

國立台灣大學理學院物理學系

碩士論文

Department of Physics

College of Science

National Taiwan University

Master Thesis



古典蒙地卡羅對三維 XY 模型非線性與異向性效應之  
研究

Classical Monte Carlo Studies on 3D XY Models with  
Effects of Nonlinearity and Anisotropy

吳柏寬

Po-Kuan Wu

指導教授：高英哲博士

Advisor: Ying-Jer Kao, Ph.D.

中華民國 103 年 7 月

July, 2014

# 國立臺灣大學研究所博、碩士學位考試試卷(口試紀錄表)

102 學年度第 2 學期

考試日期：103 年 7 月 25 日

系所組別：物理學系

考試地點：新物 R716

學號：R01222012

記錄：\_\_\_\_\_

姓名：吳柏寬

學位考試成績：A (請以等第制評分，成績評量定義詳見下列說明，研究生及格標準為 B-)

論文題目：Classical Monte Carlo Studies on 3D XY Models with Effects of Nonlinearity and Anisotropy

考試委員簽章：

高英哲  
林諭評  
陳柏中

## ※成績評量定義：Definition of Grades

- A+：所有目標皆達成且超越期望 (All goals achieved beyond expectation)
- A：所有目標皆達成 (All goals achieved)
- A-：所有目標皆達成，但需一些精進 (All goals achieved, but need some polish)
- B+：達成部分目標，且品質佳 (Some goals well achieved)
- B：達成部分目標，但品質普通 (Some goals adequately achieved)
- B-：達成部分目標，但有些缺失 (Some goals achieved with minor flaws)
- C+：達成最低目標 (Minimum goals achieved)
- C：達成最低目標，但有些缺失 (Minimum goals achieved with minor flaws)
- C-：達成最低目標但有重大缺失 (Minimum goals achieved with major flaws)
- F：未達成最低目標 (Minimum goals not achieved)
- X：因故不核予成績 (Not graded due to unexcused absences or other reasons)



## 誌謝

完成這本論文，我要感謝高英哲老師三年多的指導。讓我學會的古典蒙地卡羅方法與 GPU 程式設計一些必要的知識與技巧。指引我研究的方向之餘，也給我足夠的空間研究一些相關的領域，並予以討論與指導。也給予我很多不同的機會與其他人合作，嘗試各種模型。對於喜歡學嘗試新事物的我，是十分重要的。感謝 T. Pereg-Barnea 教授和林瑜琿教授和我合作，讓我對於各種模型的模擬更熟悉。感謝 A. Sandvik 教授與林瑜琿教授對於我的研究給予不少建議。謝謝陳柏中教授對於我論文的指導。

我還想要感謝我們組上的同學們。特別感謝謝昀達學長的帶領與指導，讓我對 GPU 和蒙地卡羅方法有更深的認識。也感謝羅雅琳學姐、林昇慶學長、楊淵榮、郭子傑、蕭維翰對於我研究及課業上的種種建議與協助。還要感謝我的高中同學與社團好友，給予我精神上的支柱。

最後，我想感謝我的父母，一直給予我精神與實質上的支持，給予我很大的自由讓我走想走的路。



## 摘要

在統計與凝態物理中，相變與臨界現象是相當重要的主題。而在研究這類問題上，晶格模型扮演了非常重要的角色。基於普適性與臨界現象的理論，理論模型的臨界行為經常能對應到真實物理系統。因此，對於理論模型的研究是了解臨界現象的關鍵之一。

蒙地卡羅方法被廣泛用於了解晶格模型的相變上。利用隨機過程，可以利用在態空間中隨機取樣得到各種熱力學性質的近似值。而基於有限尺度標度變換，從有限尺度的結果可以推估臨界指數的值，也可以幫助了解相變的性質。然而，在有限大小的結果中，不相關的場會造成一些修正項，在計算臨界指數時，這會造成系統誤差。因此，必須模擬更大的晶格。在本文中，為了在模擬大晶格模型時有更高的效率，我們利用了圖形處理器 (GPU) 將程序平行化處理。

在本論文中，利用圖形處理器上的蒙地卡羅模擬，我們研究了簡單 XY 模型，以及將交互作用項推廣為非線性的 XY 模型。簡單 XY 模型的臨界現象與  $^4\text{He}$  的  $\lambda$  相變屬於相同的普適類。而在  $q$  大於 4 時， $Z_q$  異向性都是危險不相關的。推廣的 XY 模型的情形則是不同，當自旋間的交互作用越來越接近 delta 函數，模型的行為會越來越接近 Potts 模型，相變變為一階。在參數位於某些區域時，可以明顯觀察到異向性是相干的，甚至可能因為異向性強度的增加，相變轉變為一階相變。



# Abstract

In statistical and condensed matter physics, the phase transition and the critical behavior are very important topics. To study on them, the lattice models play important roles. For the theory of universality, the behaviors of models correspond to realistic physical systems. Therefore, the studies of theoretic models are keys to understand the critical behaviors.

The Monte Carlo simulation is a widely-used way to study the phase transitions of lattice models. With the stochastic process, we can sample in the space of states and get the approximations of the thermal observables. According to the finite-size scaling, from the finite-size results, the critical exponents can be extracted and the properties of the transition can be studied. However, in finite-size cases, there are correction terms cause by the irrelevant fields, so them cause the systematic errors of exponents. To simulate the system with larger sizes more quickly, in this thesis, we parallelize the procedures of Monte Carlo simulations on GPUs.

With the GPU Monte Carlo simulations, we study the simple XY model and the generalized cases with nonlinear interactions between spins. The simple XY model is in the same universality class with the  $\lambda$ -transition of  $^4\text{He}$ , and the  $Z_q$  anisotropy is dangerously irrelevant for  $q \geq 4$ . In the generalized XY models, the behavior approaches the Potts models as the potential of interactions approaching the delta function. In some region of the parameters, the anisotropy is significantly relevant. The transitions may even become first-order as anisotropy increases.



# Contents

誌謝	ii
摘要	iii
Abstract	iv
<b>1 Introduction</b>	<b>1</b>
1.1 Phase transitions . . . . .	1
1.2 Critical behavior and scaling law . . . . .	4
1.2.1 Critical exponent . . . . .	4
1.2.2 Universality . . . . .	5
1.2.3 Scaling laws . . . . .	5
1.3 Ising Model, XY model and Potts model . . . . .	8
<b>2 Classical Monte Carlo Method with Graphics Processing Unit</b>	<b>12</b>
2.1 Classical Monte Carlo . . . . .	12
2.1.1 Importance sampling and Markov chain Monte Carlo . . . . .	13
2.1.2 Measurements . . . . .	15
2.1.3 Parallel tempering . . . . .	17

2.2	GPU architecture and CUDA framework . . . . .	18
2.3	Implementation of classical Monte Carlo simulation on GPU . . . . .	21
2.4	Multispin coding on GPU . . . . .	24
<b>3</b>	<b>Data Analysis and Finite-Size Scaling</b>	<b>27</b>
3.1	Data analysis . . . . .	27
3.2	Finite-size scaling . . . . .	30
<b>4</b>	<b>Three-dimensional XY model with <math>Z_q</math> Anisotropy</b>	<b>33</b>
4.1	3D XY model . . . . .	33
4.2	3D XY Model with $Z_q$ anisotropy and emergent U(1) symmetry . . . . .	39
<b>5</b>	<b>Nonlinear Three-dimensional XY Model with <math>Z_q</math> Anisotropy</b>	<b>45</b>
5.1	Nonlinear 3D XY model . . . . .	46
5.2	Nonlinear 3D XY model with $Z_q$ anisotropy . . . . .	51
5.2.1	From continuous to first-order . . . . .	51
5.2.2	From irrelevance to relevance . . . . .	53
<b>6</b>	<b>Summary</b>	<b>59</b>
	<b>Bibliography</b>	<b>60</b>





# List of Figures

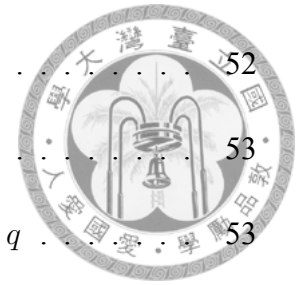
1.1	The free energy and the internal energy for the first-order and the continuous phase transitions. . . . .	3
1.2	Simplified phase diagram for water. . . . .	4
1.3	An example of the universality class. . . . .	6
1.4	Picture of block spins . . . . .	7
1.5	Ising model and XY model . . . . .	9
2.1	Correlation length of XY model. . . . .	17
2.2	Sketch map of parallel tempering and the free energy landscape . . . . .	18
2.3	GPU layout . . . . .	19
2.4	A grid of blocks, containing threads . . . . .	20
2.5	Sublattices for the checkerboard algorithm. . . . .	22
2.6	Summation over all threads in a block. . . . .	23
2.7	Allocation of values of spins. . . . .	24
2.8	Calculation of energies of bonds by a bitwise XOR. . . . .	25
2.9	CAMSC scheme to calculate the energies of every single spin. . . . .	26
3.1	The magnetization in $x$ direction versus Monte Carlo sweeps. . . . .	29





