  $i$  within the potential new community  $D$  outweighs the loss of connection strength within its original community, relative to what would be expected in a random network configuration based on node degrees. The algorithm assigns node  $i$  to the neighboring community  $D$  providing the maximum positive  $\Delta Q$ ; if no such move exists (i.e., no positive  $\Delta Q$  is found), the node remains in its original community. This procedure is repeated iteratively for all nodes, processed in a random order, until no further relocation yields a positive modularity gain, eventually signifying convergence to a local modularity optimum. The use of a random processing order implies that different executions might yield slightly varied, albeit typically congruent, community structures, due to the heuristic nature of the algorithm.

2. **Community Aggregation:** Upon convergence of the local optimization phase, a new, coarse-grained network is constructed. In this network, the nodes represent the communities identified in the preceding phase. The weight of an edge between two community-nodes corresponds to the sum of edge weights connecting the constituent nodes of these respective communities in the original network. Self-loops represent the summed internal edge weights within each community.

This two-phase process is then recursively applied to the aggregated network. This

iterative refinement and aggregation generates a hierarchical decomposition of the network, often represented as a dendrogram, where each level corresponds to a partition maximizing modularity locally. The algorithm terminates when an iteration yields no further increase in modularity for the entire network. Figure 2.1 provides a schematic representation of this iterative process.

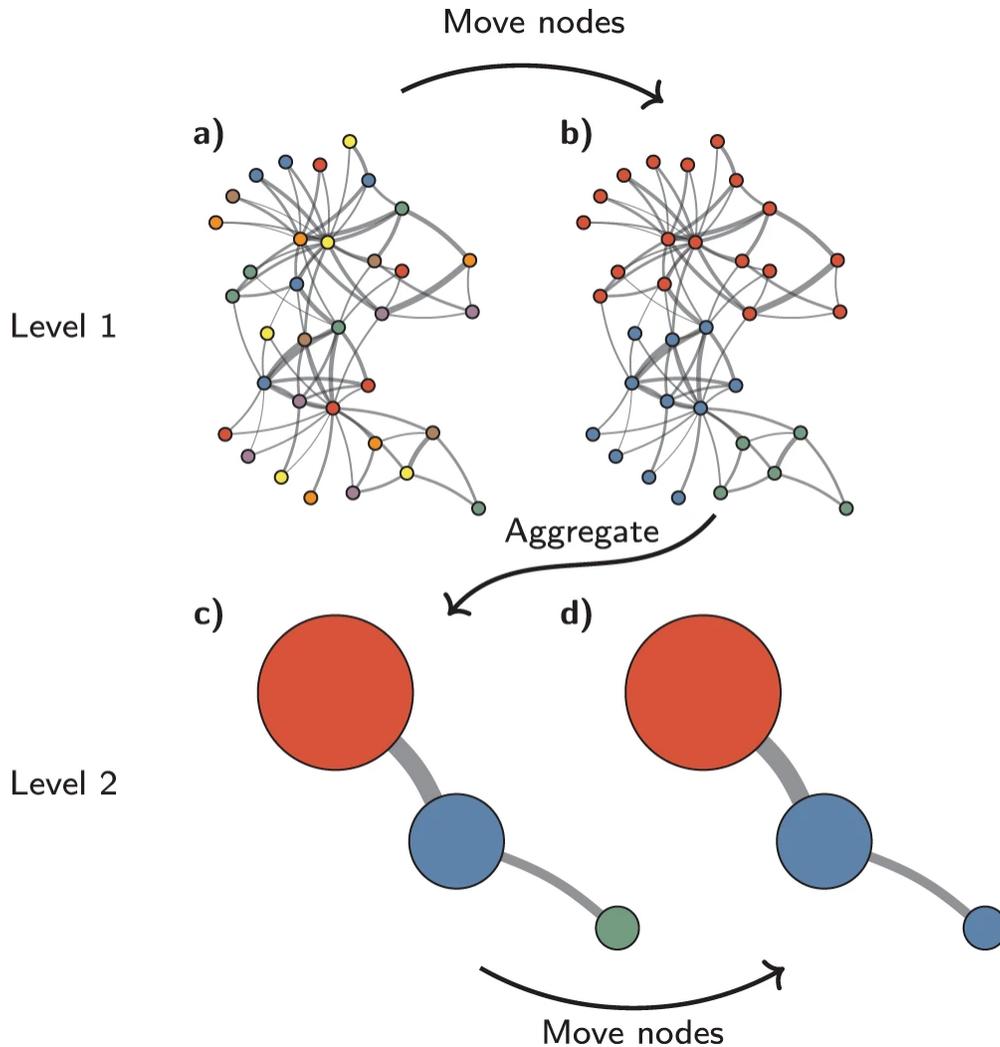


Figure 2.1: Schematic overview of the Louvain algorithm’s iterative process. (a-b) Phase 1: Nodes, initially in singleton communities, are moved between communities to optimize modularity gain. (c) Phase 2: Communities from the optimized partition are aggregated into super-nodes forming a new network. (d) Phase 1 repeats on the aggregated network. The cycle continues until convergence. Figure adapted from Traag et al. [72].

Renowned for its computational efficiency and strong empirical performance [70], Louvain’s

iterative local refinement and aggregation strategy yields high-modularity partitions at relatively low computational expense. Consequently, it was a widely adopted heuristic for community detection in large networks, particularly during the period from its introduction in 2008 until approximately the time pertinent to this study (circa 2019-2020). It is noteworthy, however, that subsequent developments in the field led to algorithms like Leiden [72], designed to address certain limitations of Louvain, such as the potential formation of poorly connected communities. The capacity for modularity-based community detection enables the parsing of intricate biological networks, often characterized by considerable scale and inherent noise, into potentially meaningful functional units, pathways, or groups of co-regulated genes. In subsequent sections, we present its application to genomic data.

### 2.2.5.1 Computational Limitations

While algorithms like Louvain offer efficient heuristics for modularity maximization, the approach itself is not without recognized limitations. Chief among these is the resolution limit problem [73], where the optimization process may fail to resolve smaller communities if merging them yields a larger increase in global modularity. This can obscure finer structural details, particularly within large networks. Furthermore, the modularity landscape itself is often rugged, possessing numerous local maxima [74]. Greedy algorithms like Louvain, while fast, are susceptible to converging to these suboptimal solutions (local maxima). This convergence can result in inconsistent partitions across different runs or yield communities that are especially poorly connected internally, potentially even becoming disconnected in some cases [72].

These challenges have motivated further algorithmic development. The Leiden algorithm, for instance, was introduced as a direct advancement over Louvain [72]. Although

still fundamentally optimizing modularity (or related functions), Leiden incorporates an intermediate refinement phase designed explicitly to guarantee that detected communities are well-connected, splitting groups that Louvain might have erroneously kept together. It also employs a more efficient node movement strategy ('fast local move') intended to better explore the partition space and reduce susceptibility to sharp local optima. Its guarantee of well-connected communities, addressing one key drawback of Louvain, has led to significant interest and adoption in bioinformatics applications. Leiden is particularly noted, at least as of the writing of this dissertation, in fields like single-cell transcriptomics, where discerning subtle cellular subpopulations is crucial. Indeed, ongoing research continues to seek methods that optimally balance computational tractability with the accurate detection of meaningful community structures in complex networks.

## 2.3 Differential Network (DN) Analysis

Differential Network (DN) analysis moves beyond studying single static network snapshots to characterizing how network connectivity patterns are altered between different conditions, states, or time points. In doing so, this approach is particularly effective in biological contexts for isolating perturbation-induced rewiring from the baseline network structure [12].

### 2.3.1 Concept and Formulation

The core idea is to compare two networks representing the system under different conditions, typically a control state and a state with perturbation. Let  $G^{control} = (V, E^{control})$  and  $G^{perturbed} = (V, E^{perturbed})$  be two networks defined on the same set of nodes  $V$ . These networks are often constructed from experimental data, such as gene co-expression networks

derived from transcriptomic profiles under control and perturbed conditions. Let  $A^{control}$  and  $A^{perturbed}$  be their respective (weighted) adjacency matrices.

As described in our work [13], one common way to define the Differential Network  $D = (V_D, E_D)$  focuses on identifying interactions gained in the perturbed condition relative to the control. The edge set  $E_D$  is then defined as the set difference between the edge sets of the perturbed and control networks:

$$E_D = E^{perturbed} \setminus E^{control} \quad (2.6)$$

This means an edge  $(i, j)$  exists in the differential network  $D$  if and only if that edge exists in  $E^{perturbed}$  but is absent from  $E^{control}$ . Figure 2.2 illustrates this process.

The determination of edge presence in  $E^{perturbed}$  and  $E^{control}$  typically involves applying a significance threshold (e.g., to correlation coefficient  $\rho_{ij}$ ) and its associated p-values) derived from experimental data under each condition [13]. Thus, an edge  $(i, j)$  is included in  $E_D$  if it is deemed significant under perturbation (e.g.,  $|\rho_{ij}^{perturbed}| \geq \tau$ ) but not significant under control (e.g.,  $|\rho_{ij}^{control}| < \tau$ ).

The node set  $V_D$  consists of all nodes incident to at least one edge in  $E_D$ . Nodes that become isolated after this edge removal process (i.e., nodes whose only connections were either present in both networks or only in the control network) are typically excluded from  $V_D$ :

$$V_D = \{i \in V \mid \exists j \text{ such that } (i, j) \in E_D\} \quad (2.7)$$

If the original networks are weighted, the weights for the edges in  $E_D$  can be directly inherited from the perturbed network,  $w_{ij}^D = A_{ij}^{perturbed} \forall (i, j) \in E_D$ .

This formulation isolates the network connections that emerge or significantly strengthen

## Differential Network Construction

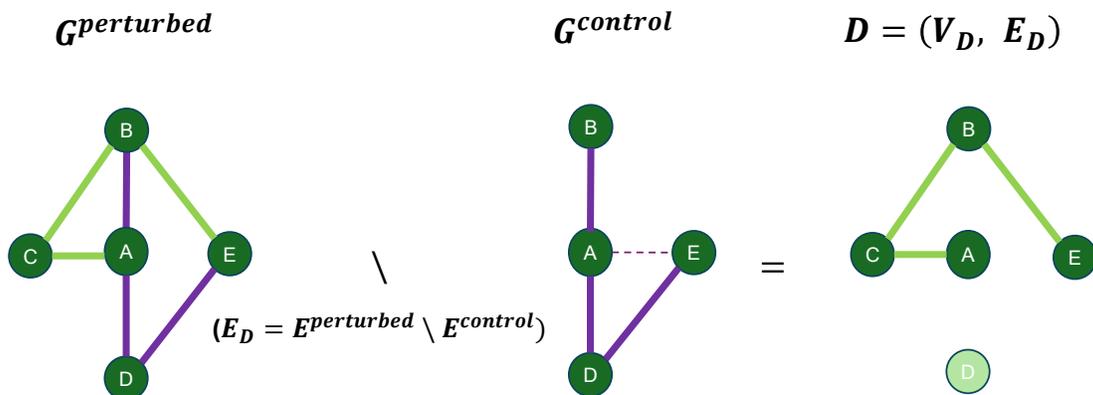


Figure 2.2: Differential Network (DN) construction using the set difference approach. Edges present in the perturbed network ( $G^{perturbed}$ ) but absent from the control network ( $G^{control}$ ) form the edges ( $E_D$ ) of the resulting differential network ( $D$ ). Nodes not incident to any differential edges are excluded from  $V_D$ .

specifically due to the perturbation. It effectively filters out constitutive or baseline interactions present in the control state.

(Note: Another common approach defines the DN via direct subtraction of adjacency matrices,  $D_{matrix} = A^{perturbed} - A^{control}$ . The entries  $D_{ij}$  then represent the change in edge weight, capturing gained ( $D_{matrix,ij} > 0$ ), lost ( $D_{matrix,ij} < 0$ ), and altered interactions. While powerful, this approach requires careful handling of positive and negative edge weights in subsequent analyses [49]. The set-difference approach used in [13] simplifies this by focusing explicitly on gained interactions).

### 2.3.2 Topological Analysis of Differential Networks

Once the differential network  $D = (V_D, E_D)$  is constructed using the chosen formalism (e.g., the set-difference approach focusing on gained edges), the derived network can then

be analyzed using standard graph-theoretic tools to understand the nature of the system's response. In other words, topological measures are applied to the differential network  $D$  itself, rather than the original static networks. This allows for the identification of structures within the change network:

- **Community Detection:** Applying algorithms like Louvain (Section 2.2.5) to the differential network  $D$  identifies modules of nodes (e.g., genes) whose connections were collectively gained or significantly strengthened in response to the perturbation. These "modules of change" often correspond to coordinated biological responses or pathways specifically activated by the stimulus [13].
- **Centrality Analysis:** Calculating centrality measures (Section 2.2) on the differential network  $D$  identifies nodes that are central within the perturbed structure. For example, high-degree nodes in  $D$  might represent key mediators or hubs in the response pathways activated by the perturbation, even if they were not hubs in the original static networks. This analysis is often performed within the detected communities of  $D$  to find local hubs driving modular responses.

By applying network analysis techniques to the differential network, we can move beyond simply listing changed interactions and gain insights into the coordinated, modular nature of the system's adaptation to perturbations.

# Chapter 3. Application of Differential Network Analysis to Salivary and B-Cell Transcriptomic Data

**Note:** Work presented in this chapter has been published as Xue S, Rogers LRK, Zheng M, He J, Piermarocchi C, Mias GI. Applying differential network analysis to longitudinal gene expression in response to perturbations. *Frontiers in Genetics*. 2022;13:1026487.

To exemplify the analytical capabilities of the Differential Network (DN) framework detailed in the preceding chapter, this chapter showcases two case studies using experimental data.

- (i) The first case study aimed to gain insights into the adaptive immune responses to PPSV23 through saliva profiling. Due to its convenience in processing relative to blood samples, saliva draws much interest for diagnostics as well as health monitoring applications. Saliva analysis can produce results in a timely manner, its collection is minimally invasive, and little training is required for saliva sampling, even for non-medically trained professionals.
- (ii) The second dataset (GSE100441) was generated from a time course experiment on primary B cells, where one set was treated with Rituximab and another used as an untreated control. Rituximab is known for its therapeutic use in targeting B cells [75] to treat cancers such as lymphomas and leukemias. This drug has a history of safe and effective usage since 1997 [76], and the World Health Organization (WHO) place Rituximab on their list of essential medicines [77]. Rituximab binds with CD20, expressed on pre-B and mature B cells, but not on stem cells, early pro-B or normal plasma cells [78]. The binding causes

perturbations to intracellular signaling and membrane structure [79], mediating the cell depletion. It is worthwhile mentioning that the B cell pathways of Rituximab activation have been experimentally validated [80, 81, 82, 83], and can be compared to the pathways identified by the DN method. Both the saliva and primary B cell experiments involve drug-treated samples (treatment sets) and untreated samples (control sets) monitored over time.

For both datasets, we started with building gene networks, one for each of the control and the treatment sets. We used gene-gene correlations between time series signals, over 24 hours in saliva and 15 hours in B cells, to evaluate pairwise gene connections. Graphically, the time series correlation networks built from the treatment sets summarized a system-wide pathway activation due to the perturbation, whereas the networks from the controls sets acted as the baseline. Within the DN analysis framework, we subtracted the baseline network from the one obtained using the treatment data, arriving at the final differential network.

### 3.1 Data Acquisition

Data for this investigation were obtained from Gene Expression Omnibus (GEO) for two time series studies using RNA-seq experiments, on Saliva (accession GSE108664) and Rituximab (GSE100441). Both sets of data are further described below. The raw RNA-seq data were mapped using Kallisto [84], with bootstrap sample parameter, `-b`, was set to 100. GENCODE[85] v28 transcripts and genome built GRCh38.p12 were used for annotation. We used Sleuth [86](with DESeq[87] adjustment of Transcripts per Million) to compile results across timepoints.

The saliva dataset was obtained from our previous study of immune responses to the PPSV23 vaccine (GSE108664) [88]. In this study, hourly saliva samples were collected from a

healthy individual over two 24 hour periods and profiled with RNA-seq every hour. The first 24 hour period provides a baseline RNA expression dataset, which we call *untreated* data. In the second 24 hour period, the same individual was monitored after receiving the PPSV23 vaccine. Saliva samples were again collected hourly over 24 hours and profiled by RNA-seq. This second step yielded the RNA expression dataset after the PPSV23 vaccination. We call these data the *treated* dataset. Both treated and untreated datasets have 24 time points of 84,647 possible expression signals using GENCODE annotation [85]. We note that all data obtained were made publicly available by the original authors, Mias et al.[88], as described therein (Michigan State University Institutional Review Board Protocol LEGACY15-071 [15-071]), and no additional institutional review board approvals were required for this investigation.

The perturbation in the primary B cell experiment was Rituximab, a monoclonal antibody drug used in the treatment of different types of lymphomas and leukemias. The experimental study (data from GSE100441) began by culturing in parallel primary B cells with and without Rituximab. During the 15 hours of Rituximab treatment, the treated and untreated primary B cells were both sampled at the same 6 time points simultaneously and profiled by RNA-seq. The untreated group provided a baseline, which we call untreated data, whereas the treated experiment produced the treated dataset. Since this study included a replicated experiment, each of the first and second duplicates was processed to generate a separate network.

### 3.1.1 Data Availability

These data can be found here: Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), accessions: (i) GSE108664 for the saliva mRNA-sequencing, and (ii) GSE100441 for the Rituximab Treatment in Primary B Cells mRNA-sequencing. Mapped

RNA-seq data have been deposited on Zenodo and are publicly available at: <http://dx.doi.org/10.5281/zenodo.7007320>. Results and code files have also been deposited on Github (<https://github.com/gmiaslab/DifferentialNetworks>) and released on Zenodo. These files are referred to as Online Data Files (ODFs) in the manuscript. DOI: <http://dx.doi.org/10.5281/zenodo.7007320>.

## 3.2 Application Workflow

In the present study, we applied a DN approach, as introduced in chapter 2, to RNA-sequencing (RNA-seq) time series datasets retrieved from two longitudinal RNA-seq experiments: (i) The first dataset (GSE108664) was generated from saliva samples from a healthy individual before and after the administration of the Pneumococcal Polysaccharide Vaccine (PPSV23) [88].

We exploited this property of the differential network to observe fine details of gene groups affected by the perturbation. That is, we employed one of the most established module detection algorithms, the Louvain method [70, 89, 90], to identify communities in our final differential network. We further explored communities by clustering heatmap and pathway enrichment analysis. The clustering heatmaps enabled us to characterize communities by their unique temporal patterns. Additionally, we identified early (and late) responding communities to perturbations and arrived at a sequential activation order for specifically the saliva DN communities. Lastly, we performed Reactome [91] pathway enrichment analysis on individual communities and annotated the results with hub genes (based on DN centrality).

Our investigation extends applications of DN to gene expression time series that include perturbative activation. The two DN applications, and particularly the community-wide

investigations provide further biological insight in gene expression changes in both Rituximab treatment, as well as pneumococcal vaccination. Specifically, each of the three investigations on DN communities provided unique perspectives on the biological response to perturbations: (i) Using our heatmap analyses, we found that each DN community can have its own temporal pattern and be used as a categorization of time-resolved gene activation. (ii) Using community enrichment, we determined the associations between the activated biological pathways and their gene clusters (communities). Combined with the community temporal patterns, our results provide a chronological order of pathway activations, and show how these may be obtained through a DN application. (iii) Lastly, our community hub analysis gave further insights into the biological functionality of individual genes in a community. These include, for example, the presence of the hub gene IL4R in one of the saliva DN communities, which suggests that the respective cluster of genes collectively activated T cells in response to the PPSV23 vaccine, and may explain a fever event in the experimental subject. Likewise, the presence of the hub gene PELI, known to be an oncogene in lymphomagenesis, in one of the B cell DN communities suggests that the entire community participates in the B cell response to Rituximab. Additional findings are summarized in the results below, and illustrate the utility of extending DN analyses to investigate time-resolved gene expression changes induced by drug and vaccine treatments.

The workflow is summarized in Figure 3.1.

### 3.2.1 Data Preprocessing

For quality control, we pre-processed the experimental data and filtered signals with multiple missing points right after importing the published data files. We coded all the data analysis in Python in this study. Using Python's pandas package [92, 93], we checked for missing

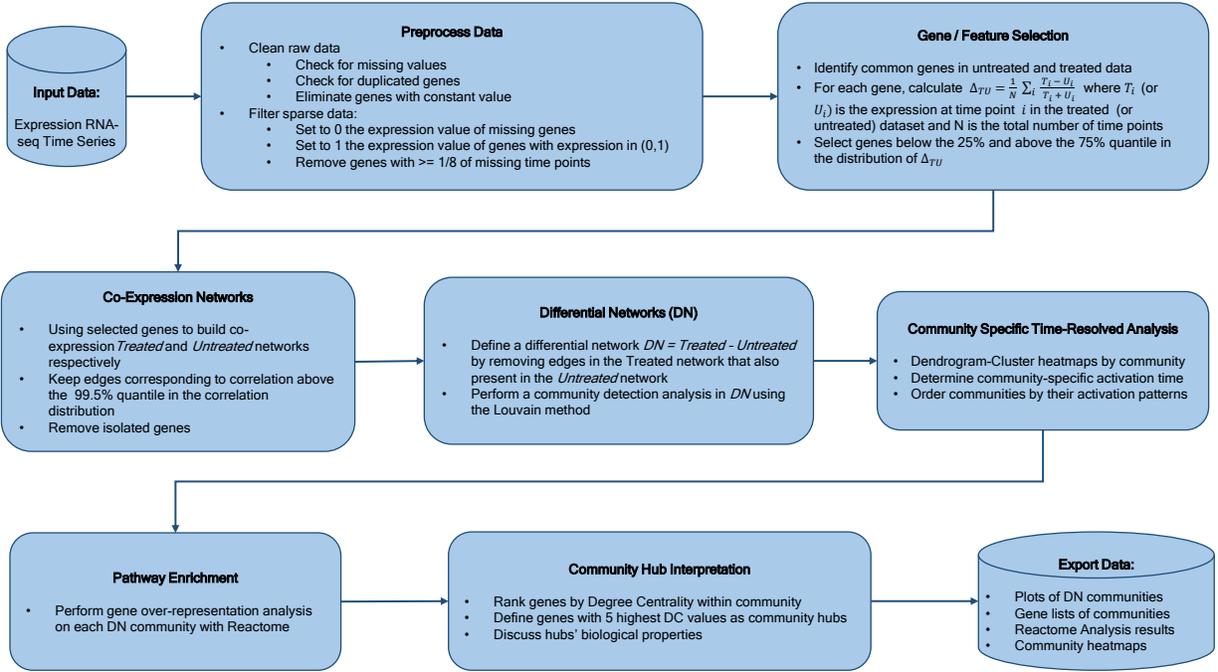


Figure 3.1: **Workflow Overview.** Our methodology starts with time course experimental data, followed by network construction, differential network determination, community detection, sequential ordering by activation pattern, pathway analyses of individual communities, community hub gene interpretation, and final results including analyses and temporal trend visualizations.

values for each gene’s expression, removing duplicate records and eliminating genes with constant values across all the 24 time points for the saliva dataset (6 time points for the B cell datasets).

We replaced missing signals with zero and also set values less than 1 to 1. Genes with zero variance in their time series were excluded from our analysis. Moreover, we considered a gene signal as sparse and removed it if its time series had missing values for more than 1/8 of the time points. The same quality control procedure was used for both the saliva and primary B cell datasets.

