WARSAW UNIVERSITY OF TECHNOLOGY

Ph.D. Thesis

Physical Sciences/
Natural Sciences

Krishnadas Mohandas , M.Sc.

Critical Properties of Structural Balance Models on Multiplex and Growing Networks

Supervisor
Prof. Janusz A. Hołyst , BEng, PhD, DSc,

WARSAW 2025

This thesis is dedicated to my beloved wife, Maneesha.

For her advice, her patience, and her faith,

because she always understood.

Acknowledgments

At this point, I would like to thank the people who contributed to the creation of this work. I want to express my sincere thanks to my Supervisor, Professor Janusz Hołyst, for all his help and substantive direction of the research throughout my PhD.

I would also like to thank Dr. Krzysztof Suchecki and Dr. Piotr Górski for their valuable tips during the course of the research work and in writing this work itself. My appreciation extends to the entire FENS Lab - Dr. Julian Sienkiewicz, Dr. Robert Paluch, Dr. Agnieszka Czaplicka, and Jolanta Sadowska for numerous discussions and all help during the doctoral thesis. I am grateful for the opportunity to collaborate with Giacomo Vaccario and George Andres, from ETH Zurich. Their input and shared expertise added meaningful dimensions to the research.

I like to express my gratitude to the Doctoral school for providing the support and benefits offered throughout my PhD program.

This research was supported by the Polish National Science Center through the Alphorn Grant No. 2019/01/Y/ST2/00058, and by the Dean's grants from the Faculty of Physics at Warsaw University of Technology for 2024 and 2025. It was also funded by the European Union under the Horizon Europe grant OMINO (Grant No. 101086321) and co-financed by the Polish Ministry of Education and Science.

Finally, I am thankful to all colleagues, collaborators, and friends who have contributed directly or indirectly to this journey.

Abstract

This thesis demonstrates that the multiplex nature of networks, where social agents maintain multiple types of relationships across distinct network layers, can substantially enhance the stability of cohesive, balanced social structures compared to single-layer networks.

Structural balance theory helps to explain the evolution and organization of social relationships characterized by friendly and hostile interactions. This thesis examines such interactions using agent-based models that simulate the dynamics of social systems and a Hamiltonian framework that incorporates the concept of social temperature. Together, these approaches provide insight into how social groups transition between ordered and disordered states.

In Chapter 2, the critical properties of structural balance on multiplex networks, where agents have correlated relations across multiple layers, are examined. In multiplex networks, introducing interlayer coupling between different types of relationships (layers), where intralayer dynamics follow Heider balance and interlayer interactions resemble Ising dynamics, will result in a richer equilibrium landscape characterized by various phases, including synchronized and disordered states. The critical temperatures governing transitions between these states will be modulated by the strength of interlayer coupling and can differ significantly from those observed in single-layer networks. Specifically, increased interlayer coupling will elevate the critical temperature for the paradise-disorder transition, and the presence of mismatched initial states across layers will introduce additional transitions related to synchronization at lower temperatures.

In Chapter 3, I explore how sparse network topology, specifically in Erdős-Rényi random graphs, affects the structural balance model. I focus on the impact of thermal noise in single-layer and bilayer networks. This work extends the study of structural balance beyond complete graphs to random networks, analyzing how network sparsity affects the stability of an entirely positive "paradise" state. I examine both single-layer and bilayer network configurations and provide a mean-field solution for the average link polarization that predicts a first-order transition where the critical temperature scales with the connection probability p as p^2 for a monolayer system and in a more complex way for a bilayer. The study also indicates that renormalizing the intralayer and interlayer interaction strengths to p^{-2} and p^{-1}, respectively, one can make the system's critical temperature

independent of the network density, mirroring the behavior observed in complete graphs under certain conditions.

Finally, in Chapter 4, the growth dynamics of a signed network governed by Heider's structural balance theory are examined, where the evolution results in the emergence of two antagonistic, fully connected cliques. New nodes (agents) make contact with a random existing agent and join one of the factions, with an attachment bias toward the group they contacted. The dynamics of this process closely follow a randomized Pólya urn model. Except for the extreme case of complete preference (full bias), the relative sizes of the two groups tend toward equal proportions, preventing either group from achieving complete dominance. The evolution of faction sizes exhibits three distinct regimes based on the bias parameter. In the anti-bias regime, the dynamics reflect contrarian behavior, driving the rapid equalization of group sizes, regardless of initial differences. In the low-bias regime, the initial size differences persist over time, obscured by fluctuations, with the faction size distribution remaining unimodal. In the high-bias regime, the system undergoes a transition where the faction size distribution becomes bimodal. The findings provide insight into how antagonistic communities evolve and self-organize under simple interaction rules.

Keywords: *structural balance, multiplex network, phase transition, growing network, agent-based models.*

Streszczenie

Niniejsza rozprawa udowadnia, że wielowarstwowa natura sieci – w której społeczni agenci utrzymują wiele typów relacji w różnych warstwach sieci – może znacząco zwiększyć stabilność spójnych, zrównoważonych struktur społecznych w porównaniu z sieciami jednowarstwowymi.

Teoria równowagi strukturalnej pomaga wyjaśnić ewolucję i organizację relacji społecznych charakteryzowanych jako przyjazne lub wrogie. Rozprawa ta rozpatruje takie interakcje, wykorzystując agentowe modele dynamiki systemów społecznych oraz podejście hamiltonowskie uwzględniające koncepcję temperatury społecznej. Łącznie podejścia te pokazują, jak grupy społeczne przechodzą między stanami uporządkowanymi i nieuporządkowanymi.

W rozdziale 2 zbadano kluczowe właściwości równowagi strukturalnej w sieciach wielowarstwowych, w których agenci mają skorelowane relacje w wielu warstwach. W sieciach takich wprowadzenie sprzężenia między różnymi typami relacji (warstwami), gdzie dynamika wewnątrzwarstwowa jest zgodna z równowagą Heidera, a interakcje międzywarstwowe przypominają dynamikę Isinga, prowadzi do bardziej złożonych stanów równowagi charakteryzującego się różnymi fazami, w tym stanami zsynchronizowanymi i nieuporządkowanymi. Temperatury krytyczne przejść między tymi stanami zależą od siły sprzężenia międzywarstwowego i mogą znacząco różnić się od temperatur obserwowanych w sieciach jednowarstwowych. W szczególności, zwiększone sprzężenie międzywarstwowe podnosi temperaturę krytyczną dla przejścia raj-nieporządek, a obecność niedopasowanych stanów początkowych w różnych warstwach wprowadza dodatkowe przejścia związane z synchronizacją w niższych temperaturach.

W rozdziale 3 badam wpływ topologii sieci, w szczególności przypadek grafów losowych Erdős-Rényi, skupiając się na analizie szumu termicznego w dynamice w sieciach jedno- i dwuwarstwowych. Niniejsza praca rozszerza badanie równowagi strukturalnej poza grafy pełne, na sieci losowe, analizując, jak gęstość sieci wpływa na stabilność całkowicie pozytywnego stanu „raju". Badam konfiguracje sieci jednowarstwowych i dwuwarstwowych, wskazując rozwiązanie średniopolowe dla polaryzacji łącz, które przewiduje istnienie przejścia pierwszego rodzaju, którego temperatura krytyczna skaluje się z prawdopodobieństwem połączenia p jak p^2 dla układu jednowarstwowego i w bardziej złożony sposób dla układu dwuwarstwowego. Badanie wskazuje również, że renormalizując siły oddziaływań

wewnątrzwarstwowych i międzywarstwowych odpowiednio do p^{-2} i p^{-1}, można uniezależnić temperaturę krytyczną systemu od gęstości sieci, odzwierciedlając zachowanie obserwowane na grafach pełnych w określonych warunkach.

Na koniec, w rozdziale 4, przeanalizowałem dynamikę wzrostu sieci ze znakiem, zgodnej z zasadami teori równowagi strukturalnej Heidera, gdzie ewolucja prowadzi do powstania dwóch antagonistycznych, w pełni połączonych klik. Nowe węzły (agenci) nawiązują losowo kontakt z dowolnym istniejącym agentem i dołączają do jednej z grup, wykazując słabą lub silną preferencję do wyboru tej grupy, z którą nawiązali pierwszy kontakt. Dynamika tego procesu jest matematycznie opisana przez losowy model urnowy P'olya. Z wyjątkiem przypadku skrajnie silnej preferencji na dołączenie do grupy wylosowanego agenta, względne rozmiary obu grup dążą do równych proporcji, uniemożliwiając którejkolwiek z nich osiągnięcie całkowitej dominacji. W ogólnym przypadku ewolucja rozmiarów grup wykazuje trzy odrębne typy w zależności od wartości parametru preferencji. W przypadku słabej preferencji dynamika odzwierciedla zachowania kontrariańskie, przyśpieszając wyrównywanie liczebności grup, niezależnie od ich początkowych wielkości. W przypadku umiarkowanej prefencji początkowe różnice liczebności grup utrzymują się w czasie, przesłonięte są jednak fluktuacjami i rozkład liczebności grup pozostaje jednomodalny. W przypadku silnej prefencji zachowanie systemu ulega jakościowej zmianie a rozkład liczebności frakcji staje się dwumodalny. Wyniki te wskazującjak ewoluować mogą antagonistyczne społeczności, samoorganizując się w ramach prostych reguł interakcji.

Słowa kluczowe: *Równowaga strukturalna, sieć multipleksowa, przejście fazowe, rosnąca sieć, modele agentowe.*

Preface

Complex networks provide a framework for understanding interconnections in natural, social, and engineered systems [1, 2]. These networks, composed of nodes and links, capture the nontrivial patterns of interactions observed in domains ranging from biological systems to human societies. By studying their topology and dynamics, researchers gain insights into phenomena as varied as disease outbreaks, information diffusion, and emergent cooperative structures. Within the landscape of complex networks, **signed** networks hold special significance. In such networks, links between individuals are represented as positive or negative signs, representing friendly or antagonistic relationships [3]. The presence of these dual types of interactions introduces rich challenges and opportunities, especially for modeling social systems where alliances and rivalries shape community organization and collective behavior.

Structural balance theory (SBT), developed by Heider [4] and formalized by Cartwright and Harary [5], suggests that signed networks evolve toward balanced configurations where relationship tensions are minimized. The fundamental unit of analysis in this theory is a triad, where structural stability emerges when triadic relationships adhere to specific balance principles such as "the friend of my friend is my friend" and "the enemy of my enemy is my friend." The implications of SBT have driven extensive research into how networks self-organize and reach equilibrium under competing social forces [6, 7]. A key consequence of structural balance dynamics is the emergence of distinct, cohesive groups with positive internal relations and negative ties to rival groups. Recent research extends these concepts to multiplex and growing networks, helping to explain phenomena such as polarization and community resilience [8, 9].

This dissertation is motivated by the following hypothesis:

1. **In multiplex networks, the introduction of interlayer coupling between different types of relationships increases the overall stability of the social structure.**

2. **A stochastic process can capture the essential dynamics of polarization and faction formation in growing networks, without relying solely on specific network properties.**

Guided by these hypotheses, the thesis pursues two main objectives:

1. To investigate how interlayer coupling influences the configuration of signed relationships across multiple layers in multiplex networks, by extending Heider balance theory and analyzing the critical properties and temperature-induced transitions that result from this coupling.

2. To determine the combined effect of probabilistic attachment bias and structural balance on faction formation and dominance, and to identify the conditions that lead to disparities in the sizes of opposing factions.

The three chapters (2, 3, and 4) examine various aspects of structural balance in signed networks, progressing from static multiplex systems to sparse networks and ultimately to growth processes. Chapter 2 examines structural balance on multiplex networks where agents maintain correlated signed relationships across multiple layers, with intralayer dynamics following Heider balance and interlayer interactions resembling Ising coupling. Chapter 3 extends this analysis from complete graphs to sparse Erdős–Rényi networks in both single-layer and bilayer settings under thermal noise, providing a mean-field solution that predicts a first-order transition with a critical temperature scaling with the connection probability for monolayer and in a much complex way for a bilayer. Chapter 4 shifts focus to growth dynamics, modeling signed networks as randomized Pólya urn processes where antagonistic, fully connected factions emerge as new agents make contact with a random existing agent and join one of the factions with a bias towards the group they made contact with.

This thesis summarizes the works of the following three articles,

[1] K. Mohandas et al., "Critical properties of Heider balance on multiplex networks", *Physical Review E* **2024**, *109*, DOI 10.1103/PhysRevE.109.044306

[2] K. Mohandas et al., "Paradise-disorder transition in structural balance dynamics on Erdős-Rényi graphs", *Physical Review E* **2025**, *111*, DOI 10.1103/PhysRevE.111.024310

[3] K. Mohandas et al., "Structurally balanced growing network as randomized Pólya urn process", *arXiv:2510.24659v1 [physics.soc-ph]* **2025**, DOI 10.48550/arXiv.2510.24659

This thesis has benefited from feedback and exposure at several conferences and workshops. I had the privilege to present and discuss parts of this research at:

- FENS (Polish Symposium on Physics in Economics and Social Sciences) 2023 - Wrocłow, Poland
- DPG (German Physical Society) 2023 - Dresden, Germany

- Complex networks workshop 2023 - Como, Italy
- ALPHORN Workshop 2024 - Zurich, Switzerland
- DPG (German Physical Society) 2025 - Regensburg, Germany
- FENS (Polish Symposium on Physics in Economics and Social Sciences) 2025 - Warsaw, Poland

List of Acronyms

BA Barabási-Albert

CTD Constrained triad dynamics

ER Erdős-Rényi

HB Heider balance

MC Monte Carlo

ME Master equation

MF Mean field

RE Rate equation

SB Structural balance

SNR Signal to noise ratio

List of Symbols

The following list contains a general notation used throughout this work. However, some models can represent different quantities by these symbols. In case of redundancy, the definition given in a chapter takes priority over this list.

α, β	Indices to denote layers
$\langle \dots \rangle$	Average
$\mathcal{H}$	Hamiltonian
$\tilde{T}$	Normalized Temperature
a	Rescaled intralayer interaction strength
$A^{(\alpha)}$	Heider/ intralayer interaction strength in layer α
A_{ij}	Adjacency matrix
d	Rescaled interlayer interaction strength
i, j, k	Indices to denote agents or nodes
K	Ising/ interlayer interaction strength
L	Number of layers in the multiplex network
L^+	Number of positive links
M	Number of triads each edge belongs to
N	Number of agents or nodes
o	Overlap metric
p	Attachment bias
T	Temperature
T_c	Critical temperature
x_{ij}^{α}	Sign of the link connecting the nodes i and j in layer α

Contents

Chapter 1

Introduction

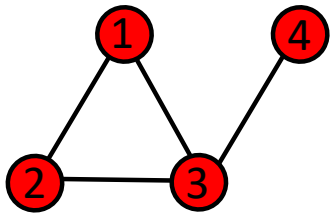
1.1 Fundamentals of networks

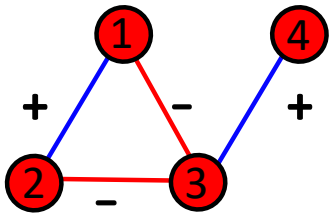
1.1.1 Complex system as a signed network

Complex systems manifest across many disciplines, encompassing social, biological, economic, and technological domains [13–24]. These systems are characterized by their nontrivial interactions, often displaying collective behavior that cannot be inferred solely from individual components. A rigorous framework encapsulating their complexity is required to effectively analyze and model such systems. Network theory provides a robust methodological approach wherein entities (nodes) and their interactions (edges) form a structured representation of complex relationships. The most basic kind of network is a simple graph: undirected, unweighted, with no self-loops or annotations on the nodes. A simple graph can then be called a binary undirected network. A simple way to represent such a network is by an adjacency matrix.
A binary undirected network with N nodes is specified by an $N \times N$ adjacency matrix:

$$A_{ij} = \begin{cases} 1, & \text{if a relationship exists between } i \text{ and } j, \\ 0, & \text{if no interaction is present between } i \text{ and } j. \end{cases} \tag{1.1}$$

In many real-world scenarios, interactions within a complex system are not merely binary (presence or absence of a link) but can carry additional attributes such as strength, directionality, or sentiment. Signed networks offer a refined extension of traditional network models by incorporating positive and negative edge weights to signify friendly and antagonistic relationships. This paradigm is particularly instrumental in modeling social structures, where alliances and conflicts coexist, biological regulatory systems exhibit activation and inhibition dynamics, and economic networks encompass competitive and collaborative interactions.





0	1	1	0
1	0	1	0
1	1	0	1
0	0	1	0

(a)

0	1	-1	0
1	0	-1	0
-1	-1	0	1
0	0	1	0

(b)

Figure 1.1: Adjacency matrix for (a) undirected network and (b) signed undirected network.

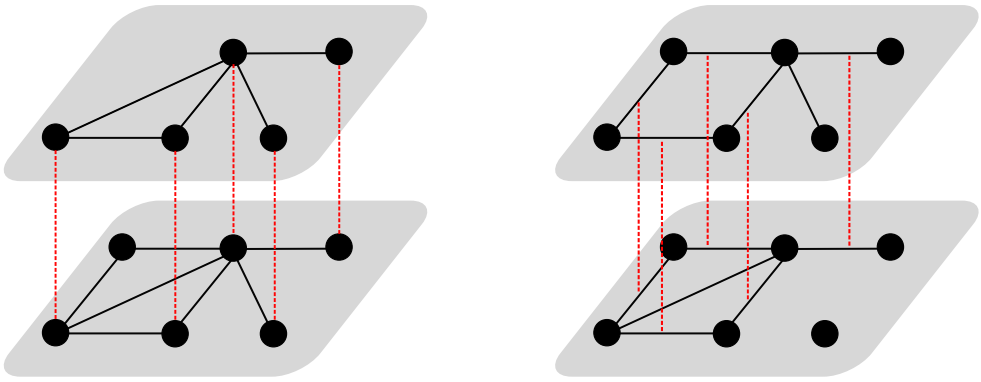
For a network with N nodes, the adjacency matrix is an $N \times N$ square matrix, where each element A_{ij} encodes the interaction between nodes i and j. Specifically, in a signed network:

$$A_{ij} = \begin{cases} +1, & \text{if a positive (friendly) relationship exists between } i \text{ and } j, \\ -1, & \text{if a negative (antagonistic) relationship exists between } i \text{ and } j, \\ 0, & \text{if no direct interaction is present between } i \text{ and } j. \end{cases} \tag{1.2}$$

1.1.2 Multiplex network

Many real-world systems do not operate in isolation. On the contrary, they are interconnected, and what happens at a single level of interaction affects the structure and function of another interconnected layer. Multiplex structures [25] find widespread application in characterizing diverse systems such as social groups, transportation networks, and biological frameworks, including protein-protein interaction systems [26–32]. Humans are inherently complex beings. When they form interconnected groups bound by relationships, the intricacy only deepens. Individuals frequently maintain a plethora of relationships, varying in nature from familial and professional ties to friendships and online connections. A multiplex network representation is invaluable in addressing this diversity of relations [28, 33, 34]

In conventional multiplex networks, interlayer edges exclusively connect the same agent



(a) (b)

Figure 1.2: Multiplex network where interlayer edges (a) connect the nodes representing the same agents and (b) connect the edges between same pair of agents.

across different layers, ensuring that interactions remain constrained to identical agents in distinct layers or contexts. This structure enables researchers to explore how various types of relationships between entities influence one another. However, in this thesis, I introduce a modified multiplex framework where interlayer connections are not established between identical agents across layers but between the edges that connect the same pair of agents in different layers (see Fig. 1.2). In this alternative approach, edges serve as links between layers rather than nodes. Consequently, if two agents do not share a connection in one layer, the corresponding interlayer connections between them will be absent across all layers. This structural adaptation provides a novel perspective on multiplex network analysis, offering deeper insights into the evolution of relationships and their influence on interconnected systems.

This layered representation of networks is pivotal for capturing the complexity of systems where multiple types of relationships coexist. For instance, in a social network, individuals may interact through various channels such as face-to-face meetings, online communication, and professional collaborations. Each type of interaction can be represented as a separate layer, yet the underlying entities (i.e., the individuals) remain consistent across layers. This approach enables the analysis of each interaction type in isolation. It provides insights into how the layers influence one another, leading to emergent properties that cannot be observed when examining a single network in isolation.

1.1.3 Random network

Random networks, originally introduced by Erdős and Rényi [35], are a fundamental concept in complex network theory. They serve as a null hypothesis in network analysis, allowing us to compare real-world networks—often marked by non-random traits such as clustering or distinct degree distributions—against a stochastic baseline to uncover meaningful patterns. Social network analysis provides a reference point for understanding how

friendships and professional relationships form compared to real-world networks [1, 36]. In biological systems, they help describe neural connectivity and protein interactions, contributing to insights into brain function and cellular processes [37–39]. In communication and technological networks, they aid in designing and optimizing data transmission protocols [40, 41]. There are two primary models for generating random networks. The first, known as the $G(N, e)$ model, consists of N labeled nodes connected by exactly e edges, where the edges are chosen randomly from the $N(N-1)/2$ possible connections. This process results in an ensemble of graphs, each with the same number of nodes and edges but different configurations. A specific random network from this model is obtained by randomly selecting one of these graphs.
The second model, $G(N, p)$, also starts with N labeled nodes, but instead of a fixed number of edges, each possible edge is included independently with probability p. This means the number of edges varies from one instance of the network to another, following a binomial distribution. Despite this variation, both models tend to exhibit similar structural properties when averaged over many instances. The expected number of edges in the $G(N, p)$ model is given by $E(e) = p\frac{N(N-1)}{2}$. This formulation allows for greater flexibility in modeling real-world networks, where connections may form probabilistically rather than being predetermined. The $G(N, p)$ model will be used in the following chapters of this thesis.

1.1.4 Growing network

The growth of networks is often modeled as a process in which new nodes are sequentially added and establish connections with existing nodes based on specific attachment mechanisms. One of the most extensively studied mechanisms is preferential attachment, as described in the Barabási-Albert (BA) model [2]. According to this model, new nodes are more likely to connect with nodes with a high degree, meaning that well-connected nodes continue accumulating links at a higher rate. This process gives rise to scale-free networks, where the degree distribution follows a power law, resulting in a few highly connected hubs that play a central role in the network structure. An extension of this growing network model with node and link creation to incorporate link directionality has also been well studied [42].

While preferential attachment is a fundamental principle in network growth, alternative mechanisms have been explored to capture the complexities observed in real-world networks. One such approach is the fitness model, where an intrinsic fitness parameter determines a node's likelihood of gaining new connections rather than just its degree [43]. This allows certain nodes to become highly connected despite starting with fewer links. Another approach considers nonlinear growth dynamics, where attachment probabilities depend on additional factors such as node age or more complex interactions between

existing connections[44]. Furthermore, some models focus on optimal growth strategies, where network evolution is guided by rules designed to shape its structural properties. For example, [45] proposed connecting new nodes preferentially to low-degree nodes can lead to highly centralized star-like structures, influencing the overall topology and function of the network.

1.2 Structural balance theory

1.2.1 Basic definition

Structural balance (SB)[1] theory (also called Heider's balance theory) was proposed by Heider[4], whose work conceptualized balance theory at the cognitive level, focusing on how individuals perceive and organize their attitudes toward people and objects to maintain psychological consistency. He proposed that cognitive systems tend toward states of balance configurations where relationships among entities are harmonious. Cartwright and Harary [5] later expanded this framework by formalizing it mathematically, transforming Heider's psychological insights into a structural theory applicable to broader signed social networks. The fundamental premise of structural balance theory rests on categorizing dyadic relationships as either positive (friendship, alliance, agreement) or negative (enmity, conflict, disagreement).

At its core, structural balance theory operates on a simple principle: a triad (three actors and their relationships) is balanced if the product of the signs of the relationships is positive [46]. This translates to a set of four axioms about combinations of friendship/enmities:

- Friends of friends are friends
- Enemies of enemies are friends
- Friends of enemies are enemies
- Enemies of friends are enemies

Balance theory posits that specific configurations of positive and negative relationships create stable social structures, while others generate tension that drives social change. Within this framework, unbalanced triads produce psychological and social tension, compelling the actors to modify their relationships to achieve balance. This dynamic process fundamentally shapes the evolution of signed networks toward increasingly stable configurations.

In signed networks, links are assigned either a positive sign (+) or a negative sign (−). The positive links (+) indicate favorable relations between the connected nodes,

[1]Structural Balance and Heider balance (HB) are used interchangeably throughout this thesis.

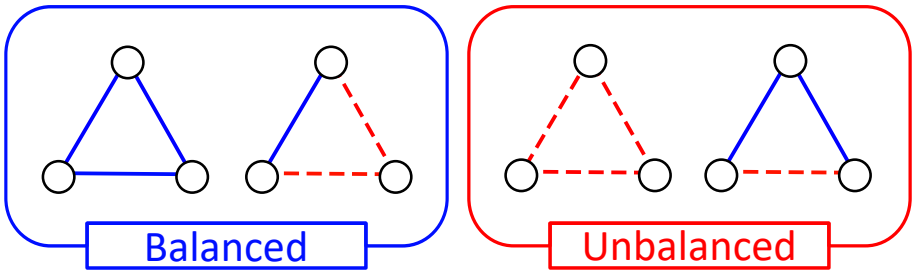


Figure 1.3: The division of triads based on structural balance. The triads on the left indicate a balanced triad with an even number of negative edges. The triads on the right are unbalanced and have an odd number of negative edges. Solid blue line indicates a positive edge, while the dashed red line represents a negative edge.

while negative links $(-)$ represent unfavorable relations. These signed connections effectively model numerous real-world social interactions across different dimensions, such as trust/distrust, like/dislike, praise/blame, and influence/negative influence ([6, 28]).

As illustrated in Figure 1.3, the first two triangle configurations are balanced, while the latter two are unbalanced. For any triangle comprising nodes i, j, and k, balance can be mathematically determined: the triangle is balanced if $x_{ij}x_{jk}x_{ki} = 1$ and unbalanced if $x_{ij}x_{jk}x_{ki} = -1$, where x_{ij} represents the sign of link ij. By convention, $x_{ij} = 1$ when link l_{ij} is positive and $x_{ij} = -1$ when the link is negative. This mathematical formulation reveals an elegant pattern: a triangle is balanced when it contains an odd number of positive links and unbalanced otherwise.

1.2.2 Statistical physics of structural balance

Statistical physics has emerged as a powerful framework for analyzing complex social systems, giving rise to fields like "sociophysics" or "social physics." By applying principles from statistical mechanics to social phenomena, researchers can model, analyze, and predict collective behaviors that emerge from individual interactions. The fundamental idea behind sociophysics is that many social systems, like physical systems, exhibit emergent properties that are not easily predictable from the individual behavior of their constituents. For example, the formation of public opinion, the spread of rumors, the dynamics of financial markets, and the evolution of social networks can be understood as collective phenomena arising from individual interactions.

One of the earliest and most influential applications of statistical physics to social systems was Granovetter's work on the threshold model of collective behavior [47]. This model explains how individual thresholds for participating in collective actions can lead to abrupt and large-scale social change. Similarly, Schelling's segregation model [48] demonstrates how even weak individual preferences for neighborhood composition can result in significant macroscopic segregation patterns.

The Ising model, a cornerstone of statistical physics, has been extensively used to

model opinion dynamics. Deffuant et al. [49] and Hegselmann and Krause [50] proposed continuous opinion models based on bounded confidence, where individuals only interact with those whose opinions are within a certain range. These models have been instrumental in understanding consensus formation and polarization in social groups. Furthermore, network theory, often rooted in statistical physics concepts like percolation and random graphs, has been used to study social networks. Barabási and Albert's scale-free network model [2] explained the emergence of power-law degree distributions in real-world networks, including social networks, highlighting the role of preferential attachment.
Sociophysics also incorporates concepts from kinetic theory to model the dynamics of human crowds and traffic flow [51]. These models help understand and predict pedestrian behavior in crowded spaces and the emergence of traffic jams. More recently, computational social science has embraced these approaches, with researchers like Lazer et al.[52] and George et al. [53] leveraging big data to validate theoretical models derived from statistical physics. While traditional social science methods focus on individual motivations and behaviors, sociophysics offers complementary insights by examining how systematic patterns emerge at the collective level through relatively simple interaction rules.

The statistical and thermodynamic formulation of structural balance theory represents another significant framework in our understanding of social dynamics. Antal et al. [54] introduced Constrained triad dynamics (CTD) where relationship signs flip according to Glauber or heat-bath algorithms at zero temperature. Motivated by studies in statistical mechanics, the CTD model can be generalized so that randomness plays a key role in changing the sign of a link, while providing a bias towards global reduction of tension. As in a CTD model, the total energy of the system is

$$\mathcal{H} = -\sum_{ijk} x_{ij}x_{jk}x_{ki}, \tag{1.3}$$

summed over all triads ijk [55]. A possible outcome of such dynamics is a *jammed states*, which are network configurations containing a finite number of unbalanced triads that cannot evolve toward complete balance through local dynamics alone. Such jammed states greatly exceed the number of balanced states, and the probability of reaching them vanishes as the system size increases. By introducing an energy landscape, the properties of such jammed states have also been studied extensively [56, 57]. Marvel et al. [56] made an approach to formalize Heider's classical balance theory within an energy landscape framework, where each network state is assigned an energy value corresponding to the number of unbalanced triads. Lee et al. [58] further developed this thermodynamic approach by drawing direct connections to the Ising model, in which temperature represents social volatility. A single parameter called social temperature represents the model incentive to seek balance. The term social temperature has been used in other models

employing this statistical physics approach to analyze social behavior [59]. Their findings suggest that at low temperatures, networks self-organize into highly ordered, balanced states, whereas at higher temperatures, social interactions become increasingly erratic, preventing the emergence of balance.

Based on the definition provided by CTD, the energy of the system $H = -1$ and $+1$ for a balanced and an unbalanced triad, respectively. At every time step, we flip a randomly chosen link with probability P, defined as,

$$P(x_{ij} \to -x_{ij}) = \frac{e^{H_{ij}/T}}{e^{-H_{ij}/T} + e^{+H_{ij}/T}} = \frac{1}{1 + e^{\Delta H_{ij}/T}} \tag{1.4}$$

where

$$H_{ij} = -x_{ij} \sum_{k \neq i,j} x_{jk} x_{ki} \tag{1.5}$$

is the Hamiltonian part dependent on the state of the link ij, and $\Delta H_{ij} = -2H_{ij}$ indicates the change in the total energy due to the link flipping every time step t. This formulation allows for the use of statistical physics tools to model how signed social networks evolve toward structurally balanced states. This thermodynamic framework has proven particularly valuable for analyzing phase transitions in large-scale networks, as demonstrated by Rabbani et al. [60], who identified critical thresholds where networks transition from predominantly balanced to unbalanced states. The statistical physics approach has also illuminated the existence of "jammed states" or "structural glasses" in social networks—configurations that remain trapped in local energy minima despite containing unbalanced triads—providing a formal explanation for the persistence of certain social tensions observed in empirical studies by Kirkley et al. [61].

1.2.3 Structural balance in multiplex network

Several models have been developed, grounded in the principles of SBT, aiming to elucidate human behavior within social networks [6, 62–65]. However, the exploration of structural balance within the realm of multiplex networks remains limited to a handful of investigations [66–70].

In the context of multiplex networks, the concept of structural balance, gains significant complexity. Górski et al.[66] directly address this by investigating the influence of interlayer coupling on Heider balance in bilayer networks. They define link multiplexes, where interlayer coupling occurs between corresponding links, and demonstrate through numerical simulations and analytical calculations that such coupling impedes the emergence of SB. These findings suggest that the multilayer nature of social interactions may explain the frequent absence of SB in real-world social systems, highlighting the necessity to consider multiplexity when studying structural balance.

While other related studies do not explicitly delve into Heider balance, they contribute

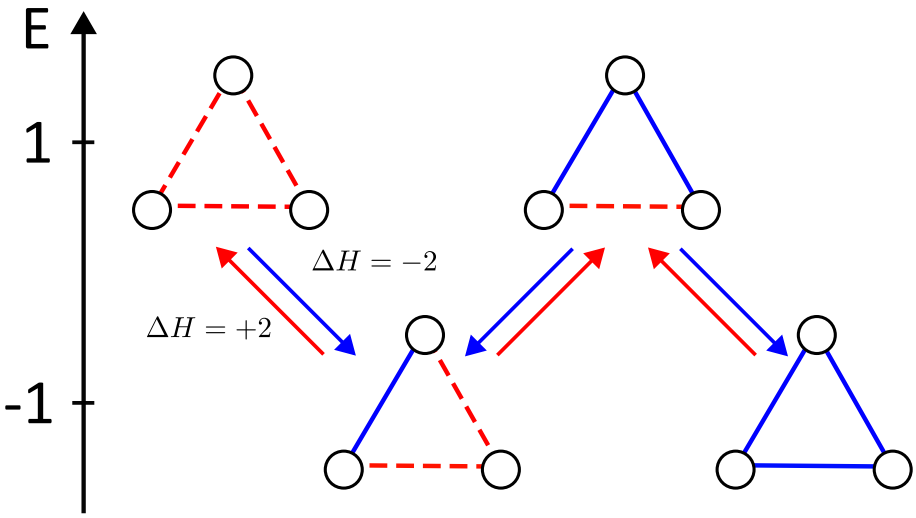


Figure 1.4: The energy landscape of signed complete network on 3 nodes. Solid blue lines show friendship links and dashed red lines represent hostility links. The triads with energy -1 $(+1)$ are balanced (imbalanced). Solid blue lines represent positive edge and dashed red lines indicate negative edge.

to understanding triadic relationships and their impact on network structure within multiplex contexts. Cozzo et al. [68] generalize transitivity to multiplex networks, emphasizing the novel degrees of freedom introduced by layered structures and observing that social networks tend to close triads at each layer. Similarly, Klimek et al. [71] employ a voter model based on triadic closure to explore community structure in online social multiplex networks, and Klimek and Thurner [72] highlight the role of triadic closure in driving scaling laws. Triadic closure, a mechanism closely linked to forming balanced states in signed networks, is a fundamental aspect of structural balance in monoplex networks.

In this dissertation, in chapters 2 and 3, I extend the Heider Balance (HB) concept to the domain of multiplex networks where we consider a similar concept of link multiplex as described in [66], wherein nodes (agents) are linked by various kinds of relations corresponding to different network layers.

1.3 Pólya urn process

The Pólya urn process is a fundamental stochastic model in probability theory and statistics that illustrates a self-reinforcing mechanism through which the probabilities of outcomes evolve based on prior events. Originating from George Pólya's work in the early 20th century, this model is often depicted through an urn containing colored balls, where the existing composition of the urn influences the likelihood of drawing a particular color. In its simplest form, an urn contains balls of different colors. When a ball is randomly drawn, it is returned to the urn along with additional balls of the same color, creating a positive feedback loop that reinforces initial conditions [73] (see Fig. 1.5). As such,

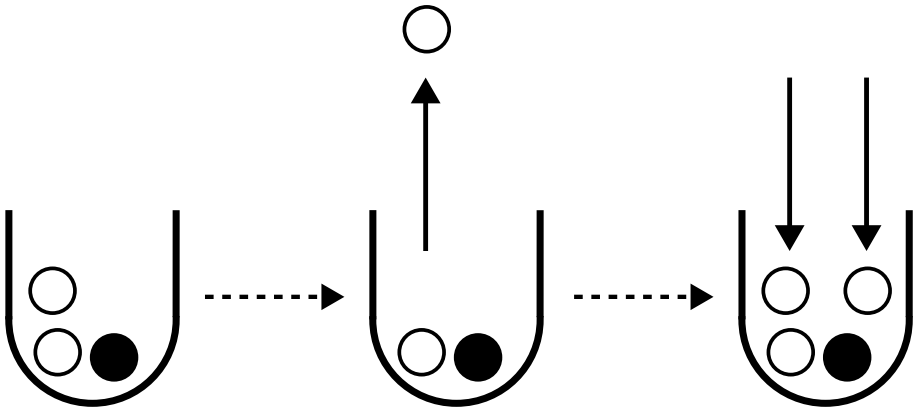


Figure 1.5: An illustration of a simple Pólya urn model where a ball is drawn at random from an urn and replaced, and an additional ball of the same color is added

the Pólya urn process serves as a powerful representation of phenomena characterized by increasing returns and preferential attachment, making it applicable in fields such as biology, economics, and social sciences [74]

The model's notable feature is its reinforcement property, where the probability of drawing a color increases as more balls of that color are drawn, leading to significant implications in the analysis of wealth distribution, social networks, and other systems with cumulative advantages. This characteristic is observed in real-world scenarios, such as the dynamics of follower accumulation on social media platforms, where initial advantages can result in exponential growth over time. This process elegantly captures the "rich-get-richer" phenomenon observed across numerous social systems. In economic contexts, the Pólya process helps explain wealth concentration dynamics. Initial advantages in resource accumulation can compound over time, leading to increasingly skewed wealth distributions [75]. Labor markets exhibit similar patterns, where early career success can lead to compounding advantages in professional advancement.

Research into the Pólya urn process has yielded a wealth of asymptotic properties, including central limit theorems that describe the behavior of the distributions of drawn colors as the process evolves. Under certain conditions, the distribution converges to a normal distribution, reinforcing the model's utility in understanding long-term outcomes in stochastic systems [76, 77].

