CRITICAL DYNAMICS IN COUPLED BOOLEAN NETWORKS WITH APPLICATIONS IN PLURIPOTENT STEM CELLS

By

CHRIS KANG

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Mathematics and Statistics

MAY 2023

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of CHRIS KANG find it satisfactory and recommend that it be accepted.

Nikolaos K. Voulgarakis, Ph.D., Chair

Elissa Schwartz, Ph.D.

Ilya Shmulevich, Ph.D.

Xueying Wang, Ph.D.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to the following people who have supported and guided me throughout this thesis in various ways:

First and foremost, my parents and grandma for their love and support. My advisor, Dr. Nikolaos K. Voulgarakis, for his invaluable guidance in the scientific work and the immeasurable patience with me in the personal life. To all the members of the dissertation committee, Dr. Elissa Schwartz, Dr. Ilya Shmulevich, and Dr. Xueying Wang, for their various contributions in this research. Emily Lewis and Dr. Lynn Schreyer for always helping me behind the scenes.

I would also like to thank the following colleagues and friends in no particular order: Dr. Jordan M. Culp, Dr. Jakob Streipel, Dr. Simon Kahan, Dr. Boris Aguilar, Joel Corser, Grace Grimm, the Amoscato Family, Ms. Maureen Poluszejko, Leon Kim, Aditya Joshi, Krittika Patil, and Kate S. Gibson.

Lastly, the Washington State University for providing the necessary resources and funding for this work throughout the years.

CRITICAL DYNAMICS IN COUPLED BOOLEAN NETWORKS WITH APPLICATIONS IN PLURIPOTENT STEM CELLS

Abstract

by Chris Kang, Ph.D.
Washington State University
May 2023

Chair: Nikolaos K. Voulgarakis

The spontaneous emergence of order is one of the fundamental features in complex systems. Embryogenesis and stem cell differentiation, for example, are highlighted by spontaneous phase transitions of disorder-order in its developmental cycle, resulting in a new functional life form. In this thesis, we examine the existence of characteristic signatures of disorder-order phase transitions in populations of isogenic cells whose gene regulatory networks are modeled as Boolean networks, motivated by the studies in pluripotent stem cells.

We present a model for coupled random Boolean network whose interaction rules are governed by the mulitilayer Ising Hamiltonian. Our approach allows for modeling multiple, biologically plausible intercellular signaling effects (paracrine, autocrine, and external fields). The model demonstrates an emergence of cell types in popula-

tions, which can be verified by three modes of analysis: (1) spectral decomposition of cell type distributions, (2) linear optimization method, and (3) a machine learning approach based on non-negative matrix factorization. Statistical analyses show that coupled random Boolean networks exhibit signatures of a second-order phase transition in cell type composition due to a combination of cell-cell cooperativity and intrinsic noise in its population. Near critical states in its parameters, stem cell populations undergo a spontaneous phenotypic transition, characterized by the symmetry-breaking events. Here, we show that this transition is possible through proper interplay of cell-cell cooperativity and intrinsic noise. Moreover, the system displays a first-order phase transition in the presence of external stimuli. We consider the effects of different sizes in control genes, specifically, the control kernel (CK) set. We see that dynamically pinning CKs with the mulitilayer Ising Hamiltonian can generate new cell types and demonstrate cell-to-cell variability in model simulations. Finally, we present that cells can collectively *self-tune* through a critical transition, which allows them to decide their fate. This behavior is seen with an internal dynamical system of a negative feedback between tissue heterogeneity and intrinsic noise, and the result is compared to recent experimental studies of mouse embryonic stem cells. Under strict conditions, the model captures experimentally-observed qualitative behaviors of multilevel transitions in cell type heterogeneity; that is, a unimodal-bimodal transition of cell states at the cellular and colony level.

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER ONE

INTRODUCTION

In the seminal work *The Origins of Order*, Stuart Kauffman challenges the *status quo* in the long-standing canons of evolutionary theory [1]. He argues that Darwin's theory of natural selection, where populations acquire and select for *useful* mutations over time, stands as an incomplete explanation of evolution, given that not all complex systems are equipped with the ability for successful adaptation. Seemingly, it suggests that there must be a characteristic feature in complex systems that allows one to perform successful natural selection, whereas others fail. He suggests that a fundamental feature of life is *spontaneous emergence of order* and proposes a new paradigm in evolutionary theory; a "marriage" of ideas between self-organization and natural selection. Borrowing from [2], we can define self-organization as the rise of nontrivial collective behaviors between interactions of multiple microscopic components[1]. Examples of studies in self-organization are commonly rooted in physics and statistical mechanics; i.e., the lattice formation of crystalline solids and the alignment of magnetic dipoles, which are systems well studied in statistical mechanics. It is no

[1]We note that historically, self-organization has a loose definition [2]. After an exhaustive survey of inconsistent use of the phrase "self-organization" in the literature, Halley and Wrinkler have proposed a unifying definition; that self-organization is a "dissipative nonequilibrium order at macroscopic levels, because of collective, nonlinear interactions between multiple microscopic components." This definition makes a distinction between self-organization and *self-assembly*, whose implication is lengthy and therefore not within the scope of this work.

surprise that Kauffman's suggestion of self-organization has a thermodynamic basis. In the context of evolution, it is reasonable to believe that complex systems capable of constraining the release of (thermodynamic) energy carry qualitatively "ordered" behaviors that can select for successful adaptation [3, 4]. To uncover the structural and behavioral mechanisms necessary for this emergence of self-organization and adaptive evolution in a system, one requires a dynamical system that includes the coupling of many components whose complexity can be easily studied. Kauffman proposed a simple model abstraction for the gene regulatory network: the Boolean network (BN) [5]. A Boolean network models the complex dynamics of a gene regulatory network with little assumption, and by analyzing *ensembles* of random Boolean networks (RBNs), high-level features can be captured. Recent decades of research in combination of *in silico* models and *in vitro* experiments have provided a solid ground for Kauffman's theory that systems best suited to adapt to evolutionary mutation and selection are those structured to exist on the phase boundary of order and chaos [6, 7, 8, 9, 10, 11, 12, 13].

Although many have focused on the dynamic complexity and phase transition of a single cell, few have studied collective behaviors of cells at a higher order. Phase transitions are the driving forces of many phenomena at the multicellular-level [14, 15, 16]. In this work, we derive biological motivation from pluripotent stem cells (PSCs). Pluripotent stem cells, such as embryonic stem cells (ESCs), form and differentiate

into specific cell lineages. Many recent experiments support the hypothesis that stem cells present heterogeneity in gene expression patterns at *population level* to differentiate [17, 18, 19]. That is, *intercellular* cooperation between microscopic agents is a critical feature in the long-term development of cellular fate in stem cells [20, 21]. To quote, MacArthur and Lemishka describe the developmental potency of stem cells "to be positively related to macrosocopic population entropy" [22] and have suggested that in order to understand the emergence in pluripotency, it is useful to consider stem cells in terms of their cellular macrostate [23, 24]. If this is indeed true, one can envision the existence of a macroscopic critical state and the characteristic phase transition in its system of PSCs (as depicted by Kauffman), and it entails further exploration.

Stem cell populations meet three necessary conditions to potentially facilitate a phase transition. First, each stem cell is an identical agent whose dynamics is determined by the molecular interactions between genetic regulators [25], which under most conditions maintain its own population [26]. Second, cells interact with each other and their extracellular environment through autocrine and paracrine signaling [27, 28, 29, 30, 31, 32]. Third, the dynamics of stem cell populations is subject to intrinsic noise owing to the molecular nature of intra- and intercellular interactions [33, 34, 35, 36]. *The essential question demanding an answer is whether it is indeed possible for stem cells to embrace and properly buffer intrinsic noise to reach*

the hypothesized critical state [23].

In our work, we observe signatures of spontaneous phase transition in multicellularity (i.e., populations of isogenic cells differentiating to specialized cells) which undergoes a series of symmetry-breaking events given different levels of complexity in its cellular cooperativity, intrinsic noise, external influence, and the structure of a BN. In our previous study, the formation of phenotypic (checkerboard) patterns, measured as long-term steady-state distributions of cells, was observed in model simulations of tissues of isogenic cells near the critical complexity (Lyapunov exponent), where cells were coupled with a linear threshold function [37]. In this work, we show that such pattern formation is not restricted to a checkerboard, as we explore macrostate behaviors of tissues through the lens of statistical mechanics. We develop a **multilayer Ising Hamiltonian** model that captures different modes of intercellular communication between BNs seen at the tissue level: paracrine, autocrine, and external field effects. Through model simulations, we show a second-order phase transition of coupled BNs in a heterogeneous cell type population. Specifically, we demonstrate that the transition from a highly excited state of *pluripotent population* [25, 38, 39] to one in the phenotypic ground state (differentiated cell types) is possible by proper buffering of intrinsic noise. We draw this parallel to the disorder-order transition in statistical mechanics. The emergence of new cell types can be verified through three different forms of analysis: (1) spectral decomposition of cell type distributions,

(2) linear optimization method, and (3) a machine learning approach based on non-negative matrix factorization. Additionally, we are able to show that coupled BNs are capable of exhibiting a first-order phase transition in different cell type populations in the presence of external stimuli.

In conjunction, we demonstrate that cells can collectively *self-tune* through a critical transition, which allows them to decide their fate. We are able to observe this collective phenotypic behavior with an internal dynamical mechanism, where there is a negative feedback between tissue heterogeneity and intrinsic noise. The model comparison shows that under strict restrictions, our results capture the qualitative behaviors of the multilevel unimodal-bimodal transition in gene expression patterns observed in experimental studies by Okamoto *et al.*. The differentiation process of mouse embryonic stem cells has been shown to require collective biophysical cooperativity and intrinsic fluctuation of cells in the colony [20].

This work is structured as follows. Chapters 2 and 3 serve as a conceptual overview of the tools necessary for the development of the coupled BN model. In Chapter 2, we review phase transitions, the Ising model, and its mean field approximation. Chapter 3 provides a background review of Boolean networks, and we the develop the notion of steady-state distributions as cell types. In Chapter 4, we introduce stochastic control kernel in BNs. In Chapter 5, we present the multilayer Ising Hamiltonian and its simulation methods. We follow with a discussion on different approaches to

analyzing the simulation outcomes of the multilayer Ising Hamiltonian, specifically cell type detection methods in Chapter 6 . In Chapter 7, we present the numerical results of the multilayer Ising Hamiltonian, and analyze phase transitions and symmetry-breaking in stem cell tissue populations with different sizes of control kernel. In Chapter 8, we show a simple *self-tuning* dynamical system which describes the negative feedback between cell type heterogeneity and intrinsic noise, and compare the model outcome with the Okamoto *et al.*'s mESC experimental results. We conclude and discuss future directions in Chapter 9. Parts of Chapters 4, 5, 6, 7, 8 have appeared in [40].

CHAPTER TWO

THE ISING MODEL

2.1 Introduction

Our long-term goal is to develop an all-encompassing intercellular dynamics for a model population of isogenic cells. Thus, we develop the tools necessary to model coupled Boolean networks.

The Ising model has had wide applications in modeling biophysical phenomena in the last two decades. In what would become the first of many energy-based models for biophysical properties of cell-cell interactions, Graner and Glazier presented a Potts Hamiltonian that replicated cellular arrangements (cell sorting) in embryogenesis [41]. This work had a significant impact on modeling morphogenetic behaviors. Subsequent work, aided by advances in computational speed, catapulted a new field of multiscale modeling known as the cellular Potts model (CPM) [42, 43]. CPM has been used to model many population-level cellular properties ranging from motility and displacement [44, 45, 46, 47] to chemical signaling (or adhesion) in tissues [48].

This chapter serves as the preliminary review of the Ising model, a special case of the Potts model, which we employ in Chapter 5. We examine the two-dimensional (2D) standard Ising model, as it is one of the simplest systems of identical agents that are coupled, whose phase transitions in the order parameter are well studied [49]. We approximate the solution of the 2D Ising model with the Weiss mean field theory

(MFT) approach[1] [50]. Finally, we briefly discuss the Potts Hamiltonian.

2.2 Ising model

The first development of the Ising model (also known as the Ising-Lenz model) dates back to 1925 when Ernst Ising derived the *partition function* for the magnetization of a one-dimensional ferromagnetic lattice in his doctoral dissertation [51]. Although Ising often received credit for his work, Wilhem Lenz, his research director, had proposed the idea five years prior. In 1941, Kramers and Wannier derived the first quantitative result for the 2D Ising model [52]. Although this was not an exact closed-form solution, it was the first of many approximation approaches to follow for the Ising model. We first discuss numerical approximations of the Ising model in Section 2.3. Then, we introduce the *Weiss mean field theory* approximation in Section 2.4[2] [53].

Consider a rectangular arrangement of atoms. For each atom, there is an associated magnetic dipole that determines the atomic spin. We consider only two directions for the spin $s_i \in \{-1, +1\}$, where $s_i = +1$ corresponds to "up" and $s_i = -1$ corre-

[1]Accompanying supplemental information related to the Landau theory of phase transitions is provided in Appendix C.

[2]Lattices of higher-dimension ($d > 2$) have been extensively studied, which are not within the scope of this discussion.

sponds to "down" for a spin site i. A *configuration* denotes the set of spin orientations of a lattice. We assume that each atom of i interacts with the nearest neighbors j's and denote all distinct coupling pairs of neighbors in the Von Neumann directions as $\langle i, j \rangle$. Then, the total energy, Hamiltonian, of the Ising model is given by

$$E = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{2.1}$$

where J is the coupling constant or the *interaction strength* between neighboring spins. Such system always attempts to evolve towards a ground state, i.e., the configuration that minimizes the energy. Examples of magnetic ground state are ferromagnetism, antiferromagnetism, and paramagnetism. For the Ising model with $J > 0$, the system prefers *ferromagnetism*, where the spins are aligned unidirectionally. With $J < 0$, the system prefers *antiferromagnetism*, where the spins are aligned in opposite directions. With $J = 0$, the spins are uncoupled.

If the system is exposed to an external field, a second term is introduced in its Hamiltonian.

$$E = -J \sum_{\langle i,j \rangle} s_i s_j - \mu H \sum_{i=1}^{N} s_i, \tag{2.2}$$

where μ is a magnetic moment, H is an external magnetic field, and N is the number of lattice sites. For convenience, we will use a single constant $\mu H = h$ to describe the

constant strength of an external field, where $h > 0$ indicates the magnetic field in the direction of positive spin. Thus,

$$E = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_{i=1}^{N} s_i. \tag{2.3}$$

The transition from a globally ordered state of ferromagnetism or antiferromagnetism to *paramagnetism*, where the spins are disordered due to thermal fluctuations (the strength of which is controlled by the magnet's temperature T), can be studied with *magnetization*,

$$m = \sum_{i=1}^{N} \frac{1}{N} s_i, \tag{2.4}$$

the excess spin (energy) proportion of a system. In the *canonical ensemble*, average spin of a site $\langle s_i \rangle$ is the thermodynamic average $\frac{\mathrm{Tr}\left(s_i \cdot e^{-\frac{E}{k_B T}}\right)}{Z}$ (k_B is the Boltzmann constant, and Z is the partition function)[3].

[3]Following the notation of [50], the details of the partition function are provided in Section 2.4.

2.3 Phase Transitions

We distinguish between two types of phase transition, also known as Ehrenfest classification[4] [54, 55]. *First-order phase transition* is characterized by a discontinuity in its free energy as it transitions from one phase to another. In contrast, *second-order phase transition* describes a continuous change of the ground state energy. The second-order phase transition is analytic under the first-derivative and discontinuous under the second-derivative. Both phase transitions are possible due to the variation of a control parameter. For the Ising model, in the first-order, the control parameter is h, while in the second-order, the control parameter is the temperature T. The second-order phase transition occurs at the critical temperature T_c also known as the Curie temperature [56]. The first-order phase transition occurs at $h = 0$, for $T < T_c$.

The 2D Ising model has been well studied for the precise reason that it demonstrates experimentally observed phase transitions. Lars Onsager's closed analytical solution to the zero-field 2D Ising model (the result being that the system undergoes a second-order phase transition at the critical temperature of $T_c = \frac{2}{\ln(1+\sqrt{2})} \approx 2.269$) provided in 1944 [57] has brought forth great advances in the theory of phase transitions [54]. For the purpose of this work, Onsager's solution is omitted here as it

[4]There is a rich history of how the original Ehrenfest classification of phase transitions was challenged and refined over the years [54]. We spare the details of scientific evolution and use colloquial definitions of phase transition.

has little relevancy in the model development. Instead, we present a more common approach to the solution to Equation 2.3: a numerical approximation with the Monte Carlo (MC) method. The Metropolis algorithm is a popularly used Markov chain Monte Carlo method to efficiently sample and sweep spin sites of large systems (Appendix A). In the Metropolis algorithm, temperature T is used in the Boltzmann distribution to either accept or reject its proposal to *flip* a spin. In other words, T serves as a thermal, *intrinsic noise* of the system.

In numerical simulations of the Ising model, the run time and size of the system are finite. This requires to approximate the *order parameter* of a system in some simulation time $\mathcal{T}$, after an equilibration time of t_{eq}, for a sufficiently large d number of simulations. For the magnetization of a configuration at time t,

$$m_t = \sum_{i=1}^{N} \frac{1}{N} s_i(t), \tag{2.5}$$

let us define the order parameter for the second-order phase transition as the ensemble average of the absolute magnetization after equilibration:

$$m = \frac{1}{d} \sum_{j=1}^{d} \frac{1}{\mathcal{T}} \sum_{t=t_{\text{eq}+1}}^{t_{\text{eq}+\mathcal{T}}} |m_t| \,. \tag{2.6}$$

Figure 2.1 shows the plot of the order parameter (m) as a function of temperature (T), where $d = 200$ simulations were performed and averaged for each temperature point

of the 2D Ising model without external field ($h = 0$). A lattice size of $N = 64 \times 64$ was used, and each simulation run was initialized with a random spin configuration. After time $t_{\text{eq}} = 2^{10}$ of the "burn-in" period for equilibration, magnetization was averaged over the span of $\mathcal{T} = 2^{11}$ MC steps. T was uniformly sampled from a range of $T \in [1.53, 3.28]$. In Figure 2.1, we observe that the magnetization drops to approximately 0 at $T_c \approx 2.269$, which is the critical temperature of the exact solution by Onsager.

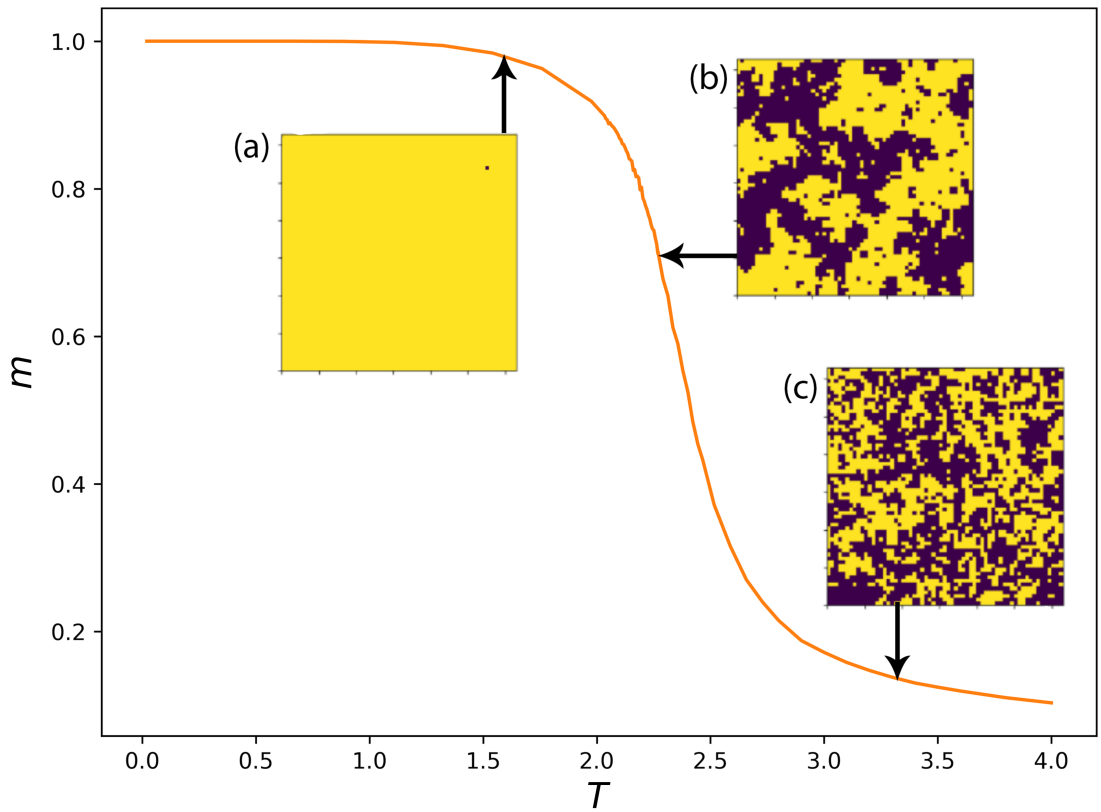


Figure 2.1: Second-order phase transition diagram (m vs. T) for a 2D Ising model with zero external field ($h = 0$). Each point corresponds to a temperature in the range of $T \in [0, 4.0]$ and the lattice size of 50×50. After a burn-in period of $t_{\text{eq}} = 2^{10}$ time steps, $\mathcal{T} = 2^{11}$ MC steps were used to compute the order parameter (m). The simulations were arbitrarily sampled from three respective temperature domains of $T < 2.269$, $T \approx 2.269$, and $T > 2.269$, and snapshots of the spin configurations were taken: (a), (b), and (c). Snapshots show different patterns in spins according to the temperature domain: (a) spins are mostly ordered (uniform orientation), (b) spins form fractal-like structure of spin islands, and (c) spins are disordered and show no decisive pattern.

Observing snapshots of spin configurations for individual simulations reveals important behaviors of the net magnetic moment of the lattice. For example, Figure 2.1 (a) shows the snapshot of a simulation with $T < T_c$ at time $t = t_{\text{eq}} + \mathcal{T}$. For low tem-

perature ("quenched state"), the system orients itself unidirectionally, as the spins prefer net ferromagnetism as the *entropy* of the system is low[5]. At $T \approx T_c$ (Figure 2.1 (b)), the spins change in pattern formation and arrange in *fractal*-like structures of spin islands [58]. With a high temperature ($T > T_c$), thus high in entropy of the system, the spins are randomly oriented (disordered) in their pattern (Figure 2.1 (c)).

When there is an external field ($h \neq 0$), the first-order phase transition can be observed. Define the order parameter for the first-order phase transition as the ensemble average of the magnetization after equilibration:

$$m = \frac{1}{d} \sum_{j=1}^{d} \frac{1}{\mathcal{T}} \sum_{t=t_{\text{eq}+1}}^{t_{\text{eq}+\mathcal{T}}} m_t. \tag{2.7}$$

Monte Carlo simulations of the 2D Ising model with varying external field (64×64 lattice) is shown in Figure 2.2. For three fixed temperature points of $T = 2, 2.269, 2.8$, and for 120 uniformly sampled values of $h \in [-1, 1]$, we see that the first-order phase transition occurs at the critical external field, $h_c = 0$ with low temperatures $T < T_c$ (i.e., $T = 2$); the net magnetization of the system jumps from $m = -1$ to $m = 1$.

Furthermore, when there is a sudden shift in the external field with little equilibration time, there is a *delay* in the switch in magnetization, which constitutes

[5]The Helmholtz free energy equation and Gibbs entropy are briefly mentioned in Section 2.4.

