

Study of Critical Phenomena with Monte Carlo and Machine Learning Techniques

Ahmadreza Azizi

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Physics

Michel Pleimling, Chair
Uwe C. Täuber
Shengfeng Cheng
Cihan Nadir Kaplan

May 7, 2020
Blacksburg, Virginia

Keywords: Critical Phenomena, Voter Model, Dynamical Scaling, Machine Learning,
Restricted Boltzmann Machine, Convolutional Neural Network

Copyright 2020, Ahmadreza Azizi

Study of Critical Phenomena with Monte Carlo and Machine Learning Techniques

Ahmadreza Azizi

(ABSTRACT)

Dynamical properties of non-equilibrium systems, similar to equilibrium ones, have been shown to obey robust time scaling laws which have enriched the concept of physical universality classes. In the first part of this Dissertation, we present the results of our investigations of some of the critical dynamical properties of systems belonging to the Voter or the Directed Percolation (DP) universality class. To be more precise, we focus on the aging properties of two-state and three-state Potts models with absorbing states and we determine temporal scaling of autocorrelation and autoresponse functions.

We propose a novel microscopic model which exhibits non-equilibrium critical points belonging to the Voter, DP and Ising Universality classes. We argue that our model has properties similar to the Generalized Voter Model (GVM) in its Langevin description. Finally, we study the time evolution of the width of interfaces separating different absorbing states.

The second part of this Dissertation is devoted to the applications of Machine Learning models in physical systems. First, we show that a trained Convolutional Neural Network (CNN) using configurations from the Ising model with conserved magnetization is able to find the location of the critical point. Second, using as our training dataset configurations of Ising models with conserved or non-conserved magnetization obtained in importance sampling Monte Carlo simulations, we investigate the physical properties of configurations generated by the Restricted Boltzmann Machine (RBM) model.

The first part of this research was sponsored by the US Army Research Office and was accomplished under Grant Number W911NF-17-1-0156.

The second part of this work was supported by the United States National Science Foundation through grant DMR-1606814.

Study of Critical Phenomena with Monte Carlo and Machine Learning Techniques

Ahmadreza Azizi

(GENERAL AUDIENCE ABSTRACT)

Physical systems with equilibrium states contain common properties with which they are categorized in different universality classes. Similar to these equilibrium systems, non-equilibrium systems may obey robust scaling laws and lie in different dynamic universality classes. In the first part of this Dissertation, we investigate the dynamical properties of two important dynamic universality classes, the Directed Percolation universality class and the Generalized Voter universality class. These two universality classes include models with absorbing states. A good example of an absorbing state is found in the contact process for epidemic spreading when all individuals are infected. We also propose a microscopic model with tunable parameters which exhibits phase transitions belonging to the Voter, Directed Percolation and Ising universality classes. To identify these universality classes, we measure specific dynamic and static quantities, such as interface density at different values of the tunable parameters and show that the physical properties of these quantities are identical to what is expected for the different universal classes.

The second part of this Dissertation is devoted to the application of Machine Learning models in physical systems. Considering physical system configurations as input dataset for our machine learning pipeline, we extract properties of the input data through our machine learning models. As a supervised learning model, we use a deep neural network model and train it using configurations from the Ising model with conserved dynamics. Finally, we address the question whether generative models in machine learning (models that output objects that are similar to inputs) are able to produce new configurations with properties similar to those obtained from given physical models. To this end we train a well known generative model, the Restricted Boltzmann Machine (RBM), on Ising configurations with either conserved or non-conserved magnetization at different temperatures and study the properties of configurations generated by RBM.

The first part of this research was sponsored by the US Army Research Office and was accomplished under Grant Number W911NF-17-1-0156.

The second part of this work was supported by the United States National Science Foundation through grant DMR-1606814.

To my beloved parents

Acknowledgments

This thesis is the peak of my Ph.D. journey that was accompanied with encouragement, enthusiasm and frustration. Many people have helped me reach this success.

First and foremost, I am sincerely thankful to my dear advisor, Dr. Michel Pleimling, who offered me his invaluable guidance, mentorship, and support. Particularly, I am deeply grateful to him for encouraging me to explore my interests and enthusiasm in different research areas.

I truly appreciate Dr. Uwe Täuber for his remarkable suggestions and discussions during our research meetings. Also, many thanks to my other committee members, Dr. Shengfeng Cheng and Dr. Nadir Kaplan, who provided me valuable feedback for my research.

During my time at Robeson Hall, I was fortunate to be in touch and have exciting discussions with many talented researchers. I am thankful for my deep conversations with Shadi Esmaili, Priyanka, Harshwardhan Chaturvedi, Bart Brown, Jacob Carroll, James Stidham, Shannon Serrao, Riya Nandi, Jason Czak, and Ruslan Mukhamadiarov. I always found our shared office as a live environment with full of energy and unforgettable fun moments.

Many thanks to our graduate coordinator Betty Wilkins and program support technician Katrina Loan for their care and patience in the administrative processes. Special thanks to our computer systems experts Roger Link, and Travis Heath, for their daily efforts to keep our machines and cluster systems running nicely.

It's my fortune to gratefully acknowledge my close and dear friends: Hadi Parsian, Arash Sarshar, Yaser Keneshlo, Ehsan Asghari Ghara, Sorour Amiri, and Arian Vezvaei for their incredible support and friendship through happy and hard moments.

Many thanks to my loving parents, Fatemeh Babashah and Abbas Azizi, whose love, endless patience, and support have always helped me stay strong in my life.

Last but not least, my deepest love and gratitude to my companion, Armaghan, for her countless sacrifices and unconditional support, especially during stressful moments.

Contents

List of Figures	ix
Preface	1
I Critical Phenomena in Presence of Symmetric Absorbing States	2
1 Introduction	3
1.1 Voter models	5
1.2 Langevin description of the Generalized Voter Model	7
1.3 Width of the interfaces between absorbing states	8
2 Models And Quantities	10
3 The Two-State Model	14
3.1 Interactions with nearest neighbors only	14
3.2 Interactions with up to third nearest neighbors	16
4 The Three-State Model	21
4.1 Interactions with nearest neighbors only	21
4.2 Interactions with up to third nearest neighbors	23
5 Conclusion I	29

6	A Microscopic Model With Tunable Parameters	31
6.1	Model definition	33
7	Numerical Phase Diagram and Universality Classes	35
7.1	Quantities	35
7.2	Phase diagram of the microscopic model	37
7.2.1	One critical line belonging to the Voter universality class	39
7.2.2	Two critical lines belonging to the Ising and Directed Percolation universality classes	41
8	Conclusion II	44
	Appendix A: Derivation of Langevin Equation	46
9	Interface Width in Systems with Absorbing States	50
9.1	Interface width scaling in the Directed Percolation universality class	54
9.2	Voter-Voter interface width in the LD model.	60
II	Application of Machine Learning Models in Physical Systems	62
10	Introduction	63
11	Supervised Learning of the Kawasaki Model Configurations	65
11.1	Model	65
11.2	Convolutional Neural Network	69
11.2.1	Structure of convolutional neural networks	70
11.2.2	CNN Model training on Ising configurations.	71
12	Restricted Boltzmann Machine	74
12.1	Architecture of the RBM	75
12.2	Training configurations	79

12.3 Learning the RBM model	82
13 Conclusion	89
Bibliography	90

List of Figures

1.1	(a) Illustration of the flip probabilities of the classical voter model. (b) The system has two different absorbing states.	5
1.2	Schematic representation of the parameter space of a-b in the Generalized Voter model [11]. The space is divided in three phases by Voter, DP and Ising critical lines. These phases are absorbing (A), ferromagnetic (F) and paramagnetic (P).	7
1.3	Configuration of two absorbing states with a shared interface where both sides of the lattice are at the DP critical point. (a) Initial condition at $t=0$. (b) Evolution of the interface after 2000 Monte Carlo steps.	9
3.1	(a) At the critical temperature $T_c = 2/\ln(3)$ the two-time autocorrelation function $C(t, s)$ for the KBB model shows a data collapse when assuming the standard aging scaling behavior (3.2) with $b = 0.120(5)$. Inset: $C(t, 0)$ exhibits a power-law tail with an exponent $\lambda = 1.00(2)$ (as indicated by the dashed line). (b) The scaling proposed in [26, 27] for the two-time autocorrelation function does not yield a data collapse nor is the expected scaling function (3.1) (dashed line) recovered. The data result from averaging over 8000 realizations for systems with linear size $L = 800$	15
3.2	(a) At the critical temperature $T_V = 1.7585$ the two-time autocorrelation function $C(t, s)$ for the LD model shows the same scaling behavior as the critical KBB model, with the same exponents. The data result from averaging over 8000 realizations for systems with linear size $L = 800$. (b) A data collapse is also encountered for the autoresponse function $\rho(t, s)$ with the same value for the scaling exponent. These simulations were done for systems with 400×400 sites and the amplitude $h = 0.05$ for the spatially random magnetic field. The data result from averaging over at least 250000 different runs.	18

3.3	Time evolution of the magnetization $M(t)$ for the critical KBB and LD models with nearest-neighbor couplings when preparing the system in a state with initial magnetization $M(0) = 0.5$. Systems of linear size $L = 400$ have been simulated, and the data result from averaging over 25000 independent runs.	19
3.4	(a) Scaling of the two-time autocorrelation and (b) of the two-time autoresponse function at the Ising critical point, with $T_I = 4.812$, of the $q = 2$ LD model with twelve nearest neighbors. The inset in (a) shows that the long-time decay of $C(t, 0)$ is in agreement with the known exponent $\lambda/z \approx 0.733$ (indicated by the dashed line). The autocorrelation data result from averaging over at least 10000 realizations, whereas for the autoresponse function at least 400000 independent runs have been made. In (b) the amplitude of the magnetic field has been set to $h = 0.05$. For both quantities systems of linear extent $L = 400$ have been simulated.	20
4.1	The steady-state density of active sites ρ_a and the steady-state magnetization M as a function of temperature for the LD model with $q = 3$ and only nearest-neighbor interactions. Both order parameters display a jump at the same temperature. The data, obtained for a system with linear extent $L = 400$, result from a time average (after having discarded the early time data) and an ensemble average over 100 independent runs. Error bars are smaller than the sizes of the symbols.	22
4.2	Two-time autocorrelation function for the three-state LD model with only nearest-neighbor interactions at the temperature $T = 1.237$ of the discontinuous phase transition. The autocorrelation is plotted as a function of the ratio t/s . The linear extent of the system is $L = 800$, and 10000 independent runs have been performed.	23
4.3	Scaling of the autocorrelation $C(t, s)$ of the equilibrium Potts model with (a) $q = 5$ states and (b) $q = 7$ (inset $q = 8$) states at their discontinuous phase transition temperature $T = [\ln(1 + \sqrt{q})]^{-1}$. A perfect aging scaling prevails at the weak first-order transition for the $q = 5$ case, whereas at the stronger discontinuous transitions of the $q = 7$ and $q = 8$ cases a simple aging data collapse can not be achieved. The dashed line in the inset of (a) indicates that $C(t, 0)$ decays for $q = 5$ algebraically with an exponent ~ 1.17 . The system size is $L = 800$, and the data result from averaging over at least 9000 independent runs.	25

4.4	The steady-state density of active sites ρ_a and the steady-state magnetization M as a function of temperature for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. The behavior of the two order parameters indicate the presence of two phase transitions that take place at different temperatures. The data, obtained for a system with linear extent $L = 400$, result from a time average (after having discarded the early time data) and an ensemble average over 150 independent runs. Error bars are of the order of the symbol sizes.	26
4.5	Number of flipped spins N_f (main figure) and survival probability P_s (inset) as a function of time for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. Both quantities display an algebraic behavior at the absorbing phase transition temperature $T_a = 3.754$. The measured exponents $\eta = 0.238(10)$ for N_f and $\delta = 0.455(6)$ for P_s agree with the values of the Directed Percolation universality class [39, 40]. The data, obtained for a system with $L = 400$, result from averaging over at least 10000 independent runs. . .	27
4.6	Main figure: Binder cumulant for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. Systems of different sizes have been simulated. Inset: finite-size scaling of the magnetization M when using the critical exponents of the $q = 3$ equilibrium Potts model. Error bars are smaller than the symbol sizes. The data result from averaging over 200 independent runs. . .	28
7.1	Phase diagram in the $b - a$ plane obtained from simulations of the microscopic model. The properties of the phase transitions taking place at particular values of the system parameters a and b (filled symbols) have been elucidated. Along the green dashed segment a special point exists where all three critical lines merge. No attempt was made to locate this special point. The full lines are guides to the eye.	38

7.2	Time dependence of various quantities for systems with $b = 0$ and different values of a . Based on these data a phase transition belonging to the Generalized Voter universality class takes place at $a = 0$. The different quantities display a typical behavior: (a) the interface density ρ decays as $1/\ln t$, (b) the average number of flipped cells N_f also decays like $1/\ln t$, which is compatible with $\eta_{GV} = 0$, and (c) the survival probability decays algebraically with an exponent $\delta_{GV} \approx 1$. The same behavior is encountered for $b < 0$. This is illustrated in the inset of panel (a) for the case $b = -8$ for which the phase transition is found to take place at $a = -4.132(2)$. The inset in panel (b) shows that at early times the average number of flipped cells increases at the critical point $a = 0$ algebraically with an exponent 0.19 (indicated by the dashed line) The data for the interface density result from averaging over at least 3000 independent runs. For the other two quantities averages were done over at least 5000 independent runs.	40
7.3	Scaling of the autocorrelation function $C(t, s)$ for $a = 0$ and $b = 0$. Simple aging scaling (7.5) with the exponents $\tilde{b} = 0.12$ and $\lambda = 1$ (indicated by the dashed line) is encountered, in agreement with the expected behavior at a voter critical point [58]. The inset shows that the same behavior prevails for $b = -8$ and $a = -4.132$. These data result from averaging over 5000 independent runs.	41
7.4	Binder cumulant U for various numbers of cells L^2 as a function of a when $b = 40$. The different lines intersect close to $a = 22.16$. The value of U at this point is close to the universal value $U_I \approx 0.61$ for the two-dimensional Ising universality class [65,67]. For these equilibrium simulations the first 300,000 time steps were discarded before sampling data for a time average. The final data result from an additional averaging over typically 1000 different runs. Error bars are comparable to the symbol sizes. The lines are guides to the eye.	42
7.5	Time dependence of (a) the average number of flipped cells N_f and (b) the survival probability P_s for $b = 40$ and different values of a . These data indicate that a Directed Percolation phase transition takes place for $a = 23.332(4)$. The data result from averaging over at least 4000 different realizations of the noise.	43
9.1	Initial configuration of LD model with width $\omega = 400$ and interface size $L = 200$. Left side (green) and right side (red) of the lattice contain 0 spins and +1 spins respectively.	50

9.2	Snapshots of the configurations at LD model Ising critical point at different time steps (a) $t = 1$, (b) $t = 100$, (c) $t = 500$, (d) $t = 1000$. Spins in the left and right sides of the system are initially set to 0 and +1 respectively. Both sides are at the Ising critical point ($T_I = 4.812$ and $z = 12$). Time evolution of the model shows the interface h does not change smoothly in space and therefore W will not be meaningful.	51
9.3	A schematic description of W as a function of time t and interface size L . (a) For interface size L_i , interface fluctuations W_i grow in time with exponent β_w and at the saturation point t_{s_i} , W_i reaches a constant value. (b) Scaling of W with respect to L . Given z , the best value for α_w is obtained when all graphs of W s collapse on to one graph.	53
9.4	(a) A lattice example of random deposition (RD). From a randomly chosen site over the surface of deposited particles, particles fall vertically with a constant rate in time and stick to the top of the column. (b) An illustration of ballistic deposition (BD). The deposition of the falling particle depends also on the neighboring columns as well. In BD, in addition to RD, particles stick to the sides of neighboring particles if possible (particle A and C).	54
9.5	Snapshots of the LD configurations at DP critical point ($T_{DP} = 4.738$) at different time steps (a) $t = 1$, (b) $t = 500$, (c) $t = 2000$, (d) $t = 5000$, (e) $t = 20000$, (f) $t = 40000$. Spins in the left and right sides of the system are initially set to 0 and +1 respectively.	56
9.6	Dynamical scaling of the fluctuation of the interface width $W(L, t)$ in the LD model for a fixed system width $w = 400$. Left graph $W(L, t)$ for different interface lengths L . As shown in the right graph, after $t \approx 5000$ the systems for different lengths of L reach the saturation regime with the scaling exponent $\alpha_w = 0.08$. Note that the exponent α_w is determined based on the large system sizes as the linear extension $L = 200$ is too small to exhibit this scaling. . . .	57
9.7	Snapshots of the MGVM with size of $L \times \omega = 200 \times 200$ where cells can take on any values in the range of $[-1, 1]$ with increment $1/100$. Color profile of cells includes RGB colors from red ($\phi = -1$) to blue ($\phi = +1$). The system initially starts with cell values $\phi = -1$ on the left and $\phi = +1$ on the right. To measure W , the actual configurations of the system (a) and (c) are transformed to (b) and (d) where cells with $\phi < 0$ respectively $\phi > 0$ are assigned the value -1 respectively $+1$. Hence measurements will be performed on (b) and (d). . . .	58

9.8	Dynamical behavior of the interface fluctuation, $W(L, t)$, in the MGVM with the system width $W = 100$ and $N = 100$. (a) For interface sizes $L \geq 200$ the growth exponent $\beta_w = 0.65$ is found which is in agreement with the β_w value obtained in the LD model. (b) We plot the scaled interface fluctuations with the value for α_w found in the LD model for $L = 50$ and $L = 100$. No stationary data are available for large values of L	59
9.9	Voter-Voter interface in the LD model with the fixed size of $L \times \omega = 100 \times 800$ at different time steps. (a) The system starts with cell values $\sigma = 0$ on the left (green cells) and $\sigma = +1$ on the right (red cells). For the other configurations the times are (b) $t = 2000$, (c) $t = 5000$, (d) $t = 10000$, (e) $t = 20000$, (f) $t = 30000$. After about 1000 time steps the interface between the two states is dissolved and there is no saturation time for the system. In the above system, the green state will fill the whole system after about 50000 time steps.	60
9.10	Log-log plot of the fluctuations of the Voter-Voter interface width versus time in the LD model with different system sizes. Fluctuations of the interface grow in time and do not reach saturation.	61
11.1	Energy values at different temperatures versus time for the Ising model with Kawasaki dynamics for system size $L=40$. Except for the data with $T=1.2$, all graphs reach their steady state after about $t = 4 \times 10^6$ (for $T=1.2$ it is about 9×10^6) Monte Carlo steps. The data are obtained from averaging over 10 different runs.	67
11.2	Configurations of the Ising model with conserved magnetization after equilibration at $T=2.0$ (a), $T=2.2$ (b), $T=2.6$ (c), $T=3.0$ (d), $T=3.4$ (e), and $T=4.0$ (f). The system size is 40×40 . These configurations are given to the CNN model as the training dataset.	68
11.3	Operation of the kernel matrix on the input image. In this example the input size is 4×4 and the kernel size is 2×2 . The kernel matrix is convolutionally multiplied by the input pixels which results in an output vector o with 6 elements. Strides values (stride_x and stride_y) show the number of steps taken in the horizontal and vertical directions.	70
11.4	An example of a Convolutional Neural Network. The above model has three convolutional layers. The first layer includes 5 filters and each filter has its own kernel matrix with size (3×4) . The last convolutional layer is connected to the dense layer with only two nodes. As an example, the nodes values are 2.55 and 0.45. These numbers are normalized to 1 by the softmax activation function.	71

11.5	Architecture of the CNN model applied to Ising configurations. A softmax function is used as the activation function for all convolutional layers. Properties of each layer are written inside of the layer rectangle. For example, the first convolutional layer has 16 filters with kernel size (3×3) , whereas the second convolutional layer has only 8 filters with the same kernel size as the first layer. Finally the dense layer has only two nodes which are normalized by the softmax activation function.	72
11.6	Output of CNN model as a function of temperature. For different values of system size L all graphs intersect with each other at the critical point. This suggests that the output of the CNN model can be considered as the order parameter.	73
12.1	A sample architecture of the RBM. The number of hidden nodes is usually smaller than the number of the visible nodes. The right hand side nodes represent bias terms in the Equation (12.2).	75
12.2	Gibbs sampling in an RBM model from the initial values V_1 to V_N . In the case of $k=1$, the hidden layer H is calculated by $P(H_1 V_1)$ and $\hat{V} = \{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n\}$ layer is obtained from $P(\hat{V} H_1)$. Same process is done to sample $\hat{V}_2$ from the initial value V_2 . These $\hat{V}$ s are the sampled configurations via Gibbs sampling.	80
12.3	Configuration of the Ising model with the system size 8×8 after equilibration at $T=1.0$ (a), $T=1.4$ (b), $T=1.8$ (c), $T=2.269$ (d) $T=2.6$ (e), $T=3.0$ (f). Spins with value +1 and 0 are colored green and red respectively.	81
12.4	Generation of new configurations with the RBM model at different temperatures. The temperatures of the configurations are $T = 1.2$ (a), $T = 1.9$ (b), $T = 2.0$ (c), $T = 2.269$ (d), $T = 2.4$ (e) and $T = 2.6$ (f).	83
12.5	Magnetization graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations. The quality of RBM graphs indicate that the model has learned to generate images with correct number of +1 and -1 spins.	84
12.6	Energy graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations.	85
12.7	Correlation functions obtained over the Monte Carlo (MC) configurations and RBM configurations (RBM) at different temperatures with (a) $L = 4$ and (b) $L = 8$. For $L = 4$ the $C(r)$ values in RBM configurations are completely consistent with MC configurations. However, for $L = 8$ a systematic error is seen at temperatures close to the critical point.	86

12.8	Magnetization graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. In both graphs, 10,000 configurations are generated by RBM. Results show that except from $T = 1.0$, at any other temperatures, magnetization in Monte Carlo configurations are outside of the error bars of magnetization values in the generated configurations.	87
12.9	Energy graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations. Unlike conserved magnetization, energy graphs in MC and RBM are consistent.	87
12.10	Correlation function measured on RBM configurations and MC configurations with system size (a) $L = 4$ and (b) $L = 8$. In both graphs, measurements over RBM configurations indicate a systematic error when compared to the Monte Carlo configurations (MC).	88

Preface

Part I

Critical Phenomena in Presence of Symmetric Absorbing States

Chapter 1

Introduction

Among non-equilibrium models there exist many systems which exhibit a phase transition from an active state to an absorbing state. By definition, an absorbing state is a state from which a system can not escape so that it remains in that state forever. An active state is a state where the system can change from one configuration to the next, in accordance with the dynamical rules.

Perhaps the most important examples of such models belong to the Directed Percolation (DP) universality class. Following the 1981 conjecture by Janssen and Grassberger, a system with short-range dynamics belongs to the DP universality class if it displays a continuous phase transition from a fluctuating active phase into a unique absorbing state [1,2]. In these systems the order parameter is a positive one-component quantity and the system has no attributes such as additional symmetries [3]. In nature, many systems with one absorbing state have been found to belong to the DP universality class. Onset of turbulence, contact process, and cellular automata phase transitions are only a few examples of systems belonging to this universality class [3].

While DP universality class behavior has been observed at many transitions, there are also ample examples of non-equilibrium models with phase transitions from an active to an absorbing phase which do not satisfy the above conditions. Therefore, these models belong to some other universality classes such as the Isotropic Percolation universality class, the pair Contact Process (PCP) universality class or the Voter universality class [4].

All the universality classes mentioned above have been studied for a long time and their different characteristics, such as critical exponents, are well understood from numerous studies using classical field theory or computational tools. However, for the Voter universality class there are still open questions regarding dynamical features of systems belonging to this universality class. While systems adhering to the Voter universality class are relatively easy to understand, some of their universal characteristics are quite unique as will be explained in Section 1.1. For example, the short term dynamical behavior at a voter critical point is

different than expected for reasons not yet well understood [5]. Recently, a field theoretical understanding of the linear Voter model has been discussed in [6] where the authors compute the values of the Voter universality class critical exponents. However no conclusion is made for other types of Voter models.

Voter models (see the description of the classical voter model in Section 1.1 and in Figure 1.1) have been first introduced in [7, 8] to describe the spreading of two voter opinions on a two-dimensional lattice using the following simple rule: a randomly selected site on a two-dimensional lattice takes the opinion of one of its neighbors with some probability p . Clearly these models have Z_2 symmetry and it is shown that although the static order parameter is discontinuous at the phase transition from the active to the absorbing state, dynamical scaling is still observed. This behavior is quite different from models in other dynamical universality classes where all quantities are continuous. Another general feature of voter models and, accordingly, the Voter universality class is that dynamics takes place at the boundaries only as there is no dynamics in bulk regions.

Knowing this background about these models, it was a motivation for us to explore and study different characteristics of voter models. Specially, as described in Chapter 2, we focus on the temporal behavior of two different voter models, one a linear model and the other a non-linear one, and investigate the autocorrelation and the autoresponse functions at the phase transition.

In 2002 [9] it was argued that there should be a fundamental connection between Ising, Voter, and DP universality classes. As described in Section 1.2, this connection was not understood until 2005, when the authors of [10] introduced the Generalized Voter Langevin equation and gave a coarse-grained description of the above phenomenon. Numerical solution of the coarse grained equation of motion reveals that lines of phase transitions belonging to the Ising and DP universality classes can merge and constitute one critical line belonging to the Voter universality class.

Despite progress in understanding the above phenomenon in a coarse-grained model, to our knowledge no model at the microscopic level has been proposed that is able to exhibit two merging critical lines, one with Ising and one with DP critical behavior, into a Voter critical line, similar to Figure 1.2. This provides us with the motivation to build a microscopic model with tunable parameters that displays this behavior. We extensively describe the properties of our proposed model in Chapter 6.

Before getting into the details of our contribution to this area, in order to provide the reader with some background information on the concepts and models mentioned here, we first explain the Voter universality class, the classical voter model, and the Generalized Voter Langevin equation in more details. In Section 1.1 as an example of a Voter phase transition, we briefly introduce the classical voter model and discuss some of its features. In Section 1.2 we give a more detailed picture of the Generalized Voter Model (GVM) Langevin equation.

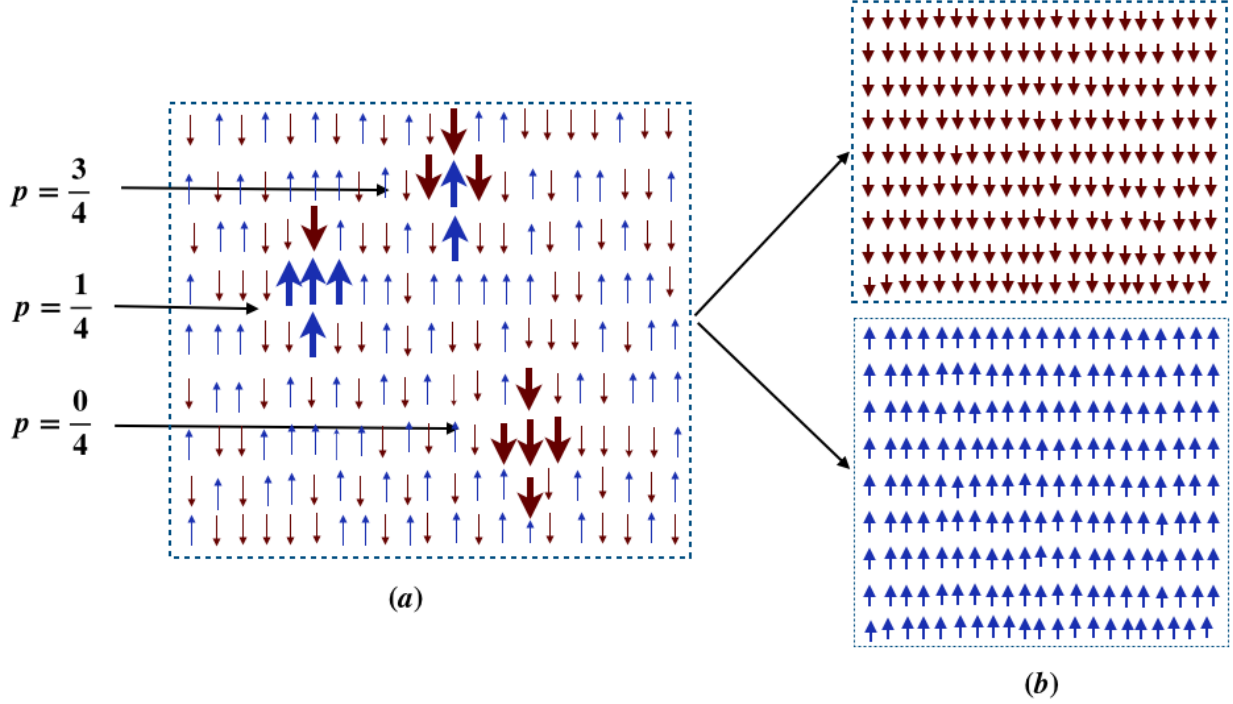


Figure 1.1: (a) Illustration of the flip probabilities of the classical voter model. (b) The system has two different absorbing states.

1.1 Voter models

Based on the definition in [10], every two-dimensional system driven by interfacial noise between two absorbing states belongs to the Voter universality class if the system has either Z_2 symmetry or conserved magnetization. Therefore, for the sake of comparison, voter models (models in the Voter universality class) have two absorbing states while DP models have only one, with a voter model exhibiting local Z_2 symmetry or a global conservation of magnetization (magnetization is defined to be the difference between the number of up and the number of down spins).

To clarify the above statement, the explicit definition of the classical voter model can be helpful. The classical voter model is an Ising spin (± 1) model on a two-dimensional lattice with a spin flip dynamics where a randomly chosen spin is flipped with probability:

$$p = \frac{\# \text{ of neighboring spins with opposite values}}{\# \text{ of neighboring spins}} \quad (1.1)$$

From this definition, as it is shown in Figure 1.1, one can immediately identify two possible absorbing states where all voters (spins) on the lattice are either up (+1) or down (−1).

With the above choice of probabilities the classical voter model undergoes an active state to absorbing state phase transition belonging to the Voter universality class. At the critical point of the model, two important phenomena happen simultaneously. On the one hand the Z_2 symmetry of the model is broken so that one spin state (up or down) dominates. On the other hand a transition to an absorbing phase takes place. Therefore at the critical point of this model, and this is true for other voter models, not only a transition to the absorbing state occurs but also the local Z_2 symmetry is broken. This is the main conceptual difference between models in the Voter universality class and models in the DP universality class.

This observation raises an interesting question: Is there a way to change the flip probabilities of the system so that the Z_2 symmetry is broken before the system experiences the absorbing phase transition?