### 3.2.2 Gene Selection

After quality control, we further processed the data to pre-screen and identify a pool of candidate genes that showed response to the perturbations (vaccine in the saliva and Rituximab in the B cell). We selected genes that are highly expressed in both untreated and treated cases. Our goal with the differential networks was to identify genes that displayed notable changes. Hence the cutoffs were selected to exclude constant signals, and signals with moderate changes when comparing corresponding paired timepoints. For each of these genes, we calculated the time-averaged relative difference between treated and untreated normalized intensities,  $\Delta_{TU}$ :

$$\Delta_{TU} = \frac{1}{N} \sum_i \frac{T_i - U_i}{T_i + U_i} \quad (3.1)$$

where  $T_i$  is the expression value at time  $i$  in the treated dataset,  $U_i$  is the expression value at time  $i$  in the untreated dataset, and  $N$  is the total number of time points. This calculation yielded a  $\Delta_{TU}$  distribution curve, from which we computed the lower and upper quartiles. As our goal was to identify time-resolved changes, genes were selected if their  $\Delta_{TUs}$  were within the bottom 25% or top 25% of the  $\Delta_{TU}$  distribution respectively. The Python Pandas package was used for all the above computations [92, 93].

### 3.2.3 Co-Expression Networks Construction

After gene selection, we calculated their pairwise Pearson correlation coefficients and built the co-expression networks. Genes were represented as nodes and were joined by edges if there was a non-zero correlation between them. We used the co-expression coefficient as a weight for

each edge. In the layout representation of the networks, the node-node distance reflects their correlation coefficients. Two genes are nearby if they have a high positive correlation. They are far apart when they have a low positive correlation or remote if negatively correlated. We used Python’s open source Networkx package [94] for network visualization and calculation of the network metrics.

We constructed the network with edges in the 99.5% quantile of the correlation distribution, and excluding singletons. The one-sided quantile cutoff essentially selects for positive correlations and is consistent with our modularity-based community analysis discussed further below. For the saliva data, we built one treated and one untreated network. Since we have data from two repeated experiments for B-cells, we built two networks for the Rituximab treatment and two networks for the untreated control. Then, we took the intersections between the two networks corresponding to the repeats to obtain a single Rituximab-treated network and one single control network.

### 3.2.4 Differential Networks Construction

We defined the DN as the control network subtracted from the treated networks both for the saliva and B cell cases. In the subtraction, we remove an edge if that edge appears both in the treated and untreated networks. Edges appearing only in the treated network and absent in the untreated are kept in the differential network. Edges appearing only in the untreated network are not included. Isolated nodes left after this procedure are discarded. We analyzed the DN’s structure using modularity [58, 59], as complex biological networks usually display a high degree of modularity [46]. Modularity is a measure to quantify relative edge densities from within the communities in comparison with those outsiders. We utilized the Louvain community detection method [70], as implemented in a published Python package [95], a

greedy algorithm for modularity maximization, to partition the entire DN into smaller clusters, also known as communities. The algorithm consists of two stages: first, individual nodes are joined into communities to achieve local maximum modularity; second, nodes within the same community are aggregated to form a new network where the node-assignment procedure is repeated until the modularity no longer increases. This graph clustering algorithm is not deterministic and can therefore result in slightly different partitions for the same graph. The partition yielded a few major components and many tiny communities of fewer than 5 genes, disconnected from the central islands. We pruned out these small communities from the DN. We found that the majority of our communities with a low number of genes yield no significant enrichment [1, 96].

### 3.2.5 Community Specific Time-Resolved Analysis

In order to investigate the time-resolved response present within the communities, we applied clustering heatmaps to each of the DN communities. For genes in the same community, we first retrieved their treated and untreated expressions, then normalized each time series by subtracting individual time points from the time 0, followed by normalization with the Euclidean norm, for both expressions. We then took the difference between the normalized treated and untreated time series. Finally, we dendrogram-clustered these series (rows) with the complete-linkage method (Farthest Point Algorithm) [2, 3]. The same procedure was repeated in all communities, and each delivered a clustering heatmap.

As the heatmaps rendered distinctive time-resolved responses in each community, we identified communities that responded quickly to perturbations and those that responded slowly. In particular, we characterized saliva communities by their peak times and arranged them in temporal order. We did not obtain an order for the B cell communities, as the B cell

heatmaps did not show dominant peak times. However, we were still able to characterize B cell communities based on 3 distinguishable temporal pattern categories.

### **3.2.6 Pathway Enrichment**

We conducted Reactome Enrichment Analysis [91] on each community to identify over-represented biological pathways within each community, using the Python package PyIOmica [97]. As the majority of our communities with a low number of genes yield no significant enrichment, we focused on results for communities with 8 or more genes [1, 96].

### **3.2.7 Community Hub Identification**

We isolated each community and identified localized hubs, only considering the communities rather than the global DN in our calculations. We adopted the standard Degree Centrality (DC) algorithm (which has been integrated with the python networkx package) and identified the genes with the top 5 DC values as the community hub genes. We examined these hubs using functional annotations (such as pathway and memberships, including from GeneCards Suite [98]) to evaluate if their biological properties could potentially elucidate the more general functionality of the module of their membership.

### **3.2.8 Results formatting and visualization**

We stored the DN nodes and edges, communities, and pathway enrichment analyses into spreadsheets that are provided in the Online Data Files (ODFs) both for the saliva and B cell data. Using Mathematica [99], we visualized the saliva and B cell DNs with their major connected components and communities.

### 3.3 Results

Our RNA-seq time series raw data were retrieved from the Gene Expression Omnibus (GEO) database under accessions GSE108664 and GSE100441 for the saliva and B cell experiments, respectively. The study of the immune response to the PPSV23 vaccine in saliva probed the expression of a potential 84647 gene identifiers (GENCODE annotation[85]) at 24 time points [88]. The other study of drug activation by Rituximab in B cells provided a dataset for 6 time points. Since gene co-expression networks rely on correlations, our network analysis could be prone to spurious correlations, which we removed as described in the Methods section.

We constructed our saliva DN by subtracting the saliva network without vaccine from the network obtained using post-vaccine data. The B cell DN in response to Rituximab was generated in a similar manner. Next, we clustered the DNs into communities using the Louvain community detection method [70]. We then conducted a Reactome Enrichment Analysis [91] using PyIOmica [97], on each community to identify significant pathways and associated genes. We also visualized the heatmaps of relative gene expression as a function of time for each community. Finally, we plotted the DNs and their major individual communities. See the Application Workflow section for additional details.

#### 3.3.1 Saliva DN

Our saliva DN contains 1144 nodes (i.e., genes) and 13,775 edges. The Louvain algorithm identified 48 communities (modules) in total. 15 of the communities have a size of at least 4 nodes, while the remaining 33 are pairs or triplets. In the global saliva DN visualization, we excluded the communities with pairs or triplets, as none of them belonged to the three major connected components of the DN network. We also filtered the network to remove

connected components with less than 4 genes. The global saliva DN is presented in Fig. 3.2a, where communities are visualized using different colors and encircled in loops. Furthermore, community labels are based on their size (largest to smallest, with C0 being the largest community, and C14 the smallest).

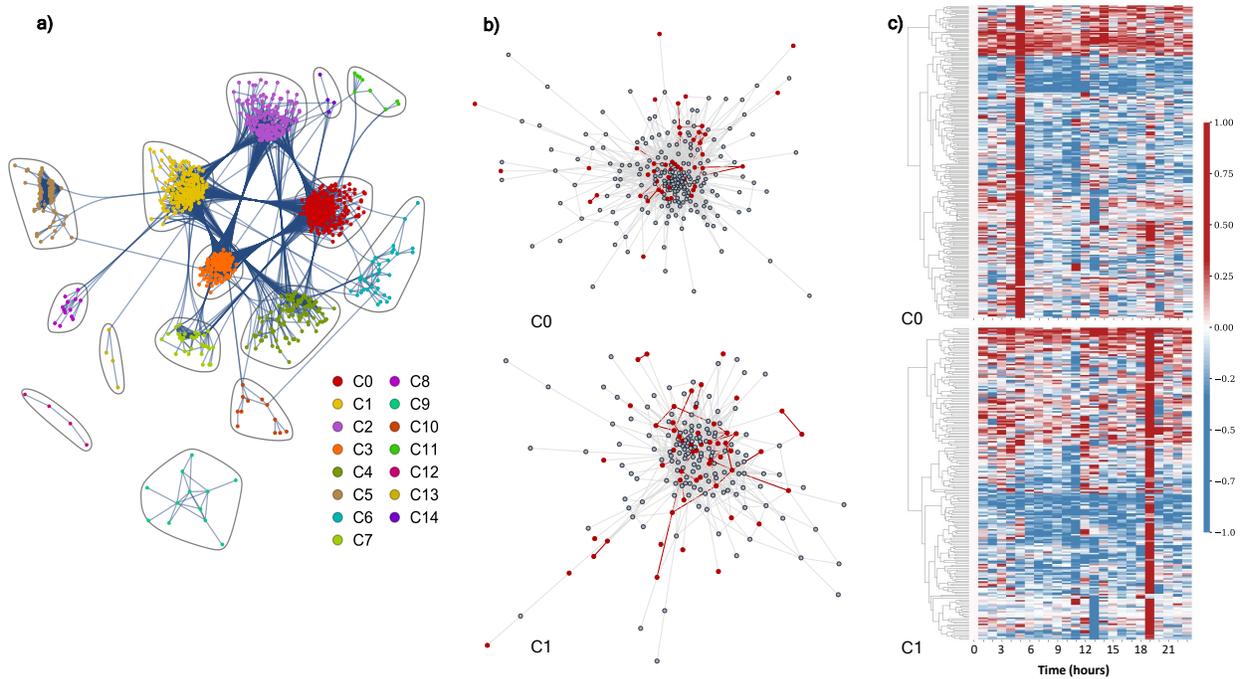


Figure 3.2: **Differential network analysis for the saliva experiment.** a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C0 (top) and C1 (bottom) communities with red highlights indicating genes found in statistically significant Reactome pathways ( $FDR \leq 0.05$ ), and their corresponding edges in the network. c) Heatmaps of C0 (top) and C1 (bottom) over 24 hours. Columns represent time points while rows denote gene identifiers. The row data show the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalizing using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (also known as the Farthest Point Algorithm) [2, 3].

### 3.3.2 Saliva Communities Temporal Visualization

We further visualized each community’s change over time with heatmaps within the DN network. This is shown for C0 and C1 in Fig.3.2c. Here, each row denotes a gene, while each column corresponds to a time point (pre and post vaccination). The values plotted in the heatmaps are rescaled gene expression differences between the treated data and the control, and indicate the expression at the particular time point relative to the first time point of the experiment, with rows normalized using Euclidean norm. Red indicates up-regulated genes, blue down-regulated genes, and white indicates unchanged expression. The hierarchical clustering dendrograms revealed relationships among genes at each time point based on the similarity of the gene expressions. The prominent red columns show that genes are upregulated together at these time points. Note that the C0 has a pronounced peak at time point 6, making it an early responding module, while C1 is a late responding cluster, with a pronounced peak at time point 19, as illustrated in Fig. 3.2c.

Here we only show heatmaps for C0 and C1 as representative communities. However, we provide the other communities’ heatmaps and with their corresponding Reactome pathway analysis in the ODFs (folder “Results/SLV\_results/network\_plots”). Our saliva DN has a clear pattern of mostly discrete punctuated gene expression response times for each community. As these punctuated response times, save for one exception (both C0 and C11 show maximized response at t5), are specific to each community, they reflect the biological signatures for individual groups. Most of our saliva DN communities have only one punctuated activation time, although C5 in the saliva DN has 3 up-regulation events at time points 15, 20, and 22 that do not overlap with those of other communities. Between the communities, we observed strong temporally-specific relationships. Our heatmaps are suggestive of the presence of

directional signaling between early-activation communities and subsequent groups, with a potential sequential activation pattern as follows: C6, C9, C8, C2, C0 and C1, C3, C4, C10, C5, C1, and finally C5. At time points from t6 to t10, t14, and from t16 to t18, no communities activated.

### 3.3.3 Pathway Enrichment of Saliva Communities

In our pathway analysis, we queried individual communities to investigate how their highly co-expressed genes are functionally related. Our analysis is based on the Reactome pathway database [91, 100, 101].

Statistically significant enrichment of pathways (with False Discovery Rate (FDR)  $< 0.05$ ) was identified in 6 communities, C0, C1, C2, C4, C8 and C9. The majority of statistically significant Reactome pathways were related to response to stimulus, immune response, and inflammatory response. Among the six communities, C0 and C1 are the two largest communities. C0 comprises of 248 genes, colored in red in the global DN shown in Fig. 3.2a, whereas C1 contains 198 genes, colored in yellow in the same panel. We display C0 and C1 in Fig. 3.2b as representative communities. Genes that belong to the statistically significant biological pathways are highlighted in red in Fig. 3.2b.

In the C0 community, the Reactome enrichment analysis identified 15 statistically significant pathways (FDR  $< 0.05$ ): (i) three pathways for interferon signaling, (ii) three related to the immune system, (iii) four related to antigen presentation, (iv) one associated with ER-Phagosomes, (v) one lymphoid-related, and (vi) three pertaining to interleukin-12 signaling. In particular, the alpha, beta, and gamma signaling pathways all appear in the interferon signaling pathways. The immune system pathways include one cytokine signaling and one related to the adaptive immune system. Among the four antigen-related pathways, two are

explicitly associated to the dependence of Class I MHC. The Endosomal / Vacuolar pathway implies the involvement of the Class I MHC and of the Antigen processing-Cross presentation. Lastly, interleukin-12 plays a crucial role in the coordination of innate and adaptive immunity [102].

In the C1 community, the Reactome analysis identified 9 statistically significant pathways (FDR < 0.05). Two of these pathways are broadly related to the immune system and cytokine signaling. Another two pathways, the NGF-stimulated transcription and the FOXO-mediated transcription pathways, modulate cell survival, growth, and differentiation. In Table 3.1 we have listed all the results of the Reactome pathway enrichment analysis for C0 and C1 with FDR < 0.05.

Table 3.1: Reactome pathway enrichment analysis. Statistically significant pathways (FDR < 0.05) are summarized for saliva DN communities C0 and C1. In the full analysis, we omitted small communities with fewer than 8 genes [1], and 12 communities (C0 to C11) qualified for the pathway analysis.

Pathway Name	Entities FDR	Submitted entities found
<i>Saliva DN: C0</i>		
Antigen Presentation: Folding, assembly and peptide loading of class I MHC	1.2E-14	HLA-B, NAA15
Endosomal/Vacuolar pathway	1.2E-14	HLA-B
Interferon gamma signaling	1.2E-14	STAT1, IRF1, HLA-B, PTPN6
Class I MHC mediated antigen processing & presentation	1.2E-14	PSMD8, TLR1, CDH1, RPN1, GBF1, HLA-B, UBR4, CYBA, NAA15, ELOC, FBXO32, FBXO11
ER-Phagosome pathway	1.2E-14	PSMD8, TLR1, RPN1, HLA-B
Interferon alpha/beta signaling	1.2E-14	STAT1, IRF1, HLA-B, PTPN6
Interferon Signaling	1.2E-14	EIF4A1, STAT1, IRF1, HLA-B, PTPN6, ARIH1
Antigen processing-Cross presentation	1.2E-14	PSMD8, TLR1, RPN1, HLA-B, CYBA
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	1.2E-14	CDH1, CD81, HLA-B, FCGR2B
Continued on next page		

Table 3.1 – continued from previous page

Pathway Name	Entities FDR	Submitted entities found
Cytokine Signaling in Immune system	1.1E-11	EIF4A1, STAT1, IRF1, HLA-B, PTPN6, ARIH1
Adaptive Immune System	4.1E-09	CD81, TCF25, RPN1, GBF1, HLA-B, UBR4, CYBA, PPP2R5D, FBXO32, FBXO11, ANKRD9, TLR1, PSMD8, CDH1, AKT2, PTPN6, ELOC, NAA15, FCGR2B, SIPA1, ARF5
Immune System	1.3E-05	CCDC71L, DDX3Y, EIF4A1, ASAH1, IL1RN, SERPINA1, TCF25, CD81, RPN1, RPLP0, UBR4, TNFAIP3, CSF2RA, PLD2, PSMD8, ANKRD9, CDH1, AKT2, OLR1, ELOC, ARIH1, SERPINB2, TNFSF14, GSTO1, STAT1, GBF1, HLA-B, CYBA, PPP2R5D, FBXO32, FBXO11, FGR, CEACAM3, CLEC4A, TLR1, IRF1, TCP1, TXNIP, PTPN6, CYSTM1, NAA15, FCGR2B, SIPA1, BIRC2, ARF5, TRIM56
Gene and protein expression by JAK-STAT signaling after Interleukin-12 stimulation	3.2E-03	SERPINB2, GSTO1, TCP1, RPLP0, ARF5
Interleukin-12 family signaling	4.6E-03	SERPINB2, GSTO1, STAT1, TCP1, RPLP0, ARF5
Interleukin-12 signaling	8.0E-03	SERPINB2, GSTO1, TCP1, RPLP0, ARF5
<i>Saliva DN: C1</i>		
Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs) bind RNA	2.7E-03	CD44
Nuclear Events (kinase and transcription factor activation)	1.8E-02	PPP2CB, TF, ID2, CHD4, FOS, DUSP6, DNM2
FOXO-mediated transcription of cell death genes	1.8E-02	BCL2L11, BCL6, NFYC
Signaling by NTRKs	2.4E-02	PPP2CB, RALA, TF, ID2, CLTA, FURIN, CHD4, FOS, DUSP6, DNM2
Signaling by NTRK1 (TRKA)	2.6E-02	PPP2CB, RALA, TF, ID2, CLTA, CHD4, FOS, DUSP6, DNM2

Continued on next page

Table 3.1 – continued from previous page

Pathway Name	Entities FDR	Submitted entities found
Immune System	2.9E-02	NAPA, RALA, CIITA, AHCYL1, RPN2, UNC93B1, JADE1, CLTA, BCL10, CFP, TANK, GNS, FCAR, STK10, PPP2CB, BCL2L11, TRIM29, ALOX5, NLRP3, FLNA, SIRPA, SLC12A6, IL6R, GBP4, RAP1GAP2, DDX17, CR1, WSB1, CISH, SH2D3C, KLHL21, FNDC3A, FOS, LILRB3, MTOR, DUSP6, VEGFA, DNM2, TF, ZNFX1, NASP, BCL6, MAN2B1, TACC2, CD300C, CALM1, CD44, LGMN
Cytokine Signaling in Immune system	3.6E-02	RALA, CIITA, CISH, RPN2, SH2D3C, FNDC3A, FOS, MTOR, DUSP6, VEGFA, PPP2CB, ZNFX1, BCL2L11, NASP, BCL6, TRIM29, ALOX5, FLNA, IL6R, GBP4, CD44
trans-Golgi Network Vesicle Budding	3.9E-02	NAPA, CPD, CLTA, GNS, CLINT1, DNM2
NGF-stimulated transcription	3.9E-02	TF, ID2, CHD4, FOS, DNM2

Of the communities we observed, the C0 community exhibits the strongest response to the stimulus and immune system, based on  $FDR \sim \mathcal{O}(10^{-14})$ . The complete pathway enrichment analysis for all communities in saliva is provided in the Online Data Files (ODFs), available on Zenodo, in the “Results/SLV\_results/reactome\_analysis” folder.

### 3.3.4 B Cell DN

Our B cell DN consists of 1,759 nodes (genes) and 10,421 edges that we classified into 145 communities using the Louvain algorithm. Similar to the saliva DN, most of these communities are small clusters on small components. Due to its larger size relative to the saliva DN and larger number of communities, our cutoff for plotting was increased to 8 nodes

both for community and component size. The global B cell DN is presented in Fig. 3.3a, with 5 components and 14 communities. Here, we omitted the remaining 130 communities since they neither belong to any of the 5 main components, nor are they large enough for Reactome enrichment analysis. Like in the saliva DN, communities were ordered in descending size (largest to smallest, from C0 to C13 respectively), designated with different colors, and encircled by loops. Fig. 3.3 has the same format of Fig. 3.2. In this case, C2 and C4 are displayed in panel b, as magnified representations of the purple cluster and the green cluster, respectively, in panel a. Panel b's magnified perspective provides details about the communities' internal structures. In Fig 3.3b, for example, we observe that some of the genes highlighted in red form a clique.

### 3.3.5 Pathway Enrichment of B Cell Communities

As for the saliva DN, we conducted a community-wise Reactome enrichment analysis for communities with at least 8 genes. 14 communities in the B cell DN were analyzed. This analysis found 9 communities with statistically significant pathway enrichment (FDR < 0.05): C2, C4, C5, C6, C7, C9, C10, C12, and C13. Most of the pathways associated with genes in these communities centered around transcriptional regulation, protein metabolism, DNA binding ability, and signaling. Among its 111 statistically significant pathways, C4 was strongly enriched with genes in the FCERI-mediated NF- $\kappa$ B activation pathway, the B cell receptor (BCR) signaling pathway, and the Fc epsilon receptor (FCERI) signaling pathway. These pathways and others relevant to Rituximab mechanism of action are listed in Table 3.2. The NF- $\kappa$ B pathway activation by FCERI leads to the production of cytokines during mast cell activation, making it important in allergic inflammatory diseases [103]. C4 also contained genes in the B cell receptor pathway, an important pathway related to B cells.

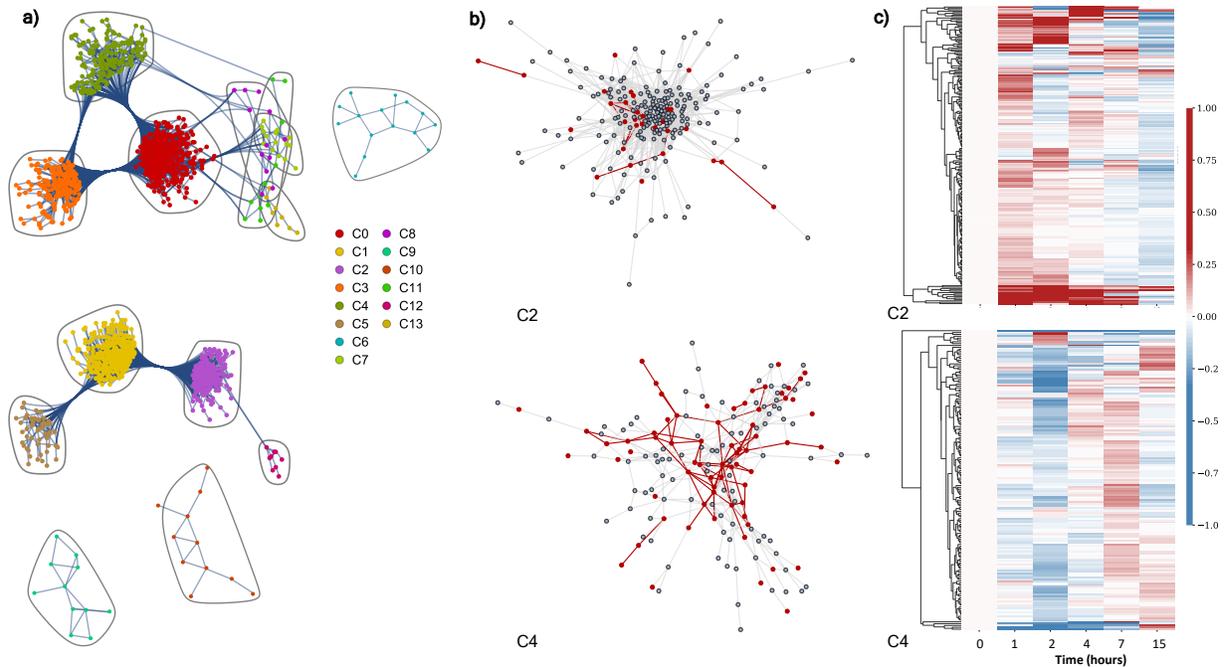


Figure 3.3: **Differential network analysis for the B cell experiment.** a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C2 (top) and C4 (bottom) communities with red highlights indicating genes found in statistically significant Reactome pathways (FDR $\leq$ 0.05), and their corresponding edges in the network. c) Heatmaps of C2 (top) and C4 (bottom) over 15 hours (6 time points). Columns represent time points while rows denote genes. These row data demonstrate the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalizing using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (Farthest Point Algorithm) [2, 3].