3.2	Data collapse of $\chi$ of the 3D XY model. . . . .	31
4.1	Energy per site v.s. $T$ of the XY model. . . . .	35
4.2	Magnetization v.s. $T$ of the XY model. . . . .	35
4.3	Specific heat of the XY model. . . . .	36
4.4	Susceptibility of the XY model. . . . .	36
4.5	Binder cumulant of the XY model. . . . .	37
4.6	Finite size scaling plot of correlation length of the XY model. . . . .	37
4.7	Finite size scaling plot of Binder cumulant of the XY model. . . . .	38
4.8	$P(m_x, m_y)$ of $Z_4$ and $Z_8$ systems. . . . .	40
4.9	$m_q$ order parameter for $q = 4$ and $h = 2.0$ . . . . .	41
4.10	$m_5$ order parameter for $Z_5$ clock model . . . . .	42
4.11	$m_q$ and $m$ versus $T$ for $q = 5, 6$ . . . . .	43
4.12	Scaling $m_q$ for $q = 4, 5, 6$ . . . . .	44
5.1	$V(\theta)$ vs. $\theta$ . . . . .	46
5.2	Binder cumulants for $P = 8$ . . . . .	47
5.3	Energy histograms for $P = 8, L = 32$ . . . . .	48
5.4	Order parameter for $P = 8$ . . . . .	48
5.5	Internal energy for $P = 8$ . . . . .	49
5.6	Micro-canonical Monte Carlo simulation result for $P = 55$ . . . . .	49
5.7	Binder cumulants for $P = 7.5$ . . . . .	50
5.8	First-order and continuous phase transitions for $q = 4$ with different $P$ and $h$ . . . . .	51

5.9	The energy histogram for $P = 2.5$ , $q = 4$ and $h = 2.0$ . . . . .	52
5.10	The order parameter distribution for $P = 2.5$ and $h = 2.0$ . . . . .	53
5.11	Existence of first-order phase transition for different $P$ and $q$ . . . . .	53
5.12	Energy versus the reduced temperature $T/T_c$ . . . . .	54
5.13	Critical exponents versus $h$ for $P = 2$ and $q = 4$ . . . . .	56
5.14	Critical exponents versus $h$ for $P = 2$ and $q = 5$ . . . . .	56
5.15	Critical exponents versus $h$ for $P = 3$ and $q = 5$ . . . . .	57
5.16	$\langle m \rangle$ and $\langle m_4 \rangle$ versus $T$ for $P = 2.0$ , $h = 2.0$ and $q = 4$ . . . . .	57
5.17	Scaling plot of $\langle m_4 \rangle$ for $P = 2.0$ , $h = 1.0$ and $q = 4$ . . . . .	58





# List of Tables

1.1	Critical exponents of Ising Models . . . . .	9
5.1	Critical exponent $\nu$ and $\gamma$ versus $P$ . . . . .	47



# Chapter 1

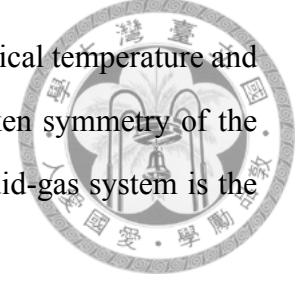
## Introduction

### 1.1 Phase transitions

In statistical mechanics, we study the macroscopic properties of materials, for example, their density or magnetization. In ordinary situations, thermal quantities change smoothly with the thermodynamic fields, such as the temperature and the magnetic field. However, at some special points of the thermodynamic fields, the properties of the material may change abruptly. It is called a phase transition.[1, 2] We can find the phenomena of phase transitions in many materials, for example, the transition from water to ice, the ferromagnet-paramagnet phase transitions at the Curie temperature and the superconductor-metal phase transitions.

Take the Curie point of a ferromagnetic system as an example: when the temperature is higher than  $T_c$ , there is no magnetization and the system is rotational invariant. In contrast, the magnetization spontaneously emerges and rotational symmetry is broken below the critical temperature. Separated by the transition temperature, there are two phases. The lower-temperature phase with reduction of the symmetry is called an ordered phase while the high-temperature phase is disordered. To describe the symmetry of a phase transition and to distinct ordered and disordered phases, an extra parameter, called *order parameter*, is defined[1, 2]. In this case, the magnetization is the order parameter, denoted by  $m$ , a

three-component vector. The order parameter tends to zero at the critical temperature and in the disordered phase. The order parameter is related to the broken symmetry of the system. For example, the single-component order parameter of liquid-gas system is the difference of density.



The correlation between particles plays an important role. The correlation function of an order parameter is defined as

$$\Gamma(r) = \langle m(r)m(0) \rangle - \langle m(r) \rangle \langle m(0) \rangle. \quad (1.1)$$

The correlation function decays exponentially with a characteristic length scale, the correlation length, denoted by  $\xi$ . According to fluctuation-dissipation theorem[1, 2], the correlation function can be asymptotically given by the Ornstein-Zernike form

$$\Gamma \propto r^{-(d-1)/2} e^{-r/\xi}, \quad (1.2)$$

where  $\xi$  is called the correlation length, an important characteristic in a phase transition.

The transition between two phases can be categorized into two types: the *first-order phase transition* and the *continuous phase transition*. According to the Ehrenfest classification[3, 4], the order of a phase transition is defined by the free energy. If the first derivative of the free energy is discontinuous, the phase transition is first-order; in contrast, the first derivatives are continuous for continuous phase transitions, as shown in Fig. 1.1.

In a first-order phase transition, at the critical temperature, the ordered and disordered phases co-exist. There is a jump for the internal energy, called latent heat. Various thermodynamic quantities are also discontinuous. Famous examples of the first-order transitions are the phase transitions between gas and liquid and the melting of solid. There are often hysteresis, i.e. we can find metastable states in some region of conjugate fields. At the transition, the correlation length  $\xi$  is finite.

In the case of continuous phase transitions, the correlation length  $\xi$  diverges at  $T_c$ . Unlike the first-order transition, the energy density and the magnetization, change smoothly.

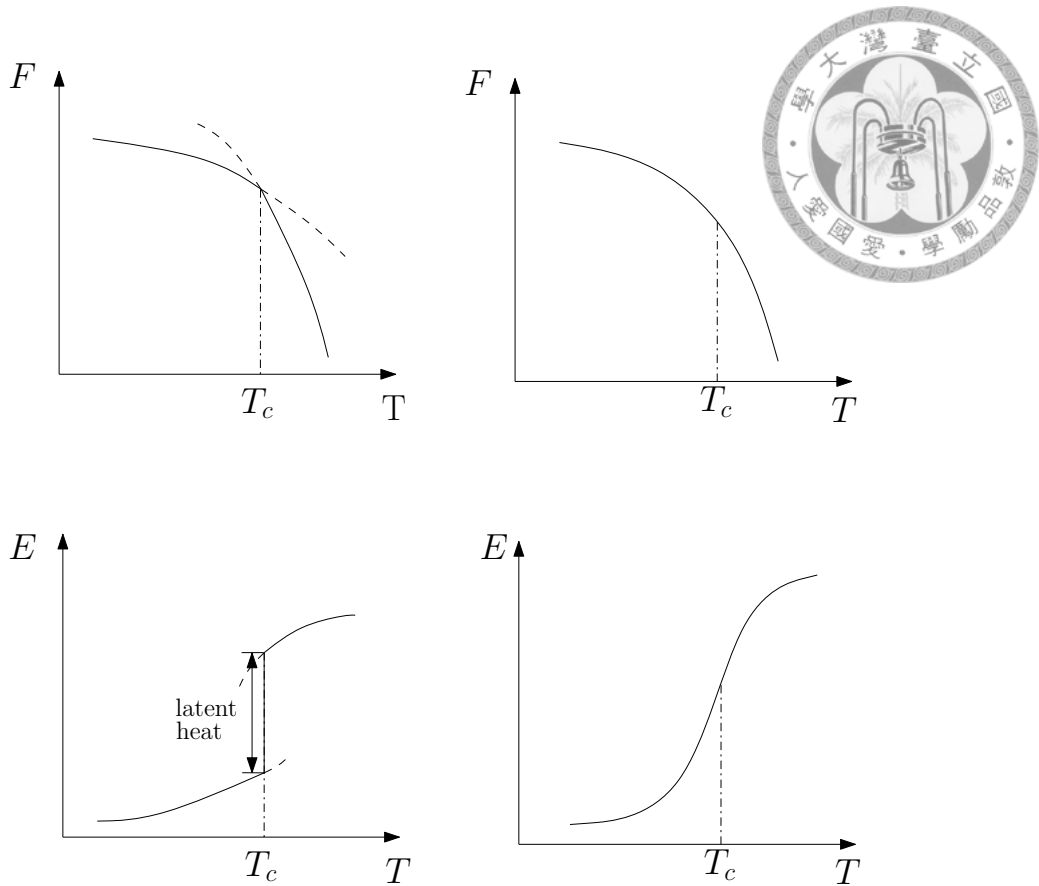


Figure 1.1: Schematic temperature dependence of the free energy and the internal energy for the first-order (left) and the continuous (right) phase transitions.