The above question was addressed through the study of voter-type models defined by Lipowski and Droz [9] which lead to a deep understanding of the relationship between Voter, DP, and Ising universality classes. As shown in [9] and extensively discussed in Chapter 2 of part I, a small modification of the Metropolis flip algorithm can lead to a new model with two absorbing states (similar to the classical voter model). Furthermore, once the interaction range increases to 12 nearest neighbors, the model exhibits two separate phase transitions. The first transition is due to the breaking of Z_2 symmetry which belongs to the Ising universality class. The second one is when the dominating spin state falls into its absorbing state, belonging to the DP universality class. This model, which we call LD from now on, is defined as follows: On a two-dimensional lattice of spins σ (with values 0 or 1), a chosen spin σ_i is not flipped if all of its nearest spins are in the same state as σ_i . But if this is not the case, then the spin is flipped according to the Metropolis probability $\min\{1, e^{-\Delta E/T}\}$ where ΔE is the change of energy E , with

$$E = - \sum_{j \in n.n} \delta_{\sigma_i, \sigma_j}, \quad (1.2)$$

whereas T is the temperature (setting $k_B = 1$). This model exhibits a Voter phase transition at $T = 1.758$ once 4 neighboring spins are considered. However, as discussed earlier, once the number of neighboring sites in the algorithm changes to 12, the model experiences separate Ising and DP phase transitions at $T = 4.812$ and $T = 4.738$ instead of one voter critical point. Considering the number of neighboring sites as a parameter in the model, one can view this as an indication that when changing the model parameters (T and the number of neighboring spins) the Ising and the DP critical points are merged into the Voter critical point. Therefore, Lipowski and Droz concluded that there seems to be a connection between the Voter, Ising, and DP universality classes. In the next subsection we explain how this interpretation becomes more clear once we look at a coarse-grained description of voter models.

It is also interesting to ask what happens to this picture in a three state microscopic spin model with $\sigma \in \{0, 1, 2\}$. This model has Z_3 symmetric states and therefore breaking of

the Z_3 symmetry might take place at the same time as a transition from an active to an absorbing phase. This question will be addressed in Chapter 4.

1.2 Langevin description of the Generalized Voter Model

According to the LD model, there seems to be a fundamental connection between Voter, DP, and Ising universality classes. This relationship was not fully understood until the authors of [11] introduced the following Langevin equation, called Generalized Voter Model (GVM) Langevin equation:

$$\frac{\partial \phi}{\partial t} = \nabla^2 \phi + (a\phi - b\phi^3)(1 - \phi^2) + \sqrt{1 - \phi^2} \eta \quad (1.3)$$

In the above equation ϕ is a coarse-grained version of the spin σ in the microscopic model. More importantly, it was shown in [11] that in the parameter space of the model (a - b), different values of a and b lead to three lines of Voter, DP, and Ising critical points (Figure 1.2).

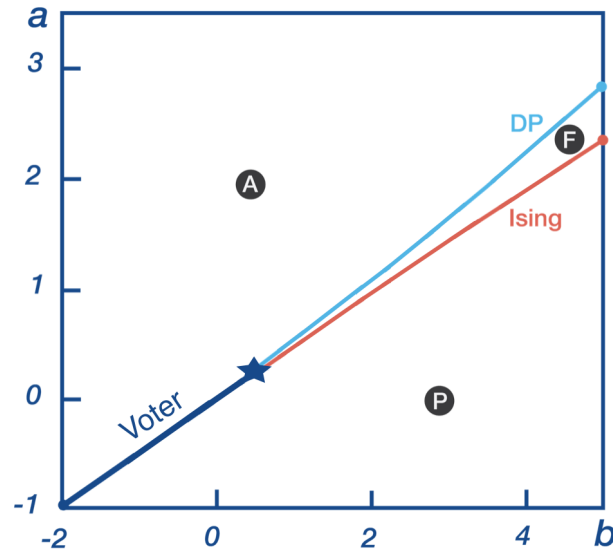


Figure 1.2: Schematic representation of the parameter space of a - b in the Generalized Voter model [11]. The space is divided in three phases by Voter, DP and Ising critical lines. These phases are absorbing (A), ferromagnetic (F) and paramagnetic (P).

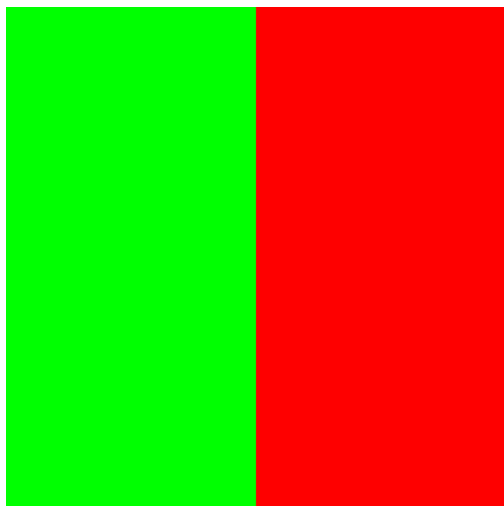
The success of the above Langevin description motivated some efforts ([12], [13], [14]) to build a microscopic model which presents properties similar to those encountered when numerically solving the GVM Langevin equation.

Chapter 6 of this thesis is devoted to our research on finding a microscopic model which describes the full features of the Langevin Equation (1.3). As it is shown in Section 6.1, we propose a microscopic model with tunable parameters that displays those three critical lines. The connection between our microscopic model and the GVM Langevin equation is elucidated by applying the Kramers-Moyal expansion of the Master Equation of the microscopic model.

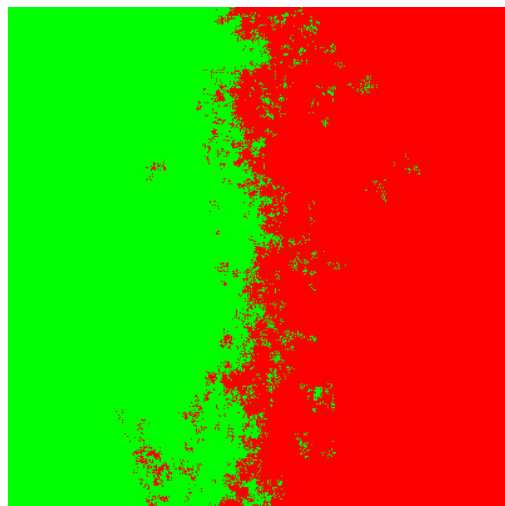
1.3 Width of the interfaces between absorbing states

As we saw in the previous section, in both LD and our microscopic model there are two absorbing states. For suitable choices of the model parameters, the phase transitions from the active to the absorbing state belongs to either the Voter or the DP universality class. With these microscopic models at hand, we can also ask other interesting questions.

Another interesting subject that we study in this thesis is the dynamical behavior of the width of an interface between two absorbing states. To make it more clear, consider the LD model. As it is shown in Figure 1.3 the system initially is set to be on one side in one of its absorbing states and on the other side in the other absorbing state. Then we let the system evolve and observe the behavior of the interface separating these two absorbing states. However only for the critical values of the parameters is dynamical scaling of the interface width obtained. Therefore, we end up with two different choices of the parameter values. One is that both sides of the model are set to be at the Directed Percolation critical point ($T = 4.738$ with 12 neighboring sites). The other one is to have on both sides the values of the Voter critical point ($T = 1.758$ with 4 neighboring sites). Similar to the LD model, in our microscopic model some system parameter values can be chosen so that different halves of the system undergo the DP or the Voter phase transitions. In the last chapter (Chapter 9) of the first part of this thesis, we study the dynamical behavior of the interface width for these three scenarios.



(a)



(b)

Figure 1.3: Configuration of two absorbing states with a shared interface where both sides of the lattice are at the DP critical point. (a) Initial condition at $t=0$. (b) Evolution of the interface after 2000 Monte Carlo steps.

Chapter 2

Models And Quantities

This chapter, with some changes, is adapted from our publication:

A. Azizi, J. Stidham, and M. Pleimling, “Dynamic critical properties of non-equilibrium Potts models with absorbing states”, J. Stat. Mech. (2018), 103208. “© SISSA Medialab Srl. Reproduced with permission. All rights reserved.”

In a series of papers [9, 15–17] Lipowski and Droz introduced a non-equilibrium model with q absorbing states which we call the LD model. As described briefly in the introduction, the model is in fact very similar to the Metropolis algorithm for the standard q -state equilibrium Potts model, but with one modification in the update scheme that yields q absorbing states, making this a non-equilibrium model. We consider the model on the square lattice with side length L where a lattice point i is characterized by a variable (Potts spin) σ_i that takes on the values $\sigma_i = 0, 1, \dots, q - 1$. The energy is given by the expression

$$E = -\frac{1}{2} \sum_{i=1}^{L^2} \sum_{j \in N_i} \delta(\sigma_i, \sigma_j) \quad (2.1)$$

where $\delta(\sigma_i, \sigma_j)$ is the Kronecker delta function, with $\delta(\sigma_i, \sigma_j) = 1$ if $\sigma_i = \sigma_j$ and zero otherwise. The sum in (2.1) is over a given neighborhood N_i around site i . We consider two different cases: the neighborhood is composed of only the four nearest neighbors or the neighborhood is formed by the twelve nearest neighbor spins.

When simulating the equilibrium model using the Metropolis update scheme, one randomly chooses a site i and then randomly selects a possible new value for the spin variable σ_i . This change is accepted with the probability $\min[1, e^{-\Delta E/T}]$ where T is the temperature of the system and ΔE is the change of energy associated with this change of the value of σ_i . The LD model has the same probability than the equilibrium model with the exception of the situation where all spins in N_i have the same value as σ_i . In that case the spin σ_i is not allowed to be updated. It is this modification that breaks detailed balance and results in a phase with q absorbing states at low temperatures [20].

In their work Lipowski and co-workers provided some information regarding the steady-state properties of the LD model on the square lattice. For $q = 2$ and only nearest-neighbor interactions they found that at $T \approx 1.7585$ the system undergoes a continuous phase transition from a high temperature paramagnetic phase to a low temperature phase with a double symmetric absorbing state [9,15]. This phase transition, at which a symmetry breaking transition between an ordered and a disordered phase takes place at the same time as an absorbing phase transition separating an active phase from an absorbing one, belongs to the Generalized Voter universality class [10–12,19], a non-equilibrium critical point at which the order parameter critical exponent $\beta = 0$.¹ For $q = 3$ and the same neighborhood a discontinuous phase transition, as witnessed by the discontinuous behavior of the steady-state density of active sites, is encountered at the temperature $T \approx 1.237$ [15]. In [9] the authors observed that for $q = 2$ the voter critical point is split into two different phase transitions when extending the neighborhood to the twelve nearest neighbors, namely a symmetry breaking phase transition belonging to the Ising universality class at high temperature, followed at a lower temperature by an absorbing phase transition belonging to the Directed Percolation universality class. This splitting of the voter critical point has been described later through Langevin equations [11] and has also been observed in other microscopic models [13,21].

The second model we consider is a generalized two-state voter model discussed by Krause, Böttcher, and Bornholdt (KBB model in the following) [18]. Introducing the number of agreeing neighbors,

$$u_i = \sum_{j \in N_i} \delta(\sigma_i, \sigma_j) , \quad (2.2)$$

i.e. the numbers of two-state variables σ_j in the neighborhood N_i of site i that have the same value as σ_i , this model is defined through the transition probabilities $p_{(u) \rightarrow (n_i - u)}$ for changing the value of the variable σ_i from 0 to 1 or from 1 to 0 in case u of the n_i neighbors in the neighborhood N_i have the same value than σ_i . For the neighborhood that contains the nearest neighbors only (which is the case considered in the following), $n_i = 4$ and the transition probabilities for the square lattice are given by [18]

$$\begin{aligned} p_{(4) \rightarrow (0)} &= 0, & p_{(0) \rightarrow (4)} &= 1 \\ p_{(3) \rightarrow (1)} &= \frac{1}{1 + \exp(2/T)}, & p_{(1) \rightarrow (3)} &= 1 - p_{(3) \rightarrow (1)} \\ p_{(2) \rightarrow (2)} &= 1/2 \end{aligned} \quad (2.3)$$

At the temperature $T = 2/\ln(3) \approx 1.8205$ [20], these probabilities are identical to those of the linear voter model where the new state of the variable σ_i is provided by one of its four neighbors chosen randomly: $p_{(4) \rightarrow (0)} = 0$, $p_{(3) \rightarrow (1)} = 1/4$, $p_{(2) \rightarrow (2)} = 1/2$, $p_{(1) \rightarrow (3)} = 3/4$, and $p_{(0) \rightarrow (4)} = 1$. The phase transition taking place at this temperature, which separates a disordered high temperature phase from a low temperature phase with two symmetric absorbing states, is therefore by construction in the Generalized Voter universality class.

¹The order-disorder phase transition encountered in models with Z_2 symmetry and two absorbing states belongs to the Generalized Voter universality class.

For comparison and for checks we also simulated the standard equilibrium Potts model with various numbers of states q larger than 4.

While our study focuses on dynamic quantities in order to elucidate the relaxation processes encountered at the different phase transitions, we also computed some standard magnetic quantities used to locate the phase transition temperatures. Defining the quantities

$$M_n = \frac{1}{(q-1)^n} \overline{\left(\frac{q}{N} N_m - 1\right)^n} \quad (2.4)$$

where q is the number of values each variable σ can take on, $N = L^2$ is the number of lattice sites, and N_m is the number of majority spins, we obtain the magnetization M , the susceptibility χ and the Binder cumulant U through the expressions

$$M = M_1 \quad (2.5)$$

$$\chi = \frac{1}{T} (M_2 - M_1^2) \quad (2.6)$$

$$U = 1 - \frac{M_4}{3M_2^2} \quad (2.7)$$

For these static quantities, $\overline{\cdots}$ indicates a time average as well as an ensemble average over runs with different initial conditions and different random number sequences.

Our dynamic quantities fall into two categories, those used for the investigation of relaxation and aging processes at symmetry-breaking transitions and those studied when characterizing absorbing phase transitions. The former one include the time-dependent magnetization $M(t)$, given by expression (2.5) with only an ensemble average over initial conditions and random number sequences, the two-time autocorrelation function

$$C(t, s) = \frac{1}{(q-1)} \left(\frac{q}{N} \sum_{i=1}^N \langle \delta(\sigma_i(t), \sigma_i(s)) \rangle - 1 \right), \quad (2.8)$$

as well as the two-time autoresponse function [22]

$$\rho(t, s) = \frac{T}{q-1} \left(\frac{q}{N} \sum_{i=1}^N \langle \delta(\sigma_i(t), r_i) \rangle - 1 \right), \quad (2.9)$$

where $\langle \dots \rangle$ indicates an average over different realizations of the noise [23]. Starting from a disordered initial state, we follow the standard protocol for calculating this response by applying for the first s time steps a spatially random field with amplitude h . The random field at site i is given by the expression $h_i = h r_i$ where the quenched random variable r_i indicates the direction of the field and takes on one of the values $0, 1, \dots, q-1$ [22]. After s time steps the field is removed and the relaxation of the system to the steady state is monitored for times $t > s$ with the help of the two-time autoresponse function (2.9). In

order to remain in the linear response regime, we choose the small value $h = 0.05$ for the field amplitude.

In order to probe the absorbing phase transition we measure the density of active sites $\rho_a(t)$, i.e. the fraction of sites for which at time t not all variables σ in the neighborhood N_i are in the same state than σ_i , the time-dependent survival probability $P_s(t)$ as well as the number of flipped spins $N_f(t)$. For the latter two quantities we prepare systems in one of the absorbing states, say state 0, and change the value of the variables σ in a small area (composed of 12 spins), setting for the $q = 2$ case $\sigma_j = 1$ for every site j in that area, whereas for the $q = 3$ case we randomly assign the value 1 or 2 to σ_j . The survival probability is then given by the fraction of systems that at time t still contain variables with values different from 0, whereas the number of flipped spins $N_f(t)$ is the average number of sites k characterized at time t by a value $\sigma_k \neq 0$.

As usual, we measure time in Monte Carlo steps, with one Monte Carlo step corresponding to L^2 proposed updates.

Chapter 3

The Two-State Model

This chapter, with some changes, is adapted from our publication:

A. Azizi, J. Stidham, and M. Pleimling, “Dynamic critical properties of non-equilibrium Potts models with absorbing states”, J. Stat. Mech. (2018), 103208. “© SISSA Medialab Srl. Reproduced with permission. All rights reserved.”

We first investigate in the following relaxation processes and aging phenomena for the two-state systems before discussing in the next Section the situation where variables can take on three different values. Having two-state variables interacting with their four nearest neighbors yields a phase transition that belongs to the Generalized Voter universality class. Consequently, our main interest in this Section is on the dynamic properties far from stationarity at a voter critical point. This is followed by a quick discussion of the situation for the case where the interaction range is extended to the twelve nearest neighbors.

3.1 Interactions with nearest neighbors only

As already mentioned in [15] and [18], the phase transitions encountered in the LD and KBB models with only nearest-neighbor interactions belong to the Generalized Voter universality class. In a series of papers de Oliveira and co-workers have investigated the dynamic properties far from the steady state for the non-equilibrium linear Glauber model [24, 25] and a non-equilibrium linear q -state model [26] that both exhibit a phase transition belonging to the Generalized Voter universality class. The linear character of these models allows for an analytical treatment, even away from stationarity. For the two-time autocorrelation function, Hase *et al.* obtain the expression

$$C(t, s) \sim \frac{s}{(t - s) \ln s} \quad (3.1)$$

when starting from a disordered initial state [26]. Cast in the standard aging scaling form [23],

$$C(t, s) = s^{-b} f_C(t/s) , \quad (3.2)$$

with the scaling function $f_C(t/s) \sim (t/s)^{-\lambda/z}$ for $t/s \gg 1$, where λ is the autocorrelation exponent and z is the dynamic exponent, this expression formally yields the exponents $b = 0$ (with logarithmic correction) and $\lambda/z = 1$.

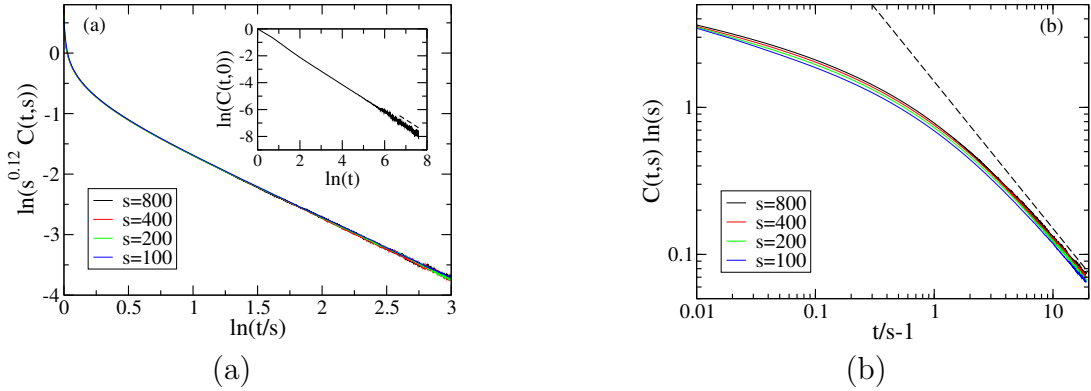


Figure 3.1: (a) At the critical temperature $T_c = 2/\ln(3)$ the two-time autocorrelation function $C(t, s)$ for the KBB model shows a data collapse when assuming the standard aging scaling behavior (3.2) with $b = 0.120(5)$. Inset: $C(t, 0)$ exhibits a power-law tail with an exponent $\lambda = 1.00(2)$ (as indicated by the dashed line). (b) The scaling proposed in [26, 27] for the two-time autocorrelation function does not yield a data collapse nor is the expected scaling function (3.1) (dashed line) recovered. The data result from averaging over 8000 realizations for systems with linear size $L = 800$.

In Figure 3.1 we probe the scaling properties of the two-time autocorrelation function (2.8) obtained for the KBB model. Panel (a) reveals that a perfect data collapse is obtained when assuming the standard aging scaling (3.2) with the scaling exponent $b = 0.120(5)$. The long-time behavior of the autocorrelation is governed by a power-law decay with an exponent $\lambda/z = 1.00(2)$, see also the inset in the figure displaying $C(t, 0)$. The scaling form (3.1) proposed in [26] does not result in a data collapse, see Figure 3.1b. In addition, the theoretical scaling function (dashed line in the figure) does not at all describe the numerical data. In Ref. [27] the aging properties at the voter critical point were briefly discussed through a study of the autocorrelation in a non-equilibrium symmetric three-state model at its voter point. These authors did not check for the possibility of a standard scaling behavior, but instead exclusively plotted the data in the way shown in Figure 3.1b. However, comparing directly the two scenarios, as done in the two panels in Figure 3.1, reveals the superiority of the standard aging scaling and the discrepancy between the theoretical prediction and the numerical data.

In Figure 3.2 we show that the LD model displays the same scaling behavior as the KBB

model, with the same two exponents $b = 0.120(5)$ and $\lambda/z = 1.00(2)$. This simple aging scaling is encountered for both the two-time autocorrelation (2.8) and the two-time autoresponse function (2.9). We note that the autoresponse in the KBB model also displays the same scaling behavior. These observations suggest a common, universal, behavior in the aging regime of models (both linear and non-linear) belonging to the Generalized Voter universality class. This universal behavior, however, is not identical to the theoretical expression (3.1) provided in [26].

Whereas the linear model (we remind the reader that the KBB model at its critical point has exactly the same transition probabilities as the original voter model) and the non-linear model display the same aging scaling for the two-time quantities, differences show up in the time-evolution of the magnetization when starting with a magnetized (but disordered) initial state. In Figure 3.3 we show $M(t)$ when preparing the system in a state with magnetization $M(0) = 0.5$. Whereas for the KBB model the magnetization is independent of time, $M(t) = 0.5$, in agreement with the analytical result obtained from the non-equilibrium linear q -state model [26], for the LD model the magnetization displays a quick initial drop before slowly approaching a limit value that differs from the initial magnetization. Our data for LD therefore agree with the earlier observation of non-conservation of the magnetization for non-linear models at their voter critical point [28].

3.2 Interactions with up to third nearest neighbors

As shown in [9], extending for the LD model the interaction neighborhood N_i of a variable σ_i to twelve sites, i.e. considering interactions up to third nearest neighbors, splits the voter critical point into an absorbing phase transition at lower temperature and a symmetry-breaking phase transition at higher temperature. Based on the symmetries involved, the absorbing phase transition is expected to belong to the (2+1) Directed Percolation universality class [3], whereas the symmetry-breaking phase transition should belong to the universality class of the two-dimensional Ising model [29]. In [9] some numerical data in support of these expectations were provided.

As already mentioned, this splitting of the voter critical point is encountered in a variety of systems, see [11, 13, 14, 21] for some examples. Whereas it is easy to show numerically the Directed Percolation nature of the absorbing phase transition, proving the Ising character of the symmetry-breaking phase transition is often challenging, due to the closeness of the two critical points. These issues are well illustrated by the controversy around the phase transitions encountered in a monomer-dimer model on a square lattice [21, 30–32].

The LD model with up to third-nearest-neighbor interactions discussed in this section is also plagued by the proximity of the two phase transitions. Whereas the symmetry-breaking phase transition takes place at $T_I = 4.812$, the absorbing phase transition is encountered at the slightly lower temperature of $T_{DP} = 4.738$ [9]. A direct consequence of this fact is the

very narrow critical region which makes the use of standard finite-size scaling approaches challenging.

In our study we focused on the question whether the measurement of dynamic quantities far from stationarity provides a viable alternative to the standard approach that aims at determining critical exponents of static quantities. Our results are mixed. Whereas the measurements of the initial-slip exponent [33], obtained from the increase of the magnetization when starting from a disordered initial state with a small magnetization, or of the exponent governing the long-time power-law decay of the autocorrelation $C(t, 0)$ (see inset in Figure 3.4a) readily yield the values expected for a critical point belonging to the two-dimensional Ising universality class, the verification of the aging scaling behavior of two-time quantities is hampered by finite-time effects that necessitate larger waiting times than usual before a clean scaling behavior with the expected Ising exponent emerges [34]. This is illustrated in Figure 3.4 for both the autocorrelation (2.8) and the autoresponse function (2.9). Plugging in the known value $2\beta/\nu z \approx 0.115$ (where β and ν are the usual static exponents and z is the critical dynamic exponent) for the scaling exponent, we remark that the longest waiting times start to exhibit an acceptable data collapse. On the other hand, as shown in the inset of Figure 3.4a, the autocorrelation function in the long-time limit does decay with the correct exponent $\lambda/z \approx 0.733$ (dashed line) [35].

Based on our data we conclude that the investigation of non-equilibrium quantities provides an alternative way of determining the universality class of a symmetry-breaking phase transition located in close vicinity to an absorbing phase transition, with some measurements, like that of the initial-slip behavior of the magnetization or the long-time decay of the autocorrelation function $C(t, 0)$, being more useful than others.

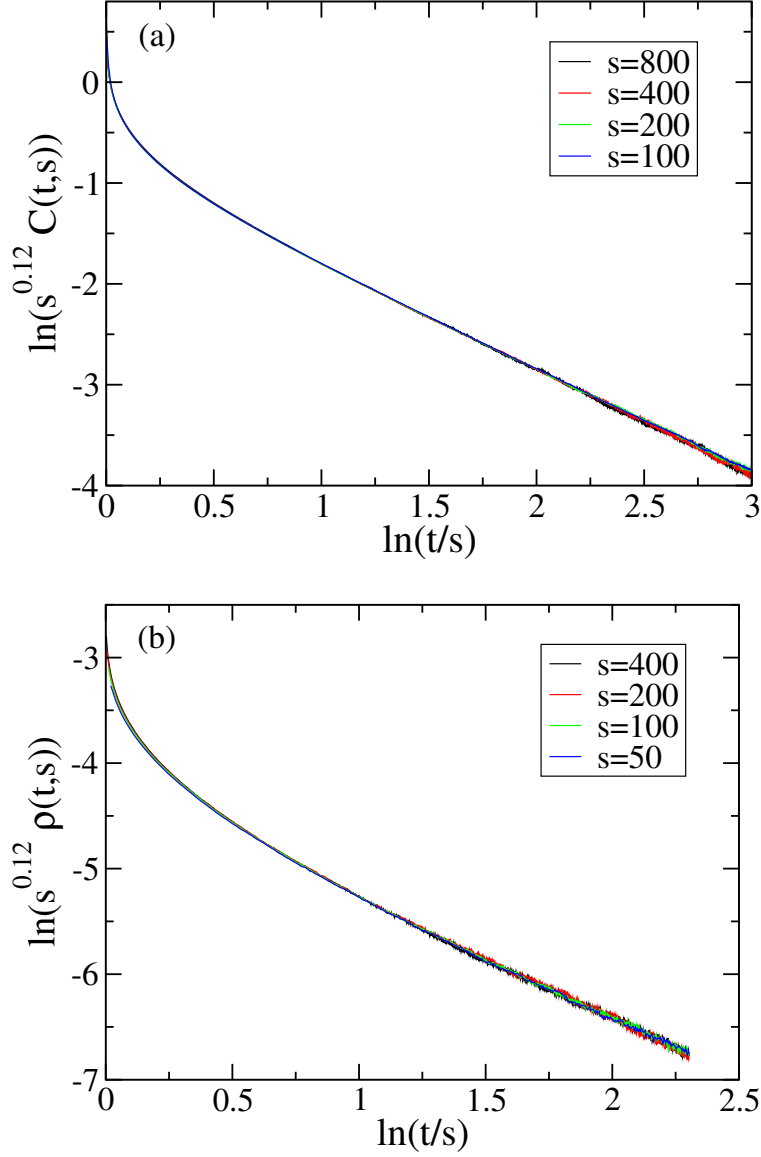


Figure 3.2: (a) At the critical temperature $T_V = 1.7585$ the two-time autocorrelation function $C(t, s)$ for the LD model shows the same scaling behavior as the critical KBB model, with the same exponents. The data result from averaging over 8000 realizations for systems with linear size $L = 800$. (b) A data collapse is also encountered for the autoresponse function $\rho(t, s)$ with the same value for the scaling exponent. These simulations were done for systems with 400×400 sites and the amplitude $h = 0.05$ for the spatially random magnetic field. The data result from averaging over at least 250000 different runs.

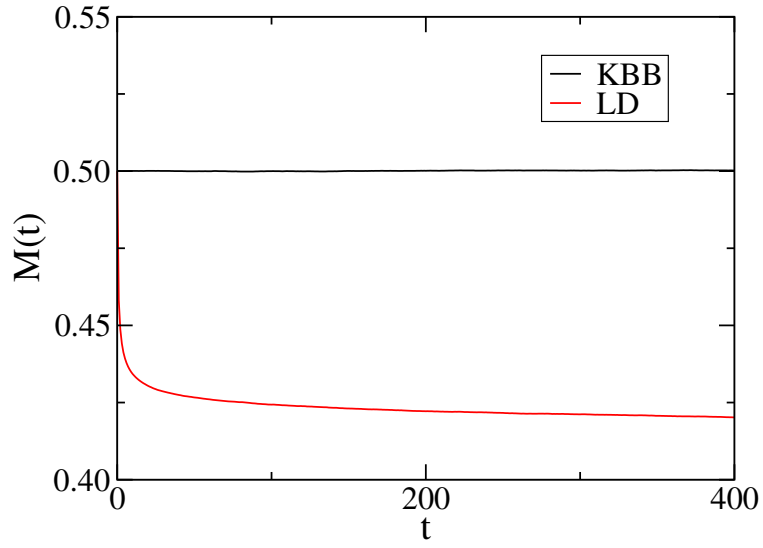


Figure 3.3: Time evolution of the magnetization $M(t)$ for the critical KBB and LD models with nearest-neighbor couplings when preparing the system in a state with initial magnetization $M(0) = 0.5$. Systems of linear size $L = 400$ have been simulated, and the data result from averaging over 25000 independent runs.