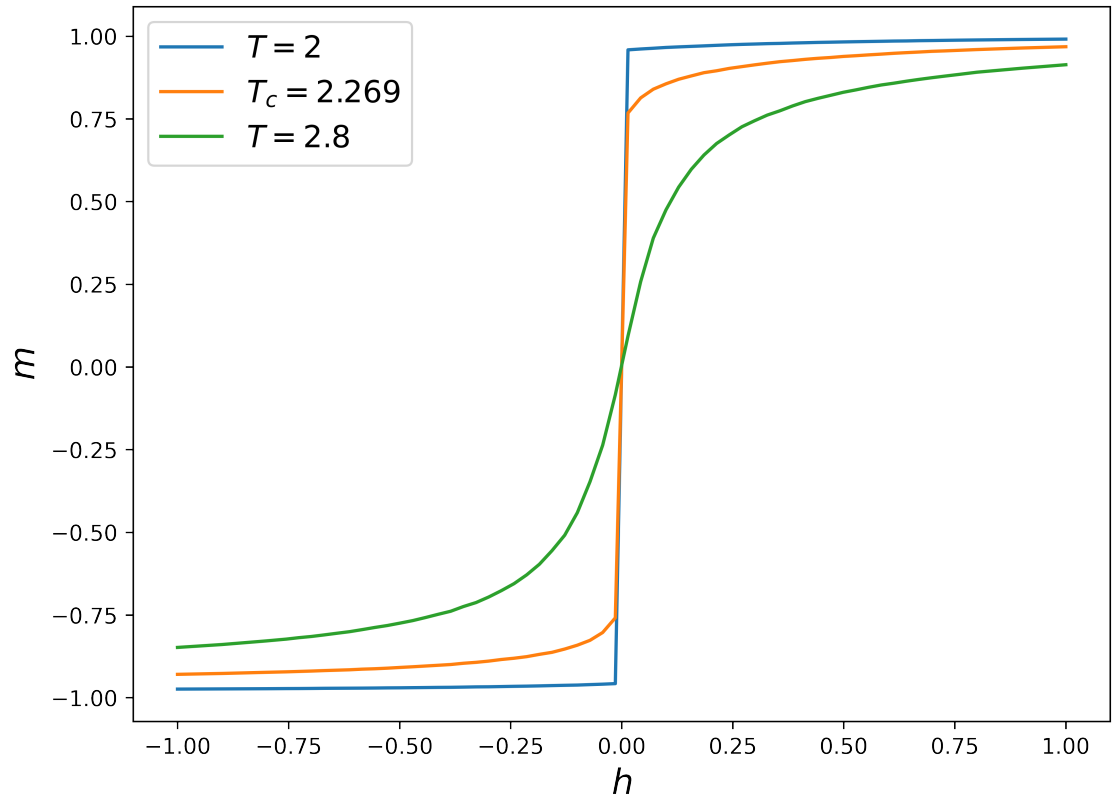


Figure 2.2: Simulations of the 2D Ising model with varying external field plotted as m vs. h. Three fixed temperature points were used: $T = 2, 2.269, 2.8$ for 120 values for h in the range of $h \in [-1, 1]$. For low temperatures $T < T_c$ (i.e., $T = 2$), the first-order phase transition can be observed at the critical external field $h_c = 0$.

hysteresis[6]. We reserve a more in-depth discussion of the first-order phase transition for Chapter 7.

We leave a brief mention of another frequently used tool to analyze the 2D Ising model: *magnetic susceptibility*. The magnetic susceptibility of the 2D Ising model

[6]See Appendix C.0.2 for further discussion on hysteresis.

can be calculated by

$$\chi = \frac{N}{T}\left(\langle m^2\rangle - \langle |m|\rangle^2\right), \tag{2.8}$$

where in this case, we define each term as the following:

$$\langle m^2\rangle = \frac{1}{d}\sum_{j=1}^{d}\frac{1}{\mathcal{T}}\sum_{t=t_{\mathrm{eq}+1}}^{t_{\mathrm{eq}+\mathcal{T}}} m_t^2 \tag{2.9}$$

and

$$\langle |m|\rangle^2 = \left[\frac{1}{d}\sum_{j=1}^{d}\frac{1}{\mathcal{T}}\sum_{t=t_{\mathrm{eq}+1}}^{t_{\mathrm{eq}+\mathcal{T}}} |m_t|\right]^2. \tag{2.10}$$

Magnetic susceptibility measures the sensitivity (or linear response) induced by a magnetic field.

2.4 Mean Field Theory

When deriving the analytical solution of an Ising model, we solve the *canonical partition function* because it provides thermodynamic information about a system, such as its free energy and average spins. The partition function for the Ising model

is

$$Z = \mathrm{Tr}\left(\exp\left(\frac{-E}{k_B T}\right)\right), \tag{2.11}$$

where E is the Hamiltonian from Equation 2.3. The derivation for Equation 2.11 is provided in Appendix B. With the solution to the partition function, it is possible to derive thermodynamic functions of the system using tools of statistical mechanics and determine whether a system undergoes a phase transition [51].

However, for many Hamiltonians, solving for the partition function is a challenging task. Specifically, the existence of *correlations* in the interactions between the spins (i.e., $-\sum_{\langle i,j\rangle} J s_i s_j$) makes it particularly difficult to solve analytically. While Onsager's close analytical solution was previously mentioned in Section 2.3, we discuss an approximation method instead, which is more suitable for complex models that cannot be solved analytically: the Weiss mean field theory approach [53]. The general strategy of MFT is to (1) take advantage of translational invariance of the magnetic dipole, (2) consider the average spin of neighboring sites, effectively decoupling the system to a single equation, and (3) finally approximating the solution to the partition function. We closely follow and owe much of the derivation to Utermohlen's work [50].

We express any degree of freedom, such as a spin variable s_i, in terms of its mean value $\langle s_i \rangle$ and its fluctuation $\Delta s_i = s_i - \langle s_i \rangle$. Thus, the spin interaction term $s_i s_j$

can be written as

$$s_i s_j = (\langle s_i \rangle + \Delta s_i) \cdot (\langle s_j \rangle + \Delta s_j) \tag{2.12}$$

$$= \langle s_i \rangle \langle s_j \rangle + \langle s_j \rangle \Delta s_i + \langle s_i \rangle \Delta s_j + \Delta s_i \Delta s_j. \tag{2.13}$$

We approximate by truncating the *quadratic fluctuation* term, $\Delta s_i \Delta s_j$. Then

$$s_i s_j \approx \langle s_i \rangle \langle s_j \rangle + \langle s_j \rangle \Delta s_i + \langle s_i \rangle \Delta s_j \tag{2.14}$$

$$= \langle s_i \rangle \langle s_j \rangle + \langle s_j \rangle (s_i - \langle s_i \rangle) + \langle s_i \rangle (s_j - \langle s_j \rangle) \tag{2.15}$$

$$= s_i \langle s_j \rangle + \langle s_i \rangle s_j - \langle s_i \rangle \langle s_j \rangle. \tag{2.16}$$

The system is translationally invariant; $\langle s_i \rangle = \langle s_j \rangle = m$. Thus,

$$s_i s_j = s_i \langle s_j \rangle + \langle s_i \rangle s_j - \langle s_i \rangle \langle s_j \rangle \tag{2.17}$$

$$= s_i m + m s_j - m^2 \tag{2.18}$$

$$= m[(s_i + s_j) - m]. \tag{2.19}$$

We can now apply mean field approximation on the Hamiltonian as the following:

$$E = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_{i=1}^{N} s_i \tag{2.20}$$

$$= -J \sum_{\langle i,j \rangle} m[(s_i + s_j) - m] - h \sum_{i=1}^{N} s_i \tag{2.21}$$

$$= -Jm \sum_{\langle i,j \rangle} [(s_i + s_j) - m] - h \sum_{i=1}^{N} s_i. \tag{2.22}$$

With the symmetry of the configurations $\sum_{\langle i,j \rangle} s_i = \sum_{\langle i,j \rangle} s_j$,

$$E = -Jm \sum_{\langle i,j \rangle} [2s_i - m] - h \sum_{i=1}^{N} s_i. \tag{2.23}$$

We can rewrite $\sum_{\langle i,j \rangle}$ as a single sum $\frac{q}{2} \sum_{i=1}^{N}$, where q is the number of neighbors in the lattice. Then,

$$E = -Jm \frac{q}{2} \sum_{i=1}^{N} [2s_i - m] - h \sum_{i=1}^{N} s_i \tag{2.24}$$

$$= -Jm \frac{q}{2} \sum_{i=1}^{N} 2s_i - \left(-Jm \frac{q}{2}\right) \cdot Nm - h \sum_{i=1}^{N} s_i \tag{2.25}$$

$$= \frac{NqJm^2}{2} - qJm \sum_{i=1}^{N} s_i - h \sum_{i=1}^{N} s_i \tag{2.26}$$

$$= \frac{NqJm^2}{2} - (h + qJm) \sum_{i=1}^{N} s_i \tag{2.27}$$

$$= \frac{NqJm^2}{2} - h_{\text{eff}} \sum_{i=1}^{N} s_i. \tag{2.28}$$

We combine the constant terms of Equation 2.27, $(h+qJm)$, to define a single effective constant term h_{eff}, the effective magnetic field influenced on the spins. Here, we have reduced our Hamiltonian from Equation 2.3 to a single site problem in Equation 2.28. The Hamiltonian has been decoupled, such that the magnetization no longer relies on the interaction of neighboring sites[7].

Now, the mean field approximation allows us to solve for the partition function. Noting that $\beta = \frac{1}{k_B T}$:

$$Z = \sum_{i=1}^{N} \exp\left(-\beta\left(-\sum_{\langle i,j\rangle} J s_i s_j - h\sum_{i=1}^{N} s_i\right)\right) \tag{2.29}$$

$$= \prod_{i=1}^{N}\left(\sum_{s_i=\pm 1}\right)\exp\left(\frac{-E}{k_B T}\right) \tag{2.30}$$

$$= \prod_{i=1}^{N}\left(\sum_{s_i=\pm 1}\right)\exp\left(\frac{-\left(\frac{NqJm^2}{2} - h_{\text{eff}}\sum_{i=1}^{N} s_i\right)}{k_B T}\right) \tag{2.31}$$

$$= \prod_{i=1}^{N}\left(\sum_{s_i=\pm 1}\right)\exp\left(-\frac{NqJm^2}{2k_B T}\right)\exp\left(\frac{h_{\text{eff}}}{k_B T}\sum_{i=1}^{N} s_i\right) \tag{2.32}$$

$$= \exp\left(-\frac{NqJm^2}{2k_B T}\right)\prod_{i=1}^{N}\sum_{s_i=\pm 1}^{N}\exp\left(\frac{h_{\text{eff}s_i}}{k_B T}\right) \tag{2.33}$$

---

[7]We note that the Ising model is one of the earliest Markov network models. In fact, it is a Markov random field with a 4-point clique (neighborhood) and thus, by the Hammersley-Clifford Theorem, it has a probability distribution of the form of a Gibbs distribution [59, 60]. While here, we used the mean field approach to decouple the spins in the Hamiltonian, other approaches such as the aforementioned use of cliques are also viable. However, the discussion for Markov random field is not within the scope of this section.

$$= \exp\left(-\frac{NqJm^2}{2k_BT}\right)\prod_{i=1}^{N}\left(\exp\frac{h_{\text{eff}}}{k_BT} + \exp-\frac{h_{\text{eff}}}{k_BT}\right) \tag{2.34}$$

$$= \exp\left(-\frac{NqJm^2}{2k_BT}\right)\left(2\cosh\frac{h_{\text{eff}}}{k_BT}\right)^N \tag{2.35}$$

Recalling the partition function, we substitute Z from Equation 2.35 into Equation B.8 to arrive at an explicit equation:

$$m = \frac{1}{N}\sum_{i=1}^{N}\langle s_i\rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{\text{Tr}\left(s_i \cdot e^{\frac{E}{k_BT}}\right)}{Z} \tag{2.36}$$

$$= \frac{1}{N}\frac{1}{Z}\sum_{i=1}^{N} s_i \exp\left(-\frac{E}{k_BT}\right) \tag{2.37}$$

$$\Rightarrow m = \tanh\frac{h + qJm}{k_BT}. \tag{2.38}$$

Equation 2.38, also known as the *self-consistency* equation[8] The self-consistency equation cannot be solved analytically. However, it can be easily graphed with a numerical software. The graph of net magnetization ($|m|$) of Equation 2.38 can be seen in Figure 2.3.

---

[8]This derivation requires a careful rearrangement of $\frac{\text{Tr}\left(s_i \cdot e^{\frac{E}{k_BT}}\right)}{Z}$ term in Equation 2.37. We again refer to [50] for the details of this step.

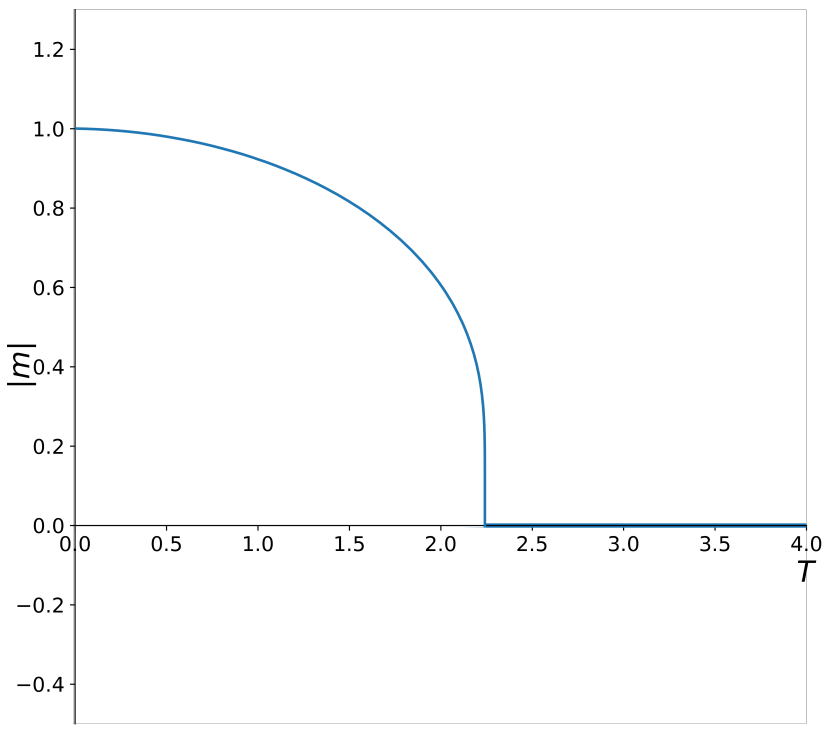


Figure 2.3: Net magnetization ($|m|$) of the self-consistency Equation 2.38. The second-order phase transition can be observed at the critical temperature, $T_c \approx 2.269$.

A major difference between the mean field approximation and the numerical approach using the Monte-Carlo method in Section 2.3 is the existence of residual magnetization (also known as the *magnetic tail*) for $T > T_c$. In Figure 2.3, $|m| = 0$ for $T > T_c$, while the magnetization in Figure 2.1 does not drop to $|m| = 0$ immediately [61].

CHAPTER THREE

RANDOM BOOLEAN NETWORKS

3.1 Introduction

In 1969, Stuart Kauffman introduced a model abstraction for gene regulatory networks (GRNs) known as Boolean networks [5, 62]. Over the years, Boolean networks have proven to be a powerful tool for modeling complex dynamics of gene regulations with little assumption. Using Boolean values to define gene activation and deactivation, and Boolean functions to represent regulatory interactions between input genes, we can generalize molecular mechanisms as a series of logic gates [63, 64, 65]. Modeling gene expression dynamics as a system of ordinary differential equations is a common practice. However, it becomes computationally inefficient and often unfeasible as the complexity of gene regulation increases; i.e., rise in the number of kinetic parameters. Fortunately, for simple regulations such as activation or inhibition, the interaction rules can be modeled as a discrete-time step function or, more conveniently, as Boolean dynamics. Diverse groups of biological systems have been successfully modeled with Boolean networks since Kauffman's introduction: notable works include yeast transcriptional factors [66, 67] and yeast cell cycle [68], segment polarity of *Drosophila melanogaster* [69], T-cell immune response [70], and various cancers [71, 72, 73, 74]. The adoption of BN modeling extends to other fields, such as supply chain networks in economics [75] and neural networks [76].

In this chapter, we provide a brief overview of Boolean networks, a discussion of the phase transition of Boolean dynamics, and the treatment of attractors as cell types. Then, we review the control kernel (CK) as a new mode of control strategy in Boolean dynamics. Finally, we provide a very brief description of probabilistic Boolean networks.

3.2 Boolean Networks

Formally, a Boolean network has n genes, $\mathbf{x} = \{x_1, \ldots, x_n\}$, and each gene can take on a binary state of "on" or "off" ($x_i \in \{0, 1\}$). Let $\mathbf{S} = \{1, ..., 2^n\}$ be the decimal representation of the Boolean states indexed from 1 to $\mathcal{N} = 2^n$ possible configurations of the state space.

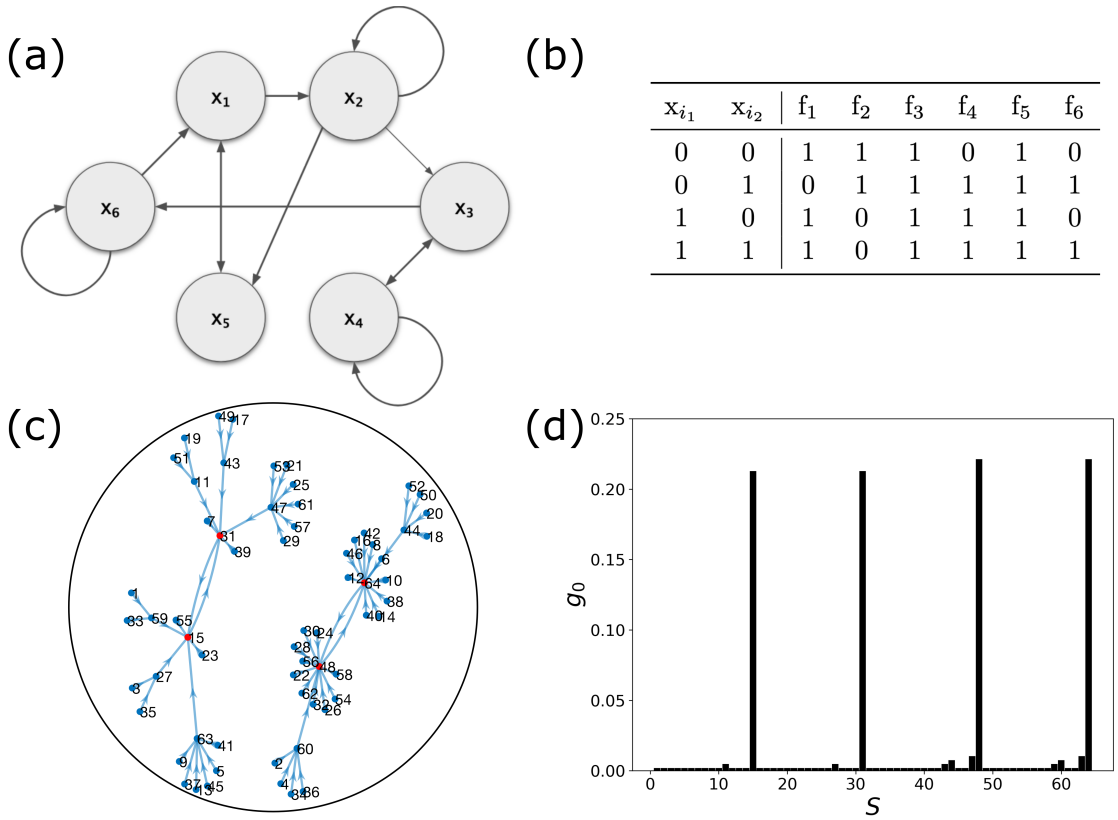
The states of these genes are "regulated" (or updated) by k_i genes, the number of input connections for the gene i, and the gene x_i is updated by its corresponding Boolean function f_i. Boolean functions are also known as *logical switching rules* and determine the state value from all possible combinations of k_i genes. Let $\mathbf{y}_i = \{x_{j_1}, x_{j_2}, ..., x_{j_{k_i}}\}$ be the set of regulating genes of x_i. Then, the state transition of the gene x_i is given by

$$x_i(t) = f_i(\mathbf{y}_i(t-1)), \tag{3.1}$$

where $f_i : \{0, 1\}^{k_i} \to \{0, 1\}$ in discrete time $t \in \{1, 2, ..., \mathcal{T}\}$, and $\mathcal{T}$ is the simulation

time. Then the Boolean *dynamics* describes the collective changes in the gene states of a network over time. Here, we note that we only discuss synchronous updating of the Boolean dynamics, where all the realizations occur at the same time. However, asynchronous Boolean networks are also well studied, as they are often considered closer to biological realism [77, 78, 79]. In random Boolean networks (RBNs), the value of f_i is assigned drawing from a probability distribution with bias of p for each input, where it takes on 1 with probability p and 0, otherwise, for all combinations of input genes $x_{j_1}(t), x_{j_2}(t), ..., x_{j_{k_i}}(t)$. With a fixed k for each gene, an RBN has 2^{2^k} possible update functions, and its *truth table*, a lookup table of the update functions with all possible input gene values ($\{f_1, \ldots, f_n\}$), is conveniently used to represent the rules of gene regulations in full.

Figure 3.1 shows different representations for a 6-gene Boolean network with connectivity of $k = 2$ and bias $p = 0.5$. Figure 3.1 (a), the *wiring diagram*, describes the assignment of the input genes, Figure 3.1 (b), the truth table, contains all information on the update rules, and Figure 3.1 (c) is the directed graph of all possible state trajectories, where each state is the decimal encoding of $\mathbf{x}$.



x_{i_1}	x_{i_2}	f_1	f_2	f_3	f_4	f_5	f_6
0	0	1	1	1	0	1	0
0	1	0	1	1	1	1	1
1	0	1	0	1	1	1	0
1	1	1	0	1	1	1	1

Figure 3.1: A 6-gene network with $k = 2$, and the bias of $p = 0.5$. (a) Wiring diagram of the network of the network. (b) Truth table of Boolean functions. (c) State transition diagram. (d) Steady-state distribution, $g_0(s)$ with perturbation $q = 0.1$.

3.3 RBN Dynamics: Ordered, Chaotic, Critical

Life forms have long been proposed to operate "at the edge of chaos" where the complexity of a dynamical system must be *robust* enough to withstand unwanted changes, but sufficiently *flexible* to adapt to external influence [6, 7, 8]. Torres-Sosa *et al.* provide a useful classification of this dualism into two analogous trade-offs in nature [9]. In the context of ontogenesis, there is a "developmental trade-off"

where cells must be robust enough to withstand environmental perturbation in gene expression patterns, but they must be flexible enough to adapt to environmental fluctuations. In Darwinian evolution, an "evolutionary trade-off" occurs when fundamental phenotypic traits must be robust enough in the face of genetic mutations, but *some* mutations are needed (and acquired) for new phenotypes to develop[1]. Torres-Sosa *et al.* show that their analogies of balance in trade-offs and the role of dynamic criticality also apply to *preservation* and *expansion* of Waddington's (phenotpyic) landscape [80]. We discuss attractor landscapes in greater detail in Section 3.4.

The balance in evolutionary properties has been a major source of motivation in studies of RBNs, and in doing so, many have explored ways to quantify stability of a network [10, 11]. The process typically begins with the question, "how much change does a small perturbation incur in the overall system?" In the context of Boolean networks, this corresponds to the impact that a bit-wise flip in gene state has on the network dynamics. The Hamming distance can be used to measure the number of bits needed to change from one string of binary values ($\mathbf{x}$) to another ($\mathbf{y}$) for binary strings of equal length: $H(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$. Then, through many numerical simulations, one can compare the average normalized Hamming distances of m pairs of randomly chosen gene states ($(\mathbf{x}, \mathbf{y})$ drawn from the state space, (S, S)) against

[1]Torres-Sosa *et al.* go on to show that dynamical criticality naturally emerges in Darwinian selection as "a consequence of evolution that favors evolvability [9]

the corresponding realizations of those states. The normalized Hamming distance of two binary gene states $(\mathbf{x}, \mathbf{y})$ is:

$$H(t) = \frac{1}{m} \sum_{(\mathbf{x},\mathbf{y}) \in (S,S)}^{m} \frac{1}{n} \sum_{i}^{n} H(\mathbf{x}(t), \mathbf{y}(t)). \tag{3.2}$$

The graph of $H(t)$ versus $H(t+1)$ is the *Derrida plot* and describes the effect of perturbation on the dynamical behaviors of Boolean networks [81]. Analysis of the Derrida plots reveals an important relation between the two key parameters of the RBNs (k and p):

$$k_c = 1/[2p(1-p)], \tag{3.3}$$

where k_c is the critical number of connectivity for critical dynamics. Equation 3.3 can be generalized as

$$s = 2kp(1-p), \tag{3.4}$$

where s is the *average sensitivity* of an RBN. Equation 3.4 describes the phase transition of ordered and chaotic regimes as a function k and p (Figure 3.2) [12]. When $s < 1$, the propagation of the perturbation has little impact on the overall behavior of the network, and its Derrida curve lies below the main diagonal of the Derrida plot

($H(t)$ vs. $H(t+1)$). When $s > 1$, the perturbation is extensive and, thus, the Derrida curve lies above the main diagonal of the Derrida plot, implying chaotic deviations. Finally, when $s = 1$, the network is in the critical domain and its Derrida curve operates near the main diagonal of the Derrida plot [77]. The network in Figure 3.1 is an example of a network in the critical regime.