A famous example of continuous phase transitions is the paramagnet to ferromagnet at the Curie temperature. The schematic phase diagram of water is shown in Fig. 1.2. The first-order liquid-gas transition becomes continuous at the end point of line of water-steam phase transition line, the so-called critical point. The critical behaviors of continuous transitions can be described by theories established on the divergence of the correlation length, which will be introduced in next section.

The behaviors of phase transitions are fruitful, and often hard to solve exactly. Therefore, numerical methods and simulations are important for understanding and classifying phase transitions. Based on stochastic process, *Monte Carlo* simulations are widely used to study both classical and quantum systems. The main purpose of this thesis is to study the phase transitions of theoretical models with classical Monte Carlo simulations.

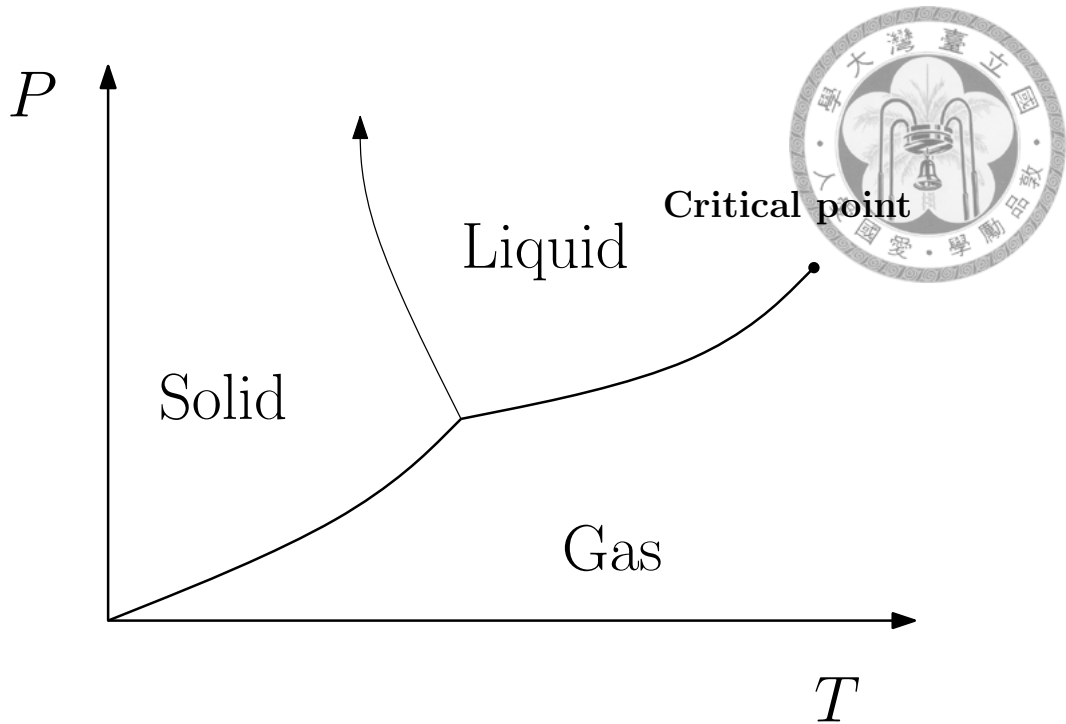


Figure 1.2: Simplified phase diagram for water.

## 1.2 Critical behavior and scaling law

### 1.2.1 Critical exponent

For continuous phase transitions, singularities of various quantities depend on the reduced distance,

$$t = \left| \frac{T - T_c}{T_c} \right|.$$

For  $t \rightarrow 0$ , we can divide the quantity into regular and singular parts. The singular part diverges at  $T_c$  and is proportional to some power of  $t$ .

Therefore, we can define the critical exponents[1]:

$$\text{Specific heat : } C_v \propto t^{-\alpha}, \quad (1.3)$$

$$\text{Order parameter : } M \propto t^\beta, \quad (1.4)$$

$$\text{Susceptibility : } \chi \propto t^{-\gamma}, \quad (1.5)$$

$$\text{Correlation length : } \xi \propto t^{-\nu}, \quad (1.6)$$

where the susceptibility  $\chi$  is defined as

$$\chi = \frac{dM}{dH}. \quad (1.7)$$



At  $T = T_c$ , there is another relation of the order parameter  $M$  and the conjugate field  $H$ , written as

$$M \propto H^{1/\delta}. \quad (1.8)$$

Considering the the correlation function at  $T_c$ , it decays as a power-law of the distance

$$\Gamma(r) \propto r^{d-2-\eta} \quad (1.9)$$

In total, there are six *critical exponents*,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\nu$ ,  $\delta$  and  $\eta$ , defined for different quantities.

## 1.2.2 Universality

If we compare the critical exponents of many different systems, we will find some of them have the same set of exponents, shown in Fig. 1.3. The theoretic three-dimensional(3D) XY model and  $\lambda$ -transition of  $^4\text{He}$ , also have the same set of critical exponents. By these universal behaviors, the systems with the set of exponents are classified into the same *universality class*. The properties, spatial dimensionality, spin dimensionality, interactions and symmetries may determine the universality class. Even systems with different order parameters, e.g. density and magnetization, may belong to a same universality class. This indicates that we can use simple models to study the critical behaviors of real systems.

## 1.2.3 Scaling laws

We have six critical exponents. However, they are not independent. Given the lattice dimensionality  $d$ , there are four relations between the exponents, and there are only two



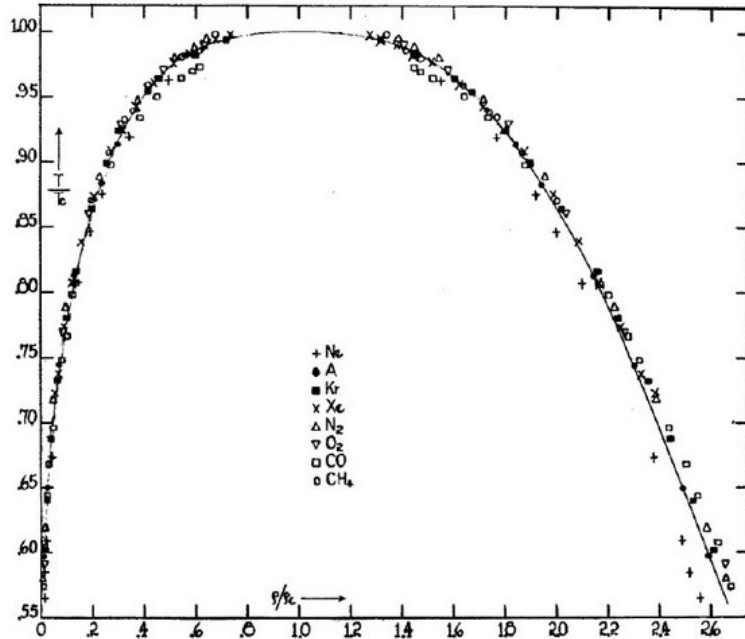
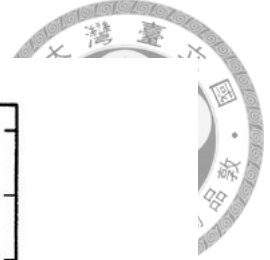


Figure 1.3: An example of the universality class. Diagram of reduced temperature and reduced density in the gas-liquid coexistence region.

independent exponents. The four relations, which are also called scaling laws, are

$$\alpha + 2\beta + \gamma = 2, \tag{1.10}$$

$$\gamma = \beta(\delta - 1), \tag{1.11}$$

$$\gamma = \nu(2 - \eta), \tag{1.12}$$

$$\nu d = 2 - \alpha. \tag{1.13}$$

We can understand these laws by the scaling hypothesis. We follow Kadanoff's picture to give an explanation[5]. First, we start from the free energy. The singular portion of it can be written as

$$G = G_{\pm}(t, h). \tag{1.14}$$

where  $h$  is proportional to the external field. It means that the free energy of a system only relative to these two variables. Shown in Fig. 1.4, when we perform the Kadanoff block-spin transformation, we regard spins in a  $l^d$  block as one spin. We assume that when  $l \ll \xi$ , the form of the free energy  $G_{\pm}(t, h)$  will not change and the free energy should