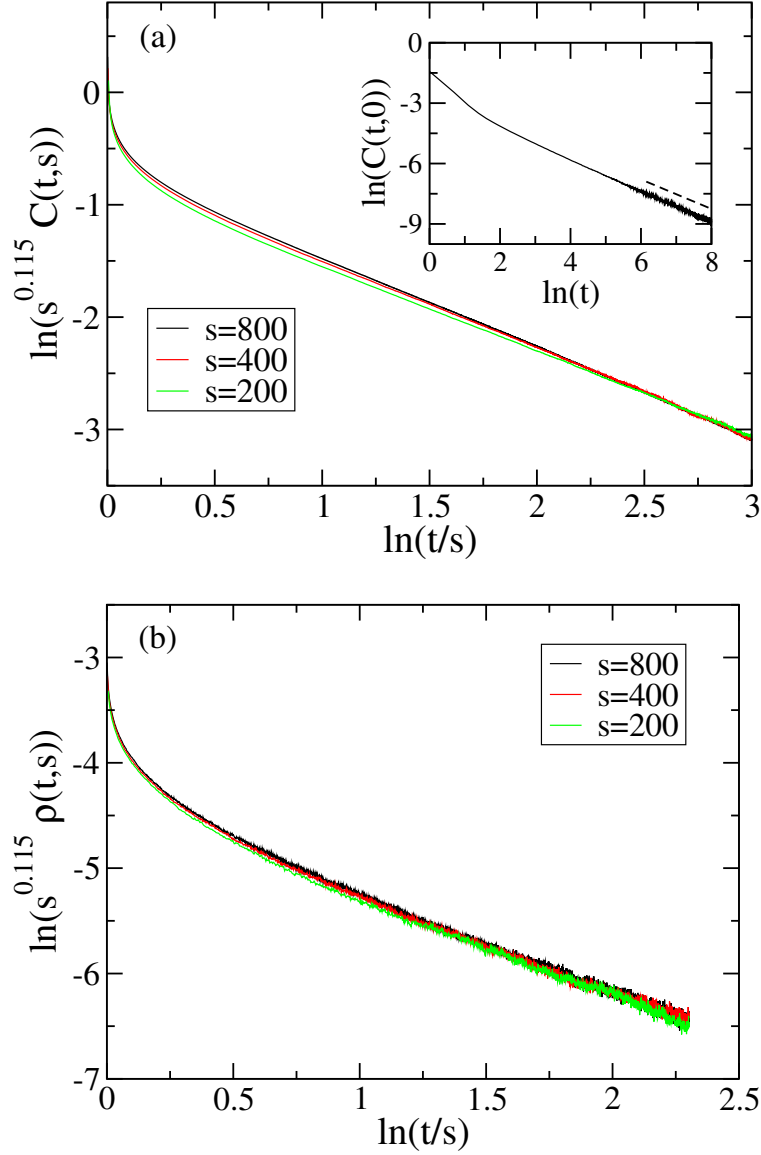


Figure 3.4: (a) Scaling of the two-time autocorrelation and (b) of the two-time autoreponse function at the Ising critical point, with $T_I = 4.812$, of the $q = 2$ LD model with twelve nearest neighbors. The inset in (a) shows that the long-time decay of $C(t, 0)$ is in agreement with the known exponent $\lambda/z \approx 0.733$ (indicated by the dashed line). The autocorrelation data result from averaging over at least 10000 realizations, whereas for the autoreponse function at least 400000 independent runs have been made. In (b) the amplitude of the magnetic field has been set to $h = 0.05$. For both quantities systems of linear extent $L = 400$ have been simulated.

Chapter 4

The Three-State Model

This chapter, with some changes, is adapted from our publication:

A. Azizi, J. Stidham, and M. Pleimling, “Dynamic critical properties of non-equilibrium Potts models with absorbing states”, J. Stat. Mech. (2018), 103208. “© SISSA Medialab Srl. Reproduced with permission. All rights reserved”

4.1 Interactions with nearest neighbors only

We now turn to the three-state LD model with the same two interaction neighborhoods. Our goal is again to characterize the different phase transitions through dynamic quantities measured far from stationary.

As already noted in [15], the two-dimensional LD model with $q = 3$ undergoes a discontinuous phase transition at $T \approx 1.273$. As shown in Figure 4.1 the discontinuity is encountered for both the symmetry-breaking order parameter (the steady-state magnetization M) and the order parameter of the absorbing phase transition (the steady-state density of active sites ρ_a). This is consistent with the observation [15] that the time-dependent survival probability displays a systematic bending without clear power-law regime.

In Figure 4.2 we plot the autocorrelation function at this discontinuous phase transition as a function of t/s . At first look the autocorrelation function seems to behave in very similar ways for the different waiting times. Close inspection of the long time behavior, however, indicates that a fit to a power-law for the larger values of t/s yields an effective exponent whose value increases with the waiting time. Dynamical scaling therefore can not be expected for this system.

In order to check whether this behavior is also observed at an equilibrium first-order phase transition, we extended our study to the equilibrium two-dimensional q -state Potts model

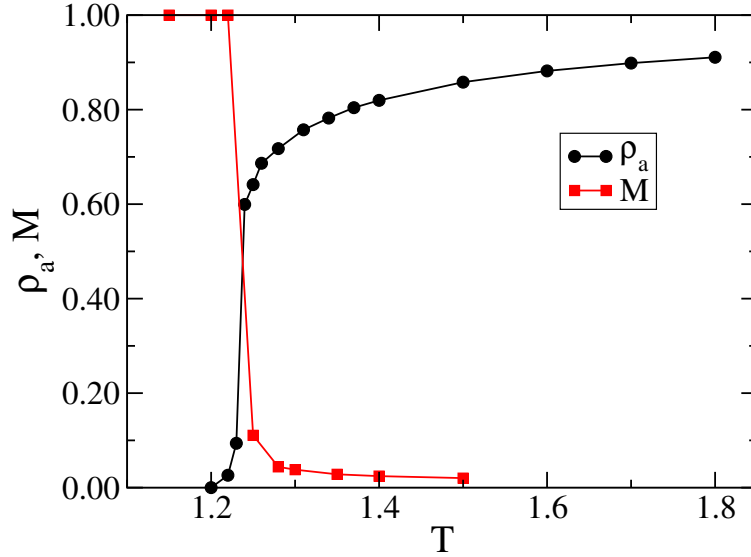


Figure 4.1: The steady-state density of active sites ρ_a and the steady-state magnetization M as a function of temperature for the LD model with $q = 3$ and only nearest-neighbor interactions. Both order parameters display a jump at the same temperature. The data, obtained for a system with linear extent $L = 400$, result from a time average (after having discarded the early time data) and an ensemble average over 100 independent runs. Error bars are smaller than the sizes of the symbols.

with $q > 4$ and computed the two-time autocorrelation at the phase transition temperature $T = [\ln(1 + \sqrt{q})]^{-1}$. In all cases with $q > 4$ the transition is known to be of first order [36]. Figure 4.3 shows a remarkable difference when comparing our data for $q = 5$ with those for $q = 7$ and $q = 8$. Whereas for $q = 7$ and $q = 8$ we have a behavior similar to that of the $q = 3$ LD model, characterized by the absence of dynamical scaling, for $q = 5$ we observe a perfect aging scaling behavior with exponents $b = 0.20(1)$ and $\lambda/z = 1.17(2)$. In order to understand these very different results for different first-order phase transitions, we remark that for $q = 5$ we are dealing with a weak first-order phase transition characterized by a pseudocritical behavior [37]. Indeed, from the exact expression for the correlation length at the transition point [38] one obtains that for $q = 5$ the stationary correlation length is $\xi = 2512.25$. As this is much larger than our system size, the observed relaxation processes do not differ from those encountered at a critical point. It is then not surprising that we observe dynamical scaling even though the system has a discontinuous transition. For the $q = 7$ respectively $q = 8$ model, the value of the stationary correlation length is $\xi = 48.10$ respectively 23.88 , i.e. much smaller than our system size. In this case, the dynamical correlation length can only grow up to the value of ξ , which results in the absence of dynamical scaling and leads to a behavior similar to that observed in Figure 4.3 for the $q = 3$ LD model.

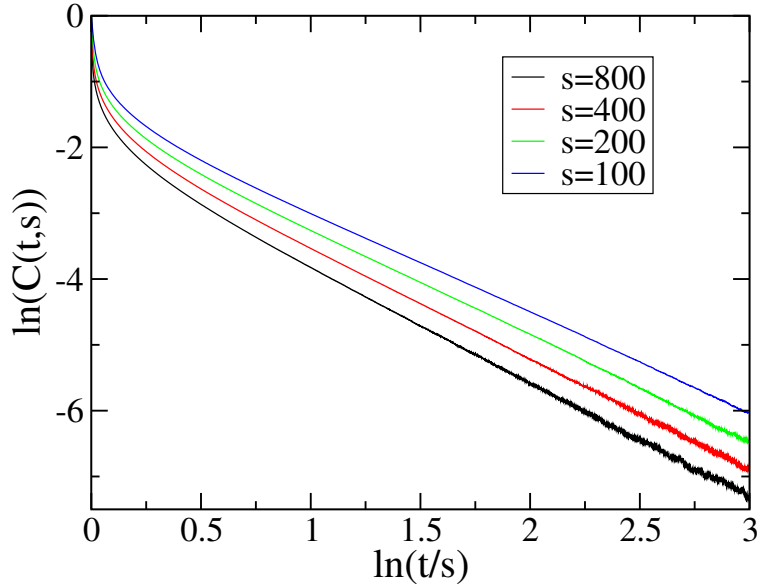


Figure 4.2: Two-time autocorrelation function for the three-state LD model with only nearest-neighbor interactions at the temperature $T = 1.237$ of the discontinuous phase transition. The autocorrelation is plotted as a function of the ratio t/s . The linear extent of the system is $L = 800$, and 10000 independent runs have been performed.

4.2 Interactions with up to third nearest neighbors

Here we discuss how the properties of the $q = 3$ LD model change when we replace the interaction neighborhood N_i with four sites by the neighborhood with twelve sites. The data we present in the following indicate that the single discontinuous transition encountered for only nearest-neighbor interactions is replaced by two continuous phase transitions when extending the interaction range to third nearest neighbors: a symmetry-breaking phase transition belonging to the universality class of the two-dimensional $q = 3$ Potts model and an absorbing phase transition belonging to the Directed Percolation universality class.

A first indication of the appearance of two phase transitions taking place at different temperatures can be seen in Figure 4.4 where we show for a system of linear extent $L = 400$ the temperature dependence of the steady-state density of active sites ρ_a and the steady-state magnetization M . The two quantities clearly approach zero at two different temperatures, thus exhibiting the same behavior as the $q = 2$ case [9].

In order to better characterize the absorbing phase transition we show in Figure 4.5 the time dependence of the number of flipped spins $N_f(t)$ as well as that of the survival probability $P_s(t)$. As already described in Section 2, we prepare for this measurement the system in one

state, say 0, before assigning to a connected cluster of twelve spins randomly the values 1 or 2. Having prepared the system in this way, we count at each time step the number of variables σ_i with a value different from 0 and derive from this the two quantities shown in Figure 4.5.

We first note that both quantities display an algebraic behavior at the temperature $T = 3.754$, whereas for temperatures slightly above or below that value clear deviations from the power-law behavior are observed. This allows us to determine the temperature of the absorbing phase transition to be $T_a = 3.754(1)$. Furthermore, the algebraic growth respectively decrease of $N_f(t) \sim t^\eta$ respectively $P_s(t) \sim t^{-\delta}$ yields the exponent $\eta = 0.238(10)$ respectively $\delta = 0.455(6)$. Comparing these values with the known values $\eta = 0.230$ and $\delta = 0.451$ for Directed Percolation [39, 40], we conclude that the absorbing phase transition encountered in the $q = 3$ LD model with $N_i = 12$ indeed belongs to the Directed Percolation universality class.

The critical temperature for the symmetry-breaking transition can be reliably determined by investigating the Binder cumulant for different system sizes, as shown in Figure 4.6. From the intersections of the different data sets we obtain $T_c = 3.813(1)$, a value that is indeed larger than the temperature T_a of the absorbing phase transition. The universality class of this symmetry-breaking phase transition can be probed through finite-size scaling. As an example we show in the inset of Figure 4.6 the scaling plot for the magnetization where we plot $ML^{\beta/\nu}$ as a function of $\varepsilon L^{1/\nu}$, with the reduced temperature $\varepsilon = \frac{T_c - T}{T_c}$. Inserting as T_c the value we obtain from the crossing of the Binder cumulant data as well as the known exponents $\beta = 1/9$ and $\nu = 5/6$ for the two-dimensional $q = 3$ Potts universality class results in the data collapse shown in the inset.

Finally, we also investigated the two-time autocorrelation and autoresponse functions for this case. As for the $q = 2$ LD model with up to third-nearest-neighbor interactions, see Figure 3.4, these quantities suffer from sizeable finite-time corrections, due to the closeness of the symmetry-breaking and the absorbing phase transitions. As a consequence one also has to go for $q = 3$ to large waiting times in order to encounter a clean data collapse with the aging exponents of the critical equilibrium three-state Potts model.

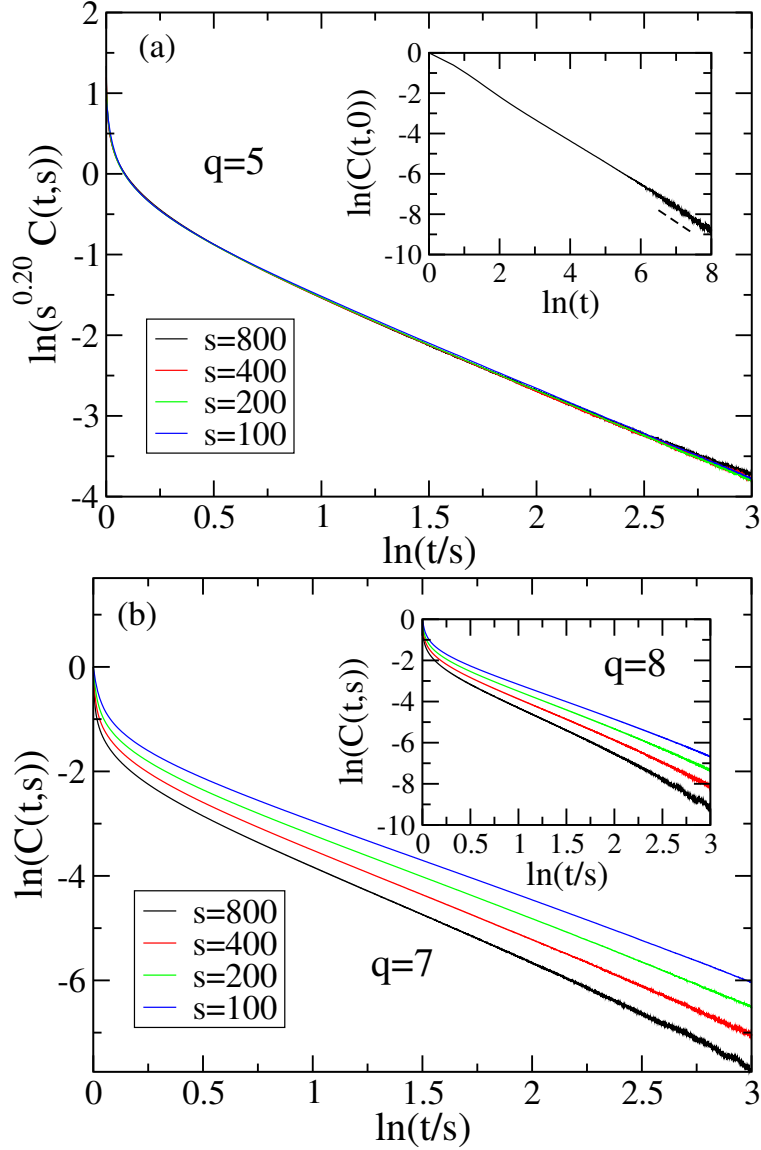


Figure 4.3: Scaling of the autocorrelation $C(t, s)$ of the equilibrium Potts model with (a) $q = 5$ states and (b) $q = 7$ (inset $q = 8$) states at their discontinuous phase transition temperature $T = [\ln(1 + \sqrt{q})]^{-1}$. A perfect aging scaling prevails at the weak first-order transition for the $q = 5$ case, whereas at the stronger discontinuous transitions of the $q = 7$ and $q = 8$ cases a simple aging data collapse can not be achieved. The dashed line in the inset of (a) indicates that $C(t, 0)$ decays for $q = 5$ algebraically with an exponent ~ 1.17 . The system size is $L = 800$, and the data result from averaging over at least 9000 independent runs.

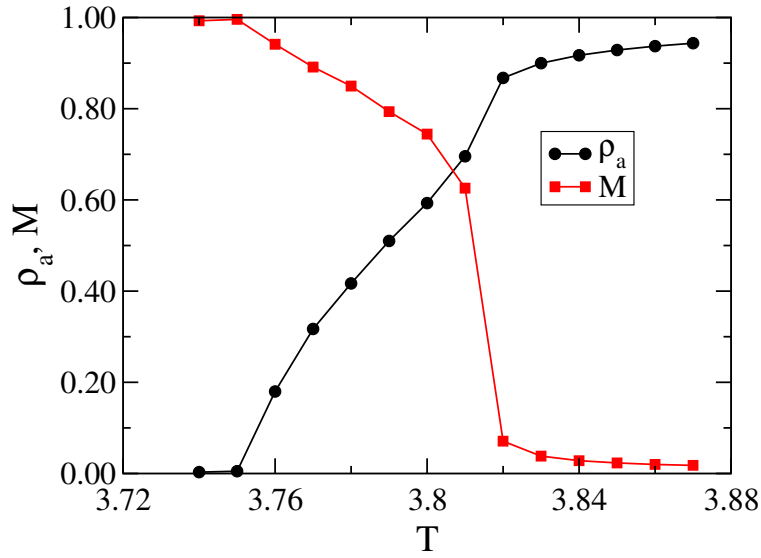


Figure 4.4: The steady-state density of active sites ρ_a and the steady-state magnetization M as a function of temperature for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. The behavior of the two order parameters indicate the presence of two phase transitions that take place at different temperatures. The data, obtained for a system with linear extent $L = 400$, result from a time average (after having discarded the early time data) and an ensemble average over 150 independent runs. Error bars are of the order of the symbol sizes.

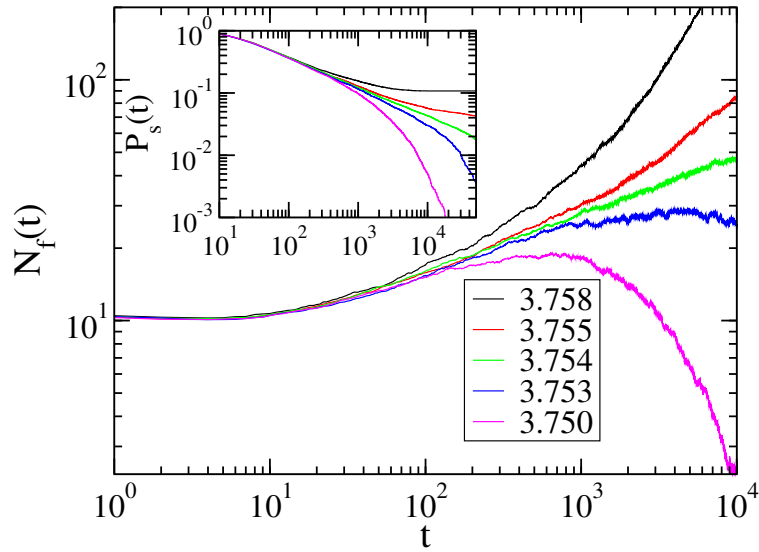


Figure 4.5: Number of flipped spins N_f (main figure) and survival probability P_s (inset) as a function of time for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. Both quantities display an algebraic behavior at the absorbing phase transition temperature $T_a = 3.754$. The measured exponents $\eta = 0.238(10)$ for N_f and $\delta = 0.455(6)$ for P_s agree with the values of the Directed Percolation universality class [39,40]. The data, obtained for a system with $L = 400$, result from averaging over at least 10000 independent runs.

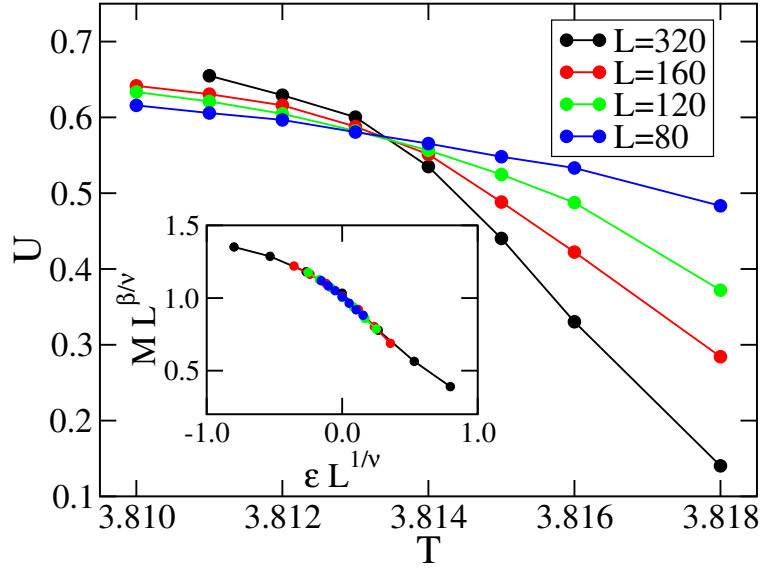


Figure 4.6: Main figure: Binder cumulant for the LD model with $q = 3$ and up to third-nearest-neighbor interactions. Systems of different sizes have been simulated. Inset: finite-size scaling of the magnetization M when using the critical exponents of the $q = 3$ equilibrium Potts model. Error bars are smaller than the symbol sizes. The data result from averaging over 200 independent runs.

Chapter 5

Conclusion I

This chapter, with some changes, is adapted from our publication:

A. Azizi, J. Stidham, and M. Pleimling, “Dynamic critical properties of non-equilibrium Potts models with absorbing states”, J. Stat. Mech. (2018), 103208. “© SISSA Medialab Srl. Reproduced with permission. All rights reserved”

The aim of previous chapters has been to further elucidate the properties of a family of non-equilibrium Potts models with absorbing states that exhibit various scenarios when changing the number of states q of the spins or the range of the spin-spin interaction [9, 15–17]. Most notably, in the presence of only nearest-neighbor interactions the model exhibits for the studied values of q a temperature at which a symmetry-breaking and an absorbing phase transition coincide, whereas for larger interaction ranges this common transition is split into a symmetry-breaking transition at higher temperature and an absorbing phase transition at lower temperature. In our work we have characterized the different phase transitions through standard steady-state quantities as well as through various dynamic quantities, including two-time correlation and response functions. In order to support some of our conclusions we have also presented results for other models such as the Generalized Voter model introduced by Krause, Böttcher, and Bornholdt [18] and the equilibrium two-dimensional Potts model with $q > 4$.

Our main results are as follows:

- LD model with $q = 2$ and only nearest-neighbor interactions.
In this case the system exhibits a single phase transition with coinciding symmetry-breaking and absorbing transitions. This phase transition is known to belong to the Generalized Voter universality class. Both for the LD model and the KBB model, whose probabilities for spin flips at the critical point are given by the transition probabilities of the linear voter model, we observe simple aging scaling for the two-time autocorrelation and autoresponse functions. The scaling exponent and the scaling function are found to differ from an analytical expression proposed in [26].

- LD model with $q = 2$ and interactions with up to third nearest neighbors.
 Extending the range of spin-spin interactions results in the splitting of the voter critical point into a symmetry-breaking phase transition, expected to belong to the equilibrium Ising universality class, and an absorbing phase transition that take place at different temperatures. These two transition temperatures are very close, which can make it challenging to determine numerically the universality class of the symmetry-breaking transition when using steady-state quantities. We discuss as an alternate approach the investigation of aging quantities and observe strong finite-time corrections. As a result long waiting times are needed before the expected aging scaling with Ising exponents is encountered.
- LD model with $q = 3$ and only nearest-neighbor interactions.
 This case is characterized by a discontinuous phase transition at which both the symmetry-breaking order parameter as well as the absorbing order parameter exhibit jumps. Two-time quantities display a behavior that is also observed at the first-order phase transition of the equilibrium Potts model with larger values of q (as verified for $q = 7$ and $q = 8$). A different, namely a simple aging scaling, behavior is encountered for the $q = 5$ equilibrium Potts model. These differences are explained by the equilibrium correlation length which for $q = 5$ is large compared to the system size, whereas in the other studied cases it is smaller than the system size.
- LD model with $q = 3$ and interactions with up to third nearest neighbors.
 As for the $q = 2$ case with up to third-nearest-neighbor interactions we also find for $q = 3$ two different phase transitions taking place at different, albeit close, temperatures. The absorbing phase transition is shown to belong to the Directed Percolation universality class, whereas the symmetry-breaking phase transition at higher temperature has the same critical exponents as the equilibrium $q = 3$ Potts model in two space dimensions.

Chapter 6

A Microscopic Model With Tunable Parameters

This chapter, with some changes, is adapted from our work submitted to Phys. Rev. E.:

A. Azizi, and M. Pleimling, “Critical phenomena in presence of symmetric absorbing states: a microscopic model with tunable parameters”.

Studies of phase transitions separating active from absorbing states have contributed tremendously to our understanding of non-equilibrium physics [4, 41]. Many systems have a unique absorbing state that can be accessed through an active-absorbing phase transition belonging to the directed percolation universality class, but this is not the only possible scenario. Examples of systems with two symmetric absorbing states (Z_2 symmetry) are found in numerous fields, ranging from opinion dynamics [42–44] to linguistics [45–47] and from catalysis [48, 49] to species competition [50, 51] and population ecology [52, 53]. Often models discussed in this context are voter models where the current state of a discrete variable (which often is an Ising spin that takes on the values ± 1) depends on the distribution of states of other variables in a specified neighborhood.

A unique feature of the voter model is the absence of bulk noise which yields coarsening without surface tension [10, 54]. The resulting slow dynamics is revealed by the fact that in two space dimensions the interface density ρ decays logarithmically with time, $\rho \sim 1/\ln t$ [48, 49, 55]. Remarkable features are also revealed at the active-absorbing phase transition in systems with two symmetric absorbing states. Indeed, in critical Z_2 symmetric models without bulk noise both the interface density and the magnetization (defined as the sum over all spins) display a logarithmic decay with time for the case that a symmetry-breaking transition between an ordered and a disordered phase takes place at the same time as an absorbing phase transition separating an active from an absorbing phase. Quantities in systems belonging to this *Generalized Voter universality* class [11] are characterized by a set of critical exponents that differs from the corresponding set of the directed percolation

universality class [4].

Simulations of a kinetic Ising model in two dimensions provided early indications that a second scenario is possible for systems with two symmetric absorbing states [9]: the splitting of the voter critical point into two separate transitions, namely an order-disorder transition belonging to the universality class of the two-dimensional Ising model and an absorbing phase transition belonging to the directed percolation universality class. Al Hammal *et al.* [11] proposed a unifying picture of a phase diagram that contains three critical lines, with the directed percolation and Ising lines merging into a Generalized Voter line. This was achieved through the investigation of the Langevin equation. For $a = b = 0$ this Langevin equation is identical to the Langevin equation for the classical voter model [56,57]. Numerical simulations of Equation (1.3) reveal that the three critical lines merge at a point that is not identical to the classical voter point $a = b = 0$.

To our knowledge no microscopic model with continuous varying parameters is known that faithfully reproduces the results obtained in the Langevin description. In numerical studies of kinetic Ising models [9,12,58] the different scenarios predicted by the Langevin equation are not found for a fixed range of the spin-spin interactions. Only when modifying the interaction range can one observe a change in scenario, with a unique Generalized Voter phase transition for nearest neighbor interactions and separate directed percolation and Ising phase transitions once the interaction distance exceeds a critical value. Other systems were shown to allow access to some parts of the phase diagram obtained from the Langevin Equation (1.3). Two-dimensional interacting monomer [31] or monomer-dimer models [21] were found to exhibit successive order-disorder and active-absorbing phase transitions belonging to the Ising and directed percolation universality classes. In [19] a reaction-diffusion system was proposed that at the coarse-grained level is described by Equation (1.3), but the physically meaningful range of parameter values does not allow to access the different critical lines. In an interesting work [13] Rodrigues *et al.* discussed a two-dimensional lattice model with freely tunable parameters. By construction their phase diagram contains a point where their model is identical to the classical voter model. They recovered the same three phases as one obtains when simulating the Langevin Equation (1.3). However, in stark contrast to the Langevin description, the model of Rodrigues *et al.* only has two critical lines as a line of discontinuous phase transitions is replacing the Generalized Voter critical line. In addition, the Ising and directed percolation critical lines end exactly at the point where the classical voter model is realized, which is different from what is obtained from the Langevin equation.

In Section 6.1 we present a microscopic model with two continuously varying parameters that exhibits a phase diagram with all the phases and all the critical lines obtained in the Langevin description. Our model is an extension of a model presented by Russell and Blythe [47] that displays a noise-induced dynamical transition between two coarsening regimes, one regime with curvature-driven coarsening (Ising-like coarsening) and one regime with coarsening in absence of surface tension (voter-like coarsening). We show that Equation (1.3) provides the correct stochastic equation of motion for our model. Through extensive numerical simulations we determine the universality classes of the three different critical

lines.

In 2011 Russell and Blythe [47] presented a stochastic microscopic spin model that displays a transition between Ising-like (i.e. curvature driven) coarsening and voter-like coarsening characterized by the absence of surface tension. They showed that the stochastic equation of motion that follows from the microscopic model is given by the Langevin Equation (1.3) with $b = 0$:

$$\frac{\partial \phi}{\partial t} = D \nabla^2 \phi + a \phi (1 - \phi^2) + \sigma \sqrt{1 - \phi^2} \eta \quad (6.1)$$

As we show in the following, this model can be extended in a way that the corresponding Langevin equation is given by Equation (1.3) with b being a tunable parameter that impacts the rate of going from one configuration to the next.

6.1 Model definition

Russell and Blythe consider square systems composed of L^2 cells with N spins in each cell. These classical spins can take on the values ± 1 . A proposed update consists in replacing a randomly selected spin with a copy from a spin taken either from the same cell or from a neighboring cell. The first move is related to the potential term in (6.1), whereas the latter move is realizing diffusion. The copy is taken from one randomly selected spin contained in one randomly selected neighboring cell with probability $q = \frac{h}{N}$ where h is a system parameter that turns out to be related to the diffusion constant D in Equation (6.1). The probability that the copy is taken from an up respectively down spin located in the same cell is $(1 - q)p$ respectively $(1 - q)(1 - p)$ where $p \neq \frac{1}{2}$ introduces a systematic bias. In [47] the choice $p = x_i + \frac{2a}{N}x_i(1 - x_i)(2x_i - 1) = \frac{1}{2}(1 + \phi_i) + \frac{a}{2N}\phi_i(1 - \phi_i^2)$ was made where $x_i = \frac{n_{i\uparrow}}{N}$ is the fraction of up spins in cell i and $\phi_i = \frac{n_{i\uparrow} - n_{i\downarrow}}{N} = 2x_i - 1$ is the local (cell) magnetization density. In addition, $n_{i\uparrow}$ respectively $n_{i\downarrow}$ is the total number of up respectively down spins in the cell and a is a model parameter. It follows from the Fokker-Planck equation in the limit $N \rightarrow \infty$ that (6.1) is the correct stochastic equation of motion for this model.

In order to obtain the full Langevin Equation (1.3) in the limit $N \rightarrow \infty$ the probability p for taking the copy from the local cell has to be modified in an appropriate way. We propose to use the expression

$$p(x_i) = x_i + \frac{2a}{N}x_i(1 - x_i)(2x_i - 1) - \frac{2b}{N}x_i(1 - x_i)(2x_i - 1)^3 \quad (6.2)$$

for this probability as a function of the fraction of up spins in the local cell, or, equivalently,

$$p(\phi_i) = \frac{1}{2}(1 + \phi_i) + \frac{a}{2N}\phi_i(1 - \phi_i^2) - \frac{b}{2N}\phi_i^3(1 - \phi_i^2) \quad (6.3)$$

in terms of the local magnetization density ϕ_i .