1.3.1 Faction formation through Pólya process and structural balance

Polarization in social networks emerges through interacting social and structural mechanisms as networks grow, notably homophily, structural balance, and tribalism, which can produce either enduring dominance by one group or coexistence of similarly sized rival factions [78–80]. Homophily fosters echo chambers and cohesion among like-minded

individuals, while structural balance organizes ties into two camps with positive relations inside groups and negative relations between them, and tribalism heightens in-group solidarity [81]. Together, these mechanisms explain the emergence and persistence of polarized factions [82, 83], though it remains unclear how much they dynamically drive polarization. It is also possible that initial impressions and visibility play decisive roles. Some studies report limited changes in alignment over time, highlighting the significant impact of early choices [84]. Rather than modeling polarization within specific network topologies through individual-level interactions, a stochastic process can capture the essential dynamics at the group level. The polarization in growing factions can then be equivalently interpreted through structural balance theory or Pólya reinforcement.

While recent works have applied Pólya urn frameworks to study network dynamics, they typically do not model the simultaneous evolution of structural balance and network growth [85, 86]. For instance, the model in Ref. [87] combines structural balance principles with Pólya processes to describe how link signs evolve under social reinforcement, incorporating mechanisms such as homophily and xenophobia. However, this approach examines dynamics among existing nodes without introducing new participants—each agent draws from their local urn comprising their neighbors' signs, but the network structure itself remains fixed.

Chapter 2

Structural balance in multiplex networks: exploring stability and phase transitions

2.1 Motivation

In this chapter, the concept of SB within the framework of multiplex networks is studied, where nodes—representing agents—are interconnected through multiple relationship types, each associated with a distinct network layer. Consider a scenario in which individuals work within the same corporate environment while simultaneously maintaining interactions in a shared sports club. The social ties established in these different contexts correspond to intralayer links within separate network layers, with the nodes remaining consistent across all layers to represent the same individuals. Each layer encapsulates signed relationships specific to a particular context, and no direct signed links span across layers.

Interpersonal relationships are inherently context-dependent—two colleagues may be professional rivals in the workplace while maintaining friendly interactions at a sports club. However, I assume that individuals tend to maintain consistency in their relationships across different contexts. A positive relationship in one setting is likely to affect interactions in another, and vice versa. This interdependence creates a natural tension when signed relations across layers do not align. To capture this phenomenon, I introduce a coupling mechanism that promotes coherence between the signed relations of the same individuals across different layers, drawing an analogy to the spin coupling in the Ising model of ferromagnets.

My main objective is to examine how interlayer coupling affects the equilibrium states of signed relations across multiple network layers. I start with an analysis of a duplex network structure, where intralayer relationships in each layer constitute a complete graph,

and interlayer interactions occur exclusively between corresponding link replicas across layers. This foundational model is subsequently generalized to an arbitrary number of layers. While the evolution of signed relations within each layer follows structural balance dynamics, the interlayer coupling dynamics resemble an Ising-like interaction, which governs the relation signs between the same pair of individuals across layers.

Subsequent sections extend the classical mean-field Mean field (MF) theory of Heider balance (HB) to multiplex networks. My theoretical framework predicts the onset of a discontinuous order-disorder transition induced by thermal fluctuations in both duplex and general multiplex network configurations. To substantiate these theoretical findings, I implement Monte Carlo (MC) simulations, offering empirical validation of the predicted phase transitions. Additionally, I conduct an in-depth analysis of HB dynamics in duplex networks with layers in different states, uncovering equilibrium states beyond the scope of the mean-field approximation. The observed dynamics reveal a temperature-driven transition leading to interlayer synchronization, followed by a secondary transition into a disordered phase. This research improves our understanding of structural balance in multiplex networks and reveals the phenomena that emerge from multiplex interactions.

2.2 Models and methods

2.2.1 Multiplex Networks and Coupling Mechanisms

The model represents individuals as agents occupying the vertices of a multiplex network, where their evolving relationships are encoded as signed edges. Each layer of the network corresponds to a specific relationship context—such as professional or personal interactions—while the set of vertices remains consistent across layers. The agents themselves are static entities without intrinsic dynamical properties; instead, the evolution of relationships is governed by two principal mechanisms: Heider structural balance dynamics, which drive intralayer interactions, and Ising-type coupling, which enforces correlation between equivalent relations across layers.

I consider a multiplex network comprising N nodes and L layers. I assume that each pair of agents is connected in every layer for analytical tractability, effectively forming complete graphs. This setup models small-scale social systems where all individuals are acquainted. Each dyadic relationship (i, j) is thus characterized by L signed edges, denoted as $x_{ij}^{(\alpha)} = \pm 1$, where i and j label the agents, and $\alpha = 1, 2, \ldots, L$ designates the layer index. These edge states evolve dynamically over time.

Figure 2.1 presents the duplex network representation of the proposed model, emphasizing how both intralayer and interlayer dynamics contribute to the evolution of signed relationships.

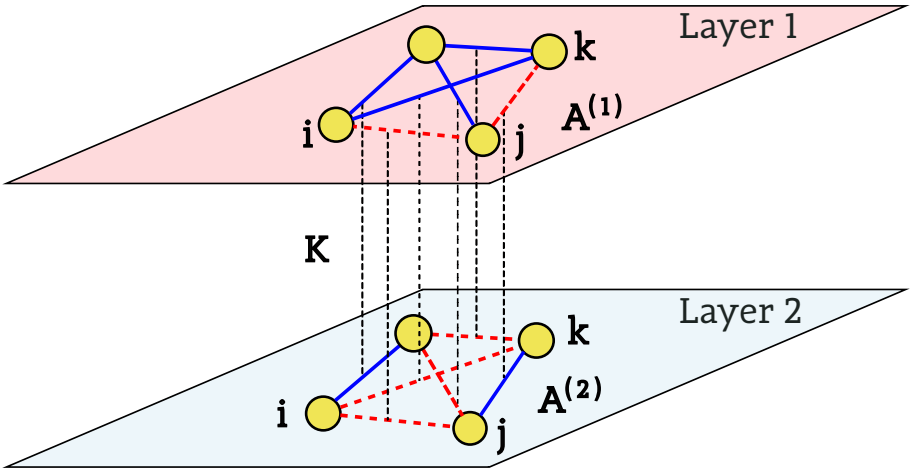


Figure 2.1: An example of interactions between edge signs (solid blue line for $+1$, broken red line for -1) on a duplex network. Edge sign changes are driven by intralayer interactions according to Heider structural balance (with strength $A^{(1)}$ or $A^{(2)}$ respectively) as well as interlayer coupling of Ising nature between the edges, with strength K. The signs of relations evolve over time, according to the heat bath dynamics with energy as expressed by Eq.2.1 and given temperature T.

2.2.2 Hamiltonian for the Dynamics of Interlayer Relationships

Having introduced the architecture of the multiplex network and the mechanisms governing both the intralayer and interlayer interactions, I will define a Hamiltonian to describe the dynamics of the system. The Hamiltonian will set the stage for subsequent analytical treatment of the system's equilibrium behavior and phase transition.

A fundamental aspect of the model is the consistency of nodes across layers, ensuring that different layers capture relationship context rather than distinct populations. Within each layer, triadic interactions play a critical role: imbalanced triads (ij, jk, ki) generate structural tension, which agents seek to minimize. This effect is quantified via associated energy: balanced triads contribute $-A^{(\alpha)}$ to the system's energy, while imbalanced triads contribute $+A^{(\alpha)}$, in accordance with Heider's structural balance theory. The coupling parameter $A^{(\alpha)} > 0$ represents the strength of the Heider interaction within each layer, promoting coherence in intralayer link signs.

Beyond intralayer effects, the model incorporates interlayer coupling, which enforces the alignment of relationships across different contexts. If a given dyadic relationship (i, j) exhibits consistent signs across layers, the associated energy contribution is $-K$; if the signs differ, the contribution is $+K$. The parameter $K > 0$ serves as a coupling strength akin to the ferromagnetic interaction in the Ising model, ensuring that the same pair of agents exhibit similar relationship patterns across contexts. Notably, this coupling applies to edge states rather than node attributes and does not manifest as explicit interlayer links.

The Hamiltonian $\mathcal{H}$ of the system is

$$\mathcal{H} = -\sum_{\alpha=1}^{L} A^{(\alpha)} \sum_{i>j>k} x_{ij}^{(\alpha)} x_{jk}^{(\alpha)} x_{ki}^{(\alpha)} - K \sum_{\alpha>\beta} \sum_{i>j} x_{ij}^{(\alpha)} x_{ij}^{(\beta)} \tag{2.1}$$

The state of the system will change according to the Hamiltonian in the presence of temperature T, which represents the uncertainty of relations or tolerance towards imbalanced relations in the social systems modeled.
The model has degenerate ground states, where any division of vertices into two groups is a ground state if all intra-group edges in all layers are positive and all inter-group edges are negative [88]. I will call this state a *two-clique state*. This also includes the so-called *paradise state*, where one group contains all vertices and the other no vertices so that all relations are positive. Due to thermal fluctuations, at any temperature $T > 0$ the system settles into an equilibrium different than the ground state and, in fact, is multistable, with the exact state reached via a dynamical process depending not only on parameters but also on the initial conditions.

2.3 Analytical Results

The system's Hamiltonian (Eq. 2.1) establishes how intralayer and interlayer coupling are encoded in the model. By narrowing the focus to duplex networks, it becomes possible to derive explicit self-consistent equations for link polarization, thereby connecting the general model to specific, solvable scenarios.

2.3.1 Mean-Field Approximation for Duplex Networks

I analyze a duplex network as a special case of the system governed by Hamiltonian (2.1) with $L = 2$, where each layer follows a complete graph topology. A mean-field approximation for a single-layer network was previously introduced in [55], in which each link x_{ij} interacts with a mean field proportional to $\langle x \rangle^2$ due to its presence in all $N - 2$ triads. Here, $\langle \dots \rangle$ represents an ensemble average over all possible configurations. This method is particularly effective in the limit $N \gg 1$, where each link interacts with a substantial fraction of all other links, justifying the mean-field assumption. Extending this approach to multiplex networks presents nontrivial challenges. A straightforward implementation of the second-layer influence as an Ising-type mean field would mistakenly imply that each edge interacts with all edges in the second layer rather than just its direct counterpart. This incorrect representation significantly distorts system behavior, particularly when interlayer interactions are strong.

To resolve this issue, I employ a mean-field approach where the fundamental interacting unit is a pair of coupled links, $\vec{x}_{ij} = [x_{ij}^{(1)}, x_{ij}^{(2)}]$, rather than a single link. This

formulation accounts for both interlayer interactions (represented as an internal energy contribution due to the coupling between $x_{ij}^{(1)}$ and $x_{ij}^{(2)}$) and intralayer interactions (modeled through the interaction of $\vec{x}_{ij}$ with a two-dimensional mean field proportional to $[\langle x^{(1)}\rangle^2, \langle x^{(2)}\rangle^2]$. Assuming thermal equilibrium at temperature T, the probability of each state $\vec{x}_{ij}$ follows the canonical ensemble:

$$P(\vec{x}_{ij}) = \frac{\exp(-h(\vec{x}_{ij})/T)}{\sum_{\vec{x}_{mn}} \exp(-h(\vec{x}_{mn})/T)}. \tag{2.2}$$

where

$$h(\vec{x}_{ij}) = -A^{(1)} \sum_{k\neq i,j} x_{ij}^{(1)} x_{jk}^{(1)} x_{ik}^{(1)} - A^{(2)} \sum_{k\neq i,j} x_{ij}^{(2)} x_{jk}^{(2)} x_{ik}^{(2)} - K x_{ij}^{(1)} x_{ij}^{(2)}. \tag{2.3}$$

For simplicity, I assume uniform interaction parameters across layers, i.e., $A^{(1)} = A^{(2)} = A$. The expected value of $\langle \vec{x}_{ij} \rangle$ is given by:

$$\langle \vec{x}_{ij} \rangle = \sum_{\vec{x}_{ij}} P(\vec{x}_{ij}) \vec{x}_{ij}. \tag{2.4}$$

In the mean-field approximation, explicit interactions between link products $x_{jk}^{(\alpha)} x_{ki}^{(\alpha)}$ are replaced by an effective interaction involving the squared mean polarization, $(x^{(\alpha)})^2$. The mean link polarization vector is defined as:

$$\vec{x} \equiv [x^{(1)}, x^{(2)}]. \tag{2.5}$$

Here, $\vec{x}$ represents the overall network polarization, while individual components $x^{(\alpha)}$ capture layer-specific polarization. The energy of a given link state can thus be rewritten as:

$$h(\vec{x}_{ij}) = -A(N-2) x_{ij}^{(1)} \left(x^{(1)}\right)^2 - A(N-2) x_{ij}^{(2)} \left(x^{(2)}\right)^2 - K x_{ij}^{(1)} x_{ij}^{(2)}. \tag{2.6}$$

By combining equations (2.6) and (2.2) into (2.4), I derive an expression for the expected link polarization $\langle \vec{x}_{ij} \rangle$. The details for finding the mean-field solution are given in section 2.3.2. Setting this expectation equal to the mean-field polarization, i.e., $\langle \vec{x}_{ij} \rangle = \vec{x}$, yields a system of self-consistent equations:

$$\begin{aligned} x^{(1)} &= f_1(x^{(1)}, x^{(2)}) \\ x^{(2)} &= f_2(x^{(1)}, x^{(2)}), \end{aligned} \tag{2.7}$$

where the functions $f_\alpha(x^{(1)}, x^{(2)})$ are defined as:

$$f_\alpha(x^{(1)}, x^{(2)}) = \frac{e^{2d} \sinh(a[\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2]) + \sinh(a(-1)^\alpha [\left(x^{(2)}\right)^2 - \left(x^{(1)}\right)^2])}{e^{2d} \cosh(a[\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2]) + \cosh(a[\left(x^{(1)}\right)^2 - \left(x^{(2)}\right)^2])}, \tag{2.8}$$

with $\alpha \in \{1, 2\}$. The parameters $a = \frac{AM}{T}$ and $d = \frac{K}{AM}$ denote the rescaled intralayer and

interlayer interaction strengths, respectively, with $M = N - 2$ representing the number of triads each edge belongs to. Note that the equations for both components differ only in the sign of the second hyperbolic sine function, expressed here as $(-1)^{\alpha}$.
Figure 2.2 presents numerical solutions for the mean polarization, derived from solving equation (2.7) across varying values of the rescaled interlayer coupling parameter d.

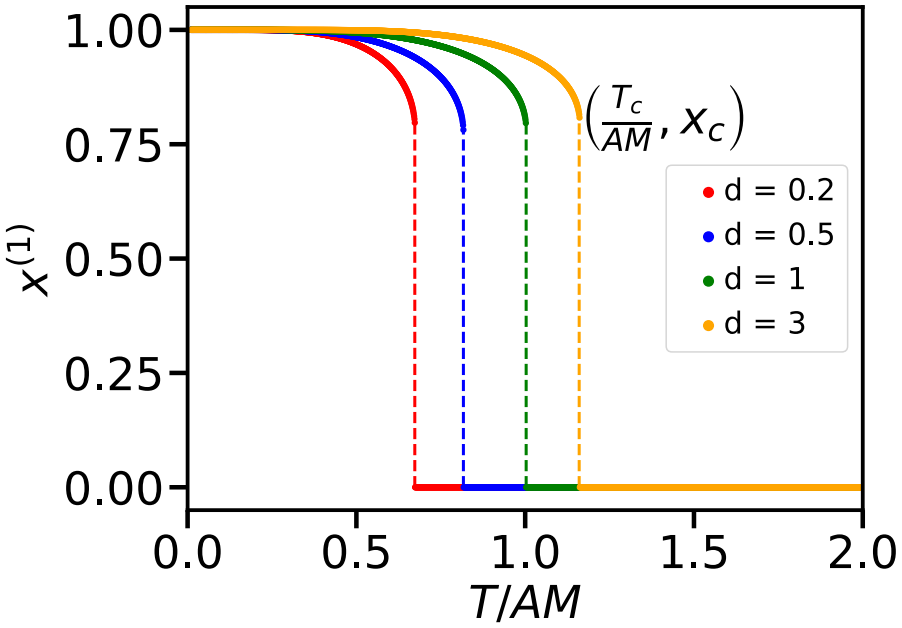


Figure 2.2: Mean field solution for average link polarization $x^{(1)}$ for various interlayer coupling strength. The coupled layers undergo a discontinuous transition at a critical temperature T_c that increases with coupling strength. The results are shown for $d = 0.2, 0.5, 1, 3$ with the transition point moving from left to right with increasing d.

As illustrated in Fig. 2.2, increasing the temperature T leads to a gradual reduction in mean polarization until a critical threshold x_c is reached, at which point a discontinuous transition to zero occurs. This corresponds to a first-order phase transition at the critical temperature T_c. The values of T_c and x_c are contingent on the interlayer coupling strength K. As K increases, T_c also increases, asymptotically approaching a saturation limit, as demonstrated in Fig. 2.6. The critical temperature T_c signifies the threshold beyond which a stable polarized state ($\vec{x} \neq 0$) ceases to exist. In the mean-field framework, the unpolarized state remains stable, leading to an irreversible transition from a polarized to an unpolarized state. Consequently, once the system transitions to an unpolarized state, reducing T below T_c does not restore polarization. However, beyond the mean-field approximation, the system exhibits more complex behavior, as discussed in Sec. 2.4.2.

From an analytical perspective, Eq. 2.7 can be framed as a recurrence relation governing the evolution of average link polarizations $[x^{(1)}, x^{(2)}]$, capturing the influence of both the mean field and the system's previous state. This approach facilitates an exploration of multistability and potential transitions between coexisting states beyond mere

equilibrium analysis. The system dynamics are defined by:

$$\begin{aligned} x^{(1)}(t+1) &= f_1(x^{(1)}(t), x^{(2)}(t)) \\ x^{(2)}(t+1) &= f_2(x^{(1)}(t), x^{(2)}(t)) \end{aligned} \tag{2.9}$$

where $f_\alpha(x^{(1)}, x^{(2)})$ represents the right-hand sides of Eq. 2.7 evaluated at time t. Fixed points of this mapping satisfy the implicit equation given by Eq. 2.7. It is important to note that $\vec{x}_0 = [0, 0]$ remains a stable fixed point of Eq. 2.9, which can coexist with a nonzero stable solution at lower temperatures. The critical temperature T_c marks the disappearance of the nonzero solution, inducing a discontinuous transition from x_c to zero when starting from a polarized state.

The values of T_c and x_c are derived from a pair of transcendental algebraic relations (Eq. 2.12 and Eq. 2.13) that characterize the fixed point and its Jacobian matrix at the transition point where the largest eigenvalue crosses unity. The detailed derivation is presented in Sec. 2.3.2. The interdependence of T_c and x_c can be encapsulated by the auxiliary variable

$$z = e^{\frac{AM}{T_c} x_c^2} \tag{2.10}$$

$$D = e^d, \tag{2.11}$$

which satisfies:

$$8 \ln z = \frac{(z^4 - 1)(z^4 D^2 + 2z^2 + D^2)}{z^2(z^4 + 2z^2 D^2 + 1)} \tag{2.12}$$

where $D = e^d$. Solving this equation numerically yields z as a function of D, which depends indirectly on the ratio of interlayer coupling K to intralayer coupling A (see Fig. 2.3 in the Sec.2.3.2 for the general behavior of this solution).

Once z is determined, the critical polarization $x_c(D, z)$ and the critical temperature $T_c(x_c, z)$ can be obtained via:

$$\frac{T_c}{AM} = \frac{x_c^2}{\ln z} = \left(\frac{1}{\ln z}\right)\left(\frac{D^2(z^4 - 1)}{D^2(z^4 + 1) + 2z^2}\right)^2. \tag{2.13}$$

The dependence of T_c on coupling strength is depicted in Fig. 2.6. In the special case $D = 1$ (corresponding to non-interacting layers), the system simplifies to the single-layer scenario, and Eq. 2.13 gives the critical temperature of a single-layer network.

2.3.2 Critical temperature calculation for duplex network

I will explicitly calculate the critical temperature at which a duplex network transitions from a polarized state to an unpolarized state. I begin by expressing the probabilities of link states using the canonical ensemble and then systematically derive analytic ex-

pressions for the mean link polarization in each layer as a function of temperature and coupling.

The probabilities of states $\vec{x}_{ij}$ with energy $h(\vec{x}_{ij})$, as specified in Eq. 2.6, adheres to the canonical ensemble formulation (Eq. 2.2):

$$P(\vec{x}_{ij}) = \frac{e^{\left(\frac{-AM}{T}[x_{ij}^{(1)}(x^{(1)})^2 + x_{ij}^{(2)}(x^{(2)})^2]\right)} e^{\left(-\frac{K}{T}x_{ij}^{(1)}x_{ij}^{(2)}\right)}}{\sum_{\vec{x}_{mn}} e^{\left(\frac{-AM}{T}[x_{mn}^{(1)}(x^{(1)})^2 + x_{mn}^{(2)}(x^{(2)})^2]\right)} e^{\left(-\frac{K}{T}x_{mn}^{(1)}x_{mn}^{(2)}\right)}} \tag{2.14}$$

where $M = N - 2$ denotes the number of triads containing the link ij. The expectation value $\langle \vec{x}_{ij} \rangle$ is given by:

$$\langle \vec{x}_{ij} \rangle = \sum_{\vec{x}_{ij}} \frac{e^{\left(\frac{-AM}{T}[x_{ij}^{(1)}(x^{(1)})^2 + x_{ij}^{(2)}(x^{(2)})^2]\right)} e^{\left(-\frac{K}{T}x_{ij}^{(1)}x_{ij}^{(2)}\right)}}{\sum_{\vec{x}_{mn}} e^{\left(\frac{-AM}{T}[x_{mn}^{(1)}(x^{(1)})^2 + x_{mn}^{(2)}(x^{(2)})^2]\right)} e^{\left(-\frac{K}{T}x_{mn}^{(1)}x_{mn}^{(2)}\right)}} \vec{x}_{ij}. \tag{2.15}$$

Since $\vec{x}_{ij}$ has only four discrete values, specifically $\vec{x}_{ij} \in \{(-1,-1),(-1,1),(1,-1),(1,1)\}$, the summation in Eq. (2.15) remains finite. Accordingly, Eq. (2.15) simplifies to:

$$\langle \vec{x}_{ij} \rangle = \frac{e^{\frac{2K}{T}}\left(e^{\frac{AM}{T}[(x^{(1)})^2+(x^{(2)})^2]} - e^{\frac{-AM}{T}[(x^{(1)})^2+(x^{(2)})^2]}\right) + \left(e^{\frac{AM}{T}(-1)^{\alpha}[(x^{(2)})^2-(x^{(1)})^2]} - e^{\frac{-AM}{T}(-1)^{\alpha}[(x^{(2)})^2-(x^{(1)})^2]}\right)}{e^{\frac{2K}{T}}\left(e^{\frac{AM}{T}[(x^{(1)})^2+(x^{(2)})^2]} + e^{\frac{-AM}{T}[(x^{(1)})^2+(x^{(2)})^2]}\right) + \left(e^{\frac{AM}{T}(-1)^{\alpha}[(x^{(2)})^2-(x^{(1)})^2]} + e^{\frac{-AM}{T}[(x^{(2)})^2-(x^{(1)})^2]}\right)} \tag{2.16}$$

Following the mean-field approximation, the expectation value $\langle \vec{x}_{ij} \rangle$ is equated to the mean-field parameter $\vec{x}$, yielding a system of self-consistent equations as previously given by Eqs. 2.7-2.8.

The system of self-consistent equations thus formulated can be interpreted as a mapping that enables the stability analysis of fixed points. Specifically, it allows for the determination of whether a given solution represents a stable equilibrium state of the system. The Jacobian matrix associated with the mean-field mapping, as described in Eq. (2.9), is defined as

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x^{(1)}} & \frac{\partial f_1}{\partial x^{(2)}} \\ \frac{\partial f_2}{\partial x^{(1)}} & \frac{\partial f_2}{\partial x^{(2)}} \end{bmatrix} \tag{2.17}$$

where the partial derivatives are given by

$$\begin{aligned}
\frac{\partial f_1}{\partial x^{(1)}} &= \frac{4ax^{(1)}e^{2d}(\cosh(2d)+\cosh(2a(x^{(2)})^2))}{\Delta},\\
\frac{\partial f_1}{\partial x^{(2)}} &= \frac{2ax^{(2)}(e^{4d}-1)}{\Delta},\\
\frac{\partial f_2}{\partial x^{(1)}} &= \frac{2ax^{(1)}(e^{4d}-1)}{\Delta},\\
\frac{\partial f_2}{\partial x^{(2)}} &= \frac{4ax^{(2)}e^{2d}(\cosh(2d)+\cosh(2a(x^{(1)})^2))}{\Delta}.
\end{aligned} \tag{2.18}$$

Here, the denominator is expressed as

$$\Delta = \left(e^{2d}\cosh(a[(x^{(1)})^2+(x^{(2)})^2])+\cosh(a[(x^{(1)})^2-(x^{(2)})^2])\right)^2 \tag{2.19}$$

To simplify the analysis, I impose symmetry in the mean-field variables since the two-dimensional maps are symmetric. By setting $x^{(1)} = x^{(2)} \equiv x$, the self-consistency equation reduces to

$$x = \frac{\sinh(2ax^2)}{\cosh(2ax^2)+e^{-2d}}. \tag{2.20}$$

In the limiting case of $d \to 0$, where interactions vanish, this equation simplifies to the mean-field equation of a single-layer network, consistent with Eq. (9a) in [55]. Likewise, for $d \to +\infty$, the same equation holds, albeit with an effective interaction coefficient of $2a$ instead of a. The stability of fixed points is determined by evaluating the eigenvalues of the Jacobian matrix. The eigenvalues corresponding to Eq. (2.18) are

$$\lambda_\pm = \frac{4ax(\cosh(2ax^2)+e^{\pm 2d})}{(e^d(\cosh(2ax^2)+e^{-2d}))^2}, \tag{2.21}$$

where λ_+ represents the largest eigenvalue. Stability is ensured if $\lambda_+ < 1$. The critical temperature $T_c/(AM)$ and the corresponding critical value x^c can be obtained from a system of transcendental algebraic equations, namely Eq. (2.20) governing the fixed point and the criticality condition $\lambda_+ = 1$ derived from Eq. (2.21). By multiplying Eq. (2.20) by the reciprocal of Eq. (2.21) at $\lambda_+ = 1$, I reformulate the dependence on x in terms of a new variable $z \equiv e^{ax^2}$, with $D \equiv e^d$. This leads to the compact transcendental equation for z:

$$8\ln z = \frac{(z^4-1)(z^4D^2+2z^2+D^2)}{z^2(z^4+2z^2D^2+1)}. \tag{2.22}$$

In general, this equation admits three solutions; however, only one is physically meaningful, corresponding to $z > 1$. The dependence of this solution on D is illustrated in Fig. 2.3. One can numerically calculate the largest solution z for a fixed value of d. Putting

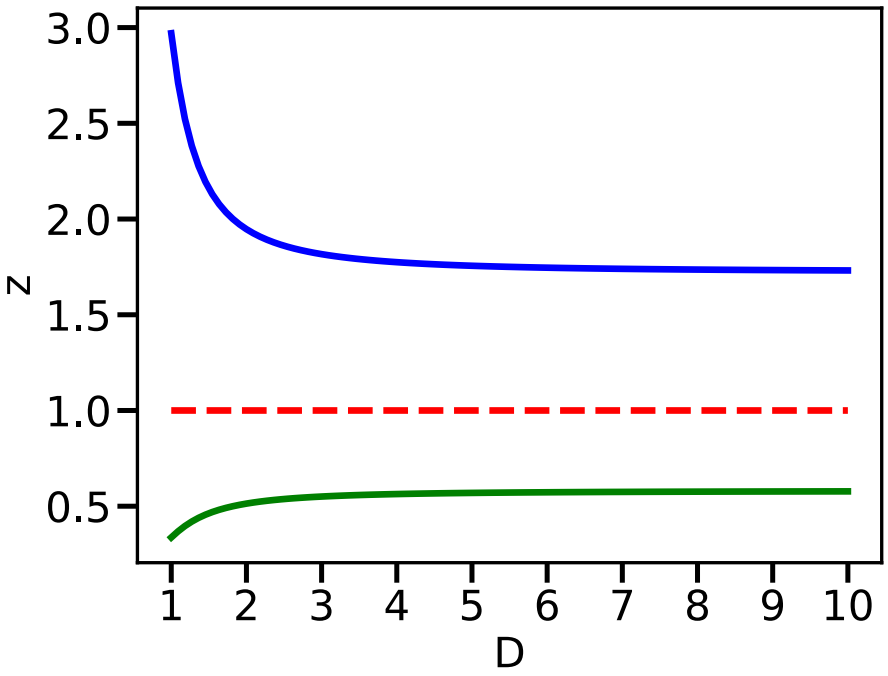


Figure 2.3: (Color online) Equation 2.22 has three solutions for auxiliary variable z that depend only on relative interlayer coupling strength D. Out of them, only solution $z > 1$ (blue top line) is physically meaningful and corresponds to a positive critical temperature T_c.

$ax^2 = \ln z$ into Eq. (2.20) gives us the link polarization x_c at the critical point

$$x_c = \frac{D^2(z^4 - 1)}{D^2(z^4 + 1) + 2z^2} \tag{2.23}$$

and having that, re-arranging definition of z gives us the inverse of a at the critical point

$$\frac{T_c}{AM} = \frac{x^2}{\ln z} \tag{2.24}$$

To understand how the saturation critical temperature relates to a number of layers, I need to define the Eq.2.24 at $d \equiv D \to +\infty$ and at $d = 0$.

$$\left.\frac{T_c}{AM}\right|_{d\to+\infty} = \frac{1}{\ln(z_\infty)} \left(\frac{z_\infty^4 - 1}{z_\infty^4 + 1}\right)^2 \tag{2.25}$$

$$\left.\frac{T_c}{AM}\right|_{d=0} = \frac{1}{\ln(z_0)} \left(\frac{(z_0^4 - 1)}{z_0^4 + 1 + 2z_0^2}\right)^2 \tag{2.26}$$

Substituting the value of z_∞=2.97 and z_0=8.823 obtained numerically from 2.22, the saturation critical temperature is,

$$\frac{T_c|_{d\to+\infty}}{T_c|_{d=0}} = \frac{1.16516\ldots}{0.58258\ldots} = 2 \tag{2.27}$$

or

$$T_c\big|_{d\to+\infty} \approx 2\ T_c\big|_{d=0} \tag{2.28}$$

Thus Eq. 2.28 shows saturation of the critical temperature at $d \to \infty$ depends on the number of layers in a duplex network.

2.3.3 Generalization to Higher-Order Multiplex Networks

To explore multilayer effects, I extend the duplex network model discussed in Sec. 2.3.1 to a more general multilayer framework. The analytical method used to calculate the critical temperature in a duplex network could be extended to a higher-order multiplex network. But this requires extensive calculation of the Jacobian, which is computationally inefficient.

I consider a system comprising L layers, where each layer represents a distinct type of relationship or communication context among the same set of nodes. Given that different relationships between the same pair of agents are correlated, edges across different layers interact with their counterparts in other layers. For simplicity, I assume that the adjacency matrices across all layers are identical, i.e., $A^{(\alpha)} = A$ for all layers α, and that the interlayer coupling strength is uniform, denoted by K.

The total energy of the L interactions between a given pair of nodes (i, j) is given by:

$$h(\vec{x}_{ij}) = -A\left(\sum_{k}^{M_{ij}}\sum_{\alpha=1}^{L} x_{ij}^{(\alpha)} x_{jk}^{(\alpha)} x_{ki}^{(\alpha)}\right) - K\sum_{\alpha>\beta} x_{ij}^{(\alpha)} x_{ij}^{(\beta)}. \tag{2.29}$$

The methodology applied to the duplex network remains valid for an arbitrary number of layers. For instance, in the case of a three-layer ($L = 3$) network, using Eq. (2.4) with $\vec{x}_{ij} = (x_{ij}^{(1)}, x_{ij}^{(2)}, x_{ij}^{(3)})$ and the energy function given in Eq. (2.29), I can derive a set of self-consistent equations for the vector order parameter $\vec{x} = (x^{(1)}, x^{(2)}, x^{(3)})$.

Applying the same approach as in the duplex network (see Sec.2.3.2), I obtain the following transcendental equation for the auxiliary variable $z = z(D) = e^{\frac{AM}{T_c}x_c^2}$:

$$8\ln(z) = \frac{(D^4z^6 + z^4 - z^2 - D^4)\,(D^4z^6 + 3z^4 + 3z^2 + D^4)}{z^2\,(D^4z^8 + 4D^4z^6 + 3D^8z^4 + 3z^4 + 4D^4z^2 + D^4)}. \tag{2.30}$$

The corresponding critical temperature for $L = 3$ is given by:

$$\frac{T_c}{AM} = \frac{1}{\ln z}\left(\frac{D^4\,(z^6-1) + z^2\,(z^2-1)}{D^4\,(z^6+1) + 3z^2\,(z^2+1)}\right)^2. \tag{2.31}$$

The dependence of this critical temperature on the interlayer coupling strength and the number of layers is illustrated in Figures 2.6 and 2.7, respectively.

While this approach remains theoretically valid for any number of layers L, the com-

plexity of the governing equations increases rapidly with L. To derive a general prediction for the critical temperature T_c beyond specific values of L, I assume statistical equivalence across all layers, allowing for a simplified analysis. Instead of tracking the full microstate $\vec{x}_{ij}$ for each pairwise interaction, I introduce a mesostate description based on the number of positive links L^+ among the L interactions. Given the statistical symmetry assumption, the precise configuration of positive and negative links within the L layers becomes irrelevant, meaning that the energy function depends solely on L^+.

For a given set of L links, where L^+ links are positive and the remaining $L^- = L - L^+$ links are negative, the mean polarization is defined as:

$$\langle x \rangle = \left\langle \frac{L^+ - L^-}{L} \right\rangle = \frac{2\langle L^+ \rangle}{L} - 1. \tag{2.32}$$

Since all layers are assumed to be statistically identical, the average polarization per layer, $\langle x \rangle$, is independent of the specific layer index α. To evaluate the statistical weight of different mesostates, I consider the number of microstates corresponding to a given value of L^+, given by the binomial coefficient:

$$\#\text{ microstates } = \binom{L}{L^+}. \tag{2.33}$$

The total energy associated with a given mesostate L^+ comprises contributions from both Heider balance interactions and interlayer Ising-like interactions. The Heider energy term is derived from Eq. (2.29), replacing explicit $x_{ij}^{(\alpha)}$ terms with their statistical representations:

$$E_{\text{Heider}}(L^+) = -A(N-2)\left(L^+\langle x \rangle^2 - (L - L^+)\langle x \rangle^2\right). \tag{2.34}$$

Care must be taken to distinguish between the macroscopic mean-field variable $\langle L^+ \rangle$ and the microscopic state variable L^+. This distinction is only neglected when computing the expected value of L^+ in the final self-consistency equation.

The Ising-like energy contribution accounts for interactions among all L links:

$$E_{\text{Ising}}(L^+) = -K\frac{L^+(L^+ - 1) + (L - L^+)(L - L^+ - 1)}{2} + KL^+(L - L^+). \tag{2.35}$$

The total energy is then given by:

$$E(L^+) = E_{\text{Heider}}(L^+) + E_{\text{Ising}}(L^+). \tag{2.36}$$

To determine the mean number of positive links, I employ a canonical ensemble approach, multiplying the factor for specific L^+ by the number of microstates such state actually represents. At this point, the mean value of microscopic L^+ is considered the same as the

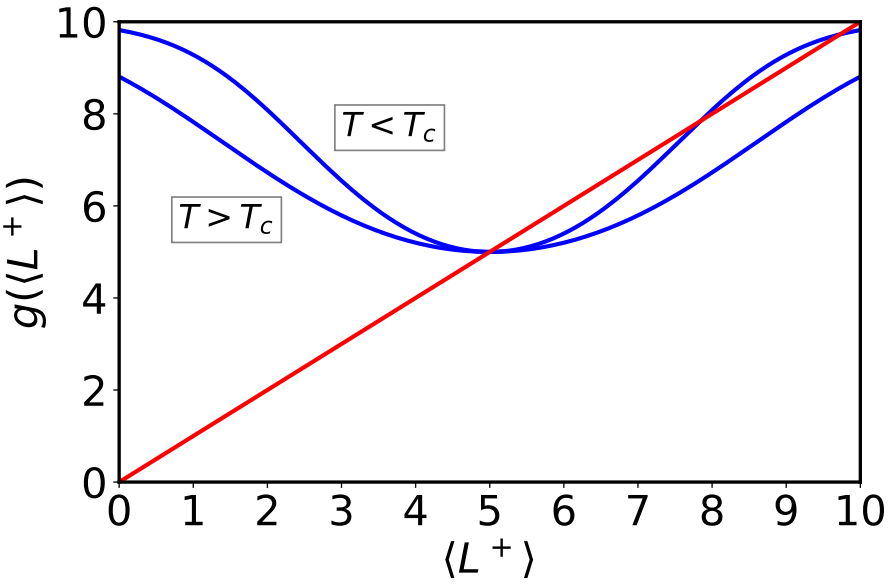


Figure 2.4: Graphical solution of the mean-field state equation (Eq. 2.37) for $L = 10$ layers. The solid blue curve is $y = g(\langle L^+ \rangle)$ and the red straight line is y=$\langle L^+ \rangle$. When $T > T_c$, there is a single stable fixed point. When $T < T_c$, there are three fixed points, and the inner one is unstable

mean-field variable $\langle L^+ \rangle$. This leads to the self-consistent equation:

$$\langle L^+ \rangle = g(\langle L^+ \rangle) = \frac{\sum_{L^+=0}^{L} L^+ \binom{L}{L^+} e^{-E(L^+)/T}}{\sum_{L^+=0}^{L} \binom{L}{L^+} e^{-E(L^+)/T}}. \tag{2.37}$$

Here $\langle L^+ \rangle = L$ is a paradise state, and $\langle L^+ \rangle = L/2$ is a completely disordered state with half the links positive and half negative. Due to using a mesostate description, instead of L equations, I have a simple scalar equation (Eq. 2.37) that could be solved using a similar stability analysis approach. Figure 2.4 shows a graphical representation of the right-hand side of Eq.(2.37), which solution is given by the intersection of the straight line L^+ and right-hand side. For $T > T_c$, there is only one stable fixed point, $L^+ = L/2$, which corresponds to a disordered state of the system, while for $T < T_c$, an additional stable fixed point appears, corresponding to an ordered state, as well as an unstable fixed point separating basins of attraction for each stable solution. When $T = T_c$, both additional points are the same, unstable fixed point tangential to the diagonal line. The critical temperature can be numerically estimated from the fixed point equation Eq.(2.37) and its derivative $dg(\langle L^+ \rangle)/d\langle L^+ \rangle = 1$.
The L-coupled layers undergo a first-order transition at T_c for any number of layers, which has been confirmed numerically for up to 10 layers using Eq.(2.37).