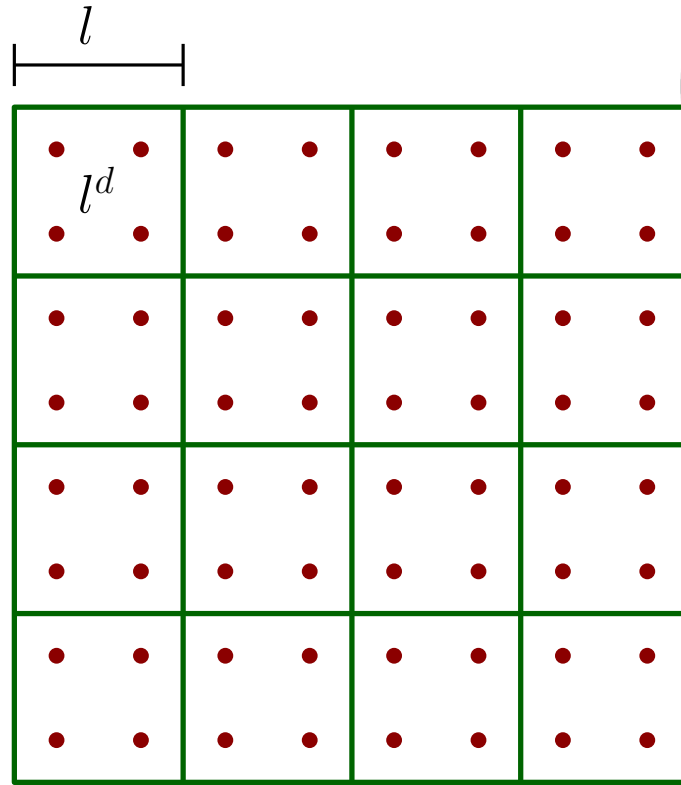


Figure 1.4: Picture of block spins

be the same after the transformation. However, after performing the coarse-graining, the correlation length looks shorter and  $t$  and  $h$  are related to it. Therefore,  $t$  and  $h$  should be rescaled into  $tl^{y_t}$  and  $hl^{y_h}$ . And, the number of spins is divided by  $l^d$ . Therefore, the scaling of free energy is written as

$$G(t, h) = l^{-d} G_{\pm}(tl^{y_t}, hl^{y_h}). \quad (1.15)$$

In a similar way, for the length between spins increases and the scaling of correlation length is given by

$$\xi(t, h) = l \xi_{\pm}(tl^{y_t}, hl^{y_h}). \quad (1.16)$$

We can obtain the scaling laws from these two homogeneous function, and the general form of scaling hypothesis. Replacing  $l$  with  $|t|^{-1/y_t}$ , and assuming  $h = 0$ , Eq. (1.16) turns into

$$\xi(t, 0) = t^{1/y_t} \xi_{\pm}(1, 0). \quad (1.17)$$

Comparing with Eq. (1.6), we get  $\nu = 1/y_t$ . And from Eq. (1.15), we get the scaling of specific heat

$$C = |t|^{-2+d\nu} \mathcal{C}_{\pm}(1, 0). \quad (1.18)$$



Comparing with Eq. (1.3), we can find the relation Eq. (1.13). In a similar way, we can get all of the scaling relations, or we can represent the critical exponents by independent variables,  $y_t$  and  $y_h$ . These relations are necessary to obtain the critical exponents and decide the universality class.

### 1.3 Ising Model, XY model and Potts model

In section 1.1, we already mention the importance of numerical methods and simulations in studying phase transitions. In many cases, experimental results are not enough or not available. By comparing with simulation data, we can examine the theoretical description. There also exist real world physical systems which are so complex that are hard to model. With simulation results, the system can be described and studied more clearly. For example, with the Monte Carlo simulations, we can study the critical behavior and find the values of the critical exponents more precisely to compare with the theoretical predictions and the experimental results. Simulation can be a support of theory and experiment, and as a tool to find new system with new phenomena. However, since the real interactions of particles are complicated, physicists try to solve or simulate some simplified model. From the renormalization group theory and the universality, though the interactions are simplified, the critical behaviors of simplified models are usually well corresponding to real physical systems.

The most famous of these theoretical models is Ising model, which was invented by Wilhelm Lenz and first studied by Ernst Ising[2, 6]. We describe the interaction by the

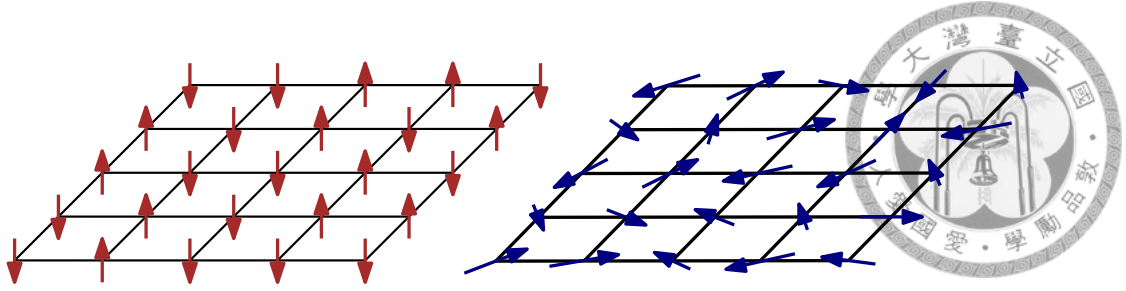


Figure 1.5: Ising model and XY model

Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i, \quad (1.19)$$

where  $s_i$  represent classical spins with values  $\pm 1$  on a chain in one-dimensional(1D) system or higher-dimensional lattice, shown in Fig. 1.5.  $\langle i, j \rangle$  means that the interaction is between nearest-neighbor sites. For  $J > 0$ , the spins tend to point to the same direction, while the nearest-neighbor spins tend to point to the opposite directions for  $J < 0$ . We can define the order parameters of these ordered phases, ferromagnetic phase and anti-ferromagnetic phase. Taking the ferromagnetic phase for example, we define the order parameter by the average magnetization per spin

$$m = \frac{\sum_i^N s_i}{N}, \quad (1.20)$$

where  $N$  is the number of spins.

Though there is no finite-temperature phase transition for the Ising model in 1D, there exist finite-temperature continuous phase transitions in 2D or 3D. The 2D Ising model is exactly solved by Onsager[7], and the 3D Ising model, which can not be solved exactly, has been widely studied by classical Monte Carlo Methods. For lattice dimensionality  $d \geq 4$ , the critical behaviors are described by mean field theory. The critical exponents of 2D and 3D Ising universality are shown in Table. 1.1[7, 8].

Table 1.1: Critical exponents of Ising Models

d	$\nu$	$\alpha$	$\beta$	$\gamma$	$\delta$	$\eta$
2	1	0	1/8	7/4	15	1/4
3	0.63005(18)	0.10985	0.32648	1.23717(28)	4.7894	0.03639
4	1/2	0	1/2	1	3	

In classical XY models, which are also called rigid-rotator model, 2-component unit vectors  $S_i = (\cos(\theta_i), \sin(\theta_i))$  are placed on every lattice sites, as shown in Fig. 1.5. The simplest XY model is described by the Hamiltonian with nearest-neighbor interactions

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j). \quad (1.21)$$

Like the Ising model, it is ferromagnetic for  $J > 0$  and antiferromagnetic for  $J < 0$ . Since the magnetization can be in any direction, we define the order parameter as the magnitude of magnetization

$$m \equiv \frac{M}{N} \equiv \frac{(M_x^2 + M_y^2)^{1/2}}{N}, \quad (1.22)$$

where

$$m_x \equiv \frac{M_x}{N} \equiv \frac{\sum_{i=1}^N \cos(\theta_i)}{N}, m_y \equiv \frac{M_y}{N} \equiv \frac{\sum_{i=1}^N \sin(\theta_i)}{N}. \quad (1.23)$$

Based on the Goldstone's theorem, the Mermin-Wagner theorem predicts that for  $d \leq 2$ , the symmetry of continuous spin systems cannot be spontaneously broken at finite temperature[2, 9]. Nevertheless, in the two-dimensional XY model, instead of ordinary continuous phase transition, there is a special transition, which is called *Kosterlitz-Thouless transition*. Below the temperature  $T_{KT}$ , the correlation function decay as a power law, and we can regard this region as a line of critical points. The KT-transition is widely studied and explained with vortices and bound vortex pairs[10].

In the 3D case, there is a continuous phase transition. The transition correspond to the  $\lambda$ -transition of  $^4\text{He}$ [11–13]. Our study will focus on this model and its extension.