The corresponding Langevin equation follows from the Fokker-Planck equation that results from the Kramers-Moyal expansion of the Master equation. Noting that a move changes x_i by the fixed value $\pm\Delta$ with $\Delta = \frac{1}{N}$, the transition probabilities entering the Master equation are

$$w(x_i + \Delta | x_i) = (1 - x_i) \left\{ \frac{h}{4N} \sum_j x_j + \left(1 - \frac{h}{N}\right) p(x_i) \right\} \quad (6.4)$$

$$w(x_i - \Delta | x_i) = x_i \left\{ \frac{h}{4N} \sum_j x_j + \left(1 - \frac{h}{N}\right) p(1 - x_i) \right\} \quad (6.5)$$

where $p(x_i)$ is given by Equation (6.2), whereas the summation is over the four nearest neighbor cells j . With these transition probabilities, a straightforward calculation yields in the limit $N \rightarrow \infty$ the Fokker-Planck equation

$$\frac{\partial P(x_i, t)}{\partial t} = -\frac{\partial}{\partial x_i} (v(x_i)P(x_i, t)) + \frac{1}{2} \frac{\partial^2}{\partial x_i^2} (d(x_i)P(x_i, t)) \quad (6.6)$$

with

$$v(x_i) = \frac{h}{4} \sum_j (x_j - x_i) + 2ax_i(1 - x_i)(2x_i - 1) - 2bx_i(1 - x_i)(2x_i - 1)^3 \quad (6.7)$$

and

$$d(x_i) = 2x_i(1 - x_i) . \quad (6.8)$$

For the readers interested in the full derivation of Equations (6.6) and (1.3), we explain the detailed calculations in the appendix A 8. Finally, using the substitution $\phi_i = 2x_i - 1$ in the Langevin correspondence yields the following stochastic equation of motion:

$$\frac{\partial \phi_i}{\partial t} = a\phi_i(1 - \phi_i^2) - b\phi_i^3(1 - \phi_i^2) + \frac{h}{4} \sum_j (\phi_j - \phi_i) + \sqrt{2} \sqrt{1 - \phi_i^2} \eta \quad (6.9)$$

which is identical to the Langevin Equation (1.3) after noting that the sum over j is proportional to the discretized version of the diffusion term $\nabla^2 \phi$.

Chapter 7

Numerical Phase Diagram and Universality Classes

This chapter, with some changes, is adapted from our work submitted to Phys. Rev. E:

A. Azizi, and M. Pleimling, “Critical phenomena in presence of symmetric absorbing states: a microscopic model with tunable parameters”.

In the following we show that the Russell-Blythe model with the probability (6.3) indeed exhibits a phase diagram with a disordered state, an absorbing state and an ordered (but not absorbing) state, separated by phase transition lines belonging to the Generalized Voter, the two-dimensional Ising and the Directed Percolation universality classes. Most of the data discussed in the following have been obtained for $L = 100$ and $N = 100$ (i.e. with a total of one million spins) and periodic boundary conditions in both directions. We fix $h = 0.5$ and vary the system parameters a and b in order to obtain the phase diagram in the $b - a$ plane. Our time step corresponds to $L \times L \times N$ updates, i.e. on average every spin is updated once during a time step.

7.1 Quantities

As we expect to encounter transition lines that either separate an active from an absorbing phase or a disordered from an ordered phase, we need to use very different quantities in order to probe the universality classes of the different transitions.

Useful quantities for locating an active-absorbing phase transition point and determining the corresponding critical universality class [4] are provided by the time-dependent average number of flipped cells $N_f(t)$ and the survival probability $P_s(t)$ of flipped cells. Indeed, exactly at the absorbing phase transition these two quantities are expected to vary algebraically with

time,

$$P_s(t) \sim t^{-\delta} \quad \text{and} \quad N_f(t) \sim t^\eta, \quad (7.1)$$

with the Directed Percolation values $\delta_{DP} = 0.45$ and $\eta_{DP} = 0.23$ [39,40] and the Generalized Voter universality class values $\delta_{GV} = 1.0$ and $\eta_{GV} = 0$ [56,62]. Starting point for our calculation of these quantities is an initial state where all spins in all the cells take on the same value (say, -1 , which results in the local (cell) magnetization density $\phi = -1$). Having prepared the system in this way, we flip all the spins in four neighboring cells to $+1$, yielding the cell magnetization density $\phi = 1$ for these cells, and then update the system using the scheme described in Section 6.1. At every time step we measure the number of cells with $\phi \neq -1$. Repeating this procedure thousands of times with different random number sequences, the average number N_f of flipped cells (i.e. cells with $\phi \neq -1$) then results from an average over these different runs. For the survival probability P_s we record at each time step t the fraction of runs that still have cells with $\phi \neq -1$.

At the voter critical point the interface density ρ displays a characteristic time dependence in the form of a logarithmic decay [10]:

$$\rho(t) \sim 1/\ln t \quad (7.2)$$

which makes it an excellent quantity for identifying the location of the critical point as well as the universality class. We prepare our system in a fully disordered initial state (i.e. every spin can take on the values 1 and -1 with the same probability $1/2$) and then calculate the interface density

$$\rho(t) = 1 - \left\langle \frac{1}{2L^2} \sum_{j=1}^L \sum_{k=1}^L \phi_{j,k}(t) (\phi_{j+1,k}(t) + \phi_{j,k+1}(t)) \right\rangle \quad (7.3)$$

where $\langle \dots \rangle$ indicates an ensemble average over different noise realizations and $\phi_{j,k}(t)$ is the cell magnetization density at time t of the cell $i = (j, k)$ where j and k are labeling the cells along the horizontal and vertical direction respectively.

Another dynamic quantity of interest is the two-time autocorrelation function

$$C(t, s) = \left\langle \frac{1}{L^2} \sum_{j=1}^L \sum_{k=1}^L \phi_{j,k}(t) \phi_{j,k}(s) \right\rangle - \left\langle \frac{1}{L^2} \sum_{j=1}^L \sum_{k=1}^L \phi_{j,k}(t) \right\rangle \left\langle \frac{1}{L^2} \sum_{j=1}^L \sum_{k=1}^L \phi_{j,k}(s) \right\rangle \quad (7.4)$$

which at a critical point is expected to display the simple aging scaling [23]

$$C(t, s) = s^{-\tilde{b}} f_C(t/s) \quad (7.5)$$

when starting from a disordered initial state. The scaling function f_C only depends on the ratio of the two times t and s and decays algebraically for large arguments with an exponent λ . In [58] we showed that this aging scaling indeed prevails at a voter critical point, with the exponents taking on the values $\tilde{b} = 0.120(5)$ and $\lambda = 1.00(2)$.

Finally, we also computed standard quantities used for investigating equilibrium second order phase transitions, notably moments of the magnetization density M and the fourth order cumulant (Binder cumulant) of the magnetization [63–65]

$$U = 1 - \frac{\overline{M^4}}{3\overline{M^2}^2} . \quad (7.6)$$

As we will see below, the Binder cumulant is very useful for finding as a function of the parameter a the location of the paramagnetic-ferromagnetic phase transition for fixed value of b . In order to obtain these equilibrium quantities, we first let the system relax for several thousand time steps before sampling data at fixed time intervals. This is repeated using different random number sequences so that $\overline{\cdots}$ indicates both a time and an ensemble average.

7.2 Phase diagram of the microscopic model

The phase diagram that results from the Langevin Equation (1.3) for systems with two absorbing states is well understood [11,12]. There exists a special value $b = b_s$ that separates the phase diagram into two parts with very different properties. For $b < b_s$ a unique phase transition takes place at a critical value of the parameter a . This phase transition, which separates a disordered, active steady state from an ordered and absorbing state, belongs to the universality class of the Generalized Voter model. For $b > b_s$, however, two consecutive phase transitions are encountered when increasing a . The system first undergoes an order-disorder phase transition that belongs to the universality class of the two-dimensional Ising model. The two phases separated by this phase transition are active phases. Increasing a further, a second phase transition to an absorbing state takes place. This second transition belongs to the Directed Percolation universality class.

Previous efforts to verify these scenarios through microscopic models have only been partially successful, due to the lack of a microscopic model with parameters that can be changed continuously. Indeed, the best studied cases are that of absorbing Ising models on a square lattice with the added voter rule that a spin is not allowed to flip if it points in the same direction as all the spins in a specified neighborhood [9,12,58]. For nearest neighbor interactions only a voter transition is observed, whereas extending the interaction range to further neighbors yields at some point a splitting of the voter critical point and the appearance of two phase transitions. However, for a system of spins on a lattice the interaction range can only be changed in discrete steps (going from nearest neighbor interactions to next nearest neighbor interactions etc.) which does not make these models convenient ones for a more systematic study.

For the model presented in the previous Section the parameters a and b , whose values can take on any real number, directly enter into the probabilities to update the system. As a result changing the values of these two parameters allows to fully explore the phase diagram

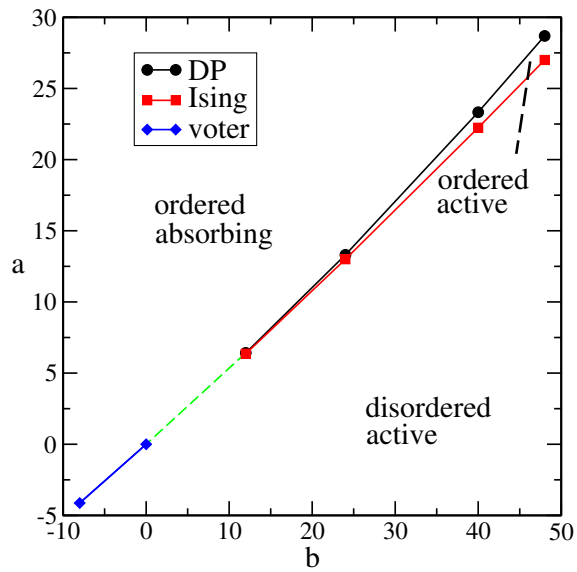


Figure 7.1: Phase diagram in the $b - a$ plane obtained from simulations of the microscopic model. The properties of the phase transitions taking place at particular values of the system parameters a and b (filled symbols) have been elucidated. Along the green dashed segment a special point exists where all three critical lines merge. No attempt was made to locate this special point. The full lines are guides to the eye.

and fine tune the properties of the model. We show in Figure 7.1 the phase diagram that we have obtained from the numerical simulations of this microscopic model. We encounter three different steady states: a disordered and active phase in the lower part of the phase diagram, an ordered and inactive phase in the upper part of the diagram, and an intermediate phase that is both ordered and active and that for values of b larger than a special value b_s is intercalated between the other two phases. As we show in the rest of this Section, the critical lines separating these phases belong to the expected three universality classes (Generalized Voter, two-dimensional Ising, and Directed Percolation) in agreement with the results obtained from the Langevin equation. These three lines should meet at a special point (b_s, a_s) located in the Figure along the green dashed line. We did not attempt to locate precisely this special point through a time-consuming systematic search.

7.2.1 One critical line belonging to the Voter universality class

We start our analysis of the character of the phase transitions by looking at the case $b \leq 0$. For these values of b we encounter a single phase transition that is expected to belong to the Generalized Voter universality class.

In order to locate the critical value of a for a fixed value of b we investigate how different quantities behave when changing a . This is shown in Figure 7.2 for the case $b = 0$. The inset in Figure 7.2a shows that the same behavior is encountered for other values of $b \leq 0$. The interface density ρ in Figure 7.2a and the survival probability P_s in Figure 7.2c directly allow to locate the phase transition and confirm the voter critical behavior through their characteristic time dependence ($1/\rho(t) \sim \ln t$ and $P_s(t) \sim t^{-\delta_{GV}}$ with $\delta_{GV} \approx 1$).

An interesting behavior is exhibited by the average number of flipped cells N_f in Figure 7.2b. The expected voter critical behavior for this quantity ($N_f(t) \sim t_{GV}^\eta$ with $\eta_{GV} = 0$) suggests that $N_f(t)$ should be constant, and this has indeed been observed in a two-dimensional absorbing Ising model [12] as well as in linear voter models [28, 58]. For our model, however, N_f shows an early algebraic growth before it decreases as $1/\ln t$ at late times. While this is formally compatible with $\eta = 0$, this behavior indicates that different scenarios are possible at a Generalized Voter critical point. As the number of cells whose cell magnetization density is not equal to -1 can be viewed as a proxy for the magnetization of a system formed by L^2 cells, we note that the logarithmic time dependence in Figure 7.2b is mirrored by a corresponding logarithmic behavior of the magnetization of non-linear models at their voter critical point [10, 28, 58]. The inset in Figure 7.2a also reveals an interesting early time behavior of N_f at the critical point. The observed power law increase with an exponent 0.19, indicated by the dashed line, is reminiscent of the initial slip behavior of the magnetization in systems with a small, non-vanishing magnetization quenched to an equilibrium critical point [33, 66]. After this initial regime, N_f reaches a maximum and then decays logarithmically with time, as discussed above.

In Figure 7.3 we verify that at the points $(b = 0, a = 0)$ and $(b = -8, a = -4.132)$ we

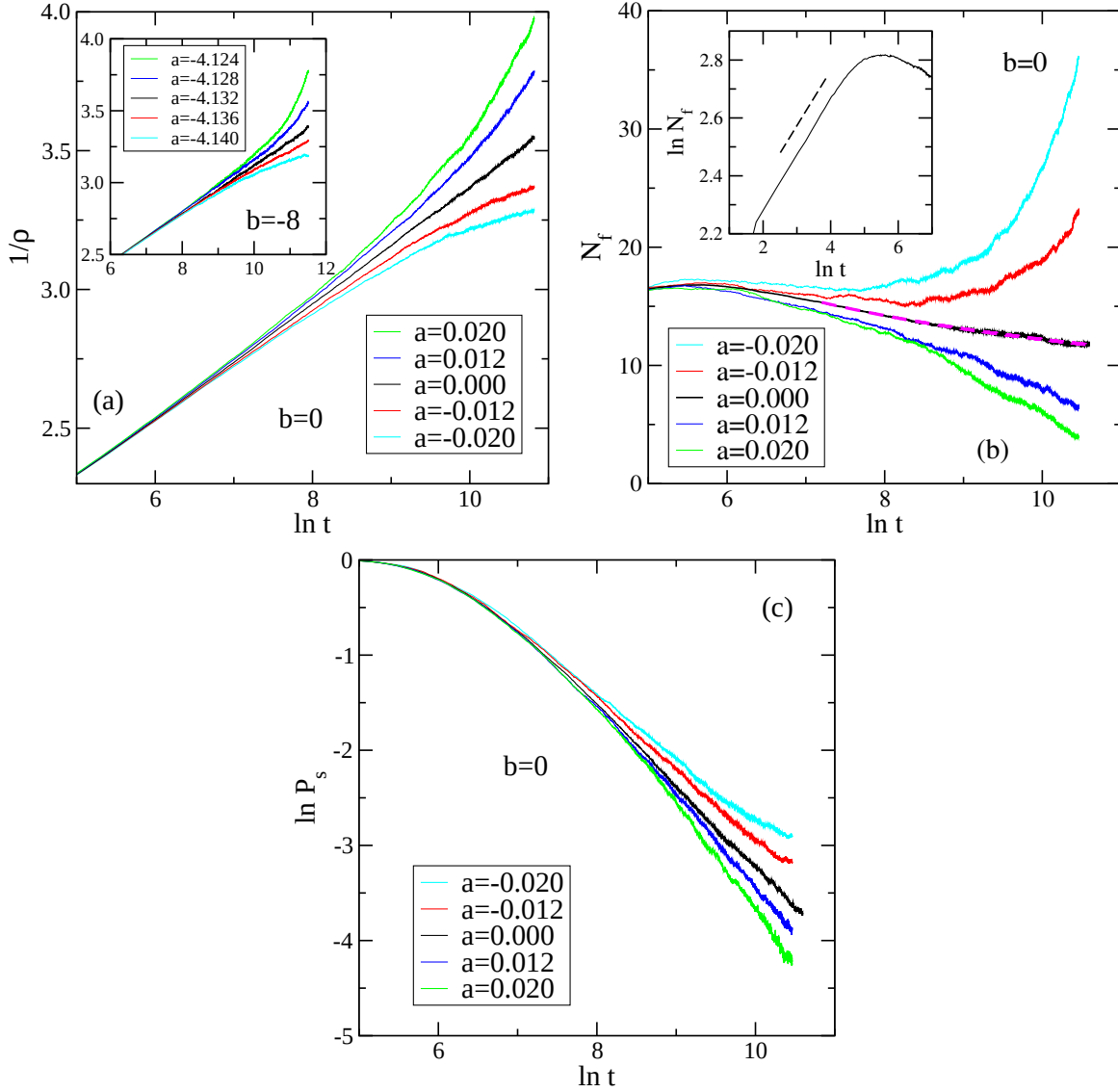


Figure 7.2: Time dependence of various quantities for systems with $b = 0$ and different values of a . Based on these data a phase transition belonging to the Generalized Voter universality class takes place at $a = 0$. The different quantities display a typical behavior: (a) the interface density ρ decays as $1/\ln t$, (b) the average number of flipped cells N_f also decays like $1/\ln t$, which is compatible with $\eta_{GV} = 0$, and (c) the survival probability decays algebraically with an exponent $\delta_{GV} \approx 1$. The same behavior is encountered for $b < 0$. This is illustrated in the inset of panel (a) for the case $b = -8$ for which the phase transition is found to take place at $a = -4.132(2)$. The inset in panel (b) shows that at early times the average number of flipped cells increases at the critical point $a = 0$ algebraically with an exponent 0.19 (indicated by the dashed line). The data for the interface density result from averaging over at least 3000 independent runs. For the other two quantities averages were done over at least 5000 independent runs.

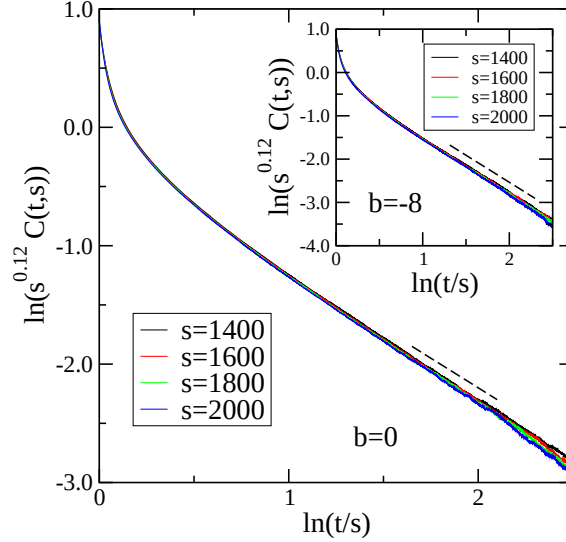


Figure 7.3: Scaling of the autocorrelation function $C(t, s)$ for $a = 0$ and $b = 0$. Simple aging scaling (7.5) with the exponents $\tilde{b} = 0.12$ and $\lambda = 1$ (indicated by the dashed line) is encountered, in agreement with the expected behavior at a voter critical point [58]. The inset shows that the same behavior prevails for $b = -8$ and $a = -4.132$. These data result from averaging over 5000 independent runs.

have for the two-time autocorrelation function the same critical aging scaling behavior (7.5) with exponents $\tilde{b} = 0.12$ and $\lambda = -1$ as encountered at other voter critical points [58]. Taken together, the data shown in Figures 7.2 and 7.3 unambiguously reveal that for $b \leq 0$ our microscopic model exhibits a line of Generalized Voter critical points separating the disordered active phase from the absorbing phase.

7.2.2 Two critical lines belonging to the Ising and Directed Percolation universality classes

For fixed $b > b_s$ the expected scenario is the appearance of two separate phase transitions when varying a . In the following we discuss the case $b = 40$, but similar results are obtained for other values of b , see Figure 7.1.

In Figure 7.4 we show the Binder cumulant (7.6) as a function of a for various numbers of cells L^2 . The dependence on the system parameters and the system size is as expected for a continuous order-disorder phase transition. The lines for the different system sizes cross in the vicinity of $a = 22.16$ which provides an estimate for the location of the phase transition. The value of U at that point is close to the theoretical value $U_I \approx 0.61$ for a phase transition belonging to the universality class of the two-dimensional Ising model. The Ising character of this phase transition is also verified through the behavior of the magnetization close to

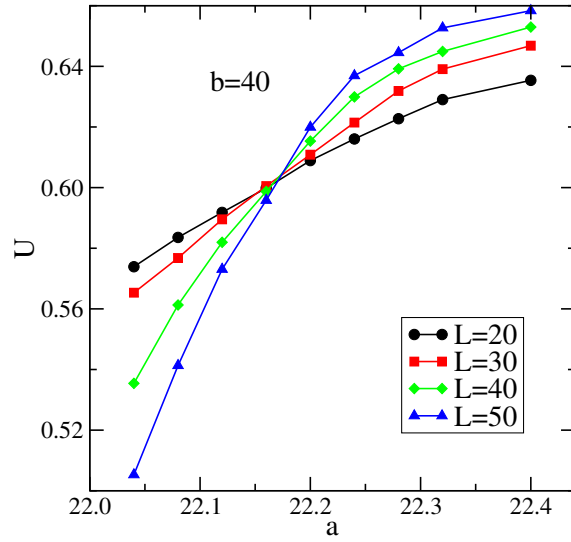


Figure 7.4: Binder cumulant U for various numbers of cells L^2 as a function of a when $b = 40$. The different lines intersect close to $a = 22.16$. The value of U at this point is close to the universal value $U_I \approx 0.61$ for the two-dimensional Ising universality class [65,67]. For these equilibrium simulations the first 300,000 time steps were discarded before sampling data for a time average. The final data result from an additional averaging over typically 1000 different runs. Error bars are comparable to the symbol sizes. The lines are guides to the eye.

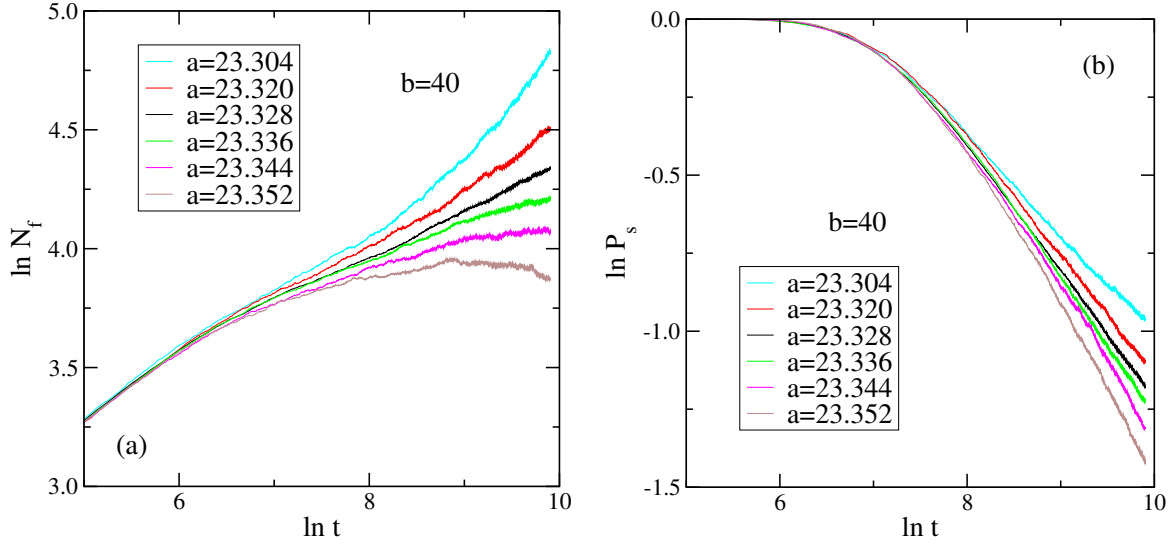


Figure 7.5: Time dependence of (a) the average number of flipped cells N_f and (b) the survival probability P_s for $b = 40$ and different values of a . These data indicate that a Directed Percolation phase transition takes place for $a = 23.332(4)$. The data result from averaging over at least 4000 different realizations of the noise.

this point.

Increasing further a results in a second phase transition, this time between an ordered but active phase, the intermediate phase shown in Figure 7.1, and the absorbing phase. Both N_f and P_s display a power-law dependence on time around $a = 23.332$, see Figure 7.5. The corresponding exponents ($\eta \approx 0.20$ for N_f and $\delta \approx 0.43$) are very close to the expected values for Directed Percolation ($\eta_{DP} = 0.23$ and $\delta_{DP} = 0.45$).

Chapter 8

Conclusion II

Systems with two symmetric absorbing states have found applications in fields ranging from population ecology to opinion dynamics. The presence of symmetric absorbing states allows for two different scenarios when crossing from the active phase to the absorbing phase. In a direct transition, where the order-disorder transition coincides with the active-absorbing transition, critical properties are those of the generalized voter universality class, characterized by the fact that after a quench to the critical point the interface density decays logarithmically with time. The transition to the absorbing state can also happen in two steps, with first an order-disorder transition belonging to the Ising universality class followed by an absorbing transition belonging to the directed percolation universality class. In the Langevin description a unified picture was provided where the Ising and directed percolation critical lines merge into the generalized voter critical line [11].

In Chapters 6 and 7 we discussed a microscopic spin lattice model with continuously tunable parameters that, to our knowledge for the first time, allows to access all three critical lines predicted by the Langevin Equation (1.3). Investigating various quantities used commonly in studies of equilibrium and non-equilibrium critical phenomena, we clearly identified the different universality classes. A generalization to a larger number of absorbing states (mirroring the extension from kinetic Ising to kinetic Potts models [9, 58]) is possible.

Our study revealed an initial slip behavior of the number of flipped cells (which can serve as a proxy for the magnetization density) when quenching the system to a voter critical point, which came as a surprise to us. Indeed, from the known properties of the generalized voter universality class, the number of flipped cells is expected to be time independent, a property indeed recovered at the critical points of linear voter models. The situation is more complex for critical non-linear models, with N_f displaying an initial power-law increase, followed by a decrease with a logarithmic dependence on time. This is consistent with the earlier observation of non-constant magnetization in critical non-linear voter models [10, 28, 58]. All this indicates that our understanding of the critical properties of non-linear voter models is incomplete and warrants a more in-depth investigation in the future.

In our current investigation we did not attempt to locate precisely the point at which the three critical lines merge. With a microscopic spin model with parameters that can be varied continuously now at hand, it is only a question of available resources whether this point can be located with precision high enough that allows an in-depth study of the critical properties at this special point. We plan to come back to this intriguing issue.

Appendix A: Derivation of Langevin Equation

In this chapter, we start with the definition of the microscopic model on the lattice and explain the steps to derive Equation (6.9). The system configuration at time t is denoted by X which represents the system information in all N cells, i.e, $X = \{x_1, x_2, \dots, x_N\}$. In the next time step, the configuration evolves to $X' = \{x_1, x_2, \dots, x_i + r, \dots, x_N\}$ where $r = \frac{-1}{N}$ or $r = \frac{1}{N}$. Therefore the Master Equation of the system becomes:

$$\frac{\partial P(x_i, t)}{\partial t} = \sum_{x'} T(x_i|x')P(x', t) - T(x'|x_i)P(x_i, t) \quad (8.1)$$

where $T(x_i|x')$ is the transition rate in the model and x_i , as defined in Chapter 6, is the fraction of up spins in cell i . Based on the update scheme explained in Chapter 6, x' is different from x_i with only r and so $T(x_i|x')$ can be written as follows:

$$T(x_i|x') = T(x', x_i - x') = T(x_i + r, -r); \quad T(x'|x_i) = T(x_i, x' - x_i) = T(x_i, r)$$

and accordingly the Equation (8.1) is written as:

$$\frac{\partial P(x_i, t)}{\partial t} = \sum_r [T(x_i + r, -r)P(x_i - r, t) - T(x_i, r)P(x_i, t)] \quad (8.2)$$

$$\begin{aligned} &= \sum_r \left[T(x_i, -r)P(x_i, t) + \sum_{m=1}^{\infty} r^m \frac{\partial}{\partial x^m} \{T(x_i, -r)P(x_i, t)\} - T(x_i, r)P(x_i, t) \right] \\ &\Rightarrow \frac{\partial P(x_i, t)}{\partial t} = \sum_r \sum_{m=1}^{\infty} (-1)^m r^m \frac{\partial}{\partial x^m} \{T(x_i, r)P(x_i, t)\} \end{aligned} \quad (8.3)$$

where the Equation ((8.3)) is the well known Kramers-Moyal expansion. Defining:

$$\alpha_m = \sum_r r^m T(x_i, r)$$

then we can make an approximation by returning only the first two terms of the expansion:

$$\frac{\partial P(x_i, t)}{\partial t} \approx -\frac{\partial}{\partial x_i}(\alpha_1 P(x_i, t)) + \frac{1}{2} \frac{\partial}{\partial x_i}(\alpha_2 P(x_i, t)) \quad (8.4)$$

where $\alpha_1 = \sum_r r T(x_i, r)$ and $\alpha_2 = \sum_r r^2 T(x_i, r)$.