Others have taken different approaches to arrive at the same result in descriptions of phase transitions of RBNs. An important work to mention is by Luque and Solé whose use of Boolean derivatives and *annealed approximation* established an order parameter: Lyapunov exponent, $\lambda = \log |2kp(1-p)|$ [82]. When $\lambda = 0$, an RBN is critical, when $\lambda < 0$, it is ordered, and when $\lambda > 0$, it is chaotic.

Experiments supporting the need for a critical phase transition in ontogenesis are now extensive [6, 9, 70, 66, 83, 84]. For more detailed background information on RBNs, the following excellent review articles are recommended: [77], [85], and [86].

3.4 Attractors as Cell Types

In most BN analyses, the *ensemble* approach is taken to extract robust features of GRNs with little assumption. Ensemble theory is a conceptual framework for statistical analysis of a collection of RBNs and allows the study of the influence of (statistically significant) structural features of the network [1]. By studying ensembles of Boolean networks, different characterizations of cell behavior can be established. In particular, through the lens of attractors and steady-state distributions of Boolean

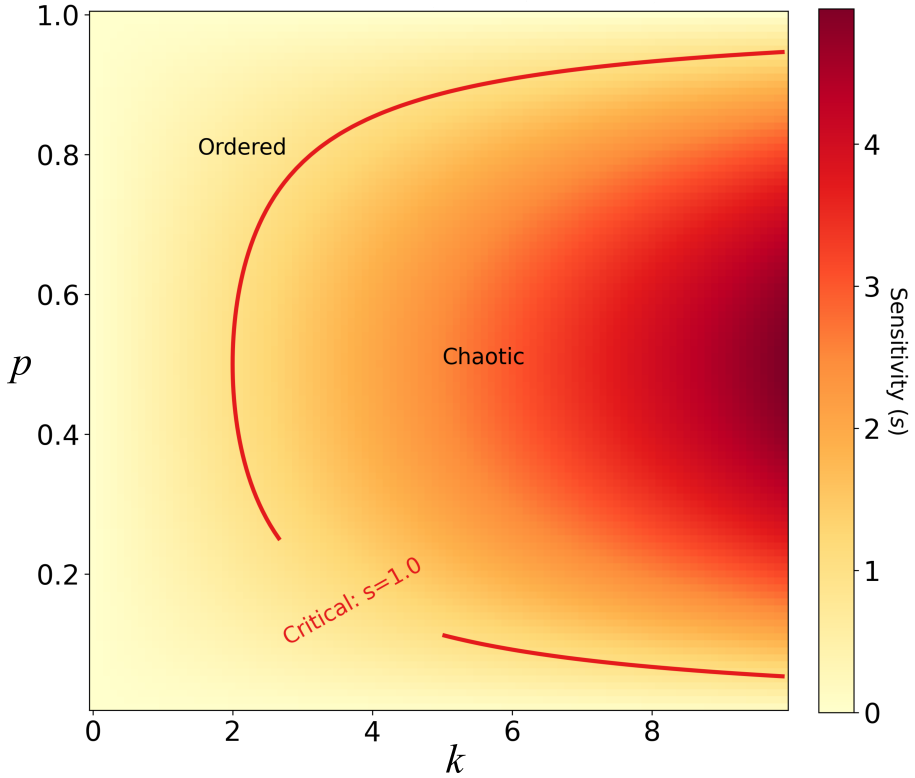


Figure 3.2: Phase diagram of RBN regimes: $s = 2kp(1-p)$. Sensitivities $s < 1$, $s > 1$, and $s = 1$ correspond to the ordered, chaotic, and critical regimes, respectively.

states, many have *defined cell types*, and this has been a source of rich discussions [87].

First, we establish definitions. An attractor is a set of static or periodically stable states of a cell. The basin of attraction is the set of states that leads to one attractor. In the case of Boolean networks, since the state space and dynamics are deterministic, the system necessarily enters the attractors with sufficient time. Significant effort has been put into finding efficient ways to find attractors for a large network through notable methods such as network reduction [78, 88]. The state transition diagram in

Figure 3.1 (c) shows two attractors, since the directed graphs of the states are two disjoint sets.

Kauffman first argued *that dynamical attractors model cell types* [1, 5, 62, 87]. This is a reasonable assumption, since no cell is defined by a single attribute, but by some combination or morphology, ontogeny, its functional and/or molecular composition, all of which are characteristics of long-term (stable-state) behaviors of a system [89]. While there is no unified clarity on how to define a cell type, many recent findings support Kauffman's thesis, and some even suggesting a more "dynamic" definition [90]. Huang *et al.* broadly characterized three *cellular states* (proliferation, differentiation, and apoptosis) as attractors [71, 91, 17]. In some ways, the cell *state* (or the phenotype) is a more useful description of what a cell *is* actually, given its long-term functional characteristics.

Huang *et al.* reasoned that stimuli on cells can lead to a reprogramming of a cell. The fact that different cell states can lead to reprogramming from one state to another suggests that there is a directional bias in cell-fate decisions [71, 91, 17, 92]. Here, the analogy to Waddington's epignetic landscape is typically drawn, where a cell differentiation process is described as a ball rolling down a series of valleys until one is stuck at some lower points [80]. Therefore, the lowest points of valleys are akin to attractors, signifying cell-fates or phenotypes. Thermodynamically this makes sense, as the lower physical locations of a landscape indicate low potential energy, while

higher locations (where Waddington's theoretical ball begins its journey) correspond to higher potential energy. At these high points, cells maintain the greatest potential (pluripotency) to differentiate and have long been associated with stem cells. In essence, cellular phenotypes are attractors in gene regulation [93]. Strong evidence corroborates this thesis with experiments in the form of cancer, capillary endothelial cells, and stem cells [71, 72, 94, 93, 95]. We note that since Boolean dynamics are discrete, it adds to the difficulty of formulating the (quasi-) potential energy landscape ($U(x)$) of attractors [96].

In the context of Boolean networks, the attractors are hypothesized to correspond to cell types. However, we look at this view with caution as, for instance, the deterministic nature of BNs can result in *attractors* that are highly sensitive to a small change[2]. One convenient way to observe the attractors of networks with a reasonably small number of genes is to numerically approximate the steady-state distribution through long-term simulations [97, 98, 99, 100, 101]. Computing steady-state distributions allows one to capture important information about a network, such as the *influence* of genes. Here, influence is the measure of long-term impact of a gene on others. We prefer this alternative definition of a cell type. Steady-state distributions characterize cell types (with mild stability assumptions), as most of the steady-state

[2]This brings about a difficult argumentation on what *are* the necessary characteristics of attractors to constitute as cell types (i.e., the length of the attractor, size of the system, etc.).

probability mass is concentrated in the attractors [99, 100]. We have previously defined steady-state distributions corresponding to RBNs with perturbation as cell type in our work [37] and adopt this for the rest of the work.

For an RBN, Monte Carlo simulations can be used to approximate the steady-state probability distribution of gene states by introducing a random perturbation into the network[3] [97, 98, 99, 100, 101, 102]. This can be implemented by considering a perturbation probability q for each time step and a random binary vector $\gamma = (\gamma_1, \gamma_2, ..., \gamma_n)$, where $\gamma_i \in \{0, 1\}$ and $P\{\gamma_i = 1\} = q$, such that

$$\begin{cases} \mathbf{x}(t) & = \gamma, \text{ with probability } (1 - (1-q)^n) \\ \mathbf{x}(t) & = \{f_1(\mathbf{y}_1(t-1)), f_2(\mathbf{y}_2(t-1)), ..., f_n(\mathbf{y}_n(t-1))\}, \text{ otherwise.} \end{cases} \tag{3.5}$$

With perturbation, there is a nonzero probability of arriving at any state, and thus the RBN is an ergodic Markov chain where the states converge after a sufficient time [97, 98]. The intrinsic noise is often present in biological systems due to the small size and molecular nature of interactions. For example, in *Escherichia coli*, the noise of gene expression modeled as stochastic dynamics has been well studied to be necessary in its regulation, fluctuation of transcription rate, and cell division [103, 104, 105, 106].

---

[3]Steady-state probabilities are often discussed in the context of probabilistic Boolean networks (PBNs). We refrain from the discussion of PBNs here, but leave with a reference [102].

The network in Figure 3.1 has the steady-state distribution $\mathbf{g}_0$, after $\mathcal{T} = 10^4$ simulation time, with a perturbation probability of $q = 0.1$. The attractor states can be identified by the peaks in $\mathbf{g}_0$, where the state frequencies are the highest.

CHAPTER FOUR

STOCHASTIC CONTROL KERNEL

4.1 Introduction

Controlling gene expression is a highly desirable action in many areas of molecular biology. For example, targeted drug delivery or the knockdown effect in gene therapy requires persistent external influence on a cell to achieve the desired gene expression. In the context of RBNs, there are numerous different control strategies. Shmulevich *et al.* initially proposed the optimal *intervention* strategy of probabilistic Boolean networks by altering Boolean functions [97]. For this type of intervention, one is often interested in finding the best candidate genes to intervene by minimizing the mean first passage time. Serra *et al.* have used the notion of knock-out or silencing of genes to analyze the effect of perturbations[1] [107, 108, 109].

In recent work, direct modification of gene states to *drive* the dynamics to desired attractors has been proposed [110]. We turn our focus to a specific notion of control, *pinning*, also known as the "node state override,", where the gene state is fixed for the entire duration of the time evolution. Pinning differs from the original formulation of intervention in that the Boolean functions are preserved for all genes except the

[1]The knock-out of a gene is here is not be confused with the *knockdown*. The knock-out refers to the complete elimination of a gene, while knockdown is the reduction in gene expression.

ones that are held static. Kim *et al.* formalized the *control kernel* of a network as "the minimal set of genes whose *pinning* reshapes the dynamics so that the basin of attraction becomes the entire configuration space [111]."

We show that by stochastically pinning the control kernel of a cell, we are able to create a new cell type, defined as a mixture of pinned steady-state distributions.

4.2 Control Kernel

Let us define a CK set, $\mathbf{z} = \{x_1, x_2, ..., x_r\}$, where $r \leq n$ is the number of pinned genes (that is, the cardinality of CK), and let $\mathbf{W}_r = \{1, ..., 2^r\}$ be the corresponding set of decimal encoding of the 2^r states in CK. Without loss of generality, CK can be arranged so that it is always the first r genes. We denote ordered sets of the pinned values of CK as $\mathbf{Z}_w$, where $w \in \mathbf{W}_r$. For example, $\mathbf{Z}_1 = \{x_1 = 0, x_2 = 0, ..., x_{r-1} = 0, x_r = 0\}$, $\mathbf{Z}_2 = \{x_1 = 0, x_2 = 0, ..., x_{r-1} = 0, x_r = 1\}$, etc... . Then, the dynamics of a BN is as follows:

$$\begin{cases} \mathbf{z}(t) & = \mathbf{Z}_w \\ x(t) & = f_i(\mathbf{y}_i(t-1)), \text{ for } i > r. \end{cases} \tag{4.1}$$

We reserve the index $w = 0$ to describe the unpinned dynamics, i.e., $\mathbf{Z}_0 = \{f_l(\mathbf{y}_l(t-1)\}_{l=1}^{r}$. In other words, when $w = 0$, Equation 4.1 is Equation 3.1. The pinning procedure naturally produces 2^r disjoint steady-state distributions, $g_w^{(r)}(s)$, whose

state spaces are equally partitioned by r. Here, we use the vector $\mathbf{g}_w^{(r)}$ to represent an individual cell type (steady-state distribution) for the CK index r. Naturally, in a pinned network, the number of "active" genes is reduced, as there is less dependency on the wiring (connectivity) of its regulators.

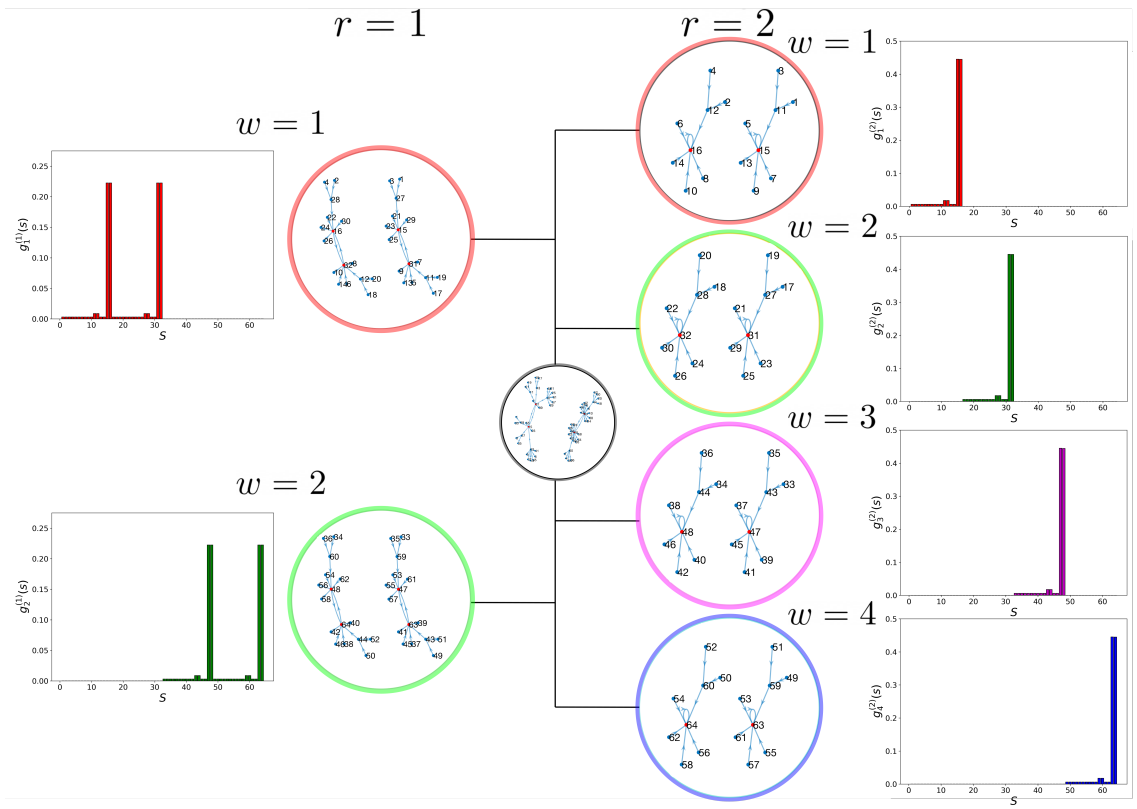


Figure 4.1: (Left) Pinned BN with $r = 1$; x_1 is either 0 or 1. The two resulting steady-state distributions $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$ have partitioned state spaces and preserve attractor points. (Right) Pinned BN with $r = 2$; x_1 and x_2 are pinned to either 0 or 1. The 4 resulting steady-state distributions $g_1^{(2)}(s)$, $g_2^{(2)}(s)$, $g_3^{(2)}(s)$ and $g_4^{(2)}(s)$, once again, have partitioned state spaces and preserve attractor points.

In Figure 4.1, we show the network (Figure 3.1) with the pin $r = 1$. By pinning the CK to $\mathbf{z}_1$ or $\mathbf{z}_2$, we generate two steady-state distributions $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$,

respectively. The resulting steady-state distributions are partitioned distributions that preserve features such as attractor states seen from the unpinned steady-state distribution of the network, $\mathbf{g}_0$. Similarly, when more than one gene is pinned (i.e., $r = 2$), the structural preservation of the features of $\mathbf{g}_0$ can be observed, albeit smaller in effect (i.e., smaller frequency in attractor state).

Figure 4.1 shows four disjoint steady-state distributions corresponding to combinations of different pin values for x_1 and x_2: $g_1^{(2)}(s)$, $g_2^{(2)}(s)$, $g_3^{(2)}(s)$, and $g_4^{(2)}(s)$. An important observation of note is that the scalar product of two distributions $g_i^{(r)}(s)$ and $g_j^{(r)}(s)$ for $i \neq j$, $\sum_{s \in S} g_i^{(r)}(s) g_j^{(r)}(s) \approx 0$, for a small $\gamma \neq 0$.

4.3 Stochastic Control Kernel

Let $\{\eta_t\}_{t \in \mathcal{T}}$, where $\eta_t \in \mathcal{W} \subseteq \{\{0\}, \mathbf{W_r}\}$, be a multilevel telegraph process that describes the transitions between the allowed states in CK. Here, we define stochastic CK as $\mathbf{z}(t) = \mathbf{Z}_{\eta_t}$, for some r. Let $g(s)$ be the new steady-state distribution as a result of $\mathbf{Z}_{\eta_t}$. Its distribution will be a mixture of all $\mathbf{g}_w^{(r)}$'s, i.e.,

$$g(s) = c_0 g_0(s) + \sum_{w \in \mathbf{W}_r} c_w^{(r)} g_w^{(r)}(s), \tag{4.2}$$

where $c_w^{(r)}$'s are the scalar weights of the respective $\mathbf{g}_w^{(r)}$'s. Here, we can consider c_w as *similarity weights* between the new cell type, $\mathbf{g}$, and $\mathbf{g}_w^{(r)}$'s.

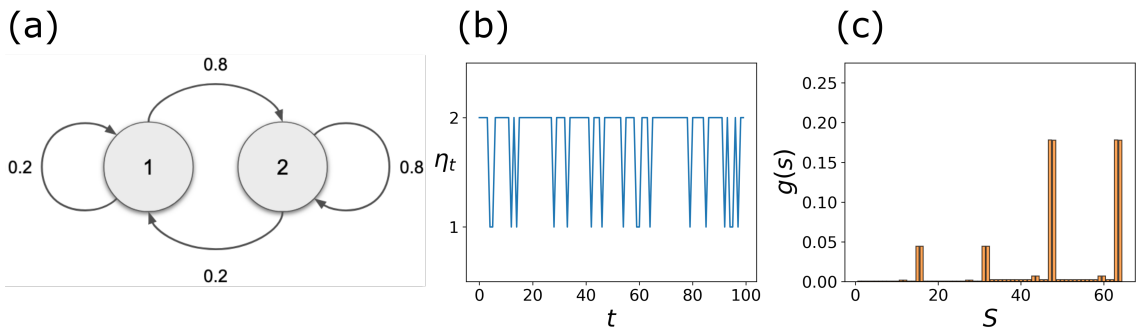


Figure 4.2: Stochastic pinning of the BN; (a) describes the transition diagram of $Q(1)$, (b) shows the first 100 sampled η_t's of the stochastic process, and (c) is the steady-state distribution of the stochastic CK, $g(s)$, obtained after $\mathcal{T} = 2^{16}$ simulation steps.

To illustrate this point, we look at the following example with the BN dynamics of the network from Figure 3.1. Consider a stochastic matrix, $\mathbf{Q}(t) = Q^t(1)$, whose entries are $\mathbf{Q}_{w,w_0}(t) = Pr(\eta_t = w, t|\eta_0 = w_0, t = 0)$, where $w, w_0 \in \mathcal{W}$, and $Q(1)$ is the one-step transition probability matrix. If

$$Q(1) = \begin{bmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{bmatrix}, \tag{4.3}$$

the CK is a two-state telegraph process for $\eta_t \in \{\mathbf{W}_1\} = \{1, 2\}$. Figure 4.2 (a) depicts the transition diagram of $Q(1)$. The given transition matrix results in higher frequencies of $\eta_t = 2$ than $\eta_t = 1$ (Figure 4.2 (b) shows the first 100 sampled η_t's). The resulting steady-state distribution of the stochastic CK dynamics, $g(s)$, is the linear combination of $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$ by Equation 4.2. Figure 4.2 (c) shows the steady-state distribution of the stochastic dynamics of CK obtained after $\mathcal{T} = 2^{16}$ time steps.

It is easy to see that $g(s)$ is "mostly" $g_2^{(1)}(s)$, which means the corresponding similarity weight $c_2 > c_1$. Here, we interpret this as that *the new cell type $g(s)$ resulting from stochastic pinning is a weighted mixture of cell types $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$.* This is true as long as the transition rates are not "too fast." In the next chapter, we take advantage of the disjoint property of the $\mathbf{g}_w$'s to develop a spectral decomposition method that provides these weights. Another method, non-negative matrix factorization (NNMF), is discussed in the case when $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$ are no longer disjoint, under different CK rules.

This brings about an important discussion on what cell differentiation is in the context of BNs. Bornholdt and Kauffman have previously stated that "if an attractor is a cell type, then differentiation is [the] flow from one to another attractor induced by noise or signal [87]." In Figure 4.2, we see a new cell type, $g(s)$, has been induced by the stochastic CK process. Certainly, this intervention by CK causes an overall shift in the distribution, creating a transitory state from one cell type, $g_0(s)$, to another, $g(s)$, which in turn is closer to $g_2^{(1)}(s)$. It is not improbable to surmise that a certain stochastic intervention, such as stochastic CK, captures the aforementioned "flow" in cell differentiation. In the next chapter, we show that this transition emerges spontaneously in coupled RBNs that operate near the critical point of a phase transition. This result, if true, supports the hypothesis that pluripotency in cell differentiation is a collective behavior in stem cell populations due to cell-cell cooperativity and strong

intrinsic fluctuations.

CHAPTER FIVE

MULTILAYER ISING HAMILTONIAN FOR COUPLED BOOLEAN NETWORKS

5.1 Introduction

While the significant majority of research on Boolean networks is focused on the regulation and expression patterns of single cells, recent efforts have been made to model multicellular systems. Serra *et al.* and Villani *et al.* have provided one of the earliest frameworks for modeling a collection of interacting Boolean networks akin to two-dimensional cellular automata, measuring average attractor periods with different levels of intensities of interaction [108, 112]. Using their model as the basis, others have explored the emergent behaviors and properties of coupled Boolean networks. Daminai *et al.* explored whether increased interaction between coupled Boolean networks contributes to expanding (or reducing) the variety of behaviors that are possible for a cell with a given genetic content [113]. Damiani *et al.* showed that short-distance interactions among Boolean models of genetic regulatory networks exhibit robust generic properties. Flann *et al.* measured information complexity (Kolomogorov complexity) of a tissue of RBNs when coupled in symmetric and orthogonal directions of interactions and determined that the formation of patterns is most information rich in the near-critical complexity domain [114]. We have previously modeled tissues of coupled Boolean networks with linear threshold functions,

where long-term steady-state patterns (phenotypes) were observed in networks of critical dynamics [37]. Finally, in the most recent study by Kim *et al.*, *multilayer* RBNs were investigated, where isogenic GRN ware coupled according to a random selection of topology *in silico* with activation rules from [114]. They showed that a multilayer RBN structure facilitated the production of a measured complexity, *antifragility*, in a system [115].

However, few have addressed the complex interplay of cell-cell cooperation and intrinsic noise required in the differentiation process of pluripotent cells. Cooperativity and intrinsic noise in stem cells have been shown to play an essential role in various behaviors, including motility, sorting, and most importantly differentiation to phenotypically heterogeneous populations of cell types [20, 21, 41]. We propose a coupled BN model whose intercellular dynamics is governed by a multilayer Ising model of interacting control kernels. The multilayer Ising model captures the cell-cell cooperativity and intrinsic noise necessary for a population of isogenic cells to self-organize to characteristically different cell types, defined by steady-state distributions[1].

In this chapter, we establish the multilayer Ising Hamiltonian and its numerical

---

[1]We make a careful distinction from the previous work on *guided* self-organization by Carlos Gershenson. In his work "Guiding the self-organization of random Boolean networks," Gershenson offers strategies to guide individual networks to attractors [116]. He summarizes a general guideline for altering BN parameters (p and k), depending on the dynamic regime to which it belongs, in order to steer the flow to desired attractors. Gershenson's self-organization is focused on fine-tuning. In Chapter 8, we show self-organization of stem cells without the aid of any tuning guidance.

simulation method. In Chapter 6, we discuss methods for detecting cell types from the coupled BN model, and in Chapter 7, we show that the multilayer Ising Hamiltonian demonstrates a second-order phase transition in the heterogeneity of the cell type population as a function of intercellular noise. Furthemore, we show that in the presence of external stimuli, hysteresis effects can be observed, indicating that a system of coupled BNs carry phenotypic memory.