The Fc epsilon gene is expressed on antigen-presenting cells, and its signaling occurs on the plasma membrane. A comprehensive list of statistically significant pathways can be found in the Online Data Files (ODFs) in the “Results/Bcell\_results/reactome\_analysis” folder.

Table 3.2: Reactome pathway enrichment analysis. Statistically significant pathways are summarized for primary B cell DN community C2 and C4. In the full analysis, we omitted small communities with fewer than 8 genes [1], and 14 communities (C0 to C13) qualified for the pathway analysis.

Pathway Name	Entities FDR	Submitted entities found
<i>B cell DN: C2</i>		
Peptide chain elongation	1.3E-06	EEF1A1, RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Response of EIF2AK4 (GCN2) to amino acid deficiency	1.3E-06	RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, FAU, RPL15, UBA52, RPL28, ATF3
Eukaryotic Translation Elongation	1.6E-06	EEF1A1, RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
GTP hydrolysis and joining of the 60S ribosomal subunit	5.9E-06	RPL4, EIF4A1, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
L13a-mediated translational silencing of Ceruloplasmin expression	5.9E-06	RPL4, EIF4A1, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	5.9E-06	RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Formation of a pool of free 40S subunits	7.7E-06	RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Eukaryotic Translation Termination	7.7E-06	RPL4, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Cap-dependent Translation Initiation	9.2E-06	RPL4, EIF4A1, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Eukaryotic Translation Initiation	9.2E-06	RPL4, EIF4A1, RPL7A, RPL27A, RPS6, RPL36, RPL14, RPS20, RPL15, FAU, UBA52, RPL28
Continued on next page		

Table 3.2 – continued from previous page

Pathway Name	Entities FDR	Submitted entities found
<i>B cell DN: C4</i>		
Metabolism of RNA	1.4E-02	SF3B4, MT-ND6, NUP205, UTP3, POP1, DDX23, CSTF2, PHAX, PLRG1, DIEXF, ZFP36L1, FTSJ3, CHERP, PSMD8, EFTUD2, PSMD9, PSMC4, PSME3, NUP35, SKIV2L2
Mitotic Anaphase	1.4E-02	PSMD8, PSMD9, NUP205, CCNB1, SPAST, PSMC4, PSME3, NUP35, SMC1A, EMD, KPNB1
Mitotic Metaphase and Anaphase	1.4E-02	PSMD8, PSMD9, NUP205, CCNB1, SPAST, PSMC4, PSME3, NUP35, SMC1A, EMD, KPNB1
FCERI mediated NF-kB activation	1.4E-02	IGLV2-11, PSMD8, PSMD9, IGKV2-29, IGKV1-16, PSMC4, PSME3, IGKV4-1
Signaling by the B Cell Receptor (BCR)	1.4E -02	IGLV2-11, PSMD8, PSMD9, IGKV2-29, IGKV1-16, PSMC4, PSME3, IGKV4-1, PIK3AP1
Fc epsilon receptor (FCERI) signaling	1.4E-02	IGLV2-11, PSMD8, PSMD9, IGKV2-29, IGKV1-16, PSMC4, PSME3, IGKV4-1
Host Interactions of HIV factors	1.4E-02	PSMD8, PSMD9, NUP205, PSMC4, PSME3, NUP35, KPNB1
G1/S Transition	1.4E-02	PSMD8, PSMD9, CCNB1, MCM7, PSMC4, PSME3, KPNB1
ABC-family proteins mediated transport	1.4E-02	PSMD8, PSMD9, PSMC4, PSME3, CSTF2, EIF2S1
Assembly of the pre-replicative complex	1.4E-02	PSMD8, PSMD9, MCM7, PSMC4, PSME3

In summary, C4 contains the highest number of responsive pathways relevant to the B cell response to Rituximab. As our representative communities, we display the C2 and C4 in Fig. 3.3 b, our two largest among the 9 communities with statistically significant pathways. Our top 10 pathways based on FDR from the Reactome enrichment analysis for C2 and C4 are listed in Table 2 3.2.

### 3.3.6 B Cell Communities Temporal Visualization

The heatmaps for the temporal behavior for the C2 and C4 communities of the B cell data are shown in Fig. 3.3c. The formatting of the heatmaps is the same as that of the saliva heatmaps; all values in the heatmaps refer to gene expression relative to time 0 in the treated dataset. The C4's blue column at time point 2 and the less prominent blue column for C2 at time point 15 identify patterns of down-regulation in the two communities. While C2 shows a trend of initial up-regulation followed by a gradual diffusion, C4 exhibits an initial down-regulation, followed by later up-regulation.

Though C2 and C4 are our representative communities, we carried out heatmap visualization for all 9 communities that demonstrated statistically significant pathway enrichment. These heatmaps are available to view in the ODFs in the “Results/Bcell\_results/network\_plots/heatmaps” folder. Overall, in the B cell community heatmaps, we recognized three types of time patterns in terms of collective behavior within an individual community. In the first pattern group, the majority of genes started with a moderate degree of down-regulation. By 7 hours, most instead displayed slight or moderate up-regulation. However, each of these timepoints contained a minority of genes with a small level of fluctuation, with the size of the deviating group differing in each heatmap. The second observed time pattern operated in reverse, with most genes beginning upregulated and shifting towards downregulation by the 15-hour mark. Finally, a third group remained consistent in its behavior, with genes trending one way or remaining unchanged across the entire time period.

### 3.3.7 Community Hubs

We examined the community hub genes for both saliva and B cell DNs, and reported the degree centrality in their respective Results tables in the ODFs with sheet name “Degree Centrality”. When two DC values are the same, the genes are tied in rankings in our consideration as community hubs.

### 3.3.8 Biological Considerations

In our results, a number of expected pathways emerged. These included pathways associated with antigen presentation and processing, Class I MHC mediated antigen processing and presentation, and ER-phagocytosis, and pathways governing the immunoregulation of interactions between Lymphoid and non-Lymphoid cells [100]. Further results indicative of the participation of immune cells, included the CLEC inflammasome pathway in C4. This pathway is associated with enabling host immune system to mount a fungal/bacterial defense using T-Helper 17 cells (TH17) [104, 105]. Interferon signaling, cytokine signaling, immune/adaptive immune, and interleukin stimulation and signaling are all part of a generalized immune response [106]. We found these more general pathways in the pathway enrichment analysis of C0, C1, C2, C9, and C10. Interferon signaling is crucial in antiviral defense, cell regulation and growth, and immune response modulation [107]. Our Reactome pathway analysis results are consistent with the results of our saliva multi-omics study [88], which observed that vaccination activates various immune response and regulation pathways, which are also identified in our present results, including ER-Phagosome pathway, Interferon alpha/beta and gamma signaling, cytokine signaling, and MHC antigen presentation.

From our community hub gene analysis, a few hub genes are suggestive of community

functionality. Notably, the community C2 hub gene, BIRC2, regulates NF- $\kappa$ B signaling as well as inflammatory signaling and immunity [108, 109]. For one of the C2 hubs, URGCP, previous findings indicate that its upregulation and downregulation are significantly involved in the molecular mechanisms of non-small cell lung cancer [110, 109]. Accordingly, the presence of the URGCP is consistent with our vaccine targeting the respiratory system. As for community C3, the IL4R gene encodes interleukin 4 and interleukin 13 to regulate IgE production, which further activates the JAK/STAT pathway [111]. This pathway orchestrates cytokine receptors, modulates T helper cell polarization, and also mediates human monocytes/macrophages [111, 112]. Lastly, in community C5, the EBF1 is known as a leading transcription factor of B-cell specification [113]. In summary, IL4R and EBF1 became the most connected gene in C3 and C5, respectively, which implies that these 2 communities are each centered around T cells (C3) and B cells (C5). Moreover, given that the experimental subject had a post-vaccination fever at hour 11 [88], at which our C3 heatmap coincidentally peaks (see figure in ODFs), the hub gene IL4R in C3 appears to relate the fever event with T cell responses.

Regarding our primary B cell results, previous work [114] has established both the biological pathways and the mechanisms of action associated with Rituximab. These previous studies have demonstrated Rituximab's ability to cause antibody-dependent and complement-dependent cellular cytotoxicity, growth inhibition and apoptosis, and regulation of the cell cycle. We also expected to observe Rituximab regulations of the B cell receptor (BCR) based on prior research. Notably, our findings included the enrichment of the nuclear factor  $\kappa$ B (NF- $\kappa$ B) pathways. According to Jazirehi et al. (2005) [82] and Bonavida (2005, 2007) [115, 116], treating NHL B cell lines with Rituximab inhibits NF- $\kappa$ B's signaling pathways by up-regulating RKIP and Raf-1 kinase inhibitors. RKIP has been found to antagonize signal transduction pathways that mediate the NF- $\kappa$ B activation [117].

Following NF- $\kappa$ B's down-regulation due to RKIP's up-regulation, the Bcl-xl expression is also down-regulated. As a result, tumor cells become more chemosensitive. Rituximab also decreased the activity of NF- $\kappa$ B-inducing kinase, I $\kappa$ B kinase, and I $\kappa$ B-a phosphorylation. Finally, the introduction of Rituximab also decreased the activity of the IKK kinase and NF- $\kappa$ B binding to DNA from 3 to 6 hours after treatment [82].

Among the more general enriched pathways observed are signaling pathways that play a role in the molecular mechanisms of chemosensitization, which are also impacted by Rituximab. In line with those effects, we anticipate impacts in the MAPK signaling pathway, the interleukin cytokine regulatory loop, and the Bcl-2 expression. Concerning the expression of genes involved in the healing process, research has uncovered Rituximab's role in affecting pathways associated with immunoglobulin production, chemotaxis, immune response, cell development, and wound healing. Rituximab can also increase existing drug-induced apoptosis [114].

In our community of C4, for example, our Reactome analysis found 5 NF- $\kappa$ B related pathways with FDR < 0.05. Of these 5 pathways, one is shown in Table 1; the remaining are displayed in the comprehensive table in the ODFs ("Results/Bcell\_results/reactome\_analysis" folder). Alongside these NF- $\kappa$ B pathways in C4 is the BCR pathway. Our results suggest that the C4 community response is highly relevant because of the activation of both NF- $\kappa$ B and BCR pathways.

Our C2 community appears to be involved with the metabolism of proteins and cellular responses to external stimuli. Rituximab targets the CD20 B cell transmembrane protein that is involved in B-cell development, activation and proliferation [114]. The C2 community captures cell development pathways included in our expectations of more generalized responses.

We also observed relevant responses in other communities. For example, the C8 community

showed activity in the RAF/MAP kinase cascade pathway. In a similar fashion, C10 demonstrated CD22 mediated BCR regulation, classical antibody-mediated complement activation, FCGR activation, antigen activation of the BCR, and initial complement triggering, etc. The pathways that emerged in our results are thus consistent and highly overlap with established pathways from previous studies.

Hub genes most pertinent to B cell/lymphocytes included PELI1 in community C5, PRDM2, MALAT1, and SND1 in C2. Other high centrality genes with similar relevance included MAPK8 in C6 and AFF3 in C1. Among these, PELI1 turned out to be closely associated with antitumor immunity in B cells, which is the therapeutic goal of the Rituxmab treatment. A previous study [118] showed that prolonged expression of PELI1 causes B cell hyperactivation, which, in turn, promotes various lymphoid malignancies. To a lesser extent, an increased expression of PELI1 can induce BCL6, an oncoprotein known for advancing lymphomagenesis, for example, B-acute lymphoblastic leukemia and chronic myeloid leukemia. This pathway has been recognized as a potential therapeutic target for treating B cell lymphoma. Out of the C2 hub genes, PRDM2 is a tumor suppressor [119], whereas the upregulation of MALAT1 is linked to tumor cell proliferation and metastasis, such as leukemia [120]. The protein encoded by gene SND1 is known to interact with Epstein-Barr virus nuclear antigen 2 (EBNA 2), which is essential for B-lymphocyte transformation [121]. As an oncogene [122], SND1 has attracted clinical investigation as a cancer treatment candidate due to its association with cell proliferation, and malignant transformation [123]. The literature also has reported that MAPK8, which we found in community C6, mediates starvation-induced BCL2 phosphorylation [124], a sign of cell apoptosis, and AFF3, found in community C1, serves a role in lymphoid maturation and oncogenesis [125]. The fact that these genes appear as hubs in the Rituximab's DN is consistent with their known important

roles in B-cell malignancies and merits further investigation.

### 3.4 Discussion

Our goal was to use a DN approach to identify the activation of biological processes caused by a perturbation in saliva and primary B cells. This study applied DN analysis, community identification and Reactome pathway analysis of the DN communities, and identified communities with highly statistically significant enrichment. In this study we implemented a modularity-based community detection, that works with positive correlations. This is a limitation of the modularity approach to the DN that may be addressed using different community detection algorithms and merits followup investigations. We analyzed the DNs of two gene expression datasets where a perturbation was applied: (i) Saliva dataset (PPSV23 vaccination as perturbation; 24 time points), (ii) Primary B-Cells dataset (*ex-vivo* Rituximab drug treatment as perturbation; 6 time points). In summary, our results from the saliva DN revealed pathway activation in immunological and inflammatory responses. In the B cell DN, statistically significant pathways were activated in the regulation of transcription, immune cell survival, activation and differentiation, and inflammatory response.

*Streptococcus Pneumoniae*'s virulence and associated host immunity have been extensively studied [126]. The PPSV23 is an inactivated vaccine that uses purified capsular polysaccharides, and is typically administered to older adults (65+) and susceptible younger individuals [127, 128, 129, 130]. In our analysis, we focused on the vaccine's potential pathways of action. Our initial saliva investigation in PPSV23 established that an immune response to the vaccination can be detected utilizing non-invasive saliva monitoring at the molecular level [88]. Since aggregate saliva was sampled, we expected that multiple immune cells

contained therein are involved in the observed patterns and associated immune responses. Based on our previous findings and general vaccine responses, we anticipated the activation of pathways involved with antigen presentation and processing, regulation of IgM and B/T cells, Lymphoid cells, MHC molecules, and phagocytosis. We also expected the activation of pathways of general immune response to stimuli or inflammation.

### **3.4.1 Perturbation Induces Temporal Responses**

Communities aid in defining the genes' collective behavior, and observing the collective behavior of communities in the entire network can clarify relative trends between these collective behaviors. The generated heatmaps for each community depicted gene regulation for individual time points, and also displayed trends over time within the identified communities. The trends we observed in our saliva data were consistent with a time-dependent regulation. The results suggest a sequence of communities activations (up- and down-regulation) at individual timepoints, indicative of sequential immune system responses due to the PPSV23 vaccination. In the primary B cell, data were less clear, as fewer time points were monitored, and also the network was more densely connected. The B-cell heatmaps still indicate overall trends associated with Rituximab activation (both up- and down-regulation) within the first 7 hrs of the treatment. Our future work will focus on the possibility of establishing a causal chained signaling response, and associated pathways across these communities.

Our analysis showed the applicability of a DN approach in evaluating time course RNA-seq data. Specifically, the DN method results in the saliva experiment data were consistent with our previous work on profiling PPSV23 vaccination responses [88]. For the primary B cell responses to Rituximab, the DN has found the same signaling pathway as numerous prior experimental results, thus helping with our validation from a computational perspective.

The DN approach complements prior studies by offering a systems-level network perspective of aggregate temporal changes due to drug activation. In future work we plan to address the identification of sequential activation of network communities, as well determining directionality/causality in such activations.

# Chapter 4. Mean Field Theory of a Non-Reciprocal Hopfield Network

## 4.1 Introduction

The Hopfield model [15] is a spin glass model introduced to describe neural networks. It addresses the issue of content-addressable, or associative, memory, i.e., how some complex extended systems are able to recover a host of memories using only partial or noisy information. The statistical properties of the Hopfield models have been extensively investigated (see, e.g., [131, 132] for a review of earlier works). In typical Hopfield models, neural interactions are symmetric, but as Hopfield pointed out [15], the introduction of asymmetric interactions can result over time in transitions between memory patterns. In addition to the Hebbian coupling,

$$J_{ij} = \frac{1}{N} \sum_{\nu=1}^p \xi_i^{\nu} \xi_j^{\nu} , \quad (4.1)$$

where  $\xi_i^{\nu}$ , with  $\nu = 1, \dots, p$  are spin memory patterns, one can introduce asymmetric interactions of the form

$$J'_{ij} = \frac{\lambda}{N} \sum_{\nu=1}^q \xi_i^{\nu+1} \xi_j^{\nu} , \quad (4.2)$$

with  $q < p$ . With this modification, some of the spin memory patterns become metastable and can be replaced in time by other patterns. This allows for the storage and retrieval of a limited number of temporal sequences of spin patterns. While incoherent asymmetry acts as a noise mechanism that can help stabilize memory retrieval [133], asymmetric interactions of the form in Eq. 4.2 enable coherent pattern evolution in time. Moreover, the addition of terms of the form in Eq. 4.2 makes the spin system non-reciprocal. Recent studies have

examined how non-reciprocity can induce novel classes of phase transitions that cannot be described using a free energy [134, 135, 19].

Dynamical spin models that can describe coherent temporal sequences, such as the class of Hopfield models above, are particularly interesting in the study of out-of-equilibrium processes. These models have recently been applied beyond modeling brain functions to many biological and biomedical systems, such as models of cell reprogramming [136, 137], classification of disease subtypes [138], or disease progression models [139, 140]. In particular, Szedlak et al. [141] used a Hopfield model with both terms in Eqs. 4.1 and 4.2 to describe the dynamics of gene expression patterns in the cell cycle of cancer and yeast cells. A key finding of the paper was the necessity of finely adjusting the model's parameters, specifically the noise level and the relative strength between symmetric and asymmetric interactions, embodied by the parameter  $\lambda$  in Eq.4.2. This adjustment guarantees that the model maintains cyclic behavior while remaining sufficiently responsive to perturbations, such as targeted inhibitions that result in observable changes. This parameter tuning aligns with the idea of operating at the "edge of chaos", where biological systems exhibit both maximal robustness and sensitivity to external conditions [142].

Here, we study two-memory Hopfield networks with  $N$ -sites, characterized by asymmetric interactions that drive the system toward a critical threshold of oscillatory instability. The non-reciprocity leads to time-reversal symmetry breaking and introduces an extended region of criticality in the phase diagram, a feature typically observed in biological systems [143, 144]. Similar behavior can be observed in other classes of non-reciprocal kinetic Ising models with on-site interactions between two different types of spin [145]. These asymmetric models can exhibit noise-induced interstate switching leading to non-equilibrium currents or oscillations [146]. Biological networks often operate far from the  $N \rightarrow \infty$  limit. For instance, the cell

cycle program only involves hundreds of genes. The role of fluctuations and their dependence on  $N$  becomes, therefore, critical in their dynamical behavior. Here, we focus on the role of fluctuations in dynamical phase transitions to limit cycles. We find that the limit cycle phase is bounded by two critical lines: a Hopf bifurcation line and a fold bifurcation line. The autocorrelation function  $C(\tau)$  on these lines scales as  $C \sim \tilde{C}(\tau/N^\zeta)$ , where  $\tilde{C}$  are universal scale-invariant functions and  $\zeta$  is a dynamical critical exponent previously introduced to characterize out-of-equilibrium critical behavior [147]. The dynamical exponent  $\zeta = 1/2$  on the Hopf line and  $\zeta = 1/3$  on the fold line. The sensitivity to an external perturbation of strength  $F$  in these two critical regions also differs. On the Hopf line, the system exhibits enhanced sensitivity to periodic perturbations resonant with the limit cycle frequency and features a response time that scales as  $|F|^{-2/3}$ . In contrast, an external bias on the fold line can only induce switches between memory patterns in a limited and controlled way, without ever pushing the state into sustained limit cycles. Moreover, the characteristic response time is faster and scales as  $|F|^{-1/2}$ . While it was established that Hopf oscillators form a dynamic universality class relevant in biology, such as in the sensitivity of hair cells in the cochlea [148, 149], the fold line identifies a distinct critical behavior that could help understanding transitions from stable points to cycles or more complex multi-step biological programs.

In Sect. 4.2, we introduce a two-memory non-reciprocal Hopfield model and analyze its phase diagram in a mean-field approximation. We show that the dynamical phase diagram is characterized by a cyclic behavior phase, bounded by two critical lines, Hopf and fold bifurcation lines. In Sect. 5.1 we examine the critical properties of the system on and near these two lines using analytical methods. We introduce an exact form of the Master Equation for the system in Sect. 6.1, explicitly accounting for the spin symmetry under pattern

exchange. Based on this Master Equation approach, we explore the system in the large  $N$  limit using a Glauber Monte Carlo procedure in Sect. 7.1 In Sect. 7.1.1, we numerically test the critical behavior and dynamic critical exponents derived analytically in Sect. 5.1. Finally, we summarize conclusions in Sect. 7.2.

## 4.2 Cyclic Hopfield networks

We consider Hopfield networks with  $N$  Ising spins  $\sigma_i = \pm 1$  interacting through non-reciprocal couplings  $J_{ij} \neq J_{ji}$ . We focus on a network encoding two memory patterns,  $\xi_i^1$  and  $\xi_i^2$ , with couplings of the form

$$J_{ij} = \frac{\lambda_+}{N} (\xi_i^1 \xi_j^1 + \xi_i^2 \xi_j^2) + \frac{\lambda_-}{N} (\xi_i^1 \xi_j^2 - \xi_i^2 \xi_j^1). \quad (4.3)$$

The term proportional to  $\lambda_+$  describes the Hebbian coupling, while  $\lambda_-$  introduces a bias between the two memory patterns. By applying the Mattis gauge transformation [150] to the spins  $\sigma_i \rightarrow \xi_i^1 \sigma_i$ ,  $J_{ij}$  reduces to

$$J_{ij} = \frac{\lambda_+}{N} (1 + \xi_i \xi_j) + \frac{\lambda_-}{N} (\xi_j - \xi_i), \quad (4.4)$$

where  $\xi_i = \xi_i^1 \xi_i^2$ , which is equivalent to setting the first memory pattern to all spin up. The symmetric case, with  $\lambda_- = 0$ , has been previously introduced and solved by Van Hemmen [151]. In this two-memory Hopfield network, the  $N$  spins separate into two sub-networks, which we call similarity ( $S$ ) and differential ( $D$ ) subnetworks [152]:  $S$  corresponding to spins with  $\xi_i = 1$  (i.e.  $\xi_i^1 = \xi_i^2$ ) and  $D$  corresponding to spins with  $\xi_i = -1$  (i.e.,  $\xi_i^1 = -\xi_i^2$ ). We

can then define two magnetizations along the two memory patterns:

$$m_1 = \frac{1}{N} \sum_{j \in S, D} \sigma_j \quad (4.5)$$

$$m_2 = \frac{1}{N} \sum_{j \in S} \sigma_j - \frac{1}{N} \sum_{j \in D} \sigma_j. \quad (4.6)$$

### 4.2.1 Mean field solution

To describe the dynamics of this Ising system we consider the probability  $p_S(\sigma, t)$  that a spin in the subnetwork  $S$  has spin  $\sigma$  at time  $t$ . This probability satisfies in the large  $N$  limit the equation [153, 154]

$$\frac{\partial p_S(\sigma, t)}{\partial t} = -p_S(\sigma)w_S(\sigma) + p_S(-\sigma)w_S(-\sigma), \quad (4.7)$$

here  $w_S(\sigma)$  is the spin-flip transition rate. Assuming the system is in a thermal reservoir with inverse temperature  $\beta = 1/k_B T$ , the spin-flip transition rates take the form

$$w_S(\sigma) = (1 - \sigma \tanh \beta h_S) / 2\tau_0, \quad (4.8)$$

where the field  $h_S$  in the subnetwork  $S$  and can be written as

$$h_S = (\lambda_+ - \lambda_-)\langle m_1 \rangle + (\lambda_+ + \lambda_-)\langle m_2 \rangle. \quad (4.9)$$

The  $\langle m_{1(2)} \rangle$  are expectation values of Eqs. 4.5 and 4.6, and  $\tau_0$  is an arbitrary constant that determines the time scale of Ising dynamics, originally introduced in one dimension by Glauber [155] and extended to higher dimensions by Suzuki and Kubo [156]. The equation

for  $p_D(\sigma, t)$ , the probability for a spin  $\sigma$  in the differential subnetwork  $D$ , is similar to the one in Eq. 4.7 but with a field:

$$h_D = (\lambda_+ + \lambda_-)\langle m_1 \rangle - (\lambda_+ - \lambda_-)\langle m_2 \rangle. \quad (4.10)$$

A mean-field system of equations for  $\langle m_1 \rangle$  and  $\langle m_2 \rangle$  can then be obtained from the equations for  $p_S(\sigma, t)$  and  $p_D(\sigma, t)$  as

$$\dot{\langle m_1 \rangle} = -\frac{\langle m_1 \rangle}{\tau_0} + \frac{1}{\tau_0} \{n_S \tanh \beta [\lambda_a \langle m_1 \rangle + \lambda_s \langle m_2 \rangle] + n_D \tanh \beta [\lambda_s \langle m_1 \rangle - \lambda_a \langle m_2 \rangle]\} \quad (4.11)$$

$$\dot{\langle m_2 \rangle} = -\frac{\langle m_2 \rangle}{\tau_0} + \frac{1}{\tau_0} \{n_S \tanh \beta [\lambda_a \langle m_1 \rangle + \lambda_s \langle m_2 \rangle] - n_D \tanh \beta [\lambda_s \langle m_1 \rangle - \lambda_a \langle m_2 \rangle]\} \quad (4.12)$$

where the coupling constants  $\lambda_a = \lambda_+ - \lambda_-$  and  $\lambda_s = \lambda_+ + \lambda_-$  account for the asymmetric and symmetric components of the interaction, respectively. The coefficients  $n_{S(D)} = N_{S(D)}/N$  take into account the number of spins in the similarity and differential networks. Unless indicated, we will assume the case  $N_D = N_S$  below. To simplify the notation, we drop the  $\langle \rangle$  and assume mean field variables.