Note that in the definition of our model r can be only $\pm\Delta = \pm\frac{1}{N}$. Therefore (8.4) is simplified to

$$\frac{\partial P(x_i, t)}{\partial t} = -\frac{1}{N} \frac{\partial}{\partial x_i} \left((T(x_i, \Delta) - T(x_i, -\Delta)) P(x_i, t) \right) + \frac{1}{2N^2} \frac{\partial^2}{\partial x_i^2} \left((T(x_i, \Delta) + T(x_i, -\Delta)) P(x_i, t) \right) \quad (8.5)$$

Recall that the above equation was written for a randomly chosen cell x_i , but to obtain the final Fokker-Planck equation we need to include all possible flip transition possibilities for all cells:

$$\frac{\partial P(X, t)}{\partial t} = \sum_{i=1}^{L^2} \left\{ -\frac{1}{N} \frac{\partial}{\partial x_i} \left((T(x_i, \Delta) - T(x_i, -\Delta)) P(X, t) \right) + \frac{1}{2N^2} \frac{\partial^2}{\partial x_i^2} \left((T(x_i, \Delta) + T(x_i, -\Delta)) P(X, t) \right) \right\} \quad (8.6)$$

With the Fokker-Planck equation of the model at hand, we can now easily calculate the drift, $v(x_i)$, and the diffusion, $D(x_i)$, terms explicitly. Noticing that the transition probabilities T are the flip probabilities w multiplied with time unit $\frac{1}{N^2}$ and that $T(x_i, \Delta) = T(x_i + \Delta|x_i)$, we then obtain:

$$D(x_i) = \frac{1}{N^2} \left(T(x_i + \Delta|x_i) + T(x_i - \Delta|x_i) \right) = \frac{N^2}{N^2} \left(w(x_i + \Delta|x_i) + w(x_i - \Delta|x_i) \right) \quad (8.7)$$

$$v(x_i) = \frac{1}{N} \left(T(x_i + \Delta|x_i) - T(x_i - \Delta|x_i) \right) = \frac{N^2}{N} \left(w(x_i + \Delta|x_i) - w(x_i - \Delta|x_i) \right) \quad (8.8)$$

Applying the flipping probability rules defined in Equation (6.4), we derive the final expression for $D(x_i)$:

$$D(x_i) = \frac{N^2}{N^2} \left((1-x_i) \left\{ \frac{h}{Nz} \sum_j x_j + \left(1 - \frac{h}{N} \right) p(x_i) \right\} + x_i \left\{ \frac{h}{Nz} \sum_j (1-x_j) + \left(1 - \frac{h}{N} \right) p(1-x_i) \right\} \right)$$

Since in our model N is considered to be as large as possible, then all terms with order $O(N^{-1})$ or higher should be neglected. Therefore the above expression is reduced to:

$$D(x_i) \approx \left((1-x_i)p(x_i) + x_i p(1-x_i) \right) = \left((1-x_i) \left(x_i + \frac{2a}{N} x_i(1-x_i)(2x_i-1) \right) + x_i \left((1-x_i) + \frac{2b}{N} (1-x_i)(x_i)(-2x_i+1)^3 \right) \right)$$

and finally:

$$D(x_i) \approx (1 - x_i)x_i + x_i(1 - x_i) = 2x_i(1 - x_i)$$

Similar to $D(x_i)$, for $v(x)$ we write:

$$\begin{aligned} v(x_i) &= \frac{N^2}{N} \left((1-x_i) \left\{ \frac{h}{Nz} \sum_j x_j + \left(1 - \frac{h}{N}\right) p(x_i) \right\} - x_i \left\{ \frac{h}{Nz} \sum_j (1-x_j) + \left(1 - \frac{h}{N}\right) p(1-x_i) \right\} \right) \\ &= \frac{N^2}{N} \left(\frac{h}{Nz} \sum_j (x_j - x_i) + \left(1 - \frac{h}{N}\right) \{ (1-x_i)p(x_i) - x_i p(1-x_i) \} \right) \end{aligned} \quad (8.9)$$

The expression in the curly brackets can be simplified as follows:

$$\begin{aligned} \{...\} &= (1-x_i) \left(x_i + \frac{2a}{N} x_i(1-x_i)(2x_i-1) - \frac{2b}{N} x_i(1-x_i)(2x_i-1)^3 \right) - \\ &\quad x_i \left((1-x_i) + \frac{2a}{N} (1-x_i)(x_i)(1-2x_i) - \frac{2b}{N} (1-x_i)x_i(1-2x_i)^3 \right) \end{aligned}$$

which is equal to:

$$\begin{aligned} \{...\} &= \frac{2a}{N} \{ x_i(1-x_i)(2x_i-1) \} - \frac{2b}{N} \{ (1-x_i)x_i(1-x_i)(2x_i-1)^3 - x_i(1-x_i)x_i(1-2x_i)^3 \} \\ &= \frac{2a}{N} \{ x_i(1-x_i)(2x_i-1) \} - \frac{2b}{N} \{ x_i(1-x_i)(2x_i-1)^3 \} \end{aligned}$$

and consequently, by substituting the above expression in Equation (8.9) (with $z = 4$), we derive the compact form:

$$v(x_i) = \frac{h}{z} \sum_j (x_j - x_i) + 2a \left(x_i(1-x_i)(2x_i-1) \right) - 2b \left(x_i(1-x_i)(2x_i-1)^3 \right)$$

In the next step, the explicit expressions of $D(x)$ and $v(x)$ will be used to obtain the Langevin equation corresponding to our microscopic model. Hence the Fokker-Planck Equation ((8.6)) is rewritten as:

$$\frac{\partial}{\partial t} P(X, t) = - \sum_{i=1}^{L^2} \frac{\partial}{\partial x_i} \left(v(x_i) P(X, t) \right) + \frac{1}{2} \frac{\partial}{\partial x_i^2} \left(D(x_i) P(X, t) \right) \quad (8.10)$$

where its corresponding Langevin equation is

$$\partial_t x_i = v(x_i) + \sqrt{D(x_i)} \eta \quad (8.11)$$

Using the explicit expressions for $v(x)$ and $D(x)$, the Equation (8.11) becomes:

$$\frac{\partial x_i}{\partial t} = \left(2a(2x_i-1)(1-x_i)x_i - 2b(2x_i-1)^3(1-x_i)x_i \right) + h \sum_j (x_j - x_i) + \sqrt{x_i(1-x_i)} \eta \quad (8.12)$$

Finally, with the transformation $x_i = \frac{1+\phi_i}{2}$, the above formula recovers the Langevin equation of the generalized voter model:

$$\frac{\partial \phi_i}{\partial t} = a\phi_i(1 - \phi_i^2) - b\phi_i^3(1 - \phi_i^2) + \frac{h}{4} \sum_j (\phi_j - \phi_i) + \sqrt{2}\sqrt{(1 - \phi_i^2)}\eta \quad (8.13)$$

Chapter 9

Interface Width in Systems with Absorbing States

In this chapter we extend our investigation to the fluctuations of interfaces in critical systems with absorbing states. Specifically, we are interested in understanding the dynamical properties of the interface width in the LD model and the MGVM at their Voter critical points and DP critical points. To be more precise about the concept of interface in these models, consider a specific initial configuration in the LD model where the left side of the lattice is occupied with 0 spins, whereas the right side is occupied with +1 spins. Hence a sharp interface is formed between these two sides (see Figure 9.1).

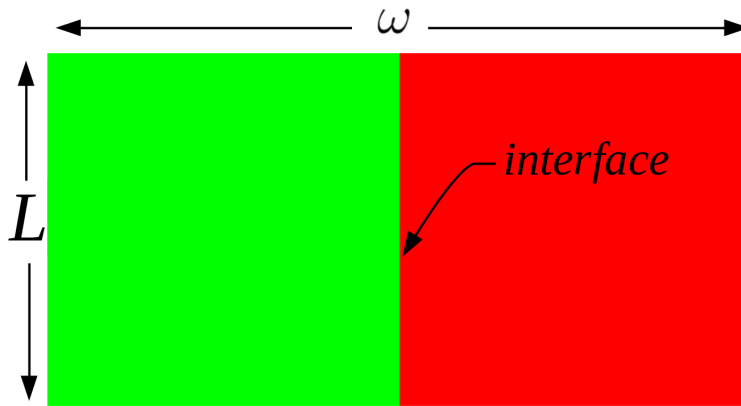


Figure 9.1: Initial configuration of LD model with width $\omega = 400$ and interface size $L = 200$. Left side (green) and right side (red) of the lattice contain 0 spins and +1 spins respectively.

Our goal is to study interface fluctuations for different scenarios. For instance, with a suitable choice of temperature and interaction range in the LD model, the two sides can be kept at the Voter critical point or at the DP critical point. We show that these two cases eventually

cause different dynamical behaviors of the interface fluctuations. Another possible scenario is to keep both sides of the lattice at an Ising critical point at which an order-disorder phase transition takes place. In this case, as shown in Figure 9.2, the interface between left side (spins 0) and right side (spins +1) will be disintegrated and many small clusters of +1 spins will emerge on the left side and vice versa. This indicates that the interface between the sides will not be stable and therefore the interface fluctuations W will not be meaningful at long time. Nevertheless, our results discussed below for DP critical points suggest a different scenario of exponents for characterizing W . At Voter critical points we show that our findings are compatible with previous works.

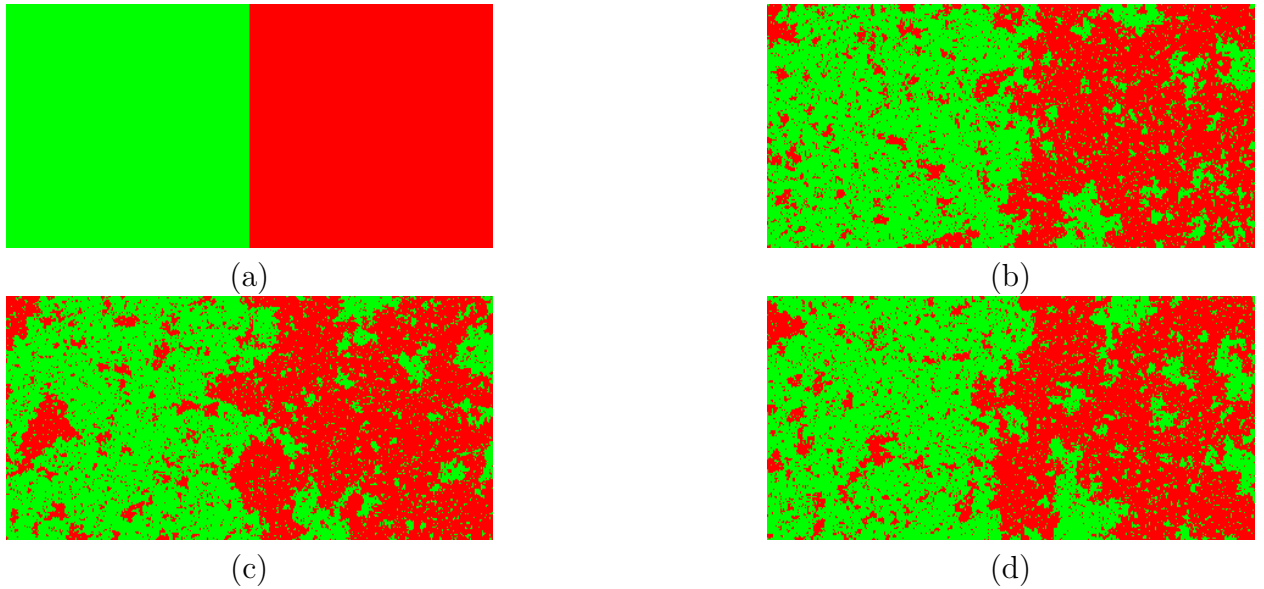


Figure 9.2: Snapshots of the configurations at LD model Ising critical point at different time steps (a) $t = 1$, (b) $t = 100$, (c) $t = 500$, (d) $t = 1000$. Spins in the left and right sides of the system are initially set to 0 and +1 respectively. Both sides are at the Ising critical point ($T_I = 4.812$ and $z = 12$). Time evolution of the model shows the interface h does not change smoothly in space and therefore W will not be meaningful.

Interface fluctuations $W(t)$

For spin models on a lattice the interface can be defined as a set of spins which determines the boundary between cluster with opposite spin states. One way to characterize the behavior of the interface is to study its fluctuations W as a function of time. Consider that $h(i, t)$ is the width height at position i and at time t (at time $t = 0$, $h(i, 0) = 0$ for every i as we start

with a straight interface). Hence, the interface width W is defined as:

$$W(L, t) = \sqrt{\frac{1}{L} \sum_{i=1}^L (h(i, t) - \bar{h}(t))^2} \quad (9.1)$$

where L is the interface length and $\bar{h}(t) = \frac{\sum_i h(i, t)}{L}$. Although dynamical behaviors of interfaces are in general dependent on many factors, in a large number of interesting physical systems it is shown that the interface fluctuation W initially grows as a power of time until it reaches the saturation point at time t_s . For times $t > t_s$, W does not grow with time any more but fluctuates around a constant value (saturation value W_s). As shown in Figure 9.3, W_s changes as a power of the interface length L as well. Therefore $W(L, t)$ can be described as follows:

$$W(L, t) \propto \begin{cases} t^{\beta_w} & t < t_s \\ L^{\alpha_w} & t > t_s \end{cases} \quad (9.2)$$

where β_w and α_w are growth and roughness exponents respectively. We note that at the saturation point t_s , W values should match when approached from two regimes, i.e, $t^{\beta_w} \propto L^{\alpha_w}$. It then follows that $t_s \propto L^z$ with the dynamical exponent $z = \frac{\alpha_w}{\beta_w}$. Hence, in a more general form, $W(L, t)$ can be written as [68]:

$$W(L, t) = L^{\alpha_w} f\left(\frac{t}{L^z}\right), \text{ with } f(u) \approx \begin{cases} u^{\beta_w} & u \ll 1 \\ \text{const.} & u \gg 1 \end{cases} \quad (9.3)$$

This form of W is observed in a broad range of physical systems, from spins on lattices to interacting particle systems [60]. For instance, interface fluctuations in the ballistic deposition (BD) models (see Figure 9.4) has the shape of Equation (9.3) with $\beta_w = 0.33$ and $\alpha_w = 0.47$, whereas for random deposition (RD) models, the growth exponent value is $\beta_w = 0.5$ with no saturation point [68].

Another approach to study the behavior of a growth process is to construct a stochastic equation as the continuum version of models defined on a lattice. A general form of a continuum growth model is defined as follows:

$$\frac{\partial h(x, t)}{\partial t} = \Phi(h, x, t) = f(h, x, t) + \eta(x, t) \quad (9.4)$$

where, in resemblance to the lattice model, $\Phi(h, x, t)$ can be regarded as a function of the number of particles deposited on the surface at position x and time t . In this stochastic process $\Phi(h, x, t)$ is written as a deterministic function $f(h, x, t)$, which contains information about properties of the interface (such as symmetries and conserved quantities), combined with the noise term $\eta(x, t)$, which can be multiplicative or additive.

In the above equation $f(h, x, t)$ is written in the most general form, but basic symmetries of the model enforces strong constraints on the final form of $f(h, x, t)$.

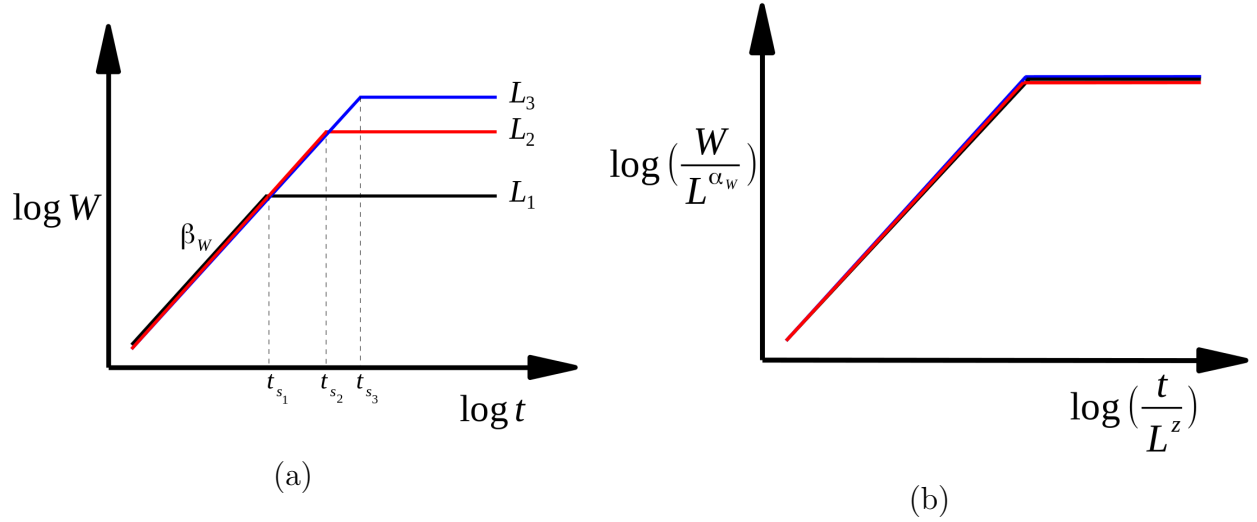


Figure 9.3: A schematic description of W as a function of time t and interface size L . (a) For interface size L_i , interface fluctuations W_i grow in time with exponent β_w and at the saturation point t_{s_i} , W_i reaches a constant value. (b) Scaling of W with respect to L . Given z , the best value for α_w is obtained when all graphs of W s collapse on to one graph.

A simple case is the Edwards-Wilkinson (EW) equation:

$$\frac{\partial h(x, t)}{\partial t} = \nu \nabla^2 h + \eta(x, t) \quad (9.5)$$

that can be solved exactly. The final form of W for the EW Equation implies [68]:

$$\alpha_w = \frac{2-d}{2}, \quad \beta_w = \frac{2-d}{4}, \quad z = 2$$

where d is the interface dimension and the exponents obey the simple relationship $z = \frac{\alpha_w}{\beta_w}$.

Finally, a more complicated growth equation called KPZ is obtained when, along with the term $\nabla^2 h$, the nonlinear term $(\nabla h)^2$ is added to the right hand side of Equation (9.5):

$$\frac{\partial h(x, t)}{\partial t} = \nu \nabla^2 h + \frac{\lambda}{2} (\nabla h)^2 + \eta(x, t) \quad (9.6)$$

for this process the roughness exponent $\alpha_w = \frac{1}{2}$ and the growth exponent $\beta_w = \frac{1}{3}$ for $d = 1$. It is worth noting that the BD model in continuum space can be mapped to the KPZ equation and consequently they belong to the same universality class [68].

In the following Sections we investigate the interface fluctuations of systems with absorbing states which are at their DP critical points or Voter critical points. We first discuss previous works and then explain the details of our simulations involving the LD model and MGVM.

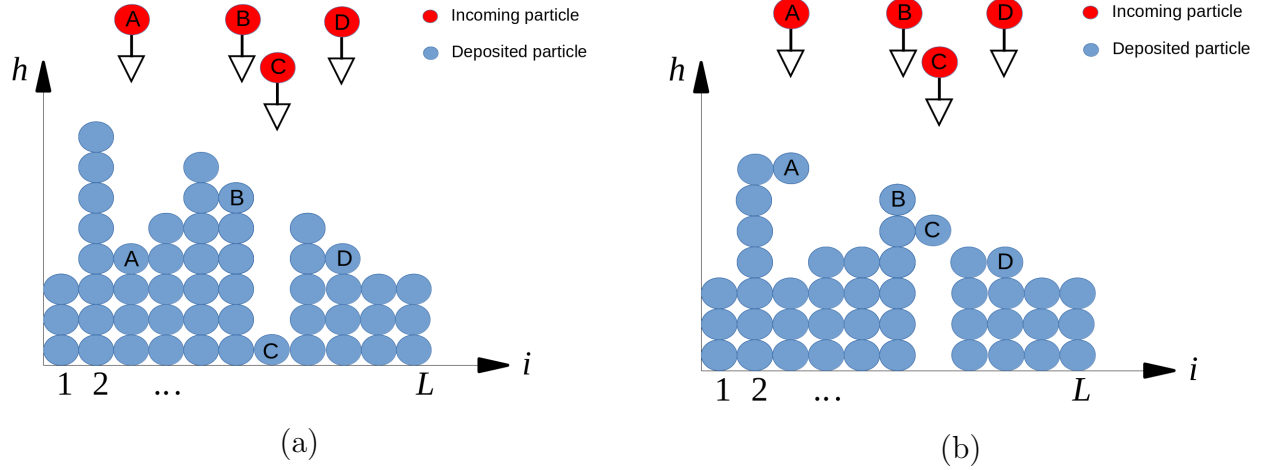


Figure 9.4: (a) A lattice example of random deposition (RD). From a randomly chosen site over the surface of deposited particles, particles fall vertically with a constant rate in time and stick to the top of the column. (b) An illustration of ballistic deposition (BD). The deposition of the falling particle depends also on the neighboring columns as well. In BD, in addition to RD, particles stick to the sides of neighboring particles if possible (particle A and C).

9.1 Interface width scaling in the Directed Percolation universality class

To the best of our knowledge, Contact Process (CP) is the only model in the DP universality class for which the interface fluctuations have been studied previously [70]. CP is defined as follows: In a d -dimensional lattice of N active (A) or inactive (I) sites, a randomly chosen pair of neighboring sites is updated per time step t ($t = \frac{1}{N}$) with the non-zero transition rates:

$$\begin{cases} w(A, I|A, A) = w(I, A|A, A) = \mu > 0 \\ w(I, I|A, I) = w(I, I|I, A) = \lambda > 0 \\ w(A, A|A, I) = w(A, A|I, A) = 1 \end{cases} \quad (9.7)$$

where $w(A, I|A, A)$ is the probability transition from a neighboring pair of active and inactive sites to both active sites. Based on numerical simulations of the DP Langevin equation, the authors of [70] measured the interface fluctuations in different dimensions. Their findings suggest that if $d < 4$, the scaling relation $\beta_w = 1 - \frac{\beta}{\nu_{||}}$ holds, and that for $d = 2$ the measured scaling exponents are:

$$\beta_w = 0.55, \quad \alpha_w = 0.97$$

Furthermore, they tried to analytically justify these values. Hence, they started with the DP stochastic differential equation:

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= \nabla^2 \rho - a\rho - b\rho^2 + \eta(x, t); \\ \langle \eta(x, t)\eta(x', t') \rangle &= \Gamma \rho(x, t) \delta(x - x') \delta(t - t')\end{aligned}\tag{9.8}$$

where ρ is the density of active sites. Defining the height of each column (in a continuum limit) as h :

$$h(x, t) = \int_0^t \rho(x, t') dt' \tag{9.9}$$

they obtained the growth equation for the models belonging to the DP universality class:

$$\begin{aligned}\frac{\partial h}{\partial t} &= \nabla^2 h - ah - b \int_0^t \left(\frac{\partial h}{\partial t'}\right)^2 dt' + \zeta(x, t); \\ \langle \zeta(x, t)\zeta(x', t') \rangle &= \Gamma h(x, \hat{t}) \delta(x - x')\end{aligned}\tag{9.10}$$

with $\hat{t} = \min(t, t')$.

Equation (9.10), in principle, describes the surface growth of models in the DP universality class, but due to the presence of the multiplicative noise and the non-linear term, the authors of [70] concluded that it is not possible to explicitly evaluate the scaling properties of W .

DP-DP interface width in the LD model

As discussed in Chapter 3, the LD model with 12 nearest neighbors exhibits a Directed Percolation critical point at $T_{DP} = 4.738$. To study the interface fluctuations, all the spins in the left half of the system are initially set to be in the 0 state and all the spins in the right half are in the state 1 (Figure 9.5). Therefore, each state constitutes one block and the dynamics of the interface between these two blocks will be studied. Interestingly, spins at the boundaries are the only active spins (at $t = 0$) and therefore determine the dynamics of the interface. Figure 9.5 shows a set of snapshots of this system at different times that illustrate the time evolution of the interface. It is important to note that for large values of t , small clusters of red spins (state 1) appear in the bulk region of the green spins (state 0), and vice versa. This phenomenon can potentially cause a finite size effect in the system if these small clusters become larger and larger. However at both sides of the system at the DP critical point there is no bulk activity. As a result, opposite spins do not grow in either of the blocks for long time and consequently do not cause a finite size effect.

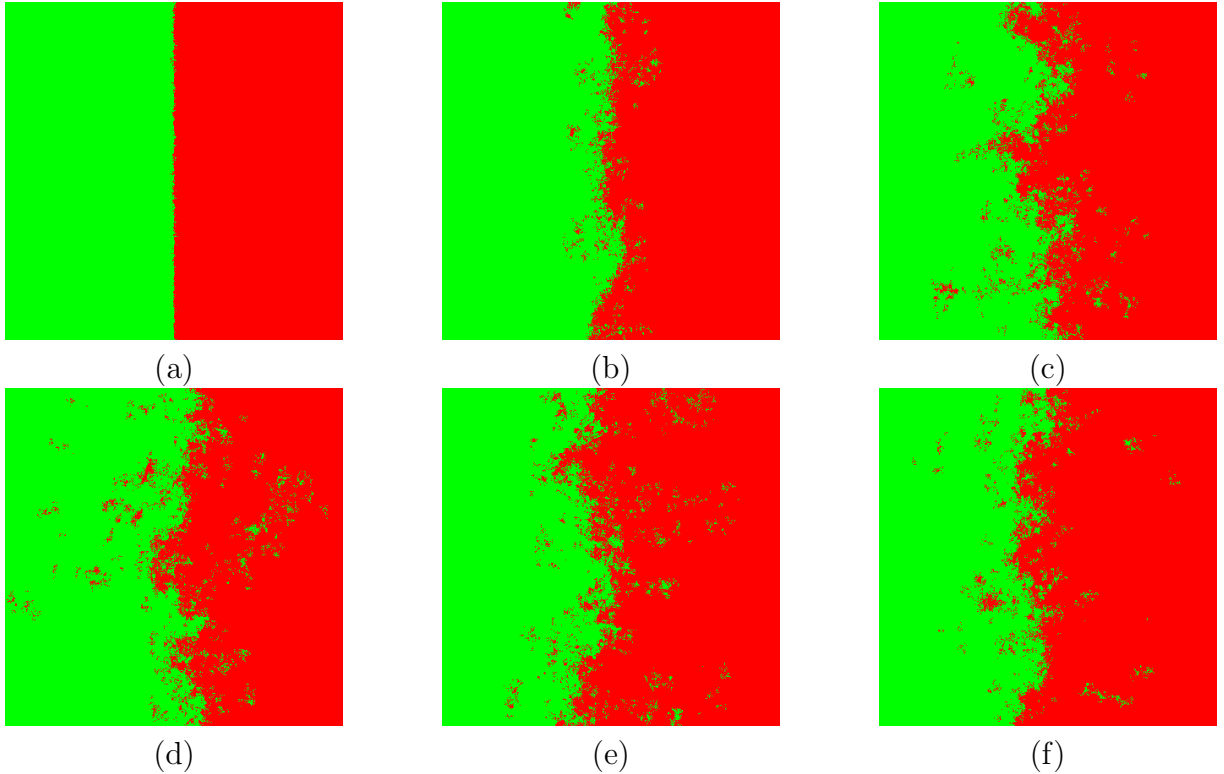


Figure 9.5: Snapshots of the LD configurations at DP critical point ($T_{DP} = 4.738$) at different time steps (a) $t = 1$, (b) $t = 500$, (c) $t = 2000$, (d) $t = 5000$, (e) $t = 20000$, (f) $t = 40000$. Spins in the left and right sides of the system are initially set to 0 and +1 respectively.

In the next step we measure the height of the interface h and its fluctuations $W(L, t)$ (Equa-

tion (9.1)) for different values of L . Based on Equation (9.2), $W(L, t)$ has a power-law dependency on time which determines the growth exponent β_w . For the roughness exponent α_w we measure $W(L, t)$ for different interface lengths L and determine the exponent α_w with which the $L^{\alpha_w} W(L, t)$ graphs collapse onto one graph at the saturation stage.

Our results show that when both sides are at the critical temperature ($T = T_{DP}$), $W(L, t)$ follows the scaling rule in Equation (9.2) with the exponents $\beta_w = 0.62$ and $\alpha_w = 0.08$. The discrepancy between these values and the ones determined in the CP model, specially the substantial difference in the value of α_w , is not yet understood. It is also interesting to note that if either side of the system is kept at a temperature $T \neq T_{DP}$, then no scaling is observed for the interface width and one of the clusters finally fills the whole system.

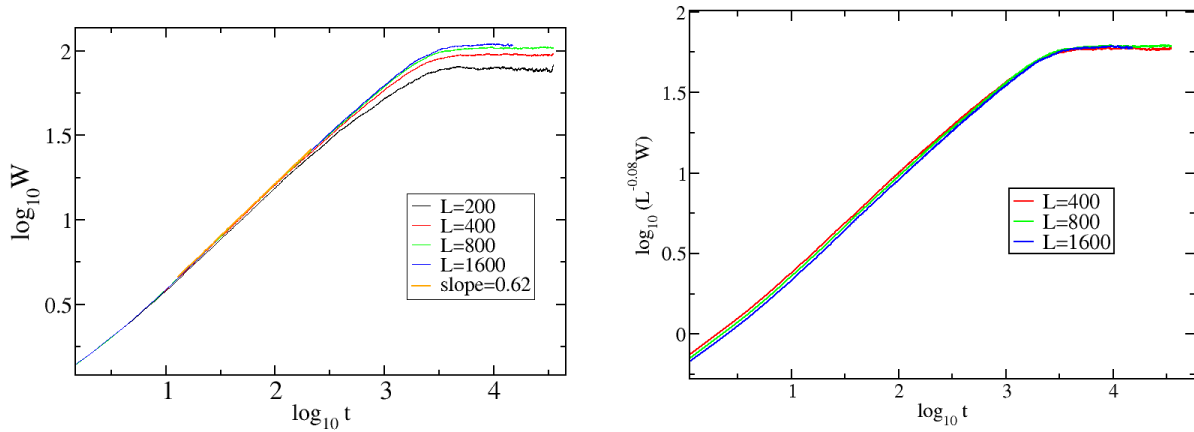


Figure 9.6: Dynamical scaling of the fluctuation of the interface width $W(L, t)$ in the LD model for a fixed system width $w = 400$. **Left** graph $W(L, t)$ for different interface lengths L . As shown in the **right** graph, after $t \approx 5000$ the systems for different lengths of L reach the saturation regime with the scaling exponent $\alpha_w = 0.08$. Note that the exponent α_w is determined based on the large system sizes as the linear extension $L = 200$ is too small to exhibit this scaling.

DP-DP interface width in Microscopic Generalized Voter model

Following the same protocol as for the LD model, we study the dynamics of the DP-DP interface in the Microscopic Generalized Voter (MGV) model. To prepare the system at the DP critical point on both sides we tune the parameters b and a to 40 and 23.328 respectively (see Chapter 6). A main difference between the LD model and MGV model is that unlike the spins in the LD model, the cells in the MGV can take on values between $\frac{-N}{N}$ to $\frac{+N}{N}$ (with increment $\frac{1}{N}$), hence there are no longer only two possible states in the system. Nevertheless, for this case we consider all cells with $\phi < 0$ as one state and all cells with $\phi > 0$ as another state which helps us measure the interface width as done in the LD model (see Figure 9.7

(b) and (d)). Note that this simplification can add effects if there exist many cells with values close to 0 (for example $\frac{-2}{N}$ or $\frac{+3}{N}$) on both ends. These cells will not be recognized in our method and therefore can cause problems. To ensure that our measurements are not affected by such effects, we monitor the growth of small clusters as time evolves and check whether the system size is large enough. Figure 9.7 shows the system configurations at the final times of the simulations where the extension of such clusters on both sides are not large enough to cause a finite size effect. To determine the scaling exponents of $W(L, t)$ in the

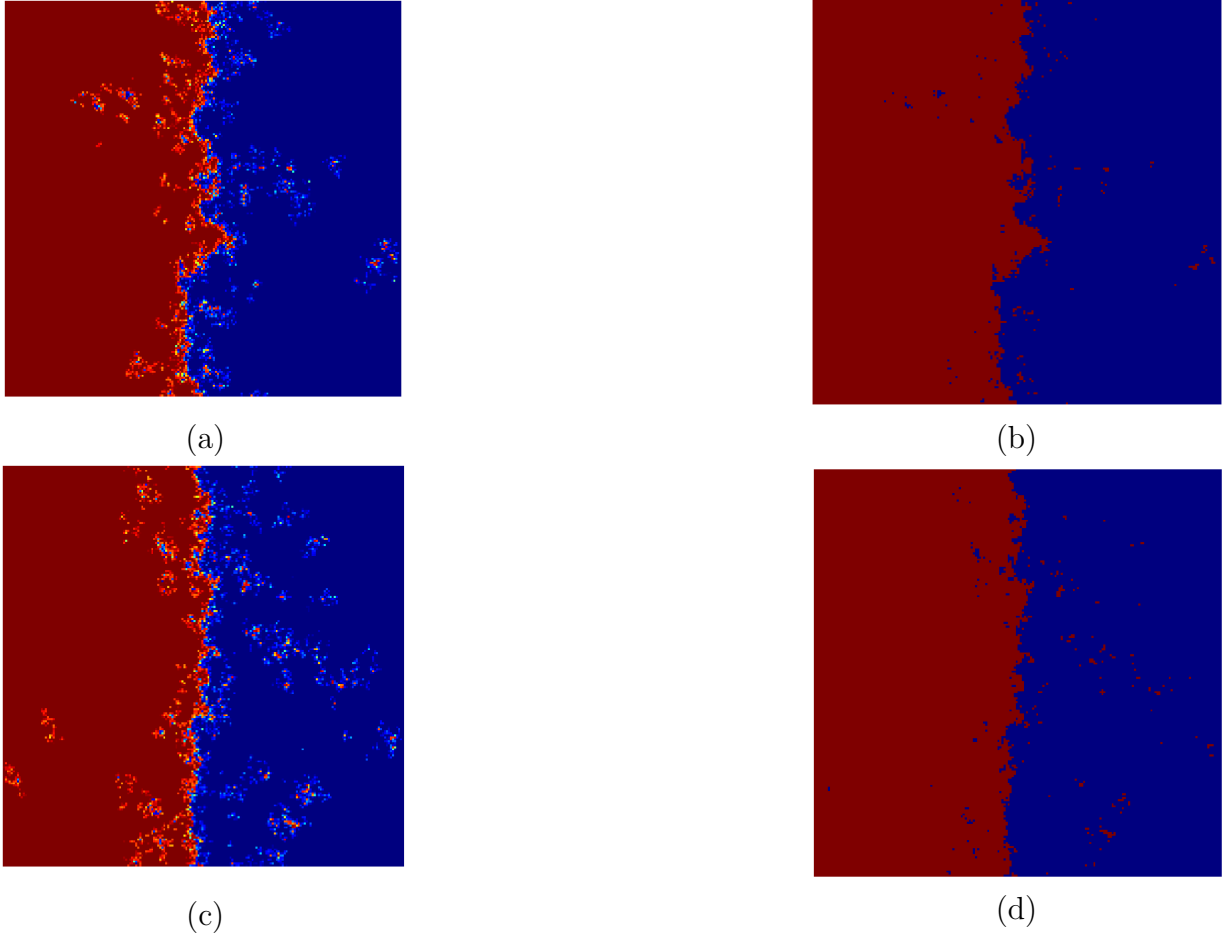


Figure 9.7: Snapshots of the MGVM with size of $L \times \omega = 200 \times 200$ where cells can take on any values in the range of $[-1, 1]$ with increment $1/100$. Color profile of cells includes RGB colors from red ($\phi = -1$) to blue ($\phi = +1$). The system initially starts with cell values $\phi = -1$ on the left and $\phi = +1$ on the right. To measure W , the actual configurations of the system (a) and (c) are transformed to (b) and (d) where cells with $\phi < 0$ respectively $\phi > 0$ are assigned the value -1 respectively $+1$. Hence measurements will be performed on (b) and (d).