5.2 Multilayer Ising Model

In multicellular systems, cellular interactions or signal transduction are autocrine or paracrine. Many have modeled Boolean networks with dedicated receptor genes reflecting this behavior [108, 117]. However, receptor genes can also be under the influence of some external field that contributes to the overall phenotypic outcome (i.e., drug delivery affecting a whole population). Furthermore, these effects on receptor genes need not be static in their gene expression.

Inspired by the stochastic pinning procedure from Section 4.2, we propose a multilayer Ising Hamiltonian for a multicellular system for $L = N \times N$ coupled Boolean networks. The Hamiltonian captures (1) interactions among neighboring cells (paracrine effect), (2) a cell's tendency to follow its own dynamics as an independent BN (i.e., feed back condition), and (3) external and/or autocrine effect (i.e., drug delivery). Each cell i is allowed to interact with a set of neighboring cells in Von Neumann directions denoted by $\mathbf{\Gamma}_i$. For each Boolean variable $\sigma \in \{0, 1\}$, define a linear trans-

formation $\bar{\sigma} = 2\sigma - 1 \in \{-1, 1\}$. We assume that the BNs interact with each other through their CK according to the multilayer Ising Hamiltonian:

$$H_r = -\sum_{m=1}^{r}\sum_{i=1}^{L}\left[\sum_{j\in\mathbf{\Gamma}_i} J_{ij}\bar{x}_{i,m}\bar{x}_{j,m} + h_0\bar{x}_{i,m}\bar{f}_{i,m} + h\bar{x}_{i,m}\bar{\psi}_m\right]. \tag{5.1}$$

Here, the CK size r determines the number of independent Ising layers employed. J_{ij} is the cell-cell interaction strength between the ith and jth CKs. h_0 and h represent strengths of *local external fields*, where they are a cell's tendency to follow its original cell's dynamics ($\bar{f}_{i,m}$) and an external intervention ($\bar{\psi}_m$) respectively.

In this work, we only consider $J > 0$ and assume $J_{ij} = J = 1$ for all CK interactions for simplicity. We focus primarily on the multilayer Ising Hamiltonian with $r = 1$ and $r = 2$, hereafter. The following summarizes the four possible cases of interaction behaviors characterized by the Hamiltonian:

1. When $h = 0$ and $h_0 = 0$, Equation 5.1 is the standard zero-field Ising model for the CK that describes paracrine signaling.

2. When $h_0 \neq 0$ and $h = 0$, Equation 5.1 is the standard Ising model with an external field. The external field could be a static or time-dependent.

3. When $h_0 = 0$ and $h \neq 0$, the cells interact with their Von Neumann neighbors as well as favoring to behave as their original uncoupled cells. When $h_0 \gg J$, Equation 5.1 behaves as independent BN dynamics.

4. When $h \neq 0$ and $h_0 \neq 0$, Equation 5.1 combines all three signaling effects of cases 1, 2 and 3.

Finally, the biological mechanism of a cell-cell communication or cell interaction with external stimuli is facilitated by diffusing molecular signals. To capture this stochastic process, we introduce an additional type of intrinsic noise in the system, denoted as T, which acts at the population level. The analog of T in standard statistical mechanics is the temperature of the system. Throughout this work, we reference and use intrinsic noise interchangeably with temperature. Thus, h_0, h, and T serve as the main control parameters for the model intercellular dynamics.

5.3 Simulation Method

The simulation was set on a 50×50 lattice with periodic boundary conditions. For each simulation run, RBNs with $n = 6$ genes with a fixed bias of $p = 0.5$ and connectivity of $k = 2$. Each network has an internal noise of $q = 0.1$. Different values of CKs ($r = 1, 2$) and external fields (h_0, h) were considered in these experiments.

A total of 65 temperature points (T) were sampled in the range of $T \in [0.01, 4.0]$ for the simulations. For each temperature point, $d = 200$ independent simulations were repeated. Each simulation had a burn-in period of equilibration for up to time $t_{\text{eq}} = 10^4$, and the magnetization data were collected after $\mathcal{T} = 10^4$ MC steps. The Metropolis Algorithm was used to simulate the multilayer Ising Hamiltonian (Ap-

pendix A).

For each MC step, there is (1) a sweep of the lattice for the CKs, according to the multilayer Ising Hamiltonian, followed by (2) updating of the gene states for all cells according to their RBN dynamics. Let the time step of the RBN dynamics be $t_{\text{RBN}} = a\Delta t$, where a is the time scale factor relative to the Monte Carlo step. Clearly, in the case where the time scale for the multilayer Ising Hamiltonian and the RBN dynamics are equal, $a = 1$. However, the time-scales of the two dynamics can be different and therefore, there may exist a time-scale separation in other forms of experiments. Without loss of generality, we employ the following simulation algorithm:

1. Begin the simulation: $t = 0$.

2. Randomly choose a cell and propose a new state(s) for the CK. This proposal state is accepted with probability $\min(1, \exp(-\Delta E/T))$, where ΔE is the energy (multilayer Ising Hamiltonian) difference between the current and the proposed state.

3. Repeat Step 2 for all other cells.

4. Update BN states for all cells according to their Boolean dynamics (Equation 4.1) for t_{RBN}.

5. Update simulation time $t = t + 1$ and repeat Steps $2 \sim 4$ until time $t = \mathcal{T}$.

Additionally, for each simulation run, CK-induced cell types $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ are computed *a priori* at $t = 0$ for post-simulation comparison. Upon the completion of a simulation, different cell type detection methods are utilized to measure the instantaneous cell type composition (Chapter 6) from a quantified order parameter (Chapter 7).

CHAPTER SIX

DETECTING CELL TYPES

6.1 Introduction

In Section 4.3, we showed that *dynamic* CKs can "probe" a steady-state distribution of a cell type, resulting in a new transitory cell type as a mixture of distributions. Let $\mathbf{G}(t)$ describe the instantaneous distribution of the gene states of *all* cells in a population induced by r dynamically pinned CKs, by the multilayer Ising Hamiltonian from Equation 5.1. Then, we can approximate this *mixture* distribution as $\tilde{\mathbf{G}}(t)$, a linear combination of the known cell types, the unpinned distribution $\mathbf{g}_0$ and the disjoint CK-induced distributions $\mathbf{g}_w^{(r)}$'s:

$$\tilde{\mathbf{G}}(t) = c_0(t)\mathbf{g}_0 + \sum_{w \in \mathbf{W}_r} c_w^{(r)}(t)\mathbf{g}_w^{(r)}, \tag{6.1}$$

where $0 \leq c_0 \leq 1$, $0 \leq c_w^{(r)} \leq 1$ and $c_0 + \sum_w c_w^{(r)} = 1$. We define the relative error of the approximation as $\epsilon(t) = ||\mathbf{G}(t) - \tilde{\mathbf{G}}(t)||/\mathcal{N}$, where $||.||$ is the Euclidean norm.

Ideally, we would like to identify $\mathbf{G}(t)$ with proportional contributions (c_0 and $c_w^{(r)}$'s) from each of the cell type $\mathbf{g}_0$ and $\mathbf{g}_w^{(r)}$'s. Depending on the terms involved in the model Equation 5.1, utilizing different detection techniques could prove to be computationally and numerically advantageous. In this section, we discuss three methods for classifying and detecting cell types from the mixture distribution $\mathbf{G}(t)$

for three different cases:

- $h_0 = 0$.
- $h_0 \neq 0$.
- When the error in approximation is high: $\epsilon(t) \gg 0$.

6.2 Spectral Decomposition: $h_0 = 0$

In the absence of the first local external field acting on the dynamic of the original cell (h_0) in Equation 5.1, the dynamics of the population is dictated by the network of interacting CKs and the external intervention. Thus, all CKs are expected to be in a state $w \in \mathbf{W}_r$. In other words, c_0, and $\mathbf{G}(t)$ can be only described by the only cell types induced by CK, $\mathbf{g}_w^{(r)}$'s. This assumption significantly simplifies finding $c_w^{(r)}$'s. Taking advantage of the non-overlapping (disjoint) property of $\mathbf{g}_w$, we define a set of orthonormalized vectors as $\hat{\mathbf{g}}_w^{(r)} = \mathbf{g}_w^{(r)} / \langle \mathbf{g}_w^{(r)}, \mathbf{g}_w^{(r)} \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product. Here, we can spectral decompose $\mathbf{G}(t)$ as the linear combination of normalized vectors, $\hat{\mathbf{g}}_w^{(r)}$'s:

$$\mathbf{G}(t) = \sum_{w \in \mathbf{W}_r} c_w^{(r)}(t) \hat{\mathbf{g}}_w^{(r)}, \tag{6.2}$$

where

$$c_w^{(r)}(t) = \langle \mathbf{G}(t) \ , \ \hat{\mathbf{g}}_w^{(r)} \rangle. \tag{6.3}$$

In summary, $\hat{\mathbf{G}}(t)$ is the normalized instantaneous distribution of the individual

cells in the population, and in the case of $h_0 = 0$, this distribution is characterized by CK induced cell types $\hat{\mathbf{g}}_w^{(r)}$'s with $c_w^{(r)}$'s, the corresponding weights. The conservation property $\sum_{w \in \mathbf{W}_r} c_w^{(r)}(t) = 1$ allows us to interpret the coefficients as the density of each cell type in the population.

We consider the special case of $r = 1$, and let N_1 and N_2 be the numbers of cells corresponding to the two cell types $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$, respectively, such that $N_1 + N_2 = L$ is the total cell population. Then, the density of each cell type population is given by $c_w^{(1)}$'s:

$$\frac{N_1}{L} = c_1^{(1)}, \quad \frac{N_2}{L} = c_2^{(1)}. \tag{6.4}$$

This allows us to quantify the relative difference of two cell-type populations at time t:

$$M_t = c_2^{(1)}(t) - c_1^{(1)}(t). \tag{6.5}$$

Clearly, $M_t \in [-1, 1]$. In other words, M_t *is the measure of cell type heterogeneity (composition of a population) at time* t*, where the scaled value* $M_t = -1$ *corresponds to a homogeneous population of a single cell type* ($\hat{\mathbf{g}}_1^{(1)}$)*, and* $M_t = 1$ *to another* ($\hat{\mathbf{g}}_1^{(2)}$). Furthermore, we can define the order parameter of Equation 5.1 as $\langle M \rangle = \langle |M_t| \rangle$, where $\langle . \rangle$ indicates the ensemble average. In this case, $\langle M \rangle$ belongs to the same

universality class as the 2D Ising model. From now on, we will refer to M_t as the instantaneous magnetization of coupled BNs.

6.3 Linear Optimization: $h_0 \neq 0$

With $h_0 \neq 0$ in Equation 5.1, the system has the tendency to follow the original Boolean dynamic, and $\mathbf{G}(t)$ is no longer a strict linear combination of disjoint orthonormal vectors $\mathbf{g}_w^{(r)}$'s induced by the CK. First, we note that when $h_0 \gg 1$, the external field dominates the multilayer Ising Hamiltonian which favors and results in activities of *uncoupled* Boolean networks. Thus, we are only interested in a reasonable range of values for h_0 (that is, $h_0 \approx 1$).

In the case where $\mathbf{G}(t)$ receives *reasonable* contributions from the control kernel set and h_0, finding the coefficients c_0 and $c_w^{(r)}$ in Equation 6.1 is a linear optimization problem with constraints:

$$\min \quad \|\mathbf{G}(t) - \hat{\mathbf{G}}(t)\| \tag{6.6}$$

$$\text{s.t.} \quad 0 \leq c_0, c_w^{(r)} \leq 1 \tag{6.7}$$

$$c_0 + \sum_w c_w^{(r)} \leq 1. \tag{6.8}$$

In cell type detection where the linear optimization method was needed, standard linear programming algorithms Interior-Point Algorithm and Sequential Quadratic

Programming (SQP) [118, 119, 120] were employed as they proved to be sufficiently fast and accurate for this work.

6.4 Non-Negative Matrix Factorization

For reasons such as a non-trivial Boolean network or strong influences of combinations of control parameters in Equation 5.1, the linear optimization method from Section 6.3 may return a poor approximation for $\mathbf{G}(t)$ with the cell types $\mathbf{g}_0$ and $\mathbf{g}_w^{(r)}$'s. Under those circumstances, we require a different method for cell type detection. To illustrate, we briefly discuss a machine learning based approach of non-negative matrix factorization (NNMF) here, however, other various decomposition methods for mixtures of distributions [121] could be applied.

NNMF is used to decompose a non-negative matrix $\mathbf{G}_{N\times\mathcal{T}}$ into a product of two non-negative matrices $\mathbf{\Omega}_{N\times K}$ and $\mathbf{E}_{K\times\mathcal{T}}$, where K is the rank of the decomposition to be determined. The matrix $\mathbf{\Omega}_{\mathcal{N}\times\mathcal{K}}$ provides $\mathcal{K}$ different steady-state distributions, which can also be interpreted as new types of cells. The corresponding time-dependent density numbers are stored in $\mathbf{E}_{\mathcal{K}\times\mathcal{T}}$.

CHAPTER SEVEN

NUMERICAL RESULTS OF THE MULTILAYER ISING HAMILTONIAN

7.1 Results and Discussions

We establish the ensemble average of absolute magnetizations over d simulations as the order parameter for the model coupled BNs for the case of $r = 1$:

$$\langle M \rangle = \frac{1}{d} \sum_{i=1}^{d} \frac{1}{\mathcal{T}} \sum_{t=t_{\mathrm{eq}}+1}^{t_{\mathrm{eq}}+\mathcal{T}} |M_t| \,. \tag{7.1}$$

With $r = 1$, $M_t = c_2^{(1)} - c_2^{(1)}$. For the order parameter $\langle M \rangle$, we utilize the spectral decomposition method from Chapter 6 to determine the tissue cell type composition, measured as population proportions or *fractions* of cell types corresponding to $\mathbf{g}_0$ and $\mathbf{g}_w^{(r)}$,s.

7.1.1 Second-order Phase Transition: $r = 1$, $h_0 = 0$, $h = 0$

Following the simulation method from Section 5.3, numerical experiments were carried out for tissue sizes of 16×16, 25×25, 50×50. 6-gene independent isogenic coupled BNs at 65 sample temperature points in the range of $T \in [0.01, 4.0]$, with $h = 0$, $h = 0$ for $r = 1$ were used. Boolean networks were constructed with parameters $p = 0.5$ and $k = 2$, and the ensemble average of $d = 200$ independent simulations was

taken for each temperature point T.

In the absence of external fields h_0 and h, and $r = 1$, Equation 5.1 is in the same universality class as the 2D Ising model; thus, with an equivalent order parameter regardless of the Boolean network of choice. With the spectral decomposition of steady-state distributions (we only have two choices of cell types $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$), we can see that the simulations have precisely the same characteristic behaviors of the second-order phase transition in the Ising model from Chapter 2 (Figure 7.1).

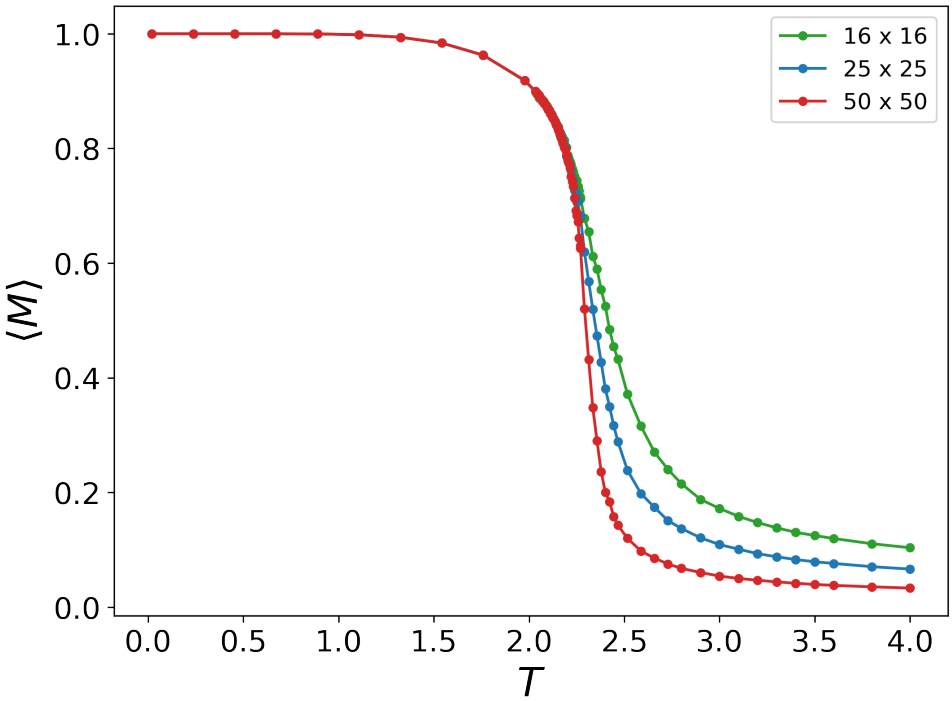


Figure 7.1: Simulation results for 6-gene, 50×50 coupled BNs for 65 temperature points for $T \in [0.01, 4.0]$, with $h = 0$, $h = 0$, $r = 1$. The ensemble averages of $d = 200$ independent simulations with BNs ($p = 0.5$ and $k = 2$) were taken to find the order parameter $\langle M \rangle$. The behavior of the order parameter $\langle M \rangle$ is exactly the same as the second-order phase transition of the 2D Ising model regardless of the choice of BNs.

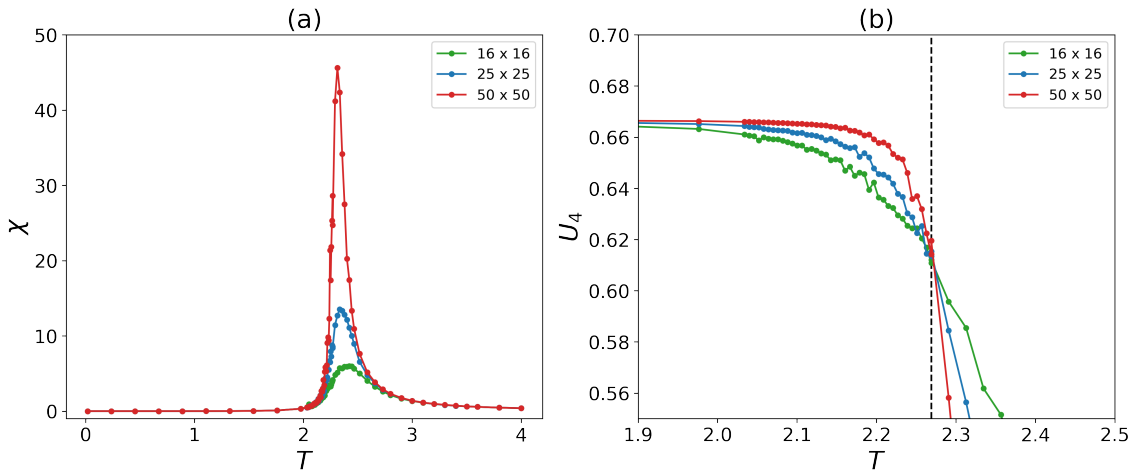


Figure 7.2: The susceptibility χ and the Binder cumulant U_4 were calculated for all simulations. (a) shows the maximum sensitivity at the critical temperature T_c, and (b) shows the intersection of Binder cumulants of three different tissue sizes, (16×16, 25×25, 50×50), also indicating the point of the critical temperature in the phase transition.

The susceptibility χ and the Binder cumulant[1] U_4 of different tissue sizes further show the exact characteristic behaviors of the second-order phase transition (Figure 7.2). As expected, maximum susceptibility is achieved at the critical temperature T_c in the standard Ising model (Figure 7.2 (a)). Another characteristic feature of the second-order phase transition is the finite-size scaling of the order parameter. This can be seen in Binder cumulants of various tissue sizes. Computed U_4's for 16×16, 25×25, and 50×50 tissue sizes intersect at the critical temperature of the Ising model: $T_c \approx 2.269$ (Figure 7.2 (b)).

[1]The Binder cumulant of the standard Ising model with zero external field is $U_4 = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$.

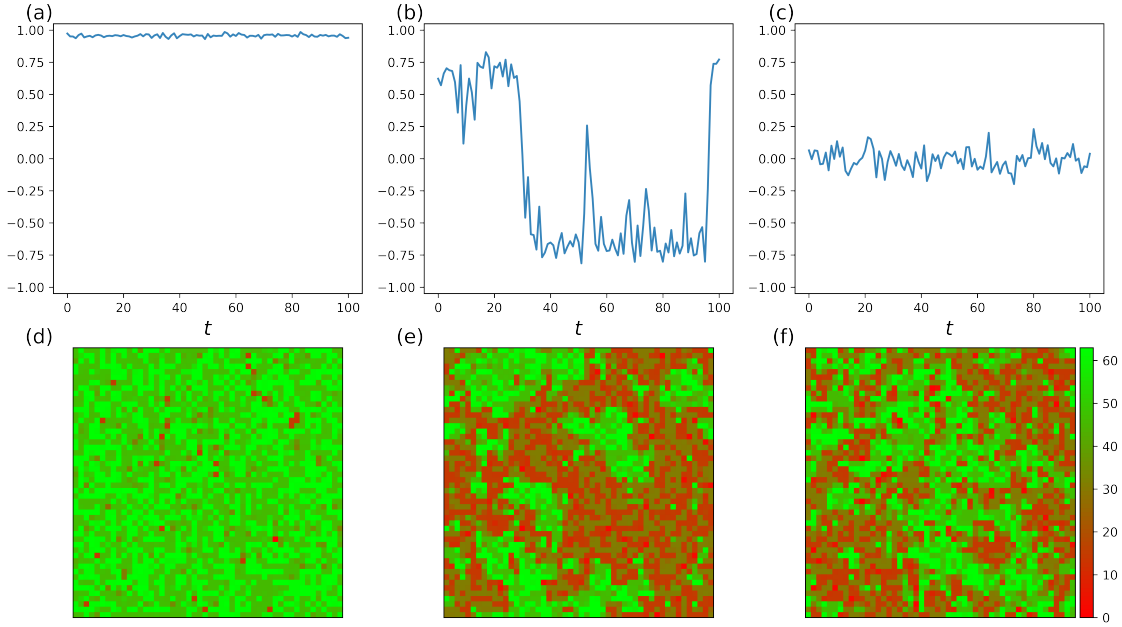


Figure 7.3: Example simulations of coupled BN simulations at fixed temperatures of $T = 1.8, 2.269, 2.8$ with $p = 0.5$, $k = 2$, $h_0 = h = 0$. First row: M_t for the duration of the post-equilibration time of $\mathcal{T} = 10^4$ MC steps. The t-axis has the time unit of 10^2 MC steps. For $T = 1.8$ (a), $M_t \approx 1$ for the duration of the simulation, there is a uniformity in the cell type population (to cell type $\mathbf{g}_2^{(1)}$). For $T = 2.269$ (b), there is a swing in M_t, as the cell population fluctuates from one cell type (i.e., $\mathbf{g}_2^{(1)}$) to another ($\mathbf{g}_1^{(1)}$). For $T = 2.8$ (c), $M_t \approx 0$ as the internal energy of the system is excited and disordered, and the cells are not fixed to either cell type ($\mathbf{g}_1^{(1)}$ or $\mathbf{g}_2^{(1)}$). In other words, it corresponds to the "pluripotent states" of stem cells. Second row: 50×50 lattice snapshots of coupled BN networks. Each cell state on the lattice ($s \in \mathbf{S}$) is color coded according to a green-to-red gradient scheme with corresponding values from 0 to $2^6 - 1$. Snapshots (d), (e), and (f) show the transition from differentiated to pluripotent cell states.