Another important model I will discuss in this thesis is the Potts model. For a  $q$ -state

Potts model, there are  $q$  states at every site. The Hamiltonian of the model is given by

$$H = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j}, \quad (1.24)$$

where  $\sigma_i = 1, 2, \dots, q$ . The energy is zero when  $\sigma_i$  and  $\sigma_j$  are in different states. It is obvious that the two-state Potts model is equal to the Ising Model. The  $q$ -state Potts models have first order phase transitions for  $q > 4$  in two dimensions and for  $q > 2$  in three dimension[14].

In this thesis, we study the transitions of the XY models and the cases with effects of nonlinearity and anisotropy. In Ch. 2, we will introduce the Monte Carlo simulations and how to implement them on GPUs. In Ch. 3, the method for data analysis and the finite-size scaling are introduced. The part of results is Ch. 4 and Ch. 5. In Ch. 4, we show our results of the simple XY model with and without the anisotropy. The results of the nonlinear XY models will be in Ch. 5.



## Chapter 2

# Classical Monte Carlo Method with Graphics Processing Unit

In this chapter, we introduce Monte Carlo simulations and how to implement Monte Carlo simulations on Graphics Processing Unit.

### 2.1 Classical Monte Carlo

In 1940s, as the development of computers, physicists invent a class of algorithms studying problems with stochastic process, called Monte Carlo methods[4, 6]. The concept of Monte Carlo methods is simple. In order to measure volumes, integrations or summations in a large space, we can estimate them by many random samples with statistical errors as the price. These methods have outstanding performance for studying spaces with large dimensionality. And, the technique of sampling plays a very important role in increasing the efficiency.

In this thesis, when we use the term "classical Monte Carlo methods", we mean the Monte Carlo algorithms for the classical systems in statistical mechanics. The widely used one is the *Metropolis algorithm*, based on the *Markov chain*. To explain how to use Monte Carlo methods on statistical systems, we begin with the partition function in the canonical

ensemble

$$Z = \sum_s e^{-\beta H(s)}, \quad (2.1)$$

where  $s$  means states and  $H(s)$  represents the energy of the state  $s$ , and we sum over all states to get the partition function. We can write the expectation value of observable  $O$  as

$$\langle O \rangle = \frac{\sum_s O(s) e^{-\beta H(s)}}{Z}, \quad (2.2)$$

where  $O(s)$  means the value of observable  $O$  of state  $s$ . For example, we can use this equation to get the energy and the magnetization.

However, when the system is very large, it is impossible to go through all micro-states to find out the thermal properties. It is time to perform Monte Carlo simulations. Though we can not sum over all states, we can choose samples randomly, and get the expectation values of observables. For instance, if we do sampling of states with uniform probability, the approaches of expectation values of observable could be written as

$$\langle O \rangle \approx \bar{O} = \frac{\sum_{i=0}^N O(s_i) e^{-\beta H(s_i)}}{\sum_{i=0}^N e^{-\beta H(s_i)}}, \quad (2.3)$$

which is similar to Eq. (2.2), but the series  $\{s_1, s_2, \dots, s_N\}$  represents states chosen randomly. As  $N$  increases,  $\bar{O}$  approaches to  $\langle O \rangle$  with decreasing statistical error. However, besides sampling more times, how to choose the random samples plays an important role to approach the real expectation. From Eq. (2.3), we find that if we choose the samples with the uniform distribution, many of them may have very small contribution to our result because of their small Boltzmann factors, thus wasting computational resources.

### 2.1.1 Importance sampling and Markov chain Monte Carlo

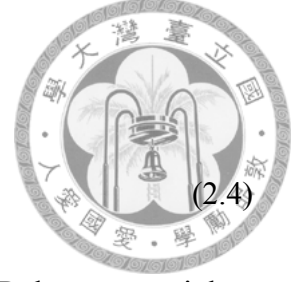
For higher efficiency, instead of choosing samples with the uniform distribution, we choose samples according to their weight. It is called the importance sampling. If choosing states with the Boltzmann weights directly, the expectation value can be written as an





average of the observable  $O(s)$

$$\bar{O} = \frac{\sum_{i=0}^N O(s)}{N}, \quad (2.4)$$



and we would not waste time to go through many states with small Boltzmann weights. Therefore, the problem is how to choose samples according to the weights. A common solution is using the Markov chain. In a Markov chain, the probabilities of next states only depend on the present states. When we perform sampling, we start from a randomly chosen state  $s_1$  and generating next state with probability only depending on present state  $P(i \rightarrow i + 1)$ . The sequence will be

$$s_1 \rightarrow s_2 \rightarrow \cdots \rightarrow s_{i-1} \rightarrow s_i \rightarrow s_{i+1} \rightarrow \dots \quad (2.5)$$

From a Markov chain, after many steps, there will be an unique steady-state probability distribution depending on the transition probabilities. Therefore, by choosing appropriate transition probability, the unique steady-state probability distribution can become equal to the Boltzmann distribution, and importance sampling can be implemented. For the Metropolis algorithm, assuming the Markov chain reversible, the transition probability should obey the condition of detailed balance

$$\pi(s_a)P(s_a \rightarrow s_b) = \pi(s_b)P(s_b \rightarrow s_a), \quad (2.6)$$

where  $\pi(s_a)$  and  $\pi(s_b)$  mean the steady-state weight of state  $s_a$  and  $s_b$ . Let the steady-state probability distribution be the Boltzmann distribution. The transition probability depends on the energy difference of the states, written as

$$\frac{P(s_a \rightarrow s_b)}{P(s_b \rightarrow s_a)} = \frac{\pi(s_b)}{\pi(s_a)} = e^{-\beta(E(s_b)-E(s_a))} = e^{-\beta\Delta E}. \quad (2.7)$$

Therefore, we generate the next state using the probability of Eq. (2.7). At first, choose a state  $s'$  as a candidate of  $s_{i+1}$ . With the acceptance rate  $P(s_i \rightarrow s')$ , if the the change accept,  $s_{i+1} = s'$ ; otherwise,  $s_{i+1} = s_i$ . For lattice systems, we perform single-spin flips, i.e. choosing the state with a spin flipped as the candidate. Following are the steps of the



algorithm:

1. Choose an initial configuration.
2. Choose a site  $p$ .
3. Calculate the energy difference  $\Delta E$  which occur if the spin at  $p$  is flipped.
4. Flip the spin with the probability  $\min(1, e^{-\beta\Delta E})$ .
5. Repeat steps 1 to 3.

A Monte Carlo sweep is the process repeating steps 1 to 3 at all sites on the lattice once, which is the unit of Monte Carlo updates. In our Monte Carlo simulations, at first, many Monte Carlo sweeps are performed as the equilibration to make the Markov chain reach the correct distribution, and then measurement is taken after every or every few Monte Carlo sweeps. After many measurements, we can finally get the expectation values of observables by Eq. 2.4.

### 2.1.2 Measurements

To perform a measurement on internal energy  $E$  and the order parameter, Eq. (1.21) and (1.22) or Eq. (1.19) and Eq. (1.20) are used for XY Models and Ising Models. Besides the properties can be measured from microstates, we are also interested in macroscopic properties measured with the thermal fluctuations, such as susceptibility and specific heat. Using the relation in thermal dynamics and statistical mechanics[1], the specific heat can be measured with the fluctuation of internal energy.

$$C_v = \frac{1}{N} \frac{dE}{dT} = \beta^2 \frac{\langle E^2 \rangle - \langle E \rangle^2}{N} = \beta^2 N \langle \mathbf{e}^2 \rangle - \langle \mathbf{e} \rangle^2, \quad (2.8)$$

where  $\mathbf{e}$  represents  $E/N$ , the energy per site. And, susceptibility is given by

$$\chi = \beta N (\langle M^2 \rangle - \langle M \rangle^2). \quad (2.9)$$

In the thermodynamic limit, as Eq. (1.3) and Eq. (1.5) show, the specific heat and the susceptibility diverge at  $T_c$ , and the critical behaviors are described by the critical exponents,  $\alpha$  and  $\gamma$ . In the simulations of finite system, instead of divergence, there are peaks near  $T_c$ , and the critical exponents can be estimated with the finite-size behaviors, which we will introduce in next chapter.