MGV, we set $N = 100$ and perform simulations with different values of L . As presented

in Figure 9.8, for $L \geq 200$ the system is indeed in the scaling regime, with a value of the exponent β_w consistent with our results in the LD model.

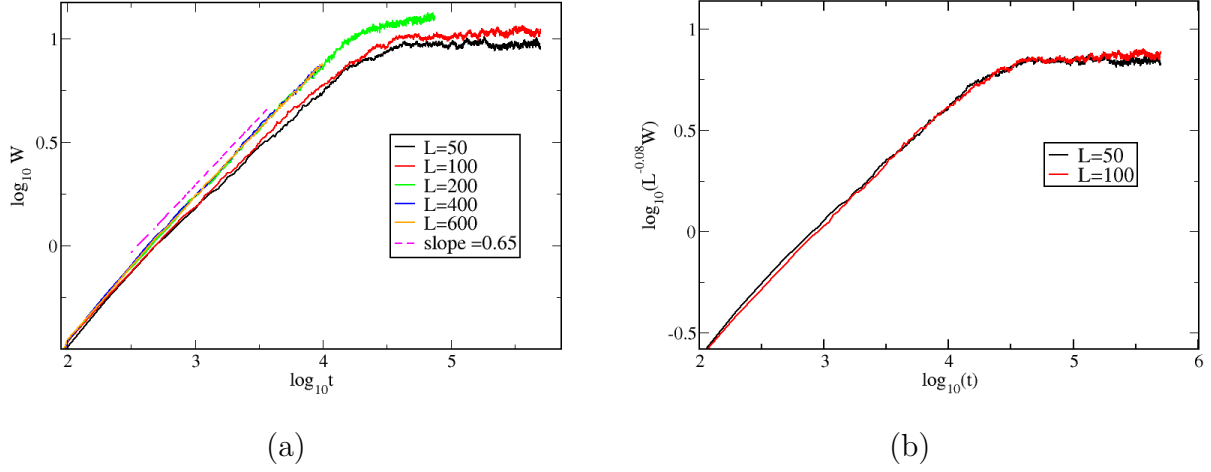


Figure 9.8: Dynamical behavior of the interface fluctuation, $W(L, t)$, in the MGVM with the system width $W = 100$ and $N = 100$. (a) For interface sizes $L \geq 200$ the growth exponent $\beta_w = 0.65$ is found which is in agreement with the β_w value obtained in the LD model. (b) We plot the scaled interface fluctuations with the value for α_w found in the LD model for $L = 50$ and $L = 100$. No stationary data are available for large values of L .

9.2 Voter-Voter interface width in the LD model.

Another interesting case is when the two sides of the system under study have opposite spin state values, but both are set to be at the voter critical point. We call this shared interface to be a Voter-Voter interface and we investigate the dynamical behavior of the fluctuations of that interface. Before we move to the details of our simulations, it is important to note that for a different model belonging to the Voter universality class the authors of [69] have studied the behavior of the interface width fluctuations. In their model, they found a power-law relationship between W and time of the form $W(t) \propto t^{\beta_w}$ with $\beta_w = 0.50(2)$. However, they did not attempt to determine the roughness exponent α_w .

In the following, we study the dynamical behavior of the interface width in the LD model set at its Voter critical point. The temperature thus is chosen to be $T_V = 1.783$ and the spins interact with only their four nearest neighbors.

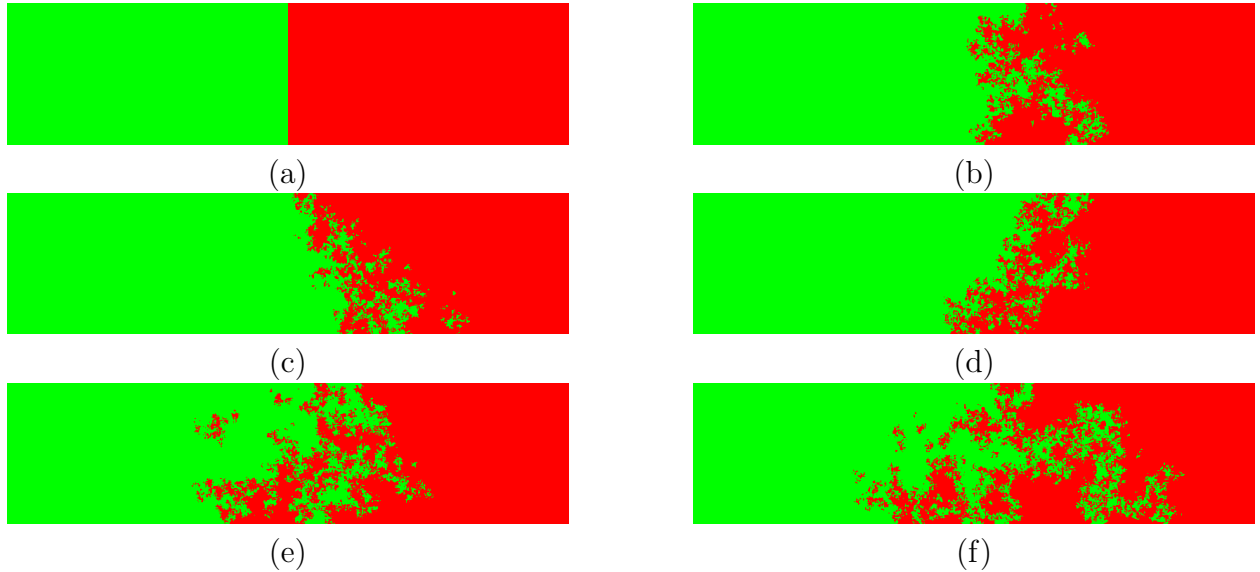


Figure 9.9: Voter-Voter interface in the LD model with the fixed size of $L \times \omega = 100 \times 800$ at different time steps. (a) The system starts with cell values $\sigma = 0$ on the left (green cells) and $\sigma = +1$ on the right (red cells). For the other configurations the times are (b) $t = 2000$, (c) $t = 5000$, (d) $t = 10000$, (e) $t = 20000$, (f) $t = 30000$. After about 1000 time steps the interface between the two states is dissolved and there is no saturation time for the system. In the above system, the green state will fill the whole system after about 50000 time steps.

Similar to the previous section, our goal here is to study dynamical behavior of $W(L, t)$ and its dependency on the interface length L . The evolution of the interface width with time indicates that $W(L, t)$ has a power-law relationship with t with the exponent $\beta_w = 0.50$ which is in agreement with the exponent reported in [69]. We also attempt to find the

roughness exponent α_w by letting the system reach its saturation stage. However, as Figure 9.9 illustrates, unlike the DP-DP interface, the Voter-Voter interface width fluctuations are drastically large. Hence the interface between the two sides of the system is dissolved and for a finite system size, one of the two states will fill out the entire lattice. This explains that the system never reaches to a saturation stage. It also explains that similar to [69], the fluctuations of the interface width grow with time in our models (Figure 9.10) for ever. Finally we note that this is similar to the interface fluctuations behavior in RD models (one has $\beta_w = 0.5$ in RD models).

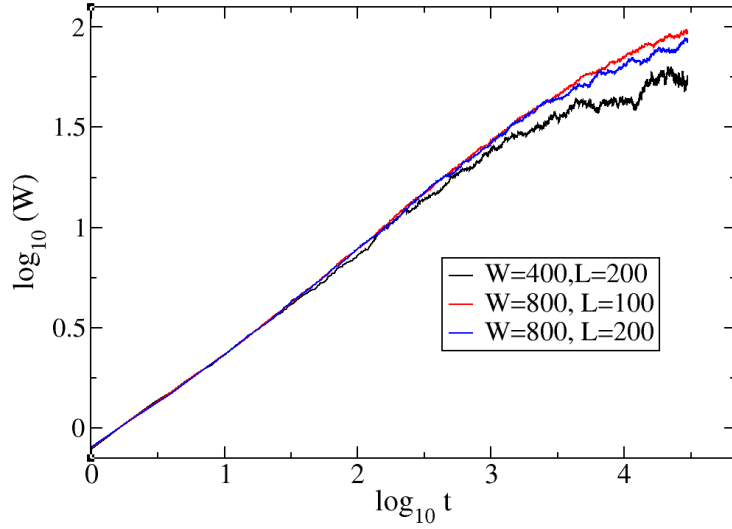


Figure 9.10: Log-log plot of the fluctuations of the Voter-Voter interface width versus time in the LD model with different system sizes. Fluctuations of the interface grow in time and do not reach saturation.

Part II

Application of Machine Learning Models in Physical Systems

Chapter 10

Introduction

Broad studies in the past years have shown that the machine learning framework is an effective tool to identify physical properties of statistical physics systems. Generally speaking, machine learning is the study of algorithms and statistical models which enables the computer system to learn and perform a specific task. Therefore almost all machine learning models include two phases; one is the learning process where the algorithm (or machine) extracts statistically important features of a given training dataset and the other one is the evaluation of its performance on a test dataset. In the process of learning, if the instances (datapoints) in the training dataset are categorized by specific labels and the machine learning task is to learn such labels, then this is called supervised learning. In fact, in supervised learning, the machine learning algorithm explores and extracts information of the instances and finds (usually a complex) relationship between those instances and their labels, i.e, predicts the labels. Once it is done, then in the next phase, the algorithm evaluates its performance on a test dataset. It is important to note that the test dataset must have the same statistical properties than the training dataset so that what the algorithm has learned from the training dataset can be validated on the test dataset. Therefore in this phase various metrics are used to evaluate the accuracy and confidence of the model in predicting the labels of the given test dataset. Besides a supervised learning, if the learning process is done over an unlabeled training/test dataset, it is called an unsupervised learning. In such processes, the algorithm extracts similar aspects of instances in the training dataset and categorize them in number of groups. For instance, one application of an unsupervised learning in physics could be categorizing the Ising model configurations and classifying them into a number of groups. So naturally the model learns to create two groups and separate the configurations based on the phase (order or disorder) they belong to. In the test phase, this algorithm takes a test configuration and determine which group it belongs to.

In the field of condensed matter, supervised and unsupervised learning methods have been used in a large number of systems from classical systems to quantum systems these applications can be grouped into two main categories.

In the first category machine learning models are applied to identify phase transitions of physical models. No matter if there is a well defined order parameter in the system or not, in many cases it has been shown that supervised and unsupervised models are able to distinguish different phases of the system under study [71–78]. Even in some cases it has been argued that for systems without explicit order parameter, machine learning can introduce new parameters to separate related phases [72]. This idea motivates us to study the configurations of the Ising model with conserved magnetization. As we know, in the Kawasaki algorithm magnetization of the system is constant and equal to the initial magnetization value. Therefore magnetization can not be a good candidate for physical order parameter. In Chapter 11 we investigate whether our neural network model output can be regarded as a quantity which behaves as an order parameter in the Kawasaki model. To this end, we produce Kawasaki model configurations as a training dataset and build a neural network with convolutional layers. The task of our model is to predict the phase (order or disorder) of a given configuration. We will argue that the output of our convolutional neural network (CNN) model can be regarded as an order parameter for the Ising configurations with conserved magnetization.

In the other category, machine learning models learn details of physical system’s configurations and accordingly generate new configurations with similar physical properties. These models are called generative models and therefore their task is to produce a dataset similar to the training dataset. To be more specific, the final goal of the generative models is to learn the hidden probability distribution functions of the training configurations and sample new configurations based on that probability distribution functions. So far in the machine learning community three generative model have been introduced: the variational auto encoder (VAE), the generative adversarial network (GAN) and finally the restricted Boltzmann machine (RBM). In the case of the Ising model authors of [80–83] have studied all of these three models and argued that the generated configurations are reliable in the sense that for a large number of generated configurations, the graphs of energy and magnetization are compatible with the original configurations. However, in Chapter 12, we argue that for the case of RBM evaluating the energy and magnetization over the generated configurations are not sufficient to conclude that they have been correctly sampled from the Boltzmann distribution. We argue that for the Ising model and generally for any condensed matter model, the correlation functions also play important roles as they reveal how clusters of the states in the system form in time and space. Therefore in order to have physically meaningful configurations, correlation function must be also taken into account. In Chapter 12 we assess the quality of this quantity in the generated configurations and evaluate whether its behavior at different temperatures is compatible with the original configurations.

Chapter 11

Supervised Learning of the Kawasaki Model Configurations

In this chapter, we study the steady state characteristics of the Ising model with conserved magnetization from the machine learning point of view. In Section 11.1 we first introduce the Ising model with Kawasaki dynamics and then simulate this model to provide a large set of Ising configurations. These configurations will be later split into training and test datasets. In Section 11.2, we explain the architecture of the machine learning model and evaluate its abilities in terms of learning the phases of the Ising model. We show that the convolutional neural networks (CNN) can identify the critical point of the Ising model with conserved magnetization. In addition to that, we discuss that the output of the CNN model can be regarded as a proxy for the order parameter of the Ising model.

11.1 Model

Here we introduce the Kawasaki algorithm and we explain how we collect the training and test data for our machine learning pipeline. With L^2 spins (with +1 or -1 for each spin) on a square lattice, the Hamiltonian of the system is given by

$$H = -J \sum_{\langle i,j \rangle}^L \sigma_i \sigma_j$$

where $\langle i,j \rangle$ means the summation is over nearest neighbors. Note that as magnetization ($M = \sum_i^{L^2} \sigma_i$) is required to be conserved through the evolution of the system, we apply the Kawasaki algorithm. For a randomly chosen pair of neighboring spins the probability of these spins to exchange is:

$$P = \min\{1, e^{\frac{-\Delta E}{\kappa_B T}}\}$$

and ΔE is the energy change of the proposed exchange and T is the temperature. Here we do not aim to study the dynamical behavior of the system, hence for the purpose of this study, from a random initial state, we let the system evolve with Kawasaki dynamics and once it reaches its steady state, we store spin configurations. To make sure that the system has reached its steady state, we monitor the change of energy as a function of time, and after the energy reaches its steady state value we start collecting the configurations (see Figure 11.1). According to the energy graphs for different temperatures in Figure 11.1, we typically start collecting the configurations of the system after 9×10^6 Monte Carlo steps (MCs), with every other 1000 MCs between configurations. Examples of spin configurations are shown in Figure 11.2. The configurations are collected from simulations with 10 different initial disordered states. For each temperature, we use 7,000 configurations to train the CNN model and 3,000 configurations as the test dataset. In these images it is clear that once the temperature is below $T = T_c$ ($T_c \approx 2.269$, with $\frac{J}{\kappa_B} = 1$), green and red clusters are separated and form on the opposite side of the lattice. After acquiring these configurations, they are stored as follows: Based on the temperature of the model, each configuration belongs to either the ordered state ($T < T_c$) or disordered state ($T > T_c$) of the Ising model. We call these states as *classes* or *labels* in the machine learning terminology. Since the labels should be transformed to numbers, we assign label 1 to the configurations in the ordered state and label 0 to those configurations in the disordered state. Note that the values of the labels are not important and that they are only used to categorize the images.

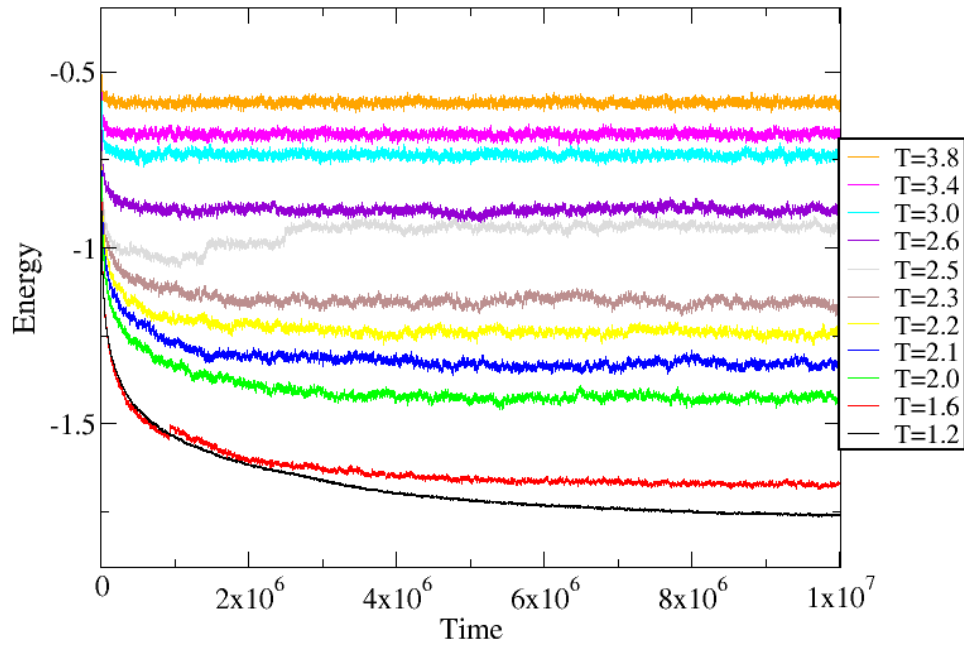


Figure 11.1: Energy values at different temperatures versus time for the Ising model with Kawasaki dynamics for system size $L=40$. Except for the data with $T=1.2$, all graphs reach their steady state after about $t = 4 \times 10^6$ (for $T=1.2$ it is about 9×10^6) Monte Carlo steps. The data are obtained from averaging over 10 different runs.

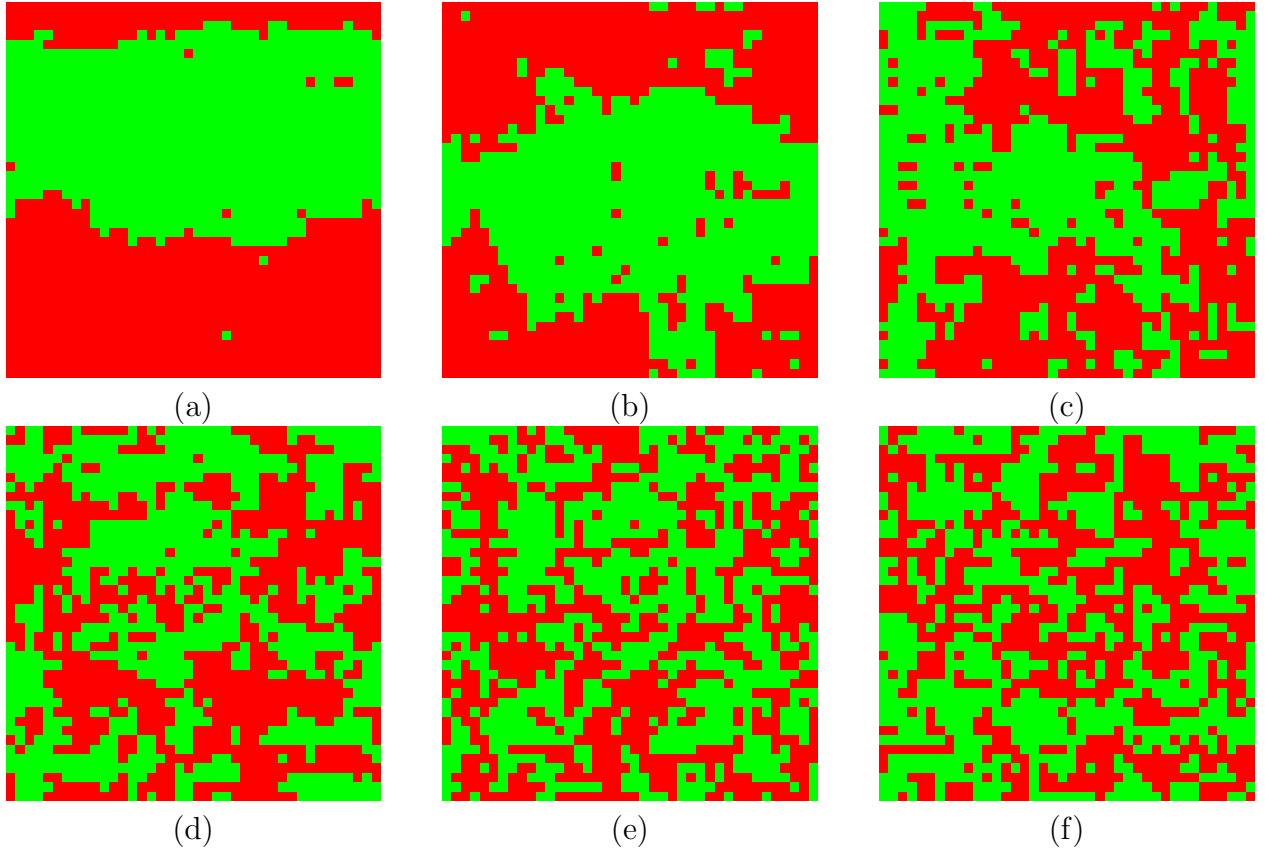


Figure 11.2: Configurations of the Ising model with conserved magnetization after equilibration at $T=2.0$ (a), $T=2.2$ (b), $T=2.6$ (c), $T=3.0$ (d), $T=3.4$ (e), and $T=4.0$ (f). The system size is 40×40 . These configurations are given to the CNN model as the training dataset.

11.2 Convolutional Neural Network

In this Section we describe the architecture of our Convolutional Neural Network (CNN) model in some details. We first explain the training and the test process of the model. Then we discuss the physical meaning of the CNN model output when it is trained on the Ising configurations.

Convolutional Neural Networks (CNN) are considered to be the most successful models in image processing problems. CNNs are able to extract important features of images, i.e, they learn boundaries of different objects in a given image and classify them accordingly. This feature of a CNN model is our motivation to apply it to the steady state configurations of the Ising model. These configurations, as it is shown in Figure 11.2, have completely different boundaries in ordered and disordered states. Therefore as a machine learning algorithm, a CNN model is a natural choice to capture the boundaries and then classify the configurations.

Here we explain how the training and test phases of our machine learning mode are carried out. The training phase is as follows: each image is the input of the CNN model. Then convolutional calculations are performed on the image (explained in details in subsection 11.2.1) and the output of the CNN model is compared with the label of each configuration through a function called *loss* function (see Section 11.2.1). Then, the parameters of the CNN model are optimized in a way that the loss function value decreases for the next images. In this way the output of the CNN model becomes closer and closer to the label value of the incoming images. This takes at least several iterations of image feeding. This is called the training phase of the model. In the test phase, new configurations are given to the classifier and it is asked to predict the label of the configurations, i.e, which phase (order or disorder) the images belongs to.

It is good to mention that studies of neural network (NN) algorithms on the configurations of the Ising model with a non-conserved magnetization reveal that even a simple fully connected neural network, with two layers, is able to learn the features of these configurations [71, 79]. Furthermore they show that the output of such models have a one-to-one relationship with the order parameter (magnetization). This observation basically implies that the above NN model learns to count the difference between the number of up and the number of down spins (which defines the magnetization) of each input configuration. However in the Ising model with Kawasaki algorithm (our case) the magnetization is constant (and 0). Hence using a fully connected neural network with few layers should not be a suitable choice here. Our investigation indeed showed that a fully connected NN model does not learn which classes the Kawasaki model configurations belongs to. This was the motivation for us to use a CNN model instead of a fully connected neural network.

Before we talk about the details of our CNN model, here we give a short introduction to CNN models.

11.2.1 Structure of convolutional neural networks

A Convolutional Neural Network is a deep learning algorithm with an input layer and an output layer, as well as multiple convolutional layers as hidden layers. Here we explain a CNN model which has the task to take in images of dog and fish and to predict their classes (there are only two classes: dog or fish). Hence for a given image, the output of the CNN model should be only two numbers which show the probability of the image being classified as dog and being classified as fish. Therefore given the image to the input layer of the CNN model, the first hidden layer with many kernel matrices (trainable parameters) takes in the input image, performs a convolutional multiplication between its kernels and input images and creates an output (see Figure 11.3).

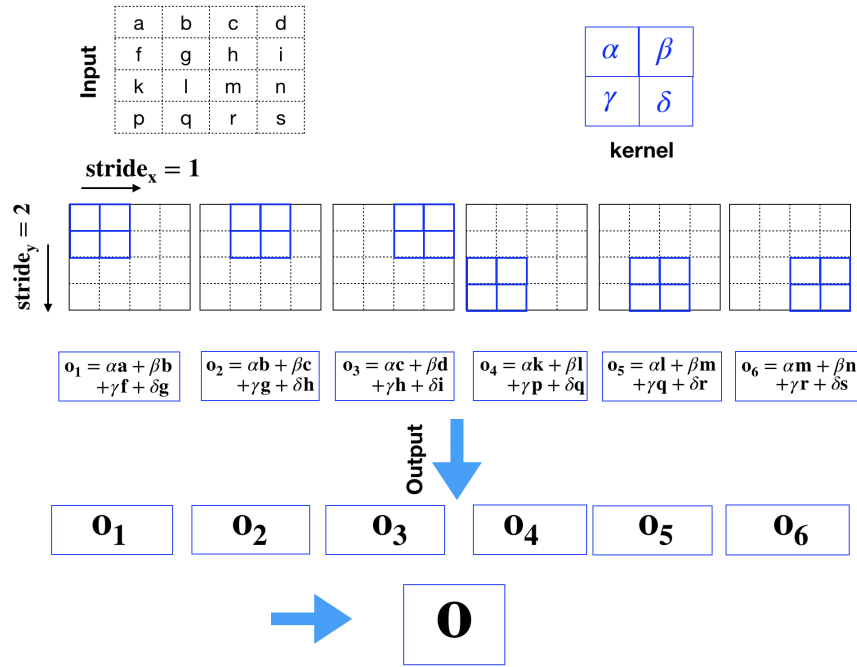


Figure 11.3: Operation of the kernel matrix on the input image. In this example the input size is 4×4 and the kernel size is 2×2 . The kernel matrix is convolutionally multiplied by the input pixels which results in an output vector o with 6 elements. Strides values (stride_x and stride_y) show the number of steps taken in the horizontal and vertical directions.

To ensure that the output matrix elements (o in the Figure 11.3) are in control, a function called *activation* function is used to normalize the output function to 1. This is basically a general rule for any other layer in a neural network model. The normalized output then is passed to the second layer as an input and another set of convolutional multiplications are performed. This process continues in all other convolutional layers until the last layer. As we mentioned earlier the task of this model is to produce two numbers as the probabilities

of the two classes (sometimes classes are called labels). Therefore the last layer decreases the number of nodes in the final convolutional layer to only two nodes. However as pointed out, another activation function should be used to normalize the sum of these two nodes values to 1, i.e, to transform them to become probability values. An example of a CNN model with the above task is shown in the Figure 11.4. It shows how the model assigns a pair of probabilities to an image as being classified in either of the classes. Finally we discuss the important steps of the training process in a CNN (or any neural network) model: Consider that the CNN model is a high non-linear function F with parameters θ which takes an image x and produces an output y , i.e, $y = F(\theta, x)$. Here θ is the set of parameters in all of the CNN layers. In our case $y = (y_1, y_2)$ is a probability vector with only two values. Now assume that images of fish are labeled as 0 and those of dogs are labeled as 1. For N training configurations, the loss function which measures the distance between the model prediction y from the image label Y (0 or 1) is defined as follows:

$$\begin{cases} \text{loss} = -\sum_i^N \log(y_1) & \text{if } Y = 0 \\ \text{loss} = -\sum_i^N \log(y_2) & \text{if } Y = 1 \end{cases}$$

The values y_1 or y_2 are the functions of the CNN parameters. Therefore in the training phase, given the loss function to an optimizer (such as gradient descent) the optimizer minimizes the loss function value by finding the optimized values of θ in F . Once the parameters θ reach their optimal values, the training phase is stopped and the model is ready to be tested.

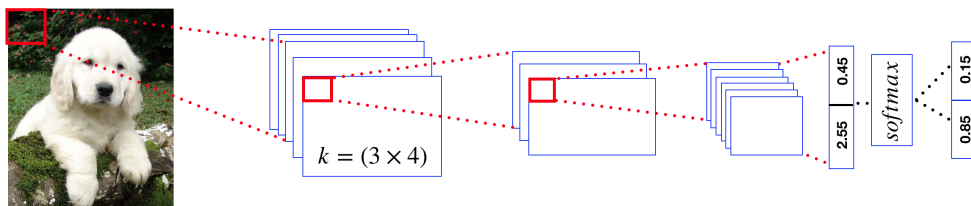


Figure 11.4: An example of a Convolutional Neural Network. The above model has three convolutional layers. The first layer includes 5 filters and each filter has its own kernel matrix with size (3×4) . The last convolutional layer is connected to the dense layer with only two nodes. As an example, the nodes values are 2.55 and 0.45. These numbers are normalized to 1 by the softmax activation function.

11.2.2 CNN Model training on Ising configurations.

The structure of the CNN model we use here is presented in Figure 11.5. As explained, the model takes in the input image with size $L \times L$ and determines with what probability the image is +1 (ordered) or 0 (disordered).