In Fig. 4.1, the phase portrait of the mean-field equations is presented. For  $\lambda_- = 0$ , a phase transition occurs at  $\beta\lambda_+ = 1$ . This transition separates the paramagnetic phase, (see orbits in **a**), from the memory retrieval phase (see orbits in **f**).

In the latter phase, either  $m_1 \neq 0$  or  $m_2 \neq 0$ , and  $|m_{1(2)}| \rightarrow 1$  when  $\beta\lambda_+ \gg 1$ . Symmetric steady-state solutions with  $m_1 = m_2 \neq 0 \leq 1/2$  are observed for  $\beta\lambda_+ > 1$ . These solutions, represented as empty circles in **f**, are mixed memory states equidistant from the two patterns. Such symmetric states are saddle point solutions and are always unstable in a two-memory scenario. In this reciprocal two-memory model, mixed asymmetric solutions are not permissi-

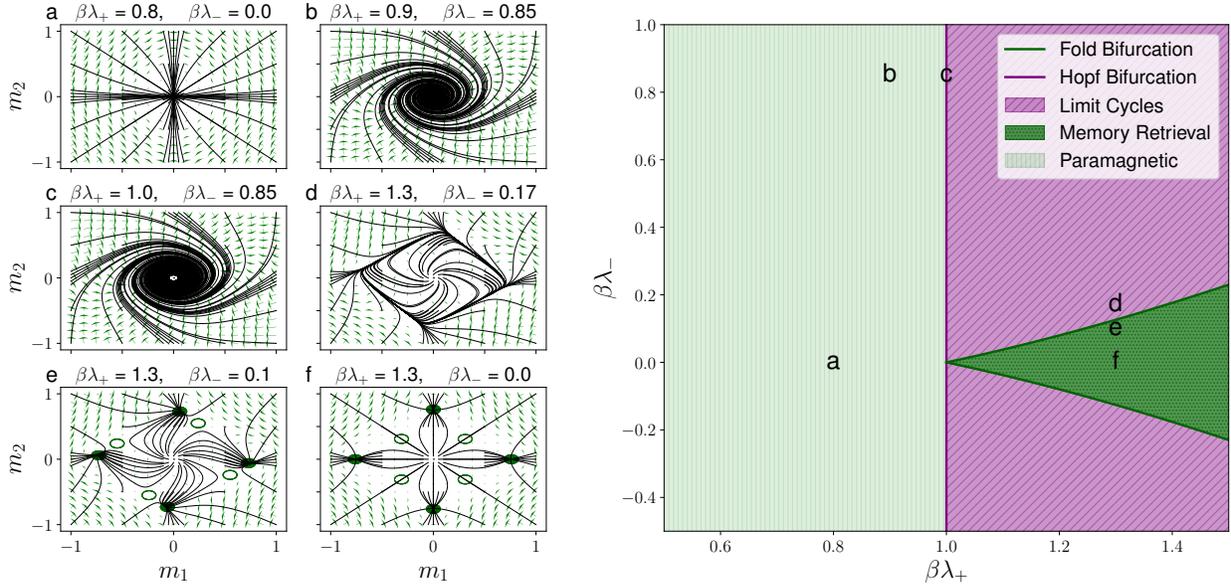


Figure 4.1: Phase portraits (left) and the phase diagram (right) for the two-memory cyclic Hopfield model. Left: The panel shows the dynamical behavior for different values of  $\lambda_+$  and  $\lambda_-$ . The six phase portraits (**a** to **f**) show the trajectory dynamics, with green arrows indicating the vector fields of the derivatives. In **e** and **f**, empty circles represent saddle points, while solid circles denote stable points (sinks). Right: The phase diagram is divided into three regions of different dynamical behavior: Limit Cycles (diagonal lines with a purple background), Memory Retrieval (dotted dark green background), and Paramagnetic (vertical stripes with a light green background). These phases are bounded by bifurcation lines: fold bifurcation lines (dark green lines) and the Hopf bifurcation (purple vertical line). The positions of the six trajectory plots (**a** - **f**) are indicated by corresponding labels on the phase diagram.

ble, in contrast to what is observed in Hopfield networks with more than two patterns, as shown by Amit et al.[157].

Next, we explore how these stable and unstable fixed points change in the presence of the asymmetric interaction  $\lambda_-$ . We rewrite Eqs. 4.11 and 4.12 in compact form as  $\dot{\mathbf{m}} = F(\mathbf{m}, \lambda_+, \lambda_-)$ , where  $\mathbf{m}$  is the vector of magnetizations. In a neighborhood of  $\overline{\mathbf{m}}$  which is a solution of  $F(\mathbf{m}, \lambda_+, \lambda_-) = 0$ , we can linearize the mean field equations as

$$\dot{\mathbf{m}} = \mathbf{A} \cdot (\mathbf{m} - \overline{\mathbf{m}}) \quad (4.13)$$

where the Jacobian matrix  $\mathbf{A}$  at  $\overline{\mathbf{m}}$  can be expressed as

$$\mathbf{A} = \begin{bmatrix} -1 + \beta\lambda_+\Delta + 2\beta\lambda_-\Gamma & \beta\lambda_-\Delta - 2\beta\lambda_+\Gamma \\ -\beta\lambda_-\Delta - 2\beta\lambda_+\Gamma & -1 + \beta\lambda_+\Delta - 2\beta\lambda_-\Gamma \end{bmatrix} \quad (4.14)$$

with  $\Delta = 1 - \overline{m_1}^2 - \overline{m_2}^2$  and  $\Gamma = \overline{m_1} \overline{m_2}$ .

The stability of steady-state solutions is determined by the eigenvalues of the Jacobian matrix  $\mathbf{A}$ . This matrix has either two real or two complex conjugate eigenvalues. In the scenario where  $\overline{\mathbf{m}} = 0$ , the eigenvalues of  $\mathbf{A}$  are  $\mu_{\pm} = \beta\lambda_+ - 1 \pm i\beta\lambda_-$ , revealing that for  $\beta\lambda_+ < 1$  and  $\lambda_- \neq 0$ , the solution  $\overline{\mathbf{m}} = 0$  is stable, with focus-type orbits (see **b**). The line in  $\beta\lambda_+ = 1$ , where the eigenvalues transition to being purely imaginary is a Hopf bifurcation line. This line is the projection of the curve defined in  $(m_1, m_2, \lambda_+, \lambda_-)$  by  $F(\mathbf{m}, \lambda_+, \lambda_-) = 0$  and  $Tr[\mathbf{A}(\mathbf{m}, \lambda_+, \lambda_-)] = 0$  on the  $(\lambda_+, \lambda_-)$  plane [158].

When  $\beta\lambda_+ > 1$ , the phase diagram splits into two distinct regions, contingent upon the existence of solutions with  $\overline{\mathbf{m}} \neq 0$ . The boundary between these regions is a fold bifurcation line (also called saddle-node bifurcation line) derived from projecting the curve

in  $(m_1, m_2, \lambda_+, \lambda_-)$ , defined by  $F(\mathbf{m}, \lambda_+, \lambda_-) = 0$  and  $\text{Det}[\mathbf{A}(\mathbf{m}, \lambda_+, \lambda_-)] = 0$  onto the  $(\lambda_+, \lambda_-)$  plane [158]. Along this line, a single real eigenvalue transitions to zero while its counterpart maintains a negative value. This behavior can be interpreted as a merging of the memory retrieval fixed points, which are stable node-type, and the mixed memory states, which are saddle points. In this system, the non-reciprocal parameter  $\lambda_-$  shifts the fixed points, causing four memory retrieval fixed points to approach the mixed memory states progressively. This convergence facilitates a circular directionality in the orbits, acting as a harbinger for the limit cycle solutions apparent above the fold bifurcation line, where only the unstable solution  $\bar{\mathbf{m}} = 0$  persists.

## 4.2.2 Near cusp dynamics

The limit cycle phase can be better described by introducing a complex variable  $z = m_1 - im_2$ . By approximating  $\tanh(x) \approx x - x^3/3$  in Eqs. 4.11 and 4.12, we obtain an equation for  $z(t/\tau_0)$ :

$$\dot{z} = (\Lambda - 1)z - \frac{\Lambda^2 \bar{\Lambda}}{2} z^2 \bar{z} + \frac{\bar{\Lambda}^3}{6} \bar{z}^3, \quad (4.15)$$

where  $\Lambda = \beta(\lambda_+ + i\lambda_-)$ , the bar indicates complex conjugation and the dot refers to derivation with respect to  $t/\tau_0$ . The last term in Eq.4.15 is an anti-resonant term, which can be eliminated using a smooth change of variables:

$$w = z - \frac{h(\Lambda, \bar{\Lambda})}{6} \bar{z}^3, \quad (4.16)$$

where  $h(\Lambda, \bar{\Lambda}) = \bar{\Lambda}^3/(3\bar{\Lambda} - \Lambda - 2)$ . By substituting Eq.4.16 into Eq.4.15 and retaining only

terms up to the cubic order in  $w$ , we obtain the Poincaré normal form:

$$\dot{w} = (\Lambda - 1)w - \frac{\Lambda^2 \bar{\Lambda}}{2} w^2 \bar{w}. \quad (4.17)$$

For  $\rho(t) = |w(t)|$ , we can then write:

$$\dot{\rho} = \rho \left[ \beta\lambda_+ - 1 - \frac{(\beta\lambda_+)^2 + (\beta\lambda_-)^2}{2} (\beta\lambda_+) \rho^2 \right], \quad (4.18)$$

which indicates that non-zero steady solutions exist for  $\beta\lambda_+ > 1$ . Focusing near the cusp point at  $\beta\lambda_+ = 1$  and  $\beta\lambda_- = 0$  and retaining terms only up to the first order in  $(\beta\lambda_+ - 1)$  and  $\beta\lambda_-$ , we find that the amplitude of the limit cycles increases with  $\beta\lambda_+$  as:

$$\rho_0^2 = 2(\beta\lambda_+ - 1). \quad (4.19)$$

To observe the change in the dynamical behavior corresponding to the fold line, we can write the equation for the phase  $\theta(t) = \arg[z(t)]$  from Eq.4.15, which keeps the anti-resonant term proportional to  $\bar{z}^3$ . Then, retaining terms up to the first order in  $(\beta\lambda_+ - 1)$  and  $\beta\lambda_-$  we have:

$$\dot{\theta} = \beta\lambda_- - \frac{\beta\lambda_+ - 1}{3} \sin 4\theta. \quad (4.20)$$

Using this equation, we can determine the period of the limit cycles as:

$$\frac{T}{\tau_0} = \frac{1}{4\beta\lambda_-} \int_0^{8\pi} \frac{d\theta}{1 - \alpha \sin \theta} = \frac{1}{\beta\lambda_-} \int_0^{2\pi} \frac{d\theta}{1 - \alpha \sin \theta} = \frac{2\pi}{\beta\lambda_-} \frac{1}{\sqrt{1 - \alpha^2}}, \quad (4.21)$$

where  $\alpha = (\beta\lambda_+ - 1)/3\beta\lambda_-$ . Near the vertical Hopf line, the period is only determined by  $\beta\lambda_-$ . As we move right in the region with  $\beta\lambda_+ > 1$ , the period increases and then diverges

when we approach the fold line, which, near the cusp point, corresponds to<sup>1</sup>

$$\beta\lambda_- = \frac{\beta\lambda_+ - 1}{3}. \quad (4.22)$$

---

<sup>1</sup>The zero temperature  $\beta = \infty$  limit of the Eqs. 4.11,4.12 can be studied by noting that  $\tanh \beta x \rightarrow \text{sgn } x$  for  $\beta \rightarrow \infty$ . The only possible values for the steady states  $m_1$  and  $m_2$  are the ones compatible with the half sum of two sign functions, which can only give 0,  $\pm 1$  or  $\pm 1/2$ . The solutions with  $|m_1| = 1$  and  $|m_2| = 0$  and vice-versa describe the perfect memory retrieval. Consider the solution  $m_1 = 1$  and  $m_2 = 0$ . By replacing these values in Eq. 4.11, we find  $1 = (\text{sgn } \lambda_a + \text{sgn } \lambda_s)/2$  which is possible only for  $\lambda_+ > \lambda_-$ . If this condition is violated, the dynamics has limit cycles [154]. The equation for the fold line, which is given by Eq. 4.22 near the cusp, changes asymptotically to  $\beta\lambda_- = \beta\lambda_+$  for  $\beta$  approaching  $\infty$ . Also, the unstable mixed memory solutions with  $|m_1| = |m_2| = 1/2$  exist in this limit only for  $\lambda_- = 0$ . This can be verified, for instance, by replacing the solution  $m_1 = m_2 = 1/2$  in Eqs. 4.11 and 4.12, which give  $\text{sgn } \lambda_+ + \text{sgn } \lambda_- = 1$  and  $\text{sgn } \lambda_+ - \text{sgn } \lambda_- = 1$ , and is possible only for  $\lambda_- = 0$ .

# Chapter 5. Critical Cyclic Behavior

## 5.1 Critical Dynamics

To study the effect of fluctuations near the critical lines we modify Eq. (4.15) as

$$\dot{z} = (\Lambda - 1)z - \frac{|\Lambda|^2 \Lambda}{2} |z|^2 z + \frac{\bar{\Lambda}^3}{6} \bar{z}^3 + \frac{1}{\sqrt{N}} \zeta(t), \quad (5.1)$$

where we have included the complex-valued white noise variable  $\zeta(t)$ ,

$$\langle \zeta(t) \bar{\zeta}(t') \rangle = D \delta(t - t'), \quad (5.2)$$

to account for noise beyond the mean-field equation. In this section, we put  $\tau_0 = 1$  to simplify the notation. The constant  $D$  is phenomenological, and the scaling with the system size  $N$  is chosen to match the standard mean-field equation plus noise for collective models (see, e.g., [147]). Let us consider the following distinct regions.

### 5.1.1 Near the cusp with $\beta\lambda_- = 0$ and $\beta\lambda_+ \approx 1$

In this case, the equation is better represented in terms of  $m_1$  and  $m_2$ . In the absence of the asymmetric term, the dynamics is governed by a free energy  $\mathcal{F}$  as

$$\dot{m}_i = -\frac{\partial \mathcal{F}}{\partial m_i} + \frac{1}{\sqrt{N}} \xi_i(t), \quad (5.3)$$

with a real white noise

$$\langle \xi_i(t) \xi_j(t') \rangle = D \delta_{ij} \delta(t - t'), \quad (5.4)$$

where

$$\mathcal{F} = \frac{r}{2}(m_1^2 + m_2^2) + u_1 m_1^4 + u_2 m_2^4 + 2u_{12} m_1^2 m_2^2. \quad (5.5)$$

Here  $r = 1 - \beta\lambda_+$  and  $u_1 = u_2 = \frac{(\beta\lambda_+)^3}{6} \approx \frac{1}{6}$ ,  $u_{12} = \frac{(\beta\lambda_+)^3}{2} \approx \frac{1}{2}$ . Note that the model exhibits a  $Z_2 \times Z_2$  symmetry. The phase diagram is determined by the sign of  $r$  and  $u_1 u_2 - u_{12}^2$ . Since  $u_1 u_2 < u_{12}^2$ , there are only three phases:  $m_1 = m_2 = 0$  when  $\beta\lambda_+ < 1$  and either  $m_1 \neq 0 = m_2$  or  $m_1 = 0 \neq m_2$  when  $\beta\lambda_+ > 1$ . This is analogous to a multi-critical point in a magnetic system where anisotropies break the  $O_n$  rotational symmetry along more than one direction (see, e.g., Sect. 4.6 of Ref. [159]).

To understand the critical behavior at the critical point  $\beta\lambda_+ = 1$  (and  $\lambda_- = 0$ ), we consider the stochastic Langevin equation

$$\dot{m}_1 = -(4u_1 m_1^3 + 4u_{12} m_1 m_2^2) + \frac{1}{\sqrt{N}} \xi(t), \quad (5.6)$$

and a similar equation for  $m_2$ . The linear term vanishes since  $r = 0$  at the critical point. Now, a rescaling of time and field variables,

$$\tilde{t} = t/N^{1/2}, \quad \tilde{m}_i = N^{1/4} m_i, \quad (5.7)$$

leads to a scale-invariant equation (i.e., independent of  $N$ ) as

$$d\tilde{m}_1/d\tilde{t} = -(4u_1 \tilde{m}_1^3 + 4u_{12} \tilde{m}_1 \tilde{m}_2^2) + \xi(\tilde{t}), \quad (5.8)$$

and similarly for  $m_2$ . This observation leads to useful scaling relations. For example, the

two-time correlation function for  $t, \tau \gg \tau_0$  can be written as

$$C_{ij}(t, \tau) = \langle m_i(t + \tau) m_j(t) \rangle \sim \delta_{ij} N^{-1/2} \tilde{C}(\tau/\sqrt{N}), \quad (5.9)$$

where  $\tilde{C}$  is a universal scaling function. The Kronecker delta function follows from the  $Z_2 \times Z_2$  symmetry of the model.

### 5.1.2 Hopf bifurcation line: $\beta\lambda_- \neq 0$ while $\beta\lambda_+ = 1$

In this case, we use the transformation  $w = e^{-i\beta\lambda_- t} z$ , and we assume that the oscillation is sufficiently fast to neglect the anti-resonant terms, which can be viewed as a rotating wave approximation. The resulting equation for  $w$  becomes

$$\dot{w} = -rw - \frac{1}{2}|w|^2 w + \frac{1}{\sqrt{N}} \zeta(t), \quad (5.10)$$

where we have replaced  $\Lambda \approx 1$  in the nonlinear term. Note that the rotating wave approximation is equivalent to the Poincaré transformation in Eq. (4.16) for large  $\beta\lambda_-$  and  $\beta\lambda_+ = 1$ . While the Poincaré method is more general and also valid in the small  $\beta\lambda_-$  limit, we will discuss this case using the rotating wave ansatz, which provides a more intuitive interpretation. Interestingly, the symmetry is now  $O(2)$  rather than  $Z_2 \times Z_2$ . We can still describe the dynamics by a free energy defined as

$$\tilde{\mathcal{F}} = r|w|^2 + \frac{1}{4}|w|^4. \quad (5.11)$$

Similar approaches have appeared before [160, 145, 161, 162]. Scaling relations similar to the ones in Eq. (5.7) at the critical point  $r = 0$

$$\tilde{t} = t/N^{1/2}, \quad \tilde{w} = N^{1/4}w, \quad (5.12)$$

leads to a scale-invariant equation, and to

$$\langle w(t)\bar{w}(0) \rangle = N^{-1/2}\tilde{C}(t/\sqrt{N}) . \quad (5.13)$$

The universal scaling function  $\tilde{C}$  differs from the previous case because the underlying symmetries and the dynamics are different. We will show in Sect. 7.1.1 that this scaling behavior is consistent with numerical simulations.

### 5.1.3 Limit cycle phase near the Hopf line

The continuous  $O(2)$  symmetry breaking in the ordered (limit cycle) phase in the regime where the rotating wave approximation applies results in a Goldstone mode, which is susceptible to noise. Defining  $w = \rho_0 e^{i\vartheta}$ , the dynamics of the phase is given by

$$\dot{\vartheta} = \frac{1}{\rho_0\sqrt{N}}\xi(t) , \quad \langle \xi(t)\xi(t') \rangle = D\delta(t-t') . \quad (5.14)$$

It then follows that  $\langle (\vartheta(t) - \vartheta(0))^2 \rangle \sim Dt/(\rho_0^2 N)$ . Therefore, the oscillations in the limit cycle phase are suppressed due to noise at any finite  $N$  as (restoring  $z = e^{i\beta\lambda - t}w$ )

$$C_z(t, \tau) = \langle z(t+\tau)\bar{z}(t) \rangle = \rho_0^2 e^{i\beta\lambda - \tau - D\tau/(\rho_0^2 N)} , \quad (5.15)$$

thus oscillations remain coherent up to a characteristic time  $T \sim \rho_0^2 N$ . Below, we will show how the exact master equation approach and Glauber simulations reproduce this damping effect for finite systems. Deep in the limit cycle phase and/or closer to the fold line, the limit cycle dynamics is not uniform (i.e., not governed by a single frequency). However, we later show that the oscillations are similarly damped.

### 5.1.4 Fold line

Fluctuations on the fold line can be studied by adding a white noise term to Eq. 4.20. As one approaches the transition line  $\beta\lambda_- = (\beta\lambda_+ - 1)/3$ , the frequency of the limit cycle vanishes, and  $\theta$  describes a soft mode. After making the transformation  $\theta \rightarrow \theta + \pi/8$  and re-scaling noise strength and time using appropriate powers of  $\beta\lambda_-$ , we obtain the stochastic equation

$$\dot{\theta} = 1 - \cos(4\theta) + \frac{1}{\sqrt{N}}\xi(t). \quad (5.16)$$

Expanding around a fixed point, say  $\theta = 0$ , we find to the first nonzero order

$$\dot{\theta} \approx 8\theta^2 + \frac{1}{\sqrt{N}}\xi(t). \quad (5.17)$$

It follows from this equation that small but negative  $\theta$  slowly converges to  $\theta = 0$  while small but positive  $\theta$  slowly diverges from  $\theta = 0$  before a quick phase slip occurs from  $0^+ \rightarrow \pi/2^-$ .

For an initial condition with  $\theta(t=0) < 0$ , the phase variable converges to  $\theta = 0$  as

$$\theta(t) \sim -\frac{1}{8t}, \quad t \rightarrow \infty. \quad (5.18)$$

The divergence for  $\theta_0 = \theta(t = 0) > 0$  is slow as well: the phase variable spends a time of the order  $t \sim 2/\theta_0$  near  $\theta = 0$  before a quick escape to a value close to, but below,  $\theta = \pi/2$ . Without noise, depending on the initial condition, the phase variable converges to one of the fixed points (in the above scenario, it would be  $\theta = 0, \pi/2$ ). However, the noise qualitatively changes this picture.