The correlation length is also measurable. Instead of using the definition based on the decay of the correlation function, which is hard to be measured in the simulations, we can define it in a way easy to practice the measurement. First, consider the Fourier transform of the correlation function, which is called the structure factor,

$$S(\mathbf{q}) = \langle s_{-\mathbf{q}} s_{\mathbf{q}} \rangle = \sum_{\mathbf{r}} e^{-i\mathbf{q}\cdot\mathbf{r}} \Gamma(\mathbf{r}) = \sum_{\mathbf{r}} \cos(\mathbf{q} \cdot \mathbf{r}) \Gamma(\mathbf{r}), \quad (2.10)$$

where  $s_{\mathbf{q}}$  is Fourier transform of the spin configuration,

$$s_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_j \sum s_j e^{-i\mathbf{q}\cdot\mathbf{r}_j}. \quad (2.11)$$

Since we study on the ferromagnetic systems, the wave-vector of the dominant correlation  $\mathbf{Q}$  is  $(0, 0, 0)$ .  $S(q)$  represents  $S(\mathbf{Q} + q\hat{u})$ , where  $\hat{u}$  denote the reciprocal-space unit vector, which can be in any direction, and for convenience, we always choose in  $x$ -direction. A correlation length  $\xi_a$  is defined with structure factor of  $q = 0$  and  $q_1 = 2\pi/L$

$$\xi_a = \frac{1}{q_1} \sqrt{\frac{S(0)}{S(q_1)} - 1}, \quad (2.12)$$

which holds a relation with the original correlation function  $\xi$

$$\xi_a = \xi \sqrt{\frac{(1+d)(3+d)}{8d}}. \quad (2.13)$$

Fig. 2.1 shows the results of  $\xi_a/L$  for XY model. Unlike the original correlation length, since this correlation length is not defined using the connected correlation function, and it goes to infinity in the ordered phase.

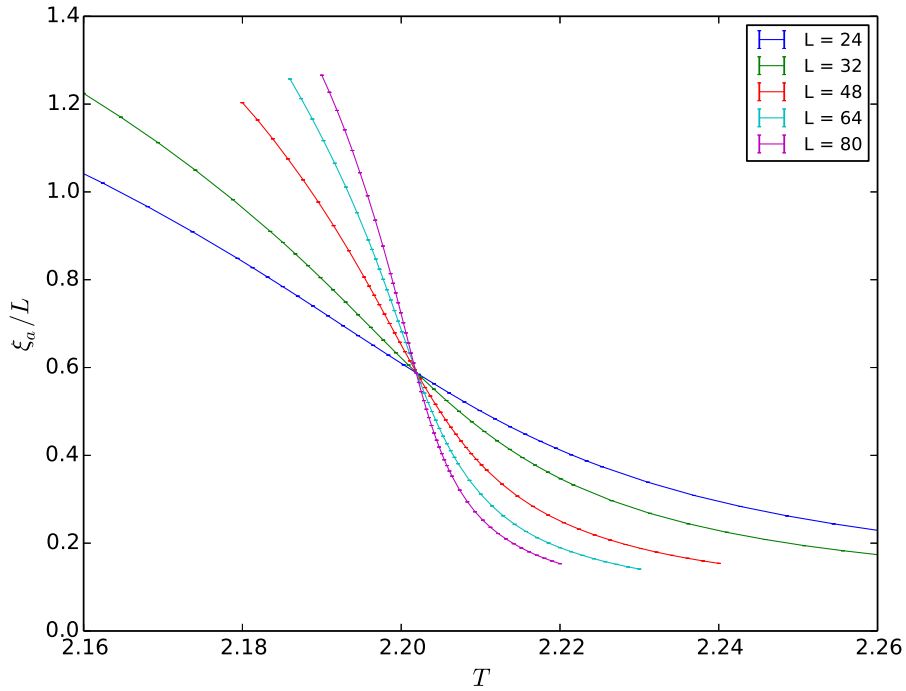
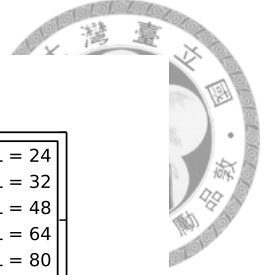


Figure 2.1: Correlation length of XY model defined by Eq. (2.12).

### 2.1.3 Parallel tempering

In many systems, for example, glassy systems, the states may be hard to reach the equilibrium, and this equilibration takes much time. The configuration may trap in a *local minimum* and hard to jump to another configuration for the rugged free energy landscape, as shown in Fig. 2.2 with the dashed red line. The probability of jumping from a local minimum configuration to another is very small.

The solution of this problem is the parallel tempering[6, 15]. For a descending temperature series  $T_1, T_2, \dots, T_M$ , simulate  $M$  replicas of the system in parallel at different temperatures. The set of these replicas is considered in a large ensemble. In this ensemble, it is allowed that replicas at different temperatures can be exchanged. As we perform the parallel tempering, the replicas at every neighbor temperatures  $T_j$  and  $T_{j+1}$  are exchanged with the acceptance probability

$$P_{C_j \leftrightarrow C_{j+1}} = \min(1, e^{(E(C_j) - E(C_{j+1}))(\beta_j - \beta_{j+1})}), \quad (2.14)$$

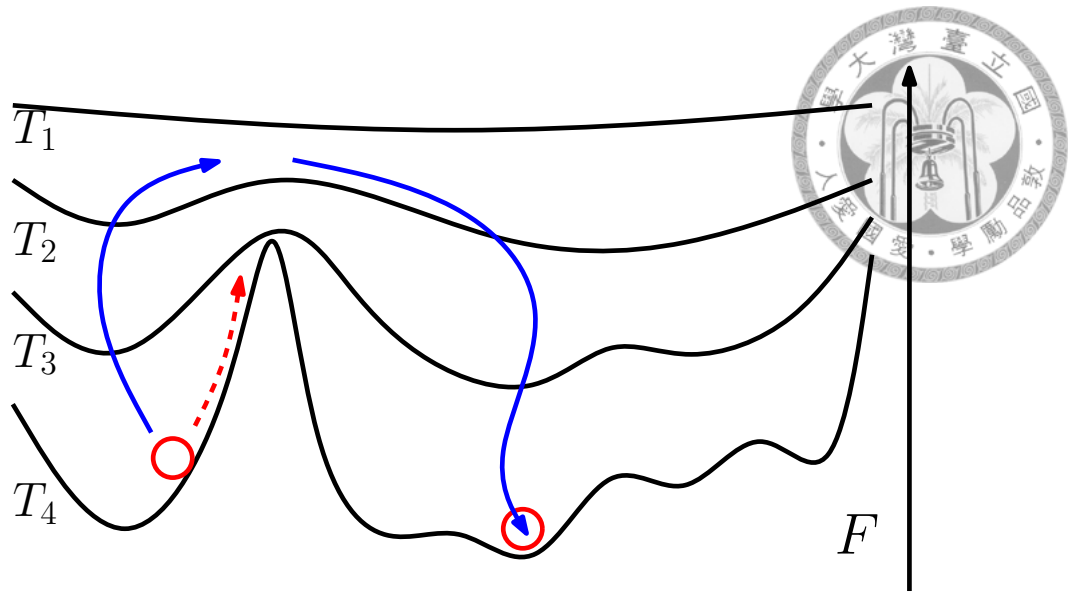


Figure 2.2: Sketch map of parallel tempering and the free energy landscape

where  $C_j$  represents the configuration of replicas. This process satisfies the detailed balance and is used to update the system. With parallel tempering, it is easier for the system to cross barriers in the landscape as the blue solid line shown in Fig. 2.2. Therefore, as we perform Monte Carlo simulations, the parallel tempering is carried out after every or every few Monte Carlo sweeps.

To perform parallel tempering, there are many configurations which should be updated in parallel. In addition, to get more precise results, we want to simulate large systems. However, as the number of spins increases, the time grows in proportion. To accelerate the simulations, graphics processing units are used.

## 2.2 GPU architecture and CUDA framework

For the demand for high-definition 3D graphics, nowadays, graphics processing units (GPU) are designed as high-parallel and multi-core processors [16]. The GPU cores focus on the calculations of floating point numbers and are normally much slower than CPU cores. With hundreds of processors, GPUs have very impressive performance on problems that can be expressed as data-parallel computations.

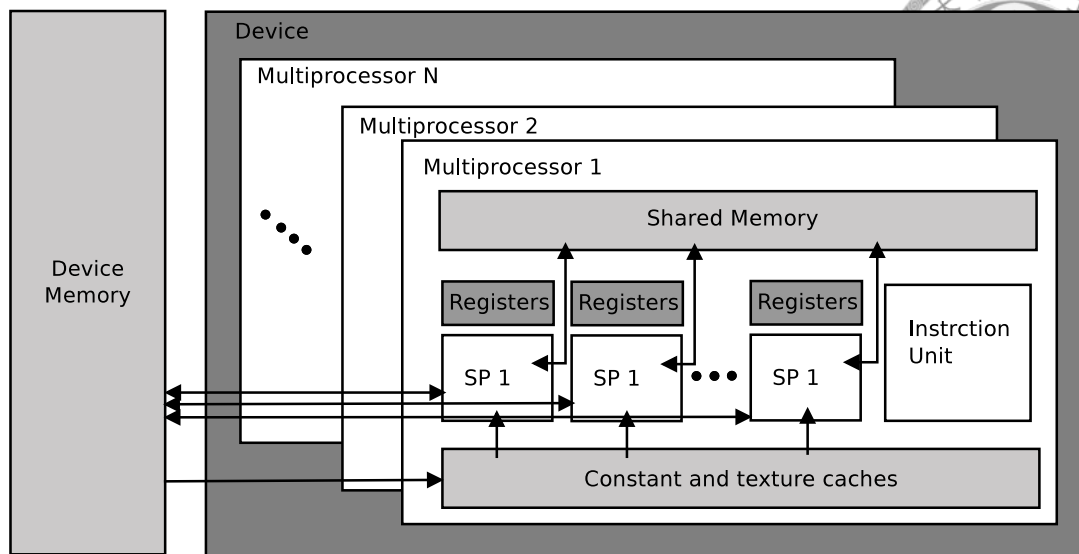


Figure 2.3: GPU layout

Take the Tesla architecture of NVIDIA as an example. It is composed of a scalable array of streaming multiprocessors(SMs), which contains several cores, called scalar processor(SP) cores. A SP core is the unit carrying out one of the parallel computations. Each core can execute a floating point instruction per clock cycle. There is global device memory, which every core in every processor can read/write, but the speed is limited by the memory bus. Texture and constant memory are also located on the device memory, but there are read-only cache for the multiprocessors. The shared memory is in every SM, and it can be read or written by cores in the same multiprocessors.