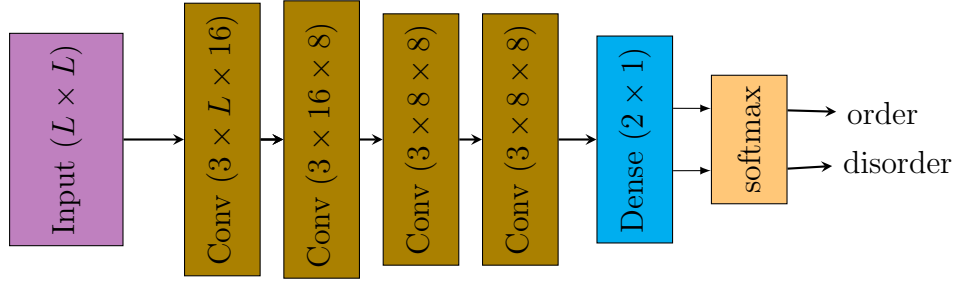


Figure 11.5: Architecture of the CNN model applied to Ising configurations. A softmax function is used as the activation function for all convolutional layers. Properties of each layer are written inside of the layer rectangle. For example, the first convolutional layer has 16 filters with kernel size (3×3) , whereas the second convolutional layer has only 8 filters with the same kernel size as the first layer. Finally the dense layer has only two nodes which are normalized by the softmax activation function.

Training of the above model is carried out with about 7000 configurations at each temperature from $T = 1.0$ to $T = 4.0$. After the training phase is done, we evaluate our model by measuring the accuracy of the model in predicting the labels of the test dataset configurations. The accuracy is defined as follows:

$$Accuracy = \frac{\text{number of configurations which are classified correctly}}{\text{total number of configurations}} \quad (11.1)$$

We test 3000 configurations of the Ising model at each temperature to measure the accuracy. Table 11.1 which shows the accuracy at different temperatures, indicates that the performance of the model in predicting the labels of the images around $T = T_c$ is not as good as for other temperature. This is understandable since the configuration images near the critical point are similar and it makes it hard for the CNN model to identify them correctly. Additionally, we define another quantity called *confidence score* which is defined as follows:

$$\text{Confidence score} = \frac{\text{number of configurations which are predicted in the ordered class}}{\text{total number of configurations}} \quad (11.2)$$

Figure 11.6 is the plot of the confidence score versus temperature. The behavior of this quantity is understandable once we notice that below and far from the critical temperature almost all configurations are predicted as label 1. The information in Table 11.1 indicates that the model performs very well in this region. Similarly, almost none of the images above and far from the critical temperature are predicted as label 1. It is interesting to note that the confidence score curves for different sizes L cross each other at the critical point $T \approx 2.26$. Note that we chose few temperatures near T_c for the training stage whereas in the previous methods of learning [72] mostly configurations at temperatures close to T_c were usual. This suggests that the confidence score can be regarded as a quantity which is able to identify

Temperature	1.0	1.2	1.4	1.6	1.8	2.0	2.1	2.2	2.3	2.4
Accuracy	1.0	1.0	1.0	1.0	0.98	0.90	0.84	0.58	0.61	0.83
Temperature	2.5	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0	4.2
Accuracy	0.94	0.99	0.998	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 11.1: Accuracy of the CNN model at different temperatures. These accuracies are measured using the test Ising configurations with $L = 20$. Close to the critical temperature $T \approx 2.26$, at $T = 2.2$ the accuracy decreases to its lowest value 0.58.

the critical point of the Ising model. Also its behavior below and above the critical point is similar to the expected behavior of an order parameter in the model.

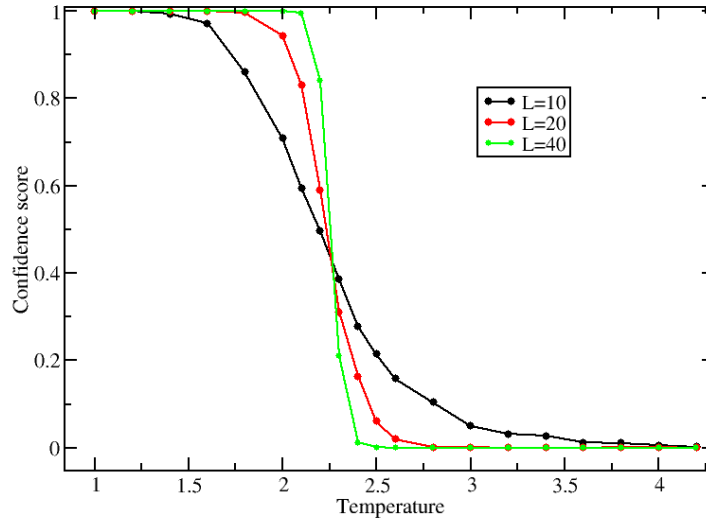


Figure 11.6: Output of CNN model as a function of temperature. For different values of system size L all graphs intersect with each other at the critical point. This suggests that the output of the CNN model can be considered as the order parameter.

Chapter 12

Restricted Boltzmann Machine

In this chapter we study the application of the restricted Boltzmann machine (RBM) for physical systems. As an unsupervised learning model, the RBM is one of the four major types of generative models that learn the unknown probability distribution function $\mathbb{P}$ from which input instances are assumed to be sampled. RBMs usually contain one layer of neurons (referred as the visible layer V) that takes in the input instances and at least one hidden layer (H) of neurons with binary values, 0 and +1. These layers are connected to each other through a matrix of parameters (W). In an RBM, the likelihood function P is constructed in the form of $P \propto e^{-E(V,H)}$ and optimized to become as similar as possible to $\mathbb{P}$. In this formula, $E(V, H)$ is called the energy function. Here we investigate the ability of the RBM in learning the underlying probability distribution function (Boltzmann distribution) which gives weights to the Ising model configurations based on their total energy values. To do so, we prepare two separated training datasets of Ising configurations, one with conserved magnetization and one with non-conserved magnetization. After training, we measure physical quantities to evaluate the quality of newly generated configurations. This is not the first time that generated configurations from the RBM are studied. In fact, the authors of [81] and [85] trained different architectures of the RBMs, such as deep Boltzmann machine (DBM) over Ising configurations with non-conserved magnetization and measured physical quantities of the generated configurations. In their results, they showed that the magnetization and the energy (as functions of temperature) in the generated configurations are in agreement with their counterparts in the original training dataset. However, here we stress that along with magnetization and energy, the correlation function in the generated dataset must be taken into account. This is essential because both the energy and magnetization are not very sensitive to the size of the clusters in the configurations. Also, we train an RBM over Ising configurations with conserved magnetization ($M = 0$) and discuss physical properties of the generated configurations. This is particularly interesting because the input configurations satisfy a constraint on the magnetization but RBM does not necessarily respect this constraint. Therefore it is an open question to what extent the RBM learns an existing constraint in the input configurations.

This Chapter is organized as follows: Section 12.1 is devoted to the details of the RBM algorithm as a machine learning generative model. In Section 12.2 we describe the process of creating training datasets and, finally, in Section 12.3 we look into the results of the RBM output after training separately over Ising configurations with either non-conserved or conserved magnetization.

12.1 Architecture of the RBM

The RBM is considered as one of the major generative models in the machine learning community. As one type of Gibbs network, the RBM was first introduced in [87] and soon after its power in generating new samples with high quality was recognized.

As a graphical model, the RBM model includes two different types of layers; one is the visible layer V , and the other one is the hidden layer H . RBM models usually have only one hidden layer, but it is quite possible to create a DBM by adding more hidden layers in the architecture (see Figure 12.1).

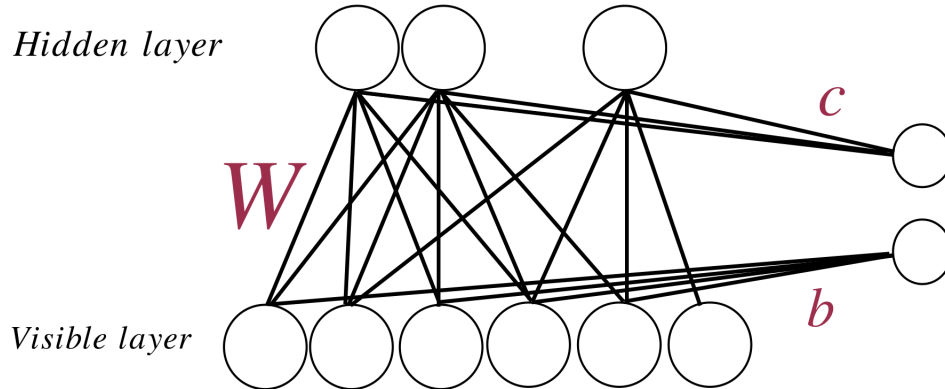


Figure 12.1: A sample architecture of the RBM. The number of hidden nodes is usually smaller than the number of the visible nodes. The right hand side nodes represent bias terms in the Equation (12.2).

We denote the nodes in the input layer as $V = \{v_1, v_2, \dots, v_n\}$ and the nodes in the hidden layer as $H = \{h_1, h_2, \dots, h_m\}$. In all versions of RBM the nodes $v_1, \dots, v_n, h_1, \dots, h_m$ take on binary values 0 and 1. The joint probability distribution in the RBM model is defined as follows:

$$P(V, H) = \frac{1}{Z} e^{-E(V, H)} \quad (12.1)$$

where $Z = \sum_{V, H} e^{-E(V, H)}$ is the partition function of the model. Based on the Equation (12.1), in the training phase the model learns the quantity $E(V, H)$. In the RBM model, the

explicit definition of the energy is given by:

$$E(V, H) = -V^T b - c^T H - V^T W H \quad (12.2)$$

where b , c , W are the vectors and the matrix which include parameters of the model that are trained through an optimization scheme. Therefore the goal of training the RBM is to find the optimal values of the above parameters. After finding these values, the process of generating new configurations is done by sampling from $P(V, H)$. Here for the sake of simplicity, we introduce $\theta = \{b, c, W\}$. Also, it is suitable to write $P(V, H)$ as a distribution over V for given θ and H , i.e, $P(V|\theta, H)$. $P(V|\theta, H)$ is called the likelihood (LL) function of the values $v_1, \dots$. In our case, its dependency on θ can be simplified more:

$$P(V|\theta, H) = \frac{1}{Z} e^{-E(V, H)} = \frac{1}{Z} e^{(V^T b + c^T H + V^T W H)} = \frac{1}{Z} e^{(V^T b)} \prod_{j=1}^M e^{(c_j h_j + v_i W_{ij} h_j)} \quad (12.3)$$

As noted, each node in the hidden layer H can take on binary values 0 and 1. Therefore summing over the values of h_i s, the Equation (12.3) becomes:

$$P(V|\theta) = \sum_{\text{for all } h} P(V|\theta, H) = \sum_{h_1=0}^1 \dots \sum_{h_M=0}^1 \left(\frac{1}{Z} e^{(V^T b)} \prod_{j=1}^M e^{c_j h_j + v_i W_{ij} h_j} \right) \quad (12.4)$$

After carrying out the above summations, the likelihood function $P(V|\theta)$ will have the following shape:

$$P(V|\theta) = \frac{1}{Z} e^{(V^T b)} \prod_j (1 + e^{c_j + V^T w_j}) \quad (12.5)$$

where w_j represents each column in W . Having the explicit form of the likelihood function, we first take the *log of the likelihood* (LLL) and then perform the optimization method over the parameter set θ . In the next Section we explicitly show how the gradients of the LLL are taken with respect to the parameters in θ .

Gradients of model parameters

In the RBM model the optimization method is based on the gradient descent. Gradient descent here determines how to update the parameters in θ in each iteration. In our case:

$$\theta_{t+1} = \theta_t - \Delta\theta, \quad \Delta\theta = \alpha \nabla \log P(V|\theta) \quad (12.6)$$

where α is called the learning rate and the gradient operator acts on all vectors and matrix elements in $\theta = b, V, W$. The explicit form of $\log P(V|\theta)$ can be written as:

$$\log P(V|\theta) = -\log(Z) - F(V, \theta) \quad (12.7)$$

where $F(V, \theta)$ is the “free energy” and equals to:

$$F(V, \theta) = \log \left(\sum_{j=1}^M (1 + e^{c_j + V^T w_j}) \right)$$

Therefore the derivative of the LLL becomes:

$$\frac{\partial \ln P(V|\theta)}{\partial \theta} = \frac{1}{Z} \frac{\partial Z(\theta)}{\partial \theta} - \frac{\partial F(V, \theta)}{\partial \theta} \quad (12.8)$$

The first term in the right hand side of the above equation is:

$$\begin{aligned} \frac{1}{Z} \frac{\partial Z(\theta)}{\partial \theta} &= \frac{1}{Z} \sum_{v'} e^{-F(v', \theta)} \frac{\partial F(v', \theta)}{\partial \theta} = - \sum_{v'} P(v', \theta) \frac{\partial F(v', \theta)}{\partial \theta} \\ &\implies \frac{1}{Z} \frac{\partial Z(\theta)}{\partial \theta} = \left\langle \frac{\partial F(v', \theta)}{\partial \theta} \right\rangle_{v'} \end{aligned}$$

This expression has an important consequence in the training of an RBM. Note that unlike the typical forms of LLL functions in machine learning models such as CNN or RNN, any averaging process is taken not over the training dataset but over all possible inputs (denoted by v'). To make our point more clear, let us provide an example. For the input configuration with the size 20×20 the variable layer size is 400. Every single neuron of those 400 can be either 0 or 1 and so there are 2^{400} possible configurations. It is therefore impossible to take the above average ($\langle \rangle_{v'}$). However, to make an approximation of this term, one can pick some of those possible inputs and do sampling over all possible configurations. This method is called Gibbs sampling and will be explained later in this Section. Considering the explicit form of the gradient of the LLL:

$$\frac{\partial \ln P(V|\theta)}{\partial \theta} = \left\langle \frac{\partial F(v', \theta)}{\partial \theta} \right\rangle_{v'} - \frac{\partial F(V, \theta)}{\partial \theta} \quad (12.9)$$

it is not hard to calculate the second term. Using the explicit form of $F(v, b, W, c)$:

$$F(v, b, W, c) = v^T b + \ln \left(\sum_{j=1}^M (1 + e^{(c_j + v^T w_j)}) \right)$$

its derivatives with respect to the parameters are:

$$\begin{aligned} \frac{\partial F}{\partial b_j} &= v_j \\ \frac{\partial F}{\partial c_j} &= \frac{e^{c_j + v^T w_j}}{\sum_k (1 + e^{c_k + v^T w_k})} = \sigma(c_j + v^T w_j) \equiv y_j(v) \\ \frac{\partial F}{\partial W_{ij}} &= \frac{v_j e^{c_j + v_i W_{ij}}}{\sum_k (1 + e^{c_k + v_i W_{ik}})} = v^T \sigma(c_j + v^T w_j) \equiv v_i y_j(v) \end{aligned}$$

So far we have only considered one training input as V in our calculations. To include the whole training dataset in the training phase, we should take an average over training configurations. With N configurations:

$$\begin{aligned}\langle \frac{\partial F}{\partial b_j} \rangle &= \frac{1}{N} \sum_v \frac{\partial F}{\partial b_j} = \langle v_j \rangle_v \\ \langle \frac{\partial F}{\partial b_j} \rangle &= \frac{1}{N} \sum_v \frac{\partial F}{\partial b_j} = \langle y_j(v) \rangle_v \\ \langle \frac{\partial F}{\partial b_j} \rangle &= \frac{1}{N} \sum_v \frac{\partial F}{\partial b_j} = \langle v_i y_j(v) \rangle_v\end{aligned}$$

Using the above expressions, the final forms of the gradient descents formula (Equation (12.6)) will be:

$$\begin{aligned}b_{n+1} &= b_n - \alpha \Delta b \\ c_{n+1} &= c_n - \alpha \Delta c \\ W_{n+1} &= W_n - \alpha \Delta W\end{aligned}$$

with:

$$\begin{aligned}\Delta b &= \frac{1}{N} \sum_v \frac{\partial \ln(P(v|\theta))}{\partial b} = \langle v \rangle_{v'} - \langle v \rangle_v \\ \Delta c &= \frac{1}{N} \sum_v \frac{\partial \ln(P(v|\theta))}{\partial c} = \langle y(v) \rangle_{v'} - \langle y(v) \rangle_v \\ \Delta W &= \frac{1}{N} \sum_v \frac{\partial \ln(P(v|\theta))}{\partial W} = \langle v y^T(v) \rangle_{v'} - \langle v y^T(v) \rangle_v\end{aligned} \tag{12.10}$$

Once the training phase is done and the optimized values of the parameters (indicated as b^*, c^*, W^*) are found, we generate new configurations. In order to do that, we find the likelihood of the configuration $v \in \{v_1, v_2, \dots, v_n\}$. From the Equation (12.3) the probability of $P(v_j|\theta, H)$ is obtained:

$$\begin{aligned}P(v_j = 1|\theta, H) &= \frac{\exp(b_j + h_i w_{ij})}{1 + \exp(b_j + h_i w_{ij})} \sum_{v_k} \prod_{k \neq j} \frac{\exp((b_k + h_i w_{ik}) v_k)}{1 + \exp(b_k + h_i w_{ik})} \\ P(v_j = 1|\theta, H) &= \frac{\exp(b_j + h_i w_{ij})}{1 + \exp(b_j + h_i w_{ij})} = \sigma(b_j + H w_j)\end{aligned} \tag{12.11}$$

To determine the value of site 'j' in a generated configuration, $P(v_j = 1|\theta, H)$ will be compared with a random probability value from $p \in (0, 1)$.

Gibbs sampling

As described earlier, the first term in the right hand side of Equation (12.9) needs to be sampled from all possible inputs v' . The sampling procedure is shown in Figure 12.2 and here are the steps which should be taken:

1. Take a configuration from the training dataset as V with elements $(v_1, \dots, v_n)$.
2. Based on the probability $P(h_j = 1|\theta, V)$, determine the nodes in the hidden layer H . Similar calculation as in the Equation (12.11) reveals the explicit form of this probability for nodes h_j :

$$P(h_j = 1|\theta, V) = \frac{\exp(c_j + w_{ij}v_i)}{1 + \exp(c_j + w_{ij}v_i)} = \sigma(c_j + w_j V^T) \quad (12.12)$$

3. To create a sampled configuration, for each v_j calculate probability $P(v_j = 1, \theta, H)$
4. Use the configuration in step 3 and go back to step 2.

This procedure now has a loop of actions and can be continued after step 3. Until step 3 there is only one iteration of the loop and so $k=1$ (see Figure 12.2). Gibbs sampling can be done with different initial values of V to produce many sampled configurations over v' . This process is not computationally expensive and the sampled configurations can be considered as a reliable representation of all possible configurations v' in the Equation (12.9). Having these samples of configurations helps us to take the average in the first term of the Equation (12.9).

12.2 Training configurations

We explain the Monte Carlo simulation process of collecting the training datasets. These dataset will be used to train our restricted Boltzmann machine.

Ising configurations with conserved magnetization

We train our RBM on two separated datasets of Ising configurations. The first dataset includes the Monte Carlo configurations of the Ising model with conserved magnetization. The process of creating these configurations is explained in details in Section 11.1, hence we do not repeat it here. We only add that these configurations are created in two system sizes $L = 4$ and $L = 8$.

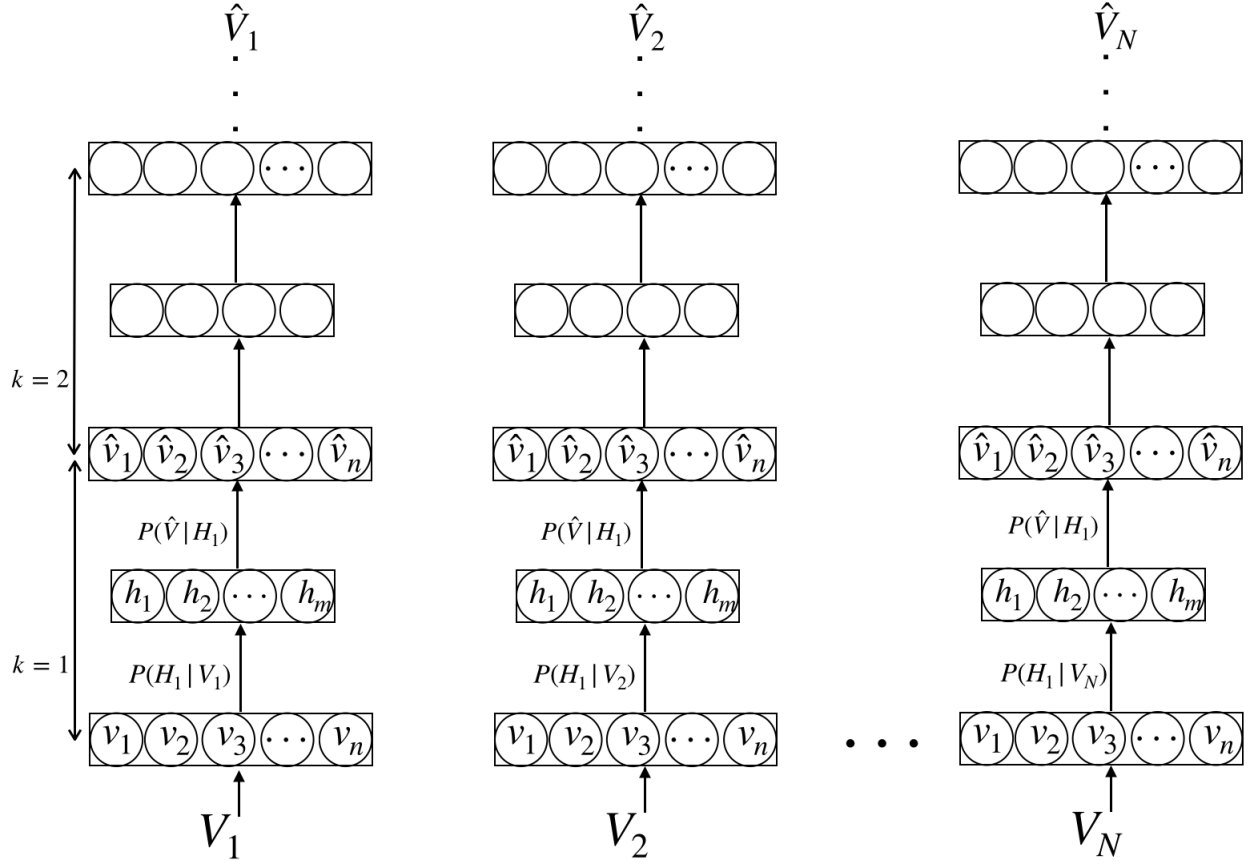


Figure 12.2: Gibbs sampling in an RBM model from the initial values V_1 to V_N . In the case of $k=1$, the hidden layer H is calculated by $P(H_1|V_1)$ and $\hat{V} = \{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n\}$ layer is obtained from $P(\hat{V}|H_1)$. Same process is done to sample $\hat{V}_2$ from the initial value V_2 . These $\hat{V}$ s are the sampled configurations via Gibbs sampling.

Ising configurations with non-conserved magnetization

The second dataset of configurations with non-conserved magnetization is defined as a two dimensional square spin lattice with the size of L . These spins can take on two values -1 or $+1$. The interaction between these spins on the lattice is defined by the Hamiltonian H , defined as:

$$H = - \sum_{i,j \in nn(i)} S_i S_j \quad (12.13)$$

where $nn(i)$ are the first nearest neighbor sites of i . To simulate this model, we use the Metropolis algorithm. In the Metropolis algorithm a randomly chosen spin is flipped to its

new state with the flip probability P :

$$P = \min\{1, \exp(-\Delta H/\kappa_B T)\} \quad (12.14)$$

with the units chosen so that $\kappa_B = 1$. We simulate the system with two different system sizes, one with $L = 4$ and one with $L = 8$. In both cases we initialize the system with a random configuration. The system evolves with the flip probability rule P . After about $t = 5000$ (Monte Carlo time steps) the system reaches its equilibrium state (steady state) and we start taking data between $t = 10,000$ and $t = 20,000$. To remove any unwanted correlation between the configurations, we pick every 50th configurations, so as a result, we will have 200 configurations. We also run the simulation for 50 times and therefore we end up with 10,000 configurations. Finally we repeat this experiment for different temperatures T in the range of 1.4 to 4.0. Figure 12.3 shows a sample of these Monte Carlo configurations that will be used as the training input of the RBM.

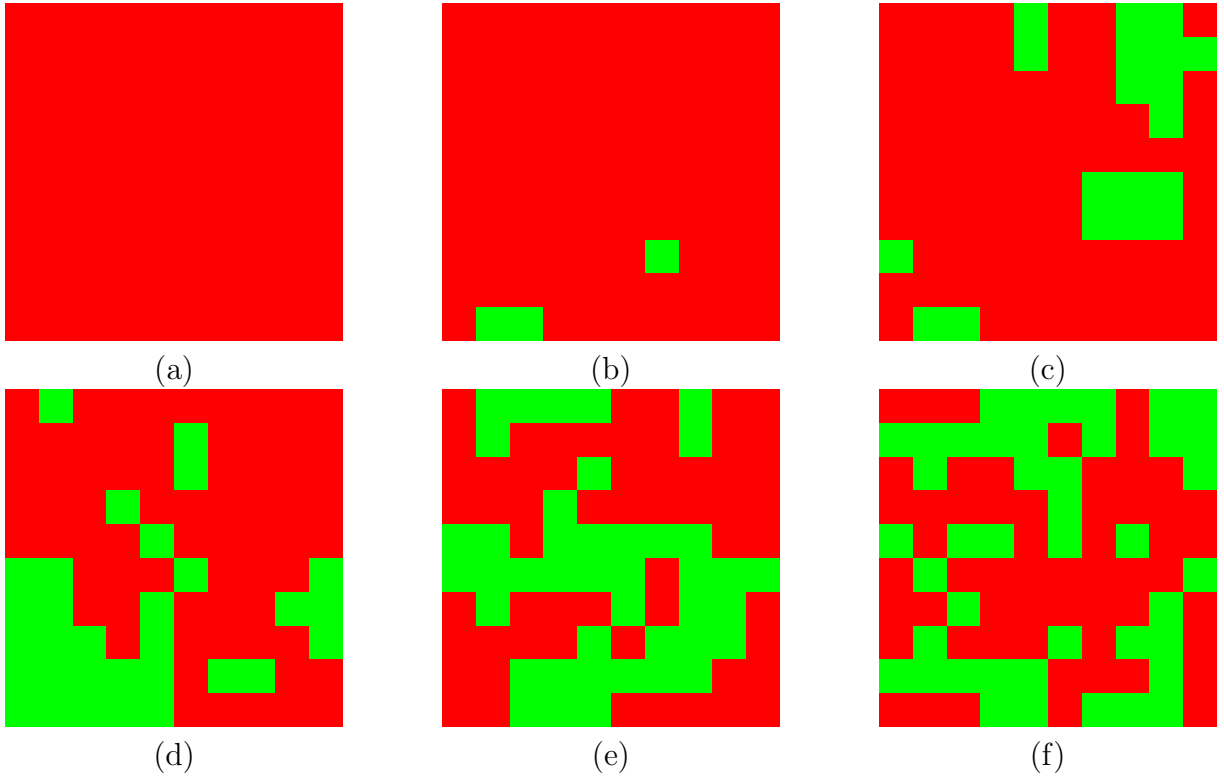


Figure 12.3: Configuration of the Ising model with the system size 8×8 after equilibration at $T=1.0$ (a), $T=1.4$ (b), $T=1.8$ (c), $T=2.269$ (d) $T=2.6$ (e), $T=3.0$ (f). Spins with value +1 and 0 are colored green and red respectively.

12.3 Learning the RBM model

In order to train the RBM model, we choose the visible layer nodes with size $L \times L$ that matches with the size of the training images. As the nodes take on only values 1 or 0, the configurations with spin values 1 to -1 are transformed to 1 and 0. As shown in [80,81], the hidden layer dimension is convenient to be identical to the visible layer, hence we set it to L^2 . Also our initial results indicate no improvements with larger hidden layers. Once the visible layer is fed with input configurations, the quantities in Equation (12.10) will be evaluated and the weights (parameters) of the model updated. Note that to evaluate $y_j(v)$, there is no need to know the values of the hidden layer (h_i s), however for the Gibbs sampling process the h_i s are required (h_i s are calculated as explained in Section 12.1). In our simulation, the Gibbs sampling iteration $k = 1$ and the learning rate $\alpha = 5 \times 10^{-3}$. We continue the training phase for 5×10^3 epochs (iterations).

When the model parameters $\{b, c, w\}$ are updated to their optimal values $\{b^*, c^*, w^*\}$, the training phase should be stopped. However the optimal values are not known and so they can not be used to stop the training process. In fact the lack of a good stopping criteria for training is a major issue in RBMs. In other generative models, typically one monitors the LLL function defined in the models and when it converges to a stable value, the training phase is stopped. This idea is not practical in RBMs as the LLL function (Equation (12.5)) includes the function Z and as we discussed, it is impossible to accurately calculate Z . Another approach introduced in [88] suggests to closely watch when the model parameters do not change significantly, i.e $\Delta b \approx 0$, $\Delta c \approx 0$, and $\Delta W \approx 0$. This method is useful but not ideal as the second terms in all Δ 's (Equation (12.10)) are computed over a sample of possible configurations and hence is inaccurate. Nonetheless, we use the latter idea as the stopping criteria of the training phase. After training, we start generating new configurations from the model as follows: we first create a random configuration with the size of $L \times L$ and after feeding this to the visible layer, the values of nodes in H are calculated using Equation (12.12). Again, with the known hidden layer H , the node values of the visible layer V are updated to their new values according to the Equation (12.11). We continue this loop of updating nodes in the hidden and visible layers for 1000 times and store the final values of the visible layer. This stored layer is now considered as a generated configuration. We also repeat the experiment 10000 times to generate the same number of configurations as we produced to train the model. Figure 12.4 displays the initial and final configuration of a generated configuration via RBM model for a few temperatures.