As shown above, without nonlinearity, the noise will induce a mean square displacement given by  $\langle(\theta(t) - \theta_0)^2\rangle = 2Dt/N$ . Therefore, even with  $\theta_0 < 0$ , noise would induce excursions to  $\theta > 0$ , followed by a long plateau, and then a quick slip to  $\pi/2^-$ . This is again followed by a noise-induced excursion to  $\pi/2^+$  slightly above  $\pi/2$ , another long plateau, and then a phase slip to  $\pi^-$ , and so on and so forth; see Fig. 5.1. The resulting effect is a slow net rotation of the complex order parameter. Note that this rotation disappears as  $N \rightarrow \infty$  since the noise is suppressed. The following argument gives the dynamical scaling behavior in  $N$ : Suppose we are close to  $\theta = 0$ . At short times, the nonlinearity is unimportant, while the noise induces a displacement of the order of  $\theta(t)^2 \sim t/N$ . At a sufficiently long time  $t_*$ ,

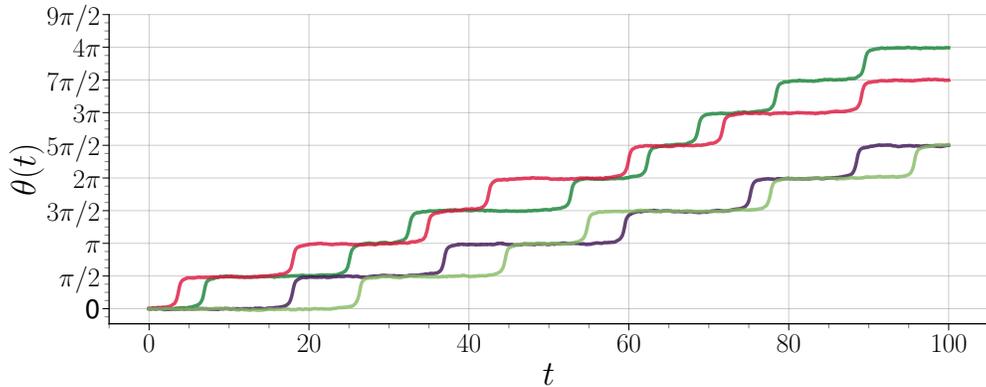


Figure 5.1: Representative trajectories at the fold transition as a function of  $t$  with  $N = 1300$  and  $D = 1$ . One can notice several features that are absent (deep) in the limit cycle phase: First, there is a larger variation between different trajectories. This highlights the role of noise in inducing phase slips. Second, the interval between jumps is rather long, and scales as  $T \sim N^{1/3}$ , in this case roughly of the order  $T \sim 10$  compared to deep in the limit cycle phase where it is of order 1. Again this is due to noise as the period should diverge when  $N \rightarrow \infty$ .

when  $\theta$  is sufficiently large, and importantly also positive, the nonlinearity becomes relevant, making the phase variable diverge from  $0^+$ . A slow dynamics of the order  $1/\theta(t_*)$  is followed by a quick phase slip before arriving at  $\pi/2^-$ . The time scale  $t_*$  (or rather  $\theta_* = \theta(t_*)$ ) is determined by minimizing (dropping constant factors for simplicity)

$$t_{\text{tot}} = N\theta_*^2 + \frac{1}{\theta_*}. \quad (5.19)$$

It follows that  $\theta_* \sim N^{-1/3}$  and

$$t_* \sim N^{1/3}. \quad (5.20)$$

This means that the frequency of oscillations (at the critical point) goes to zero as  $N^{-1/3}$ . This behavior is also reflected in the correlation function, which for  $t, \tau \gg \tau_0$  scales as

$$C_z(t, \tau) \sim \rho_0^2 \tilde{C}(\tau/N^{1/3}) \quad (5.21)$$

with  $\tilde{C}$  being a scale-free function. This scaling behavior is verified, and the form of the scaling functions is calculated numerically in Sect. 7.1.1.

### 5.1.5 Near the fold line

Our discussion has focused on the phase transition exactly at the fold line. We next discuss what happens slightly away from this line into either the limit cycle or the fixed memory retrieval phase. In the limit cycle phase, another scale appears away from the phase transition,  $\epsilon = \beta\lambda_- - (\beta\lambda_+ - 1)/3$  described by a modification of Eq. 5.16:

$$\dot{\theta} = 1 + \epsilon - \cos(4\theta) + \frac{1}{\sqrt{N}}\xi(t). \quad (5.22)$$

The motion is highly non-harmonic and resembles a step-wise rather than a smooth linear increase of the phase variable (hence, it is not described by a single frequency). Next we investigate whether the argument leading to Eq. 5.15 still follows and a damping with a characteristic time  $T \sim N$  appears. The noise-less version of Eq. 5.22 admits an exact solution. While the precise form of the equation is not directly used in the following discussion, we report it for completeness:

$$\theta_0(t) = 2 \tan^{-1} \left( \frac{\epsilon \tan \left( 2\sqrt{\epsilon(\epsilon+2)}(t-t_0) \right)}{\sqrt{\epsilon(\epsilon+2)}} \right). \quad (5.23)$$

The characteristic oscillation frequency can then be extracted as  $\omega_{LC} \propto 2\sqrt{\epsilon(\epsilon+2)}$ . To describe small fluctuations around this (noiseless) solution, we can take  $t_0 \rightarrow -f(t)$  and expand the equation of motion to the first order in  $f(t)$ . Since  $f(t) = \text{const}$  is an exact solution, the expansion only involves the time derivative, and we obtain

$$\frac{4\epsilon(\epsilon+2)}{\cos \left( 4t\sqrt{\epsilon(\epsilon+2)} \right) + \epsilon + 1} \dot{f} + \mathcal{O}(f^2) = \frac{1}{\sqrt{N}} \xi(t). \quad (5.24)$$

Also, we note that

$$\theta(t) = \theta_0(t) + \frac{4\epsilon(\epsilon+2)}{\cos \left( 4t\sqrt{\epsilon(\epsilon+2)} \right) + \epsilon + 1} f(t) + \mathcal{O}(f^2). \quad (5.25)$$

The same prefactor, which we denote by  $\theta_1(t)$ , appears in both equations above. We can then show that, up to a phase factor due to limit cycle oscillations, autocorrelation function scales as

$$|C_z(t, \tau)| \sim \rho_0^2 e^{-\frac{D}{2N} \theta_1^2(t)} \int_t^{t+\tau} dt' 1/\theta_1^2(t'). \quad (5.26)$$

One can then see that the above term decays with time roughly exponentially (when coarse graining the features over each cycle) approximately as  $\exp(-D\tau/2N)$ , as in Eq. 5.15. We conclude that the latter equation is more general than the assumptions that were used to derive it, and is likely valid throughout the limit cycle phase. Indeed the above equation suggests that the anharmonicity in the oscillations can be made uniform by reparametrizing the time as  $d\tilde{t} = dt/\theta_1^2(t)$ , leading to an equation similar to Eq. 5.14 that describes the dynamics of  $\theta(\tau)$ .

Near the fold line, the limit cycle frequency scales as  $\omega_{LC} \sim \sqrt{\epsilon}$ . Comparing this with the behavior on the critical line, where  $\omega \sim N^{-1/3}$ , a rescaled variable  $\epsilon N^{2/3}$  emerges governing the crossover between the two limits. This scaling follows from an application of the Arrhenius law on the other side of the phase transition where the point memory retrieval phase emerges. To this end, we consider  $\epsilon < 0$  describing the point memory phase. We can approximate the dynamics by introducing a tilted Sine-Gordon effective potential (see Fig. 5.2) :

$$V(\theta) = -(1 - |\epsilon|)\theta + \frac{1}{4} \sin 4\theta. \quad (5.27)$$

For  $\epsilon < 0$ , a small barrier emerges, whose height scales<sup>1</sup> as  $\Delta V \sim |\epsilon|^{3/2}$ . Now according to the Arrhenius law, we find the decay rate given by  $\Gamma \sim \exp(-\beta_{\text{eff}}\Delta V)$ , where the effective temperature, characterizing the noise strength, scales as  $\beta_{\text{eff}} \sim N$ . Therefore,  $\Gamma \sim \exp(-AN|\epsilon|^{3/2})$  and

$$C_z(t, \tau) \sim e^{-\Gamma\tau}. \quad (5.28)$$

Note that the same scaling variable ( $|\epsilon|^{3/2}N$ ) governs both sides of the fold transition. This is a known characteristic dependence for fluctuation-induced transitions in nonlinear systems

---

<sup>1</sup>More precisely,  $V(-x_0) - V(x_0) \sim |\epsilon|^{3/2}$  where  $V'(\pm x_0) = 0$ .

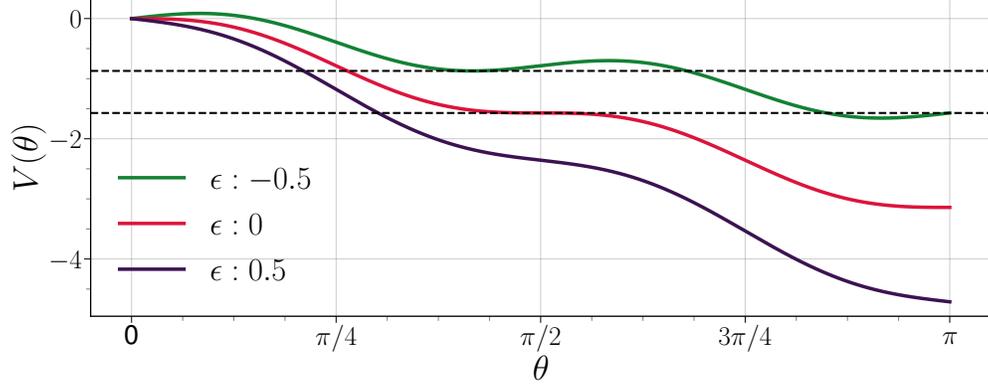


Figure 5.2: Effective potential  $V(\theta)$  for different values of  $\epsilon$ . Depending on its sign,  $\epsilon$  alters the steepness in  $V(\theta)$ , thereby affecting system's stability. A negative  $\epsilon$  initiates an uphill start in  $V$ , which poses a potential hill for  $\theta$  as a metastable state till it overcomes the hill. For  $\epsilon = 0$ ,  $V$  starts flat, then slips down at a faster rate than in the negative  $\epsilon$  scenario. A positive  $\epsilon$  triggers an immediate downhill movement in  $V$ , swiftly driving the system into the oscillatory phase at an even faster rate. The two horizontal dashed lines mark the 1<sup>st</sup> local minima of the  $\epsilon = -0.5$  and  $\epsilon = 0$  cases.

near bifurcations [163].

### 5.1.6 External drive

Let us now consider the effect of an external drive on systems at the two critical lines. We assume that the external drive has the form  $F e^{i\omega t}$ , where  $\omega$  is nearly resonant with the cycle frequency, determined by  $\lambda_-$  near the Hopf line, and approaching zero near the fold line. On the Hopf line, we can shift to a rotating frame by setting  $w = e^{-i\omega_{LC}t} z$ , which gives an equation similar to Eq. 5.10

$$\dot{w} = -i\delta w - \frac{1}{2}|w|^2 w + F \quad (5.29)$$

with the detuning  $\delta = \beta\lambda_- - \omega_{LC}$  while dropping the noise term. By rescaling to units  $\tilde{t} = tF^{2/3}$  and  $\tilde{w} = wF^{-1/3}$ , Eq. (5.29) leads to

$$w(t) \propto F^{1/3} \tilde{w}(tF^{2/3}, \delta F^{-2/3}), \quad (5.30)$$

where  $\tilde{w}$  is a parameter-free function. This suggests that, at  $\delta = 0$ , the response gain  $w/F \sim F^{-2/3}$  diverges for small perturbations. Therefore, near the Hopf line, the system behaves like a filter with larger gain for weaker perturbations. This enhanced sensitivity at criticality is known to be relevant in biological functions, such as in the auditory sensitivity of hair cells in the cochlea [149]. The scaling analysis also shows that the dynamics at criticality is slowed down by a factor  $F^{-2/3}$ , so while smaller perturbations give enhanced gain, it also takes longer for the oscillator to respond to the external drive.

The response behavior on the fold line is qualitatively different. Since, without noise, the system is frozen on the fold line, we consider a constant complex drive  $F$  with a fixed phase. Retaining only the phase dynamics, we find a modified Eq. 5.22

$$\dot{\theta} = 1 + \epsilon - \cos(4\theta) + \text{Im}(Fe^{-i\theta}). \quad (5.31)$$

The term  $e^{-i\theta}$  can only be  $\pm 1$  or  $\pm i$  except during a fast switch between the memory pattern. Consider then a system initially frozen on the fold line (without noise) or in the memory retrieval phase with a small  $\epsilon < 0$ . The last term in Eq. 5.31 can shift the value of  $\epsilon$  by  $\pm \text{Re}F$  or  $\pm \text{Im}F$ , and a switch happens only if the result is positive. Since the sign of the shift is state-dependent, the maximum number of memory switches is limited to two. In other words, a static drive  $F$  will never be able to push the system at criticality into a phase with sustained limit cycles. Such a drive can only switch between memory patterns in a limited and controlled way. Finally, near an equilibrium position, a scaling analysis shows that

$$\theta(t) \propto F^{1/2} \Theta(tF^{1/2}, \epsilon/F) \quad (5.32)$$

at short times before the switch; here,  $\Theta$  is a parameter-independent scaling function. This suggests that the response time on the fold line scales as  $F^{-1/2}$  and is faster than the  $F^{-2/3}$  dependence on the Hopf line.

# Chapter 6. Master Equation and Spectral Analysis

## 6.1 Master Equation

We now introduce a formulation for the Master Equation to describe the full dynamics of the network, allowing us to explore exactly the critical behavior studied in the previous section. Taking into account the separation of the full network into similarity and differential networks, we can rewrite the probability distribution at time  $t$  for a given configuration of all spins  $(\sigma_1, \sigma_2, \dots, \sigma_N) = \{\sigma_i\}$  as:

$$P(\{\sigma_i\}, t) = \tilde{P}(M_S, M_D, t), \quad (6.1)$$

where the variables

$$M_{S(D)} \in \left[ -N_{S(D)}, -N_{S(D)} + 2, \dots, N_{S(D)} \right]$$

identify the sum of the spin configuration  $\{\sigma_i\}$  over the subnetworks  $S$  and  $D$ . Each value of  $(M_S, M_D)$  is associated with a number of equivalent spin configurations given by:

$$g(M_S, M_D) = \binom{N_S}{N_{M_S}^+} * \binom{N_D}{N_{M_D}^+},$$

where  $N_{M_{S(D)}}^\pm = (N_{S(D)} \pm M_{S(D)})/2$  indicate the number of spins up or down for a given  $M_{S(D)}$ . This degeneracy can be taken into account by defining a probability distribution

$$P(M_S, M_D, t) = g(M_S, M_D) * \tilde{P}(M_S, M_D, t), \quad (6.2)$$

which satisfies:

$$\sum_{M_S, M_D} P(M_S, M_D, t) = 1, \quad (6.3)$$

and its dynamics are determined by the Master Equation:

$$\frac{\partial P(M_S, M_D, t)}{\partial t} = I_{in} - I_{out}, \quad (6.4)$$

where

$$\begin{aligned} I_{in} &= N_{M_S+2}^+ w_S^+(M_S+2, M_D) P(M_S+2, M_D, t) + N_{M_S-2}^- w_S^-(M_S-2, M_D) P(M_S-2, M_D, t) + \\ &\quad + N_{M_D+2}^+ w_D^+(M_S, M_D+2) P(M_S, M_D+2, t) + N_{M_D-2}^- w_D^-(M_S, M_D-2) P(M_S, M_D-2, t), \\ I_{out} &= \left[ N_{M_S}^+ w_S^+(M_S, M_D) + N_{M_S}^- w_S^-(M_S, M_D) + N_{M_D}^+ w_D^+(M_S, M_D) + N_{M_D}^- w_D^-(M_S, M_D) \right] P(M_S, M_D, t), \end{aligned} \quad (6.5)$$

with  $I_{in(out)}$  as the flux into (out of) the state  $(M_S, M_D)$ , and the  $\pm$  spin-flip transition rates defined as

$$\begin{aligned} w_S^\pm(M_S, M_D) &= \frac{1}{2\tau_0} \left( 1 \mp \tanh \frac{2}{N} [\beta\lambda_+(M_S \mp 1) - \beta\lambda_- M_D] \right) \\ &= \frac{1}{2\tau_0} \left( 1 \mp \tanh \frac{2\beta}{N} h_S^\pm(M_S, M_D) \right), \end{aligned} \quad (6.6)$$

$$\begin{aligned} w_D^\pm(M_S, M_D) &= \frac{1}{2\tau_0} \left( 1 \mp \tanh \frac{2}{N} [\beta\lambda_+(M_D \mp 1) + \beta\lambda_- M_S] \right) \\ &= \frac{1}{2\tau_0} \left( 1 \mp \tanh \frac{2\beta}{N} h_D^\pm(M_S, M_D) \right). \end{aligned} \quad (6.7)$$

The terms  $M_{S(D)} \mp 1$  in Eqs. 6.6 and 6.7 take into account the exclusion of the spin self-interaction. Previous studies have explored the effect of including versus omitting self-interaction terms in Hopfield dynamics [164, 165] and exclusion of self-interactions has been shown to lead to larger information storage capacities [165]. The effect of the  $\pm \frac{1}{N}$  is irrelevant

in the mean-field solutions discussed above, and for the remainder of the paper, we focus on the case that omits self-interaction.

The single spin flip rates in Eqs. 6.6 and 6.7 can be rewritten in terms of the local energy change  $\delta\epsilon_{S(D)}^\pm$  due to a spin-flip:

$$w_{S(D)}^\pm = \left( 1 + e^{\beta\delta\epsilon_{S(D)}^\pm} \right)^{-1}, \quad (6.8)$$

where  $\delta\epsilon_{S(D)}^\pm = -h_{S(D)}^\pm(M_S, M_D)\delta M_{S(D)}^\pm$  with  $\delta M_{S(D)}^\pm = \mp 2$ . Any cyclic process that starts from a given spin configuration and involves flipping only spins within either subnetwork  $S$  or  $D$  conserves the total energy, resulting in a net energy change of zero. However, when processes involve spins from both subnetworks  $S$  and  $D$ , the energy change depends on the cycle path. Consider, for instance, the two-spin cycle

$$\begin{aligned} (M_S, M_D) &\rightarrow (M_S - 2, M_D) \rightarrow (M_S - 2, M_D - 2) \\ &\rightarrow (M_S, M_D - 2) \rightarrow (M_S, M_D), \end{aligned} \quad (6.9)$$

where two spins up are sequentially flipped down and then back up, with the  $S$  spin flipped before the  $D$  spin. The total energy change in this case is  $\delta\epsilon = -16\lambda_-/N$ . In contrast, the time-reversed process in which the spin in  $D$  is flipped before the one in  $S$  results in  $\delta\epsilon = +16\lambda_-/N$ . This path dependence implies the violation of Kolmogorov's criterion for the transition rates [166, 167] and, therefore, breaking of the detailed balance principle.

## 6.2 Numerical diagonalization of Liouvillian

By enumerating the states using a single index  $k = (M_S, M_D)$  we can rewrite the Master Equation in Eq. 6.4 as:

$$\dot{P}(k, t) = - \sum_{k'} \mathcal{L}_{k, k'} P(k', t) \quad (6.10)$$

where  $\mathcal{L}$  is the Liouvillian matrix. The all-ones vector is always a left eigenvector of the nonsymmetric matrix  $\mathcal{L}$  with eigenvalue  $\Lambda_1 = 0$ , which guarantees the probability conservation in Eq. 6.3, and, for finite  $N$ , all the eigenvalues of the Liouvillian have a positive real part. To study the system's phase diagram, we focus in Fig. 6.1 on the second smallest eigenvalue  $\Lambda_2$  and its dependence as a function of  $N$ . Note that the real part of  $\Lambda_2$  remains nonzero

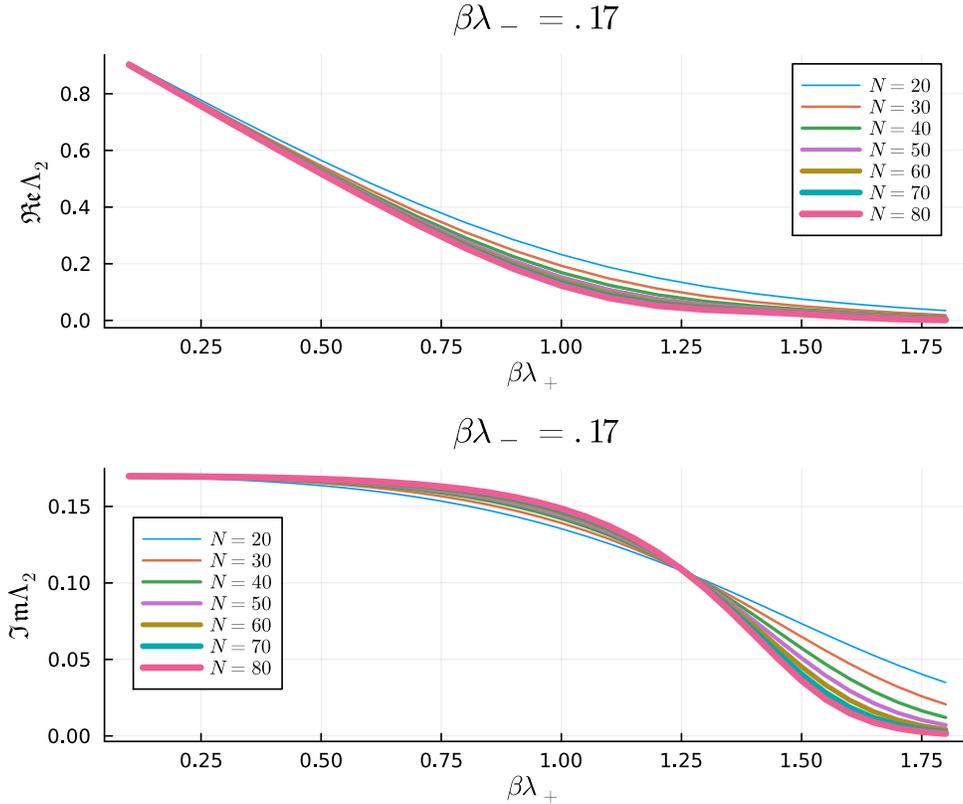


Figure 6.1: Real and Imaginary part of the second smallest eigenvalue of the Liouvillian matrix,  $\Lambda_2$ , as a function of  $\beta\lambda_+$  for a fixed value of  $\beta\lambda_- = 0.17$  and  $N_S = N_D = N/2$ .

in the region  $\beta\lambda_+ < 1$  of the phase diagram in Fig. 4.1, corresponding to the paramagnetic phase. For  $\beta\lambda_+ > 1$ , the real part of  $\Lambda_2$  converges to zero, allowing for the memory retrieval of a constant magnetization value as  $N \rightarrow \infty$ . The imaginary part of  $\Lambda_2$ , on the other hand, changes its behavior as a function of  $N$  for  $\beta\lambda_+ \sim 1.3$ , which is near the fold line of the mean-field model, separating the limit cycle and the memory retrieval phases where the oscillations disappear. Observing the sharp features of the diagram in Fig. 4.1 by analyzing the eigenvalues of the Liouvillian is computationally demanding, and even for  $N = 80$ , resulting in an  $\mathcal{L}$  of dimensions 1681 by 1681, the transitions in Fig. 6.1 are not sharply defined. Below, we will implement a Glauber Monte Carlo algorithm that allows us to explore significantly larger  $N$ .