2.3.4 Saturation of critical temperature

The normalized critical temperature, $\frac{T_c}{AM}$, depends on d, the relative coupling strength, and increases as d increases, eventually saturating at high d values (see Fig. 2.6).
In the strong coupling limit, $D \to +\infty$, Eqs. 2.12 and 2.13 simplify to

$$8 \ln z = \frac{(z^8 - 1)}{2z^4} \implies z \approx 1.72346\ldots \tag{2.38}$$

$$\frac{T_c}{AM} = \left(\frac{1}{\ln z}\right)\left(\frac{(z^4 - 1)}{(z^4 + 1)}\right)^2 \implies \frac{T_c}{AM} \approx 1.16516\ldots \tag{2.39}$$

For comparison, in the single-layer case, the critical temperature is $T_c/AM \approx 0.58258$, which is consistent with previous results [55]. This leads to the following relationship:

$$T_c\big|_{L=2,d\to+\infty} \approx 2 \cdot T_c\big|_{L=1} \tag{2.40}$$

To generalize this result, I extend the analysis to a system with three layers ($L = 3$). In the limit $D \to +\infty$, Eqs. 2.30 and 2.31 reduce to

$$8 \ln z = \frac{(z^{12} - 1)}{3z^6} \implies z \approx 1.43749\ldots \tag{2.41}$$

$$\frac{T_c}{AM} = \left(\frac{1}{\ln z}\right)\left(\frac{(z^6 - 1)}{(z^6 + 1)}\right)^2 \implies \frac{T_c}{AM} \approx 1.74516\ldots \tag{2.42}$$

Thus, the saturation critical temperature for a three-layer system satisfies

$$T_c\big|_{\alpha=3,d\to+\infty} \approx 3 \cdot T_c\big|_{\alpha=1} \tag{2.43}$$

For a general system with L coupled layers, the transcendental equation in the strong coupling limit ($D \to +\infty$) takes the form

$$8 \ln z = \frac{\left(z^{4L} - 1\right)}{Lz^{2L}} \tag{2.44}$$

$$\frac{T_c}{AM} = \left(\frac{1}{\ln z}\right)\left(\frac{(z^{2L} - 1)}{(z^{2L} + 1)}\right) \tag{2.45}$$

This result confirms the general scaling relation for the saturation critical temperature:

$$T_c\big|_{\alpha=L,K\to+\infty} \approx L \cdot T_c\big|_{\alpha=1} \tag{2.46}$$

Thus, for sufficiently large coupling strength, the saturation critical temperature increases proportionally with the number of layers L.

2.4 Numerical results

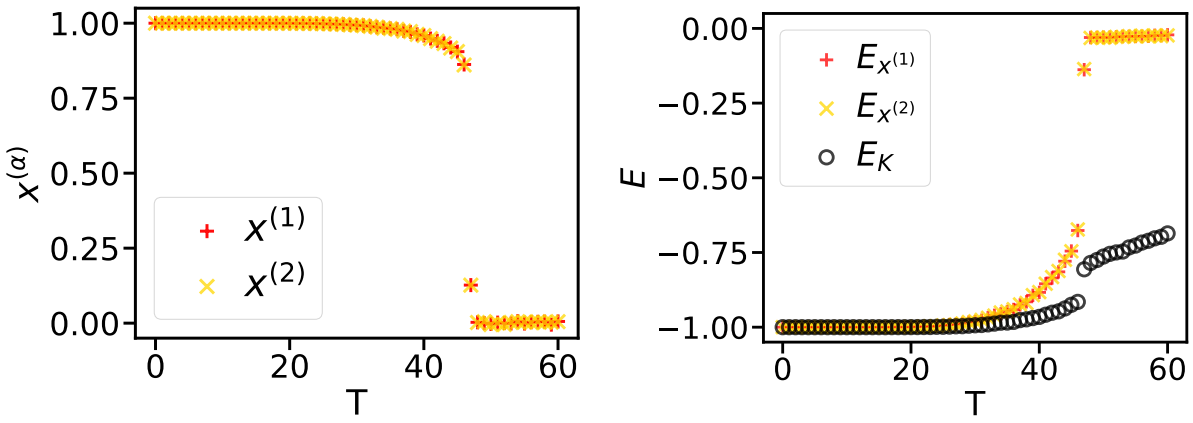
2.4.1 Numerical Simulations of paradise state stability

To validate the analytical results presented in the previous section, numerical MC simulations were carried out for the model. The heat bath algorithm was employed, where each update modifies the entire set of links between a randomly chosen pair of agents, i and j, with updates occurring asynchronously. A single time step consists of $N(N-1)/2$ elementary updates. During each update, a pair of agents ij is selected uniformly at random, and the link signs of all corresponding parallel links, $\vec{x}_{ij}$, are updated to one of 2^L possible states. The probability of selecting a given state follows the canonical ensemble distribution given by $P(\vec{x}_{ij}) = \frac{1}{Z} \exp\left(-\frac{h(\vec{x}_{ij})}{k_B T}\right)$. The energy $h(\vec{x}_{ij})$ is computed based on the current configuration of the entire system, meaning all links $\vec{x}_{kl}$ where $kl \neq ij$, following Eq. 2.29.

The simulations are initialized with a fully connected graph in the *paradise* state, where all links are positive. The heat bath algorithm was preferred over the Metropolis algorithm due to the latter's extremely slow convergence when applied to systems with a larger number of layers and strong inter-layer coupling. Notably, for $L = 2$, both algorithms produced consistent results, with any differences attributable to stochastic fluctuations inherent to MC methods.

In a fully connected network, each pair of nodes has $M = N - 2$ pairwise neighbors, and I set $A^{(1)} = A^{(2)} = 1$. The intralayer energy was rescaled relative to the average Heider energy per triad, while the interlayer energy was normalized to the average energy per interacting link pair. Simulations were performed across a range of temperatures, starting from $T = 0$ and increasing in increments of ΔT, to determine the highest temperature at which $\langle x_{ij}^{(\alpha)} \rangle$ remains positive. The true critical temperature T_c lies within the interval $[T^*, T^*\Delta T]$, and is estimated as $T_c = T^* + \frac{\Delta T}{2}$. Starting from the paradise state, the system transitions to a disordered state for temperatures above T_c. As shown in Figure 2.5, both layers exhibit similar, synchronized behavior. Above the critical temperature, the intralayer energies $E_{x^{(1)}}, E_{x^{(2)}}$ approach zero, while the mean polarization values fluctuate around zero. The interlayer energy E_K remains negative both below and above T_c, indicating that link signs across different layers retain partial alignment even in the disordered phase.

The critical temperature T_c obtained from MC simulations is compared with the predictions from the analytically evaluated mean-field theory, as shown in Fig. 2.6. While there is a slight discrepancy in the transition temperature between the numerical simulations and the mean-field approximation, the analytical approach successfully captures the overall behavior of the system. The dependence of T_c on the number of layers L is further explored in Fig. 2.7, which demonstrates that the simulation results align with



(a) (b)

Figure 2.5: Mean link polarization $x^{(1)} = \langle x_{ij}^{(1)} \rangle$ and $x^{(2)} = \langle x_{ij}^{(2)} \rangle$ (2.5a) and corresponding energies (2.5b) for a duplex network as a function of temperature T for the complete graphs with $N = 50$ nodes and a coupling strength $K = 50$ (corresponding to d = 1). The results show averages over 50 independent simulations. Initial conditions were paradise states at both layers, where the overlapping red plus and yellow cross symbols correspond to layers one and two, respectively. In contrast to the intralayer energies (overlapping red plus and yellow cross), the interlayer energy (black circle) increases slowly when $T > T_c$, with layers remaining partially synchronized even when intralayer disorder sets in.

Eq. 2.46 in the saturation limit. However, some discrepancies arise for large values of L. According to the analytical predictions, as L increases, the critical temperature should exhibit a nonlinear increase, eventually saturating at a fixed value T_c, independent of K. Even for small K, the critical temperature is expected to exhibit this behavior at sufficiently large L, which is particularly evident in Fig. 2.7 for $K = 10$ and $K = 20$.

Contrary to these predictions, MC simulations reveal that such a transition does not occur. Instead, the critical temperature increases linearly with L for all values of K, with a slope that depends on K. This indicates that the mean-field approach does not fully capture the behavior of the system when multiple weakly interacting layers are present. Nevertheless, the mean-field theory remains qualitatively accurate. Specifically, it correctly predicts a linear increase of T_c with L for small K and L, as well as the saturation of T_c at high K.

2.4.2 Critical behavior of the system with layers in different states

The critical temperature, as analyzed in prior studies, is defined as the threshold at which the nonzero stable state vanishes. Analytical treatments predominantly address the existence of the paradise state, evaluating its stability and the transition to disorder. However, due to the system's inherent multistability, within the parameter regime where the paradise state remains stable, additional equilibrium states may also materialize. The

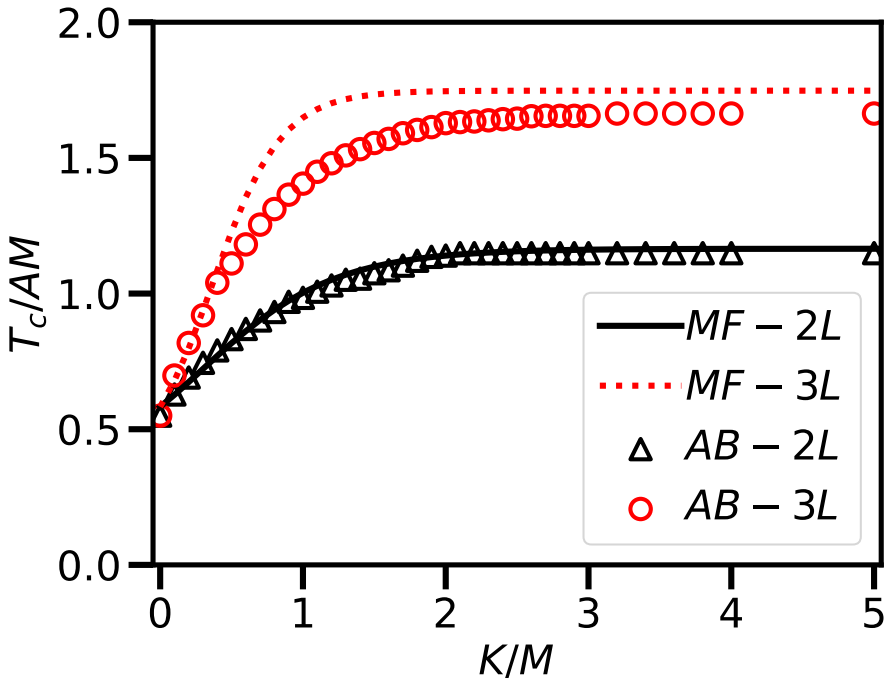


Figure 2.6: (Color online) Critical temperature T_c increases with coupling strength K and saturates to $T \approx L \cdot T_{c,L=1}$ for high K/M values. The solid black line and black triangles show mean-field prediction (Eq. 2.13) and numerical results for $L = 2$, while the dotted red line and red circles show mean-field prediction (Eq. 2.31), and numerical results for $L = 3$.

system's final configuration is contingent on initial conditions, potentially converging to distinct stable states.

This section extends the investigation to multiplex networks, with a particular emphasis on scenarios where the two layers adopt different configurations. Within the framework of HB dynamics, multiple fully balanced states exist, which can broadly be categorized as a "two-clique state." This state is defined by a bipartition of the vertices into two groups, wherein intra-group links are exclusively positive, while inter-group links are uniformly negative. The specific realization of the two-clique state varies depending on the relative sizes of the two groups, spanning configurations from an equal division to an extreme case where one group encompasses all vertices and the other is empty, which is termed the "paradise state." Among all possible balanced states, a near-equally partitioned two-clique structure is the most probable outcome of HB dynamics, as evidenced in [88]. Conversely, the paradise state, though theoretically significant, is exceedingly improbable as it represents an idealized state where all agents maintain positive relationships, akin to consensus in opinion dynamics.

For the purposes of this study, I constrain the analysis to three equilibrium states: the two-clique state with equally sized groups, the paradise state, and the disordered state. Other balanced states characterized by asymmetric group sizes are not considered. In the

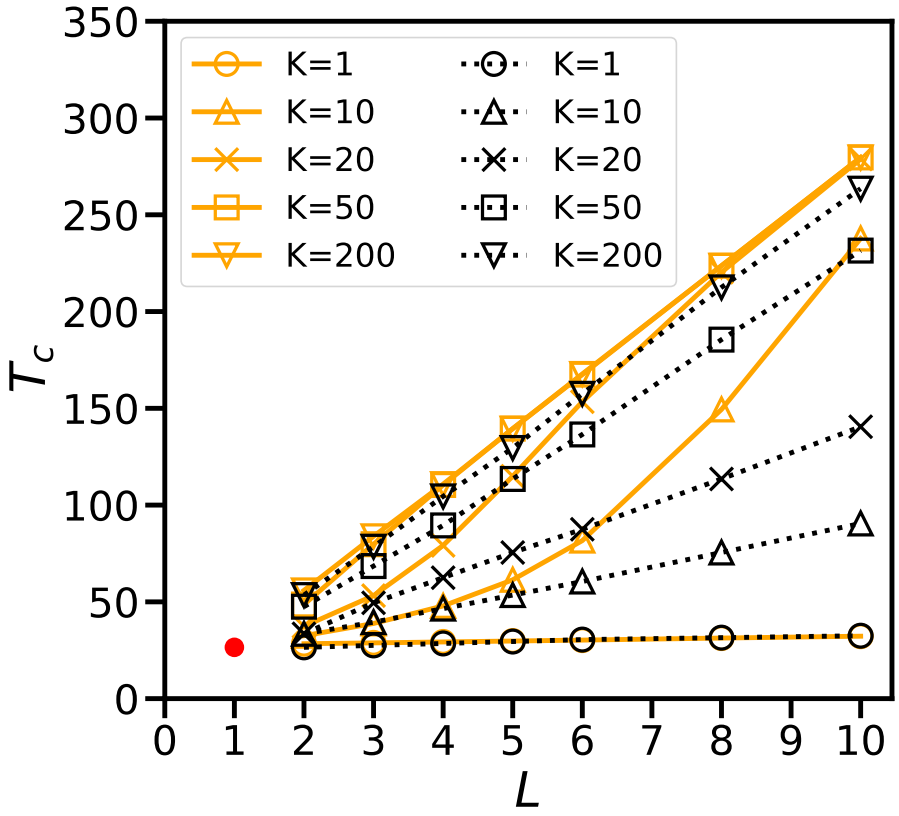


Figure 2.7: The critical temperature T_c increases with the number of layers L when coupling K is positive. Black symbols and broken lines show results of MC simulations, while orange symbols and solid lines show mean-field prediction (iterating Eq. 2.37 as a map) at different interlayer coupling strengths for $N = 50$ and $A = 1$. The critical temperature for a single layer is shown as a red full circle at $T_c = 26.5$.

context of a duplex network, the equilibrium states that satisfy intralayer HB conditions (excluding asymmetric two-clique structures) include three principal configurations: (i) both layers adopting the paradise state, (ii) one layer in the paradise state while the other forms a two-clique structure, and (iii) both layers assuming two-clique states. When both layers are in the paradise state, all links are positive, ensuring complete alignment of edge signs across layers. By contrast, when one layer assumes the paradise state and the other a two-clique configuration, approximately half of the inter-group links in the two-clique state are negative, while their corresponding edges in the paradise layer remain positive. This results in an interlayer energy that approximates to zero.

A particularly compelling scenario arises when both layers exhibit two-clique states, as the degree of correlation between the cliques across layers can vary. The alignment between the two-clique structures may range from complete congruence (“matching order”) to partial correlation (“partial matching order”) and, at the extreme, to minimal overlap (“mismatched order”), as depicted in Fig. 2.8.

To quantify the structural similarity between two-clique states in different layers, I introduce the ”overlap” metric, denoted as o. This metric represents the fraction of vertices

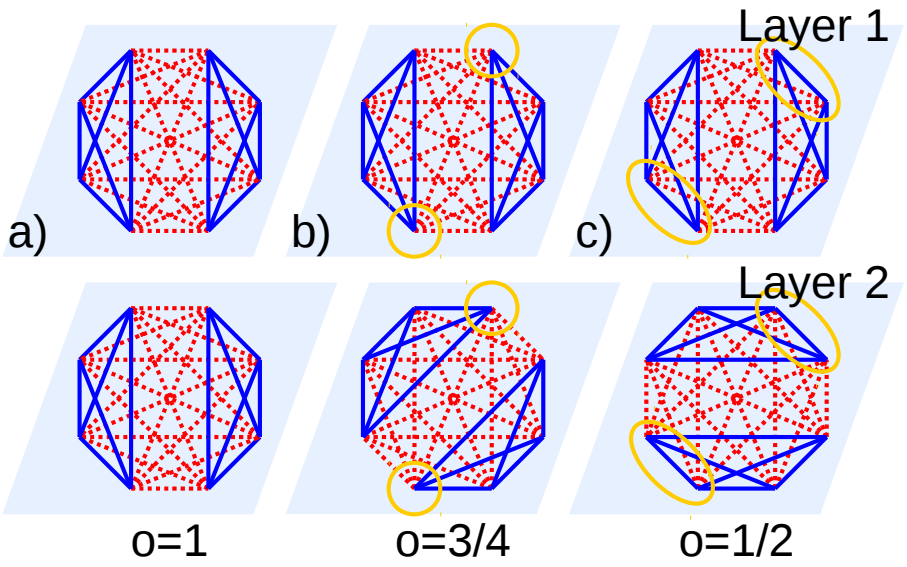


Figure 2.8: If both layers of a duplex are in two-clique states, the overlap between the layers can vary from full overlap (panel a, matching order, overlap $o = 1$), through partial overlap (panel b, partial matching order, overlap $o = 3/4$ in this example) to minimal overlap displayed by "orthogonal" two-clique states (panel c, mismatched order, overlap $o = 1/2$). The same nodes exist in both layers, and interlayer interactions exist only between the same link replicas with coupling strength K. Solid blue lines represent positive links, dotted red lines represent negative links, and yellow circles enclose agents belonging to different groups in both layers.

assigned to the same groups across both layers, effectively capturing the degree of consistency in clique assignments. Since cliques themselves lack intrinsic identities, I define overlap to always reflect the maximal alignment, thereby constraining its values to the range $[\frac{1}{2}, 1]$. This overlap metric directly correlates with the Spearman correlation between an agent's group assignments in both layers, expressed as $c = 2o - 1$. Additionally, it influences the interlayer energy of fully Heider-balanced states (neglecting fluctuations), given by $E_{inter} = \frac{KN}{2}(1 - N(1 - 2o)^2)$, where N is the number of agents and K denotes the interlayer coupling strength introduced in Sec. 2.2. Due to the discrete nature of agent assignments, the set of possible overlap values o for a given network size N is restricted to discrete points, yet these values always fall within the theoretical bounds of $[\frac{1}{2}, 1]$. In a duplex system at equilibrium, the possible configurations depend on temperature and interaction dynamics. The system can adopt a *matching order* state, where both layers either reside in the paradise state or exhibit identical two-clique state with an overlap of $o = 1$. Alternatively, it may exist in a *partial matching order* state, where the two-clique state displays intermediate degrees of overlap, $o \in (\frac{1}{2}, 1)$. In contrast, a *mismatched order* state arises when two-clique states are minimally correlated, with $o = \frac{1}{2}$, or when one layer assumes the paradise state while the other forms a two-clique state. Finally, the system may reside in a *disordered* state, wherein neither layer achieves an internally balanced structure.

For a multiplex network with more than two layers, the diversity of possible states increases. Different pairs of layers can exhibit varying degrees of overlap, leading to configurations such as one matching pair alongside a non-matching third layer, three mutually mismatched layers, or a mixture of partial overlaps. Certain topological constraints regulate these configurations; for instance, if layers 1 and 2 are perfectly matched and layers 2 and 3 are matched, then layers 1 and 3 must also match. This transitivity condition prohibits inconsistencies in layer matching but still allows for multiple mutually mismatched layers, depending on the number of agents N. In the simplest case where $N = 2^n$, the system can accommodate up to $1 + \binom{N/2}{N/4}$ two-clique states with zero mutual overlap. The analytical framework outlined in Sec. 2.3.1 does not incorporate the presence of two-clique states, except for the trivial case of the paradise state. Consequently, numerical simulations presented in this study reveal deviations from mean-field predictions, particularly when two-clique states are involved. A significant discrepancy arises from the mean-field theory's assertion that the disordered state remains stable at all temperatures. However, numerical simulations demonstrate that under appropriate conditions, the system instead transitions to a Heider-balanced two-clique state, which, despite its structured nature, exhibits zero polarization similar to the disordered state. Given that the analytical model neglects the existence of two-clique states, it is unable to capture the critical phenomena associated with them. In particular, the transition temperatures T_s, T_d, T_o, and T_m, which are introduced in subsequent sections, fall outside the predictive capability of the mean-field approach.

To systematically examine the impact of initial conditions on system evolution, I conduct numerical investigations for five scenarios in which layers are initially mismatched:

- One layer is initialized in the paradise state, while the other adopts a two-clique balanced configuration.
- Both layers initialized in mutually mismatched two-clique balanced states.
- Both layers are initialized in two-clique states with partial overlap.
- A three-layer network where one layer starts in the paradise state while the remaining two are in perfectly matching two-clique states.
- A three-layer network where one layer is in the paradise state, and the other two are in mismatched two-clique states.

By analyzing the steady-state behavior emerging from these initial conditions, alongside simulations of initially disordered configurations, I characterize the stability and transitions between equilibrium states. These findings enable the construction of a comprehensive phase diagram that delineates the parameter regimes in which different system states emerge.

2.4.3 Coupling between a paradise and two-clique state

In a multiplex network, a mismatched order state can emerge when one layer assumes a paradise configuration while the other forms a two-clique structure. Consider a system of size $N = 2m$, where each antagonistic group consists of m nodes. Each node maintains $(m-1)$ positive intragroup connections and m negative intergroup connections, maintaining balance in the network. The system's macroscopic behavior is dictated by temperature and interlayer interaction strength. Numerical simulations, as illustrated in Fig. 2.9, reveal the presence of four out of the five characteristic or critical temperatures that govern the system's phase transitions and stability dynamics.

The temperature T_s defines the threshold above which the mismatched order state becomes unstable. For $T > T_s$, the initially mismatched layers synchronize, transitioning into a matching order state. However, for $T > T_o$, the mismatched layers may instead evolve into a disordered state rather than achieving a matching configuration. At higher interaction strengths K, T_o converges with T_d (Fig. 2.9c), which represents the minimum temperature at which a disordered state remains stable. Conversely, as $K \to 0$, these temperatures diverge: T_o asymptotically approaches the single-layer critical temperature T_c, where order loses stability, whereas T_d aligns with the single-layer critical temperature T_d, where disorder becomes stable, as depicted in Fig. 2.14. Above T_o, the system's final state—whether it settles into a matching order state or a disordered state—becomes probabilistic, influenced by temperature T and layer overlap. The temperature T_m serves as the upper bound beyond which two mismatched layers can no longer evolve into a matching order state. For $T > T_m$, mismatched layers inevitably transition into disorder. Notably, $T_m < T_c$, meaning that an initially mismatched order destabilizes at a lower temperature than an initially matching order, which remains stable up to T_c (the only temperature not illustrated in Fig. 2.9). I classify T_s and T_d as critical temperatures, as they delineate stability transitions. In contrast, T_o and T_m are not associated with stability shifts and are therefore referred to as characteristic temperatures. The synchronization process of an initially mismatched order state exhibits intrinsic stochasticity. In the specific scenario of a mismatched paradise and two-clique state, the final configuration may either be a matching paradise or a matching two-clique state. Furthermore, the resulting two-clique state need not replicate the initial configuration of the original two-clique layer; the clique sizes and their spatial arrangement may differ, indicating that both layers may undergo structural reconfiguration rather than one layer imposing its structure on the other. The mean link polarization of a two-clique state depends on the clique sizes m_1 and m_2:

$$\langle x_{ij} \rangle = \frac{\binom{m_1}{2} + \binom{m_2}{2} - m_1 \cdot m_2}{\binom{N}{2}} \tag{2.47}$$

For significantly asymmetric clique sizes $m_1 \gg m_2$, $\langle x_{ij} \rangle$ approaches 1, whereas for ap-

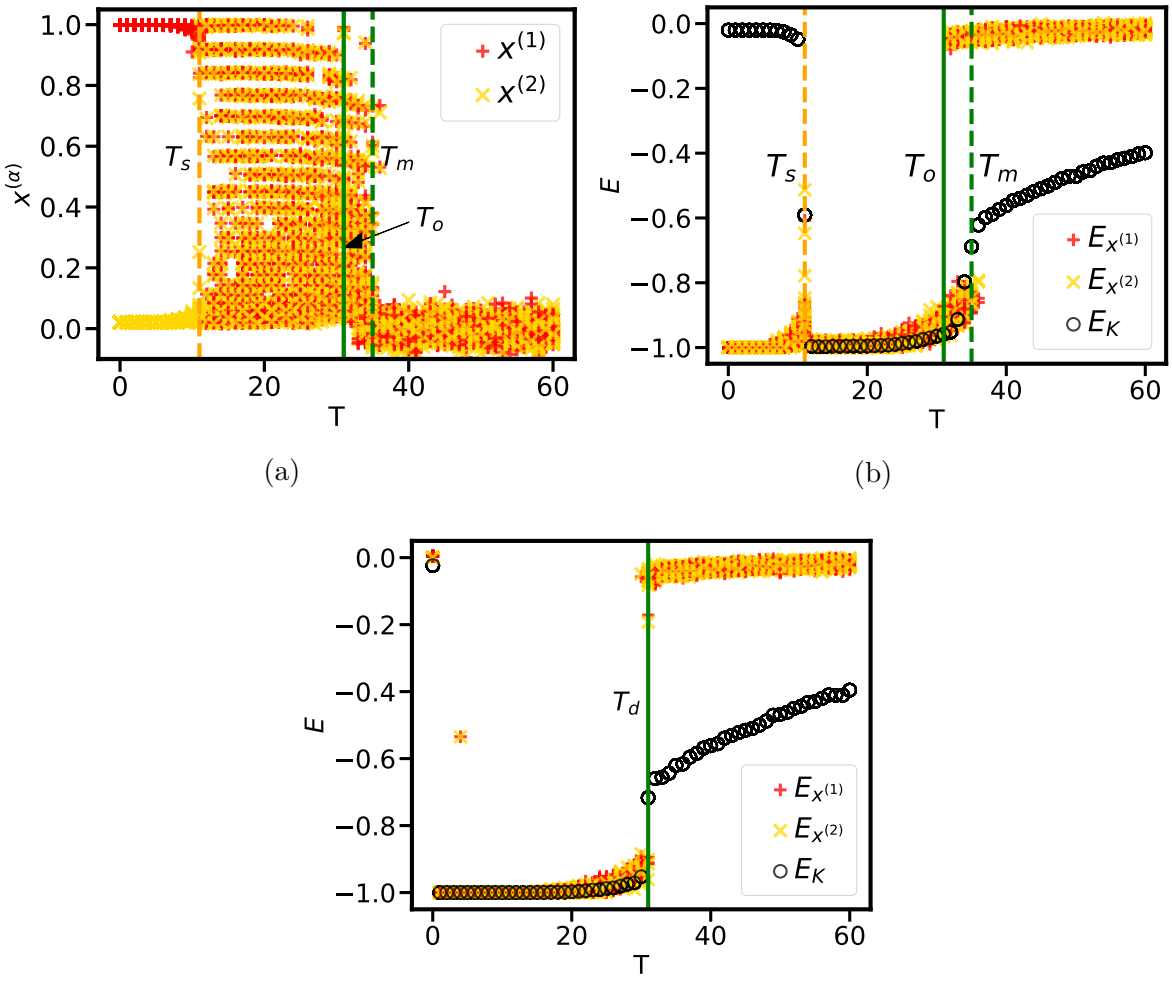


Figure 2.9: In the case of coupled paradise and two-clique states, the system shows two transitions: synchronization of layers starting at T_s, transition to disorder starting at T_o, and ending at additional characteristic temperature T_m. Panel (a) shows the mean link polarization of 50 independent simulations over a range of temperatures for $N = 50$ nodes and a coupling strength $K = 25$, with one layer starting in the paradise state (red plus) and the second layer starting in the two-clique state (yellow cross). Panel (b) shows the corresponding mean intralayer energy (red and yellow) and interlayer energy (black circle). Panel (c) shows intralayer and interlayer energy for a duplex network *starting in a disordered state*, showing that spontaneous order appears at any temperature below T_d.

proximately equal clique sizes $m_1 \approx m_2$, it nears zero. A quantitative analysis was conducted to assess the frequency f of mean link polarization across both network layers following synchronization. The distribution of mean link polarization is presented in Fig. 2.10. Although the paradise state $\langle x_{ij} \rangle = 1$ remains the most probable outcome following

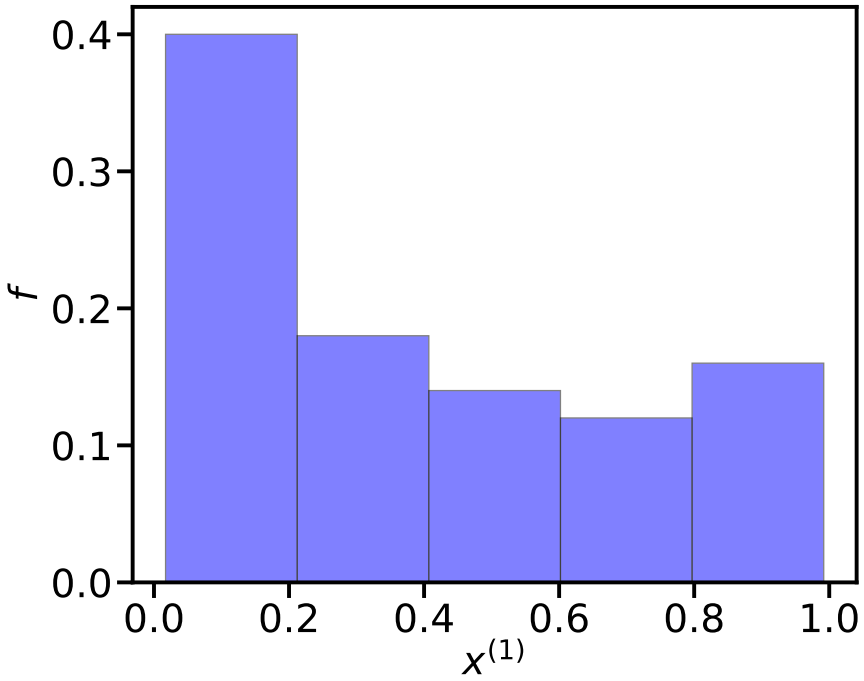


Figure 2.10: When paradise and two-clique states synchronize, it results most often in either paradise $x^{(1)} \approx 1$ but two-clique states $x^{(1)} < 1$ are also a possible outcome. The graph shows the distribution of mean links polarization after paradise and two-clique states synchronize at the temperature T where $T_s < T < T_d$.

synchronization, the formation of two-clique states is also a plausible result.

To systematically investigate the dependence of the critical temperatures T_d and T_s, as well as the characteristic temperatures T_o and T_m, on model parameters, I performed simulations across a range of interlayer coupling strengths K. The results of these simulations have been synthesized into the phase diagram of the two-layered system, provided in Sec. 2.4.6.

Notably, even when the system evolves into a disordered state, partial synchronization is retained due to interlayer Ising interactions. This effect is analogous to paramagnetic partial ordering induced by an external field, where, in this case, one layer serves as the external field for the other. Consequently, the interlayer energy $E(K)$ in Fig. 2.9 only asymptotically approaches zero as temperature increases, in contrast to the intralayer energy, which reflects ordering within individual layers.

An important aspect of synchronization dynamics is the temporal evolution of link polarization between the two layers. Figure 2.11 presents representative trajectories of the mean link polarization during relaxation processes. Beginning from an initial configu-

ration in which one layer is in the paradise state while the other adopts a two-clique structure with equally sized cliques, the system synchronizes into one of three possible matching order states: a paradise state, a two-clique state with equally sized cliques, or a two-clique state with cliques of unequal sizes (see also Fig. 2.11).

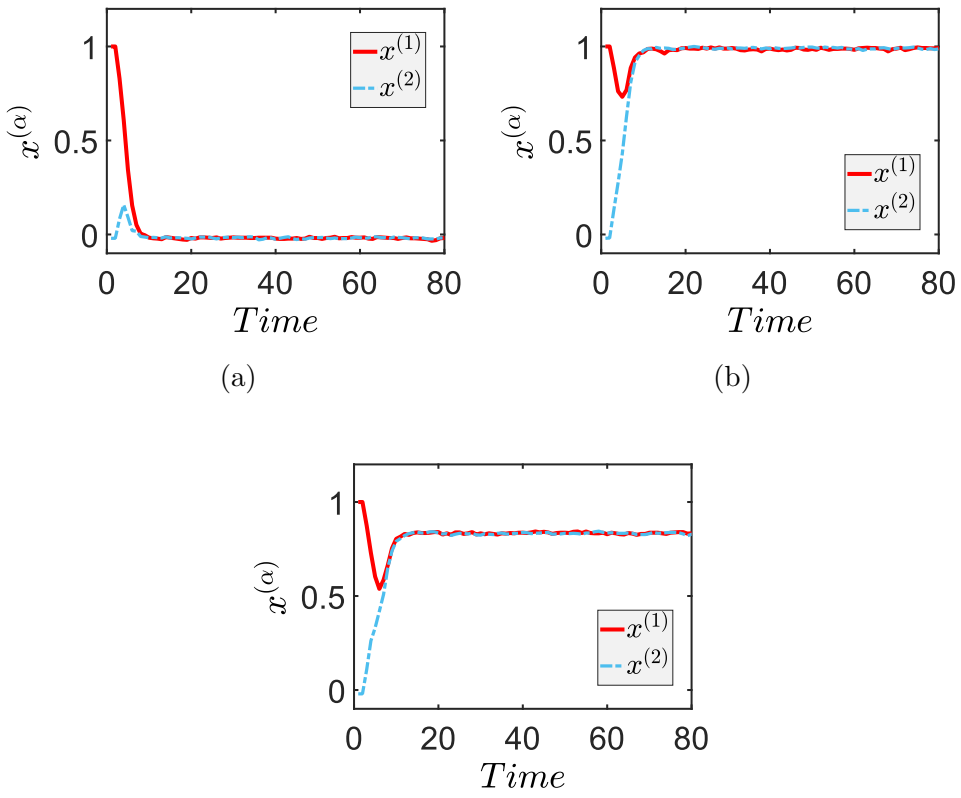


Figure 2.11: The final state of the synchronization between two mismatched layers is partially random due to thermal fluctuations. Example trajectories for the synchronization scenario between paradise (red line) and two-clique (blue dashed) layers for N=50, K = 25, and T=20. The synchronized state can be either a two-clique state similar to the initial condition of the second layer (panel (a)), a paradise state (panel (b)), or a two-clique state of different sizes, matching neither of the two initial layer states (panel (c)). Each shown scenario is a result of independent agent-based simulation under the same initial conditions and parameters.