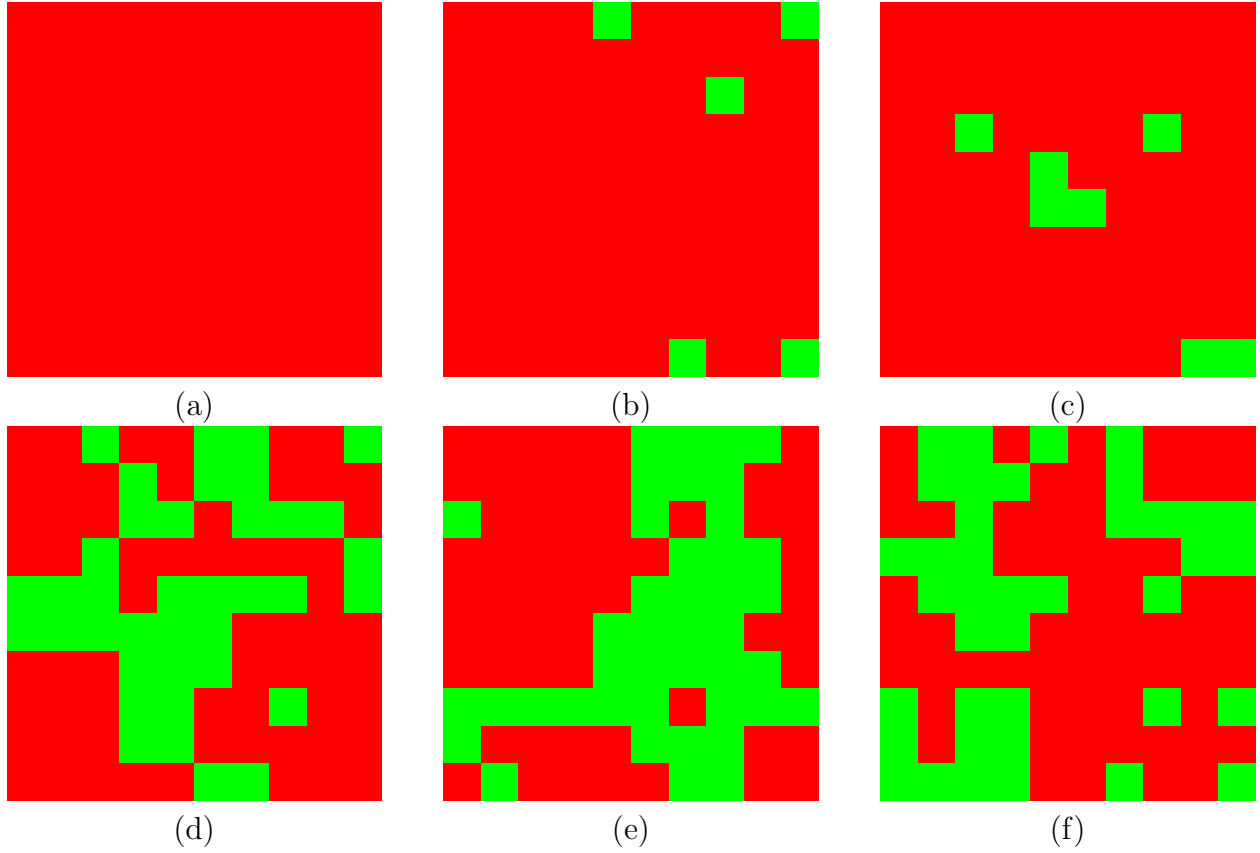


Figure 12.4: Generation of new configurations with the RBM model at different temperatures. The temperatures of the configurations are $T = 1.2$ (a), $T = 1.9$ (b), $T = 2.0$ (c), $T = 2.269$ (d), $T = 2.4$ (e) and $T = 2.6$ (f).

After obtaining the generated dataset, we measure three quantities, magnetization, energy and correlation function and compare them with their values over the Monte Carlo configurations. These quantities are defined as follows:

$$\begin{aligned}
 M &= \frac{1}{N} \langle \sum_{i=1} S_i \rangle \\
 E &= - \sum_{i=1} \sum_{j \in \text{nn}(i)} \langle S_i S_j \rangle \\
 C(r) &= \frac{1}{N} \sum_i \langle S_i S_{i+r} \rangle - \langle S_i \rangle \langle S_{i+r} \rangle, \quad |i - j| = r
 \end{aligned} \tag{12.15}$$

with $N = L^2$.

Learning configurations with non-conserved magnetization

In Figure 12.5, we measure the magnetization M in the generated datasets as well as the input datasets for two system sizes $L = 4$ and $L = 8$. As we would expect from the previous studies, there is a good agreement on the value of magnetization at different temperatures.

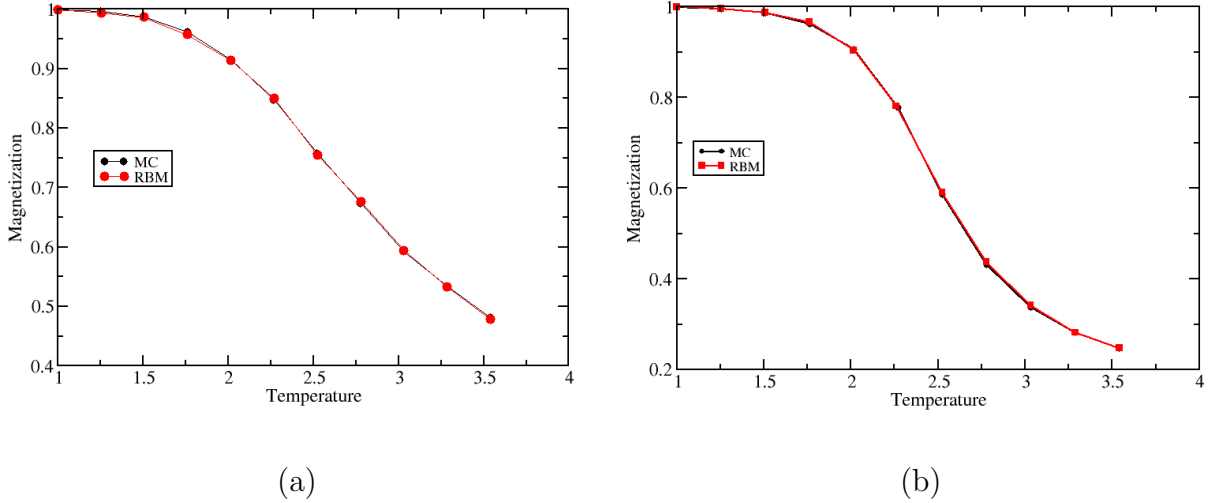
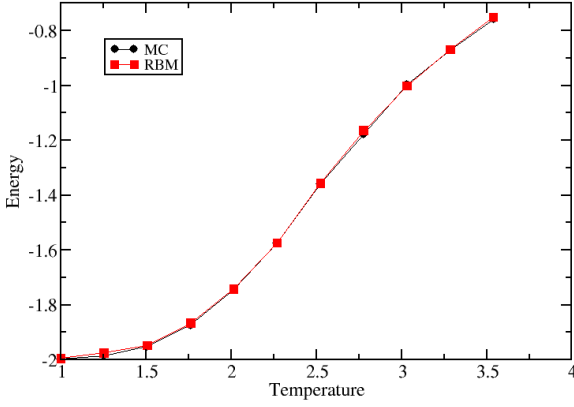
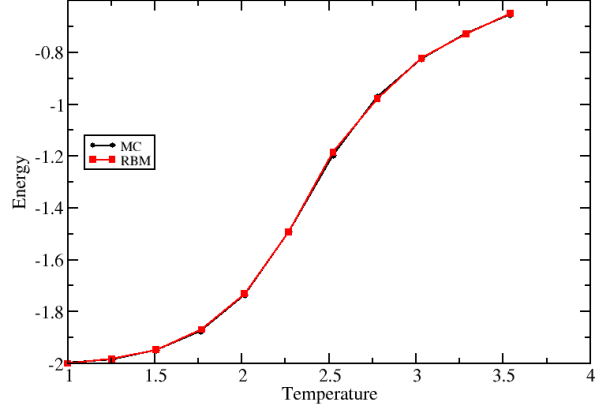


Figure 12.5: Magnetization graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations. The quality of RBM graphs indicate that the model has learned to generate images with correct number of +1 and -1 spins.

Energy is the second quantity we measured. Our results for both system sizes are presented in Figure 12.6 which again imply a nice agreement between the generated data and input data at all temperatures. Considering the energy as a special two point function, unlike the magnetization, it depends on the form of the clusters in the configurations. Therefore energy values close to the critical temperature might lead us to this conclusion that RBM has learned important physical information about the spins clusters. However we show in the following that this is not completely correct.



(a)



(b)

Figure 12.6: Energy graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations.

For $L = 4$, the correlation function $C(r)$ measurements on generated data (shown in Figure 12.7) are completely consistent with same measurements on the input data. Therefore one can conclude that generated configurations for $L = 4$ are statistically consistent with the physical configurations produced by Monte Carlo simulation. However, for $L = 8$, this consistency does not hold. In fact we observe a systematic error between the correlation function values in the generated data and the input data. Specially, at temperatures close to the critical point ($T \approx 2.269$), this gap becomes larger. We thus conclude that although the generated configurations are physically reliable for $L = 4$ but this is not always true for larger system size $L = 8$. As a result, we stress that full measurements over relevant physical quantities are required to evaluate the quality of generated configurations.

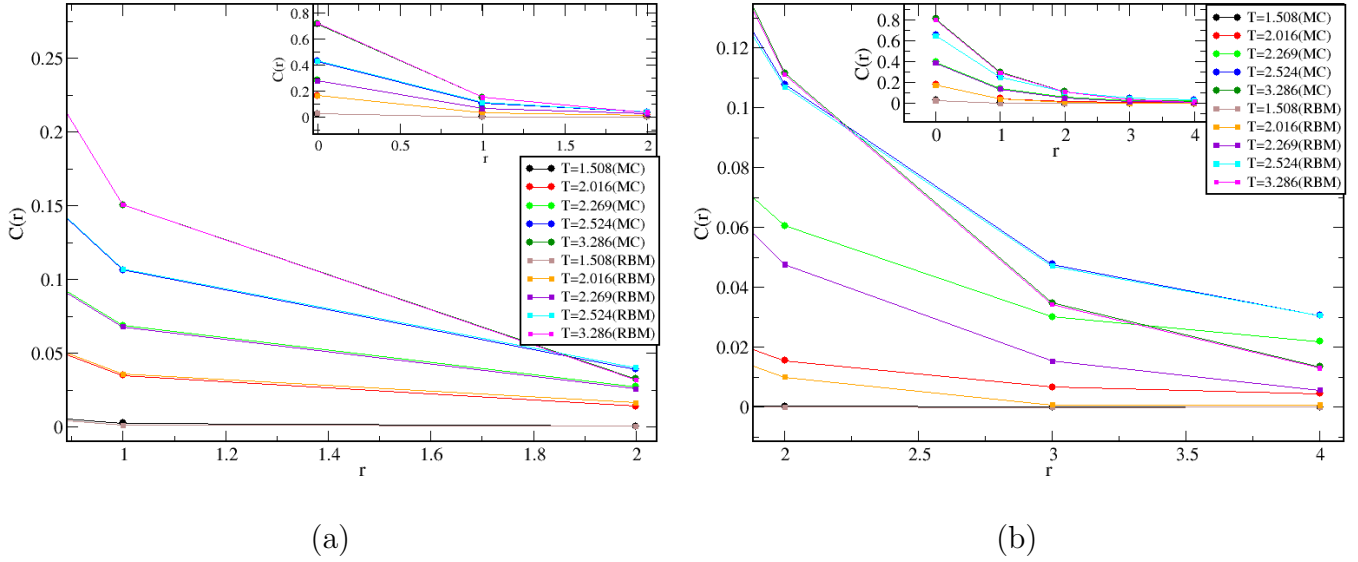
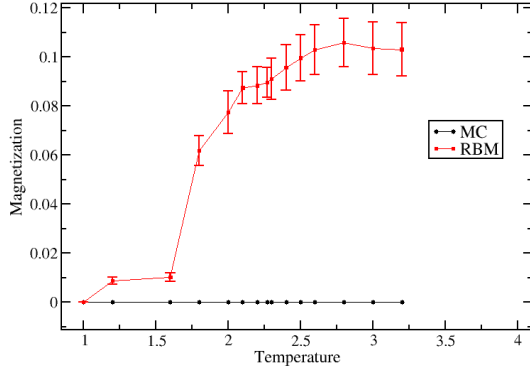


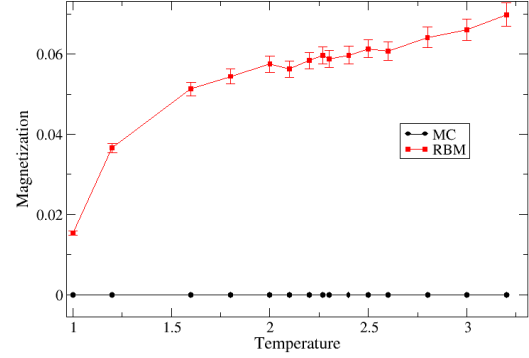
Figure 12.7: Correlation functions obtained over the Monte Carlo (MC) configurations and RBM configurations (RBM) at different temperatures with (a) $L = 4$ and (b) $L = 8$. For $L = 4$ the $C(r)$ values in RBM configurations are completely consistent with MC configurations. However, for $L = 8$ a systematic error is seen at temperatures close to the critical point.

Learning configurations with conserved magnetization

Training RBM over these configurations is more challenging as the magnetization is restricted to satisfy $m = 0$. The RBM architecture is not generally structured to respect this restriction and therefore the RBM can possibly learn it through the training process. However, our results for both system sizes of $L = 4$ and $L = 8$ indicate that generated configurations do not hold this constraint. As shown in Figure 12.8, magnetization graphs of generated configurations are not consistent with the conserved magnetization in Monte Carlo configurations. Therefore we conclude that the RBM does not learn to strictly produce the same number of spins with values $+1$ and -1 . But what exactly does the RBM learn? Interestingly we observe that in both system sizes, energy graphs derived from generated configurations and from input configurations are compatible to each other. One explanation of this observation is that the RBM is an energy-based machine learning model and structured to minimize the energy in the system. Therefore regardless of producing exact numbers of up spins ($+1$) and down spins (-1) it decreases the energy, as long as generated configurations are almost similar to the input configurations.

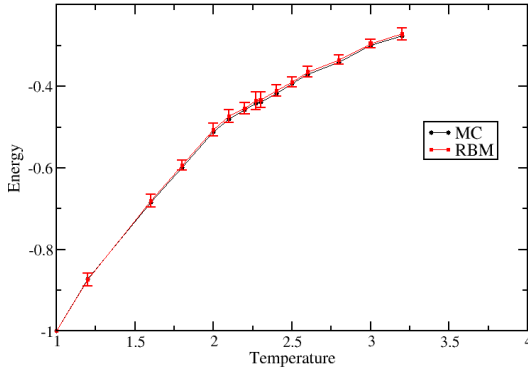


(a)

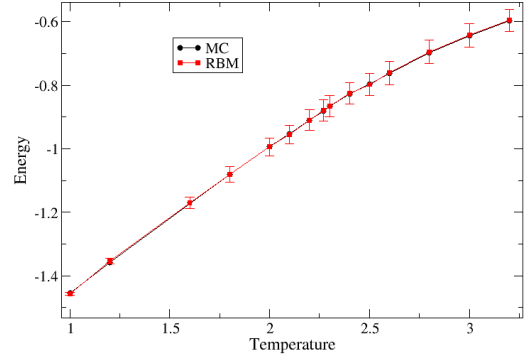


(b)

Figure 12.8: Magnetization graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. In both graphs, 10,000 configurations are generated by RBM. Results show that except from $T = 1.0$, at any other temperatures, magnetization in Monte Carlo configurations are outside of the error bars of magnetization values in the generated configurations.



(a)



(b)

Figure 12.9: Energy graphs of the Monte Carlo configurations (MC) and generated configurations (RBM) as functions of temperature with (a) $L = 4$ and (b) $L = 8$. All plots are averaged over 10,000 configurations. Unlike conserved magnetization, energy graphs in MC and RBM are consistent.

Furthermore, we measured $C(r)$ in the generated configurations as well as the input configurations (see Figure 12.10). Our results show a systematic error in the values of correlation functions in generated configurations. Again we conclude that when RBM is trained over

Ising configurations with conserved magnetization, it learns the energy values of the configurations very well but does not learn the magnetization constraints in addition to the correlation function values.

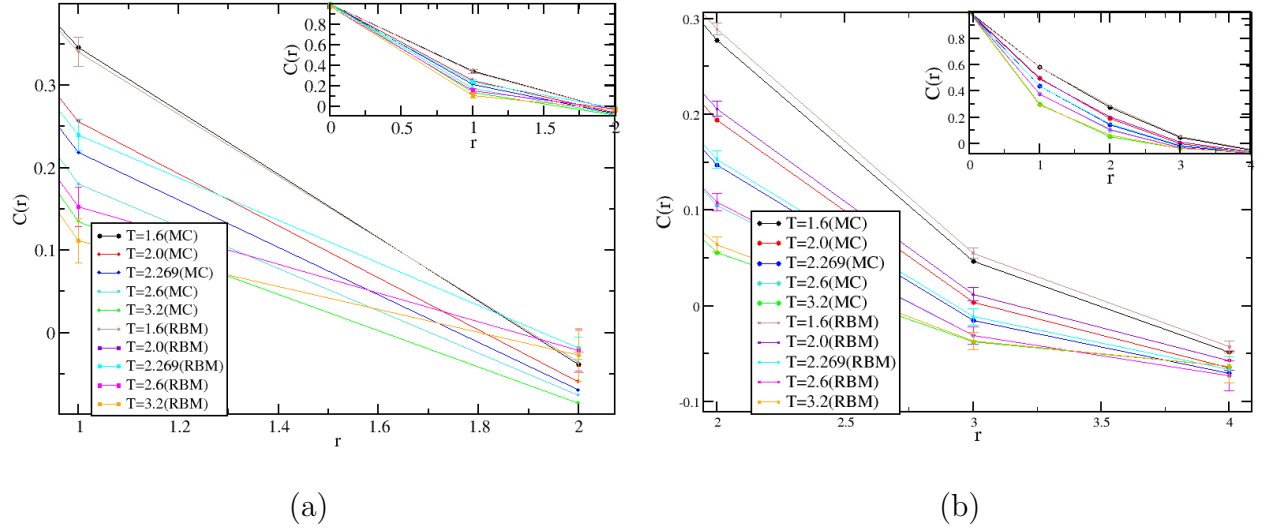


Figure 12.10: Correlation function measured on RBM configurations and MC configurations with system size (a) $L = 4$ and (b) $L = 8$. In both graphs, measurements over RBM configurations indicate a systematic error when compared to the Monte Carlo configurations (MC).

Chapter 13

Conclusion

Here we conclude our study on the applications of machine learning models to physical systems. In Chapter 11 we explained that Convolutional Neural Networks can identify the critical temperature of the Ising model with conserved magnetization. Also we discussed that the prediction accuracy of the model over the ordered phase can play the role of the order parameter. This idea can be generalized as well. In condensed matter, there are models which exhibit discontinuous phase transitions and therefore it is not easy to find the transition point. Given that, our results imply that training deep classifiers such as CNNs can be helpful to identify the phase transition point accurately.

We also studied learning process of Restricted Boltzmann Machines as an energy based generative model. We trained the RBM over the Ising model configurations with non-conserved magnetization and conserved magnetization. In the first case, we discussed that measuring magnetization and energy from the generated configurations does not allow the conclusions that the generated configurations always contain same physical properties as in the input data. We measured the correlation function over the generated configurations which supported our conclusion. Furthermore, we showed that RBM does not learn the restriction on the magnetization for Ising configurations with conserved magnetization.

Bibliography

- [1] H. K. Janssen, Z. Phys. B 42, 151 (1981). (Cited on page 3.)
- [2] P. Grassberger, Z. Phys. B 47, 365–374 (1982). (Cited on page 3.)
- [3] H. Hinrichsen, Adv. Phys. **49**, 815 (2000). (Cited on pages 3 and 16.)
- [4] G. Ódor, *Universality in Nonequilibrium Lattices Systems: Theoretical Foundations* (World Scientific, Singapore, 2008). (Cited on pages 3, 31, 32, and 35.)
- [5] C. Castellano and R. Pastor-Satorras, Phys. Rev. E **86**, 051123 (2012). (Cited on page 4.)
- [6] R. Garcia-Millan, arXiv:1910.02315 (2019). (Cited on page 4.)
- [7] T. M. Liggett, *Interacting Particle Systems*, (Springer, New York, 1985). (Cited on page 4.)
- [8] R. Durrett, *Lecture notes on particle systems and percolation* (Wadsworth, Pacific Grove, CA) (Cited on page 4.)
- [9] M. Droz, A. L. Ferreira, and A. Lipowski, Phys. Rev. E **67**, 056108 (2003). (Cited on pages 4, 6, 10, 11, 16, 23, 29, 32, 37, and 44.)
- [10] I. Dornic, H. Chaté, J. Chauve, and H. Hinrichsen, Phys. Rev. Lett. **87**, 045701 (2001). (Cited on pages 4, 5, 11, 31, 36, 39, and 44.)
- [11] O. Al Hammal, H. Chaté, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **94**, 230601 (2005). (Cited on pages ix, 7, 11, 16, 31, 32, 37, and 44.)
- [12] F. Vázquez and C. López, Phys. Rev. E **78**, 061127 (2008). (Cited on pages 7, 11, 32, 37, and 39.)
- [13] Á. L. Rodrigues, C. Chatelain, T. Tomé, and M. J. de Oliveira, JSTAT (**2015**), P01035. (Cited on pages 7, 11, 16, and 32.)
- [14] C. Castellano, M. A. Muñoz, and R. Pastor-Satorras, Phys. Rev. E **80**, 041129 (2009). (Cited on pages 7 and 16.)

- [15] A. Lipowski and M. Droz, Phys. Rev. E **65**, 056114 (2002). (Cited on pages 10, 11, 14, 21, and 29.)
- [16] A. Lipowski and M. Droz, Phys. Rev. E **66**, 016106 (2002). (Cited on pages 10 and 29.)
- [17] M. Droz and A. Lipowski, Braz. J. Phys. **33**, 526 (2003). (Cited on pages 10 and 29.)
- [18] S. M. Krause, P. Böttcher, and S. Bornholdt, Phys. Rev. E **85**, 031126 (2012). (Cited on pages 11, 14, and 29.)
- [19] L. Canet, H. Chaté, B. Delamotte, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **95**, 100601 (2005). (Cited on pages 11 and 32.)
- [20] Strictly speaking, T is not a temperature in the non-equilibrium models discussed in this work. Still, for convenience we will continue denoting the parameter T as temperature in the following. (Cited on pages 10 and 11.)
- [21] S.-C. Park, JSTAT (**2015**), P10009. (Cited on pages 11, 16, and 32.)
- [22] E. Lorenz and W. Janke, EPL **77**, 10003 (2007). (Cited on page 12.)
- [23] M. Henkel and M. Pleimling, *Non-Equilibrium Phase Transitions, Volume 2: Ageing and Dynamical Scaling Far From Equilibrium* (Springer, Heidelberg, 2010). (Cited on pages 12, 15, and 36.)
- [24] M. J. de Oliveira, Phys. Rev. E **67**, 066101 (2003). (Cited on page 14.)
- [25] M. O. Hase, S. R. Salinas, T. Tomé, and M. J. de Oliveira, Phys. Rev. E **73**, 056117 (2006). (Cited on page 14.)
- [26] M. O. Hase, T. Tomé, and M. J. de Oliveira, Phys. Rev. E **82**, 011133 (2010). (Cited on pages ix, 14, 15, 16, and 29.)
- [27] C. Chatelain, T. Tomé, and M. J. de Oliveira, JSTAT (**2011**), P02018. (Cited on pages ix and 15.)
- [28] C. Castellano and R. Pastor-Satorras, Phys. Rev. E **86**, 051123 (2012). (Cited on pages 16, 39, and 44.)
- [29] G. Grinstein, C. Jayaprakash, and Y. He, Phys. Rev. Lett. **55**, 2527 (1985). (Cited on page 16.)
- [30] K. Nam, S. Park, B. Kim, and S. J. Lee, JSTAT (**2011**), L06001. (Cited on page 16.)
- [31] S.-C. Park, Phys. Rev. E **85**, 041140 (2012). (Cited on pages 16 and 32.)
- [32] K. Nam, B. Kim, and S. J. Lee, JSTAT (**2014**), P08011. (Cited on page 16.)

- [33] H.-K. Janssen, B. Schaub, and B. Schmittmann, *Z. Physik B* **73**, 539 (1989). (Cited on pages 17 and 39.)
- [34] M. Henkel, M. Pleimling, C. Godrèche, and J.-M. Luck, *Phys. Rev. Lett.* **87**, 265701 (2001). (Cited on page 17.)
- [35] M. Pleimling and A. Gambassi, *Phys. Rev. B* **71**, 180401(R) (2005). (Cited on page 17.)
- [36] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982). (Cited on page 22.)
- [37] P. Peczak and D. P. Landau, *Phys. Rev. B* **39**, 11932 (1989). (Cited on page 22.)
- [38] E. Buffenoir and S. Wallon, *J. Phys. A: Math. Gen.* **26**, 3045 (1993). (Cited on page 22.)
- [39] P. Grassberger and Y.-C. Zhang, *Physica A* **224**, 169 (1996). (Cited on pages xi, 24, 27, and 36.)
- [40] C. M. Voigt and R. M. Ziff, *Phys. Rev. E* **56**, R6241 (1997). (Cited on pages xi, 24, 27, and 36.)
- [41] M. Henkel, H. Hinrichsen, and S. Lübeck, *Non-Equilibrium Phase Transitions, Volume 1: Absorbing Phase Transitions* Springer/Canopus, Dordrecht/Bristol, 2008). (Cited on page 31.)
- [42] F. Vazquez, P. L. Krapivsky, and S. Redner, *J. Phys. A* **36**, L61 (2003). (Cited on page 31.)
- [43] C. Castellano, S. Fortunato, and V. Loreto, *Rev. Mod. Phys.* **81**, 591 (2009). (Cited on page 31.)
- [44] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, *Phys. Rev. Lett.* **112**, 158701 (2014). (Cited on page 31.)
- [45] D. M. Abrams and S. H. Strogatz, *Nature* **424**, 900 (2003). (Cited on page 31.)
- [46] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. Mckane, *Language Variation and Change* **21**, 257 (2009). (Cited on page 31.)
- [47] D. I. Russell and R. A. Blythe, *Phys. Rev. Lett.* **106**, 165702 (2011). (Cited on pages 31, 32, and 33.)
- [48] P. L. Krapivsky, *Phys. Rev. A* **45**, 1067 (1992). (Cited on page 31.)
- [49] L. Frachebourg and P. L. Krapivsky, *Phys. Rev. E* **53**, R3009 (1996). (Cited on page 31.)
- [50] P. Clifford and A. Sudbury, *Biometrika* **60**, 581 (1973). (Cited on page 31.)
- [51] T. Shimaya and K. A. Takeuchi, *Phys. Rev. E* **99**, 042403 (2019). (Cited on page 31.)

- [52] C. Borile, M. A. Muñoz, S. Azaele, J. R. Banavar, and A. Maritan, Phys. Rev. Lett. **109**, 038102 (2012). (Cited on page 31.)
- [53] C. Borile, A. Maritan, and M. A. Muñoz, J. Stat. Mech. **(2013)**, P04032. (Cited on page 31.)
- [54] A. Tartaglia, L. F. Cugliandolo, and M. Picco, Phys. Rev. E **92**, 042109 (2015). (Cited on page 31.)
- [55] M. Scheucher and H. Spohn, J. Stat. Phys. **53**, 279 (1988). (Cited on page 31.)
- [56] R. Dickman and A. Yu. Tretyakov, Phys. Rev. E **52**, 3218 (1995). (Cited on pages 32 and 36.)
- [57] M. A. Muñoz, G. Grinstein, and Y. Tu, Phys. Rev. E **56**, 5101 (1997). (Cited on page 32.)
- [58] A. Azizi, J. Stidham, and M. Pleimling, J. Stat. Mech. **(2018)**, 103208. (Cited on pages xii, 32, 36, 37, 39, 41, and 44.)
- [59] P. Meakin, *Fractals, Scaling and Growth far from Equilibrium* (Cambridge University Press, New York, 1998) (Not cited.)
- [60] F. Family, and T. Vicsek, J. Phys. A **18**, L75 (1985). (Cited on page 52.)
- [61] M. Kardar, G. Parisi, and Y. C. Zhang, Phys. Rev. Lett. **56**, 889 (1986). (Not cited.)
- [62] G. Ódor, Rev. Mod. Phys. **76**, 663 (2004). (Cited on page 36.)
- [63] K. Binder, Z. Physik B **43**, 119 (1981). (Cited on page 37.)
- [64] K. Binder, Phys. Rev. Lett. **47**, 693 (1981). (Cited on page 37.)
- [65] W. Selke, Eur. Phys. J. B **51**, 223 (2006). (Cited on pages xii, 37, and 42.)
- [66] H. K. Janssen, in *From Phase Transitions to Chaos. Topics in Modern Statistical Physics*, eds. G. Györgyi, I. Kondor, L. Sasvári, and T. Tél (World Scientific, Singapore, 1992). (Cited on page 39.)
- [67] G. Kamieniarz and H. W. J. Blöte, J. Phys. A: Math. Gen. **26**, 201 (1993). (Cited on pages xii and 42.)
- [68] A.-L. Barabási, and H. E. Stanley *Fractal concepts in surface growth* (Cambridge university press) (Cited on pages 52 and 53.)
- [69] C. M. Bordogna, and E. V. Albano, Phys. Rev. E **83**, 046111 (2011). (Cited on pages 60 and 61.)

- [70] R. Dickman, and M. A. Munoz, Phys. Rev. E **62**, 7632 (2000). (Cited on pages 54 and 55.)
- [71] J. Carrasquilla, and R. G. Melko, Nature **13**, 431 (2017) (Cited on pages 64 and 69.)
- [72] L. Wang, Phys. Rev. B **94**, 195105 (2016). (Cited on pages 64 and 72.)
- [73] S. J. Wetzel, Phys. Rev. E **96**, 022140 (2017). (Cited on page 64.)
- [74] W. Hu, R. R. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017). (Cited on page 64.)
- [75] W. Zhang, L. Jiayu, and W. Tzu-Chieh, Phys. Rev. E **99**, 032142 (2019). (Cited on page 64.)
- [76] C. Alexandrou, A. Athenodorou, C. Chrysostomou, and S. Paul, arXiv:1903.03506 (2019). (Cited on page 64.)
- [77] A. Tanaka, and A. Tomiya, J. Phys. Soc. Japan **86**, 063001 (2017). (Cited on page 64.)
- [78] C. Wang, and H. Zhai, Phys. Rev. B **96**, 144432 (2017). (Cited on page 64.)
- [79] D. Kim, and D. H. Kim, Phys. Rev. E **98**, 022138 (2018). (Cited on page 69.)
- [80] G. Torlai, and R. G. Melko, Phys. Rev. B, **94**, 165134 (2016). (Cited on pages 64 and 82.)
- [81] A. Morningstar, and R. G. Melko, The Journal of Machine Learning Research **18**, 5975 (2017). (Cited on pages 64, 74, and 82.)
- [82] Z. Liu, S. P. Rodrigues, and W. Cai, arXiv:1710.04987 (2017). (Cited on page 64.)
- [83] M. Cristoforetti, G. Jurman, A. I. Nardelli, and C. Furlanello, arXiv:1705.09524 (2017). (Cited on page 64.)
- [84] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature, **549**, 195 (2017). (Not cited.)
- [85] L. Huang, and L. Wang, Phys. Rev. B **95**, 035105 (2017). (Cited on page 74.)
- [86] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B, **95** 041101.(2017). (Not cited.)
- [87] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, Cognitive science, **9(1)**, 147 (1985). (Cited on page 75.)
- [88] G. E. Hinton, *A practical guide to training restricted Boltzmann machines. In Neural networks: Tricks of the trade, (pp. 599-619)* (Springer, Berlin, Heidelberg, 2010). (Cited on page 82.)