### 6.2.1 Exact expectations and correlation functions

The Liouvillian matrix can be used to calculate exact expectation and correlation functions. For instance, given a probability distribution at  $t = 0$ ,  $P(k, 0)$ , the average magnetization along the first memory pattern as a function of time can be calculated as

$$\langle m_1(t) \rangle = \frac{1}{N} \sum_k [M_1]_k P(k, t) , \quad (6.11)$$

where

$$P(k, t) = \sum_{k'} \left[ e^{-\mathcal{L}t} \right]_{k,k'} P(k', 0) , \quad (6.12)$$

and  $M_1 = M_S + M_D$ . Similarly, the two-time correlation function for  $M_1$  can be defined as

$$C_{1,1}(t, \tau) = \frac{1}{N^2} \sum_{k, \bar{k}} [M_1]_{\bar{k}} [M_1]_k P(\bar{k}, t + \tau; k, t). \quad (6.13)$$

The joint probability  $P(\bar{k}, t + \tau; k, t)$  can be rewritten as

$$P(\bar{k}, t + \tau; k, t) = P(\bar{k}, t + \tau | k, t) P(k, t), \quad (6.14)$$

where  $P(\bar{k}, t + \tau | k, t)$  is the conditional probability of the system to be in state  $\bar{k}$  at time  $t + \tau$ , given it was in state  $k$  at time  $t$ . This conditional probability can be calculated by shifting the initial condition  $t \rightarrow 0$  and using

$$P(\bar{k}, t + \tau | k, t) = \sum_{k'} \left[ e^{-\mathcal{L}\tau} \right]_{\bar{k}, k'} P(k', 0), \quad (6.15)$$

with the initial probability set to  $P(k', 0) = \delta_{k', k}$ . The two-time correlation can then be expressed as [156]

$$C_{1,1}(t, \tau) = \frac{1}{N^2} \sum_k \langle M_1(\tau) \rangle_k [M_1]_k P(k, t), \quad (6.16)$$

where

$$\langle M_1(\tau) \rangle_k = \sum_{\bar{k}} [M_1]_{\bar{k}} \left[ e^{-\mathcal{L}\tau} \right]_{\bar{k}, k} \quad (6.17)$$

is the expectation of  $M_1$  at  $\tau$  given having been in configuration  $k$  at  $t = 0$ . Similar averages and two-time correlations can be defined for other quantities such as  $M_2$  and  $Z = M_1 - iM_2$ . Fig. 6.2 shows the exact  $\langle m_2(t) \rangle$  for different values of  $N$  calculated using the Liouvillian. The initial state was configured such that  $m_1(0) = 1$  and  $m_2(0) = 0$ , with the parameters  $\beta\lambda_+ = 1.3$  and  $\beta\lambda_- = 0.17$ . This positions the system slightly above the fold line in the phase diagram of Fig. 4.1. Fig. 6.3 shows the two-time correlation function for  $M_2$ , denoted as  $C_{2,2}(t, \tau)$ . The function is dependent on  $N$  as well. In smaller systems ( $N = 50, 100$ ),  $C_{2,2}(t, \tau)$  quickly drops to zero, indicating that  $M_2(t + \tau)$  becomes uncorrelated with  $M_2(t)$

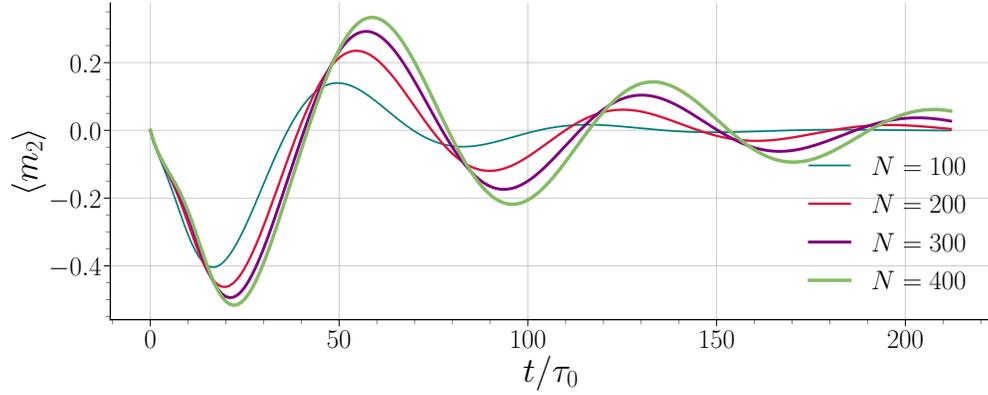


Figure 6.2: Exact  $\langle m_2(t) \rangle$  solved from the master equation for different system sizes  $N$  with  $\beta\lambda_+ = 1.3$  and  $\beta\lambda_- = 0.17$ . As  $N$  increases, the oscillations become slower and more pronounced.

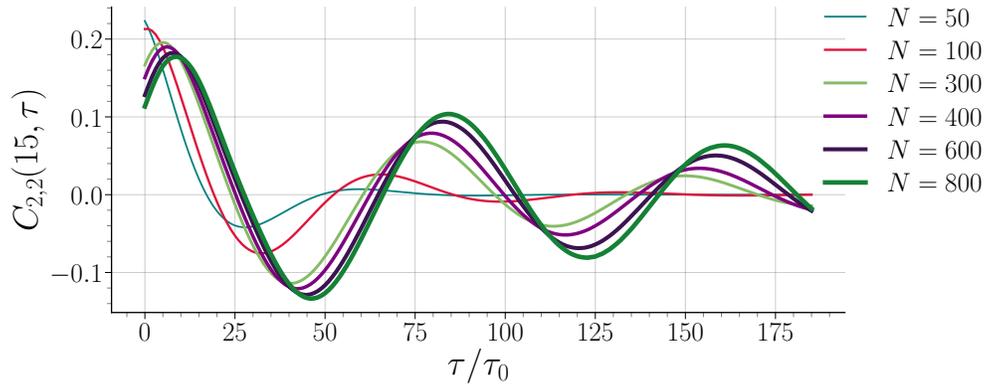


Figure 6.3: Two-time correlation function  $C_{2,2}(t, \tau)$  for  $M_2$  at  $t = 15$ ,  $\beta\lambda_+ = 1.3$  and  $\beta\lambda_- = 0.17$ , for various  $N$ . As  $N$  increases,  $C_{2,2}(t, \tau)$  exhibits defined oscillations.

as  $\tau$  increases due to fluctuations, while in larger systems, oscillations in  $C_{2,2}(t, \tau)$  emerges.

endfigure

# Chapter 7. Monte Carlo Validation with Glauber Dynamics

## 7.1 Glauber Dynamics

In parallel with deriving the master equations for  $P(M_S, M_D, t)$ , we implemented a Glauber dynamics that utilizes the division into subnets, rather than relying on random spin flips across the entire network. This adaptation not only provides a direct comparison with the predictions of the master equations but also allows us to examine much larger systems. Specifically, in our implementation we consider the total magnetizations  $M_1 = M_S + M_D$  and  $M_2 = M_S - M_D$ . Each Monte Carlo step involves a probabilistic decision to flip a spin within one of the two subnets, with the selection between  $S$  and  $D$  being randomized. The corresponding transition rates, as defined in Eqs. 6.6 and 6.7, incorporate the effects of  $\lambda_+$  and  $\lambda_-$ . Our implementation tracks these magnetizations at intervals of  $N$  iterations. Below we show results from our simulations where we varied network sizes  $N$ , with additional adjustments in interaction strengths  $\lambda_+$  and  $\lambda_-$ . We focused on assessing the system's finite-size effects and convergence towards the mean field solutions.

Fig. 7.1 illustrates the convergence of the Glauber dynamics toward the mean field solution for  $N \rightarrow \infty$ . The observations are consistent in both  $m_1(t)$  and  $m_2(t)$ . At  $N = 100$ , deviations from the mean field solution are notable, particularly in the oscillation frequency and noise levels. As  $N$  increases to 1,000 and 10,000, the discrepancies between the simulations and mean field solutions decrease, with progressively smoother magnetization dynamics. At  $N = 10,000$  and 100,000, stochastic effects significantly recede. In these larger systems, the dynamics closely resemble those of an infinite, continuous medium.

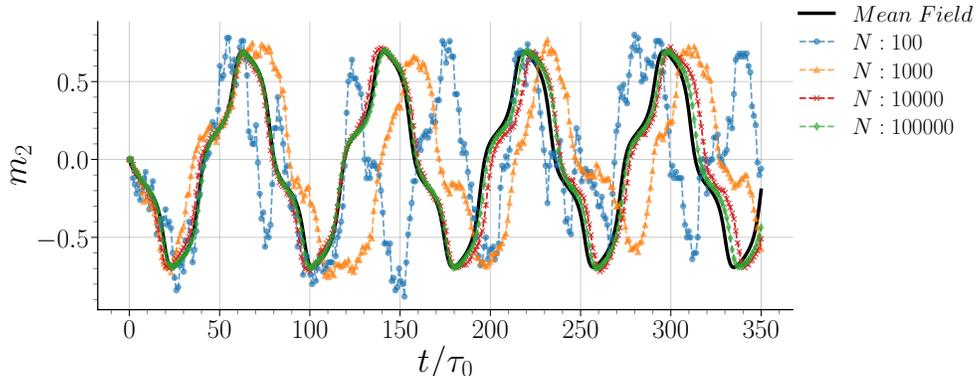


Figure 7.1: Magnetization  $m_2(t)$  in Glauber dynamics as it approaches the mean field solution. Each dashed line with markers represents a single realization for system sizes  $N = 100$  (blue circles), 1,000 (orange triangles), 10,000 (red crosses), and 100,000 (green diamonds) at  $\beta\lambda_+ = 1.3$  and  $\beta\lambda_- = 0.17$ , compared to the mean field solution (solid black line). As  $N$  increases, the simulations match the mean field predictions, with larger systems nearly overlapping with the mean field curve.

While individual realizations of Glauber dynamics for very large  $N$  align well with the mean-field solution, smaller systems exhibit significant variability. A comparison between ensemble-averaged simulations and the mean-field solution reveals damping as a net result of averaging over realizations. In Fig. 7.2, the averaged  $m_2(t)$  displays oscillation damping, even in relatively large systems. This damping effect due to the ensemble average remains pronounced even in a relatively larger system with  $N = 1,000$ . As expected, larger systems recover the mean field oscillation amplitude and maintain persistent oscillations over an extended range.

Fig. 7.3 contrasts the ensemble-averaged magnetization  $m_2(t)$  from Glauber dynamics simulations with the exact Liouvillian solution of Fig. 6.2. As the sampling increases, the stochastic ensemble mean converges towards the Liouvillian dynamics, demonstrating the equivalence between the statistical expectations of stochastic processes and the deterministic predictions derived from the master equation. For larger sampling (purple and green trajectories), the averaging of individual dynamics leads to destructive interference and damped oscillations. A single realization (blue trajectory) still preserves the characteristic oscillation

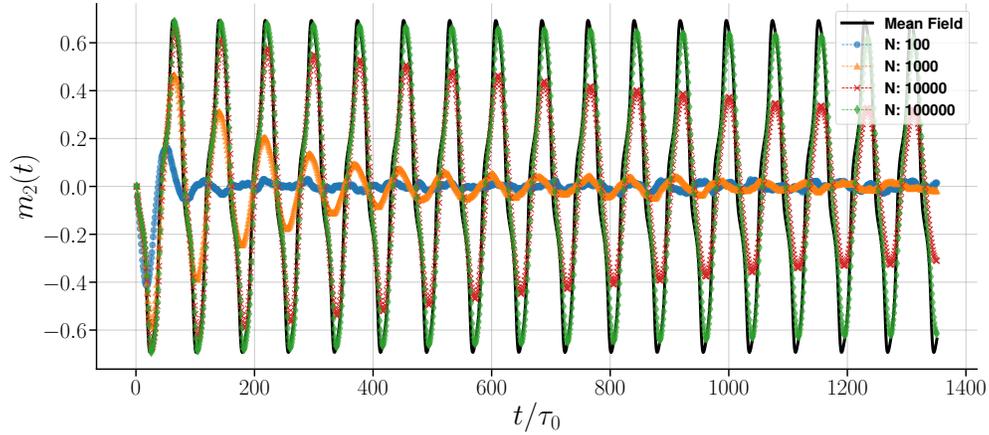


Figure 7.2: Average magnetization,  $m_2(t)$  in Glauber dynamics over 1,000 realizations. The  $N = 100$  system shows a significant decay within approximately two periods. At  $N = 1,000$ , the oscillation initially matches the mean field period but begins to shorten around  $t/\tau_0 = 400$  while damping out. A larger system exhibits prolonged oscillation persistence, yet still with a noticeable damping. The largest  $N$  (the green curve) approximates an infinite system and more closely recovers the oscillations of the mean field solution.

within the limit cycle regime, albeit with inconsistent periods. This can be attributed to the high susceptibility to noise in smaller systems. Our simulation was limited to  $N = 200$ , a relatively small configuration, due to the computational expense associated with the Liouvillian matrix calculation, as discussed above.

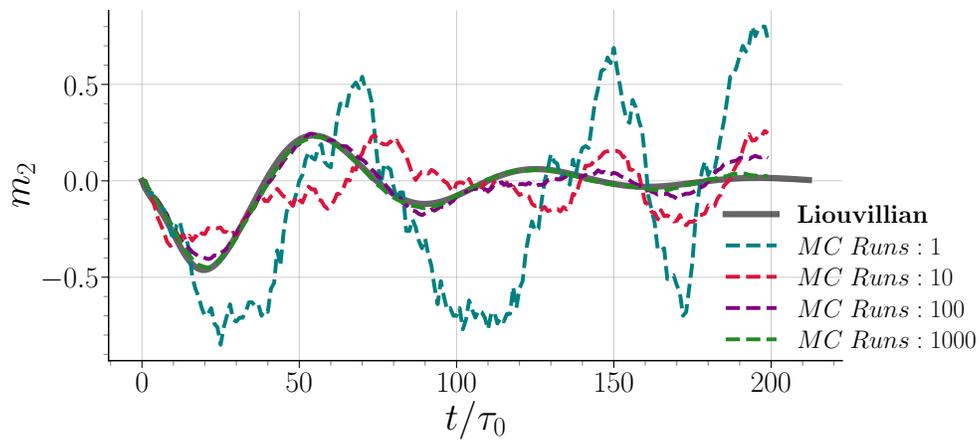


Figure 7.3: Equivalence of the master equation solution and averages of Glauber dynamics. System parameters and initial conditions are same of those in Fig. 7.1 and 7.2:  $\beta\lambda_+ = 1.3$ ,  $\beta\lambda_- = 0.17$ ,  $m_1(0) = 1$ , and  $m_2(0) = 0$ . The system has 100 spins in each subnet  $S$  and  $D$ , totaling  $N = 200$ . Colored dashed curves represent the averages from Monte Carlo simulation runs. As the sampling size increases, the average trajectory of all stochastic paths converges to the Liouvillian solution.

### 7.1.1 Numerical tests of critical behavior

In this last section, we test the predictions obtained using Langevin's equations in Sect. 5.1 with Glauber numerical simulations. We focus first on predictions related to the fold line. The first observation from Eq. 5.21 is that the oscillations of the autocorrelation function for a system exactly on the fold line are purely driven by fluctuations and are characterized by a period that scales as  $N^{1/3}$ . We show this behavior in Fig. 7.4, where after rescaling the delay time  $\tau$  by  $N^{1/3}$ , the autocorrelation functions calculated numerically with  $N$  ranging from  $N = 1000$  to  $N = 50000$  collapse to a single universal function. The autocorrelation is calculated starting at  $t = 100\tau_0$  to remove transients related to the choice of the initial conditions. The expected scaling behavior is observed for the real and imaginary components of the autocorrelation of  $z = m_1 - im_2$ .

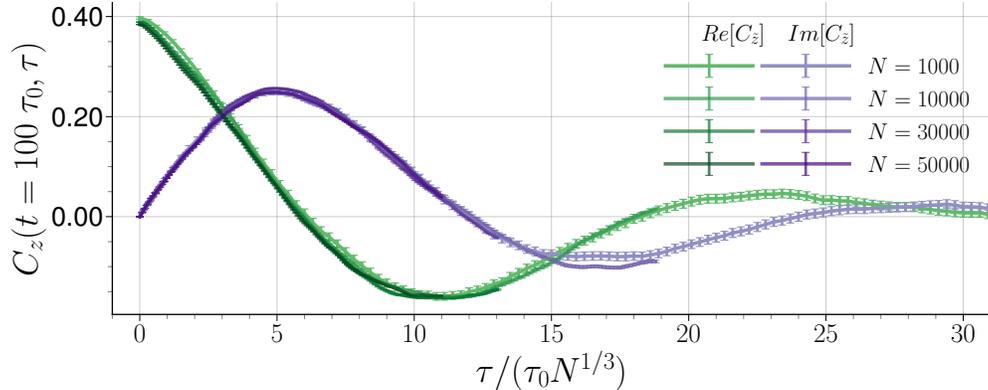


Figure 7.4: Autocorrelation on the fold line at  $\beta\lambda_+ = 1.25$  and  $\beta\lambda_- = 0.1025$ . For different values of  $N$ . The time axis is scaled according to Eq. 5.21 to show collapsing into a single function. The absolute value of the autocorrelation decays purely exponentially, while its real and imaginary components show underdamped oscillations.

The second prediction relates to the response of a system on the fold line to an external drive and its dependence on the strength of the drive,  $F$ . According to Eq. 5.32, we expect that in the limit of large  $N$  where the noise-induced switching is suppressed, the characteristic

time for switching scales as  $F^{-1/2}$ . We tested this behavior in Fig. 7.5, where we show the rotation of the angle  $\theta = \arctan m_2/m_1$  right after the activation of a constant field  $F$  in a system initially at  $m_2 = 1$ . The constant  $F$  pushes the state towards  $m_1$ , and the amplitude of rotation and its time dependence scale as predicted by Eq. 5.32 in the limit of small  $\theta$ .

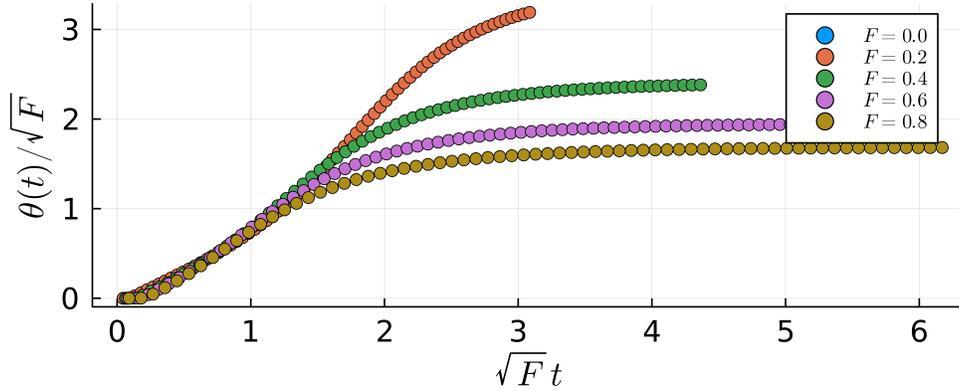


Figure 7.5:  $\theta$  rotations following the activation of a constant  $F$  for a system of  $N = 10^6$  on the fold line with  $\beta\lambda_+ = 1.25$  and  $\beta\lambda_- = 0.1025$ . Time and angles are rescaled according to Eq. 5.32, which is valid for  $\theta \ll 1$ .

We also find  $N$ -dependent damped oscillations for the autocorrelation function on the Hopf bifurcation line. This is consistent with the scaling relation obtained in Eq. 5.13 using a rotating wave  $\tilde{z} = e^{-i\beta\lambda_-t}z$ . Fig. 7.6 shows how, by rescaling the autocorrelation in amplitude and time, simulation runs for networks of different sizes  $N$  collapse into a universal function. We have verified that this behavior holds for different values of  $\beta\lambda_-$  along the Hopf line.

Finally, we numerically studied the behavior of the autocorrelation functions slightly outside the critical lines, identifying two distinct behaviors. Near the Hopf line and above the fold line, the damping of the autocorrelation is associated with a characteristic time  $T$  that scales linearly with  $N$ , following the analytical predictions of Eqs. 5.15 and 5.26. In contrast, below the fold line, in the regime where memory retrieval is effective, the characteristic time

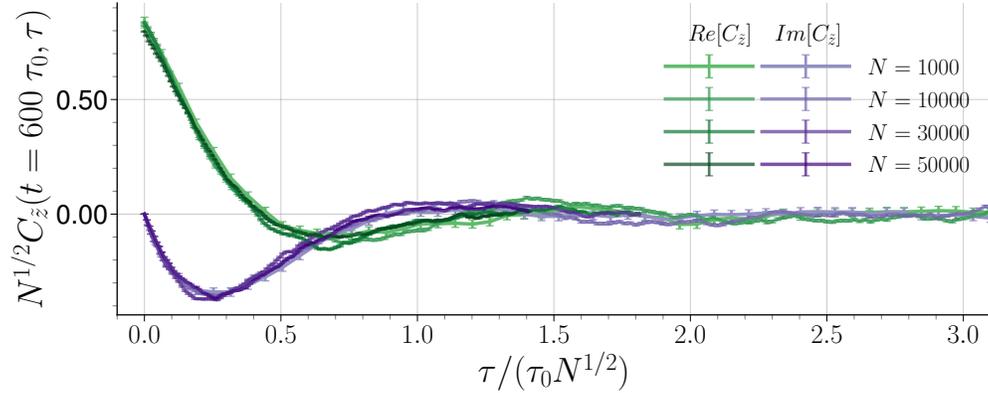


Figure 7.6: Hopf critical exponent:  $\beta\lambda_+ = 1.0$ ,  $\beta\lambda_- = 1.7$ . The time axis and autocorrelation are scaled according to Eq. 5.13 to show collapsing into a single function.

increases exponentially with  $N$ , as described by Eq. 5.28. Fig. 7.7 presents the results of numerical simulations where the decay of the autocorrelation function was fitted to an exponential model with a characteristic time  $T$ .

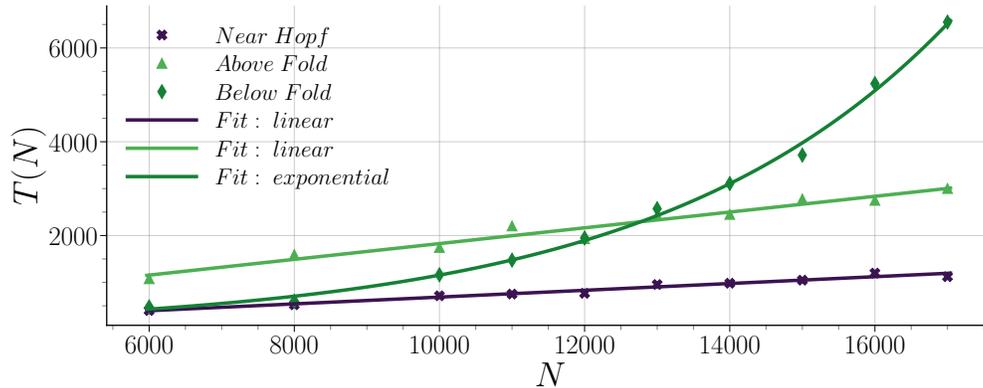


Figure 7.7: Characteristic decay times of autocorrelation near the Hopf line and above the Fold line grow linearly with  $N$ , following predictions of Eqs. 5.15 and 5.26. Below the Fold line, the decay time grows exponentially with  $N$ , following Eq. 5.28. Errors in the data are too small to be visible and have been omitted for clarity. The fits for all three regimes indicate strong statistical agreement between the data and the respective fitted models, with  $R^2$  values of 0.969 (near Hopf), 0.980 (above Fold), and 0.993 (below Fold).