CUDA is a scalable parallel programming model and software environment. As an extension of C language, it provide methods to parallelize calculations into GPU cores. The basic idea is organizing parallelized calculation into *blocks* and *threads*. In implementation, we write a GPU kernel function in C which will be executed by every thread. Several threads are grouped in a block, and they share the same shared memory, but the number of threads in a single block has a upper bound determined by the device. Every thread and block has its unique ID. We use the thread and block ID to perform the calculation of different data for the same kernel function. For example, the thread ID can be used to decide the data position. Since the reading or writing of data of different thread may interrupt each other and it may cause the race condition, after changing the data, there

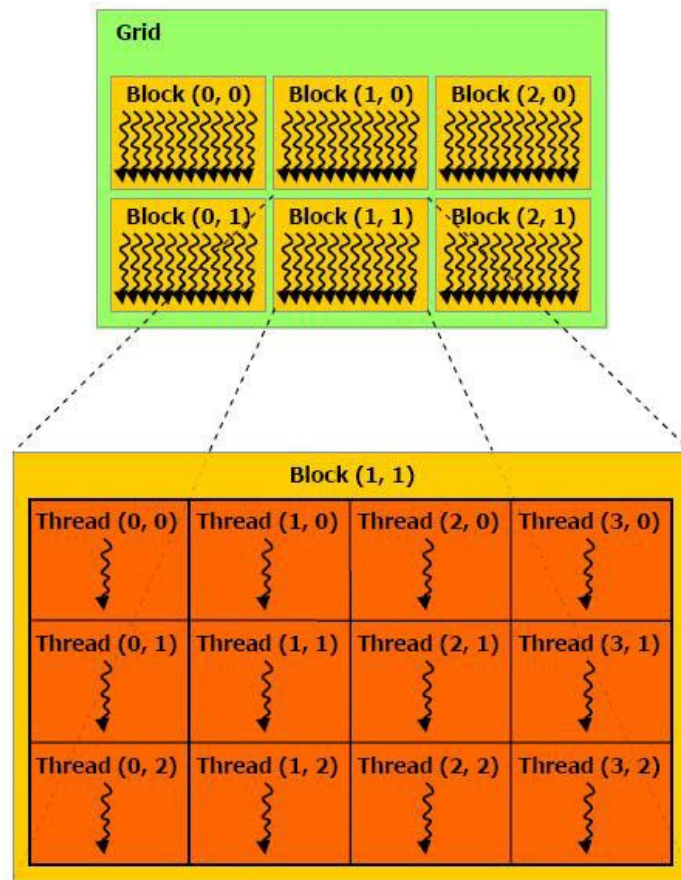


Figure 2.4: A grid of blocks, containing threads

should be synchronization to wait for all threads reach the same point. To communicate with host function executed on CPU, duplication between device memory and host memory is necessary. Therefore, how to allocate data on different types of memory is dominant in performance.



The architecture for managing large number of threads running on multiprocessors is called single-instruction multiple thread (SIMT). The multiprocessors will map every thread into one SP core, and each thread executes independently. The SIMT group 32 threads into a warp, which are handled in the same SM with the same instructions. If there is any condition expression in the kernel function, it will result in warp divergence. The warp execute each conditional branch with the thread not on the branch ignored. It implies that it takes the summation of the time taking on all branches, so it is better to avoid using condition expressions and loops in kernel functions.

## **2.3 Implementation of classical Monte Carlo simulation on GPU**

To implement Monte Carlo simulations on a GPU, the first problem is how to parallelize the process of Monte Carlo simulations. There must exist some steps executed over and over again that can be parallelized and executed in many threads. The single-spin flip on a site (step 2 to step 4 in Metropolis method) seems to be accord with these conditions. However, the interactions between spins are changed while spins flip. Trying to flip two spins with interaction simultaneously will cause the race condition and the inconsistency between the acceptance rates and the configuration of spins. Therefore, updating all spins with parallel thread in the same time is invalid.

Our scheme is a method called checkerboard update algorithm[17]. The basic idea is dividing spins into several sublattices, in one of which spins in the same sublattice have no interaction with each other. Take the system with nearest-neighbor interactions as an example, shown in Fig. 2.5. The internal energy of interactions of a spin in the red sublattice





tice only depends on the spins in the white sublattice and itself, so we can update all spins in the red or white sublattice in parallel with no race condition. Therefore, a Monte Carlo sweep is broken into one red sublattice update and one white sublattice update. Rather than update every single spin in sequence, GPU cores update half of spins simultaneously, and it takes much less time with a large number of GPU cores.

As the blue dotted line shown in Fig. 2.5, rather than handle only one spin, a kernel function may handle several spins. In our program, a kernel function updates spins in one sublattice in a  $2 \times 2 \times 2$  cubic. For convenience, take the 2D case as an example. The lattice is divided into many  $2 \times 2$  squares, and every thread handles one of them. In the kernel function for red sublattice, every thread updates left-top and right-bottom spins in its own square, while in the other kernel function, every thread updates left-top and right-bottom spins. As an example, in our 3D XY simulations for a  $48 \times 48 \times 48$  lattice, since a thread handle a  $2 \times 2 \times 2$  cubic, there is  $24 \times 24 \times 24$  threads. We group  $8 \times 8 \times 4$  threads as a block. There are 54 blocks, and each of them govern  $16 \times 16 \times 8$  spins.

The measurements of energy and magnetization can also perform on the GPU with parallelized process. We use the same assignment of spins, threads and blocks. The first step is that each thread calculates the local summation of observables, and restore it in

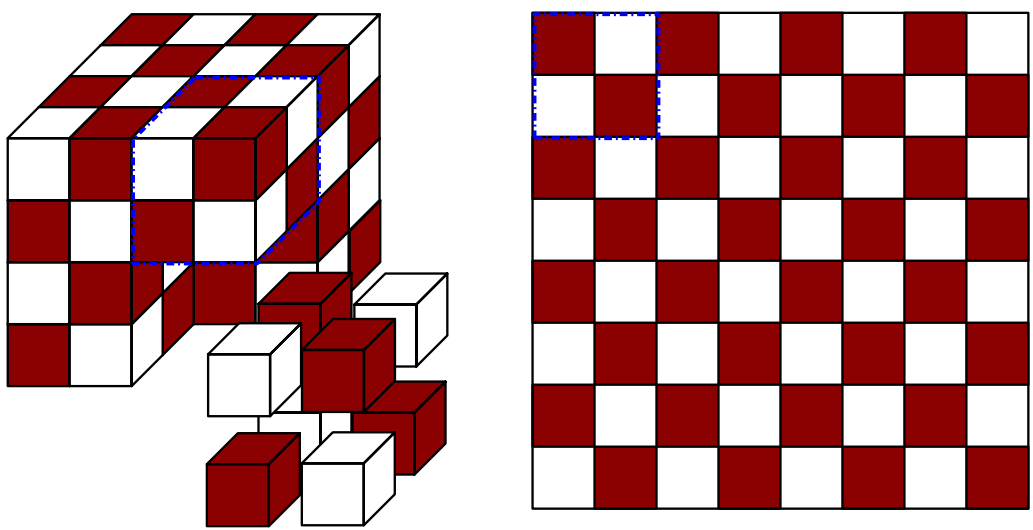


Figure 2.5: Sublattices of 2D and 3D system with nearest-neighbor interaction for the checkerboard algorithm.