2.4.4 Coupling between different two-clique states

In the case where both layers adopt a two-clique configuration, the extent of overlap between the layers can vary between 1 (indicating a perfectly matching order) and 1/2 (corresponding to a fully mismatched order). For a fully matched configuration, the system's behavior is analogous to that of the initial paradise-paradise state. However, unlike the paradise configuration, the mean link polarization does not necessarily reach 1; rather, it depends on the relative sizes of the cliques, approaching approximately 0 when

the cliques are of equal size. The critical temperature governing the transition to disorder remains T_c, consistent with that of the matching paradise state.
In contrast, when the two-clique structures in the layers are mismatched, with an overlap close to 1/2, the system dynamics resemble those observed in the mismatched paradise–two-clique scenario discussed earlier. As in that case, the system is characterized by two critical temperatures: T_s, which marks synchronization, and T_d, which signifies the onset of disorder. Additionally, the characteristic temperatures T_o and T_m persist, where T_o denotes the temperature above which the system may become disordered, and T_m represents the threshold beyond which disorder becomes the inevitable outcome instead of synchronization, as illustrated in Fig. 2.12. While the specific values of T_s, T_o, and T_m depend on the particular configuration, the temperatures T_d and T_c remain fixed, as they are intrinsically tied to the stability of the disordered and matching order states, being determined solely by the model parameters.

For configurations in which the two-clique states exhibit partial alignment, the qualitative behavior remains unchanged, though the exact values of T_s, T_o, and T_m depend on the degree of overlap o between the layers. In the limiting case where $o \to 1$, the synchronization temperature T_s approaches zero, while both T_o and T_m asymptotically converge to T_c.

2.4.5 Triplex network with mismatched layer pairs

As the number of layers in the network increases beyond two, the range of possible system states expands accordingly. In this analysis, I consider a triplex network in which one layer is in the paradise state, while the remaining two layers exhibit two-clique structures. These two-clique layers may either be perfectly aligned, corresponding to an overlap of $o = 0$, or mismatched with an overlap of $o = 13/25$, consistent with the configuration examined in the preceding subsection. In the first scenario, two of the layers form a matching configuration while remaining mismatched with the third layer. In the second scenario, all three layers are mutually mismatched. While the overall system dynamics retain fundamental similarities to those observed in the duplex network, the introduction of an additional layer introduces notable distinctions arising from the increased structural and interactional complexity of the system.

The most notable distinction between the two considered configurations lies in their respective synchronization temperatures, denoted as $T_s^{(pcc)}$ and $T_s^{(pcc')}$. Each configuration exhibits a distinct synchronization threshold, as evidenced by the simulation results presented in Fig. 2.13. This difference is expected, given that in the scenario where two layers are in a matching state, these layers are inherently more stable and exert a stronger, coherent influence on the third, mismatched layer. Consequently, synchronization occurs at a lower temperature $T_s^{(pcc)}$. Conversely, in the case where all three

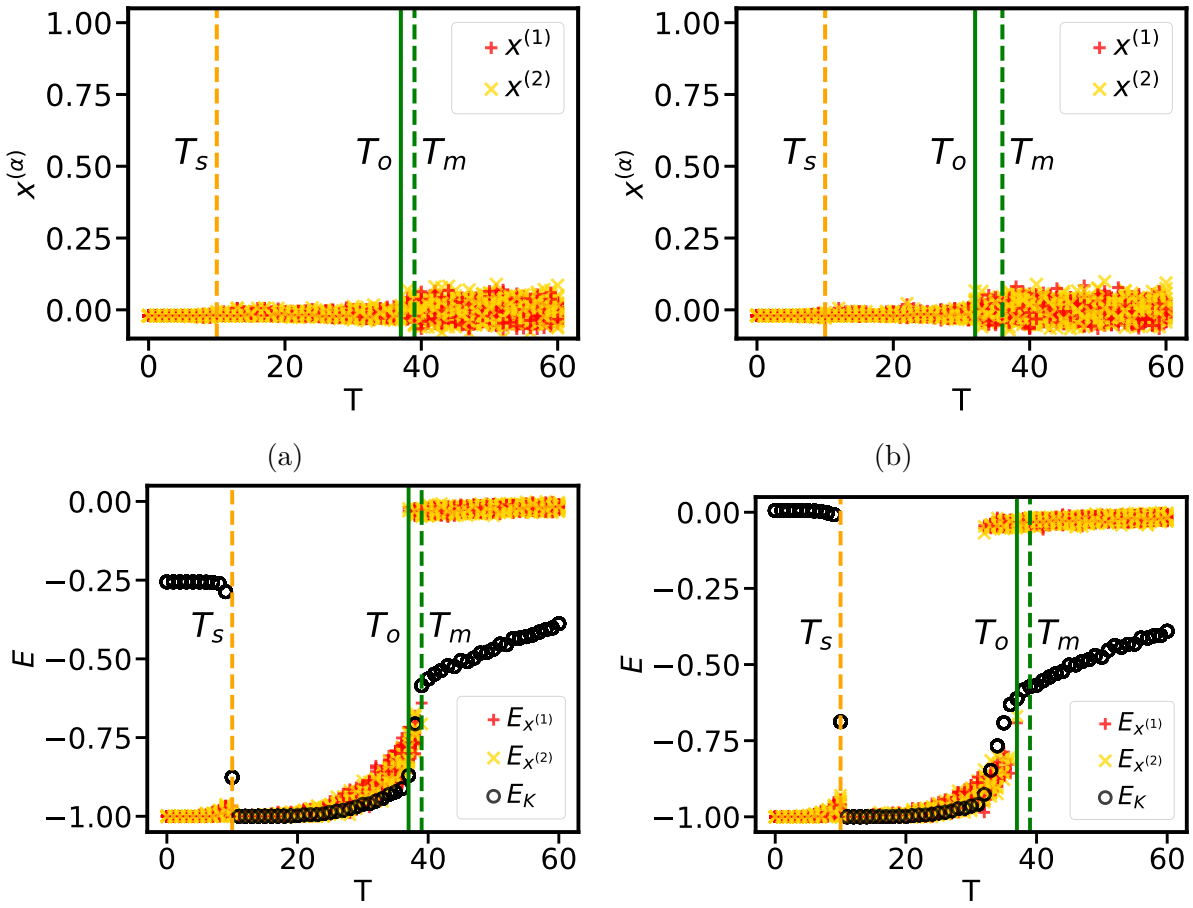


Figure 2.12: The behavior of the system of two partially matching (a, c) and mismatched (b, d) layers in two-clique states for $N = 50$ and $K = 25$, closely resembles the mismatched order of paradise and two-clique layers seen in Fig. 2.9. The top panels (a, b) show the variation of mean link polarization for two orthogonal two-clique layers over temperature for 50 independent simulations. The bottom panels (c, d) show the corresponding variation of mean intralayer (red plus and yellow cross) and interlayer (black circle) energy over temperature.

layers are mutually mismatched, no single layer experiences a dominant, coherent influence. As a result, individual layer configurations remain stable up to a higher transition temperature $T_s^{(pcc')}$. A further distinction between these two scenarios is the nature of the transition itself. In the mutually mismatched configuration, the system retains a degree of symmetry, leading to a transition behavior analogous to that observed in the two-layer mismatched state. Since none of the layers dominate the others, the system can evolve into states with varying clique sizes, which may not necessarily align with any of the initial configurations—though such deviations occur less frequently than in a duplex network. Consequently, as the number of layers increases, so does the number of possible stable states, leading to a corresponding increase in the diversity of state-specific transition temperatures, including T_s, T_o, and T_m.

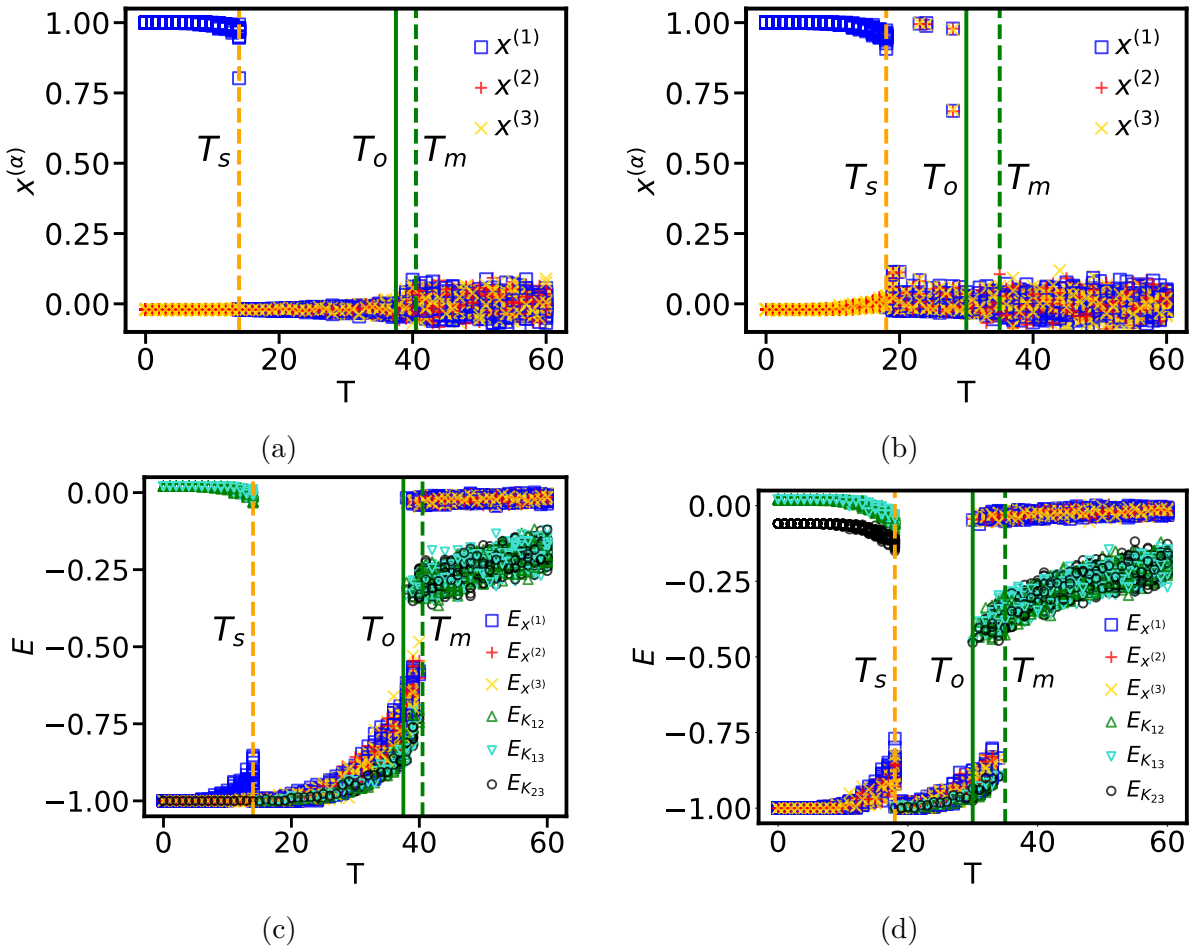


Figure 2.13: In a triplex network, there are two classes of states – two matching layers and a third mismatched (PCC) (a,c), or all three mutually mismatched layers (PCC') (b,d). These two state classes show different synchronization temperatures T_s. The top panels (a,b) show the mean polarization (blue square, red plus, and yellow cross) of each of the three layers, while the bottom panels (c,d) show intralayer energies (blue square, red plus, and yellow cross) as well as interlayer energy (green triangle, inverted triangle and black circle) tied to each pair of layers for $N = 50$ and $K = 10$. The PCC state consists of the paradise state in one layer and matching two-clique states in the other two, the PCC' state consists of paradise in one layer and mismatched two-clique states in the other two.

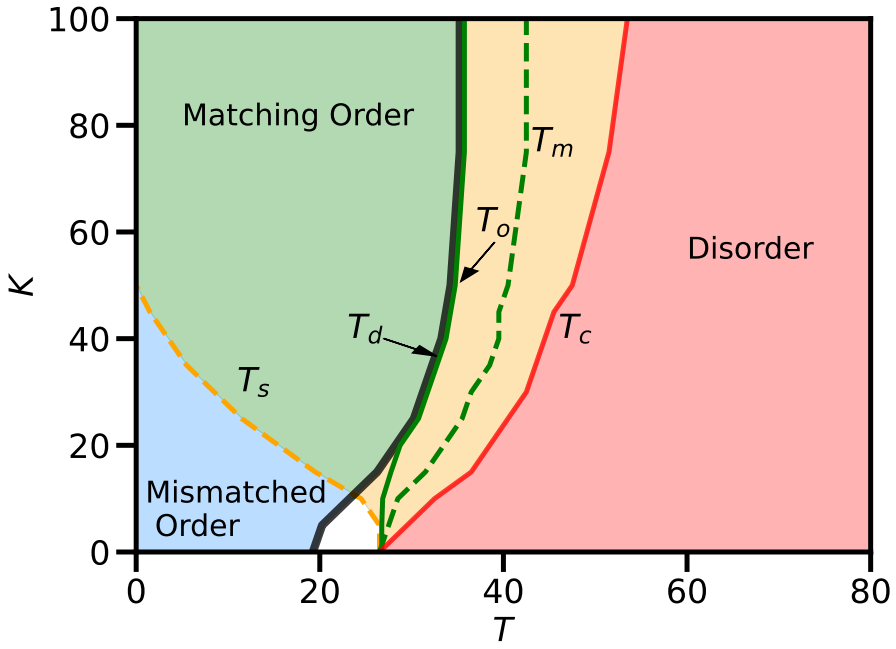


Figure 2.14: Phase diagram of the duplex system featuring multistability. The five colored areas are: blue (below T_s) – a region where both mismatched and matching order states are stable, green (between T_s and T_d) – a region where only matching order state is stable, yellow (between T_d and T_c) – a region where both matching order and disordered states are stable, red (above T_c) – a region where only disordered state is stable, white (between T_d and T_s) – a region where mismatched order, matching order and disordered state are all stable. The boundaries T_d and T_c are fixed, but T_s, T_o and T_m vary for different states. The results were obtained from agent-based simulations for $N = 50$ and $A = B = 1$ with 50 realizations.

2.4.6 Phase diagram

The findings presented in the previous subsections, along with additional results obtained for a range of parameter K, enable us to construct an experimental phase diagram for the system. This analysis allows us to identify the possible system states, determine their stability conditions, and characterize the transitions between them. Figure 2.14 illustrates the phase diagram as a function of temperature and interlayer coupling strength.

The phase diagram delineates five principal parameter regions. In the blue region, corresponding to temperatures below T_s, the system can exist in either a matching or mismatched order state. The green region, spanning T_s to T_d, permits only the matching order state. In the yellow region, between T_d and T_c, the system may exhibit either a matching order or a disordered state. The red region, representing temperatures above T_c, exclusively allows a disordered state. Lastly, the white region, situated between T_d and T_s, allows for the coexistence of mismatched order, matching order, and disorder. A crucial observation is that while the boundaries defined by T_d and T_c mark the lower stability limit of disorder and the upper stability limit of matching order, respectively, these thresholds remain fixed. However, the boundary determined by T_o is not universally constant, as

its precise value depends on the overlap o between mismatched layers. Consequently, T_o varies with the degree of overlap. When considering a partially matching order state, the system follows a qualitatively similar behavior, where T_s decreases with increasing overlap o, ultimately reaching $T_s = 0$ for fully matching layers. The characteristic temperatures T_o and T_m further define transition thresholds: T_o represents the temperature above which a system that initially exists in a mismatched order state may transition into disorder, while T_m denotes the temperature beyond which disorder is inevitable. Similar to T_s, these characteristic temperatures are state-dependent and influenced by the overlap o between layers in the initial configuration. In Fig. 2.14, the depicted phase boundaries correspond to an initial mismatched order state consisting of one layer in the paradise state and the other in a two-clique state. As the overlap between layers increases, both T_o and T_m increase, ultimately converging at T_c when $o = 1$, at which point the layers are fully matching. For sufficiently strong interlayer coupling strength K, T_o, as shown for the mismatched order state, coincides with T_d, implying that the system can transition to disorder as soon as disorder becomes a stable configuration. However, for lower values of K, the mismatched order state persists beyond T_d, transitioning to disorder only when $T_o > T_d$.

While the phase diagram in Fig. 2.14 identifies the stability regions of different states under varying parameter values, it does not explicitly illustrate the transitions between these states. To analyze the possible transitions and further characterize T_o and T_m, I consider a fixed K and examine the stable states as a function of temperature, as depicted in Fig. 2.15. Figure 2.15 schematically illustrates the stability of distinct network states across varying temperature regimes. At lower temperatures, both mismatched and matching order states remain stable, characterized by intralayer ordering in either a two-clique configuration or a paradise state. Partially matching states exhibit behavior akin to mismatched states but undergo transitions at temperatures specific to their initial configurations, primarily dictated by the overlap o between layers. The synchronization temperature T_s delineates the upper stability limit for the mismatched order, while T_d defines the lower bound for disorder stability, and T_c signifies the upper stability threshold for the matching order state.

When the temperature of a system in a matching order state exceeds T_c, it undergoes a transition into a disordered state, as indicated by the solid gray arrow in the figure. Conversely, if the temperature of a disordered state drops below T_d, the system transitions into a matching order state, also represented by a solid gray arrow. Notably, in larger systems, the probability of a self-organized matching order leading to a paradise state is negligibly small. Consequently, transitions predominantly result in both layers adopting two-clique configurations.

For a system initially in a mismatched order state, the nature of its transition upon exceeding T_s is contingent upon the final temperature. If the temperature falls within

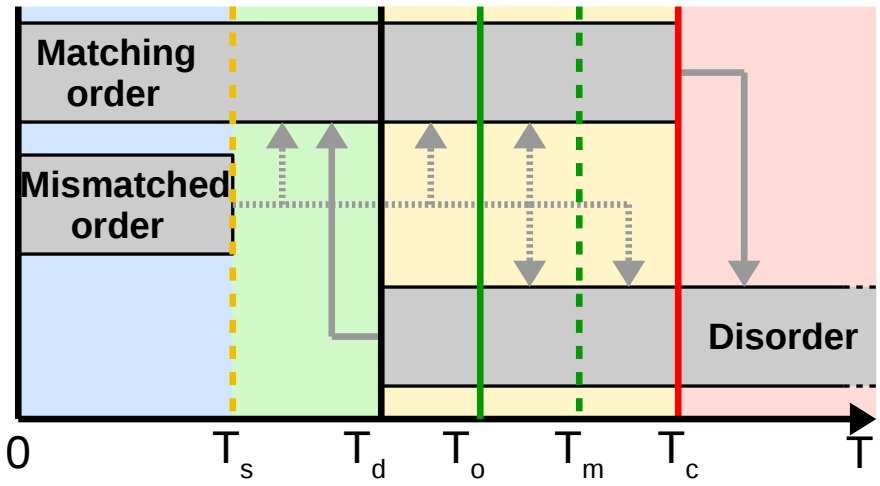


Figure 2.15: The Heider balance dynamics on a duplex network is multistable, with different states possible depending on the temperature. Grey boxes show a range of temperature where a given state is stable for a fixed K value, with grey arrows showing possible transitions between these states. Five characteristic temperatures can be distinguished: synchronization at T_s, the appearance of disorder at T_d, the lower limit of transitions to disorder T_o, the upper limit of transitions to matching order T_m, and the disappearance of order at T_c. Note that the figure is schematic and does not depict actual values for better clarity. The temperature values T_s depend on the exact states they correspond to, since partially matching states with lower overlap have lower temperatures T_s. Characteristic temperatures T_o and T_m depend on the initial state and its overlap between the layers, since higher overlap means higher temperature values, up to T_c for $o = 1$.

the range $T_s < T < T_o$, the system consistently evolves into a matching order state. Within the range $T_o < T < T_m$, the system may transition into either a matching order or a disordered state, subject to stochastic fluctuations. Beyond T_m, disorder becomes the inevitable outcome despite the continued stability of the matching order state up to $T_c > T_m$. This indicates that for temperatures within $T_m < T < T_c$, an already established matching order may persist, yet thermal fluctuations are too substantial for such a state to emerge spontaneously from a mismatched order configuration. It is essential to underscore that this analysis specifically pertains to a duplex network where $A^{(1)} = A^{(2)}$. If the Heider interaction strengths differ across layers, the resultant dynamics and phase transitions may become significantly more complex.

2.5 Discussion and Conclusions

The chapter discusses the extension of Heider's structural balance theory to multiplex networks, where agents have correlated relationships across multiple layers. The dynamics of these signed networks, considering intralayer interactions governed by Heider balance dynamics and interlayer interactions modeled as Ising coupling that promotes consistency between relations of the same pair of agents across different layers, were investigated.

Studying HB dynamics in multiplex networks highlights how social relationships, characterized by both friendly and hostile connections, influence network evolution. My findings indicate that the balance theory, traditionally applied to single-layer networks, extends meaningfully into the multiplex realm, enabling a deeper understanding of social dynamics where individuals maintain multiple relationships across varied contexts.

Through agent-based simulations and an analytical mean-field approach, I demonstrated that these multiplex configurations exhibit unique patterns of order-disorder transitions. Notably, the presence of a two-clique configuration revealed critical behaviors that were not predictable through standard mean-field analyses. The emergence of synchronized states between mismatched layers highlights the complexity of human social interactions, where relationship dynamics across different contexts influence one another. Specifically, it was observed that as the temperature increased, the system could abruptly transition from a balanced state to disorder, illustrating the sensitivity of signed social networks to external influences. This phenomenon mirrors findings in statistical physics, drawing a parallel between the tension of imbalanced triads and energy excitation above the system's ground state [55, 89].

The system exhibits multistability, with various stable states possible depending on temperature and initial conditions. These states include matching order (e.g., all layers in paradise or matching two-clique states), mismatched order (layers in different balanced configurations like paradise and a two-clique), and partially matching order (two-clique states with partial overlap). The mean-field approach primarily focuses on the paradise state and its stability, not fully capturing the behavior involving two-clique states.
The chapter also explores scenarios where one layer exhibits a two-clique configuration instead of a paradise state. This introduces additional transitions, including the synchronization of interlayer relations and a transition to a disorder occurring at a different, lower temperature compared to matching paradise states. This reveals the nuanced dynamics between HB and interactions across multiple layers.

Three different critical temperatures and two characteristic temperatures were identified:

- Critical Temperature (T_c): The critical temperature above which no ordered states are stable increases with the strength of positive interlayer coupling (K) and saturates for high K values. For L-coupled layers with strong coupling, the saturation critical temperature is approximately L times the critical temperature of a single layer. This was predicted analytically and observed in simulations. However, for many weakly interacting layers, the mean-field approach overestimates the rise of T_c.

- Synchronization Temperature (T_s): For systems starting in mismatched or partially matching order states, there is a synchronization temperature (T_s) below which

these states can be stable. Above T_s, thermal fluctuations tend to drive the system towards a matching order state. T_s depends on the specific mismatched state, particularly the overlap between two-clique configurations in different layers.

- Disorder Stability (T_d): There is a minimum temperature (T_d) below which a disordered system will spontaneously order into a matching two-clique state.

- Characteristic Temperatures (T_o and T_m): The study identifies two characteristic temperatures related to transitions from mismatched/partially matching order to a disordered state. To is a lower bound for this transition (for high enough interlayer coupling, To can coincide with T_d), and T_m is an upper bound above which a mismatched order will always transition to disorder, even if a matching order state is still stable. These temperatures also depend on the specific initial mismatched state and the overlap between layers.

Note that all the critical temperatures T_s,T_d,T_c, as well as the characteristic temperatures T_o and T_m depend on model and system parameters, in particular on network size N, intralayer interaction strength A and interlayer coupling K. In addition, as stated before, critical temperature T_s and characteristic temperatures T_o and T_m are determined for a specific mismatched or a partially matching order state and therefore there are many such temperatures for a given system.

Chapter 3

Structural balance on sparse random networks: single and multilayer analysis

3.1 Motivation

In the previous chapter, which focused on structurally balanced dynamics under dense or fully connected assumptions, I now turn to a more realistic and theoretically revealing regime of sparse connectivity. While Chapter 2 demonstrated how triadic interaction generates ordered and disordered phases in idealized topologies, those findings leave unresolved the fundamental question of whether—and how—these phenomena persist when network connections are scarce and locally clustered rather than globally present. This chapter addresses that gap by examining Heider balance dynamics on Erdős-Rényi (ER) networks and subsequently on multilayer networks to determine that network sparsity and interlayer coupling fundamentally reshape critical behavior, alter the stability thresholds of the paradise state, and influence the emergence of disorder.

ER random graphs, where every pair of nodes is connected with fixed probability p [35], are well-researched. The intrinsic structural properties of ER graphs—including motif distributions [90–92], percolation thresholds [93], and their implications for dynamic processes [94–98] have been extensively studied. However, the influence of network sparsity on the emergent dynamics of triadic interaction models, particularly HB, remains insufficiently explored. A critical research question in this domain concerns how network sparsity p influences key properties of structural balance models. While two-body interaction models, such as the Ising model [94, 95, 99], exhibit well-characterized scaling behaviors with network density p, the impact of sparse connectivity on three-body interactions, as in Heider balance models, remains less understood. Prior investigations have predominantly examined phase transitions between ordered and disordered states within

fully connected graphs, but their manifestation in sparse random graphs remains an open problem. Recent studies [100, 101] have begun addressing this gap by identifying the critical temperature at which the disorder emerges as a stable state. However, an equally pertinent yet under-explored aspect is the identification of the upper-temperature limit at which the ordered state, commonly referred to as the paradise state, remains resilient in the presence of stochastic fluctuations.

This chapter advances the theoretical and computational understanding of HB dynamics by moving beyond the conventional complete graph assumption. Specifically, it examines the conditions under which phase transitions occur in Erdős-Rényi random graphs and assesses whether the critical phenomena observed in fully connected networks persist under sparse connectivity constraints. The primary goal is to derive an analytical approximation for the upper stability threshold of the ordered state using a mean-field approach, supplemented by validation through numerical MC simulations. While alternative equilibrium states, such as split cliques, may arise in balanced networks, this research focuses primarily on the paradise state's stability properties.

Additionally, this study extends the analysis to multilayer network structures [10, 66], wherein each layer represents an independently evolving Erdős-Rényi random graph.

3.2 Models and methods

Consider an Erdős-Rényi (ER) network composed of N agents, where each pair of nodes (i, j) is linked with probability p. Structural balance is introduced by assigning a sign x_{ij} to each existing edge, representing the polarization of interactions between agents. I assume a binary interaction model, where $x_{ij} = \pm 1$ denotes positive or negative interactions, and $x_{ij} = 0$ signifies a missing link rather than a neutral state. The network topology remains fixed (quenched disorder), and the total number of links is given by $|E|$.

According to structural balance theory, stable configurations are those in which all triads are balanced. The energy associated with a balanced triad is $-A$, where $A > 0$ quantifies Heider's interaction strength. As derived in [102], the system's Hamiltonian is given by:

$$\mathcal{H} = -A \sum_{i<j<k}^{N} x_{ij} x_{jk} x_{ki}. \tag{3.1}$$

The degenerate ground states of this Hamiltonian correspond to fully balanced configurations, where each triad (ijk) satisfies one of two conditions: (i) all links are positive, or (ii) two links are negative while one is positive (e.g., $x_{ij} = x_{jk} = -1$, $x_{ki} = 1$). In both cases, the product $x_{ij}x_{jk}x_{ki}$ is positive, minimizing the system's energy contribution to $-A$ per triad [4, 5, 54, 56, 88, 103]. Conversely, unbalanced triads (containing either one negative link and two positive ones, or three negative links) yield an energy contribution of $+A$, as

their polarization product is negative. These states are considered excited configurations [60, 102]. Triads involving missing links contribute zero energy to the Hamiltonian due to their incomplete connectivity.

In a bilayer extension of the ER network, two network layers share the same set of N nodes. However, a link (ij) present in one layer may be absent in the other, where the interlayer coupling exists only between link states of the same node pairs in different layers (see Fig. 3.1). The energy of the system incorporates a term that accounts for discrepancies in interlayer link signs. If link polarizations match in both layers, the energy decreases by $K > 0$; if they differ, the energy increases by K.
The bilayer Hamiltonian, incorporating layer-specific interaction strengths $A^{(1)}$ and $A^{(2)}$, is expressed as:

$$\mathcal{H} = -A^{(1)} \sum_{i<j<k}^{N} x_{ij}^{(1)} x_{jk}^{(1)} x_{ki}^{(1)} - A^{(2)} \sum_{i<j<k}^{N} x_{ij}^{(2)} x_{jk}^{(2)} x_{ki}^{(2)} - K \sum_{i<j}^{N} x_{ij}^{(1)} x_{ij}^{(2)}. \tag{3.2}$$

The ground states of this bilayer system occur when all existing triads in each layer are balanced ($x_{ij}^{(\alpha)} x_{jk}^{(\alpha)} x_{ki}^{(\alpha)} = 1$) and when link polarizations are identical across layers ($x_{ij}^{(1)} = x_{ij}^{(2)}$). The model does not incorporate intrinsic node attributes; instead, the system's evolution is entirely governed by fluctuations in link signs $x_{ij}^{(\alpha)}$. Thermal noise, represented by a temperature parameter T, introduces randomness into the dynamics, capturing the stochastic nature of human interactions. At nonzero temperatures $T > 0$, the system reaches an equilibrium distinct from the ground state. A special ground state termed the "paradise" state, occurs when all existing links are positive ($x_{ij}^{(\alpha)} = 1$). Other equilibrium states exist, characterized by internally cohesive groups with hostile intergroup relations [54, 88, 102], though this study does not focus on these alternative configurations.

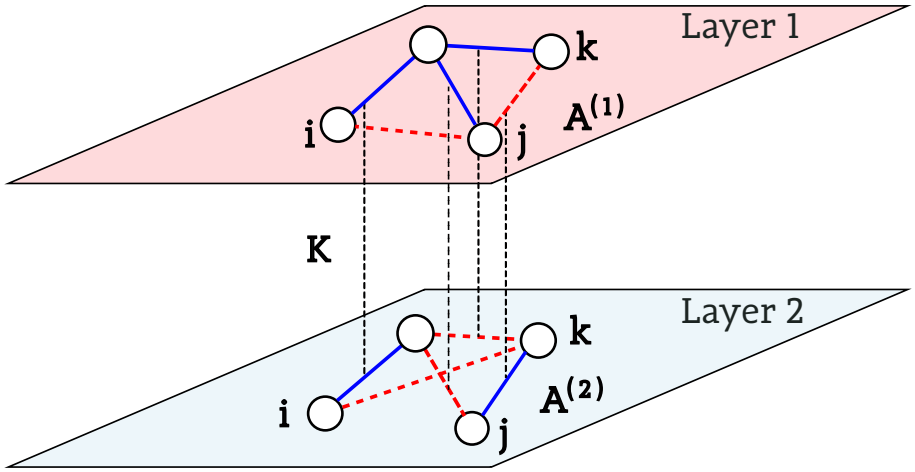


Figure 3.1: An illustration of interactions between signed links of a bilayer ER network. Positive links are represented by solid blue lines, and negative links are represented by dashed red lines. The structural balance drives the network to a balanced state, which is governed by the intralayer interactions as well as the interlayer interactions. When the interlayer coupling strength $K = 0$, the system is reduced to two independent single-layer ER networks.

3.3 Analytical results

3.3.1 Mean-field approximation for single ER network

I begin the analysis with a monolayer Erdős-Rényi (ER) network of size N, where each pair of nodes is connected with probability p. The system's dynamics are governed by the Hamiltonian defined in Eq. (3.1). Prior studies have formulated the mean-field description of structural balance for a fully connected network [102], demonstrating that this framework inherently identifies only the paradise state as the ground state. However, it does not capture the two-clique states, which are also ground states, due to intrinsic limitations in describing internally structured states within a mean-field framework. Consequently, analytical predictions are restricted to the transition from the paradise state to disorder and do not encompass transitions between disordered and two-clique states. Notably, these transitions occur at distinct temperatures [10, 102].

This study extends the analytical approach to networks with stochastic topology. As a result, the analysis remains focused on the transition from a perturbed paradise state to a fully disordered state. While the paradise state is improbable in a spontaneous evolution from disorder, it constitutes a stable configuration rather than a metastable one, as it is separated from alternative configurations by substantial energy barriers composed of multiple unbalanced triads. Furthermore, initializing the system in the paradise state is justified when investigating the threshold at which predominantly amicable social interactions become unstable. In contrast to the fully connected network examined in [102], the model introduces missing links with a fraction of $1 - p$. Given that an ER network

lacks correlations and possesses fewer edges and triads than a complete graph when $p < 1$, I anticipate similar qualitative behavior but with reduced effective interaction strength and a correspondingly lower critical temperature.

Consider an edge between two nodes, i and j. The edge polarization x_{ij} is a stochastic variable subject to thermal fluctuations. At a fixed temperature T, the probability of the system residing in a microstate $x_{ij} = \pm 1$ follows the Boltzmann distribution:

$$P(x_{ij}) = \frac{\exp[-h(x_{ij})/T]}{\sum_{x_{ij}=\pm 1} \exp[-h(x_{ij})/T]}, \tag{3.3}$$

where the Hamiltonian governing the local interaction is given by

$$h(x_{ij}) = -A x_{ij} \sum_{k \neq i,j}^{N} x_{jk} x_{ik}. \tag{3.4}$$

The expected value of x_{ij} is then:

$$\langle x_{ij} \rangle = \sum_{x_{ij}=\pm 1} P(x_{ij}) x_{ij}. \tag{3.5}$$

Following the mean-field approach in [102], I define the mean link polarization across the network as:

$$x = \frac{1}{|E|} \sum_{i<j}^{N} \langle x_{ij} \rangle \,. \tag{3.6}$$

This summation accounts only for existing links ($x_{ij} = \pm 1$), excluding absent links, which are formally set to $x_{ij} = 0$. Hence, the normalization factor is $|E|$ rather than $N(N-1)/2$. The polarization x thus quantifies the global mean link sign of the network.

To approximate interactions in Eq.(3.1) via a mean-field approach, I replace pairwise interactions with their statistical averages. Assuming $x_{jk}x_{ki} \approx x^2$ and statistical equivalence across all links such that $\langle x_{ij} \rangle \equiv x$, I recognize that the summation in Eq. (3.1) extends over all node pairs jk and ki. However, since an ER network is characterized by structural randomness, a fraction $1 - p^2$ of these terms vanishes due to missing edges. The probability that both links jk and ki exist is p^2, given their independent formation. Averaging over both existing and missing links, I obtain the mean-field energy of a given link x_{ij}:

$$h(x_{ij}) \overset{\text{MF}}{=} -A x_{ij} (N-2) p^2 x^2. \tag{3.7}$$

This result mirrors the mean-field formulation for a complete graph [102], differing only by the rescaling factor p^2, which accounts for interaction dilution due to missing edges. Using Eqs. (3.3), (3.5), and (3.7), I derive the self-consistent equation governing x as a

function of temperature:

$$x = \tanh[a(T)x^2], \tag{3.8}$$

where $a = \dfrac{Ap^2M}{T}$ and $M = (N-2)$ represents the number of triads containing a given edge in a complete graph. The tangency condition characterizes the transition from order to disorder:

$$\frac{d}{dx}\tanh(a_c x_c^2) = 1. \tag{3.9}$$

Eqs. (3.8) and (3.9) constitute a system of algebraic equations determining x_c and a_c at the discontinuous phase transition. While these transcendental equations lack closed-form solutions, their primary distinction from the complete graph case is the rescaling of interaction strength a by p^2.