## 7.2 Conclusions

Several biological processes evolve through multi-step sequential transitions. Hematopoiesis, for instance, is a multi-step cascade that starts with stem cells and progresses through oligopotent and lineage-committed progenitors. Similarly, central pattern generators are neural circuits producing rhythmic or periodic functions such as breathing or walking. Another example is the cell cycle, which consists of a finely-tuned sequence of cellular phases. From a theoretical perspective, developing and understanding effective models that can address questions related to these biological sequential transitions is important. For instance, are these transitions controlled by intrinsic or extrinsic factors? What is the role of stochasticity, and how does it scale with the number of involved components? Are there critical regions that separate phases of different behaviors and exhibit some scale invariance properties? Are there critical regions of the phase space with enhanced sensitivity to external perturbations?

The two-memory non-reciprocal Hopfield model studied here addresses many of the above questions. Switching is encoded through non-reciprocal interactions that modify Hebbian coupling. In this  $N$ -body system, we explore the effects of the number of components,  $N$ , and noise. We found that two distinct regions of critical behavior emerge at the interface of different dynamical phases. We identified and studied these regions, which correspond to Hopf bifurcations and fold bifurcations. Previous studies have explored the hypothesis that some biological systems operate at Hopf bifurcation criticality. However, behaviors near the fold line could explain other biological phenomena involving state switching. The dynamic scaling behavior, marked by different critical exponents  $\zeta$  in the autocorrelation function, suggests these two regimes are qualitatively distinct. Furthermore, we showed that sensitivity to external signals varies significantly. Specifically, in the Hopf bifurcation line, the system is

sensitive to perturbations resonant with the limit cycle frequency. In contrast, perturbations to a system in the fold line do not induce sustained limit cycles but enable controlled state switching. The time required to respond to perturbations also differs, scaling faster in the fold line than in the Hopf line.

The model studied here can be generalized to more than two patterns. For a system encoding  $p$  patterns, the  $N$  spins partition into  $2^p - 1$  subnetworks, analogous to the division into similarity ( $S$ ) and differential ( $D$ ) spins introduced earlier. For instance, in Ref. [141], four patterns were used to extend the  $Z_2 \times Z_2$  ( $C_4$ ) model considered here to the  $C_8$  symmetry case. A modification of the interaction using a Moore-Penrose pseudoinverse matrix of spins and patterns [165] was also used in that paper to reduce errors due to correlation among the memory patterns. However, for larger  $p$ , the model's ability to recover sequences of patterns quickly diminishes [132]. One way to address this limitation involves introducing a delay in the switching term [21], which could be realized through a modulation of the interaction, as recently explored by Herron et al. [168]. Hopfield networks do not need to be complete networks for memory retrieval. For instance, in random asymmetric networks, memory retrieval is preserved when the average network connectivity is above a critical value [169]. This property can be exploited to integrate the models with additional biological information. For example, in Ref. [152], the wiring of gene regulatory networks was combined with the memory retrieval property of the Hopfield model to identify bottleneck genes more susceptible to cell state switching. Another exciting extension involves defining branching points for memory patterns. Instead of cycles or fixed points, one can represent dynamics in which a memory pattern  $\xi_1$  can transition into  $\xi_2$  or  $\xi_3$  patterns. This can be implemented by adding a random switch in the Glauber dynamics that randomly chooses between  $\xi_2$  and  $\xi_3$  in the dynamics. This approach was implemented in Ref. [140] to model the random switching

between clonal states in disease progression.

While the present study has been motivated by biological questions, Hopfield networks with dilute memory patterns (i.e.,  $p < \log_2(N)$ ) have been explored in the presence of a transverse field on the  $x$ -axis, which renders the system quantum mechanical [170]. Although non-reciprocity in physical systems is less common than in biological settings, integrated photonics systems can be engineered to exhibit real space asymmetric coupling [171]. Non-reciprocity resulting from quantum mechanical effects in coupled parametric oscillators has also been recently demonstrated [171]. Studying these physical systems in critical regions near oscillatory instability could help understand the effects of noise and driving in truly out-of-equilibrium systems. Hopfield networks and their modern improvements [172] have also received renewed attention due to their connection to machine learning and artificial intelligence. For instance, new message-passing algorithms for Restricted Boltzmann Machines (RBM) have been proposed based on the mapping of Hopfield networks to RBM by a Hubbard-Stratonovich Gaussian transformation [173, 174]. Moreover, Hopfield networks have been suggested as a better alternative to the attention mechanism used in transformers [175]. Since the attention mechanism is the key innovation of the transformer architecture [176], a fundamental understanding of the properties of symmetric and asymmetric Hopfield neural networks could suggest more powerful architectures for AI applications.

### 7.2.1 Code Availability

The code developed for this manuscript is available at <https://github.com/shuyue13/non-reciprocal-Hopfield>

# Chapter 8. Conclusions

Many complex adaptive systems, from immune networks to neural circuits, rely on modular architectures and non-reciprocal interactions to generate context-specific responses to external stimuli. This dissertation has addressed the fundamental question: *How do modular structure and asymmetric couplings jointly shape the dynamics of such networks, and what theoretical framework can elucidate the underlying mechanisms?*

To answer this, we pursued two complementary research tracks—(I) a data-driven *Differential Network* (DN) analysis of high-throughput gene-expression time series, and (II) a theoretical study of non-reciprocal Hopfield networks. Together, these tracks illuminate both the observable modular reorganization and the underlying dynamical mechanisms that could drive such behavior.

Having laid the empirical foundations in Chapters 2–3 (Track I) and the theoretical foundations in Chapters 4–7 (Track II), we now synthesize their key findings and contributions and outline future directions that could further deepen our understanding of modular networks and their adaptive responses.

## 8.1 Track I: key findings and contributions

Track I employed DN analysis to investigate how biological networks reorganize after external perturbations. The saliva dataset tracked innate immune responses to PPSV23 vaccination, whereas the B-cell dataset monitored responses to Rituximab treatment. Louvain community detection, combined with temporal ordering and pathway enrichment, yielded the following insights.

- **Modular temporal organization.** In the saliva network we retained 15 communities of size  $\geq 4$ . Ranking each module by the peak time of its  $\Delta$ -expression heat-map revealed a clear cascade  $C6 \rightarrow C9 \rightarrow C8 \rightarrow C2 \rightarrow C0, C11 \rightarrow \dots \rightarrow C1$ , suggesting that early-responding modules may trigger later ones. In contrast, the B-cell network (14 communities retained after component and size filtering) exhibited *three* broad temporal pattern classes rather than a single cascade, hinting at heterogeneous sensitivity to Rituximab.
- **Functional enrichment confirmation.** Reactome and KEGG analyses showed that each community possesses a coherent biological signature (e.g. innate-immune signalling, metabolic regulation, or stress responses) highlighting how modular organization enables specialized yet coordinated functionality.
- **Structure–function correlations.** Hub-gene analysis pinpointed clinically relevant drivers. In saliva, **IL4R** peaks in community C3 exactly when the subject recorded fever, spotting gene-level changes to physiology. In B cells, the oncogene **PELI1** serves as a hub in community C5; as an E3 ubiquitin ligase that amplifies TRAF-mediated NF- $\kappa$ B activation, its prominence suggests that Rituximab may also modulate this pro-survival axis.
- **Conceptual insights into inter-modular dynamics.** We propose that network modules engage in a sophisticated choreography of cross-talk rather than acting in isolation. This conceptual framework, validated by observing staggered activation patterns and module-specific pathway enrichments, offers a novel lens for understanding how biological systems mount targeted responses while preserving essential functions.
- **Methodological contribution.** We released a fully reproducible DN-analysis pipeline

(authored by Shuyue Xue): <https://github.com/gmiaslab/DifferentialNetworks> which integrates alignment-free quantification, community detection, heat-map ordering, and enrichment statistics — advancing the methodological toolkit available for future multi-condition time-series studies.

## 8.2 Track II: key findings and contributions

While DN analysis documents *what* changes, it cannot by itself explain *how interactions among subnetworks drive the global network’s dynamic responses or dictate their stability and transitions*. Motivated by the empirical observations of Track I, Track II employed techniques from statistical mechanics and dynamical-systems theory to build a **non-reciprocal Hopfield network** — an infinite-range Ising model with asymmetric couplings that naturally supports multiple attractors and limit-cycle dynamics. This theoretical paradigm allowed us to probe how internal differentiation (similarity vs. differential subnetworks), stochastic noise, external drive, and coupling asymmetry shape collective behavior.

In this theoretical pursuit, the dissertation makes the following contributions to Hopfield theory and, more broadly, to systems with non-reciprocal interactions. Both are significant area within theoretical condensed matter physics and statistical mechanics.

- **Phase topology and bifurcations.** Mean-field fixed-point analysis unveils a phase diagram with three dynamical regimes — paramagnetic, memory retrieval, and limit-cycle — separated by a Hopf line ( $\beta\lambda_+ = 1$ ) and a fold line given by  $\det A = 0$  where  $A$  is the Jacobian of the mean-field map at the fixed point; both symmetric ( $\lambda_+$ ) and asymmetric ( $\lambda_-$ ) couplings control the transitions.
- **Critical behaviour (beyond mean-field).** A stochastic centre-manifold treatment

of the critical lines yields dynamical exponents  $\zeta = \frac{1}{2}$  (Hopf) and  $\zeta = \frac{1}{3}$  (fold), with response-time scalings  $|F|^{-2/3}$  and  $|F|^{-1/2}$ , respectively, and predicts a limit-cycle coherence time  $T \sim N$  within the oscillatory phase.

- **Exact finite- $N$  kinetics.** A Master-Equation/Liouvillian diagonalisation exact system-size-dependent relaxation spectra and two-time correlation functions, bridging microscopic spin-flip rules to macroscopic observables.
- **Algorithmic acceleration & validation.** We devised a subnetwork-partitioned Glauber Monte-Carlo algorithm that alternately updates spins in the similarity subnetwork and in the differential subnetwork, which exploits the resulting block structure in the coupling matrix. This reduces per-step complexity and enabled simulations up to  $N = 10^5$  and even larger in scale. The resulting trajectories confirm all analytical scaling laws and capture the damping of limit-cycle coherence at finite  $N$ .

### 8.3 Future Directions

The central question that threads this dissertation together is how to fuse data-driven network structure with the parameters and dynamics of theoretically-grounded models like the non-reciprocal Hopfield network investigated here. A natural future direction is therefore to build a more explicit unifying framework that maps empirical differential-network communities onto model architecture. Or conversely, using the phase diagram to generate testable predictions about which parameter should be tuned.

Equally important is broadening the theoretical model space toward greater biological realism. Incorporating nested or overlapping modules, hierarchical organization, or adaptive learning rules would allow the model to capture developmental plasticity and long-term

re-wiring. Such extensions promise not only richer dynamical phenomena but also a tighter feedback loop with experiments, where hypotheses about specific interaction mechanisms can be validated or refined. Taken together, these avenues move us toward an integrative framework in which empirical network analysis and non-equilibrium theory mutually inform and constrain one another.

# BIBLIOGRAPHY

## BIBLIOGRAPHY

- [1] J. Reimand, R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Rostamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, *et al.*, “Pathway enrichment analysis and visualization of omics data using g: Profiler, gsea, cytoscape and enrichmentmap,” *Nature protocols*, vol. 14, no. 2, pp. 482–517, 2019.
- [2] D. Defays, “An efficient algorithm for a complete link method,” *The Computer Journal*, vol. 20, no. 4, pp. 364–366, 1977.
- [3] J. A. Hartigan, “Statistical theory in clustering,” *Journal of classification*, vol. 2, no. 1, pp. 63–76, 1985.
- [4] W. R. Ashby, *An Introduction to Cybernetics*. London: Chapman & Hall, 1956.
- [5] K. O. Ellefsen, J.-B. Mouret, and J. Clune, “Neural modularity helps organisms evolve to learn new skills without forgetting old skills,” *PLoS computational biology*, vol. 11, no. 4, p. e1004128, 2015.
- [6] G. P. Wagner, M. Pavlicev, and J. M. Cheverud, “The road to modularity,” *Nature Reviews Genetics*, vol. 8, no. 12, pp. 921–931, 2007.
- [7] H. A. Simon, “The architecture of complexity,” *Proceedings of the American Philosophical Society*, vol. 106, no. 6, pp. 467–482, 1962.
- [8] G. P. Wagner and L. Altenberg, “Complex adaptations and the evolution of evolvability,” *Evolution*, vol. 50, no. 3, pp. 967–976, 1996.
- [9] D. M. Lorenz, A. Jeng, and M. W. Deem, “The emergence of modularity in biological systems,” *Physics of life reviews*, vol. 8, no. 2, pp. 129–160, 2011.
- [10] J. Clune, J.-B. Mouret, and H. Lipson, “The evolutionary origins of modularity,” *Proceedings of the Royal Society b: Biological sciences*, vol. 280, no. 1755, p. 20122863, 2013.
- [11] D. Yu, M. Kim, G. Xiao, and T. H. Hwang, “Review of biological network data and its applications,” *Genomics & Informatics*, vol. 11, no. 4, pp. 200–210, 2013.
- [12] T. Ideker and N. J. Krogan, “Differential network biology,” *Molecular systems biology*, vol. 8, no. 1, p. 565, 2012.
- [13] S. Xue, L. R. Rogers, M. Zheng, J. He, C. Piermarocchi, and G. I. Mias, “Applying differential network analysis to longitudinal gene expression in response to perturbations,” *Frontiers in Genetics*, vol. 13, p. 1026487, 2022.

- [14] R. Albert, “Network inference, analysis, and modeling in systems biology,” *Annual Review of Biomedical Engineering*, vol. 9, pp. 37–53, 2007.
- [15] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities.,” *Proceedings of the national academy of sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.
- [16] M. Smart, *Collective Dynamics of Interacting Cell Types*. PhD thesis, University of Toronto (Canada), 2022.
- [17] A. Szedlak, *Structures and Boolean Dynamics in Gene Regulatory Networks*. 2017.
- [18] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, “Hierarchical group dynamics in pigeon flocks,” *Nature*, vol. 464, no. 7290, pp. 890–893, 2010.
- [19] M. Fruchart, R. Hanai, P. B. Littlewood, and V. Vitelli, “Non-reciprocal phase transitions,” *Nature*, vol. 592, no. 7854, pp. 363–369, 2021.
- [20] E. Montbrió and D. Pazó, “Kuramoto model for excitation-inhibition-based oscillations,” *Physical review letters*, vol. 120, no. 24, p. 244101, 2018.
- [21] H. Sompolinsky and I. Kanter, “Temporal association in asymmetric neural networks,” *Physical review letters*, vol. 57, no. 22, p. 2861, 1986.
- [22] H. Sompolinsky, A. Crisanti, and H. J. Sommers, “Chaos in random neural networks,” *Physical Review Letters*, vol. 61, no. 3, p. 259, 1988.
- [23] J. H. Han and S. Y. Park, “Nonreciprocal frustration: Time crystalline order-by-disorder phenomenon and a spin-glass-like state,” *Physical Review X*, vol. 14, no. 1, p. 011029, 2024.
- [24] E. Blumenthal, J. W. Rocks, and P. Mehta, “Phase transition to chaos in complex ecosystems with nonreciprocal species-resource interactions,” *Physical review letters*, vol. 132, no. 12, p. 127401, 2024.
- [25] G. Bianconi, “Interdisciplinary and physics challenges of network theory,” *EPL (Europhysics Letters)*, vol. 111, no. 5, p. 56001, 2015.
- [26] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.
- [27] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.
- [28] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, 2003.

- [29] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports*, vol. 424, no. 4–5, pp. 175–308, 2006.
- [30] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*. Cambridge University Press, 2007.
- [31] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, 2010.
- [32] S. Fortunato, “Community detection in graphs,” *Physics reports*, vol. 486, no. 3-5, pp. 75–174, 2010.
- [33] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [34] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [35] M. E. Newman, “Assortative mixing in networks,” *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.
- [36] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, “Large-scale topological and dynamical properties of the internet,” *Physical Review E*, vol. 65, no. 6, p. 066130, 2002.
- [37] S. N. Dorogovtsev, A. V. Goltsev, and J. F. Mendes, “Critical phenomena in complex networks,” *Reviews of Modern Physics*, vol. 80, no. 4, pp. 1275–1335, 2008.
- [38] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.
- [39] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, “Controllability of complex networks,” *Nature*, vol. 473, pp. 167–173, 2011.
- [40] G. Menichetti, L. Dall’Asta, and G. Bianconi, “Control of multilayer networks,” *Physical Review Letters*, vol. 113, no. 7, p. 078701, 2014.
- [41] J. Park and M. E. Newman, “Statistical mechanics of networks,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 70, no. 6, p. 066117, 2004.
- [42] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, “The statistical physics of real-world networks,” *Nature Reviews Physics*, vol. 1, no. 1, pp. 58–71, 2019.
- [43] M. Newman, *Networks*. OUP Oxford, 2018.

- [44] U. Alon, *Network motifs: theory and experimental approaches*, vol. 8. Nature Publishing Group, 2007.
- [45] M. E. Newman and M. Girvan, “Finding and evaluating community structure in networks,” *Physical review E*, vol. 69, no. 2, p. 026113, 2004.
- [46] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, “From molecular to modular cell biology,” *Nature*, vol. 402, no. 6761, pp. C47–C52, 1999.
- [47] T. B. Martin, *Theoretical tools for network analysis: Game theory, graph centrality, and statistical inference*. Phd thesis, University of Michigan, 2016.
- [48] S. Bandyopadhyay, M. Mehta, D. Kuo, M.-K. Sung, R. Chuang, E. J. Jaehnig, B. Bodenmiller, K. Licon, W. Copeland, M. Shales, *et al.*, “Rewiring of genetic networks in response to dna damage,” *Science*, vol. 330, no. 6009, pp. 1385–1389, 2010.
- [49] M. J. Ha, V. Baladandayuthapani, and K.-A. Do, “Dingo: differential network analysis in genomics,” *Bioinformatics*, vol. 31, no. 21, pp. 3413–3420, 2015.
- [50] T.-H. Hsiao, Y.-C. Chiu, P.-Y. Hsu, T.-P. Lu, L.-C. Lai, M.-H. Tsai, T. H.-M. Huang, E. Y. Chuang, and Y. Chen, “Differential network analysis reveals the genome-wide landscape of estrogen receptor modulation in hormonal cancers,” *Scientific reports*, vol. 6, no. 1, pp. 1–16, 2016.
- [51] K. Mitra, A.-R. Carvunis, S. K. Ramesh, and T. Ideker, “Integrative approaches for finding modular structure in biological networks,” *Nature Reviews Genetics*, vol. 14, no. 10, pp. 719–732, 2013.
- [52] A. de la Fuente, “From ‘differential expression’ to ‘differential networking’—identification of dysfunctional regulatory networks in diseases,” *Trends in genetics*, vol. 26, no. 7, pp. 326–333, 2010.
- [53] L. Cabusora, E. Sutton, A. Fulmer, and C. V. Forst, “Differential network expression during drug and stress response,” *Bioinformatics*, vol. 21, no. 12, pp. 2898–2905, 2005.
- [54] E. T. Jaynes, “Information theory and statistical mechanics,” *Physical review*, vol. 106, no. 4, p. 620, 1957.
- [55] A.-L. Barabási, N. Gulbahce, and J. Loscalzo, “Network medicine: a network-based approach to human disease,” *Nature reviews genetics*, vol. 12, no. 1, pp. 56–68, 2011.
- [56] A. A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. D. Faveria, and A. Califano, “Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context,” in *BMC bioinformatics*, vol. 7, pp. 1–15, Springer, 2006.

- [57] B. Zhang and S. Horvath, “A general framework for weighted gene co-expression network analysis,” *Statistical applications in genetics and molecular biology*, vol. 4, no. 1, 2005.
- [58] M. Girvan and M. E. Newman, “Community structure in social and biological networks,” *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [59] M. E. J. Newman, “Modularity and community structure in networks,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [60] M. A. Porter, J.-P. Onnela, P. J. Mucha, *et al.*, “Communities in networks,” *Notices of the AMS*, vol. 56, no. 9, pp. 1082–1097, 2009. Available at: <https://www.ams.org/journals/notices/200909/rtx090901082p.pdf> (Accessed Aug 19, 2022).
- [61] A. W. Rives and T. Galitski, “Modular organization of cellular networks,” *Proceedings of the national Academy of sciences*, vol. 100, no. 3, pp. 1128–1133, 2003.
- [62] N. Gulbahce and S. Lehmann, “The art of community detection,” *BioEssays*, vol. 30, no. 10, pp. 934–938, 2008.
- [63] J.-D. J. Han, “Understanding biological functions through molecular networks,” *Cell research*, vol. 18, no. 2, pp. 224–237, 2008.
- [64] G. Calderer and M. L. Kuijjer, “Community detection in large-scale bipartite biological networks,” *Frontiers in Genetics*, vol. 12, p. 520, 2021.
- [65] A. Clauset, M. E. Newman, and C. Moore, “Finding community structure in very large networks,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 70, no. 6, p. 066111, 2004.
- [66] M. E. Newman, “Modularity and community structure in networks,” *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [67] B. Karrer and M. E. Newman, “Stochastic blockmodels and community structure in networks,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 83, no. 1, p. 016107, 2011.
- [68] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, “Inference and phase transitions in the detection of modules in sparse networks,” *Physical Review Letters*, vol. 107, no. 6, p. 065701, 2011.
- [69] J. Duch and A. Arenas, “Community detection in complex networks using extremal optimization,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 72, no. 2, p. 027104, 2005.