To illustrate further, three simulations of tissue size 50×50 at fixed low temperature ($T = 1.8$), medium temperature ($T = 2.269$), and high temperature ($T = 2.8$) are shown in Figure 7.3. Snapshots of the lattice were taken at $t = t_{eq} + \mathcal{T}/2$, where

the cell state $s \in \mathbf{S}$ was color coded according to a green-to-red gradient scheme (Figure 7.3 second row). Visually, the snapshots of the three temperatures resemble Ising spin configurations at different temperature domains (Figure 2.1): for $T = 1.8$, the pattern of cell states is uniform, for $T = 2.269$, the pattern forms fractal-like structures of spin islands, and for $T = 2.8$, the pattern is well mixed. The time evolution of M_t for the three simulations are shown in the first row of Figure 7.3. For $T = 1.8$, M_t remains consistent throughout the simulation, indicating that there is no change in homogeneity in the population of cell types. For $T = 2.269$, there is a swing in the range for $M_t \in [-1, 1]$. Here, as the temperature approaches T_c, the system reaches a *critical juncture*, where the cell differentiation is undecided between the two possible ground states. Thus, there are two different cell types at play, where there is a swing in dominance from one cell type population to another. Finally, for $T = 2.8$, M_t remains $M_t \approx 0$, indicating a failure to identify one of the two specific cell types. Here, the internal energy of this system is buffered to be "excited", and the cells are not fixed to either cell type ($\mathbf{g}_1^{(1)}$ or $\mathbf{g}_2^{(1)}$). In other words, for $T > T_c$, there is a persistent *heterogeneity* in the population of cell types, and it could be interpreted that $\langle M \rangle$ corresponds to the "pluripotent states" of stem cells.

For a statistical mechanics interpretation of this biologically motivated (and posed) numerical result, recall Waddington's epigenetic landscape from Section 3.3. Cells at the peak of a valley are in the highest potential state in the differentiation process. As

they move toward positions of lower states, cells end up in ground-state energy, eventually differentiating to specific phenotypes. $\langle M \rangle$ for $T > T_c$, precisely correspond to the "pluripotent states" of stem cells, as they are in highly excited cellular states induced by the system's intrinsic noise. *If* the temperature is reduced, the pluripotent state *flows* through a spontaneous transition (which has the signatures of the second-order phase transition) and eventually arrives at a differentiated state. Thus, the heterogeneity in the cell population describes a pluripotent state, which transitions to a homogeneous population of cell types with proper buffering of the internal noise. In this context, ***the disorder-order second-order phase transition in condensed matter physics is equivalent to the pluripotent-differentiated phase transition in the early stages of cell differentiation***. Much of this result supports MacArthur *et al.*'s view; to quote, "it is useful to think of pluripotency as a statistical property, similar to a macrostate in statistical physics. [22]"

7.1.2 First-order Phase Transition and Hysteresis: $r = 1$, $h_0 = 0$, $h \neq 0$

A new set of experiments was performed for coupled BNs under the influence of an external field $h \neq 0$. Simulations were carried out for 120 uniformly distributed values of $h \in [-1, 1]$ with three fixed temperatures: low temperature ($T = 1.8$), medium temperature ($T = 2.269$), and high temperature ($T = 2.8$). For each value of h, $d = 40$ simulations were repeated, where for each simulation run, a 6-gene RBN in the critical regime ($p = 0.5$, $k = 2$) was used. $d = 40$ simulations were run for each

h from $h = -1$ to $h = 1$ first in increasing order and repeated in reverse order. Here, we define the ensemble average of the magnetization (without absolute value) as the order parameter for the first-order phase transition:

$$\langle M \rangle = \frac{1}{d} \sum_{i=1}^{d} \frac{1}{\mathcal{T}} \sum_{t=t_{\mathrm{eq}}+1}^{t_{\mathrm{eq}}+\mathcal{T}} M_t. \tag{7.2}$$

$\langle M \rangle$ is plotted as a function of h (Figure 7.4).

Figure 7.4 shows that at $h = 0$, there is a spontaneous phase transition induced by the external field, where cells collectively switch from one cell type to another (observed by the transition in $\langle M \rangle$ from -1 to 1). Furthermore, for $T = 1.8$ and $T = 2.269$, there is a jump discontinuity, while, for $T = 2.8$, there is a continuous transition. The jump discontinuities for $T \leq T_c$ are characteristic features of the first-order phase transition.

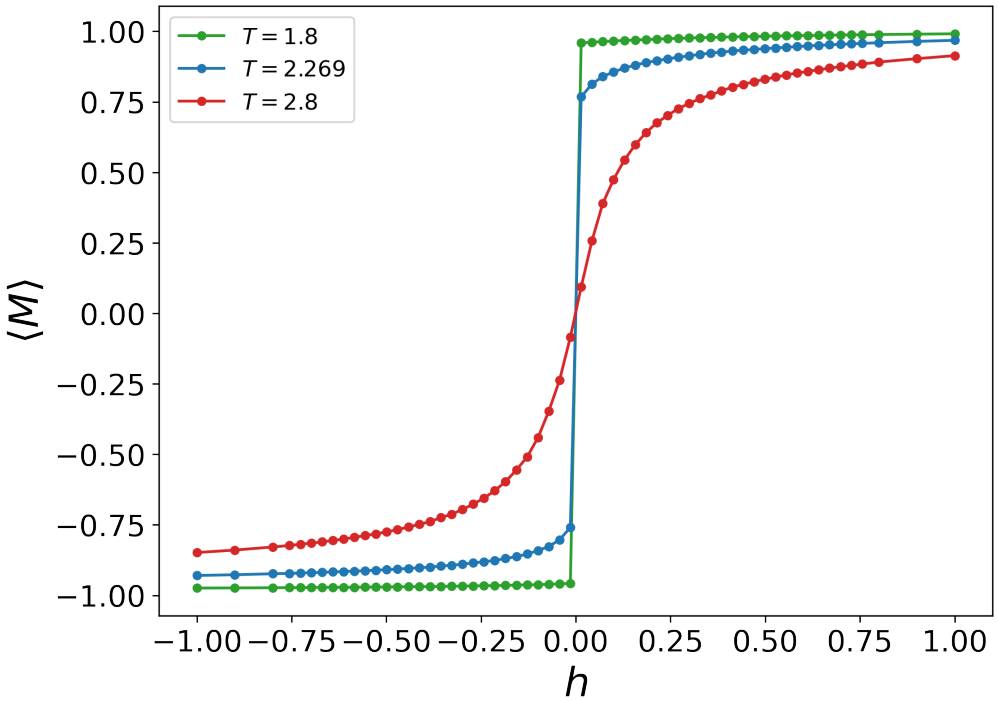


Figure 7.4: Simulation results for 6-gene, 50×50 coupled BNs in the critical regime with $p = 0.5$ and $k = 2$. 120 uniformly distributed points were drawn from the range $h \in [-1, 1]$, with $h = 0$ and fixed temperatures of $T = 1.8, 2.269, 2.8$. The ensemble average of $d = 40$ independent simulations was taken for each h for the critical RBNs to compute $\langle M \rangle$.

The same set of experiments was then carried out with no equilibration time and much shorter simulation time of $\mathcal{T} = 100$ MC steps. Once again, fixed temperature points of $T = 1.8$, $T = 2.269$, and $T = 2.8$ were used for 120 uniformly distributed values of $h \in [-1, 1]$. Simulations were run with increasing values of $h = -1$ to $h = 1$ first and then repeated in reverse order (decreasing values from $h = 1$ to $h = -1$). $\langle M \rangle$ is plotted as a function of h (Figure 7.5).

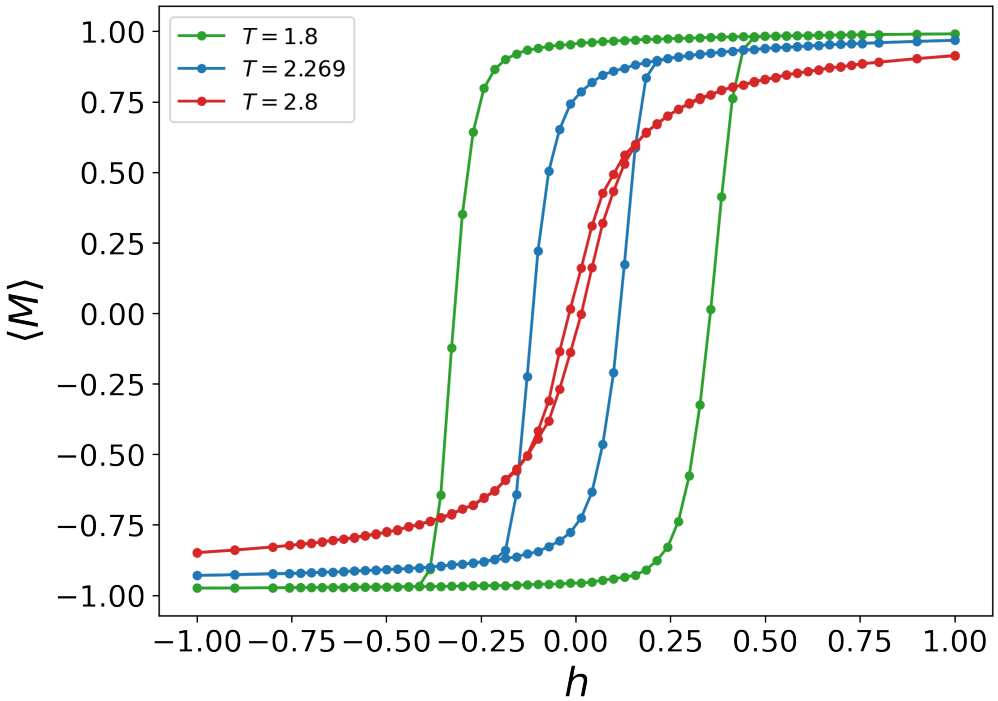


Figure 7.5: The ensemble average of $d = 40$ independent simulations was taken for each h with a much shorter simulation time of $\mathcal{T} = 100$ MC steps. The figure shows the delay in the first-order phase transition, characterized by hysteresis loops. The delay in the magnetization switch can be seen for all three fixed temperature simulations in both increasing and decreasing directions of $h \in [-1, 1]$.

Figure 7.5 shows the delay in magnetization for the three fixed-temperature simulations in increasing and decreasing directions of $h \in [-1, 1]$. The change and delay in magnetization, as a response to an external field, describes the toggle switch-like behavior of hysteresis loops in magnetic materials in a B-H curve[2]. This is precisely the behavior of the lagging 2D Ising model under external field. In the context of

[2]Magnetic flux vs. magnetic force curve.

pluripotent tissue of cells, the existence of hysteresis could implicate *phenotypic memory*. In electric circuits, hysteresis loops are associated with magnetic retention (or memory) in a magnetic field, and ferromagnets behave as *lagging* toggle switches. Here, the analogy of hysteresis loops could be made as a phenotypic memory in cell self-regulation [122, 14]. A possible avenue of application for this system is a drug delivery process in a tissue of cells, where the pharmacokinetics could be modeled as an external field in the multilayer Ising Hamiltonian. For instance, upon successful exposure to a drug, it is plausible that there is a toggle switch-like change across the tissue from undesirable cell states (cells requiring treatment) to desirable states (healthy cells). Ideally, this cell treatment process should be difficult to reverse. This robustness is a required feature of phenotypic memory at the tissue-level.

We can further understand these simpler cases of phase transitions with the Landau free energy equations provided in Appendix C. The first-order phase transition at critical $h = 0$ is qualitatively similar to the breaking of the symmetry in the energy of the quadratic expansion with a linear term in the Landau free energy. The hysteresis loops in Figure 7.5 are qualitatively similar to the cusp bifurcation (Appendix C.0.2).

7.1.3 Spontaneous Cell Differentiation: $r = 1$, $h_0 \neq 0$, $h = 0$

We now consider Equation 5.1 with $h_0 \neq 0$ and zero external field ($h = 0$), and consider the change in compositions of cell types. For the case of $r = 1$, we have $\mathbf{g}_0$, $\mathbf{g}_1^{(1)}$, and $\mathbf{g}_2^{(1)}$ possible cell types, where the corresponding composition fractions

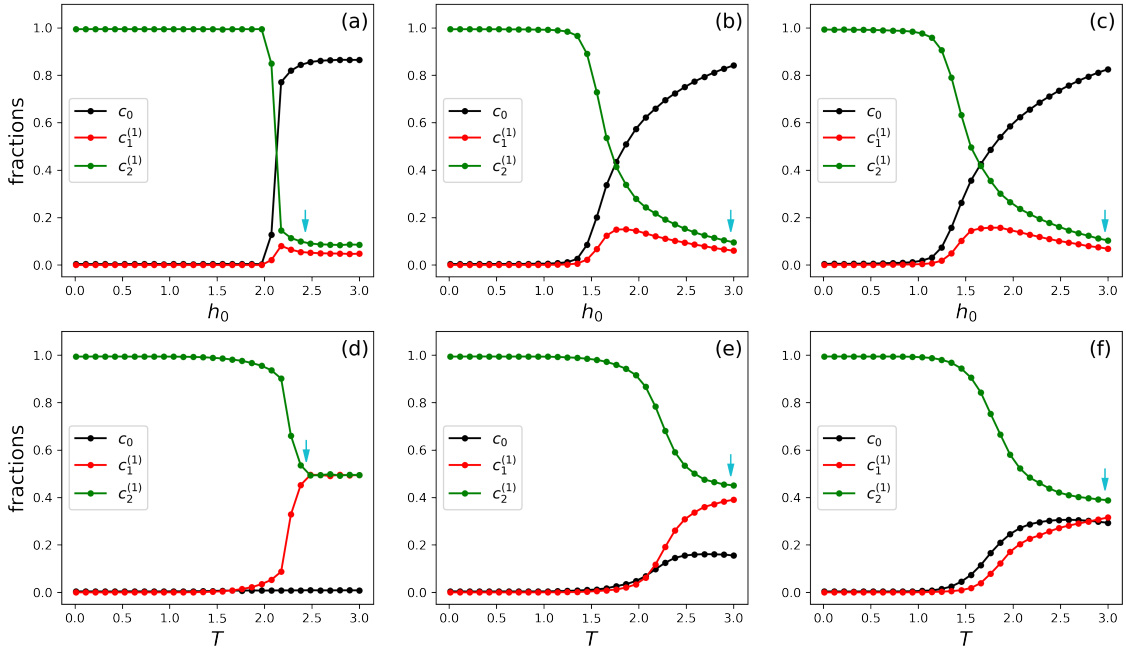


Figure 7.6: Spontaneous differentiation process for 50×50 tissues with $r = 1$. First row: Spontaneous differentiation process as a function of $h \in [0, 3]$ for (a) $T = 0.25$, (b) $T = 1$, and (c) $T = 1.25$. Second row: Spontaneous differentiation process as a function of $T \in [0, 3]$ for (d) $h_0 = 0$, (e) $h_0 = 0.5$, and (f) $h_0 = 1$. Cyan arrows indicate approximate starting points of spontaneous differentiation.

are $c_0 + c_1^{(1)} + c_2^{(1)} = 1$. We have simulated the multilayer Ising Hamiltonian for 30 uniformly distributed temperature points for $h_0 \in [0.01, 3]$ in descending order with fixed T values of $T = 0.25, 1, 1.25$ respectively (Figures 7.6 (a), (b), (c)). Similarly, we simulated 30 uniformly distributed $T \in [0.01, 3]$ in descending order with fixed h_0 values of $h_0 = 0, 0.5, 1$ respectively (Figures 7.6 (d), (e), (f)). For these simulations, we case study the Boolean network in Figure 3.1, and apply the linear optimization method from Section 6.3 to find c_0, $c_1^{(1)}$, and $c_2^{(1)}$.

For the first three cases where the temperature is fixed to $T = 0.25, 1, 1.25$ (Fig-

ures 7.6 (a), (b), (c)) and with a high initial value of $h_0 = 3$, we expect and see that the contribution of the original cell type $\mathbf{g}_0$ is dominant in Equation 5.1; thus c_0 is comparably high. Most of the cells in the tissue act as independent BNs. When $h_0 = 3$ drops to $h_0 = 0.01$, c_0, $c_1^{(1)}$, and $c_2^{(1)}$ pass through critical values, where the dominant contribution of c_0 begins to decrease and the three cell types are in play, and finally the tissue reaches a spontaneous *differentiation* point, where only one of the two cell types ($\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$) is chosen. This transition is an example of a *symmetry-breaking* event in which a disordered state, say pluripotent stem cells, has spontaneously differentiated into an ordered state of specialized cells. This disorder-order phase transition further corroborates past findings and characterization of pluripotent (multipotent) stem cells as a "balanced, undecided state" of multiple gene expression patterns [17].

Figures 7.6 (d), (e), and (f) show changes in cell type composition with respect to temperature with fixed values of $h_0 = 0, 0.5, 1$. The trivial case of $h_0 = 0$ describes the standard 2D Ising model seen in Section 7.1.1 (Figure 7.6 (d)). Only two cell types $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ are in play as differentiated cell types. In the beginning of the simulation with a high temperature of $T = 3$, there is an even distribution of cell types, thus $c_1^{(1)} = c_2^{(1)} = 0.5$. We see that with different h_0 values of $h_0 = 0.5$ and $h_0 = 1$ (Figures 7.6 (e) and (f)), the initial cell type compositions vary. As the temperature is reduced, the c_0, $c_1^{(1)}$, and $c_2^{(1)}$ go through critical transitions where the three systems differentiate spontaneously. Cyan arrows are placed to indicate

the approximate starting points of spontaneous differentiation for all simulations in Figure 7.6.

We carefully note that in all these experiments, the system always differentiates to $\mathbf{g}_2^{(1)}$. This decision is related to the structure of the original BN (Figure 3.1). The steady-state distribution of the original cell type $\mathbf{g}_0$ is not perfectly balanced in the state space, resulting in differentiation to $\mathbf{g}_2^{(1)}$ as the system parameters change. We expect that with a different BN construction, the system can expect to choose alternate paths in cell type differentiation.

7.1.4 Spontaneous Cell Differentiation: $r = 2$, $h_0 \neq 0$, $h = 0$

We consider the case of $r = 2$, where there are potentially 5 cell types $\mathbf{g}_0$, $\mathbf{g}_1^{(2)}$, $\mathbf{g}_2^{(2)}$, $\mathbf{g}_3^{(2)}$, and $\mathbf{g}_4^{(2)}$, are in play, and the corresponding composition fractions are $c_0 + c_1^{(1)} + c_2^{(1)} = 1$. Again, we have simulated the multilayer Ising Hamiltonian for 30 uniformly distributed temperature points for $h_0 \in [0.01, 3]$ in descending order with fixed T values of $T = 0.25, 1, 1.25$ respectively (Figures 7.7 (a), (b), (c)), and then $T \in [0.01, 3]$ in descending order with fixed h_0 values of $h_0 = 0, 0.5, 1$ respectively (Figures 7.7 (d), (e), (f)). The same Boolean network in Figure 3.1, and the linear optimization method from Section 6.3 was used to find composition fractions.

Here, we again see in Figures 7.7 (a), (b), and (c) that with a high initial value of $h_0 = 3$, c_0 is dominant. The other four cell types ($\mathbf{g}_1^{(2)}$, $\mathbf{g}_2^{(2)}$, $\mathbf{g}_3^{(2)}$, and $\mathbf{g}_4^{(2)}$) are also present, however, they contribute little to the overall composition. As $h_0 = 3$

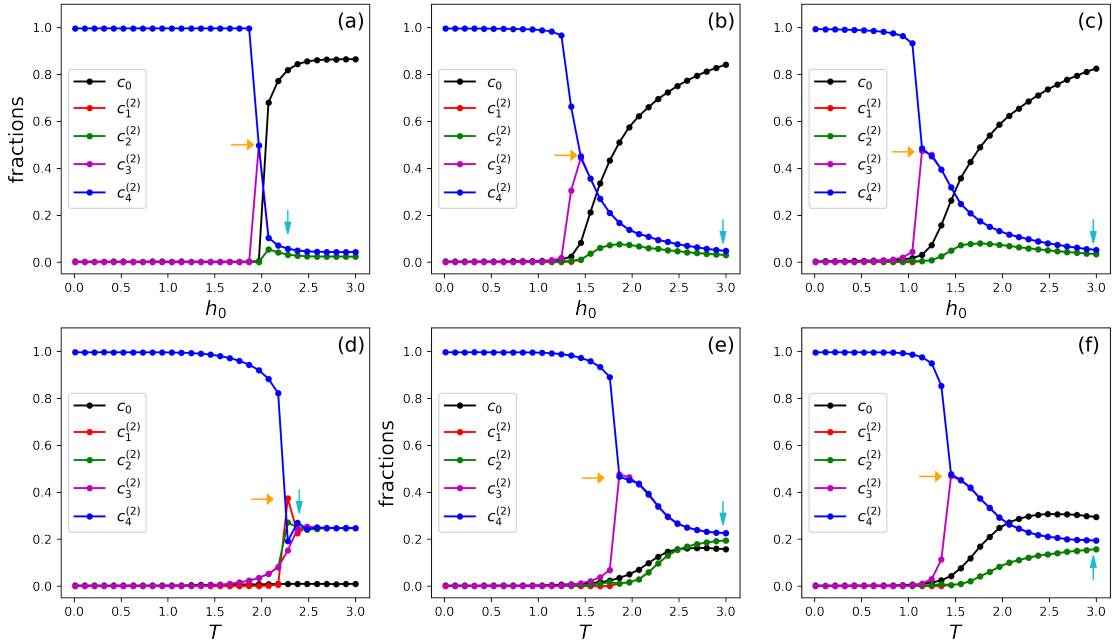


Figure 7.7: Spontaneous differentiation process for 50×50 tissues with $r = 1$. First row: Spontaneous differentiation process as a function of h_0 for (a) $T = 0.25$, (b) $T = 1$, and (c) $T = 1.25$. Second row: Spontaneous differentiation process as a function of T for (d) $h_0 = 0$, (e) $h_0 = 0.5$, and (f) $h_0 = 1$. Cyan and orange arrows indicate approximate starting points of the two spontaneous differentiations. Cyan is the initial differentiation, and orange is the latter differentiation.

is reduced to $h_0 = 0.01$, we witness two sets of symmetry-breaking events occurring as $\mathbf{g}_0$ differentiates initially into two separate groups $\left\{\mathbf{g}_1^{(2)}, \mathbf{g}_2^{(2)}\right\}$ and $\left\{\mathbf{g}_3^{(2)}, \mathbf{g}_4^{(2)}\right\}$, followed by a terminal differentiation to one of the four CK-induced cell types.

Figures 7.7 (d), (e), and (f) show similar changes in cell type composition (with respect to temperature) as Figures 7.6 (d), (e), and (f). With $h_0 = 0$ and the high temperature of $T = 3$, the tissue is in a pluripotent state with all four cell types equally present. We see that with different h_0 values of $h_0 = 0.5$ and $h_0 = 1$ (Figures 7.7

(e) and (f)), the initial cell type composition varies. As the temperature is reduced to $T = 0.01$, c_0 and $c_w^{(2)}$' go through two critical phase transitions where the systems differentiate spontaneously in two symmetry-breaking sets. Cyan and orange arrows mark the approximate starting points of the two spontaneous differentiation processes, where cyan is the initial differentiation, and orange is the latter differentiation.

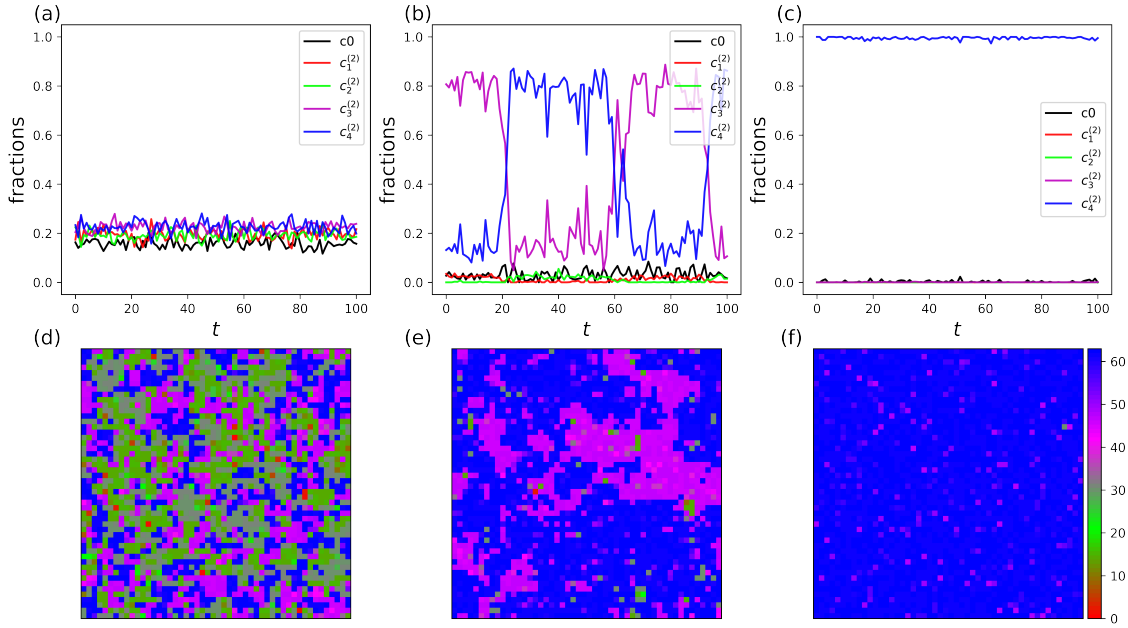


Figure 7.8: Example simulations of cell-to-cell variability in coupled BN simulations for $r = 2$ with a fixed $h_0 = 0.5$. First row: time evolution of cell type compositions as fractions at three different temperatures of $T = 3, 1.85, 1$, respectively, for time $t_{eq} = \mathcal{T} = 10^4$ MC steps. Only the post-equilibration time evolutions are shown, where the t-axis has the time unit of 10^2 MC steps. For the evolutions at $T = 3$ and $T = 1$, the fractions of the cell types remain consistent. For $T = 1.85$, the dominant fractions of cell types switch regularly between $\mathbf{g}_3^{(2)}$ and $\mathbf{g}_4^{(2)}$. Second row: snapshots of the three temperature points taken at time $t = 16 \times 10^3$.