Figure 3.2a presents the numerical solution for mean polarization in a single-layer ER network, obtained from Eqs. (3.8) and (3.9) for various p values. Assuming an initial paradise-state condition, I expect that increasing temperature drives equilibrium polarization downward from $x = 1$ until reaching x_c, where a discontinuous transition at T_c leads to a disordered state with $x = 0$. Numerical analysis [102] yields $x_c \approx 0.79638$, invariant under model parameters. The critical temperature T_c is given by:

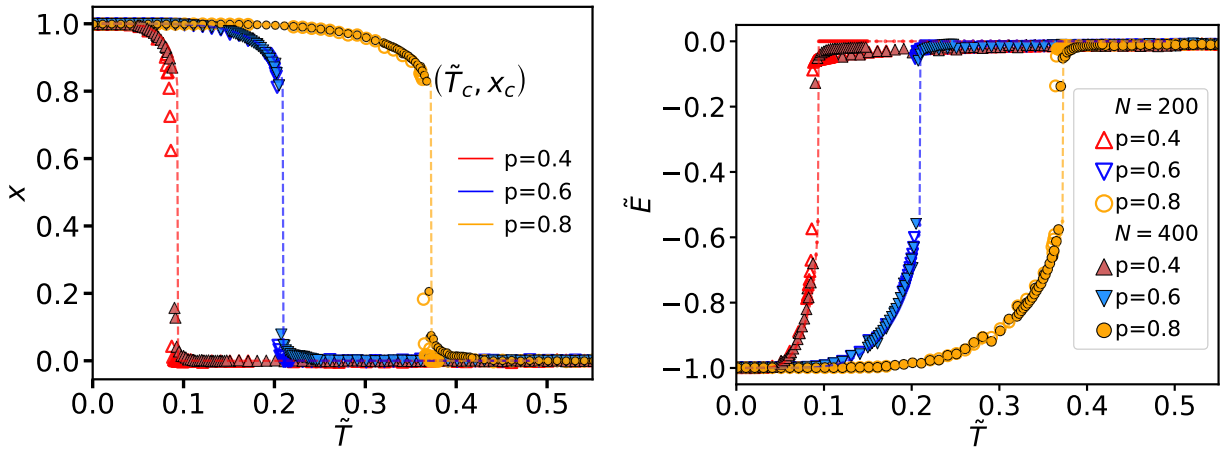
$$T_c = \left(\frac{1}{a_c}\right) Ap^2M, \tag{3.10}$$

where $a_c \approx 1.71649$ remains independent of model parameters. Defining the normalized critical temperature as $\tilde{T}_c = T_c/AM$, I obtain:

$$\tilde{T}_c = (1/a_c)p^2 \approx 0.58258p^2. \tag{3.11}$$

3.3.2 Mean-field approximation for bilayer ER network

The extension of mean-field approach to a bilayer network composed of two complete graph layers is discussed in [10], where a pair of coupled links $\vec{x}_{ij} = [x_{ij}^{(1)}, x_{ij}^{(2)}]$ is considered as the elementary subsystem interacting with the mean-field, instead of a single link x_{ij}. Applying the same methodology for a bilayer random graph is not trivial because not all node pairs ij have links in both layers. If I consider two independent random graphs with a probability p of each potential link existing and treat them as two layers of the same duplex graph, then I have $\frac{N(N-1)}{2}p^2$ node pairs that have links in both layers, $\frac{N(N-1)}{2}2p(1-p)$ node pairs featuring a link only in one of the layers, and $\frac{N(N-1)}{2}(1-p)^2$ node pairs without any links (see Fig. 3.3). Within the mean-field approximation, the lack of link correlations in (ER) random graphs permits the neglect of specific link configurations in the network. As a result, unpaired links in layers 1 and 2 can be reorganized into virtual, non-interacting link pairs. Consequently, the network consists of $\frac{N(N-1)}{2}p^2$ link pairs interacting with a



(a) (b)

Figure 3.2: Mean-field solution obtained analytically (lines) and the simulation results (triangles and circles) for a single layer ER network where a first-order transition is observed at the normalized critical temperature $\tilde{T}_c$. The normalized critical temperature $\tilde{T}_c$ decreases with the decreasing value of connection probability p since the network is sparser and every link is influenced by a smaller number of triads. Panel (b) shows the corresponding normalized energy (Eq.3.16) of the simulated network. The results are displayed for $p = 0.4, 0.6, 1$ and $N = 200, 400$, where the transition point moves from left to right for increasing values of p. Both panels share the same legend. It is worth noting that normalized critical temperature values $\tilde{T}$ are independent of network size N.

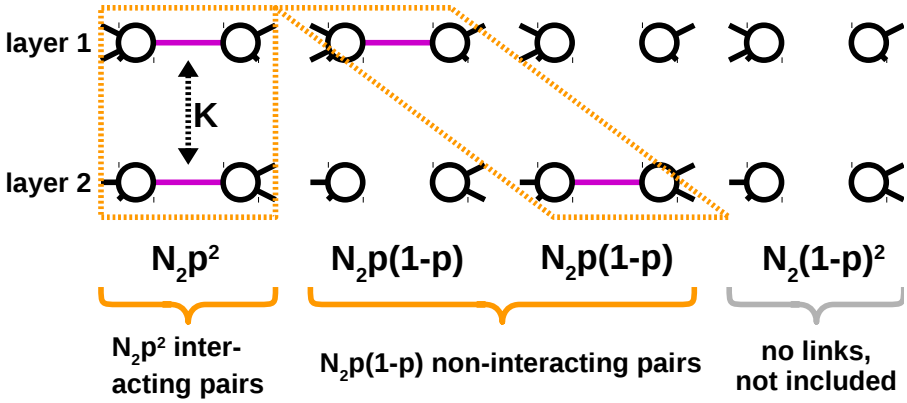


Figure 3.3: In bilayer Erdős-Rényi graph with N nodes and probability p for each link to exist, there are N_2p^2 pairs of nodes that have links in both layers, $2N_2p(1-p)$ pairs with link in only one layer and $N_2(1-p)^2$ node pairs without any links, where $N_2 = N(N-1)/2$ is the total number of node pairs. I can arrange the single links that have no equivalent links in the other layer into pairs, ignoring the actual topology of the graphs into $N_2p(1-p)$ link pairs that have no coupling between them. This arrangement results in N_2p^2 coupled link pairs with coupling strength K and $N_2p(1-p)$ link pairs that are not coupled (both link pair types encircled in orange). Since the arranged link pairs do not interact (no coupling) and I consider mean-field approximation only, the exact original placement of these links in the network does not matter. This allows us to consider bilayer random graphs in terms of link pairs only, out of which a fraction p is coupled and does not have to consider single unpaired links in the mean-field approach.

coupling strength K, alongside $\frac{N(N-1)}{2}p(1-p)$ virtual, non-interacting link pairs where the effective coupling is $K = 0$. Since the mean-field approach considers only link states, node pairs without any connections are disregarded.

This approximation holds under the assumption that both layers share an identical link density p, ensuring an equivalent mean number of links in each layer. Under this condition, the methodology established for bilayer complete graphs [10] can be applied, with the crucial distinction that only a fraction p of link pairs exhibit interlayer interactions. Consequently, the mean interlayer interaction strength is reduced from K to Kp.

Assuming layer symmetry, such that $A^{(1)} = A^{(2)} = A$, and defining the mean link pair polarization as:

$$\vec{x} \equiv [x^{(1)}, x^{(2)}] = \left[\frac{\sum_{i,j}^{N} x_{ij}^{(1)}}{\sum_{i,j}^{N} |x_{ij}^{(1)}|}, \frac{\sum_{i,j}^{N} x_{ij}^{(2)}}{\sum_{i,j}^{N} |x_{ij}^{(2)}|} \right], \tag{3.12}$$

I approximate the energy associated with the state $\vec{x}_{ij}$ as:

$$h(\vec{x}_{ij}) \overset{\text{MF}}{=} -A \left(x_{ij}^{(1)} + x_{ij}^{(2)} \right) Mp^2 x^2 - Kp x_{ij}^{(1)} x_{ij}^{(2)} \tag{3.13}$$

where $M = N - 2$.

This formulation indicates that, within the mean-field framework, the energy of a link pair in bilayer ER graphs retains the structure observed in bilayer complete graphs [10], with two key adjustments: the intralayer interaction strength is rescaled by p^2, consistent with the single-layered ER network case, while the interlayer interaction strength is reduced by p due to the fraction of link pairs that engage in interlayer interactions. Since both interaction strengths undergo a straightforward rescaling, the system behavior should resemble that of the complete graph, albeit with a modified critical temperature.

According to the results of [10], the critical temperature for the bilayer ER graph is expressed as:

$$T_c = Ap^2 M \left(\frac{1}{\ln z} \right) \left(\frac{D^2(z^4 - 1)}{D^2(z^4 + 1) + 2z^2} \right)^2, \tag{3.14}$$

where $D = \exp[K/(ApM)]$ and $z = z(D)$ satisfies the transcendental equation:

$$8 \ln z = \frac{(z^4 - 1)(z^4 D^2 + 2z^2 + D^2)}{z^2(z^4 + 2z^2 D^2 + 1)}. \tag{3.15}$$

This equation maintains the same form as in the complete graph case, with the distinction that D now explicitly depends on p.

The normalized critical temperature, defined as $\tilde{T}_c = T_c/AM$, exhibits a dependence on p. While an initial analysis suggests a quadratic relationship, implying a p^2 scaling of the critical temperature, the dependency of D on p introduces further complexity. In the limiting case where $D = 1$ (corresponding to $K = 0$), the layers behave as independent systems, and $\tilde{T}_c$ follows a direct quadratic dependence on p.

In summary, within the mean-field paradigm, if the intralayer interaction strength A is rescaled to A/p^2 and the interlayer interaction strength K is adjusted to K/p, the statistical properties of the system remain independent of p. Under this transformation, the normalized critical temperature $\tilde{T}_c$ remains invariant and corresponds to the result obtained for the complete graph case (i.e., $p = 1$). This invariance is further substantiated in the inset of Fig. 3.10.

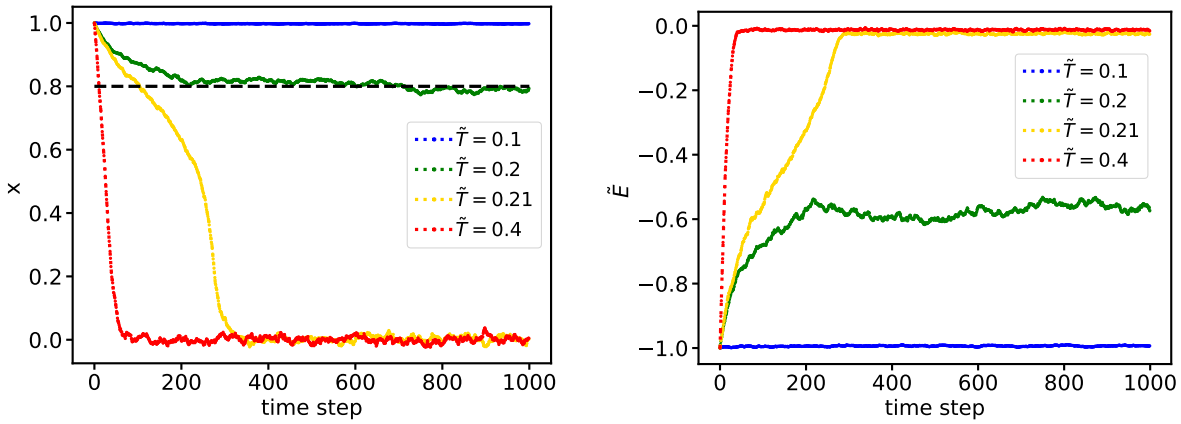
3.4 Numerical simulation

Numerical MC simulations were conducted to validate the analytical results derived for both single-layer and bilayer ER networks. The simulations employed an asynchronous Metropolis algorithm, wherein each time step comprised $pN(N-1)/2$ elementary updates. For the single-layer ER network, each update involved selecting a random pair of connected nodes (i, j), computing the energy difference ΔE associated with flipping the link sign, and executing the inversion with probability $\min(1, e^{-\Delta E/T})$.
In the bilayer setting, each update entailed selecting two independent link pairs, (i, j) and (k, l), one from each layer. The energy difference ΔE was calculated between the current configuration and the configuration where both link signs were flipped, incorporating both intralayer Heider interactions and interlayer coupling effects. The link signs $x_{ij}^{(1)}$ and $x_{kl}^{(2)}$ were then inverted with probability $\min(1, e^{-\Delta E/T})$.

The simulations were performed over 50 independent realizations of ER networks, spanning a range of link densities from sparsely connected networks ($p = 0.1$) to the fully connected limit ($p = 1$). The system temperature $\tilde{T}$ was incrementally increased from zero in discrete steps of $\Delta\tilde{T}$ to identify the normalized critical temperature $\tilde{T}_c$, defined as the point where both the mean link polarization x and the normalized intralayer energy vanish. The precise value of $\tilde{T}_c$ was determined as the peak of the slope in the energy profile averaged over 50 realizations, marking the transition from an ordered ground state to a disordered higher-energy state (see Sec. 3.4.2). Standard errors were computed across all realizations to quantify statistical variations.
To analyze the simulation outcomes, the normalized intralayer and interlayer energies were computed from the Hamiltonian (3.2). The intralayer energy for each layer was normalized to reflect the average energy per existing triad:

$$\tilde{E}_\alpha = \frac{-A^{(\alpha)} \sum_{i<j<k}^{N} x_{ij}^{(\alpha)} x_{jk}^{(\alpha)} x_{ki}^{(\alpha)}}{A^{(\alpha)} N_\Delta^{(\alpha)}} \tag{3.16}$$

where $A^{(\alpha)}$ represents the intralayer interaction strength, and $N_\Delta^{(\alpha)}$ denotes the number of triads present in layer α, with a mean value given by $\frac{N(N-1)(N-2)}{6}p^3$. Throughout the simulations, the intralayer interaction strength was fixed at $A^{(\alpha)} = 1$. Similarly, the



(a) (b)

Figure 3.4: Examples of the time evolution of the average values of all links' polarization for temperatures below (blue and green lines) and above (yellow and red lines) the normalized critical temperature ($\tilde{T}_c$) for a network size $N = 200$ with connection probability $p = 0.6$. For $\tilde{T} = 0.2$, fluctuations are observed around values that are close but always larger than the critical solution x_c (dashed black line), however for a slightly higher temperature, the asymptotic values of mean link polarization show a discontinuous transition. The starting point of all simulations is the paradise state. The normalized critical temperature $\tilde{T}_c$ is found between the two temperature values $\tilde{T} = 0.2$ and $\tilde{T} = 0.21$.

interlayer energy was normalized to represent the average energy per interacting link pair:

$$\tilde{E}_K = \frac{-K \sum_{i<j}^{N} x_{ij}^{(1)} x_{ij}^{(2)}}{K N_K} \tag{3.17}$$

where K denotes the interlayer interaction strength, and N_K represents the number of coupled link pairs, with a mean value of $\frac{N(N-1)}{2} p^2$.

3.4.1 Critical behavior of a single layer ER network

For the paradise-state initial configuration, the system transitions to a disordered state once the temperature surpasses the normalized critical temperature $\tilde{T}_c$. Figure 3.4 illustrates the temporal evolution of x for various normalized temperatures $\tilde{T}$ in a system of size $N = 200$. The dashed black line indicates the critical threshold x_c, as defined in Section 3.3.1.

Figure 3.2 presents the mean link polarization and the corresponding normalized intralayer energy $\tilde{E}^{(1)}$, computed from Eq. 3.16, for varying connection probabilities p across network sizes $N = 200$ and $N = 400$. The results indicate that the normalized critical temperature $\tilde{T}_c$ increases monotonically with p. Beyond the critical temperature, the intralayer energy asymptotically approaches zero, and the mean polarization exhibits fluctuations around zero. The critical temperature values extracted from simulations exhibit strong agreement with the predictions of the mean-field approximation.

A comparative analysis of the normalized critical temperature $\tilde{T}_c$ derived from the analytical mean-field approach (Eq. 3.14) and those obtained from MC simulations is depicted in Figure 3.5. The system is initialized in the paradise state, and the transition to a disordered regime occurs at $\tilde{T}_c$, beyond which the ordered state becomes unstable. The simulated critical temperature follows a power-law dependence, with an exponent $b \approx 2.1$, determined through fitting a power function $\tilde{T}_c(p) = cp^b$ to the simulation data. This scaling is in close alignment with mean-field theoretical predictions, which suggest that $\tilde{T}_c$ should scale as p^2. A detailed account of the methodology employed to extract the critical temperature from simulation data is provided in Sec. 3.4.2. Although minor discrepancies emerge for transition points when $p < 0.5$ in comparisons between numerical simulations and the mean-field approximation, the analytical framework remains broadly valid. The reduced accuracy for sparse networks is expected, given that the mean-field approximation is optimally suited for complete graphs. Notably, for $p = 0.1$, the characteristic sharp first-order transition observed in denser networks is absent, suggesting a qualitative change in the nature of the transition. This phenomenon aligns with prior observations reported by Malarz et al. [101], indicating that numerical results and analytical predictions should be interpreted cautiously in this regime. A more detailed discussion on this issue is included in the Sec. 3.4.3.

When the connection probability p is varied while simultaneously rescaling the interaction strength to A/p^2, the normalized critical temperature $\tilde{T}_c$ remains invariant. This behavior is demonstrated in the inset of Figure 3.5, where additional simulations performed with rescaled interaction strength A/p^2 confirm that the observed $\tilde{T}_c$ remains approximately constant and corresponds to the critical temperature of the complete graph with interaction strength A. Consistent with the primary dataset, the accuracy of analytical predictions diminishes as p decreases.

3.4.2 Finding critical temperature from numerical simulations

The critical temperature derived from MC simulations is determined by locating the point of maximum slope in the system's normalized energy profile, as defined in Eq. (3.16), which signifies the transition from an ordered ground state to a disordered state characterized by higher energy. The normalized energy $\tilde{E}$ is selected as the primary observable for identifying the critical temperature due to its relative robustness against stochastic fluctuations, making it a more reliable metric than mean polarization.

For individual simulation realizations, the system undergoes a discontinuous transition. However, due to the finite system size, intrinsic fluctuations introduce noise in the numerical results, leading to a less distinct transition. Averaging over multiple simulation runs yields a smoother and more continuous transition curve that better aligns with theoretical predictions, effectively mitigating finite-size effects, particularly for lower values

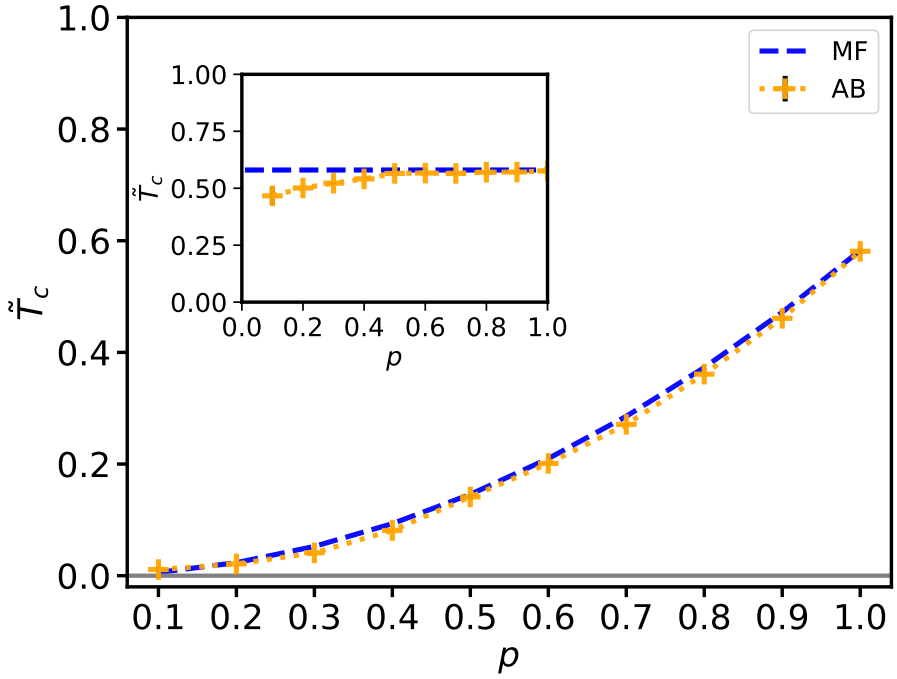
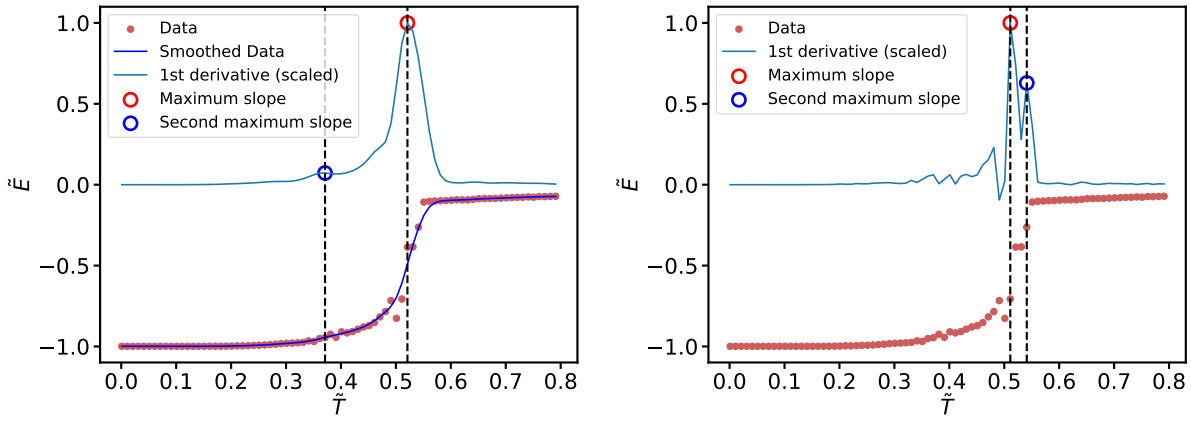


Figure 3.5: Critical temperature increases as p^2 for a single layer ER network started from the paradise state (shown for $N = 200$). The dashed blue line represents the analytically predicted normalized critical temperature $\tilde{T}_c$, and the orange pluses correspond to $\tilde{T}_c$ values from the simulation. The inset shows the normalized critical temperature observed in simulations where interaction strength is normalized to A/p^2. The normalized critical temperature $\tilde{T}_c$ remains approximately constant for all values of p and has the same value as for the complete graph.

of p, as depicted in Figure 3.2. It is also crucial to recognize that, for a fixed parameter p, each simulation instance corresponds to a different network topology due to the random nature of the ER graph model. As a result, simulations represent an ensemble of distinct ER network realizations, contributing to the observed noise in the transition behavior across numerical studies.

To refine the estimation of the critical temperature, I employ Gaussian kernel smoothing on the normalized energy data, averaging over 50 independent realizations. The selection of an appropriate smoothing parameter σ, which determines the width of the Gaussian kernel, plays a crucial role in enhancing result clarity and accuracy. This consideration becomes particularly important for small values of p, where the transition behavior is inherently more noisy. To determine the optimal value of σ, I systematically evaluate a range of values starting from zero. For each σ, I identify the two most prominent maxima of the slope. If only a single maximum is detected or if the ratio between the highest and second-highest maxima exceeds 10 (a threshold determined empirically), I conclude that the smoothing process is optimal and that the identified maximum corresponds to the critical temperature. If this criterion is not met, I increment σ by 0.1 and repeat the procedure until the specified conditions are satisfied. Figure 3.6 illustrates the pro-



(a) (b)

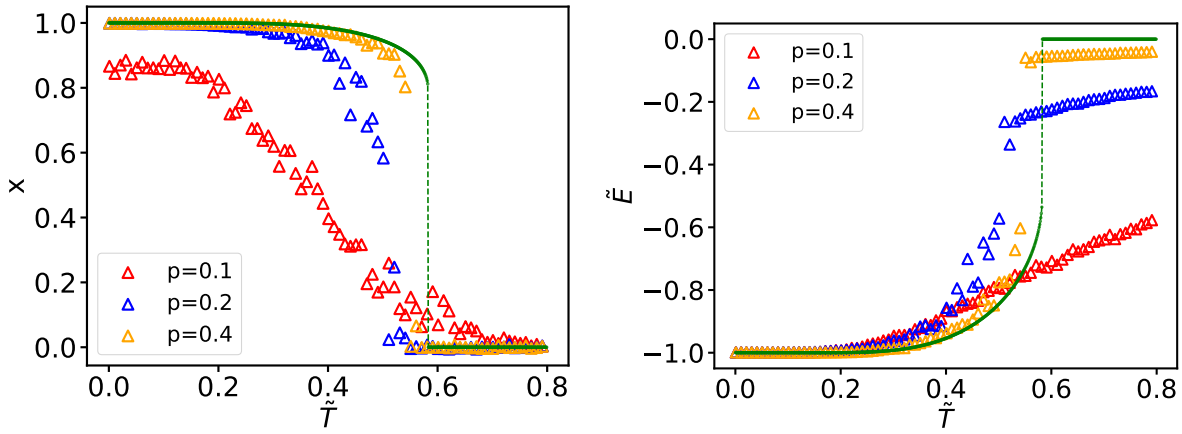
Figure 3.6: The critical temperature for a single-layer ER network (normalized interaction strength) of size $N = 200$ and connection probability $p = 0.3$ is identified by examining the maximum of the slope in the dependence of normalized intralayer energy $\tilde{E}$. This temperature is derived from the maximum of the first derivative, where the ratio of the slopes (red and blue circles) differs between the smoothed data (a) and the unsmoothed data (b). The optimal σ for the smoothed data is chosen by examining the ratio of the absolute values of the largest and second-largest first derivatives of smoothed data. In this case, the smoothed data has an optimal σ of 0.51.

cess of identifying slope maxima for both unsmoothed and smoothed data. Given the discrete nature of the original dataset, the smoothed data retains this discrete character. To accurately determine slope maxima, numerical derivatives of the smoothed dataset are computed using central differences for interior points and first differences at the boundaries.

3.4.3 Continuity of transition for sparse graphs

In the simulation data for low values of p, the discontinuity in the transition from the ordered to the disordered state is not immediately discernible. This effect is particularly challenging to observe in the primary dataset presented in Figure 3.5 due to the proximity of critical temperatures to zero. However, it becomes more pronounced in the re-scaled inset data. Figure 3.7 illustrates the behavior of polarization x and normalized energy $\tilde{E}$ for sparse ER networks under the constraint $A/p^2 = \text{const.}$. Notably, for $p = 0.1$, the transition between ordered and disordered states does not exhibit a distinct discontinuity, even when accounting for the finite-size effects inherent in numerical simulations. While the critical temperature T_c for $p = 0.1$ can still be determined using the data shown in Figure 3.7b, and while these values remain relatively close to those predicted by the analytical approach, the nature of the transition deviates from expectations. This discrepancy necessitates a cautious interpretation of the numerical results. It is also important to emphasize that, although the observed effect bears similarities to findings reported by

Malarz et al. [101] in the context of a disorder-to-order transition, the nature of the transitions investigated in the two studies differs fundamentally. Consequently, the present findings neither constitute a direct confirmation of the results obtained by Malarz et al. nor can their conclusions be regarded as a validation of this study.



(a) (b)

Figure 3.7: Simulation results for the normalized interaction strength A/p^2 for a single-layer ER network of size $N = 200$ and different connection probability p. A distinct first-order transition is not observed for $p = 0.1$ in normalized energy $\tilde{E}$ and mean polarization x data. Green curves show the mean-field results that are independent of the parameter p.

3.4.4 Distribution of triads

To rigorously substantiate the interpretation of the system state in terms of polarization x and energy E, I conducted a comprehensive analysis of the frequency distribution of distinct triad types within a single-layer ER network. Triads are categorized based on the number of negative links they contain and are designated as Δ_l, where the subscript l denotes the number of negative edges. The structural balance of a triad is dictated by the product of the signs of its constituent edges. Specifically, triads of type Δ_0 (all positive links) and Δ_2 (two negative links) are classified as structurally balanced, whereas triads of type Δ_1 (one negative link) and Δ_3 (three negative links) are unbalanced. When the ER network is initialized in the paradise state ($x = 1$), findings reveal that at sufficiently low temperatures (as illustrated in Fig. 3.8), the network is overwhelmingly dominated by the balanced Δ_0 configuration, reinforcing the stability of the paradise state. As long as the temperature remains below the normalized critical threshold $\tilde{T}_c$, the majority of triads persist in the Δ_0 state, although thermal fluctuations introduce a measurable fraction of Δ_1 triads. Above the critical temperature $\tilde{T}_c$, the triad distribution exhibits a pronounced shift. In this high-temperature regime, entropy-driven dynamics govern the system, yielding a triad distribution wherein Δ_1 and Δ_2 appear approximately three times more frequently than Δ_0 and Δ_3, a pattern consistent with the expectation for

randomly assigned link signs. However, a residual influence of Heider's balance theory persists, as structurally balanced triads (Δ_0 and Δ_2) remain slightly more prevalent than their unbalanced counterparts (Δ_1 and Δ_3).

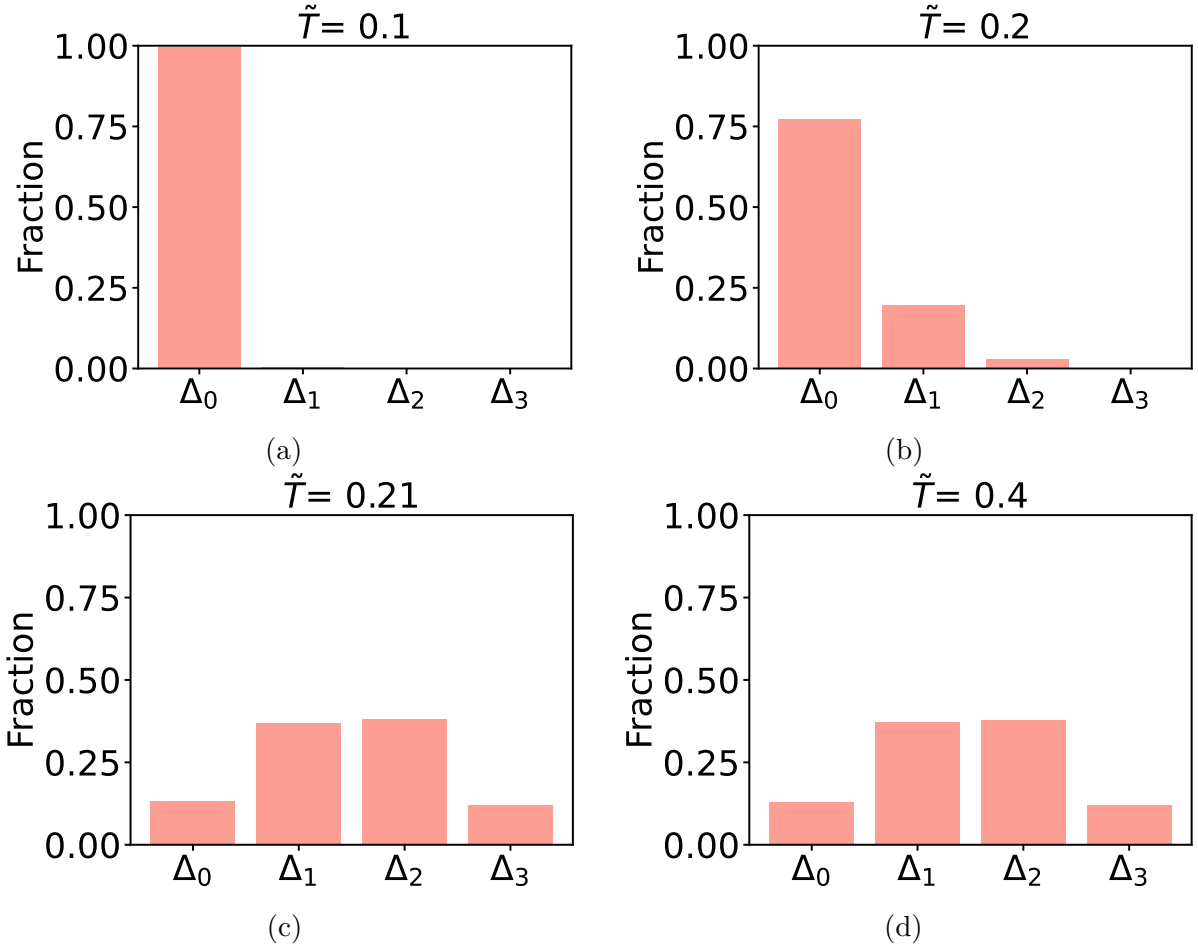
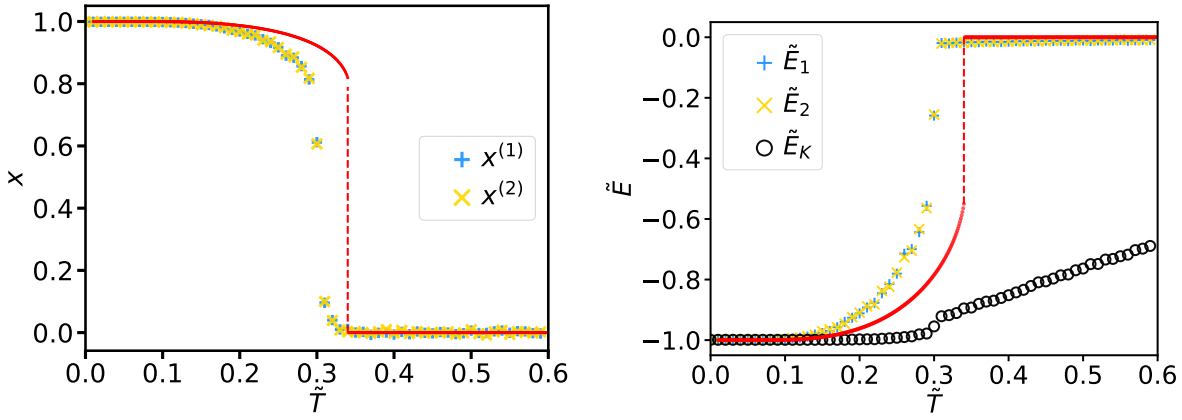


Figure 3.8: The distribution of triads for the ER graph initialized from the paradise state representing the network's asymptotic state depicted in Fig.3.4 at different temperatures for size $N = 200$ and connection probability $p = 0.6$. The ER network started from the paradise state has all its triads in the balanced state (Δ_0) at low temperatures (a). As the temperature remains below $\tilde{T}_c \approx 0.2097$, the distribution retains a high proportion of stable Δ_0 triads (b). Above the critical temperature, the distribution shifts dramatically to $\approx \left(\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}\right)$ (c and d).

3.4.5 Effect of coupling for a bilayer ER network

In a bilayer ER network, the system is considered balanced when both layers exist in the paradise state. This configuration is characterized by exclusively positive intralayer interactions, where $\langle x_{ij}^{(1)} \rangle = \langle x_{ij}^{(2)} \rangle = 1$, resulting in a fully cooperative and internally consistent network structure across both layers. In the simulations, the interaction parameters within each layer are fixed at $A^{(1)} = A^{(2)} = 1$, ensuring uniform interaction strength. Up to a critical temperature $\tilde{T}_c = \tilde{T}_c(K, p)$, the system remains in a polarized

state with $\langle x_{ij}^{(1)} \rangle \approx \langle x_{ij}^{(2)} \rangle > 0$. However, when the temperature surpasses $\tilde{T}_c$, the system undergoes a discontinuous transition to a disordered phase, where $\langle x_{ij}^{(1)} \rangle \approx \langle x_{ij}^{(2)} \rangle \approx 0$. This transition is illustrated in Fig. 3.9, which shows a strong alignment across both layers below $\tilde{T}_c$, whereas, above this threshold, the mean polarization collapses to near zero, indicating a loss of order. The critical temperature estimated from these numerical simu-



(a) (b)

Figure 3.9: Mean link polarization (a) and corresponding normalized energies (b) for a duplex ER network as a function of the normalized temperature $\tilde{T} = T/(AM)$ for a system of size $N = 200$ nodes, with connection probability $p = 0.6$ and a coupling strength $K = 25$. The results show averages of 50 independent simulations. Initial conditions were paradise states at both layers, where the overlapping blue plus and yellow cross symbols correspond to layers (1) and (2), respectively. The red points at both panels show the analytical results for mean polarization and intralayer energies. While normalized intralayer energies $\tilde{E}_1$ and $\tilde{E}_2$ (overlapping blue plus and yellow cross symbols) display a discontinuous transition to around zero, the interlayer energy E_K (black circle) shows only a small change. Then, it increases slowly when $T > \tilde{T}_c$, indicating that layers remain partially synchronized even in a disordered state.

lations is lower than the values predicted by mean-field theory. Concurrent with the loss of polarization, a discontinuous shift is observed in the normalized intralayer energies $\tilde{E}_1$ and $\tilde{E}_2$, both of which approach zero. Despite the disordering transition, some localized ordering persists, causing the mean energy to remain slightly negative. The normalized interlayer energy $\tilde{E}_K$ also undergoes a transition but remains negative both above and below $\tilde{T}_c$, indicating that partial alignment of link signs across layers persists even in the disordered phase. This behavior is analogous to that observed in fully connected graphs ($p = 1$), as discussed in Ref. [10]. Figure 3.10 compares the normalized critical temperature $\tilde{T}_c = T_c/(AM)$ obtained from mean-field approximations with values derived from MC simulations, where both layers are initialized in the paradise state. The results show that $\tilde{T}_c$ increases with p and exhibits strong agreement with analytical predictions at high p. However, deviations emerge at lower p, likely due to the limitations of mean-field theory in describing sparse network behavior. If the connection probability p is varied while

simultaneously rescaling the intralayer interaction strength to A/p^2 and the interlayer interaction strength to K/p, the critical temperature predicted by analytical mean-field calculations (Eq. 3.14) remains invariant with respect to p. This theoretical prediction is confirmed through numerical simulations, as demonstrated in the inset of Fig. 3.10.