- [70] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.
- [71] H. Kitano, “Biological robustness,” *Nature Reviews Genetics*, vol. 5, no. 11, pp. 826–837, 2004.
- [72] V. A. Traag, L. Waltman, and N. J. van Eck, “From louvain to leiden: guaranteeing well-connected communities,” *Scientific Reports*, vol. 9, no. 1, p. 5233, 2019.
- [73] S. Fortunato and M. Barthelemy, “Resolution limit in community detection,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41, 2007.
- [74] B. H. Good, Y.-A. de Montjoye, and A. Clauset, “Performance of modularity maximization in practical contexts,” *Physical Review E*, vol. 81, no. 4, p. 046106, 2010.
- [75] J. C. Edwards, L. Szczepański, J. Szechiński, A. Filipowicz-Sosnowska, P. Emery, D. R. Close, R. M. Stevens, and T. Shaw, “Efficacy of b-cell-targeted therapy with rituximab in patients with rheumatoid arthritis,” *New England Journal of Medicine*, vol. 350, no. 25, pp. 2572–2581, 2004.
- [76] X. Bosch, M. Ramos-Casals, and M. A. Khamashta, *Drugs targeting B-cells in autoimmune diseases*. Springer, 2014.
- [77] World Health Organization, “World Health Organization model list of essential medicines: 21st list 2021,” technical documents, World Health Organization, 2021. Available at: <https://www.who.int/publications/i/item/WHO-MHP-HPS-EML-2021.02> (Accessed Aug 19, 2022).
- [78] T. Shaw, J. Quan, and M. Totoritis, “B cell therapy for rheumatoid arthritis: the rituximab (anti-cd20) experience,” *Annals of the rheumatic diseases*, vol. 62, no. suppl 2, pp. ii55–ii59, 2003.
- [79] P. Johnson and M. Glennie, “The mechanisms of action of rituximab in the elimination of tumor cells,” in *Seminars in oncology*, vol. 30, pp. 3–8, Elsevier, 2003.
- [80] R. A. Clynes, T. L. Towers, L. G. Presta, and J. V. Ravetch, “Inhibitory fc receptors modulate in vivo cytotoxicity against tumor targets,” *Nature medicine*, vol. 6, no. 4, pp. 443–446, 2000.
- [81] A. R. Jazirehi, X.-H. Gan, S. De Vos, C. Emmanouilides, and B. Bonavida, “Rituximab (anti-cd20) selectively modifies bcl-xl and apoptosis protease activating factor-1 (apaf-1) expression and sensitizes human non-hodgkin’s lymphoma b cell lines to paclitaxel-induced apoptosis,” *Molecular cancer therapeutics*, vol. 2, no. 11,

- pp. 1183–1193, 2003. [Eprint]. Available at: <https://aacrjournals.org/mct/article-pdf/2/11/1183/1865287/1183-1193.pdf> (Accessed Aug 19, 2022).
- [82] A. R. Jazirehi, S. Huerta-Yepez, G. Cheng, and B. Bonavida, “Rituximab (chimeric anti-cd20 monoclonal antibody) inhibits the constitutive nuclear factor- $\kappa$ b signaling pathway in non-hodgkin’s lymphoma b-cell lines: role in sensitization to chemotherapeutic drug-induced apoptosis,” *Cancer research*, vol. 65, no. 1, pp. 264–276, 2005.
- [83] M. I. Vega, A. R. Jazirehi, S. Huerta-Yepez, and B. Bonavida, “Rituximab-induced inhibition of yy1 and bcl-xl expression in ramos non-hodgkin’s lymphoma cell line via inhibition of nf- $\kappa$ b activity: role of yy1 and bcl-xl in fas resistance and chemoresistance, respectively,” *The Journal of Immunology*, vol. 175, no. 4, pp. 2174–2183, 2005.
- [84] N. L. Bray, H. Pimentel, P. Melsted, and L. Pachter, “Near-optimal probabilistic RNA-seq quantification,” *Nature biotechnology*, vol. 34, no. 5, pp. 525–527, 2016.
- [85] J. Harrow, A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans, F. Kokocinski, B. L. Aken, D. Barrell, A. Zadissa, S. Searle, *et al.*, “Gencode: the reference human genome annotation for the encode project,” *Genome research*, vol. 22, no. 9, pp. 1760–1774, 2012.
- [86] H. Pimentel, N. L. Bray, S. Puente, P. Melsted, and L. Pachter, “Differential analysis of rna-seq incorporating quantification uncertainty,” *Nat Methods*, vol. 14, no. 7, pp. 687–690, 2017.
- [87] S. Anders and W. Huber, “Differential expression analysis for sequence count data,” *Genome Biol*, vol. 11, no. 10, p. R106, 2010.
- [88] G. I. Mias, V. V. Singh, L. R. Rogers, S. Xue, M. Zheng, S. Domanskyi, M. Kanada, C. Piermarocchi, and J. He, “Longitudinal saliva omics responses to immune perturbation: a case study,” *Scientific Reports*, vol. 11, no. 1, pp. 1–20, 2021.
- [89] A. Lancichinetti and S. Fortunato, “Community detection algorithms: a comparative analysis,” *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 80, no. 5 Pt 2, p. 056117, 2009.
- [90] Z. Yang, R. Algesheimer, and C. J. Tessone, “A comparative analysis of community detection algorithms on artificial networks,” *Sci Rep*, vol. 6, p. 30750, 2016.
- [91] D. Croft, G. O’kelly, G. Wu, R. Haw, M. Gillespie, L. Matthews, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, *et al.*, “Reactome: a database of reactions, pathways and biological processes,” *Nucleic acids research*, vol. 39, no. suppl\_1, pp. D691–D697, 2010.
- [92] The Pandas development team, “pandas-dev/pandas: Pandas,” feb 2020.

- [93] Wes McKinney, “Data Structures for Statistical Computing in Python,” in *Proceedings of the 9th Python in Science Conference* (Stéfan van der Walt and Jarrod Millman, eds.), pp. 56 – 61, 2010.
- [94] A. A. Hagberg, D. A. Schult, and P. J. Swart, “Exploring network structure, dynamics, and function using networkx,” in *Proceedings of the 7th Python in Science Conference* (G. Varoquaux, T. Vaught, and J. Millman, eds.), (Pasadena, CA USA), pp. 11 – 15, 2008. Available at: <https://www.osti.gov/biblio/960616> (Accessed Aug 19, 2022).
- [95] T. Bonald, N. de Lara, Q. Lutz, and B. Charpentier, “Scikit-network: Graph analysis in python,” *Journal of Machine Learning Research*, vol. 21, no. 185, pp. 1–6, 2020. Available at: <http://jmlr.org/papers/v21/20-412.html> (Accessed Aug 19, 2022).
- [96] Gene Set Enrichment Analysis, “GSEA User Guide,” 2022. Available at: <https://www.gsea-msigdb.org/gsea/doc/GSEAUserGuideFrame.html> (Accessed Sep 12, 2022).
- [97] S. Domanskyi, C. Piermarocchi, and G. I. Mias, “Pyiomica: longitudinal omics analysis and trend identification,” *Bioinformatics*, vol. 36, no. 7, pp. 2306–2307, 2020.
- [98] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, T. I. Stein, R. Nudel, I. Lieder, Y. Mazor, *et al.*, “The genecards suite: from gene data mining to disease genome sequence analyses,” *Current protocols in bioinformatics*, vol. 54, no. 1, pp. 1–30, 2016.
- [99] Wolfram Research, Inc., “Mathematica, Version 12.0.0,” 2021. [Software]. Available at: <https://www.wolfram.com/mathematica> (Accessed Aug 19, 2022).
- [100] G. Joshi-Tope, M. Gillespie, I. Vastrik, P. D’Eustachio, E. Schmidt, B. de Bono, B. Jassal, G. Gopinath, G. Wu, L. Matthews, *et al.*, “Reactome: a knowledgebase of biological pathways,” *Nucleic acids research*, vol. 33, no. suppl\_1, pp. D428–D432, 2005.
- [101] L. Matthews, G. Gopinath, M. Gillespie, M. Caudy, D. Croft, B. de Bono, P. Garapati, J. Hemish, H. Hermjakob, B. Jassal, *et al.*, “Reactome knowledgebase of human biological pathways and processes,” *Nucleic acids research*, vol. 37, no. suppl\_1, pp. D619–D622, 2009.
- [102] W. T. Watford, M. Moriguchi, A. Morinobu, and J. J. O’Shea, “The biology of il-12: coordinating innate and adaptive immune responses,” *Cytokine & growth factor reviews*, vol. 14, no. 5, pp. 361–368, 2003.
- [103] S. Klemm, J. Gutermuth, L. Hültner, T. Sparwasser, H. Behrendt, C. Peschel, T. W. Mak, T. Jakob, and J. Ruland, “The bcl10–malt1 complex segregates fc $\epsilon$ ri-mediated nuclear factor  $\kappa$ b activation and cytokine production from mast cell degranulation,” *The Journal of experimental medicine*, vol. 203, no. 2, pp. 337–347, 2006.

- [104] S. I. Gringhuis, T. M. Kaptein, B. A. Wevers, B. Theelen, M. Van Der Vlist, T. Boekhout, and T. B. Geijtenbeek, “Dectin-1 is an extracellular pathogen sensor for the induction and processing of il-1 $\beta$  via a noncanonical caspase-8 inflammasome,” *Nature immunology*, vol. 13, no. 3, pp. 246–254, 2012.
- [105] S.-C. Cheng, F. L. van de Veerdonk, M. Lenardon, M. Stoffels, T. Plantinga, S. Smeekens, L. Rizzetto, L. Mukaremera, K. Preechasuth, D. Cavalieri, *et al.*, “The dectin-1/inflammasome pathway is responsible for the induction of protective t-helper 17 responses that discriminate between yeasts and hyphae of candida albicans,” *Journal of leukocyte biology*, vol. 90, no. 2, pp. 357–366, 2011.
- [106] A. K. Abbas, A. H. Lichtman, and S. Pillai, *Cellular and molecular immunology E-book*. Elsevier Health Sciences, 2014.
- [107] C. A. Bonjardim, P. C. Ferreira, and E. G. Kroon, “Interferons: signaling, antiviral and viral evasion,” *Immunology letters*, vol. 122, no. 1, pp. 1–11, 2009.
- [108] W. Liu-Mares, Z. Sun, W. R. Bamlet, E. J. Atkinson, B. L. Fridley, S. L. Slager, M. de Andrade, and E. L. Goode, “Analysis of variation in nf- $\kappa$ b genes and expression levels of nf- $\kappa$ b-regulated molecules,” in *BMC proceedings*, vol. 1, pp. 1–5, Springer, 2007.
- [109] T. U. Consortium, “Uniprot: the universal protein knowledgebase in 2021,” *Nucleic acids research*, vol. 49, no. D1, pp. D480–D489, 2021.
- [110] J. Cai, R. Li, X. Xu, L. Zhang, S. Wu, T. Yang, L. Fang, J. Wu, X. Zhu, M. Li, *et al.*, “Urgcp promotes non-small cell lung cancer invasiveness by activating the nf- $\kappa$ b-mmp-9 pathway,” *Oncotarget*, vol. 6, no. 34, p. 36489, 2015.
- [111] A. Bhattacharjee, M. Shukla, V. P. Yakubenko, A. Mulya, S. Kundu, and M. K. Cathcart, “Il-4 and il-13 employ discrete signaling pathways for target gene expression in alternatively activated monocytes/macrophages,” *Free Radical Biology and Medicine*, vol. 54, pp. 1–16, 2013.
- [112] F. Seif, M. Khoshmirsafa, H. Aazami, M. Mohsenzadegan, G. Sedighi, and M. Bahar, “The role of jak-stat signaling pathway and its regulators in the fate of t helper cells,” *Cell communication and signaling*, vol. 15, no. 1, pp. 1–13, 2017.
- [113] C.-Y. Yang, S. Ramamoorthy, S. Boller, M. Rosenbaum, A. R. Gil, G. Mittler, Y. Imai, K. Kuba, and R. Grosschedl, “Interaction of ccr4—not with ebf1 regulates gene-specific transcription and mrna stability in b lymphopoiesis,” *Genes & development*, vol. 30, no. 20, pp. 2310–2324, 2016.

- [114] N. Seyfizadeh, N. Seyfizadeh, J. Hasenkamp, and S. Huerta-Yepez, “A molecular perspective on rituximab: a monoclonal antibody for b cell non hodgkin lymphoma and other affections,” *Critical reviews in oncology/hematology*, vol. 97, pp. 275–290, 2016.
- [115] A. R. Jazirehi and B. Bonavida, “Cellular and molecular signal transduction pathways modulated by rituximab (rituxan, anti-cd20 mab) in non-hodgkin’s lymphoma: implications in chemosensitization and therapeutic intervention,” *Oncogene*, vol. 24, no. 13, pp. 2121–2143, 2005.
- [116] B. Bonavida, “Rituximab-induced inhibition of antiapoptotic cell survival pathways: implications in chemo/immunosresistance, rituximab unresponsiveness, prognostic and novel therapeutic interventions,” *Oncogene*, vol. 26, no. 25, pp. 3629–3636, 2007.
- [117] K. C. Yeung, D. W. Rose, A. S. Dhillon, D. Yaros, M. Gustafsson, D. Chatterjee, B. McFerran, J. Wyche, W. Kolch, and J. M. Sedivy, “Raf kinase inhibitor protein interacts with nf- $\kappa$ b-inducing kinase and tak1 and inhibits nf- $\kappa$ b activation,” *Molecular and cellular biology*, vol. 21, no. 21, pp. 7207–7217, 2001.
- [118] H.-Y. Park, H. Go, H. R. Song, S. Kim, G.-H. Ha, Y.-K. Jeon, J.-E. Kim, H. Lee, H. Cho, H. C. Kang, *et al.*, “Pellino 1 promotes lymphomagenesis by deregulating bcl6 polyubiquitination,” *The Journal of clinical investigation*, vol. 124, no. 11, pp. 4976–4988, 2014.
- [119] G. Steele-Perkins, W. Fang, X.-H. Yang, M. Van Gele, T. Carling, J. Gu, I. M. Buyse, J. A. Fletcher, J. Liu, R. Bronson, *et al.*, “Tumor formation and inactivation of riz1, an rb-binding member of a nuclear protein–methyltransferase superfamily,” *Genes & development*, vol. 15, no. 17, pp. 2250–2262, 2001.
- [120] J.-L. Huang, W. Liu, L.-H. Tian, T.-T. Chai, Y. Liu, F. Zhang, H.-Y. Fu, H.-R. Zhou, and J.-Z. Shen, “Upregulation of long non-coding rna malat-1 confers poor prognosis and influences cell proliferation and apoptosis in acute monocytic leukemia,” *Oncology reports*, vol. 38, no. 3, pp. 1353–1362, 2017.
- [121] X. Tong, R. Drapkin, R. Yalamanchili, G. Mosialos, and E. Kieff, “The epstein-barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with tfie,” *Molecular and cellular biology*, vol. 15, no. 9, pp. 4735–4744, 1995.
- [122] M. A. Blanco, M. Alečković, Y. Hua, T. Li, Y. Wei, Z. Xu, I. M. Cristea, and Y. Kang, “Identification of staphylococcal nuclease domain-containing 1 (snd1) as a metadherin-interacting protein with metastasis-promoting functions,” *Journal of Biological Chemistry*, vol. 286, no. 22, pp. 19982–19992, 2011.
- [123] B. Ochoa, Y. Chico, and M. J. Martínez, “Insights into snd1 oncogene promoter regulation,” *Frontiers in Oncology*, vol. 8, p. 606, 2018.

- [124] Y. Wei, S. Pattingre, S. Sinha, M. Bassik, and B. Levine, “Jnk1-mediated phosphorylation of bcl-2 regulates starvation-induced autophagy,” *Molecular cell*, vol. 30, no. 6, pp. 678–688, 2008.
- [125] M. Hiwatari, T. Taki, T. Taketani, M. Taniwaki, K. Sugita, M. Okuya, M. Eguchi, K. Ida, and Y. Hayashi, “Fusion of an af4-related gene, *laf4*, to *mll* in childhood acute lymphoblastic leukemia with t (2; 11)(q11; q23),” *Oncogene*, vol. 22, no. 18, pp. 2851–2855, 2003.
- [126] L. R. Brooks and G. I. Mias, “Streptococcus pneumoniae’s virulence and host immunity: aging, diagnostics, and prevention,” *Frontiers in immunology*, vol. 9, p. 1366, 2018.
- [127] US Food and Drug Administration, “Pneumovax 23 Prescribing Information (Accessed Sept 18th, 2021),” *The Food and Drug Administration (FDA)*, 2014.
- [128] J. C. Butler, R. F. Breiman, J. F. Campbell, H. B. Lipman, C. V. Broome, and R. R. Facklam, “Pneumococcal polysaccharide vaccine efficacy. an evaluation of current recommendations,” *JAMA*, vol. 270, no. 15, pp. 1826–31, 1993.
- [129] P. Smit, D. Oberholzer, S. Hayden-Smith, H. J. Koornhof, and M. R. Hilleman, “Protective efficacy of pneumococcal polysaccharide vaccines,” *JAMA*, vol. 238, no. 24, pp. 2613–6, 1977.
- [130] M. W. Pletz, U. Maus, N. Krug, T. Welte, and H. Lode, “Pneumococcal vaccines: mechanism of action, impact on epidemiology and adaption of the species,” *International journal of antimicrobial agents*, vol. 32, no. 3, pp. 199–206, 2008.
- [131] J. Hertz, A. Krogh, R. G. Palmer, and H. Horner, “Introduction to the theory of neural computation,” 1991.
- [132] D. Amit, *Modeling brain function: The world of attractor neural networks*. Cambridge: Cambridge University Press, 1989.
- [133] J. Hertz, G. Grinstein, and S. Solla, “Irreversible spin glasses and neural networks,” in *Heidelberg Colloquium on Glassy Dynamics: Proceedings of a Colloquium on Spin Glasses, Optimization and Neural Networks Held at the University of Heidelberg June 9–13, 1986*, pp. 538–546, Springer, 1987.
- [134] J. T. Young, A. V. Gorshkov, M. Foss-Feig, and M. F. Maghrebi, “Nonequilibrium fixed points of coupled ising models,” *Phys. Rev. X*, vol. 10, p. 011039, Feb 2020.
- [135] J. T. Young, A. V. Gorshkov, and M. Maghrebi, “Nonequilibrium universality of the nonreciprocally coupled  $O(n_1) \times O(n_2)$  model,” *arXiv:2411.12680*, 2024.

- [136] R. Hannam, A. Annibale, and R. Kühn, “Cell reprogramming modelled as transitions in a hierarchy of cell cycles,” *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 42, p. 425601, 2017.
- [137] A. H. Lang, H. Li, J. J. Collins, and P. Mehta, “Epigenetic landscapes explain partially reprogrammed cells and identify key reprogramming genes,” *PLoS computational biology*, vol. 10, no. 8, p. e1003734, 2014.
- [138] L. Cantini and M. Caselle, “Hope4Genes: a Hopfield-like class prediction algorithm for transcriptomic data,” *Scientific reports*, vol. 9, no. 1, p. 337, 2019.
- [139] A. Taherian Fard and M. A. Ragan, “Modeling the attractor landscape of disease progression: a network-based approach,” *Frontiers in genetics*, vol. 8, p. 48, 2017.
- [140] S. Domanskyi, A. Hakansson, G. Paternostro, and C. Piermarocchi, “Modeling disease progression in multiple myeloma with Hopfield networks and single-cell RNA-seq,” in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2129–2136, IEEE, 2019.
- [141] A. Szedlak, S. Sims, N. Smith, G. Paternostro, and C. Piermarocchi, “Cell cycle time series gene expression data encoded as cyclic attractors in Hopfield systems,” *PLoS computational biology*, vol. 13, no. 11, p. e1005849, 2017.
- [142] S. A. Kauffman and S. Johnsen, “Coevolution to the edge of chaos: Coupled fitness landscapes, poised states, and coevolutionary avalanches,” *Journal of theoretical biology*, vol. 149, no. 4, pp. 467–505, 1991.
- [143] G. Longo and M. Montévil, *Perspectives on organisms*. Springer, 2014.
- [144] T. Mora and W. Bialek, “Are biological systems poised at criticality?,” *Journal of Statistical Physics*, vol. 144, pp. 268–302, 2011.
- [145] Y. Avni, M. Fruchart, D. Martin, D. Seara, and V. Vitelli, “The non-reciprocal Ising model,” *arXiv preprint arXiv:2311.05471*, 2024.
- [146] C. Han, M. Wang, B. Zhang, M. Dykman, and H. Chan, “Coupled parametric oscillators: From disorder-induced current to asymmetric Ising model,” *Physical Review Research*, vol. 6, no. 2, p. 023162, 2024.
- [147] D. A. Paz and M. F. Maghrebi, “Driven-dissipative Ising model: An exact field-theoretical analysis,” *Physical Review A*, vol. 104, no. 2, p. 023713, 2021.
- [148] A. Hudspeth, F. Jülicher, and P. Martin, “A critique of the critical cochlea: Hopf—a bifurcation—is better than none,” *Journal of neurophysiology*, vol. 104, no. 3, pp. 1219–1229, 2010.

- [149] S. Camalet, T. Duke, F. Jülicher, and J. Prost, “Auditory sensitivity provided by self-tuned critical oscillations of hair cells,” *Proceedings of the national academy of sciences*, vol. 97, no. 7, pp. 3183–3188, 2000.
- [150] D. Mattis, “Solvable spin systems with random interactions,” *Physics Letters A*, vol. 56, no. 5, pp. 421–422, 1976.
- [151] J. L. van Hemmen, “Classical spin-glass model,” *Physical Review Letters*, vol. 49, no. 6, p. 409, 1982.
- [152] A. Szedlak, G. Paternostro, and C. Piermarocchi, “Control of asymmetric Hopfield networks and application to cancer attractors,” *PloS one*, vol. 9, no. 8, p. e105842, 2014.
- [153] M. Shiino, “Stochastic analyses of the dynamics of generalized Little-Hopfield-Hemmen type neural networks,” *Journal of statistical physics*, vol. 59, pp. 1051–1075, 1990.
- [154] M. Shiino, H. Nishimori, and M. Ono, “Nonlinear Master Equation Approach to Asymmetrical Neural Networks of the Hopfield-Hemmen Type,” *Journal of the Physical Society of Japan*, vol. 58, no. 3, pp. 763–766, 1989.
- [155] R. J. Glauber, “Time-dependent statistics of the Ising model,” *Journal of mathematical physics*, vol. 4, no. 2, pp. 294–307, 1963.
- [156] M. Suzuki and R. Kubo, “Dynamics of the Ising model near the critical point. I,” *Journal of the Physical Society of Japan*, vol. 24, no. 1, pp. 51–60, 1968.
- [157] D. J. Amit, H. Gutfreund, and H. Sompolinsky, “Spin-glass models of neural networks,” *Physical Review A*, vol. 32, no. 2, p. 1007, 1985.
- [158] Y. A. Kuznetsov, *Elements of applied bifurcation theory*. Springer, New York, 1998.
- [159] P. M. Chaikin and T. C. Lubensky, *Principles of condensed matter physics*. Cambridge University Press, Cambridge, 1995.
- [160] C.-K. Chan, T. E. Lee, and S. Gopalakrishnan, “Limit-cycle phase in driven-dissipative spin systems,” *Physical Review A*, vol. 91, no. 5, p. 051601, 2015.
- [161] R. Daviet, C. P. Zelle, A. Rosch, and S. Diehl, “Nonequilibrium Criticality at the Onset of Time-Crystalline Order,” *Physical Review Letters*, vol. 132, no. 16, p. 167102, 2024.
- [162] C. P. Zelle, R. Daviet, A. Rosch, and S. Diehl, “Universal phenomenology at critical exceptional points of nonequilibrium  $O(N)$  models,” *Physical Review X*, vol. 14, no. 2, p. 021052, 2024.

- [163] M. Dykman and M. Krivoglaz, “Fluctuations in nonlinear systems near bifurcations corresponding to the appearance of new stable states,” *Physica A: Statistical Mechanics and its Applications*, vol. 104, no. 3, pp. 480–494, 1980.
- [164] L. Personnaz, I. Guyon, and G. Dreyfus, “Information storage and retrieval in spin-glass like neural networks,” *Journal de Physique Lettres*, vol. 46, no. 8, pp. 359–365, 1985.
- [165] I. Kanter and H. Sompolinsky, “Associative recall of memory without errors,” *Physical Review A*, vol. 35, no. 1, p. 380, 1987.
- [166] F. P. Kelly, *Reversibility and stochastic networks*. Cambridge University Press, 2011.
- [167] U. C. Täuber, *Critical dynamics: a field theory approach to equilibrium and non-equilibrium scaling behavior*. Cambridge University Press, 2014.
- [168] L. Herron, P. Sartori, and B. Xue, “Robust retrieval of dynamic sequences through interaction modulation,” *PRX Life*, vol. 1, no. 2, p. 023012, 2023.
- [169] B. Derrida, E. Gardner, and A. Zippelius, “An exactly solvable asymmetric neural network model,” *Europhysics Letters*, vol. 4, no. 2, p. 167, 1987.
- [170] R. Xie and A. Kamenev, “Quantum Hopfield model with dilute memories,” *Physical Review A*, vol. 110, no. 3, p. 032418, 2024.
- [171] O. E. Orsel, J. Noh, P. Zhu, J. Yim, T. L. Hughes, R. Thomale, and G. Bahl, “Giant non-reciprocity and gyration through modulation-induced Hatano-Nelson coupling in integrated photonics,” *arXiv preprint arXiv:2410.10079*, 2024.
- [172] D. Krotov and J. J. Hopfield, “Dense associative memory for pattern recognition,” *Advances in neural information processing systems*, vol. 29, 2016.
- [173] A. Barra, A. Bernacchia, E. Santucci, and P. Contucci, “On the equivalence of Hopfield networks and Boltzmann machines,” *Neural Networks*, vol. 34, pp. 1–9, 2012.
- [174] M. Mézard, “Mean-field message-passing equations in the Hopfield model and its generalizations,” *Physical Review E*, vol. 95, no. 2, p. 022117, 2017.
- [175] H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, T. Adler, L. Gruber, M. Holzleitner, M. Pavlović, G. K. Sandve, *et al.*, “Hopfield networks is all you need,” *arXiv preprint arXiv:2008.02217*, 2020.
- [176] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.