We can observe the cell-to-cell variability in the multilayer Ising Hamiltonian

by examining the time evolution of cell type composition and tissue snapshots of simulations at a specific time. Figure 7.8 shows three independent simulations with $r = 2$, $h_0 = 0.5$, and three different temperature points of $T = 3, 1.85, 1$. The first row shows the time evolution of the cell type fractions for the 5 cell types. The second row shows the snapshots of the three temperature points taken at time $t = 16 \times 10^3$. In Figure 7.8 (a), where $T = 3$, there is an even distribution of cell types and this remains consistent throughout. The cell states, marked with a continuous color-scheme for the designated cell type colors, show a strong mixture and co-existence of cell types at the time of the snapshot (Figure 7.8 (d)). For $T = 1.85$, which according to Figure 7.7 (e) falls in the range between the first spontaneous differentiation and the second spontaneous differentiation, regularly changes in the composition of the dominant cell type between $\mathbf{g}_3^{(2)}$ and $\mathbf{g}_4^{(2)}$ (Figure 7.8 (b)). Furthermore, the snapshot indicates the formation of fractal-like islands for the cell type $\mathbf{g}_3^{(2)}$. These signatures of the second-order phase transition are consistent with the standard 2D Ising model. Figure 7.7 (e) shows that with $T = 1$, the tissue dynamics has passed the second spontaneous differentiation, resulting in a consistently homogeneous cell type (Figure 7.8 (c)). The snapshot further supports this homogeneity of cell types (Figure 7.8 (f)).

CHAPTER EIGHT

SELF-TUNED CELL DIFFERENTIATION

8.1 Introduction

In the development of mouse embryonic stem cells (mESCs), cells behave collectively and synchronously at the population level to transition from a pluripotent state to differentiated [123, 124]. Kalmar *et al.* have argued that pluripotency in mESCs is a state of *dynamic heterogeneity* of a population [34]. Through stochastic modeling of the three main transcription factors involved in the regulation of pluripotent states of embryonic cells (Sox2, Oct4, and Nanog) or, as described by them as *transcripitional noise*, Kalmar *et al.* have corroborated the well-supported view that the foundation of pluripotency lies in "the maintenance of a poise state for differentiation, a ground state" [38]. It is clear that one of the primary purposes of transcriptional noise in a regulatory network is to generate dynamic heterogeneities at the population level. Furthermore, heterogeneity in cell populations forms the basis for pattern formation in embryogenesis. Many others have expressed similar supporting evidence on the importance of intrinsic noise in maintaining cell-type population heterogeneity in a wide range of contexts [21, 122, 125].

Another characteristic feature of stem cells is their ability to maintain dual capacity for self-renewal (maintenance) and differentiation at the population level through a series of regulatory mechanisms and metabolic pathways [126, 127, 128]. An ex-

ample of a self-tuning mechanism is the AKT-mTor nutrient sensing pathway in stem cells. AKT, also known as Protein kinase B (PKB), and mTOR, which regulates cell proliferation and apoptosis, play a crucial role in the regulation of stem cell energy production by simultaneously suppressing certain processes (oxidative phosphorylation) and boosting others (such as glycolysis) [128]. This balance, which is self-regulated, has been shown to be essential for the maintenance of stem cells in various tissues, and it is *self-tuned* and coordinated at the population level.

Here, we develop a simple internal ODE feedback model that *self-tunes* one's pluripotent tissue to differentiated in response to its instantaneous population heterogeneity (M_t) and the intrinsic noise of the system (T). We identify Okamoto *et al.*'s observation of mESCs and their *unimodal to a bimodal* transition in the gene expression levels of the Nanog and Oct4 states as a potential avenue of application for the developed model.

8.2 Model Implementation

In Chapter 7, we showed that by manually tuning T and h_0 through the critical values, the system undergoes a series of symmetry-breaking events. In this section, we demonstrate that cells can collectively *self-tune* through a critical state that allows them to decide their fate. For simplicity, we study $r = 1$ only and neglect the local fields in Equation 5.1. This limit is equivalent to the standard Ising model. We recently showed that a mean-field approach of the Ising model with a negative

feedback mechanism drives the system through a supercritical pitchfork bifurcation that can be interpreted as a cell fate decision [129]. We apply this approach to the full Ising Hamiltonian EquationEquation 5.1. Here, we utilize the instantaneous magnetization

$$M_t = c_2^{(1)}(t) - c_1^{(1)}(t) \tag{8.1}$$

to measure the heterogeneity of a tissue as a dynamical system. As discussed in Chapter 7, a perfect mixture of $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ ($M_t = 0$) corresponds to the pluripotent state of the cell described by Huang *et al.* [17]. $M_t = -1$ and $M_t = 1$ correspond to homogeneous populations of cell types $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$, respectively.

We consider an internal mechanism that allows the heterogeneity of a population, measured in terms of instantaneous magnetization of the tissue, M_t, to regulate the intrinsic noise of the population (temperature T):

$$\frac{dT}{dt} = |M_t| - \alpha T, \tag{8.2}$$

where α is the relaxation coefficient. That is to say, Equation 8.2 captures the negative feedback response between cell-cell cooperativity and its intrinsic gene expression noise.

The model Hamiltonian for coupled Boolean networks (Equation 5.1) with $h = 0$

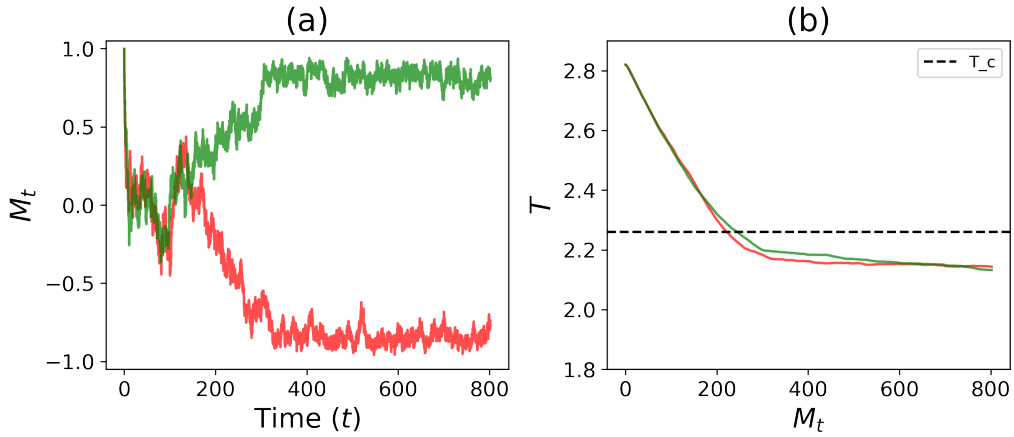


Figure 8.1: Simulations of a 32×32 Ising model with a self-tuning feedback equation (Equation 8.2). Here, $h = 0$, $h_0 = 0$, and $\alpha = 0.8$. (a) Magnetization trajectories show two systems are driven to $M_t \approx 0$ immediately upon initialization and eventually *self-tunes* to two different homogeneous cell types ($M_t \approx -1$ or $M_t \approx 1$). (b) The simulations show that the temperature drop slows as it descends below the critical temperature (T_c) and eventually reaches a steady temperature. Combining (a) and (b), we see that as the feedback temperature reaches the Ising critical temperature (T_c), the magnetizations begin to diverge and tissue differentiates homogeneously to one of the two possible cell types. The time unit is 10^2 MC steps with $N_{MC} = 10$.

and $h_0 = 0$, combined with the internal temperature-magnetization feedback mechanism (Equation 8.2) was simulated on tissues of size 32×32. The system was set to evolve for up to $t = 8 \times 10^5$ MC steps where $N_{MC} = 10$. Equation 8.2 was solved using the Euler method with a step size of $\Delta\tau = 5 \times 10^{-7}$, and the parameters $\alpha = 0.8$ were fixed. The model Boolean network from Figure 3.1 was instantiated for all cells with internal noise of $q = 0.02$.

The simulation of the model begins with random gene states for all cells in the tissue, except the CK node, which, without loss of generality, is set: $\mathbf{x}_c = \{1\}$. We

choose a high starting temperature of $T = 2.8$ for the system because it generates a natural state of hypothesized heterogeneity in pluripotent cells in the early stage of the embryonic stem cell cycle. Through negative feedback on instantaneous magnetization, Equation 8.2 then self-tunes the system towards the critical and then sub-critical temperatures, where symmetry breaking triggers spontaneous differentiation.

Figure 8.1 (a) shows the time evolution of magnetization (M_t) for two independent simulations. Here, it can be seen that both simulations begin with $M_t = 1$, which quickly approaches 0 over time. As the system evolves, the simulations decide and bifurcate in their magnetization paths ($M_t \approx 1$ and $M_t \approx -1$), resulting in differentiation of population cell types (which correspond to $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$). Figure 8.1 (b) shows the time evolution of the temperature trajectories of the two simulations. Here, the trajectories begin at $T = 2.8$, and with sufficient time, the temperature drops below the critical temperature and eventually reaches an equilibrium sub-critical temperature. Combined, we see that as the temperature reaches the Ising critical temperature (T_c), the tissue magnetizations diverge with an equal chance of the system choosing one of the two cell types.

8.3 Unimodal to Bimodal Transition in Cell States

Observing tissue-level statistics provides additional insight into collective behaviors in pluripotent cells transitioning to two possible cell types. One hundred independent and identical tissue simulations of Figure 8.1 were carried out and instantaneous

state distributions $\mathbf{G}(t)$ and the mean gene state was collected at each time step. At the beginning of the simulation, $t = 1$, the system was initialized with a high temperature point ($T = 2.8$) and with the CK fixed to $\mathbf{x}_c = \{1\}$ as in Figure 8.1. Trivially, all cell states are in $\mathbf{g}_2^{(1)}$, and the average gene state, which describes the tissue-level distribution, is unimodal (Figure 8.2 (a)). At time $t = 1 \times 10^5$ where the system reaches $M_t \approx 0$, two different cell types are probed, resulting in a mixture of $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ cell types, while the average state at the colony level remains unimodal, centered between $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ (Figure 8.2 (b)). With time, the temperature self-tunes and reaches an equilibrium point below the critical temperature (T_c). At time $t = 6 \times 10^5$, approximately half of the tissues form a homogeneous cell type of $\mathbf{g}_1^{(1)}$, and the other half form a cell type $\mathbf{g}_2^{(1)}$. The tissue-level states reach a split bimodal distribution (Figure 8.2 (c)). The system describes a full transition from a population of pluripotent tissue to two differentiated cell types. At the tissue level, this *unimodal-bimodal* transition at the critical junction of the phase transition occurs in several areas from mouse embryogenesis [20] to the development of the cancer cell line [21].

As an application, *self-tuned differentiation* can replicate Okamoto *et al.*'s experimental work on collective differentiation of mouse embryonic stem cells (mESCs) under strict conditions [20]. Okamoto *et al.* have observed the gene expression levels of key transcription factors in mESC, Nanog, and Oct4, in the early stage of differentiation. According to the immunofluorescence markers of Venus and mKate2, which

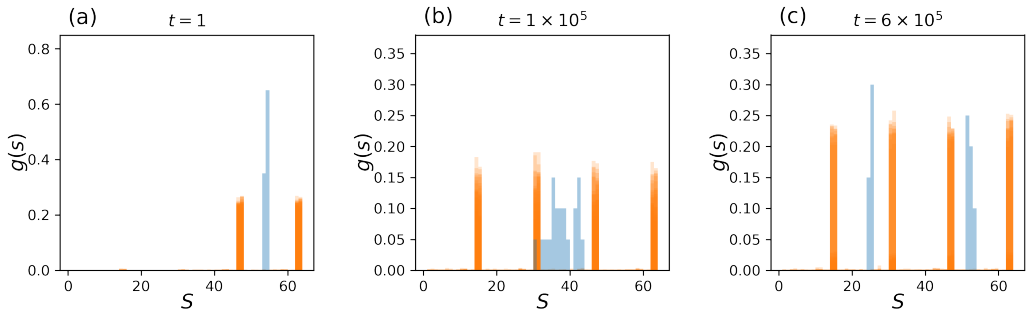


Figure 8.2: One hundred independent tissue simulations of a 32×32 Ising Hamiltonian (Equation 5.1) with the temperature-magnetization feedback mechanism (Equation 8.2) are shown: (a) At time $t = 1$, where the initial temperature is high ($T = 2.8$), all cells are of cell type $\mathbf{g}_2^{(1)}$. (b) At time $t = 1 \times 10^5$, where the tissues are $M_t \approx 0$, there is a mixture of $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$ cell types from Figure 4.1 (left), and the average cell state at the colony-level remains unimodal, centered between the distributions $g_1^{(1)}(s)$ and $g_2^{(1)}(s)$. (c) At time $t = 6 \times 10^5$, the tissues decide on the fate of the cells with a drop in temperature to critical and subcritical points, and hence, $M_t \approx -1$ or $M_t \approx 1$. This results in a split, bimodal distributions of gene states at the cellular and colony-level.

report Nanog and Oct4 gene expressions, respectively, colonies of mESCs exposed to leukemia inhibitory factor (LIF) demonstrated a high intensity of fluorescence, thus exhibiting single-state behavior in Nanog and Oct4. Here, LIF acts to enhance Nanog heterogeneity, in other words, to maintain the pluripotent state in the stem cell population. However, in the absence of LIF, high and low levels of Venus and mKate2 fluorescence were observed in cells, and cells are free to transition from pluripotent to differentiated state. What was observed was a multilevel *unimodal-bimodal transition* in the heterogeneity of gene expression levels at the single-cell level, followed by a delayed transition at the colony level.

In this work, Okamoto *et al.* investigated the role of *noise* in the multilevel transition from a unimodal to a bimodal distribution of mESC gene expression and classified two main sources of behaviors: stochastic intrinsic fluctuation of cell state in a cell (intrinsic noise) and deterministic extrinsic regulation with a network of neighboring cells (cell-cell cooperativity) [20]. Of the two sources, Okamoto *et al.* concluded that *dynamic change in cell-cell cooperativity* (in the case of the experiment, the collective behavioral effect of LIF on the gene expressions of mESCs) is necessary to observe this transition.

The cellular and colony-level unimodal-bimodal transition in the mESC distributions of *pluripotent population* [25, 38, 39] is characterized by the multilayer Ising Hamiltonian (Equation 5.1) with the temperature-magnetization feedback mechanism

(Equation 8.2) when pluripotent cells and differentiated cells are assumed to be *complementary* in gene states (i.e., they are induced by a fixed CK of $r = 1$). Then, pluripotent and differentiated cell types form $\mathbf{g}_1^{(1)}$ and $\mathbf{g}_2^{(1)}$, which exhibit multilevel unimodal-bimodal transitions of cell states with the self-tuning mechanism, as seen in Figure 8.2.

CHAPTER NINE

DISCUSSION AND CONCLUSION

In this work, we began by establishing the background knowledge necessary to develop the intercellular rules that govern the coupled BNs. Specifically, we have provided a surface review of the Ising model and random Boolean networks, and we have defined cell types as long-term steady-state behaviors of BNs. We introduced an intervention strategy known as the stochastic control kernel. Finally, we established the model paradigm for coupled BNs, where the interactions of the cells are governed by the multilayer Ising Hamiltonian, which captures three different effects of paracrine signaling, isolated independent BN dynamics, and autocrine signaling and/or external interventions. We have explored different modes of cell type detection method, which, depending on the model Hamiltonian, prove to be numerically accurate (and efficient) in practice. We showed through model simulations of coupled BNs with multilayer Ising Hamiltonian that (1) cell populations can undergo characteristic second-order phase transitions in the composition of the cell types under different levels of cooperativity and intrinsic noise. Furthermore, (2) in the presence of external stimuli, cells demonstrate phenotypic memory. These characteristic behaviors of coupled RBNs show that disorder-order, second-order phase transition in condensed matter physics is equivalent to the symmetry-breaking events involved in pluripotent-differentiated phase transition in the early stages of stem cell differ-

entiation. Through numerical simulations, we showed that tissues of coupled BNs undergo a series of symmetry-breaking events that alter the dynamics of the original cell, resulting in spontaneous differentiation to various cell types with critical control parameters of values T and h_0. Finally, we have implemented a negative feedback response between cell type composition and intrinsic noise in the population, which *self-tunes* a population of pluripotent cells to differentiated cell types and captures the qualitative unimodal-bimodal transition in cell state distributions, as observed experimentally by Okamoto *et al.* under strict conditions.

The coupled BN model with multilayer Ising Hamiltonian and the methods of simulation and detection in this work have many avenues for improvement and exploration. For instance, while each signaling term in Equation 5.1 was discussed in detail and numerically simulated for selected parameter values, a model experiment with full ranges of effects of h_0, and h was not assessed. Computing the full spectrum of parameter values would provide a complete characterization of spontaneous symmetry-breaking events in cell type differentiation.

Additionally, we note that the analysis in spontaneous phenotypic differentiation (Sections 7.1.3 and 7.1.3) depends greatly on the use of the model construction of a BN. The 6-gene BN exemplified throughout this work (Figure 3.1) is a conveniently critical network that exhibits "near symmetric" steady-state distributions in state space due to equal attractor sizes. We surmise that with more complex and sizable

gene regulatory networks, the cell type detection methods discussed may prove inadequate, since the number of different types of cells that each BN can generate strongly depends on the complexity of the cell-cell interaction and the structure of the original BN. Consequently, it may prove more challenging to observe spontaneous differentiation in ordered and chaotic networks using cell type detection methods such as the linear optimization approach, as these methods are sensitive to the sizes of attractors.

In the self-tuning model of the multilayer Ising Hamiltonian, we have illustrated a cellular and population-level transition in cell states or differentiation, and Okamto *et al.*'s experimental served as a model application under strict conditions. It is worth pointing out that the general model assumption in Okamoto *et al.*'s work is that cells transition between two attractor states (pluripotent and differentiated) of a double-well potential landscape. The potential barrier (whether it is skewed in one direction or symmetric) is lowered in the absence of the stimuli, LIF+. This is a contrasting interpretation of the life cycle of a cell depicted in Waddington's epigenetic landscape. The multilayer Ising Hamiltonian shows a differentiation from one state (pluripotent) to two distinct cell types (differentiated) that more closely reflects Waddington's epigenetic landscape and Huang *et al.*'s multipotent stem cell interpretation as a "balanced, undecided state" of multiple gene expression patterns [17]. This difference in interpretation of the potential landscape results in a less than one-to-one match with the model experiment.

Plans are in development to improve the model Hamiltonian and associated method tools and address the interpretive issues mentioned in the aforementioned article on experimental results in the future. We believe that further studies in the critical dynamics, phase transitions, and symmetry breaking of coupled BNs will provide an improved model closer to the biological realism of a stem cell cycle [129].

BIBLIOGRAPHY

[1] Stuart A Kauffman et al. *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA, 1993.

[2] Julianne D Halley and David A Winkler. Consistent concepts of self-organization and self-assembly. *Complexity*, 14(2):10–17, 2008.

[3] Herman Haken. Synergetics. *Physics Bulletin*, 28(9):412, 1977.

[4] Mikhail Prokopenko, Fabio Boschetti, and Alex J Ryan. An information-theoretic primer on complexity, self-organization, and emergence. *Complexity*, 15(1):11–28, 2009.

[5] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467, 1969.

[6] Enrique Balleza, Elena R Alvarez-Buylla, Alvaro Chaos, Stuart Kauffman, Ilya Shmulevich, and Maximino Aldana. Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS One*, 3(6):e2456, 2008.

[7] Matti Nykter, Nathan D Price, Maximino Aldana, Stephen A Ramsey, Stuart A Kauffman, Leroy E Hood, Olli Yli-Harja, and Ilya Shmulevich. Gene expression dynamics in the macrophage exhibit criticality. *Proceedings of the National Academy of Sciences*, 105(6):1897–1900, 2008.

[8] Matti Nykter, Nathan D Price, Antti Larjo, Tommi Aho, Stuart A Kauffman, Olli Yli-Harja, and Ilya Shmulevich. Critical networks exhibit maximal information diversity in structure-dynamics relationships. *Physical review letters*, 100(5):058702, 2008.

[9] Christian Torres-Sosa, Sui Huang, and Maximino Aldana. Criticality is an emergent property of genetic networks that exhibit evolvability. *PLOS Computational Biology*, 8(9):1–18, 09 2012.

[10] Michael D Stern. Emergence of homeostasis and "noise imprinting" in an evolution model. *Proceedings of the National Academy of Sciences*, 96(19):10746–10751, 1999.

[11] Ney Lemke, CM Jose'e, and Bardo EJ Bodmann. A numerical investigation of adaptation in populations of random boolean networks. *Physica A: Statistical Mechanics and its Applications*, 301(1-4):589–600, 2001.

[12] Ilya Shmulevich and Stuart A Kauffman. Activities and sensitivities in boolean network models. *Physical review letters*, 93(4):048701, 2004.

[13] Ilya Shmulevich, Stuart A Kauffman, and Maximino Aldana. Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proceedings of the National Academy of Sciences*, 102(38):13439–13444, 2005.

[14] Marc Weber and Javier Buceta. The cellular ising model: a framework for phase transitions in multicellular environments. *Journal of The Royal Society Interface*, 13(119):20151092, 2016.

[15] Thomas H Hraha, Matthew J Westacott, Marina Pozzoli, Aleena M Notary, P Mason McClatchey, and Richard KP Benninger. Phase transitions in the multi-cellular regulatory behavior of pancreatic islet excitability. *PLoS computational biology*, 10(9):e1003819, 2014.

[16] Paul CW Davies, Lloyd Demetrius, and Jack A Tuszynski. Cancer as a dynamical phase transition. *Theoretical Biology and Medical Modelling*, 8(1):1–16, 2011.

[17] Sui Huang. Reprogramming cell fates: reconciling rarity with robustness. *Bioessays*, 31(5):546–560, 2009.

[18] Nicola K Wilson, David G Kent, Florian Buettner, Mona Shehata, Iain C Macaulay, Fernando J Calero-Nieto, Manuel Sánchez Castillo, Caroline A Oedekoven, Evangelia Diamanti, Reiner Schulte, et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell stem cell*, 16(6):712–724, 2015.

[19] Lars Velten, Simon F Haas, Simon Raffel, Sandra Blaszkiewicz, Saiful Islam, Bianca P Hennig, Christoph Hirche, Christoph Lutz, Eike C Buss, Daniel Nowak, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nature cell biology*, 19(4):271–281, 2017.

[20] Kazuko Okamoto, Arno Germond, Hideaki Fujita, Chikara Furusawa, Yasushi Okada, and Tomonobu M Watanabe. Single cell analysis reveals a biophysical aspect of collective cell-state transition in embryonic stem cell differentiation. *Scientific reports*, 8(1):1–13, 2018.

[21] Masa Tsuchiya, Alessandro Giuliani, Midori Hashimoto, Jekaterina Erenpreisa, and Kenichi Yoshikawa. Emergent self-organized criticality in gene expression dynamics: Temporal development of global phase transition revealed in a cancer cell line. *PLoS One*, 10(6):e0128565, 2015.