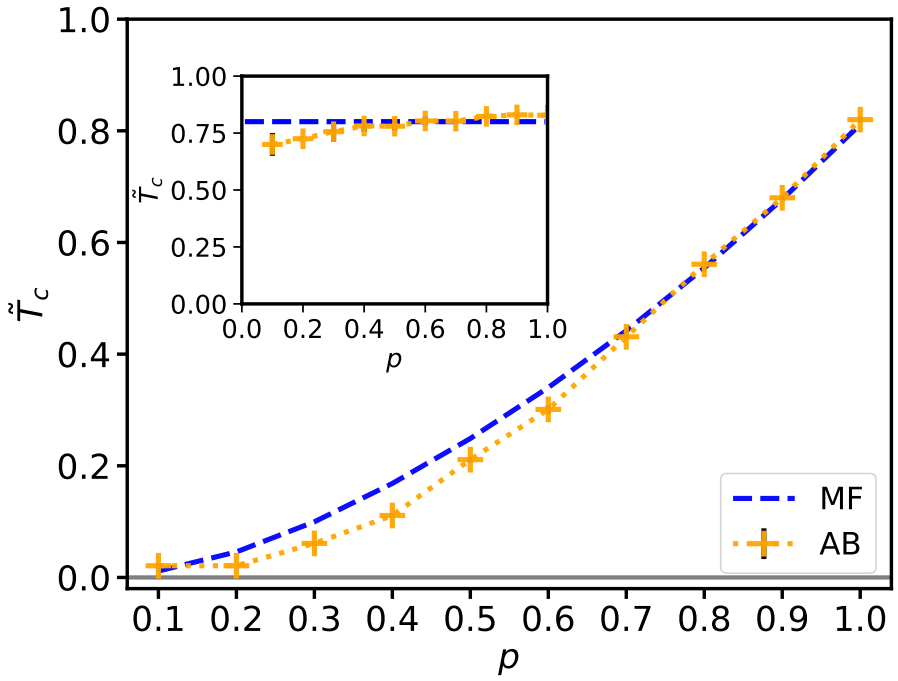


Figure 3.10: The normalized critical temperature $\tilde{T}_c = T_c/(AM)$ for a bilayer ER network increases with connection probability p (shown for $N = 200$ and $K = 100$). The blue dashed line shows the analytical results, and the orange pluses correspond to the results of agent-based simulation. There is some discrepancy in the critical temperature for $p = 0.1$ and $p = 0.2$. The inset shows the situation when intralayer Heider coupling strength is re-scaled to A/p^2 and interlayer Ising coupling to K/p. In such case, the normalized critical temperature $\tilde{T}_c$ remains approximately constant, as predicted analytically.

3.5 Discussion and Conclusions

In this chapter, I look beyond the conventional complete graph framework that is employed in structural balance theory. Complete graph structure may be sufficient to model small social groups, but given the human limitations on the number of relations humans maintain [104], assuming all-to-all relations in larger social groups is impossible. While ER graphs may not capture all the intricacies of social communities, just as Heider balance theory does not fully encapsulate the complexities of opinion dynamics, I aimed to investigate whether the phase transitions observed in complete graphs also occur in random graphs. In this study, I analyze how the HB model behaves in Erdős-Rényi (ER) random graphs, focusing on how it transitions from a well-ordered "paradise" state to disorder in both single-layer and bilayer networks. My results show that, in a single-layer network,

the critical temperature of this transition scales with network density as p^2. However, in a bilayer network, this relationship becomes much more complex.

I find that the Heider balance model in an ER network with Heider interaction strength A and interlayer interaction strength K behaves similarly to a complete graph with interaction strengths Ap^2 and Kp, as long as the network density p is not too small. This scaling behavior is different from typical two-body interaction models, like the Ising model, where the critical temperature scales linearly with p [94, 95]. The key reason behind this difference is that the number of triads a given link is part of increases faster with link density than the probability that a link in one layer has a corresponding link in the other layer.

To describe this behavior, I use a mean-field MF approach and confirm the theoretical predictions with numerical MC simulations. However, it is to be noted that the mean-field predictions become less accurate for networks with low link densities. Previous research on structural balance in ER graphs [100, 101] has also studied critical temperature transitions. For example, [100] focused on the temperature at which disorder becomes stable in networks that start in a disordered state. However, due to system multistability, as also discussed in [100], this transition temperature differs from the temperature $\tilde{T}_c$ at which the ordered state loses stability, which is the focus of my work. Likewise, [101] studied the conditions under which disorder remains stable but did not examine the point where order becomes unstable. Their results suggest that when network density is very low, the transition from order to disorder becomes continuous. This aligns with the numerical results for $p = 0.1$, where the transition from paradise to disorder appears to lose its discontinuous nature.

I find that the normalized critical temperature $\tilde{T}_c$ scales approximately as p^2, with numerical fits suggesting a slightly larger exponent. In contrast, the critical temperature T_d found in [101], which marks the stability of disorder, follows a slower scaling pattern with an exponent of roughly 1.72. My results indicate that despite the structural randomness of ER networks, spontaneous order in structural balance can still emerge below a critical temperature. Additionally, I show that a properly adapted mean-field approach can effectively describe the system's critical properties, similar to methods used for complete graphs.

Overall, this study offers both theoretical and numerical insights into structural balance in ER networks. These findings may serve as a useful reference for future research on structural balance in more complex network topologies and real-world social systems.

Chapter 4

Structurally balanced growing network as randomized Pólya urn process

4.1 Motivation

In the previous two chapters, I employed a Hamiltonian formulation on the structural balance models that explore multi-layered networks with interlayer coupling and the resulting phase transitions between various balanced and disordered states. In this chapter, I focus on the network growth and evolution of faction sizes using Pólya urn process.

Understanding when and how polarization emerges in growing social systems is crucial for explaining the persistence of opposing factions and the conditions under which one group dominates or balanced coexistence prevails [105–107]. Many existing models capture aspects of polarization—such as homophily, tribalism, or structural balance- but often within fixed-size networks or with mechanisms operating at the microscopic level (link or individual) [82, 83]. This leaves open how collective, path-dependent dynamics shape faction sizes as the system grows.

Among these, homophily—the tendency of individuals to form positive connections with those who share similar opinions—plays a central role. Structural balance theory implies that polarization emerges in the presence of cycles where the product of polarities of all links along a cycle is positive [108]. When individuals preferentially connect among like-minded individuals (homophily), they create echo chambers that amplify existing divisions and reinforce group cohesion [109]. Such processes of selective association can lead to the development of social structures that can be described by structural balance theory, where the network can split into two factions: strong, positive relationships (friendships) are maintained within each group, while negative relationships (hostilities) dominate between groups [110].

The phenomena described above share a similarity in generating factions or factions. However, the extent to which the underlying processes dynamically affect the system, eventually driving polarization, remains unclear. It is also possible that initial impressions and visibility play decisive roles. Some studies report limited changes in alignment over time, highlighting the significant impact of early choices [84]. Building on these observations and motivated by the fundamental similarities across polarization phenomena, I propose the following abstraction: instead of modeling polarization within specific network topologies through individual-level interactions, I introduce a stochastic process that captures the essential dynamics of polarization at the group level. My approach is inspired by the path-dependent Pólya urn process and time-invariant Bernoulli trials. Together, these elements provide a minimal yet general framework for understanding the emergence, persistence, and relative dominance of opposing factions during network growth.

4.2 Models and methods

I consider the evolution of a network comprising undirected signed links, where each link $x_{i,j}$ between nodes i and j denotes either a friendly ($x_{i,j} = 1$) or hostile ($x_{i,j} = -1$) relationship. The network initialization begins at time $t = t_0$ with t_0 fully connected nodes labeled $i = 1, 2, \ldots, t_0$, arranged such that the entire graph satisfies the condition of structural balance. In this initial configuration, the agents are partitioned into two mutually antagonistic factions: intra-faction connections are positive (friendship), while inter-faction links are negative (hostility). As a special case, one of the factions may initially contain no nodes, resulting in a network with exclusively positive relations. Subsequently, at each time step $t > t_0$, a new node labeled $i = t$ is introduced. Its first connection is formed with a randomly selected existing node j, and the sign of this initial link is positive with probability $p \in [0, 1]$, or negative with probability $1 - p$. This parameter p governs the likelihood of forming a friendly initial relation and thus acts as a bias in the attachment process. Following this initial contact, the new node connects to all other existing nodes in the network. The signs of these new links are determined such that every triad formed with the new node is structurally balanced. Specifically, if the relation to node j is friendly, then the new node aligns positively with j's friends and negatively with j's enemies. Conversely, if j is perceived as hostile, the new node adopts the opposite pattern of alignment. As a result, the network remains a complete graph, with each new node establishing links to all others, and the structural balance condition is preserved at every time step. The network thereby continues to evolve as two hostile factions, internally connected by positive links and externally by negative ones.

This dynamic bears a strong resemblance to a generalized Pólya urn process. In this analogy, the two factions correspond to two colors of balls in the urn (e.g., white and black). At each step, one ball is drawn (i.e., one existing node is selected), and a new ball

is added: it matches the color of the drawn ball with probability p, or is of the opposite color with probability $1 - p$. In this framework, the attachment bias p serves as the reinforcement parameter, influencing the preferential expansion of the factions. Figure 4.1 represents a schematic illustration of this process.

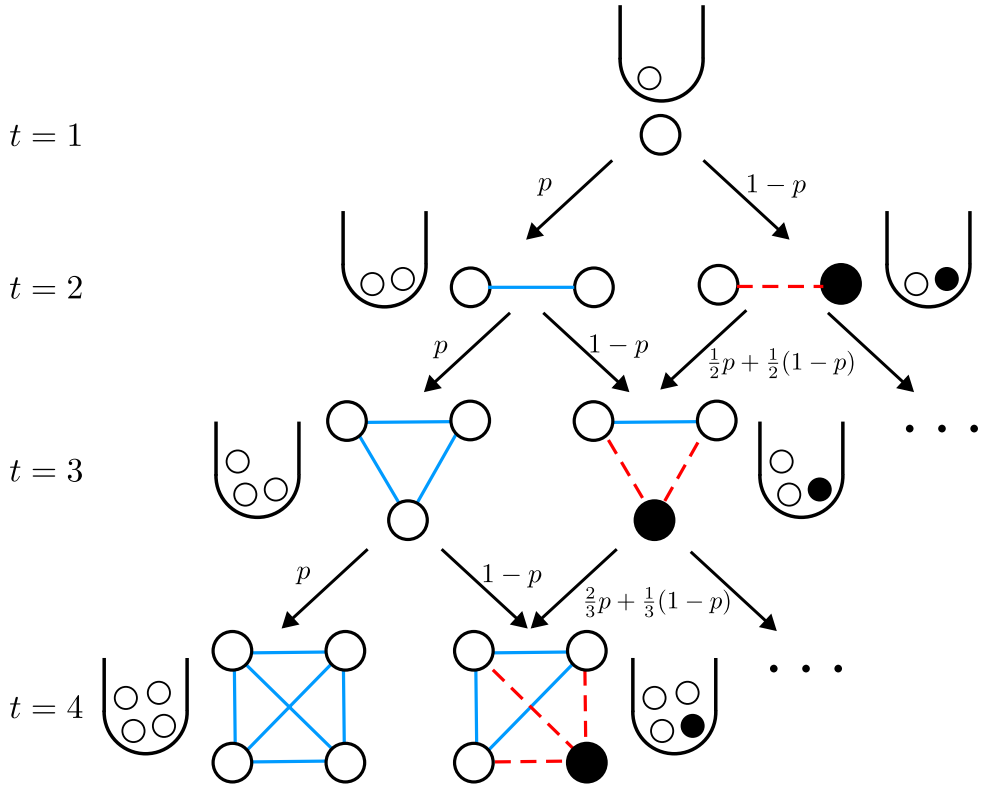


Figure 4.1: A schematic representation of the growing network is shown. The signed network grows according to structural balance and follows a generalized Pólya urn process. We start with a single node (white ball). At each time t, the complete graph divides into two mutually hostile factions. Solid blue lines represent friendly links, while dashed red lines indicate hostile links.

In the limiting case of $p = 1$, the process becomes partially deterministic: the new node invariably joins the faction of the selected reference node j. Over time, the initially larger faction is likely to maintain or increase its dominance. This behavior corresponds exactly to the classic Pólya urn process, where the long-term distribution of faction sizes converges to a beta distribution [111, 112], shaped by the initial configuration of the system. In contrast, setting $p = 0.5$ results in a fully unbiased process. Each faction has an equal probability of attracting the newly introduced node, leading to symmetric stochastic growth. For values $p < 0.5$, the model exhibits contrarian dynamics [113], whereby new agents are more inclined to oppose the stance of their initial contact. This *anti-bias* introduces a corrective tendency that favors the growth of the minority faction. In the intermediate regime, where $0.5 < p < p^{\text{ch}}$, the system displays a modest bias toward the dominant faction, where new nodes are more likely to align with the majority. As p approaches 1, the system enters a high-bias phase in which the reinforcement mechanism

strongly amplifies the growth of the dominant faction. The specific characteristics and implications of this high-bias regime, particularly for $p \in (0.5, 1)$, are examined in detail in the subsequent sections.

4.3 Evolution of mean faction sizes

In the model, the evolving network can always be partitioned into two antagonistic factions [108]. By construction, each faction contains only positive (friendly) links internally, while all links between the two factions are negative (hostile). One special case, often referred to as "paradise", occurs when one faction contains all nodes and the other is empty.

Let us denote by $m(t)$ the size of a given faction at time t. There are exactly two ways in which this faction can gain a new member when a node arrives at time $t + 1$:

1. A new agent's first (randomly chosen) link is **positive** and attaches to one of the m existing nodes in this faction.

2. A new agent's first link is **negative** and attaches to a node in the other faction, which has size $(t - m)$; in such a case, the structural-balance rule dictates that the new agent will form positive links to all members of the first faction and thus join it.

Since there are t nodes total at time t, the probability of case (1) is $p \cdot \frac{m}{t}$, and the probability of case (2) is $(1 - p) \cdot \frac{t - m}{t}$. Accordingly, the expected increment in the faction size satisfies

$$m(t + 1) - m(t) = p\frac{m(t)}{t} + (1 - p)\frac{t - m(t)}{t}. \tag{4.1}$$

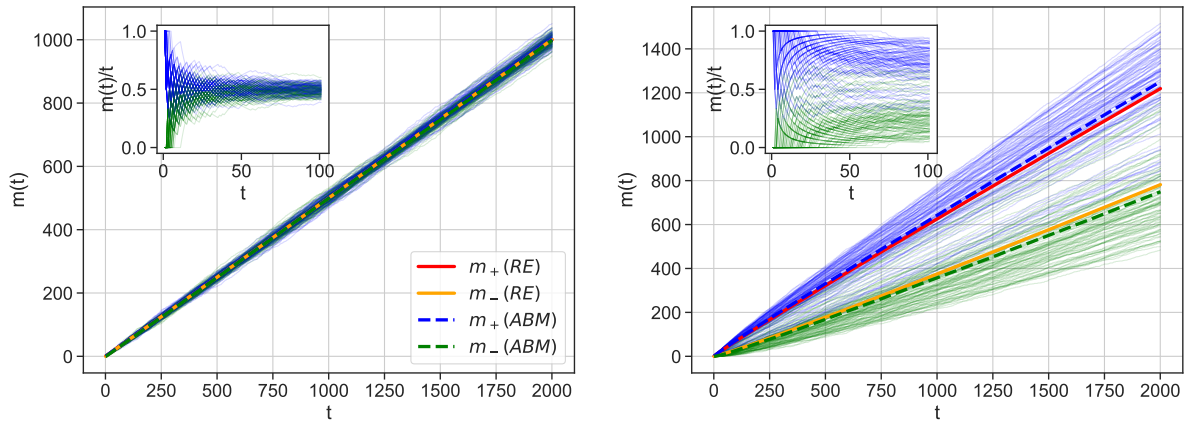
Passing to a continuous-time approximation (treating $m(t+1) - m(t) \approx \frac{dm}{dt}$), one obtains the differential equation whose general solution can be written as

$$m(t) = \frac{t}{2} + Ct^{2p-1}, \tag{4.2}$$

where C is an integration constant determined by the initial configuration. In other words, equation (4.2) describes the expected size of one faction as a function of time, and the constant C encodes how the network was seeded.

The value of the integration constant is determined by fixing the initial conditions. Without loss of generality, consider a network initialized at time t_0, consisting of two factions, $m_+(t_0)$ and $m_-(t_0)$, satisfying $m_+(t_0) + m_-(t_0) = t_0$. Given these initial conditions, the constant C from Eq. (4.2) is given by:

$$m_\pm(t_0) = \frac{t_0}{2} + Ct_0^{2p-1} \implies C = \frac{2m_\pm(t_0) - t_0}{2t_0^{2p-1}}. \tag{4.3}$$



(a) (b)

Figure 4.2: Evolution of sizes of two factions $m_\pm(t)$ for different attachment biases p. Expected faction sizes given by the rate equation (RE, solid lines) closely match the respective means of the agent-based model (ABM, dashed lines) simulation. (a) For a small value of $p = 0.3$, the mean sizes of the factions are equal. (b) For a large value $p = 0.9$, the means $\langle m_\pm(t)\rangle$ diverge. Thin lines show individual trajectories of 100 agent-based simulations of the two faction sizes. The initial condition was a single starting node belonging to m_+. The inset in each panel shows the corresponding normalized mean over time, which converges over a long time.

Thus, the expected faction size grows as

$$m(t) = \frac{t}{2} + \left(m_\pm(t_0) - \frac{t_0}{2}\right)\left(\frac{t}{t_0}\right)^{2p-1} \tag{4.4}$$

(see Fig. 4.2) and the expected difference in faction sizes is given by:

$$\Delta m(t) \equiv |m_+(t) - m_-(t)| = \Delta m(t_0)\left(\frac{t}{t_0}\right)^{2p-1}, \tag{4.5}$$

where $\Delta m(t_0) = |m_+(t_0) - m_-(t_0)|$ is the initial difference in faction sizes.

Another way of estimating the faction mean sizes exploits the connection between the presented stochastic process and the Pólya process. The standard Pólya process, in which the ball drawn from the urn determines the color of the newly added ball, has been generalized in various ways. A particularly relevant generalization for this work is described as follows [77]. After selecting a ball of color i, a set of additional balls $[c_1^i, c_2^i]$ is returned to the urn, where c_j^i denotes the number of balls of color j added in response to drawing the color i. In [77], the numbers c_j^i are integers sampled from a given distribution. I instead assume abstractly a continuous (*fractional*) growth version of the process, in which at each step, *fractional* numbers of balls (nodes), p and $1-p$, are added to the respective factions. I refer to this as the *fractional growth process*. Under this

assumption, with a single node initially, the expected mean faction sizes evolve as

$$m_{\pm}(t) = \frac{t}{2} \pm \frac{t^{2p-1}}{2\Gamma(2p)}, \tag{4.6}$$

where Γ is the gamma function.

This second estimate of the mean faction size better approximates mid-term and long-term behavior compared to the rate equation Rate equation (RE). Fig. 4.3 illustrates this comparison by showing the normalized mean $(\langle m_+ \rangle / t$, computed using Eq. (4.6), Eq.((4.2), and the exact mean obtained from the master equation Master equation (ME) (4.9) (derived in the following section). For anti-bias values of p, i.e., when $p < 0.5$, the mean rapidly converges to 0.5, whereas for larger p, it approaches this value asymptotically. The rate equation aligns with the master equation at short timescales, whereas the fractional growth Pólya process better approximates long-term behavior. It is worth noting, however, that Eq. (4.2) gives a general solution for any initial distribution of nodes (see Fig. 4.4), while Eq. (4.6) applies specifically to the case of a single initial node.

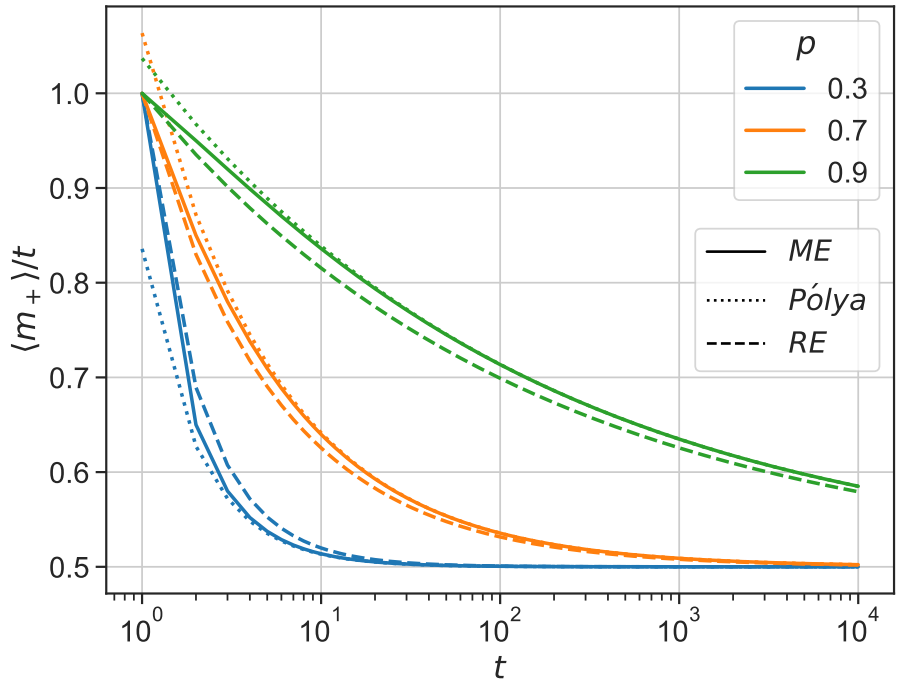


Figure 4.3: In time, the normalized group sizes approach equality. Convergence rates depend on the bias p. The system starts at $t_0 = 1$ with a single node belonging to group m_+. Exact results derived from the master equation (ME) (4.9) are shown as solid lines, while approximate values from the Pólya process formula (4.6) and the rate equation (4.4) (dashed) are shown as dotted and dashed lines, respectively.

Both expressions for $\Delta m(t)$—that in (4.5) and the one derived from (4.6)—exhibit the same scaling, $\Delta m \propto t^{2p-1}$. This leads to two interesting observations about the role of the

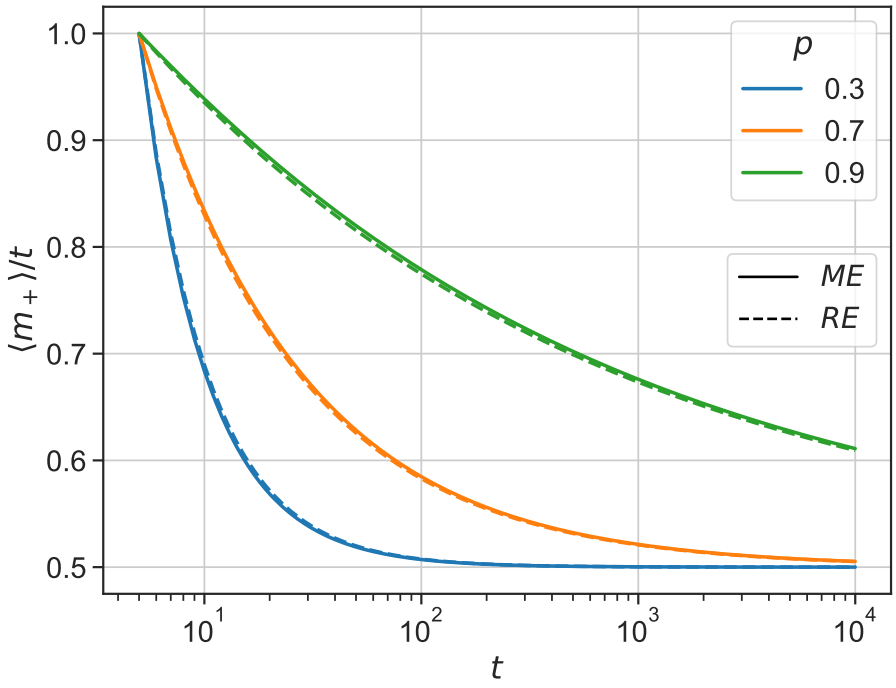


Figure 4.4: The rate equation (RE) closely matches results derived from the master equation (ME), also when initially, the system consists of five nodes in a single faction. This shows that the rate equation is general for any initial conditions. The plot shows the time evolution of the means for three different attachment bias p values.

control parameter p. First, for $p < 0.5$, Δm decreases to zero as the system size t increases (see Fig. 4.2a), whereas for $p > 0.5$, it diverges as t grows (see Fig. 4.2b). This indicates that $p = 0.5$ constitutes the first critical point governing the system behavior. Second, for all biases $p < 1$, the normalized value of this difference, $\delta m = \Delta m/t$, tends to 0 when $t \to \infty$. This can be seen in Fig. 4.3, where all expected normalized means converge universally to 0.5. This result has one very interesting implication: When the system size grows, it becomes harder and harder to determine which is the dominating faction. To explore the origin of this phenomenon, I next examine the *probability distributions* of the faction sizes and their evolution.

4.4 Size distribution of evolving factions

The expression given in Eq. (4.4) provides the expected growth trajectory for the size of each faction over time. However, since the model incorporates stochastic elements—stemming from both the random selection of attachment targets and the probabilistic bias parameter p—this mean-field description is insufficient to capture the full dynamics of the system. To gain a more comprehensive understanding, I now turn our attention to the underlying probability distributions governing faction sizes during network evolution. This approach parallels methodologies employed in related contexts, such as the analysis

of opinion dynamics within the Weidlich model [114]. Let $P(m,t)$ denote the probability that a given faction has size m at time t. To evaluate how this distribution evolves, I consider two mutually exclusive events through which a faction can reach size m at time $t+1$. This can occur in two ways: (1) the faction size was $m-1$ at the previous time step, and a new node is added at the next time step, or (2) the faction size was m at time t, and no nodes are added to the faction. In the first case, the new node is added with a rate $w(m|m-1)$ and in the second case – with a rate $w(m+1|m)$. Thus probability $P(m,t+1)$ is given by the sum of

- the probability there were $m-1$ nodes at time t and the probability to increase the faction size by one in Δt:

$$P(m-1,t)\cdot w(m|m-1)\Delta t \tag{4.7}$$

- the probability there were m nodes at time t and the probability that no increase in the faction size will occur:

$$P(m,t)\cdot\left(1-w(m+1|m)\right)\Delta t \tag{4.8}$$

Then, the master equation for the process gives the probability $P(m,t+1)$:

$$\begin{aligned} P(m,t+1) = {} & w(m\mid m-1)\cdot P(m-1,t) \\ & + w(m\mid m)\cdot P(m,t), \end{aligned} \tag{4.9}$$

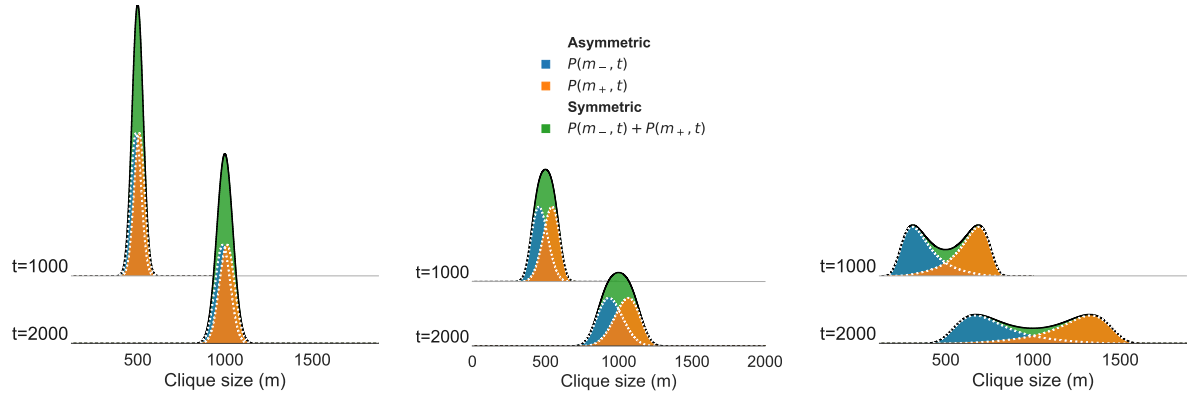
where the transition rates are given by,

$$\begin{aligned} w(m+1\mid m) &= \frac{m(2p-1)}{t}+1-p, \\ w(m\mid m) &= 1-w(m+1\mid m). \end{aligned} \tag{4.10}$$

The master equation (4.9) governs the evolution of the probability distribution, which remains inherently non-stationary due to the ongoing growth process. The shape of this distribution depends both on the parameter p and the chosen initial conditions. In the following discussion, I explore how different initial setups lead to either unimodal or bimodal distributions. In particular, I distinguish between **asymmetric** initial conditions, where the two initial factions differ in size, and **symmetric** conditions, where they begin statistically equivalent.

I begin with the asymmetric initial condition $P(m,t=1)=\delta_{m,1}$, which corresponds to the process starting with a single node in one faction while the other remains empty. The initially larger faction is denoted as m_+, and the smaller as m_-. Under this condition, the resulting distributions $P(m_+,t)$ and $P(m_-,t)$ remain unimodal across all values of p

(as indicated by the dashed orange and blue curves in Fig. 4.5).



(a) $p = 0.7$ (b) $p = 0.8$ (c) $p = 0.9$

Figure 4.5: Evolution of the three faction size distributions for three values of the attachment bias parameter p. Each panel shows the exact distributions derived from the master equation, starting from a single-node system. Two distributions (with dotted outlines) in each panel correspond to an asymmetric initial condition, where the starting node belongs to the faction m_+, while the other group m_- is empty. The blue and orange fill are used for the resulting distributions for the initially smaller faction $P(m_-, t \mid m_-^{t_0=1} = 0)$ and the larger one $P(m_+, t \mid m_+^{t_0=1} = 1)$, respectively. Solid outline with green fill represents the distribution under a symmetric initial condition, where the first node is equally likely to join any faction. It also reflects the case where no specific group is observed, and one asks for the probability that any faction is of size m. That is why it is presented as $P(m_-, t) + P(m_+, t)$. (a) At $p = 0.7$, all distributions are unimodal and approximately Gaussian. (b) At $p = 0.8$, although the asymmetric distributions diverge, the combined result remains unimodal, and no bimodal behavior is observed. (c) At $p = 0.9$, the symmetric distribution becomes distinctly bimodal, with the two peaks gradually diverging as new nodes are added.

Next, I examine the symmetric initial condition $P(m, t = 1) = \frac{1}{2}\delta_{m,1} + \frac{1}{2}\delta_{m,0}$, representing a system where the first node joins either faction with equal probability. Here, I observe one particular faction without specifying whether it is initially dominant. Using the analogy of two factions, "black" and "white," the first agent joins one of them randomly, resulting in a distribution symmetric around $t/2$. As the system evolves, this inherent uncertainty can lead to bimodality, where each peak represents a distinct dominant outcome. For sufficiently large p, these peaks become well separated (see Fig. 4.5).

An alternative interpretation of the symmetric case is that it reflects the probability of observing a faction of size m in a system originally initialized asymmetrically. In other words, starting from the asymmetric setup, allowing the system to evolve, and then randomly selecting one faction (black or white) yields the symmetric distribution. Since one faction evolves from $P(m, t = 1) = \delta_{m,1}$ and the other from $P(m, t = 1) = \delta_{m,0}$, their combined, equally weighted distribution reproduces the symmetric case. Thus, the symmetric distribution can be seen as the sum of the two asymmetric cases (see the legend in Fig. 4.5).

By comparing the distributions at $t = 1000$ and $t = 2000$ (top vs bottom panels in Fig. 4.5), I observe the continuous growth of the factions as new nodes are added. Under asymmetric initial conditions (blue and orange curves), the peaks shift rightward over time, indicating an increase in the mode of each distribution. The overlap between them gradually decreases, suggesting that one faction grows faster than the other. In the symmetric case (green curve), the distribution appears unimodal for small p and becomes distinctly bimodal for larger p. For small p, the peaks shift rightward but remain merged [Fig. 4.5(a),(b)], while for large p, two clearly separated peaks emerge and move apart as time progresses [Fig. 4.5(c)].

These results reveal two noteworthy observations. First, under asymmetric initial conditions, the mean size difference between factions, derived from Eq. 4.5, increases over time when $p > 0.5$. This indicates that the mean sizes of the larger and smaller factions diverge as p surpasses this threshold, which might suggest the onset of bimodality. However, the normalized difference between their mean sizes vanishes for all $p < 1$. Second, for symmetric initial conditions, the distribution (in both absolute and relative terms) is expected to become bimodal for $p > 0.5$, but this transition is not apparent until p is sufficiently large, as shown in Fig. 4.5. To reconcile these findings, I must establish a quantitative criterion for identifying bimodality and examine how expected faction sizes relate to their fluctuations.

4.4.1 Statistical measures of the faction size distribution

The evolution of faction sizes in the model can be understood through their underlying statistical properties. As the system grows and the bias parameter p varies, the faction size distributions undergo qualitative changes that can be quantified using measures such as skewness and kurtosis. These statistical descriptors provide insight into how asymmetry and modality emerge in the system, reflecting transitions from balanced to polarized states.

Because the two opposing factions evolve symmetrically around the midpoint $t/2$, their size distributions satisfy the relation $P(m_+, t) = P(t - m_-, t)$. This symmetry allows the construction of a unified probability distribution that characterizes the likelihood of observing a faction of a given size, independent of its identity. In other words, the combined distribution represents the overall probability that a randomly chosen faction—regardless of group affiliation—attains a specific size. It can be obtained by summing the two individual distributions and, if necessary, normalizing by a factor of $\frac{1}{2}$ to preserve total probability (see Fig. 4.6). This approach leads to the same functional form as the distribution derived under symmetric initial conditions, $P(m, t = 1)$. Figure 4.5 illustrates this correspondence, showing the distributions for the larger and smaller factions, as well as twice the combined distribution. For sufficiently large values of the model parameter p,

the combined distribution becomes bimodal, developing two distinct peaks (see Fig. 4.5c).

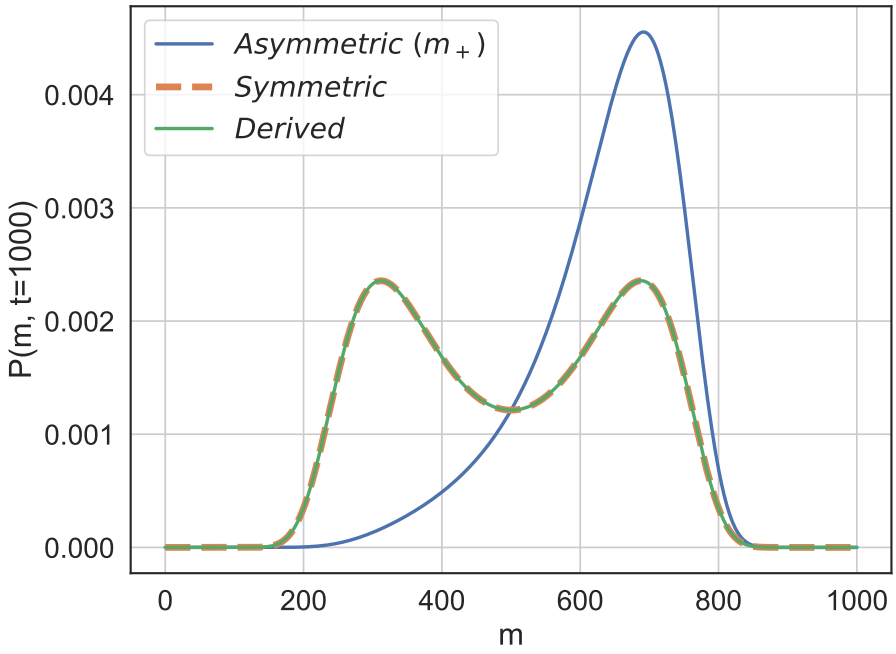


Figure 4.6: A bimodal distribution can be formed by adding the individual distributions of both factions through a reflection-based transformation. At time t, the distribution of one faction (blue line) is reflected about the midpoint $m/2$ to represent the distribution of the opposing faction. The two are then averaged, scaling the sum by a factor of $\frac{1}{2}$ to preserve normalization, resulting in a bimodal distribution (green line) with two distinct peaks. For comparison, the dashed orange line shows the exact distribution obtained under the symmetric initial condition, demonstrating the equivalence of both construction methods.

Both unimodal and bimodal forms of the distribution emerge dynamically as new nodes are added to the system. In the unimodal regime, the peak gradually shifts to larger sizes, indicating a steady growth of the dominant faction. Conversely, in the bimodal regime, the probability mass separates into two modes, reflecting polarization between the groups. The increasing asymmetry of the unimodal distribution with higher values of p can be quantified using the third standardized moment (skewness), denoted by α_3 [115]:

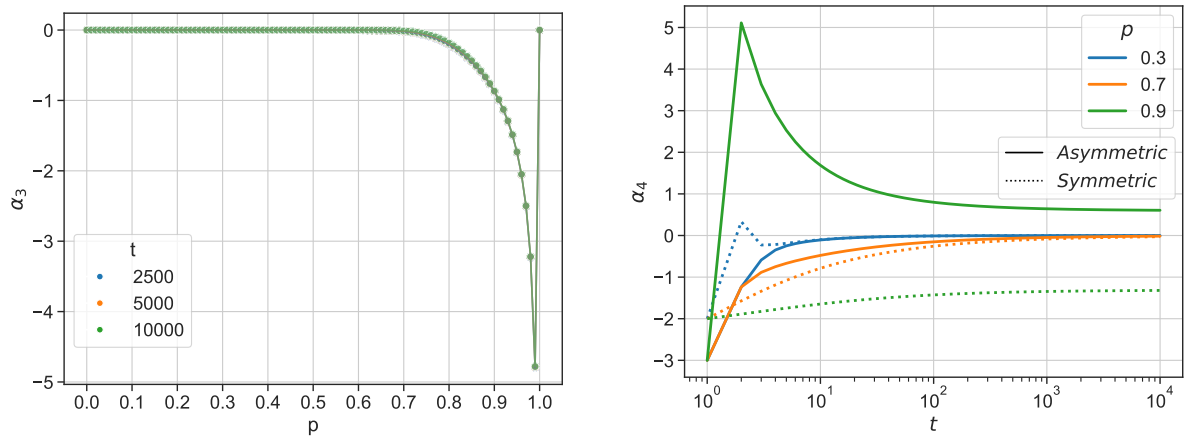
$$\alpha_3 = \mathbb{E}\left[\left(\frac{X - \langle m_+ \rangle)}{\sigma_+}\right)^3\right], \tag{4.11}$$

where σ_+ is the standard deviation of $P(m_+, t)$. Because of the mirror symmetry $P(m_+, t) = P(t - m_-, t)$, the skewness of the smaller faction is $-\alpha_3$. This reciprocal relationship highlights the complementary evolution of the two distributions and facilitates the formation of a symmetric, identity-independent combined distribution. As shown in the Fig. 4.7a, α_3 is increasingly negative with the increase of bias. It becomes visibly nonzero for $p \approx 0.75$, yet α_3 is never exactly 0, but it is a monotonously decreasing function of p.