[22] Ben D MacArthur and Ihor R Lemischka. Statistical mechanics of pluripotency. *Cell*, 154(3):484–489, 2013.

[23] Elisabet Pujadas and Andrew P Feinberg. Regulated noise in the epigenetic landscape of development and disease. *Cell*, 148(6):1123–1131, 2012.

[24] Jordi Garcia-Ojalvo and Alfonso Martinez Arias. Towards a statistical mechanics of cell fate decisions. *Current opinion in genetics & development*, 22(6):619–626, 2012.

[25] Graziano Martello and Austin Smith. The nature of embryonic stem cells. *Annual review of cell and developmental biology*, 30:647–675, 2014.

[26] Raymond Schofield. The stem cell system. *Biomedicine & pharmacotherapy= Biomedecine & pharmacotherapie*, 37(8):375–380, 1983.

[27] Laralynne Przybyla and Joel Voldman. Probing embryonic stem cell autocrine and paracrine signaling using microfluidics. *Annual review of analytical chemistry (Palo Alto, Calif.)*, 5:293, 2012.

[28] Atta Behfar, Leonid V Zingman, Denice M Hodgson, Jean-Michel Rauzier, Garvan C Kane, Andre Terzic, and Michel Pucéat. Stem cell differentiation requires a paracrine pathway in the heart. *The FASEB Journal*, 16(12):1558–1566, 2002.

[29] Xiaoting Liang, Yue Ding, Yuelin Zhang, Hung-Fat Tse, and Qizhou Lian. Paracrine mechanisms of mesenchymal stem cell-based therapy: current status and perspectives. *Cell transplantation*, 23(9):1045–1059, 2014.

[30] Anna Janowska-Wieczorek, Marcin Majka, Janina Ratajczak, and Mariusz Z Ratajczak. Autocrine/paracrine mechanisms in human hematopoiesis. *Stem cells*, 19(2):99–107, 2001.

[31] Massimiliano Gnecchi, Zhiping Zhang, Aiguo Ni, and Victor J Dzau. Paracrine mechanisms in adult stem cell signaling and therapy. *Circulation research*, 103(11):1204–1219, 2008.

[32] Celia Sid-Otmane, Louis P Perrault, and Hung Q Ly. Mesenchymal stem cell mediates cardiac repair through autocrine, paracrine and endocrine axes. *Journal of Translational Medicine*, 18(1):1–9, 2020.

[33] James E Till, Ernest A McCulloch, and Louis Siminovitch. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proceedings of the National Academy of Sciences*, 51(1):29–36, 1964.

[34] Tibor Kalmar, Chea Lim, Penelope Hayward, Silvia Munoz-Descalzo, Jennifer Nichols, Jordi Garcia-Ojalvo, and Alfonso Martinez Arias. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS biology*, 7(7):e1000149, 2009.

[35] Vahid Shahrezaei and Peter S Swain. The stochastic nature of biochemical networks. *Current opinion in biotechnology*, 19(4):369–374, 2008.

[36] Marc Sturrock, Andreas Hellander, Anastasios Matzavinos, and Mark AJ Chaplain. Spatial stochastic modelling of the hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. *Journal of The Royal Society Interface*, 10(80):20120988, 2013.

[37] Chris Kang, Boris Aguilar, and Ilya Shmulevich. Emergence of diversity in homogeneous coupled boolean networks. *Physical Review E*, 97(5):052415, 2018.

[38] Jose Silva and Austin Smith. Capturing pluripotency. *Cell*, 132(4):532–536, 2008.

[39] Jamie A Hackett and M Azim Surani. Regulatory principles of pluripotency: from the ground state up. *Cell stem cell*, 15(4):416–430, 2014.

[40] Chris Kang, Madelynn McElroy, and Nikolaos K Voulgarakis. Emergent criticality in coupled boolean networks. *Entropy*, 25(2):235, 2023.

[41] François Graner and James A Glazier. Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical review letters*, 69(13):2013, 1992.

[42] Nicholas J Savill and Paulien Hogeweg. Modelling morphogenesis: from single cells to crawling slugs. *Journal of theoretical biology*, 184(3):229–235, 1997.

[43] Athanasius FM Marée, Verônica A Grieneisen, and Paulien Hogeweg. The cellular potts model and biophysical properties of cells, tissues and morphogenesis. In *Single-cell-based models in biology and medicine*, pages 107–136. Springer, 2007.

[44] Anja Voss-Böhme. Multi-scale modeling in morphogenesis: a critical analysis of the cellular potts model. *PLoS One*, 7(9):e42852, 2012.

[45] Noriyuki Bob Ouchi, James A Glazier, Jean-Paul Rieu, Arpita Upadhyaya, and Yasuji Sawada. Improving the realism of the cellular potts model in simulations of biological cells. *Physica A: Statistical Mechanics and its Applications*, 329(3-4):451–458, 2003.

[46] Elisabeth G Rens and Leah Edelstein-Keshet. From energy to cellular forces in the cellular potts model: An algorithmic approach. *PLoS computational biology*, 15(12):e1007459, 2019.

[47] Nara Guisoni, Karina I Mazzitello, and Luis Diambra. Modeling active cell movement with the potts model. *Frontiers in Physics*, 6:61, 2018.

[48] Stephen Turner and Jonathan A Sherratt. Intercellular adhesion and cancer invasion: a discrete simulation using the extended potts model. *Journal of theoretical biology*, 216(1):85–100, 2002.

[49] Hugo Duminil-Copin. Lectures on the ising and potts models on the hypercubic lattice. In *PIMS-CRM Summer School in Probability*, pages 35–161. Springer, 2017.

[50] Franz Utermohlen. Mean field theory solution of the ising model, 2018.

[51] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.

[52] Hendrik A Kramers and Gregory H Wannier. Statistics of the two-dimensional ferromagnet. part i. *Physical Review*, 60(3):252, 1941.

[53] Pierre Weiss. L'hypothèse du champ moléculaire et la propriété ferromagnétique. *J. Phys. Theor. Appl.*, 6(1):661–690, 1907.

[54] Gregg Jaeger. The ehrenfest classification of phase transitions: introduction and evolution. *Archive for history of exact sciences*, 53(1):51–81, 1998.

[55] Paul Ehrenfest. *Phasenumwandlungen im ueblichen und erweiterten Sinn, classifiziert nach den entsprechenden Singularitaeten des thermodynamischen Potentiales*. NV Noord-Hollandsche Uitgevers Maatschappij, 1933.

[56] Karl Fabian, Valera P Shcherbakov, and Suzanne A McEnroe. Measuring the curie temperature. *Geochemistry, Geophysics, Geosystems*, 14(4):947–961, 2013.

[57] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review*, 65(3-4):117, 1944.

[58] Manfred Schroeder. *Fractals, chaos, power laws: Minutes from an infinite paradise*. Courier Corporation, 2009.

[59] David Pollard. Hammersley-clifford theorem for markov random fields. *Yale University, New Haven*, 2004.

[60] Jarek Duda. Nearly accurate solutions for ising-like models using maximal entropy random walk. *arXiv preprint arXiv:1912.13300*, 2019.

[61] Richard Fitzpatrick. Computational physics. *Lecture notes, University of Texas at Austin*, 2006.

[62] Stuart Kauffman. The large scale structure and dynamics of gene control circuits: an ensemble approach. *Journal of Theoretical Biology*, 44(1):167–190, 1974.

[63] Leon Glass. Combinatorial and topological methods in nonlinear chemical kinetics. *The Journal of chemical physics*, 63(4):1325–1335, 1975.

[64] Sanjeev Kumar and Peter J Bentley. An introduction to computational development. *On Growth, Form and Computers*, 1:1–44, 2003.

[65] Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature reviews Molecular cell biology*, 9(10):770–780, 2008.

[66] Stuart Kauffman, Carsten Peterson, Björn Samuelsson, and Carl Troein. Random boolean network models and the yeast transcriptional network. *Proceedings of the National Academy of Sciences*, 100(25):14796–14799, 2003.

[67] Ronaldo Fumio Hashimoto, Henrique Stagni, and Carlos Henrique Aguena Higa. Budding yeast cell cycle modeled by context-sensitive probabilistic boolean network. In *2009 IEEE International Workshop on Genomic Signal Processing and Statistics*, pages 1–4. IEEE, 2009.

[68] Maria I Davidich and Stefan Bornholdt. Boolean network model predicts cell cycle sequence of fission yeast. *PloS one*, 3(2):e1672, 2008.

[69] Réka Albert and Hans G Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster. *Journal of theoretical biology*, 223(1):1–18, 2003.

[70] Julio Saez-Rodriguez, Luca Simeoni, Jonathan A Lindquist, Rebecca Hemenway, Ursula Bommhardt, Boerge Arndt, Utz-Uwe Haus, Robert Weismantel, Ernst D Gilles, Steffen Klamt, et al. A logical model provides insights into t cell receptor signaling. *PLoS computational biology*, 3(8):e163, 2007.

[71] Sui Huang, Ingemar Ernberg, and Stuart Kauffman. Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. In *Seminars in cell & developmental biology*, volume 20, pages 869–876. Elsevier, 2009.

[72] Sui Huang and Stuart Kauffman. How to escape the cancer attractor: rationale and limitations of multi-target drugs. In *Seminars in cancer biology*, volume 23, pages 270–278. Elsevier, 2013.

[73] Herman F Fumia and Marcelo L Martins. Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes. *PloS one*, 8(7):e69008, 2013.

[74] Panuwat Trairatphisan, Monique Wiesinger, Christelle Bahlawane, Serge Haan, and Thomas Sauter. A probabilistic boolean network approach for the analysis of cancer-specific signalling: A case study of deregulated pdgf signalling in gist. *PloS one*, 11(5):e0156223, 2016.

[75] David C Earnest and Ian F Wilkinson. An agent based model of the evolution of supplier networks. *Computational and Mathematical Organization Theory*, 24(1):112–144, 2018.

[76] KE Kurten. Correspondence between neural threshold networks and kauffman boolean cellular automata. *Journal of Physics A: Mathematical and General*, 21(11):L615, 1988.

[77] Julian D Schwab, Silke D Kühlwein, Nensi Ikonomi, Michael Kühl, and Hans A Kestler. Concepts in boolean network modeling: What do they all mean? *Computational and structural biotechnology journal*, 18:571–582, 2020.

[78] Assieh Saadatpour, István Albert, and Réka Albert. Attractor analysis of asynchronous boolean models of signal transduction networks. *Journal of theoretical biology*, 266(4):641–656, 2010.

[79] René Thomas. Regulatory networks seen as asynchronous automata: a logical description. *Journal of theoretical biology*, 153(1):1–23, 1991.

[80] Conrad Hal Waddington et al. Organisers and genes. *Organisers and genes.*, 1940.

[81] Bernard Derrida and Yves Pomeau. Random networks of automata: a simple annealed approximation. *EPL (Europhysics Letters)*, 1(2):45, 1986.

[82] Bartolo Luque and Ricard V Solé. Lyapunov exponents in random boolean networks. *Physica A: Statistical Mechanics and its Applications*, 284(1-4):33–45, 2000.

[83] P Rämö, J Kesseli, and O Yli-Harja. Perturbation avalanches and criticality in gene regulatory networks. *Journal of Theoretical Biology*, 242(1):164–170, 2006.

[84] Bryan C Daniels, Hyunju Kim, Douglas Moore, Siyu Zhou, Harrison B Smith, Bradley Karas, Stuart A Kauffman, and Sara I Walker. Criticality distinguishes the ensemble of biological regulatory networks. *Physical review letters*, 121(13):138102, 2018.

[85] Barbara Drossel. Random boolean networks. *Reviews of nonlinear dynamics and complexity*, pages 69–110, 2008.

[86] Carlos Gershenson. Introduction to random boolean networks. *arXiv preprint nlin/0408006*, 2004.

[87] Stefan Bornholdt and Stuart Kauffman. Ensembles, dynamics, and cell types: Revisiting the statistical mechanics perspective on cellular regulation. *Journal of theoretical biology*, 467:15–22, 2019.

[88] Ulrike Münzner, Tomoya Mori, Marcus Krantz, Edda Klipp, and Tatsuya Akutsu. Identification of periodic attractors in boolean networks using a priori information. *PLOS Computational Biology*, 18(1):e1009702, 2022.

[89] Hans Clevers, Susanne Rafelski, Michael Elowitz, Allon Klein, Jay Shendure, Cole Trapnell, Ed Lein, Emma Lundberg, Matthias Uhlen, Alfonso Martinez-Arias, et al. What is your conceptual definition of "cell type" in the context of a mature organism? *Cell Systems*, 4(3):255–259, 2017.

[90] Roberto Serra, Marco Villani, Alessia Barbieri, Stuart A Kauffman, and Annamaria Colacci. On the dynamics of random boolean networks subject to noise: attractors, ergodic sets and cell types. *Journal of theoretical biology*, 265(2):185–193, 2010.

[91] Sui Huang. Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery. *Journal of molecular medicine*, 77(6):469–480, 1999.

[92] Tariq Enver, Martin Pera, Carsten Peterson, and Peter W Andrews. Stem cell states, fates, and the rules of attraction. *Cell stem cell*, 4(5):387–397, 2009.

[93] Sui Huang, Gabriel Eichler, Yaneer Bar-Yam, and Donald E Ingber. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical review letters*, 94(12):128701, 2005.

[94] Sui Huang and Donald E Ingber. Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. *Experimental cell research*, 261(1):91–103, 2000.

[95] Rui Chang, Robert Shoemaker, and Wei Wang. Systematic search for recipes to generate induced pluripotent stem cells. *PLoS computational biology*, 7(12):e1002300, 2011.

[96] Joseph Xu Zhou, MDS Aliyu, Erik Aurell, and Sui Huang. Quasi-potential landscape in complex multi-stable systems. *Journal of the Royal Society Interface*, 9(77):3539–3553, 2012.

[97] Ilya Shmulevich, Edward R Dougherty, Seungchan Kim, and Wei Zhang. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.

[98] Ilya Shmulevich, Ilya Gluhovsky, Ronaldo F Hashimoto, Edward R Dougherty, and Wei Zhang. Steady-state analysis of genetic regulatory networks modelled by probabilistic boolean networks. *Comparative and functional genomics*, 4(6):601–608, 2003.

[99] Marcel Brun, Edward R Dougherty, and Ilya Shmulevich. Steady-state probabilities for attractors in probabilistic boolean networks. *Signal Processing*, 85(10):1993–2013, 2005.

[100] Ranadip Pal, Ivan Ivanov, Aniruddha Datta, Michael L Bittner, and Edward R Dougherty. Generating boolean networks with a prescribed attractor structure. *Bioinformatics*, 21(21):4021–4025, 2005.

[101] Wai-Ki Ching, Shuqin Zhang, Michael K Ng, and Tatsuya Akutsu. An approximation method for solving the steady-state probability distribution of probabilistic boolean networks. *Bioinformatics*, 23(12):1511–1518, 2007.

[102] Ilya Shmulevich and Edward R Dougherty. *Probabilistic Boolean networks: the modeling and control of gene regulatory networks*. SIAM, 2010.

[103] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[104] Jonathan M Raser and Erin K O'Shea. Control of stochasticity in eukaryotic gene expression. *science*, 304(5678):1811–1814, 2004.

[105] Ying Chen, Ido Golding, Satoshi Sawai, Ling Guo, and Edward C Cox. Population fitness and the regulation of escherichia coli genes by bacterial viruses. *PLoS biology*, 3(7):e229, 2005.

[106] Xue Lei, Wei Tian, Hongyuan Zhu, Tianqi Chen, and Ping Ao. Biological sources of intrinsic and extrinsic noise in ci expression of lysogenic phage lambda. *Scientific reports*, 5(1):1–12, 2015.

[107] Roberto Serra, Marco Villani, and Alessandro Semeria. Genetic network models and statistical properties of gene expression data in knock-out experiments. *Journal of theoretical biology*, 227(1):149–157, 2004.

[108] Roberto Serra, Marco Villani, Chiara Damiani, Alex Graudenzi, A Colacci, and SA Kauffman. Interacting random boolean networks. In *Proceedings of ECCS07: European Conference on Complex Systems*, pages 1–15. Citeseer, 2007.

[109] Roberto Serra, Marco Villani, Alex Graudenzi, and SA Kauffman. Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data. *Journal of theoretical biology*, 246(3):449–460, 2007.

[110] Daizhan Cheng and Hongsheng Qi. Controllability and observability of boolean control networks. *Automatica*, 45(7):1659–1667, 2009.

[111] Junil Kim, Sang-Min Park, and Kwang-Hyun Cho. Discovery of a kernel for controlling biomolecular regulatory networks. *Scientific reports*, 3(1):1–9, 2013.

[112] Marco Villani, Roberto Serra, P Ingrami, and Stuart A Kauffman. Coupled random boolean network forming an artificial tissue. In *International Conference on Cellular Automata*, pages 548–556. Springer, 2006.

[113] Chiara Damiani, Roberto Serra, Marco Villani, SA Kauffman, and A Colacci. Cell–cell interaction and diversity of emergent behaviours. *IET systems biology*, 5(2):137–144, 2011.

[114] Nicholas S Flann, Hamid Mohamadlou, and Gregory J Podgorski. Kolmogorov complexity of epithelial pattern formation: the role of regulatory network configuration. *Biosystems*, 112(2):131–138, 2013.

[115] Hyobin Kim, Omar K Pineda, and Carlos Gershenson. A multilayer structure facilitates the production of antifragile systems in boolean network models. *Complexity*, 2019, 2019.

[116] Carlos Gershenson. Guiding the self-organization of random boolean networks. *Theory in Biosciences*, 131(3):181–191, 2012.

[117] Hyobin Kim and Hiroki Sayama. How criticality of gene regulatory networks affects the resulting morphogenesis under genetic perturbations. *Artificial Life*, 24(02):85–105, 2018.

[118] John E Dennis, Jr and Jorge J Moré. Quasi-newton methods, motivation and theory. *SIAM review*, 19(1):46–89, 1977.

[119] John E Dennis Jr and Robert B Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM, 1996.

[120] Paul T Boggs and Jon W Tolle. Sequential quadratic programming. *Acta numerica*, 4:1–51, 1995.

[121] Xinhua Zhuang, Yan Huang, Kannappan Palaniappan, and Yunxin Zhao. Gaussian mixture density modeling, decomposition, and applications. *IEEE Transactions on Image Processing*, 5(9):1293–1302, 1996.

[122] Najme Khorasani and Mehdi Sadeghi. A computational model of stem cells' decision-making mechanism to maintain tissue homeostasis and organization in the presence of stochasticity. *Scientific Reports*, 12(1):1–17, 2022.

[123] Eszter Posfai, Oliver H Tam, and Janet Rossant. Mechanisms of pluripotency in vivo and in vitro. *Current topics in developmental biology*, 107:1–37, 2014.

[124] Sarah JL Graham and Magdalena Zernicka-Goetz. The acquisition of cell fate in mouse development: how do cells first become heterogeneous? *Current Topics in Developmental Biology*, 117:671–695, 2016.

[125] Hadiseh Safdari, Ata Kalirad, Cristian Picioreanu, Rouzbeh Tusserkani, Bahram Goliaei, and Mehdi Sadeghi. Noise-driven cell differentiation and the emergence of spatiotemporal patterns. *Plos one*, 15(4):e0232060, 2020.

[126] Irving L Weissman, David J Anderson, and Fred Gage. Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. *Annual review of cell and developmental biology*, 17:387–403, 2001.

[127] Jun Seita and Irving L Weissman. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(6):640–653, 2010.

[128] Keisuke Ito and Toshio Suda. Metabolic requirements for the maintenance of self-renewing stem cells. *Nature reviews Molecular cell biology*, 15(4):243–256, 2014.

[129] Madelynn McElroy and Nikolaos K. Voulgarakis, in preparation, 2023.

[130] Takeo Matsubara. A new approach to quantum-statistical mechanics. *Progress of theoretical physics*, 14(4):351–378, 1955.

[131] Indrani Bose and Sayantari Ghosh. Bifurcation and criticality. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(4):043403, 2019.

[132] Indika Rajapakse and Steve Smale. The pitchfork bifurcation. *arXiv preprint arXiv:1609.05996*, 2016.

[133] Indika Rajapakse and Stephen Smale. Emergence of function from coordinated cells in a tissue. *Proceedings of the National Academy of Sciences*, 114(7):1462–1467, 2017.

[134] Vladimir A Dobrushkin. *Applied Differential Equations: The Primary Course*. CRC Press, 2018.

[135] Hong Qian, Ping Ao, Yuhai Tu, and Jin Wang. A framework towards understanding mesoscopic phenomena: Emergent unpredictability, symmetry breaking and dynamics across scales. *Chemical Physics Letters*, 665:153–161, 2016.

[136] Peter D Olmsted. Lectures on landau theory of phase transitions. *University of Leeds, Department of Physics and Astronomy*, 2000.

[137] Paul Fendley. Modern statistical mechanics. *The University of Virginia*, 2014.

[138] Mario Reis. *Fundamentals of magnetism*. Elsevier, 2013.

[139] Evgenii Mikhailovich Lifshitz and Lev Petrovich Pitaevskii. *Statistical physics: theory of the condensed state*, volume 9. Elsevier, 2013.

APPENDIX A

METROPOLIS ALGORITHM

Here, we provide a basic schematics for the Metropolis Algorithm.

1. Choose an arbitrary site i, whose spin is s_i.

2. Flip the spin to get the candidate proposal: s_i'.

3. Determine the change in the new local energy ΔE.

4. Accept/Reject proposal:

   (a) If $\Delta E < 0$, accept the proposal.

   (b) Otherwise, accept the proposal according to the Boltzmann distribution:

      i. $p \sim e^{\frac{-\Delta E}{T}}$

5. Repeat steps $1 \sim 4$ for each Monte Carlo step.

APPENDIX B

CANONICAL ENSEMBLE, BOLTZMANN DISTRIBUTION, AND THE PARTITION FUNCTION

In statistical mechanics, there are three major thermodynamic ensembles describing particle microstates (s): the canonical ensemble, microcanonical ensemble, and grand canonical ensemble. Each ensemble has fixed macroscopic conditions in combinations of the number of particles (N), volume (V), energy (E), chemical potential (μ) or temperature (T).

In the canonical ensemble, the macroscopic conditions N, V, and T are kept fixed. Thus, the thermodynamic potential of the system, F, (Helmholtz free energy) is given by $F(T, V, N) = E - TS$, where E is the internal energy of the system, and S is the Gibbs entropy. In the canonical ensemble, the probability (P_s) to find the system in a microstate (s) with energy ($E(s)$) can be expressed in terms of the Boltzmann distribution:

$$P_s = \frac{e^{-\frac{E(s)}{k_B T}}}{\sum_s e^{-\frac{E(s)}{k_B T}}}. \tag{B.1}$$

Then, statistical ensemble P is the density matrix (we refer to [50] for the details of

this step):

$$P = e^{\frac{(F-E)}{k_B T}}. \tag{B.2}$$

The Helmholtz free energy is determined by the probability normalization condition that the density matrix has a trace of one, $P = 1$:

$$\Rightarrow P = \frac{e^{\frac{(F)}{k_B T}}}{e^{\frac{(E)}{k_B T}}} \tag{B.3}$$

$$\Rightarrow \frac{1}{e^{\frac{(F)}{k_B T}}} P = \frac{1}{e^{\frac{(E)}{k_B T}}} \tag{B.4}$$

$$\Rightarrow \exp\left(\frac{-F}{k_B T}\right) = \mathrm{Tr}\left(\exp\left(\frac{-E}{k_B T}\right)\right). \tag{B.5}$$

Here, the right-hand side of equation (B.5) is the canonical partition function Z:

$$Z = \mathrm{Tr}\left(\exp\left(\frac{-E}{k_B T}\right)\right). \tag{B.6}$$

The outcome of Equation B.6 is that the Helmholtz free energy equation can be expressed in terms of the partition function,

$$F = -k_B T \ln Z. \tag{B.7}$$

In the case of the Ising model, the thermodynamics is described by the *canonical*

ensemble. Then, from Equation B.5, the average spin of the states, $\langle s_i \rangle$, can be written as a function of the partition function

$$\langle s_i \rangle = \frac{\mathrm{Tr}\left(s_i \cdot e^{\frac{E}{k_B T}}\right)}{Z}, \tag{B.8}$$

where E is now the Hamiltonian of the Ising model (Equation 2.3). This is known as the Matsubara formalism in thermal quantum field theory (or statistical field theory) [130].

APPENDIX C

LANDAU FREE ENERGY

In Section 2.4, we saw the mean field treatment of the zero-field two-dimensional Ising model shows the second-order phase transition at the critical temperature. It turns out the dynamical behavior of $T \to T_c$ qualitatively equates to *supercritical pitchfork bifurcation* [131, 50].

We wish to establish the similarities and differences between bifurcation and phase transition, as they become a recurring theme in this section. It is often useful to understand the stability of a dynamical system in terms of its free energy or potential function $V(x)$. The behaviors of dynamical systems are usually characterized by their equilibrium states (fixed points), which correspond to the extrema of $V(x)$. When equilibrium states of a dynamical system can be created or destroyed (or such that, when the stabilities of the system change) by a "smooth" change in the system parameter, bifurcation points arise [131]. In dynamical systems, bifurcation occurs in a *finite*-dimensional state space, and the asymptotic behavior of time, $t \to \infty$, is required for the states of a system to reach its equilibrium (or attractor).

Rajapakse and Smale provide a fascinating example of an observable bifurcation in morphogenesis in their recent work [132, 133]. They postulate that the emergence of a new cell type from its progenitor, that is, through natural means of differentiation, reprogramming, or cancer, is the result of *pitchfork bifurcation*. Rajapakse and Smale

justify that symmetric and asymmetric divisions (either in the same lineage or in different cell types) to daughter cells are reflected by pitchfork bifurcation, as the cell requires a "departure" from a fixed-point stable equilibrium to undergo cell division of new (multiple) fixed points, the defining characteristic of pitchfork bifurcation [132, 133].

On the other hand, the macroscopic state of a thermodynamic system corresponds to the global minimum (ground state) of $V(x)$. When the ground state shifts from one state to another due to a change in the macroscopic variable (i.e., temperature, pressure, and magnetic field), there is a phase transition [134]. In *thermodynamic systems*, the limiting behaviors with respect to time (t) *and* and the size of the system (N) are assumed: $t \to \infty$ and $N \to \infty$ [131, 135].

Finally, we introduce *the Landau Theory of Phase Transition.* The Landau Theory of Phase Transition is a tool to describe a phase transition as a continuous free energy form with a Taylor expansion about an order parameter [136]. It is a *phenomenological*, expression of free energy, F, where we are concerned with the "lowering of the symmetry" in F [137]. It does not care for the microscopic behavior of a system, such as the dipole moment of a spin. This requires two assumptions. That (1), F is analytic (differentiable everywhere), and (2) it obeys the symmetry of the

Hamiltonian[1] [138].

We expand these assumptions for the Landau theory of phase transition to the following:

1. For the macroscopic variable T, the order parameter m is assumed to be small at points of the critical phenomenon (T_c).

2. The free energy of a system can be approximated by the low-order expansion in the powers of the order parameter m:

$$F(m) \approx F_0 + F_1(m) + F_2 m^2 + F_3 m^3 + ... \tag{C.1}$$

3. The coefficients F_i depend on the critical junctures of the macroscopic variables T and h (if present), where h is the strength of the external field. This means that we consider $F_i = f\left((T - T_c), (h - h_c)\right)$, and h_c is the critical point of the external field.

In the following sections, we discuss two simple cases of Landau free energy: (1) the quadratic expansion and (2) the quadratic expansion with a linear term. In these analyses, we see that equilibrium solutions to Landau expansions correspond to

[1]The language for these two assumptions are borrowed from *Fundamentals of Magnetism* by Mario Reis [138].

bifurcation points.

C.0.1 Quadratic Expansion of Landau Free Energy

Under systems where the order parameter is bound by the parity symmetry, such as a ferromagnetic system, the parity inversion is invariant: $m \mapsto -m$. In this chapter, we assume parity symmetry in all of the systems considered. This allows a simplification of Equation C.1 to a Taylor expansion of even-powered terms of m's only. Without the presence of any external field ($h = 0$), the simplest example of such free energy F is the quadratic expansion:

$$F(m) = F_2 m^2 + F_4 m^4. \tag{C.2}$$

Our goal is to derive the *magnetization dynamics*,

$$\dot{m} = f(r, m), \tag{C.3}$$

from Equation C.2, which we can *nondimensionalize*. We wish to obtain the free energy form of the dynamical system that is a function of dimensionless parameters r and M, with respect to some $\tau = f(t)$:

$$\frac{dM}{d\tau} = f(r, M). \tag{C.4}$$

By Assumption 3, F_i' s depend on the critical junctures of the macroscopic variables: $(T - T_c)$ and $(h - h_c) = 0$. Then, let $F_2 = \frac{a}{2}(T - T_c)$ and $F_4 = \frac{1}{4}b$ in Equation C.2, for some constants $a, b > 0$. By substituting F_2 and F_4,

$$F(m) = \frac{a}{2}(T - T_c)m^2 + \frac{b}{4}m^4 \tag{C.5}$$

$$= -\frac{A}{2}\left(1 - \frac{T}{T_c}\right)m^2 + \frac{b}{4}m^4, \tag{C.6}$$

where $A = aT_c$. We take the first derivative with respect to m from Equation C.6 to find the magnetization dynamics.

$$\dot{m} = -r\frac{dF}{dm} = -r\left(-A\left(1 - \frac{T}{T_c}\right)m + bm^3\right) \tag{C.7}$$

$$= rA\left(1 - \frac{T}{T_c}\right)m - rbm^3. \tag{C.8}$$

We apply a change of variables to $\tau = (rA)t$ and let $M = \frac{m}{\sqrt{A/b}}$ and $r = (1 - T/T_c)$. Then r is the *scaled* measure of how far the system is from the critical temperature T_c. We note that nondimensionalization of the Landau expansion allows us to analyze systems only in terms of two parameters (M and r) without loss of information. Now, the magnetization dynamic (Equation C.8) can be nondimensionalized with

dimensionless parameters M, r, and τ:

$$\frac{dM}{d\tau} = rM - M^3. \tag{C.9}$$

Equation C.9 is a dimensionless vector field, and from it we can easily obtain *quadratic Landau free energy expansion* of Equation C.9 (the potential function):

$$F(M) = -\int f(M)\ dM = F_0 - \frac{1}{2}rM^2 + \frac{1}{4}M^4. \tag{C.10}$$

For simplicity, let us assume $F_0 = 0$.

Equilibrium Points and Bifurcation Diagram

A quick inspection of Equation C.9 tells us that the dynamical system is an example of *supercritical pitchfork bifurcation*. Equilibrium solutions can be easily solved:

$$0 = rM^* - (M^*)^3 \tag{C.11}$$

$$\Rightarrow 0 = M^*(r - (M^*)^2) \tag{C.12}$$

$$\Rightarrow 0 = M^*(\sqrt{r} - M^*)(\sqrt{r} + M^*). \tag{C.13}$$

$M^* = 0$ for $\forall r \in \mathbb{R}$ and $M^* = \pm\sqrt{r}$ for $r > 0$. We are interested in the stabilities at the equilibrium points

$$f'(M^*) = r - 3(M^*)^2. \tag{C.14}$$

At $M^* = 0$ and $r < 0$, $f'(M^*)$ is stable. At $M^* = 0$ and $r > 0$, $f'(M^*)$ is unstable. At $M^* = \pm\sqrt{r}$ and $r > 0$, $f'(M^*)$ is stable. We have a supercritical pitchfork bifurcation at $M = 0$, $r = 0$.

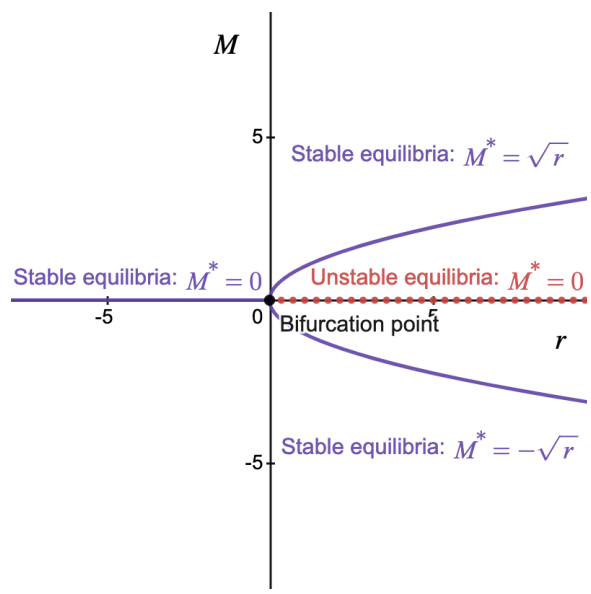


Figure C.1: Bifurcation diagram of the quadratic Landau free energy expansion: M vs. r. Supercritical pitchfork bifurcation occurs at $r = 0, M = 0$.

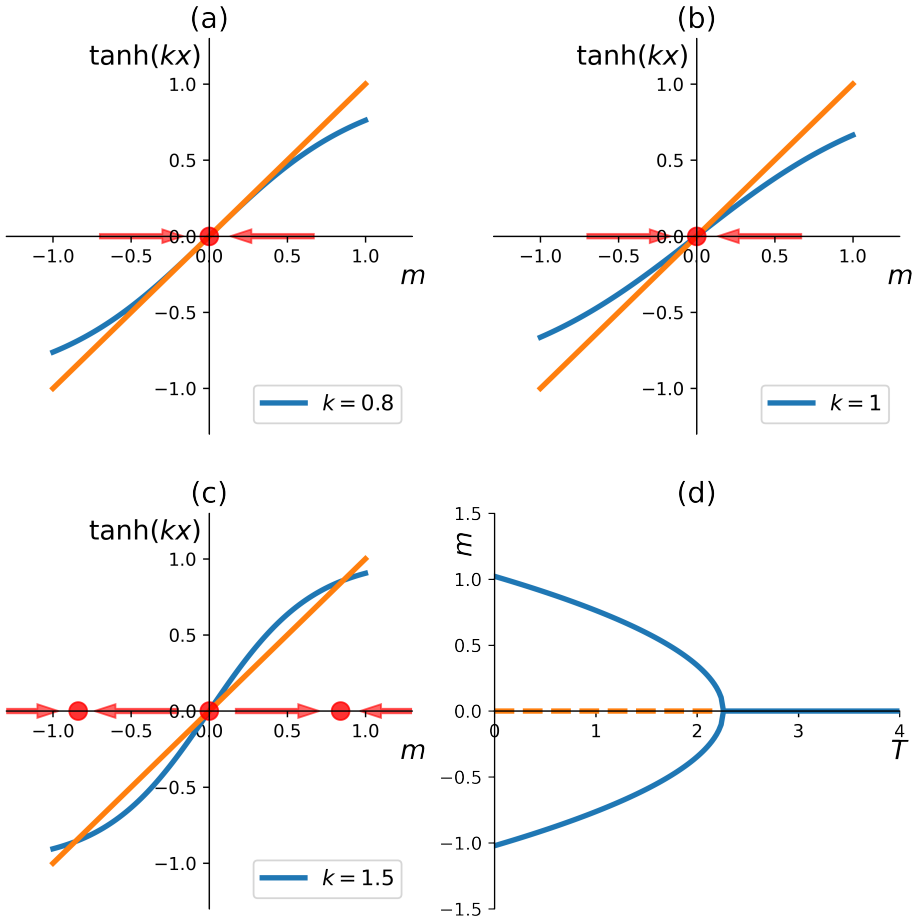


Figure C.2: Stability analysis for m from the self-consistency equation (Equation 2.38) near the critical temperature T_c. With $k = q/(k_B T)$, stabilities of m at varying value of $k = 0.8, 1.0, 1.5$ are shown in (a), (b), and (c). (d) shows that the second-order phase transition at T_c is qualitatively similar to supercritical pitchfork bifurcation in Figure C.1.

For the zero-field 2-D Ising model, the phase transition at T_c is qualitatively similar to the supercritical pitchfork bifurcation in Figure C.1. Recall the self-consistency

equation (Equation 2.38). Let $k = q/(k_B T)$. The stability analysis can be performed for $m = \tanh(km)$ near the critical temperature $T \to T_c$ (Figure C.2).

Ground States and Phase Transition

Ground states are the global minima of free energy. We plot Equation C.10 to find the ground states of the system. Figure C.3 plots the quadratic expansion with three different values of $r = -1, 0, 1$. The ground states are $M^* = 0$ for $\forall r \leq 0$ and $M^* = \pm\sqrt{r}$ for $r > 0$.

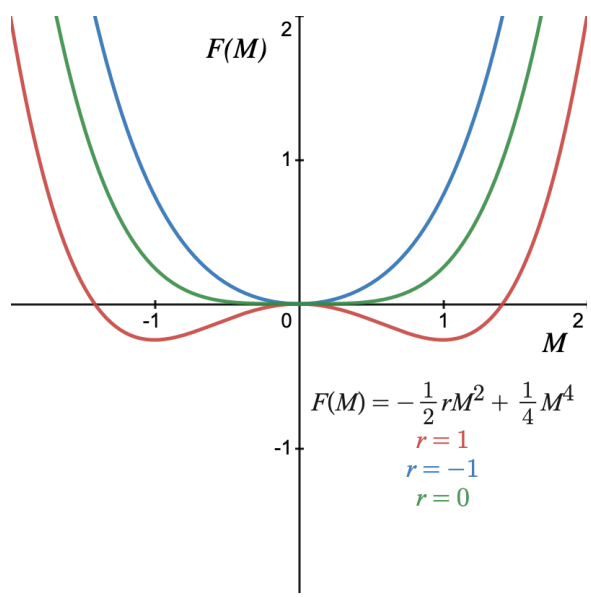


Figure C.3: Plot of the quadratic Landau free energy expansion with values of $r = -1, 0, 1$: $F(M)$ vs. M. Second-order phase transition occurs at $M = 0$, as with $r < 0$, there is a continuous change in ground state(s).

It is easy to see graphically in Figure C.3 that Equation C.10 exhibits a *second-order* phase transition. Recall that we characterize the second-order phase transition

by differentiability near the critical point, where there is a sudden shift in the ground state(s). When $r \to 0^+$, there is a singular ground state at $M^* = 0$. There is a continuous change in M until the critical point of $r = 0$, where the ground state is no longer at $M^* = 0$, but has shifted to two possible states of $M^* = \pm\sqrt{r}$. Interestingly, for quadratic expansion, the second-order phase transition is described by the bifurcation diagram itself (Figure C.1).

C.0.2 Quadratic Expansion of Landau Free Energy with a Linear Term

We can perturb a system with an external field h that couples linearly with the order parameter. An example of such a linear coupling is the classical dipole moment, where the energy shift in the coupling due to the applied field is given by the linear term $|h|M$ [139]. Consider the quadratic expansion of Landau free energy from the previous section, now with the inclusion of the external field (or the linear term):

$$F(M) = |h|M - \frac{1}{2}(r)M^2 + \frac{1}{4}M^4 \tag{C.15}$$

Its magnetization dynamic is

$$\frac{dM}{d\tau} = |h| + rM - M^3. \tag{C.16}$$

Equilibrium Points and Bifurcation Diagram

We wish to find the equilibrium points (M^*, h_c), where h_c is the critical point of h. We take the first derivative of Equation C.16 and find the equilibrium values of M^*:

$$0 = r - 3(M^*)^2 \tag{C.17}$$

$$\Rightarrow M^* = \sqrt{\frac{r}{3}}. \tag{C.18}$$

We use M^* to Equation C.16 to find the set of values of h_c that describe the three-dimensional surface regions of the equilibrium points (M^*, r, h):

$$0 = |h_c| + rM^* - (M^*)^3 \tag{C.19}$$

$$\Rightarrow h_c = \pm\left(r\sqrt{\frac{r}{3}} - \sqrt{\frac{r}{3}}^3\right) \tag{C.20}$$

$$\Rightarrow h_c = \pm\frac{2r}{3}\sqrt{\frac{r}{3}} \tag{C.21}$$

Here, we have identified two sets of equilibrium points for M and h. For $r > 0$, $h_c = \pm\frac{2r}{3}\sqrt{\frac{r}{3}}$ divide the parameter space into two regions. In the region bounded by functions $h_c = \pm\frac{2r}{3}\sqrt{\frac{r}{3}}$ and $r > 0$, the parameter space has three fixed points. For the other region, the system has only one fixed point. Finally, when $h = 0$ and $r = 0$, $M = 0$, there is a triple root, which is the *cusp*. Figure C.4 visualizes the parameter

space of the equilibrium points as M vs. h. The discriminant[2] (Δ) of Equation C.19 can also tell us regions of the three-dimensional parameter space, where there are different numbers of fixed points (M^*, r, h_c).

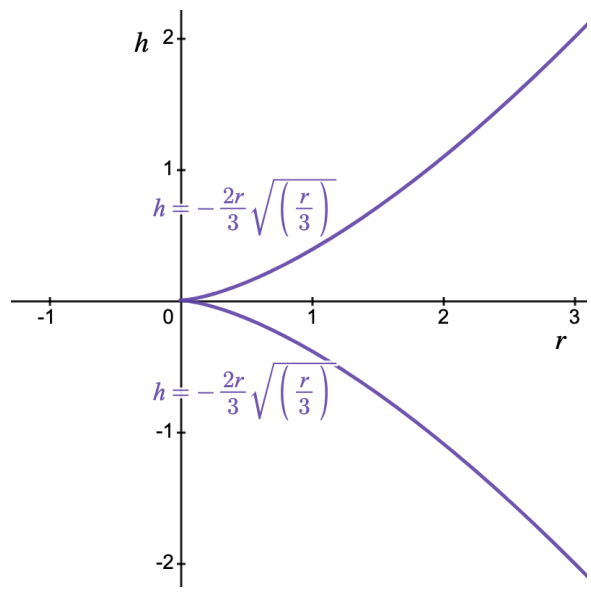


Figure C.4: For the quadratic Landau free energy expansion with a linear term, the parameter space of equilibrium points is bounded by $h_c = \pm\frac{2r}{3}\sqrt{\frac{r}{3}}$ (purple lines). Within this bounded region (h_c), the system demonstrates three equilibrium points. Outside of the region has just one equilibrium point.

The bifurcation points of Equation C.16 are a three-dimensional surface as a function of h and r (Figure C.5). The bifurcation analysis of M vs. r shows that when $h > 0$, there is a saddle-node bifurcation. When $h = 0$, there is a pitchfork bifurcation. When $h < 0$, there is again a saddle-node bifurcation. When different bifurcations

[2]Cubic equations of the form $x^3 + px + q = 0$ have the discriminant, $\Delta = -4p^3 - 27q^2$.

come together as such, there is a "fold" in the bifurcation curve. This "bifurcation of bifurcations" is called *cusp catastrophe*, and they occur at the spontaneous moment of geometric "fold" in the parameter space and the order parameter.

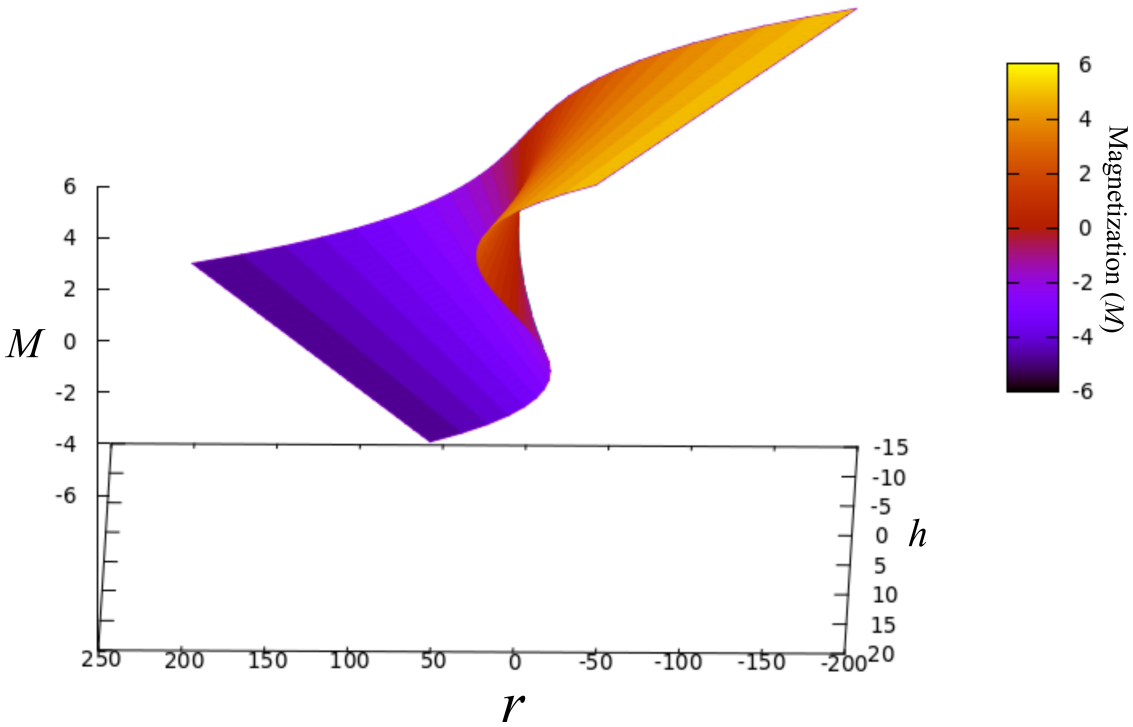


Figure C.5: Bifurcation diagram of the quadratic expansion of free energy with a linear term is a three-dimensional surface: M vs. h vs. r. *Cusp catastrophe* is observed at the imperfect bifurcation point of $r = 0$, $h = 0$. At this point, bifurcation of bifurcation points come together.

Ground States and Phase Transition

When $h = 0$, the system is trivially equal to Equation C.10 and has a double well potential (Figure C.3) [3]. With $h \neq 0$, the symmetry of the quadratic free energy starts to break. Symmetry breaking occurs when a small fluctuation acting on the system near the critical point alters the fate of the system. The plot of the quadratic Landau free energy expansion with a linear term (Figure C.6) demonstrates the symmetry breaking with non-zero $|h|$. Additionally, the shift in ground states can be seen with different non-zero r.

---

[3]We highlight that the double well potential is a field of great interest in many various areas of statistical physics and will serve as an important tool in analyzing phase transition of tissue phenotypes in Chapters 5 and 8.

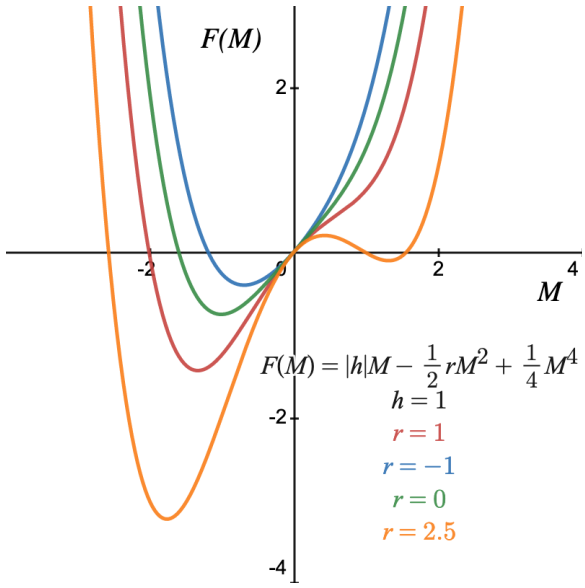


Figure C.6: Plot of the quadratic Landau free energy expansion with values of $r = -1, 0, 1$: $F(M)$ vs. M. There is a clear symmetry breaking with non-zero values of h. Furthermore, there are two different second-order phase transitions.

We note that when $r > 0$ is fixed, there are bistable steady states for certain values of h. For instance, as h increases in the lower branch of stability (of the equilibrium surface (r, h, M)), at the point of saddle-node bifurcation, h must "jump up" to the higher branch stable branch. Likewise, for the steady state to return to the lower branch of stability, h must "jump down" on the other end of the saddle-node bifurcation. These "jump" discontinuities are *first-order phase transitions*, and characterize *hysteresis*. Hysteresis is encountered in countless areas of physics and biology (i.e. thermal cycling and electric displacement fields of a ferroelectric material). They are considered the memory effect of a system, as there is a delay in "jump" discontinuities from different directions. For the 2D Ising model with external field, the *hysteresis*

is observed, when there is a fast thermal change in fluctuation, resulting in the delay of phase transition.