The shape of the faction size distribution can further be characterized by its kurtosis, which measures the sharpness or flatness of the distribution relative to a Gaussian form. When the system transitions from a unimodal to a bimodal state, the probability mass becomes concentrated around two distinct peaks, effectively flattening the central region of the distribution [116]. This effect can be partially quantified using the excess kurtosis α_4 [117, 118], defined as

$$\alpha_4 = \mathbb{E}\left[\left(\frac{X - \langle m \rangle)}{\sigma_0}\right)^4\right] - 3. \tag{4.12}$$

where $\langle m \rangle = t/2$ is the mean and σ_0 is the standard deviation of the symmetric distribution. Negative values of α_4 typically indicate a broad or flat-topped shape and may suggest the emergence of bimodality, though not conclusively. In this model, increasing bias p drives the distribution toward larger absolute values of α_4, reinforcing the development of distinct modes as the peaks become more pronounced [116, 118]. The transition from unimodal to bimodal behavior occurs near a critical value of p, which is analyzed in later sections. Figure 4.7b presents the variation of excess kurtosis with the bias parameter p under both symmetric and asymmetric initial conditions. For asymmetric initialization, α_4 remains close to zero at low or anti-bias levels but becomes positive as p increases, indicating sharper, more peaked distributions. Under symmetric conditions, however, α_4 becomes negative at high bias, a signature commonly associated with bimodal behavior [117]. Although negative kurtosis alone does not definitively establish bimodality, in this symmetric model, it provides a useful statistical indicator of the onset of polarization [119].



(a) (b)

Figure 4.7: Statistical measures quantify key features of asymmetric and symmetric distributions. Panel (a) shows that skewness α_3 signals asymmetry in the high-p regime, becoming appreciable only for $p \gtrsim 0.75$, with minimal time dependence. Results are obtained from the master equation under asymmetric initial conditions. Panel (b) shows excess kurtosis α_4 converging over time, influenced by p and initial conditions. For symmetric initial conditions, negative α_4 indicates the emergence of bimodality.

4.4.2 Bimodality criterion

Under symmetric initial conditions, the faction size distribution is theoretically expected to exhibit bimodal characteristics for values of $p > 0.5$. However, as illustrated in Fig. 4.5, this bimodality does not emerge immediately and becomes apparent only at significantly higher p values. To understand this discrepancy, the distribution can be viewed as a combination of two overlapping components corresponding to the smaller (m_-) and larger (m_+) factions. This mixture representation provides a more accurate description of the overall behavior observed under symmetric initial conditions.

In the context of Gaussian mixtures, the emergence of bimodality can be evaluated based on the separation of component means relative to their spread (standard deviation). According to the classical result of [120], a mixture of two Gaussian distributions with means μ_1 and μ_2 and standard deviations σ_1 and σ_2 becomes bimodal if:

$$|\mu_1 - \mu_2| > 2 \min(\sigma_1, \sigma_2). \tag{4.13}$$

Although this relation does not strictly apply to the considered model, a qualitatively similar condition determines whether bimodality becomes visible. Specifically, the degree of separation between the mean faction sizes, compared to the distribution's overall width, dictates whether two distinct peaks can be resolved. In this case, the component distributions are symmetric about $t/2$ and share an identical standard deviation, $\sigma_+ = \sigma_- \equiv \sigma$.

From theory, the mean sizes of the larger and smaller factions start to diverge once $p > 0.5$, as indicated by Eq. 4.4. This produces an increasing difference $\Delta m = |m_+ - m_-|$, which would appear to favor bimodality. However, this divergence alone is insufficient to yield a bimodal distribution. In practice, solutions to the master equation under symmetric initial conditions remain unimodal for a wide range of $p > 0.5$. Bimodality arises only when the mean separation exceeds a certain multiple of the standard deviation. Using a heuristic form of Eq. (4.13), I obtain the condition $|m_+ - m_-| > 2\sigma$, with the midpoint $t/2 = (m_+ + m_-)/2$ representing the center between peaks. Since the network growth follows a Poisson-like process, the standard deviation scales as $\sigma \sim \sqrt{t}$. Thus, the bimodality condition becomes:

$$t^{(2p-1)} > 2\sqrt{t} \quad \Leftrightarrow \quad t^{2p-3/2} > 2. \tag{4.14}$$

In the asymptotic limit of large t, this simplifies to:

$$2p - \frac{3}{2} > 0 \quad \Rightarrow \quad p > \frac{3}{4}. \tag{4.15}$$

Therefore, for $p > 0.75$, the peak separation grows faster than the width of the distribu-

tion, leading to clearly visible bimodality. For $p < 0.75$, the separation remains too small relative to σ, resulting in a unimodal distribution. Consequently, $p = 0.75$ serves as an approximate analytical threshold separating unimodal and bimodal regimes. Nevertheless, as shown in subsequent analysis, this value underestimates the actual point at which bimodality becomes numerically evident.

The parameter p quantifies the tendency of new nodes to reinforce existing factional structures. When $p = 1$, the system evolves into a beta distribution of faction sizes determined by the initial conditions. If the network begins with two equally sized groups, the final distribution remains centered, implying that the system shows no preference for either faction. Conversely, for $p < 1$, the group sizes tend to converge over time, with the convergence rate depending on p (see Fig. 4.3).

A straightforward way to test for bimodality in the symmetric case is to evaluate the concavity of the distribution between the two potential maxima. This can be done by computing the second derivative of the distribution at the midpoint between the peaks. A positive second derivative at this midpoint indicates a local minimum, confirming bimodality. As shown in Fig. 4.8, the midpoint curvature becomes positive for $p \gtrsim 0.836$, signaling the emergence of bimodality. The precise point where the second derivative changes sign shifts slightly with time.

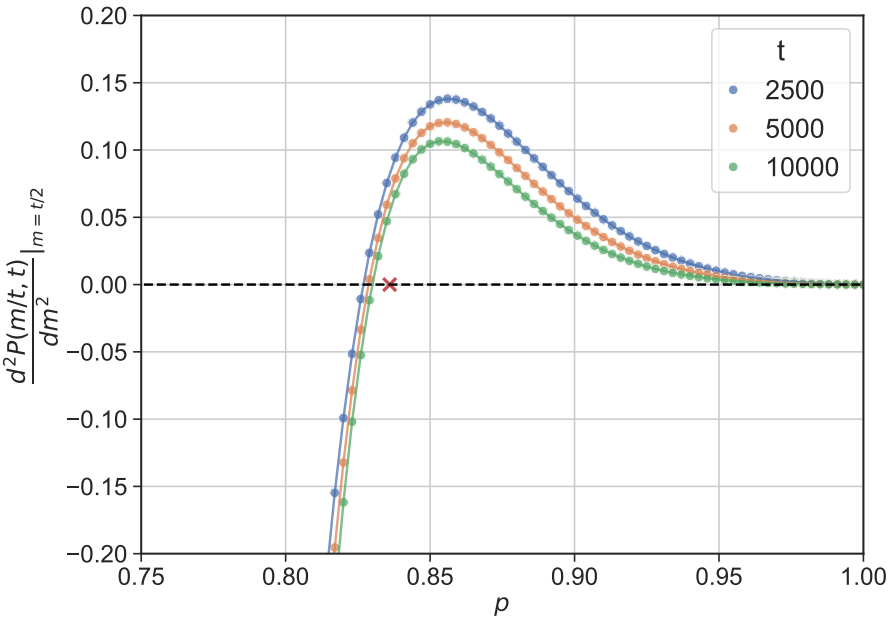


Figure 4.8: Observing bimodality. The second derivative of the distribution evaluated at $m = t/2$ serves as an indicator of bimodality; a positive value signifies the presence of a bimodal distribution. For symmetric initial conditions, bimodality persists in the limit $t \to \infty$ when the attachment bias p exceeds the characteristic value $p^{ch} \approx 0.836$ indicated by a red cross.

The discrepancy between the analytical estimate ($p = 0.75$) and the numerically observed onset of bimodality raises an important question: does $p = 0.75$ represent a physi-

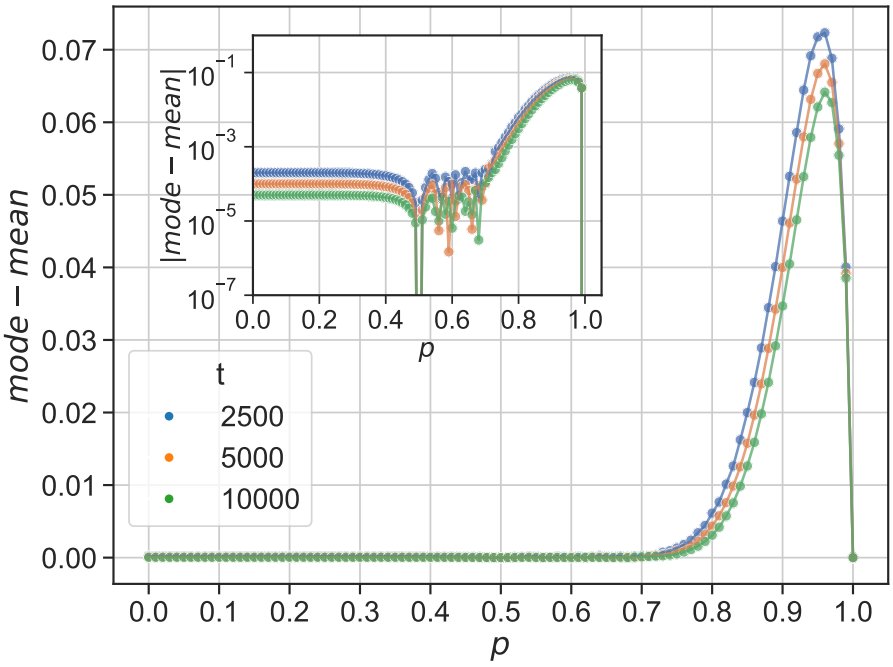


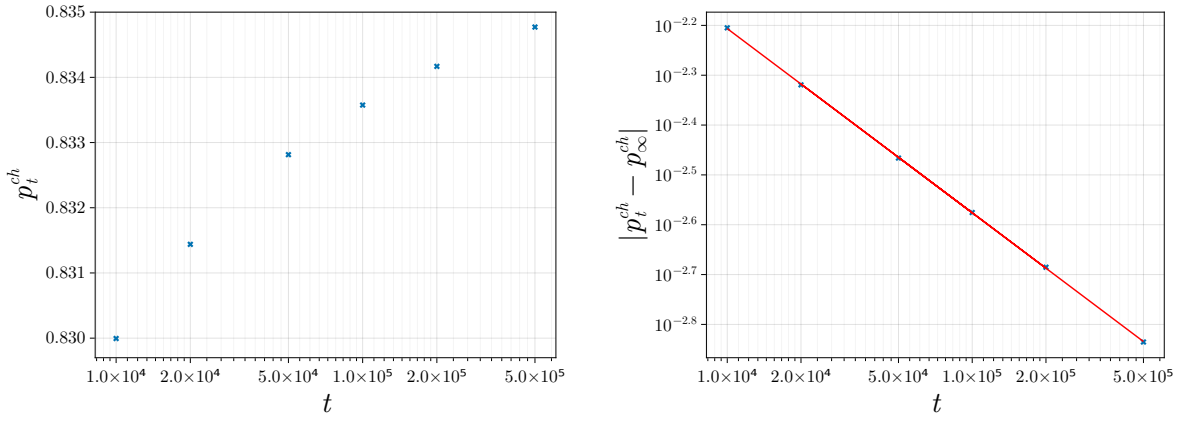
Figure 4.9: Distribution asymmetry is a signature of the high p regime. The figure shows the difference between the normalized mode and the normalized mean, with the inset presenting its absolute value on a logarithmic scale. For $p \gtrsim 0.75$, the difference becomes significantly larger than 0, signaling the beginning of the high-bias regime. All three distinct regimes of p are clearly evident in the inset. For $p < 0.5$, the normalized mode-mean difference behaves smoothly, remaining close to $1/2t$ due to the discrete nature of the distribution, whereas for $0.5 < p < 0.75$, it exhibits fluctuations. These trends are approximately independent of time.

cally meaningful threshold, or is it simply an analytical artifact? Figure 4.9 suggests the former. While clear bimodality is only observed for $p \gtrsim 0.836$, qualitative changes in the distribution's structure already appear around $p = 0.75$, including increasing asymmetry in faction sizes—features not captured by bimodality alone. Figure 4.9 further emphasizes this asymmetry through the normalized difference between the mode and mean of the distribution. This difference remains negligible for $p \lesssim 0.75$, then rises sharply, peaking near $p \approx 0.95$, before dropping back to zero at $p = 1$, where all new nodes join the same initial faction. In the anti-bias regime, this difference theoretically vanishes but remains slightly positive due to finite-size effects. For $p \lesssim 0.75$, the mode and mean are nearly coincident; however, on a logarithmic scale, an additional behavioral transition emerges between the anti-bias and low-bias regions. For $p < 0.5$, asymmetry follows a smooth, discrete-like behavior near $1/2t$, whereas in the intermediate range $0.5 < p < 0.75$, fluctuations appear without a consistent trend.

Collectively, these findings indicate that $p \approx 0.75$ does not mark the true onset of bimodality but instead signals the beginning of a gradual transition in system behavior. The difference between the mode and mean effectively captures all three regimes of the bias parameter p, delineating the shift from symmetric convergence through intermediate

asymmetry to strong factional reinforcement.

Asymptotic characteristic bias



(a) (b)

Figure 4.10: Convergence of characteristic bias p^{ch} over time. Panel (a) shows that the characteristic bias p_t^{ch} increases monotonically. Panel (b) depicts the relation between the absolute difference $|p_t^{ch} - p_\infty^{ch}|$ and time t on a double logarithmic scale. The fitted line follows a power-law scaling, $\log|p_t^{ch} - p_\infty^{ch}| \sim -\log t$. The asymptotic value p_∞^{ch} was determined by maximizing the coefficient of determination R^2 of the fit. For the obtained fit, $R^2 > 0.9999$ and $p_\infty^{ch} > 0.835$. The results are shown when starting from a single-node symmetric initial condition.

The second derivative of the distribution offers a direct and effective criterion for detecting the emergence of bimodality by assessing the concavity at the center of the distribution. Specifically, evaluating the second derivative at the midpoint $m = t/2$, i.e.,

$$\left.\frac{d^2P(m,t)}{dm^2}\right|_{m=t/2},$$

allows us to distinguish between unimodal and bimodal distributions. A negative second derivative at this point indicates concavity (a local maximum), consistent with unimodality, while a positive value suggests convexity (a local minimum), indicating the presence of two separate peaks—hence, bimodality. For each time step t, I estimate the corresponding characteristic bias p_t^{ch} by identifying the values of p at which the second derivative changes sign. Specifically, p_t^{ch} is obtained via linear interpolation between the two adjacent values of p that bracket the zero-crossing of the second derivative. The resulting sequence of values for p_t^{ch}, computed for various values of t, is plotted in Fig. 4.10a. To determine the asymptotic characteristic bias p_∞^{ch}, I analyzed the convergence of p_t^{ch} by fitting a power-law relation

$$\log|p_t^{ch} - p_\infty^{ch}| \sim -\log t, \tag{4.16}$$

where p_∞^{ch} is chosen to maximize the coefficient of determination R^2 of the linear fit. As a result, the value of the asymptotic characteristic bias is $p_\infty^{ch} \approx 0.836$. The quality of the fit and the resulting p_∞^{ch} are shown in Fig. 4.10b.

Characteristic bias based on observability criterion

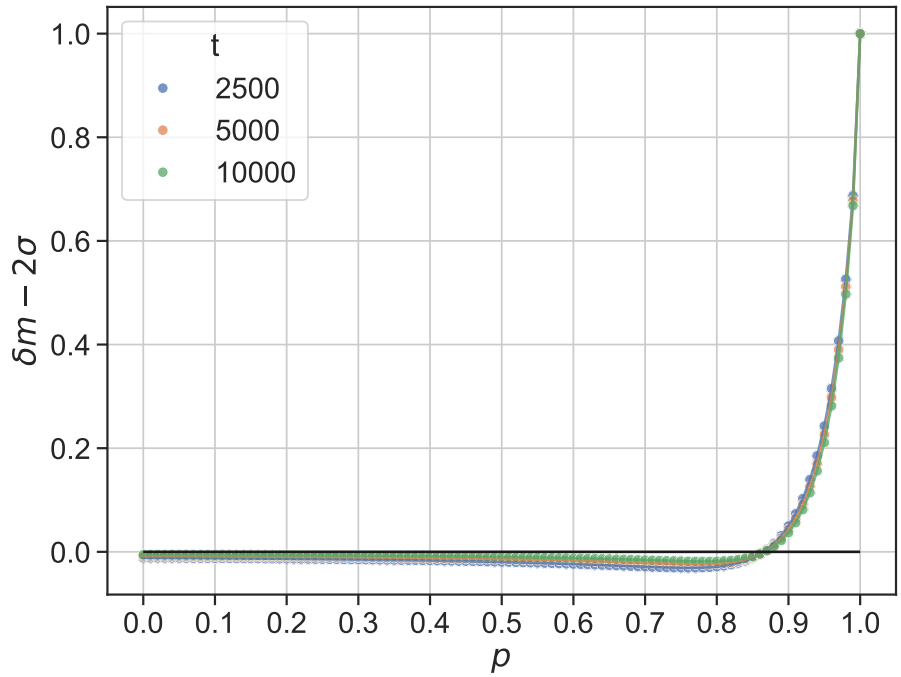


Figure 4.11: A criterion based on Gaussian approximation (4.13) predicts the emergence of bimodality when the separation between the normalized peaks $\delta m = (m_+ - m_-)/t$ exceeds twice the standard deviation σ of the normalized distribution under the asymmetric initial condition. For values of attachment bias $p > 0.86$, this condition is met, suggesting the potential for a bimodal distribution if symmetric initial conditions are applied. When the difference is negative, the peaks are expected to merge, resulting in a unimodal distribution.

A criterion discussed in Sec.4.4.2 offers an approximate method for determining the modality (unimodal vs. bimodal) of the combined distribution constructed from two independent distributions. The core idea relies on analyzing the sign of the expression $\delta m - 2\sigma$, where $\delta m = \Delta m/t$ represents the normalized difference between the sizes of the larger and smaller factions, and σ is the standard deviation of the normalized distribution (see Fig.4.11). When $p \gtrsim 0.86$, this expression becomes positive, suggesting that the two peaks are separated sufficiently to avoid overlap. In such cases, the distribution becomes bimodal. Conversely, for $p \lesssim 0.86$, the expression is negative, indicating that stochastic fluctuations are large enough to blur the separation between the modes, resulting in an effectively unimodal distribution. This result, however, appears slightly inconsistent with the characteristic value $p^{ch} \approx 0.836$ obtained numerically in Sec. 4.4.2, where a detailed analysis of the second derivative of the distribution identified the onset of bimodality at

a lower threshold. This discrepancy stems from the fact that the $\delta m - 2\sigma$ criterion is based on an approximation valid primarily for Gaussian distributions. In contrast, the true distributions in the model deviate from Gaussianity, especially near the transition regime. Despite this limitation, the criterion remains practically useful.

4.4.3 Fluctuations around mean sizes of factions

In deriving the equations connecting the Pólya process with the growing network model, I broadly adopt the formalism developed in [77]. In this framework, a replacement matrix A is introduced, where each element a_{ij} specifies the number of balls of color j added upon drawing a ball of color i. In general, these elements may be random variables drawn from a specified distribution. For the two-color case relevant to the model, the matrix A is given by:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

The long-term behavior of the process is governed by the eigenvalues of this matrix. The largest eigenvalue λ_1 is real and positive, which determines the leading-order growth of the system. For branching-type processes such as the Pólya urn, the variance and limiting distribution depend sensitively on whether the second eigenvalue λ_2, which satisfies $\mathrm{Re}(\lambda_2) > \lambda_1/2$. In such cases, the mean grows as t^{λ_1}, while the variance scales as $t^{2\lambda_2}$.

The growing network model considered in this study is equivalent to a two-color Pólya process with symmetric cross-color interactions. That is, the probability of adding a ball of the opposite color—relative to the one drawn—is the same for both colors. This corresponds to the replacement matrix:

$$A = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix},$$

where $p \in (0, 1)$ is the attachment bias.

This setting corresponds to the "Randomized play-the-winner" scheme (Example 7.3 in [77]), originally introduced in [121]. The eigenvalues of this matrix are $\lambda_1 = 1$ and $\lambda_2 = 2p - 1$. Applying the condition $\lambda_2 > \lambda_1/2$ identifies a threshold at $p = 0.75$, above which the variance increases more rapidly and the limiting distribution changes qualitatively.

For this class of processes, Theorems 3.22 and 3.23 in [77] provide asymptotic results for $p \leq 0.75$, yielding equal expected faction sizes and enabling closed-form expressions for the variance.

For $p > 0.75$, the process does not admit a closed-form limit distribution. In this regime, I apply Theorem 3.24 from [77], using the deterministic setting described in Example 3.27, where all entries of the replacement matrix A are constants (i.e., the

number of balls added at each step is deterministic rather than random). This framework enables the derivation of the limiting distribution and corresponding expected values as presented in the main text:

$$t^{-(2p-1)}\left(m_{\_} \pm -\frac{t}{2}\right) \xrightarrow{\text{a.s.}} \pm\frac{1}{2}\hat{Z}, \tag{4.17}$$

where $\hat{Z}$ is a distribution. When the process begins with a single ball of a given color, the moments of $\hat{Z}$ can be computed using Theorem 3.26 in [77]:

$$\mathbb{E}\hat{Z}^k = \frac{1}{\Gamma\left(1+k(2p-1)\right)}\mathbb{E}Z^k, \tag{4.18}$$

where following Example 3.13 of the same work [77]. For this case, it holds that $\mathbb{E}[Z] = 1$, and the variance is given by $\mathrm{Var}(Z) = \dfrac{(2p-1)^2}{4p-3}$. This allows one to compute the second moment $\mathbb{E}[Z^2]$, and subsequently derive $\mathbb{E}[\hat{Z}]$, $\mathbb{E}[\hat{Z}^2]$, and finally $\mathrm{Var}(\hat{Z})$.

Summing up, the expected variance of faction sizes as a function of the attachment bias p and time t is given by:

$$\sigma^2(p,t) = \begin{cases} \dfrac{t}{4(3-4p)}, & \text{for } p < 0.75, \\ \dfrac{t\log t}{4}, & \text{for } p = 0.75, \\ \dfrac{t^{2(2p-1)}}{2}B, & \text{for } p > 0.75. \end{cases} \tag{4.19}$$

where $B = \frac{1}{\Gamma(4p-1)} \cdot \frac{4p^2-2}{4p-3} - \left(\frac{1}{\Gamma(2p)}\right)^2$. Note that the estimates for $p < 0.75$ and $p > 0.75$ should not be used close to $p = 0.75$. These estimates diverge at $p = 0.75$ while the simulated values are continuous and very close to the analytical one computed for this specific p.

Figure 4.12 compares these theoretical predictions with the exact variances obtained from the master equation. For $p \lesssim 0.65$, the theoretical expressions match the true variance closely. However, a noticeable divergence emerges as p approaches the threshold value of 0.75, where the analytical solutions for the $p < 0.75$ and $p > 0.75$ regimes no longer align smoothly. Consequently, discrepancies between theory and exact values increase near this point. Interestingly, the expression derived specifically for $p = 0.75$ provides variance estimates that deviate only slightly from the exact results, comparable in magnitude to the discrepancies observed for larger values such as $p \gtrsim 0.8$. The discrepancy between analytics and simulation results arises from the breakdown of assumptions made in the theoretical approximations. For $p < 0.75$, the analytical expression relies on a central limit theorem-type behavior, assuming that fluctuations scale linearly with time,

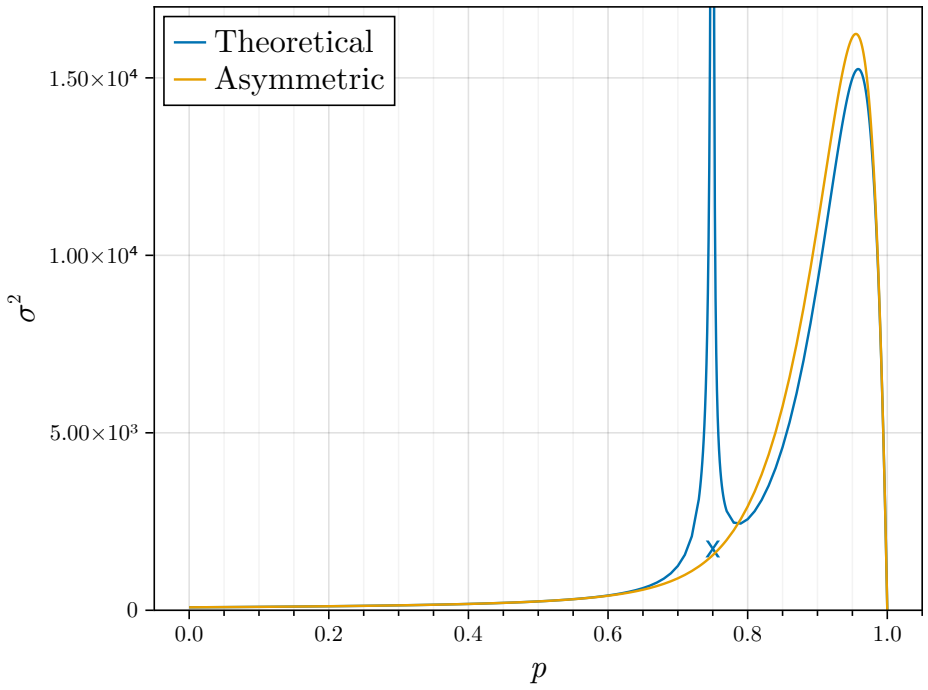


Figure 4.12: Variance of the faction size under asymmetric initial conditions with a single starting node for different values of p. The theoretical variance (blue line) calculated using Pólya process analysis (Eq. 4.19) diverges at $p \to 0.75$, which is inconsistent with the true variance (orange line) obtained using the master equation. The plot shows the variance after $t = 1000$ steps.

i.e., as t. For $p > 0.75$, a different scaling is assumed, with fluctuations growing superlinearly as $t^{2(2p-1)}$. Both assumptions hold well away from the critical point. However, near the "critical" value $p = 0.75$, they break down: the transition between the linear and superlinear scaling is not abrupt but smooth, and neither scaling form captures the correct fluctuation behavior precisely. Therefore, $p = 0.75$ should not be interpreted as a true critical point of the system, but rather as a threshold marking the onset of a gradual transition in the system's behavior.

Using the estimates for the variance (4.19), I compute how the signal-to-noise ratio (Signal to noise ratio (SNR)), i.e., $\Delta m(t)/\sigma(t)$, evolves over time:

$$\frac{\Delta m(t)}{\sigma(t)} \sim \begin{cases} t^{2p-1.5}, & \text{for } p < 0.75, \\ \dfrac{1}{\sqrt{\log t}}, & \text{for } p = 0.75, \\ D, & \text{for } p > 0.75. \end{cases} \tag{4.20}$$

where $D = 2/(\Gamma(2p)B)$, and Δ_m was obtained from (4.6). From these scaling relations, I find that for $p < 0.75$, the signal-to-noise ratio goes to zero for $t \to \infty$ as $2p - 1.5 < 0$.

This indicates that for $p < 0.75$, it becomes harder and harder to detect any difference in the faction sizes as the system grows. For $p = 0.75$ a similar finding holds, but the decrease in the signal-to-noise ratio is very slow. For $p > 0.75$, the signal-to-noise ratio is a constant, indicating that the difference in the faction sizes are more stable and detectable.

These analytical expectations are confirmed by simulations (see Fig. 4.13) and have one interesting consequence. Figure 4.9 shows the normalized difference between the distribution's mode and mean. Recalling that the normalized mean converges to 0.5 (see Fig. 4.3), this difference tells us how bigger (or smaller) is the dominating (dominated) faction. This difference remains negligible for $p \lesssim 0.75$, then increases sharply, peaking near $p \approx 0.95$, before falling back to zero at $p = 1$, where all new nodes join the single, starting one. Moreover, in the anti-bias regime $p < 0.5$, the difference is theoretically expected to approach zero, but, due to finite-size effects, it remains slightly greater than zero. Although the mean and mode appear almost equal below $p \approx 0.75$, the logarithm scale (see inset of Fig. 4.9) reveals another change of behavior between anti-bias and low-bias values of p. For $p < 0.5$, the asymmetry behaves smoothly, remaining close to $1/2t$ due to the discrete nature of the distribution, whereas fluctuations without any clear trend occur within the range $0.5 < p < 0.75$.

Recall, however, that to observe bimodality under symmetric initial conditions, the control parameter must satisfy $p > p^{\mathrm{ch}} = 0.836 > 0.75$. For asymmetric initial conditions, bimodality does not occur. Instead, for $p > 0.75$, the mode, which is the most probable size, exceeds the mean, representing the expected size. Taken together, these observations suggest that $p \approx 0.75$ does not mark the onset of bimodality, but rather signals a gradual change in the system's behavior.

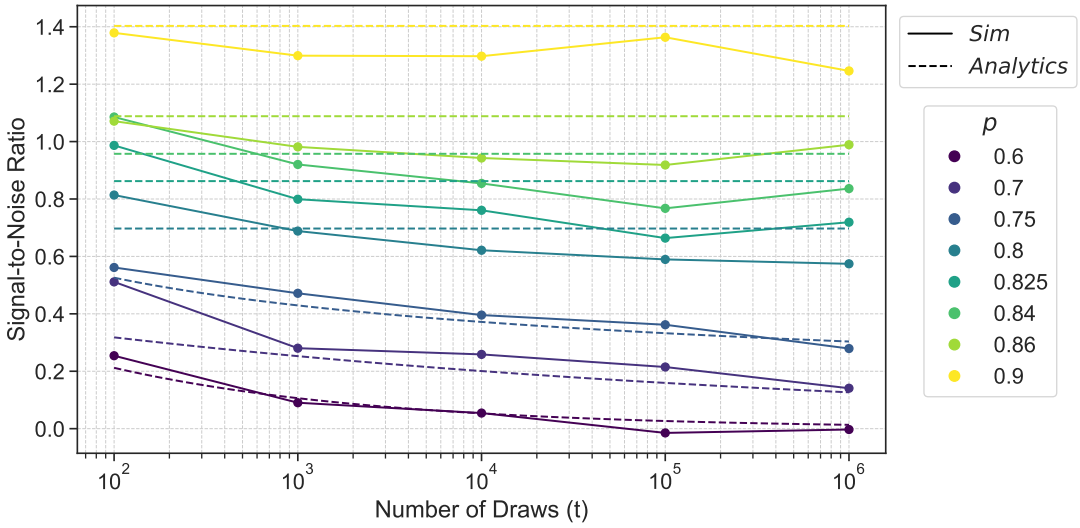


Figure 4.13: SNR (4.20) as a way to estimate the relationship between the mean faction sizes and fluctuations. In the low bias regime $p < 0.75$, it becomes increasingly difficult to detect any difference in the faction sizes as the system grows. At $p = 0.75$, the SNR decreases, but this decrease is very slow. In the high Bias Regime ($p > 0.75$), SNR is a constant. This constancy indicates that in this regime, the difference in faction sizes is more stable and detectable

4.5 Discussion and Conclusion

In this chapter, I have explored the model of two opposing growing factions, in which new agents join either the group of a randomly chosen agent, or with probability $1 - p$, the antagonistic group. I demonstrated that the probabilistic evolution of faction sizes aligns with the framework of a randomized Pólya urn process. A central insight is that for all $p < 1$, the system gradually trends toward an equilibrium where faction sizes are comparable in relative terms, although the rate and manner of this convergence vary with the value of p. Specifically, a persistent asymmetry in faction sizes is sustained when the bias parameter exceeds the characteristic threshold $p^{ch} = 0.836$.

The attachment bias parameter p represents a feedback mechanism triggered by a new individual's interaction with the first encountered group. The positive feedback leads the individual to join this group, while the negative one results in choosing the opposing faction. While there are many potential reasons why people join one of many groups—e.g., access to resources, internal group structure, productivity, safety, or personal development [122–124]—the model does not attempt to account for all such factors explicitly. Instead, the bias parameter serves as a proxy for these influences, capturing their cumulative effect through a single probabilistic mechanism. Although this specific formulation of attachment bias is unique to the considered model, its outcomes can be meaningfully related to results in other studies. For instance, the emergence of group size inequality for $p > p^{ch}$ is conceptually similar to the productivity-driven inequalities described in [123]. In another model [125], the overall group size equality, similar to the case of $p < p^{ch}$ in

my framework, has been achieved through the introduction of the third possible state of not belonging to any of the competing groups.

This study proposes a dual perspective on faction formation: one grounded in network-specific structural properties and the other in underlying stochastic processes. Traditionally, network science has emphasized structural features, such as clustering, modularity, and structural balance measures, as primary explanatory variables for the emergence and persistence of polarized groups [7, 126, 127]. These approaches typically center on individuals whose decisions shape the formation of network connection patterns, thereby generating a divided system. A similar viewpoint can be adopted in the model when interpreted at the level of individuals forming links. However, such a detailed, connection-specific description is not necessary. The dynamics can instead be represented purely from a group-level perspective, where individuals join one of the two factions. From this higher-level viewpoint, the system evolves via two stochastic mechanisms: the Pólya urn process and Bernoulli trials. This abstraction eliminates the need to model explicit connections between all agents while still capturing essential polarization dynamics. Because the individual-based and group-based descriptions are consistent with each other, my framework enables drawing analogies and insights across both levels of analysis. My approach highlights the importance of stochastic processes in shaping social and network dynamics, which were observed in other studies [9, 128]. The combination of individual and group perspectives has also been explored in more complex models involving hypergraphs, where interactions extend beyond pairwise links to group-level relations [129]. In this sense, my model can be interpreted as a limiting case of hypergraph-based models that capture polarization and group formation [130–132].

Three different methods were used to examine the dynamics of the system growth: the master equation, the rate equation, and the Pólya process formula. The master equation gives an exact distribution. It is not closed-form, but it can be used to derive the mean faction size. The rate equation aligns with the master equation at short timescales. The accuracy decreases slightly for larger t, but the discrepancy remains small and constant. For the analyzed initial conditions with the single starting node, the Pólya process formula better approximates the long-term behavior compared to the rate equation. Obtaining similar formulae for other initial conditions might be possible, yet it is not trivial. On the other hand, the rate equation is general for any initial conditions.

My results reveal that for all $p < 1$, the network evolves toward relatively balanced factions. However, perfect symmetry is not guaranteed. I identified three distinct behavioral regimes in the evolution of faction sizes, each characterized by qualitatively different statistical properties.

1. **Anti-bias regime** ($p < 0.5$): Here, the network swiftly evolves toward balance, with both the mean and mode of faction sizes converging rapidly to $t/2$, regardless of initial conditions. The symmetry emerges robustly, and transient fluctuations

dissipate quickly.

2. **Low-bias regime** ($0.5 < p < p^{ch}$): In this intermediate range, the mean sizes of the two factions begin to diverge, yet the fluctuations are sufficient to produce a unimodal distribution. The distinction between factions becomes obscured, especially under symmetric initial conditions, where the mode stabilizes quickly. Under asymmetric conditions, convergence occurs more slowly but ultimately yields indistinguishable group sizes in the limit.

3. **High-bias regime** ($p > p^{ch} \approx 0.836$): Beyond the characteristic threshold, the system enters a qualitatively new regime. When initialized symmetrically, the faction size distribution becomes distinctly bimodal, indicating the emergence of two stable configurations associated with persistent group asymmetry. Below this threshold, fluctuations remain too pronounced to permit clear peak separation.

These regimes can be differentiated not only by the modality of the distribution but also by the relationship between statistical descriptors such as the mean and mode of faction sizes. Notably, this mode-mean separation increases markedly in the high-bias regime and thus serves as a diagnostic tool for detecting regime shifts.

An earlier analytical estimate, derived from scaling arguments in the Pólya process literature, suggested a transition at $p = 0.75$. However, the numerical results based on the master equation reveal no substantive transition at this point. Instead, the scaling assumptions that underlie this analytical prediction break down near the vicinity of $p = 0.75$, and no sharp phase transition is observed. Rather, this value corresponds to the onset of noticeable skewness in the distribution—a gradual structural change rather than a critical point. True bimodality emerges only when p exceeds the empirically determined threshold $p^{ch} \approx 0.836$.

An interesting takeaway is the fact that unless $p = 1$, the normalized group sizes equalize. Consider a scenario in which an uninformed individual encounters information online from one side of a debate (for example, anti-vaccine arguments). If there is even a slight chance that this individual is not persuaded and instead ends up joining the opposing camp ($p < 1$), then the two opposing camps will tend toward the same size, without one or the other winning. This is in stark contrast to most majority opinion dynamics. The key distinction is that, in my model, opinions do not convert opponents but rather compete to attract those who have yet to form an opinion, and thus are growing. While in hindsight this result is logical, as $p = 1$ represents proportional growth where relative sizes of each group remain constant, and any $p < 1$ means actual advantage for the minority, the outcome of group size equalization is not immediately obvious. It would be intriguing to extend this framework by integrating it with traditional opinion dynamics models such as the voter model [133] or majority rule model [134]. Such a combination would capture both the internal opinion shifts among individuals already holding views

and the external recruitment of new agents, offering a more comprehensive perspective on the formation and evolution of public opinion.

This work assumed that agents form relations with all those already present, and they strictly adhere to the principles of the structural balance, which forces the system into a permanent, balanced two-clique state (two opposing factions). A natural extension would be to consider either some noise in the link sign choice or new agents probabilistically forming relations with neighboring agents as described in [135]. This could lead to more complex behavior, potentially involving the formation of multiple hostile groups or a transition towards structurally unbalanced network configurations.

Summing up, this work advances our understanding of how structural balance principles interact with randomized attachment dynamics, which might offer insights into the self-organization of growing social alliances, adversarial communities, and other polarized systems.

Chapter 5

Summary

This dissertation investigates the dynamics, stability, and emergent structural properties of signed social networks grounded in the principles of Heider's structural balance theory. The investigations employ two primary methodological approaches: Statistical physics approaches are used to analyze equilibrium states and phase transitions in structural balance models across different network configurations, including multiplex networks and sparse random graphs. While complementary, stochastic modeling techniques are employed for understanding the emergence, persistence, and relative dominance of opposing factions during network growth.

One significant aspect of this thesis explores the extension of the Heider balance theory to multiplex networks. In Chapter 2, I delve into scenarios where agents maintain relationships across multiple interconnected layers, reflecting the multifaceted nature of social interactions. The model incorporates intralayer dynamics based on Heider balance and interlayer correlations inspired by Ising interactions, with temperature representing the tolerance for imbalanced relationships. Investigating scenarios where agents maintain relationships across interconnected layers, the analysis reveals how interlayer correlations, modeled akin to Ising interactions, coupled with intralayer balance dynamics (influenced by temperature as noise), shape network stability and phase transitions. Key findings demonstrate that the multiplex nature of the network enhances the stability of the ordered "paradise" state against thermal noise compared to single-layer networks. Furthermore, interactions between the layer states give rise to complex phase behaviors, characterized by distinct temperatures for interlayer synchronization (T_s) and the transition to disorder (T_d). A comprehensive phase diagram for duplex systems shows the stability regions of ordered and disordered states based on temperature and interlayer coupling strength. It identifies multiple characteristic transition points revealed by mean-field theory and agent-based simulations.

Recognizing that many real-world social networks exhibit sparse connectivity, the thesis subsequently examines structural balance dynamics on ER random graphs, again employing statistical physics methods. This investigation focuses on the transition from an

all-positive "paradise" state to disorder under thermal fluctuations within sparse topology. A key theoretical contribution is the derivation of the dependence of interaction strengths on connection probability– renormalized as p^{-2} for intralayer Heider interaction and p^{-1} for interlayer Ising coupling. This renormalization recovers mean-field behavior similar to that of complete graphs within the sparse ER setting. Analysis of triad distributions confirms the shift towards entropy-dominated configurations in the disordered phase. Overall, it has been demonstrated that although the ER network structure contains a topological disorder, the spontaneous order in structural balance for ER graphs can still exist below a critical temperature. Such a study establishes that the critical properties of balance dynamics on random graphs can be described using a properly adapted mean-field approach adapted to network topology.

Shifting focus from the equilibrium properties of static network structures (both multiplex and sparse) to the processes of network formation, the final part of the dissertation investigates the faction formation in a growing network using a stochastic modeling approach. A model based on a generalized randomized Pólya urn process is introduced. In this model, new nodes attach sequentially, forming signed links that are influenced by an intrinsic attachment bias —representing the baseline probability of forming a positive (friendly) link—while simultaneously being constrained to maintain structural balance. Mathematical analysis using rate and master equations to track faction size evolution shows that the combined influence of stochastic attachment bias and structural balance causes the network to self-organize into two opposing, cohesive factions. Three regimes describe the system's behavior. In the anti-bias regime, group sizes equalize rapidly. The dynamics reflect contrarian behavior [113], causing the system to reinforce the minority group. In the low bias regime, which exceeds 0.5 but remains below a critical threshold, mean faction sizes diverge, but fluctuations mask this asymmetry, keeping the distribution unimodal. For the high bias regime, bimodality emerges, reflecting the stable coexistence of unequal factions. A notable crossover occurs before reaching the critical threshold, corresponds to a change in fluctuation scaling rather than a true phase transition.

In conclusion, this thesis examines multiplex networks, where agents have multiple types of relationships across distinct layers. Structurally balanced configurations of such networks are more stable and resilient than those of single-layer networks. This thesis also demonstrates that stable group divisions can emerge from a simple, group-level stochastic process, in which a new individual supports or opposes a randomly chosen person's group.

Bibliography

[1] M. E. Newman, D. J. Watts, S. H. Strogatz, "Random graph models of social networks", *Proceedings of the national academy of sciences* **2002**, *99*, 2566–2572, DOI 10.1073/pnas.012582999 (cit. on pp. 11, 24).

[2] A.-L. Barabási, R. Albert, "Emergence of scaling in random networks", *science* **1999**, *286*, 509–512, DOI 10.1126/science.286.5439.509 (cit. on pp. 11, 24, 27).

[3] J. Leskovec, D. Huttenlocher, J. Kleinberg, "Signed networks in social media", *Proc. SIGCHI Conf. Human Factors in Computing Systems* **2010**, *2*, 1361–1370, DOI 10.1145/1753326.1753532 (cit. on p. 11).

[4] F. Heider, "Attitudes and Cognitive Organization", *Journal of Psychology* **1946**, *21*, 107–112, DOI 10.1080/00223980.1946.9917275 (cit. on pp. 11, 25, 66).

[5] D. Cartwright, F. Harary, "Structural balance: a generalization of Heider's theory", *Psychol. Rev.* **1956**, *63*, 277–293, DOI 10.1037/h0046049 (cit. on pp. 11, 25, 66).

[6] G. Facchetti, G. Iacono, C. Altafini, "Computing global structural balance in large-scale signed social networks", *Proceedings of the National Academy of Sciences* **2011**, *108*, 20953–20958, DOI 10.1073/pnas.1109521108 (cit. on pp. 11, 26, 28).

[7] P. Doreian, A. Mrvar, "A partitioning approach to structural balance", *Social networks* **1996**, *18*, 149–168, DOI 10.1016/0378-8733(95)00259-6 (cit. on pp. 11, 105).

[8] X. He, H. Du, M. Cai, M. W. Feldman, "The evolution of cooperation in signed networks under the impact of structural balance", *PLoS one* **2018**, *13*, e0205084 (cit. on p. 11).

[9] C. Stadtfeld, K. Takács, A. Vörös, "The emergence and stability of groups in social networks", *Social Networks* **2020**, *60*, 129–145, DOI 10.1016/j.socnet.2019.10.008 (cit. on pp. 11, 105).

[10] K. Mohandas, K. Suchecki, J. A. Hołyst, "Critical properties of Heider balance on multiplex networks", *Physical Review E* **2024**, *109*, DOI 10.1103/PhysRevE.109.044306 (cit. on pp. 12, 66, 68, 70, 72, 80).

[11] K. Mohandas, K. Suchecki, J. A. Hołyst, "Paradise-disorder transition in structural balance dynamics on Erdős-Rényi graphs", *Physical Review E* **2025**, *111*, DOI 10.1103/PhysRevE.111.024310 (cit. on p. 12).

[12] K. Mohandas, P. J. Górski, K. Suchecki, G. Andres, G. Vaccario, J. A. Hołyst, "Structurally balanced growing network as randomized Pólya urn process", *arXiv:-2510.24659v1 [physics.soc-ph]* **2025**, DOI 10.48550/arXiv.2510.24659 (cit. on p. 12).

[13] C. Mesjasz, "Complexity of social systems", *Acta Physica Polonica A* **2010**, *117*, 706–715, DOI 10.12693/APhysPolA.117.706 (cit. on p. 21).

[14] J. Johnson, "The future of the social sciences and humanities in the science of complex systems", *Innovation–The European Journal of Social Science Research* **2010**, *23*, 115–134, DOI 10.1080/13511610.2010.518422 (cit. on p. 21).

[15] D. Daems, *Social complexity and complex systems in archaeology*, Routledge, **2021**, DOI 10.4324/9781003036968 (cit. on p. 21).

[16] K. Kaneko, *Life: an introduction to complex systems biology*, Springer, **2006**, DOI 10.1007/978-3-540-32667-0 (cit. on p. 21).

[17] A. Ma'ayan, "Complex systems biology", *Journal of the Royal Society Interface* **2017**, *14*, 20170391, DOI 10.1098/rsif.2017.0391 (cit. on p. 21).

[18] L. d. F. Costa, F. A. Rodrigues, A. S. Cristino, "Complex networks: the key to systems biology", *Genetics and Molecular Biology* **2008**, *31*, 591–601, DOI 10.1590/S1415-47572008000400001 (cit. on p. 21).

[19] J. Foster, "From simplistic to complex systems in economics", *Cambridge Journal of Economics* **2005**, *29*, 873–892, DOI 10.1093/cje/bei083 (cit. on p. 21).

[20] O. Gomes, M. Gubareva, "Complex systems in economics and where to find them", *Journal of Systems Science and Complexity* **2021**, *34*, 314–338, DOI 10.1007/s11424-020-9149-1 (cit. on p. 21).

[21] H. Deguchi, *Economics as an agent-based complex system: toward agent-based social systems sciences*, Springer Science & Business Media, **2004**, DOI 10.1007/978-4-431-53957-5 (cit. on p. 21).

[22] A. J. Seely, P. T. Macklem, "Complex systems and the technology of variability analysis", *Critical care* **2004**, *8*, 1–18, DOI 10.1186/cc2948 (cit. on p. 21).

[23] I. P. McCarthy, "Technology management–a complex adaptive systems approach", *International Journal of Technology Management* **2003**, *25*, 728–745, DOI 10.1504/IJTM.2003.003134 (cit. on p. 21).

[24] S. A. Sheard, A. Mostashari, "Principles of complex systems for systems engineering", *Systems Engineering* **2009**, *12*, 295–311, DOI 10.1002/sys.20124 (cit. on p. 21).

[25] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, M. A. Porter, "Multilayer networks", *Journal of Complex Networks* **2014**, *2*, 203–271, DOI 10.1093/comnet/cnu016 (cit. on p. 22).

[26] Y. Murase, J. Török, H.-H. Jo, K. Kaski, J. Kertész, "Multilayer weighted social network model", *Physical Review E* **2014**, *90*, DOI 10.1103/PhysRevE.90.052810 (cit. on p. 22).

[27] A. Solé-Ribalta, S. Gómez, A. Arenas, "Congestion Induced by the Structure of Multiplex Networks", *Physical review letters* **2016**, *116*, DOI 10.1103/PhysRevLett.116.108701 (cit. on p. 22).

[28] M. Szell, R. Lambiotte, S. Thurner, "Multirelational organization of large-scale social networks in an online world", *Proceedings of the National Academy of Sciences* **2010**, *107*, 13636–13641, DOI 10.1073/pnas.1004008107 (cit. on pp. 22, 26).

[29] S. Boccaletti, G. Bianconi, R. Criado, C. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, M. Zanin, "The structure and dynamics of multilayer networks", *Physics reports* **2014**, *544*, 1–122, DOI 10.1016/j.physrep.2014.07.001 (cit. on p. 22).

[30] A. I. E. Hosni, K. Li, S. Ahmad, "Minimizing rumor influence in multiplex online social networks based on human individual and social behaviors", *Information Sciences* **2020**, *512*, 1458–1480, DOI 10.1016/j.ins.2019.10.063 (cit. on p. 22).

[31] R. Singh, J. Xu, B. Berger, "Global alignment of multiple protein interaction networks with application to functional orthology detection", *Proceedings of the National Academy of Sciences* **2008**, *105*, 12763–12768, DOI 10.1073/pnas.0806627105 (cit. on p. 22).

[32] J. Gómez-Gardeñes, I. Reinares, A. Arenas, L. M. Floria, "Evolution of cooperation in multiplex networks", *Scientific Reports* **2012**, *2*, DOI 10.1038/srep00620 (cit. on p. 22).

[33] M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, D. Montesi, "Spreading Processes in Multilayer Networks", *IEEE Transactions Network Science and Engineering* **2015**, *2*, 65–83, DOI 10.1109/TNSE.2015.2425961 (cit. on p. 22).

[34] N. Arinik, R. Figueiredo, V. Labatut, "Multiple partitioning of multiplex signed networks: Application to European parliament votes", *Social Networks* **2020**, *60*, 83–102, DOI 10.1016/j.socnet.2019.02.001 (cit. on p. 22).

[35] P. Erdös, A. Rényi, "On Random Graphs I", *Publicationes Mathematicae Debrecen* **1959**, *6*, 290–297 (cit. on pp. 23, 65).

[36] M. Shumate, E. T. Palazzolo, "Exponential random graph (p*) models as a method for social network analysis in communication research", *Communication Methods and Measures* **2010**, *4*, 341–371, DOI 10.1080/19312458.2010.527869 (cit. on p. 24).

[37] P. Sah, L. O. Singh, A. Clauset, S. Bansal, "Exploring community structure in biological networks with random graphs", *BMC bioinformatics* **2014**, *15*, 1–14, DOI 10.1186/1471-2105-15-220 (cit. on p. 24).

[38] F. Slanina, M. Kotrla, "Random networks created by biological evolution", *Physical Review E* **2000**, *62*, 6170, DOI 10.1103/PhysRevE.62.6170 (cit. on p. 24).

[39] G. Grinstein, R. Linsker, "Synchronous neural activity in scale-free network models versus random network models", *Proceedings of the National Academy of Sciences* **2005**, *102*, 9948–9953, DOI 10.1073/pnas.0504127102 (cit. on p. 24).

[40] F. Wu, "Random graphs and network communication", *Journal of Physics A: Mathematical and General* **1982**, *15*, L395, DOI 10.1088/0305-4470/15/8/004 (cit. on p. 24).

[41] R. Elsässer, L. Gasieniec in Proceedings of the seventeenth annual ACM symposium on Parallelism in algorithms and architectures, Association for Computing Machinery, **2005**, pp. 309–315, DOI 10.1145/1073970.1074021 (cit. on p. 24).

[42] P. L. Krapivsky, G. J. Rodgers, S. Redner, "Degree distributions of growing networks", *Physical Review Letters* **2001**, *86*, 5401, DOI 10.1103/PhysRevLett.86.5401 (cit. on p. 24).

[43] K.-k. Shang, B. Yang, J. M. Moore, Q. Ji, M. Small, "Growing networks with communities: A distributive link model", *Chaos: An Interdisciplinary Journal of Nonlinear Science: An Interdisciplinary Journal of Nonlinear Science* **2020**, *30*, DOI 10.1063/5.0007422 (cit. on p. 24).

[44] M. Golosovsky, S. Solomon, "Growing complex network of citations of scientific papers: Modeling and measurements", *Physical Review E* **2017**, *95*, 012324, DOI 10.1103/PhysRevE.95.012324 (cit. on p. 25).

[45] M. Small, Y. Li, T. Stemler, K. Judd, "Growing optimal scale-free networks via likelihood", *Physical Review E* **2015**, *91*, 042801, DOI 10.1103/PhysRevE.91.042801 (cit. on p. 25).

[46] J. A. Davis, "Clustering and structural balance in graphs", *Human relations* **1967**, *20*, 181–187, DOI 10.1177/001872676702000206 (cit. on p. 25).

[47] M. Granovetter, "Threshold models of collective behavior", *American journal of sociology* **1978**, *83*, 1420–1443, DOI 10.1086/226707 (cit. on p. 26).

[48] T. C. Schelling, "Dynamic models of segregation", *Journal of mathematical sociology* **1971**, *1*, 143–186, DOI 10.1080/0022250X.1971.9989794 (cit. on p. 26).

[49] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, "Mixing beliefs among interacting agents", *Advances in Complex Systems* **2000**, *3*, 87–98, DOI 10.1142/S0219525900000078 (cit. on p. 27).

[50] H. Rainer, U. Krause, "Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation", *Journal of Artificial Societies and Social Simulation* **2002**, *5* (cit. on p. 27).

[51] D. Helbing, "Traffic and related self-driven many-particle systems", *Reviews of modern physics* **2001**, *73*, 1067, DOI 10.1103/RevModPhys.73.1067 (cit. on p. 27).

[52] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al., "Computational social science", *Science* **2009**, *323*, 721–723, DOI 10.1126/science.1167742 (cit. on p. 27).

[53] G. George, M. R. Haas, A. Pentland, Big data and management, **2014**, DOI 10.5465/amj.2014.4002 (cit. on p. 27).

[54] T. Antal, P. L. Krapivsky, S. Redner, "Dynamics of social balance on networks", *Physical Review E* **2005**, *72*, DOI 10.1103/PhysRevE.72.036121 (cit. on pp. 27, 66, 67).

[55] K. Malarz, J. A. Hołyst, "Mean-field approximation for structural balance dynamics in heat bath", *Physical Review E* **2022**, *106*, DOI 10.1103/PhysRevE.106.064139 (cit. on pp. 27, 36, 41, 46, 63).

[56] S. A. Marvel, S. H. Strogatz, J. M. Kleinberg, "Energy Landscape of Social Balance", *Physical review letters* **2009**, *103*, DOI 10.1103/PhysRevLett.103.198701 (cit. on pp. 27, 66).

[57] G. Facchetti, G. Iacono, C. Altafini, "Exploring the low-energy landscape of large-scale signed social networks", *Physical Review E—Statistical Nonlinear and Soft Matter Physics* **2012**, *86*, 036116, DOI 10.1103/PhysRevE.86.036116 (cit. on p. 27).

[58] K.-M. Lee, E. Shin, S. You, "An economic model of friendship and enmity for measuring social balance in networks", *Physica A: Statistical Mechanics and its Applications* **2017**, *488*, 205–215, DOI 10.1016/j.physa.2017.07.012 (cit. on p. 27).

[59] K. Kacperski, J. A. Hołyst, "Phase transitions and hysteresis in a cellular automata-based model of opinion formation", *Journal of Statistical Physics* **1996**, *84*, 169–189, DOI 10.1007/BF02179581 (cit. on p. 28).

[60] F. Rabbani, A. H. Shirazi, G. R. Jafari, "Mean-field solution of structural balance dynamics in nonzero temperature", *Physical Review E* **2019**, *99*, DOI 10.1103/PhysRevE.99.062302 (cit. on pp. 28, 67).

[61] A. Kirkley, G. T. Cantwell, M. E. Newman, "Balance in signed networks", *Physical Review E* **2019**, *99*, 012320, DOI 10.1103/PhysRevE.99.012320 (cit. on p. 28).

[62] T. H. Summers, I. Shames, "Active influence in dynamical models of structural balance in social networks", *EPL* **2013**, *103*, DOI 10.1209/0295-5075/103/18001 (cit. on p. 28).

[63] W. Xia, M. Cao, K. H. Johansson, "Structural Balance and Opinion Separation in Trust-Mistrust Social Networks", *IEEE Transactions Control Network Systems* **2016**, *3*, 46–56, DOI 10.1109/TCNS.2015.2437528 (cit. on p. 28).

[64] M. Yang, X. Wang, L. Ma, Q. He, M. Huang, "A promotive structural balance model based on reinforcement learning for signed social networks", *Neural Computing and Applications* **2022**, *34*, 16683–16700, DOI 10.1007/s00521-022-07298-y (cit. on p. 28).

[65] P. Doreian, A. Mrvar, "Structural balance and signed international relations", *Journal of Social Structure* **2015**, *16*, 1–49, DOI 10.21307/joss-2019-012 (cit. on p. 28).

[66] P. J. Górski, K. Kułakowski, P. Gawroński, J. A. Hołyst, "Destructive influence of interlayer coupling on Heider balance in bilayer networks", *Scientific Reports* **2017**, *7*, DOI 10.1038/s41598-017-15960-y (cit. on pp. 28, 29, 66).

[67] R. P. Kundu, S. Pandey, "Balance and Cooperation in Signed Networks: A Model of Multiplex Formation", *SSRN* **2022**, DOI 10.2139/ssrn.4499176 (cit. on p. 28).

[68] E. Cozzo, M. Kivelä, M. D. Domenico, A. Solé-Ribalta, A. Arenas, S. Gómez, M. A. Porter, Y. Moreno, "Structure of triadic relations in multiplex networks", *New J. Phys.* **2015**, *17*, DOI 10.1088/1367-2630/17/7/073029 (cit. on pp. 28, 29).

[69] K. Burghardt, Z. Maoz, "Imbalance in Multiplex Networks", *SSRN* **2018**, DOI 10.2139/ssrn.3192890 (cit. on p. 28).

[70] K. Kułakowski, P. Gawroński, P. Gronek, "The Heider balance: A continuous approach", *International Journal of Modern Physics C* **2005**, *16*, 707–716, DOI 10.1142/S012918310500742X (cit. on p. 28).

[71] P. Klimek, M. Diakonova, V. M. Eguíluz, M. San Miguel, S. Thurner, "Dynamical origins of the community structure of an online multi-layer society", *New journal of Physics* **2016**, *18*, 083045, DOI 10.1088/1367-2630/18/8/083045 (cit. on p. 29).

[72] P. Klimek, S. Thurner, "Triadic closure dynamics drives scaling laws in social multiplex networks", *New Journal of Physics* **2013**, *15*, 063008, DOI 10.1088/1367-2630/15/6/063008 (cit. on p. 29).

[73] H. Mahmoud, *Pólya urn models*, Chapman and Hall/CRC, **2008**, DOI 10.1201/9781420059847 (cit. on p. 29).

[74] R. Golman, D. Hagmann, J. H. Miller, "Polya's bees: A model of decentralized decision-making", *Science Advances* **2015**, *1*, e1500253, DOI 10.1126/sciadv.1500253 (cit. on p. 30).

[75] S. Bowles, *The new economics of inequality and redistribution*, Cambridge University Press, **2012**, DOI 10.1017/S0266267114000078 (cit. on p. 30).

[76] P. L. Lau, T. T. Koo, C.-L. Wu, "Spatial distribution of tourism activities: A polya urn process model of rank-size distribution", *Journal of Travel Research* **2020**, *59*, 231–246, DOI 10.1177/0047287519829251 (cit. on p. 30).

[77] S. Janson, "Functional limit theorems for multitype branching processes and generalized Pólya urns", *Stochastic Processes and their Applications* **2004**, *110*, 177–245, DOI 10.1016/j.spa.2003.12.002 (cit. on pp. 30, 87, 100, 101).

[78] P. Siedlecki, J. Szwabiński, T. Weron, "The interplay between conformity and anticonformity and its polarizing effect on society", *JASSS* **2016**, *19*, DOI 10.18564/jasss.3203 (cit. on p. 30).

[79] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, M. Starnini, "Modeling echo chambers and polarization dynamics in social networks", *Physical review letters* **2020**, *124*, 048301, DOI 10.1103/PhysRevLett.124.048301 (cit. on p. 30).

[80] S. C. Phillips, J. Uyheng, K. M. Carley, "A high-dimensional approach to measuring online polarization", *Journal of Computational Social Science* **2023**, *6*, 1147–1178, DOI 10.1007/s42001-023-00227-6 (cit. on p. 30).

[81] S. Whitt, A. B. Yanus, B. McDonald, J. Graeber, M. Setzler, G. Ballingrud, M. Kifer, "Tribalism in America: Behavioral experiments on affective polarization in the Trump era", *Journal of Experimental Political Science* **2021**, *8*, 247–259, DOI 10.1017/XPS.2020.29 (cit. on p. 31).

[82] P. Dandekar, A. Goel, D. T. Lee, "Biased assimilation, homophily, and the dynamics of polarization", *Proceedings of the National Academy of Sciences* **2013**, *110*, 5791–5796, DOI 10.1073/pnas.1217220110 (cit. on pp. 31, 83).

[83] A. Asikainen, G. Iñiguez, J. Ureña-Carrión, K. Kaski, M. Kivelä, "Cumulative effects of triadic closure and homophily in social networks", *Science Advances* **2020**, *6*, eaax7310, DOI 10.1007/s42001-023-00227-6 (cit. on pp. 31, 83).

[84] M. Szell, R. Lambiotte, S. Thurner, "Multirelational organization of large-scale social networks in an online world", *Proceedings of the National Academy of Sciences* **2010**, *107*, 13636–13641, DOI 10.1073/pnas.1004008107 (cit. on pp. 31, 84).

[85] J. Lu, H. Du, X. He, "A hypernetwork-based urn model for explaining collective dynamics", *PLoS one* **2023**, *18*, 1–20, DOI 10.1371/journal.pone.0291778 (cit. on p. 31).

[86] R. Marcaccioli, G. Livan, "A pólya urn approach to information filtering in complex networks", *Nature communications* **2019**, *10*, 745, DOI 10.1038/s41467-019-08667-3 (cit. on p. 31).

[87] X. He, J. Lu, H. Du, X. Jin, "Social Influence in Signed Networks", *IEEE Transactions on Computational Social Systems* **2024**, *11*, 330–339, DOI 10.1109/TCSS.2022.3220944 (cit. on p. 31).

[88] T. Antal, P. Krapivsky, S. Redner, "Social balance on networks: The dynamics of friendship and enmity", *Physica D* **2006**, *224*, 130–136, DOI 10.1016/j.physd.2006.09.028 (cit. on pp. 36, 49, 66, 67).

[89] R. Shojaei, P. Manshour, A. Montakhab, "Phase transition in a network model of social balance with Glauber dynamics", *Physical Review E* **2019**, *100*, DOI 10.1103/PhysRevE.100.022303 (cit. on p. 63).

[90] M. Winkler, J. Reichardt, "Motifs in triadic random graphs based on Steiner triple systems", *Physical Review E* **2013**, *88*, DOI 10.1103/PhysRevE.88.022805 (cit. on p. 65).

[91] S. Itzkovitz, U. Alon, "Subgraphs and network motifs in geometric networks", *Physical Review E* **2005**, *71*, DOI 10.1103/PhysRevE.71.026117 (cit. on p. 65).

[92] H. Shi, L. Shi, "Identifying emerging motif in growing networks", *PLoS one* **2014**, *9*, DOI 10.1371/journal.pone.0099634 (cit. on p. 65).

[93] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, "Network Robustness and Fragility: Percolation on Random Graphs", *Physical review letters* **2000**, *85*, 5468–5471, DOI 10.1103/PhysRevLett.85.5468 (cit. on p. 65).

[94] S. Dorogovtsev, A. Goltsev, J. Mendes, "Ising model on networks with an arbitrary distribution of connections", *Phys Rev E* **2002**, *66*, DOI 10.1103/PhysRevE.66.016104 (cit. on pp. 65, 82).

[95] A. Dembo, A. Montanari, "Ising models on locally tree-like graphs", *Ann. Appl. Probab.* **2010**, *20*, 565–592, DOI 10.1214/09-AAP627 (cit. on pp. 65, 82).

[96] D. Vilone, C. Castellano, "Solution of voter model dynamics on annealed small-world networks", *Physical Review E* **2004**, *69*, DOI 10.1103/PhysRevE.69.016109 (cit. on p. 65).

[97] L. F. Pereira, F. B. Moreira, "Majority-vote model on random graphs", *Physical Review E* **2005**, *71*, DOI 10.1103/PhysRevE.71.016123 (cit. on p. 65).

[98] A. Pournaki, E. Olbrich, S. Banisch, K. Klemm, "Order-disorder transition in the zero-temperature Ising model on random graphs", *Physical Review E* **2023**, *107*, DOI 10.1103/PhysRevE.107.054112 (cit. on p. 65).

[99] M. W. Macy, B. K. Szymanski, J. A. Hołyst, "The Ising model celebrates a century of interdisciplinary contributions", *npj Complexity* **2024**, *1*, 10, DOI 10.1038/s44260-024-00012-0 (cit. on p. 65).

[100] R. Masoumi, F. Oloomi, S. Sajjadi, A. Shirazi, G. Jafari, "Modified Heider balance on Erdös-Rényi networks", *Physical Review E* **2022**, *106*, DOI 10.1103/PhysRevE.106.034309 (cit. on pp. 66, 82).

[101] K. Malarz, M. Wołoszyn, "Thermal properties of structurally balanced systems on classical random graphs", *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2023**, *33*, DOI 10.1063/5.0141019 (cit. on pp. 66, 75, 78, 82).

[102] K. Malarz, J. A. Hołyst, "Mean-field approximation for structural balance dynamics in heat bath", *Physical Review E* **2022**, *106*, DOI 10.1103/PhysRevE.106.064139 (cit. on pp. 66–70).

[103] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, S. H. Strogatz, "Continuous-time model of structural balance", *Proceedings of the National Academy of Sciences* **2011**, *108*, 1771–1776, DOI 10.1073/pnas.1013213108 (cit. on p. 66).

[104] R. Dunbar, "The Anatomy of Friendship", *Trends in Cognitive Sciences* **2018**, *22*, 32–51, DOI 10.1016/j.tics.2017.10.004 (cit. on p. 81).

[105] A. Bramson, P. Grim, D. J. Singer, S. Fisher, W. Berger, G. Sack, C. Flocken, "Disambiguation of social polarization concepts and measures", *The Journal of Mathematical Sociology* **2016**, *40*, 80–111, DOI 10.1080/0022250X.2016.1147443 (cit. on p. 83).

[106] J. Liu, S. Huang, N. M. Aden, N. F. Johnson, C. Song, "Emergence of polarization in coevolving networks", *Physical Review Letters* **2023**, *130*, 037401, DOI 10.1103/PhysRevLett.130.037401 (cit. on p. 83).

[107] P. Guerra, W. Meira Jr, C. Cardie, R. Kleinberg in Proceedings of the international AAAI conference on web and social media, *Vol. 7*, **2013**, pp. 215–224, DOI 10.1609/icwsm.v7i1.14421 (cit. on p. 83).

[108] D. Cartwright, F. Harary, "Structural balance: a generalization of Heider's theory.", *Psychological Review* **1956**, *63*, 277 (cit. on pp. 83, 86).

[109] R. Axelrod, "The dissemination of culture: A model with local convergence and global polarization", *Journal of conflict resolution* **1997**, *41*, 203–226, DOI 10.1177/0022002797041002001 (cit. on p. 83).

[110] K. Malarz, J. A. Hołyst, "Mean-field approximation for structural balance dynamics in heat bath", *Physical Review E* **2022**, *106*, 064139, DOI 10.1103/PHYSREVE.106.064139/FIGURES/6/MEDIUM (cit. on p. 83).

[111] N. L. Johnson, S. Kotz, *Urn Models and Their Application*, John Wiley & Sons, New York, **1977** (cit. on p. 85).

[112] S. Kotz, H. Mahmoud, P. Robert, "On generalized Pólya urn models", *Statistics & Probability Letters* **2000**, *49*, 163–173, DOI 10.1016/S0167-7152(00)00045-6 (cit. on p. 85).

[113] S. Galam, "Contrarian deterministic effects on opinion dynamics: "the hung elections scenario"", *Physica A: Statistical Mechanics and its Applications* **2004**, *333*, 453–460, DOI https://doi.org/10.1016/j.physa.2003.10.041 (cit. on pp. 85, 110).

[114] W. Weidlich, "Synergetic modelling concepts for sociodynamics with application to collective political opinion formation", *The Journal of Mathematical Sociology* **1994**, *18*, 267–291, DOI 10.1080/0022250X.1994.9990129 (cit. on p. 90).

[115] R. B. D'Agostino, *Goodness-of-fit-techniques*, *Vol. 68*, CRC press, **1986**, DOI 10.5555/19293 (cit. on p. 93).

[116] J. B. Freeman, R. Dale, "Assessing bimodality to detect the presence of a dual cognitive process", *Behavior research methods* **2013**, *45*, 83–97, DOI 10.3758/s13428-012-0225-x (cit. on p. 94).

[117] R. B. D. and, "Is Kurtosis Really "Peakedness?"", *The American Statistician* **1970**, *24*, 19–22, DOI 10.1080/00031305.1970.10478885 (cit. on p. 94).

[118] K. P. Balanda, H. L. M. and, "Kurtosis: A Critical Review", *The American Statistician* **1988**, *42*, 111–119, DOI 10.1080/00031305.1988.10475539 (cit. on p. 94).

[119] T. Wyszomirski, "Detecting and displaying size bimodality: kurtosis, skewness and bimodalizable distributions", *Journal of Theoretical Biology* **1992**, *158*, 109–128, DOI 10.1016/S0022-5193(05)80649-6 (cit. on p. 94).

[120] J. Behboodian, "On the Modes of a Mixture of Two Normal Distributions", *Technometrics* **1970**, *12*, 131–139, DOI 10.1080/00401706.1970.10488640 (cit. on p. 95).

[121] L. Wei, S. Durham, "The Randomized Play-the-Winner Rule in Medical Trials", *Journal of the American Statistical Association* **1978**, *73*, 840–843 (cit. on p. 100).

[122] D. L. Poepsel, D. A. Schroeder, "Joining groups: How resources are to be divided matters", *Group Dynamics: Theory Research and Practice* **2013**, *17*, 180–193, DOI 10.1037/a0033846 (cit. on p. 104).

[123] B. Simpson, O. Aksoy, "Cumulative advantage in collective action groups: How competition for group members alters the provision of public goods", *Social Science Research* **2017**, *66*, 1–21, DOI https://doi.org/10.1016/j.ssresearch.2017.03.001 (cit. on p. 104).

[124] Z. Hohman, O. Kuljian, "Why People Join Groups" in *Oxford Research Encyclopedia of Psychology*, Oxford University Press, **2021**, DOI 10.1093/acrefore/9780190236557.001.0001/acrefore-9780190236557-e-257 (cit. on p. 104).

[125] H. Pérez-Martínez, F. J. Bauzá, D. Soriano-Paños, J. Gómez-Gardeñes, L. M. Floría, "Emergence, survival, and segregation of competing gangs", *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2022**, *32*, 083114, DOI 10.1063/5.0084972 (cit. on p. 104).

[126] M. E. Newman, "Modularity and community structure in networks", *Proceedings of the national academy of sciences* **2006**, *103*, 8577–8582, DOI 10.1073/pnas.0601602103 (cit. on p. 105).

[127] S. Fortunato, "Community detection in graphs", *Physics reports* **2010**, *486*, 75–174, DOI 10.1016/j.physrep.2009.11.002 (cit. on p. 105).

[128] T. M. Pham, A. C. Alexander, J. Korbel, R. Hanel, S. Thurner, "Balance and fragmentation in societies with homophily and social balance", *Scientific reports* **2021**, *11*, 17188, DOI 10.1038/s41598-021-96065-5 (cit. on p. 105).

[129] S. Majhi, M. Perc, D. Ghosh, "Dynamics on higher-order networks: a review", *Journal of the Royal Society Interface* **2022**, *19*, 20220043, DOI 10.1098/rsif.2022.0043 (cit. on p. 105).

[130] W. Xu, Z. Zhang, "Linear opinion dynamics model with higher order interactions", *IEEE Transactions on Computational Social Systems* **2023**, *11*, 3627–3636, DOI 10.1109/TCSS.2023.3324144 (cit. on p. 105).

[131] N. W. Landry, J. G. Restrepo, "Opinion disparity in hypergraphs with community structure", *Physical Review E* **2023**, *108*, 034311, DOI 10.1103/PhysRevE.108.034311 (cit. on p. 105).

[132] A. Pister, M. Barthelemy, "Stochastic block hypergraph model", *Physical Review E* **2024**, *110*, 034312, DOI 10.1103/PhysRevE.110.034312 (cit. on p. 105).

[133] R. A. Holley, T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model", *The annals of probability* **1975**, 643–663, DOI 10.1214/aop/1176996306 (cit. on p. 106).

[134] S. Galam, "Minority opinion spreading in random geometry", *The European Physical Journal B-Condensed Matter and Complex Systems* **2002**, *25*, 403–406, DOI 10.1140/epjb/e20020045 (cit. on p. 106).

[135] P. Li, J. Cheng, Y. Chen, H. Wang, "A model of social network formation under the impact of structural balance", *Int. J. Mod. Phys. B.* **2016**, *30*, 1650051 (cit. on p. 107).