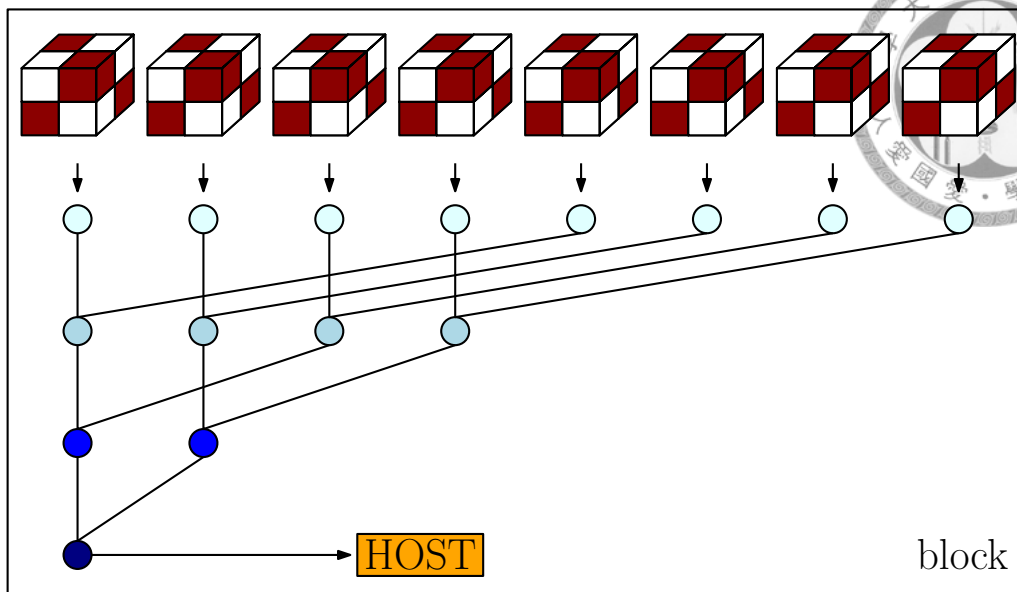


Figure 2.6: Summation over all threads in a block.

shared memory. There is no need to consider sublattices, because there is no race condition here. As shown in Fig. 2.6, the next step is summing up the local energy or magnetization in a block using a tree. With this method, only  $\log_2 N$  summations are carried out for each thread, and we only need to transfer the summation of the observables of spins in blocks to the host.

To implement the parallel tempering, all replicas are updated simultaneously, and the temperatures in different threads are determined by the thread ID and block ID. When carrying out the replica-exchange, we exchange the order of temperatures on CPU and copy them to device memory. Therefore, the only two situations to transfer the spin configurations between host and kernel are just the begin and the end of the simulation. Minimizing the data transferred is important for the speed-up.

Comparing with the performance of CPU, the speed-up is significant. In our simulation on XY models, for a update, a Monte Carlo sweep and a parallel tempering are carried out, and we measure helicity, magnetization, energy and correlation length. On average, about  $3 \times 10^8$  spins can be updated and measured per second.

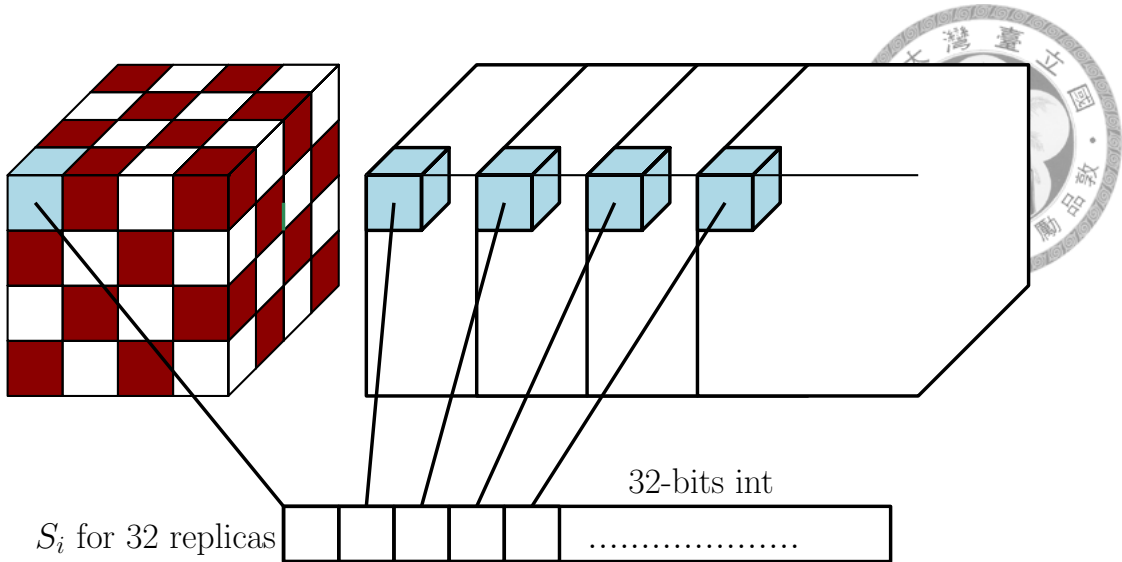


Figure 2.7: Allocation of values of spins.

## 2.4 Multispin coding on GPU

For the Ising model, the value of a spin  $S^o \in \{-1, 1\}$  can be just represented in one bit with  $S \in \{0, 1\}$ . However, the size of a variable in computers can not be less than 1 byte. In Ref. 18, to make the best usage of the memory, the values of many spins are stored in one integer. For example, in a 32-bit integer, 32 values are stored. This method is called multispin coding.

To perform the multispin coding, the first step is to allocate the value of spins. In Ref. 18, they show that the unified allocation has the good efficiency. According to the unified allocation, the value of spins on the same site in many replicas are stored in one integer, as shown in Fig. 2.7. The replicas can be at different temperatures for parallel tempering or different copies for glassy systems.

Besides reduction of the memory use, the calculation of spins within an integer can be done together. For a single-spin flip, the calculations for internal energies of every single spin are necessary. These calculations can be accelerated by replacing the multiplication of 32 bonds by a bitwise XOR, as shown in Fig. 2.8. Since we represent the interaction  $J^o \in \{-1, 1\}$  and  $S^o$  by  $J \in \{0, 1\}$  and  $S$ , there are a shift and rescale between the real energy and the result gotten by the bitwise XOR. However, since we find the probabilities

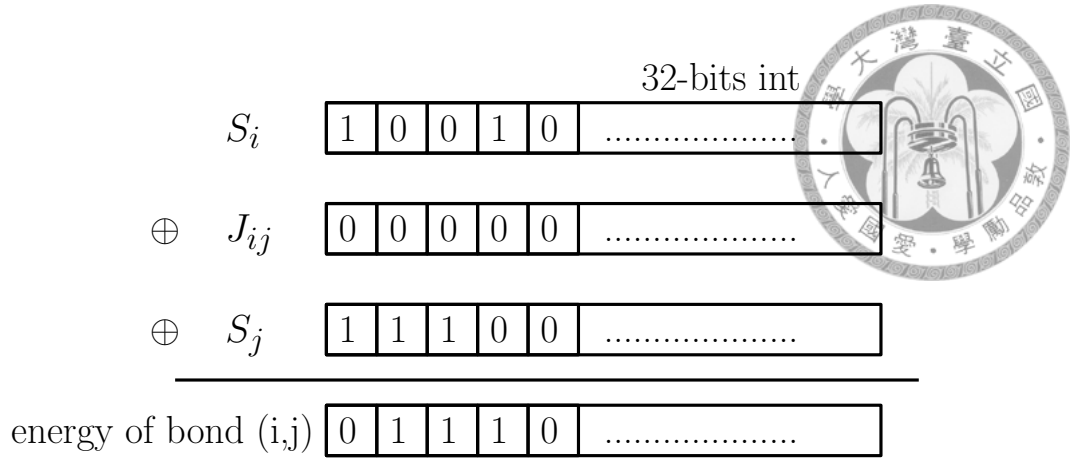


Figure 2.8: Calculation of energies of bonds by a bitwise XOR.

with the look-up table, this is not a problem.

To perform a single-spin flip, we should sum over all the bond energies of a spin, and this value can not be stored in one bit. Therefore, there are many different schemes to do the bitwise XOR and the summation. In Ref. 18, they show the scheme, called Compact Asynchronous Multispin Coding(CAMSC), with the best efficiency.

As Fig. 2.9 demonstrates, at first we get the bond energies by the bitwise XOR. Since we need three bits for the storage of the local energy of a spin, the replicas are divided into four groups. For the  $x - th$  group, we extract the energies for  $4n + x th$  replicas by the bitwise AND. Now, there are four bits to store the local energy, so we can sum the energies of bonds directly and get the local energies of the replicas in the  $x - th$  group. Even though the calculations are broken into four groups, the calculations are much faster than individual calculations of the energy of single spins. In my implementation, comparing with the program without multispin coding, the single-spin flip performed with CAMSC is four times faster.

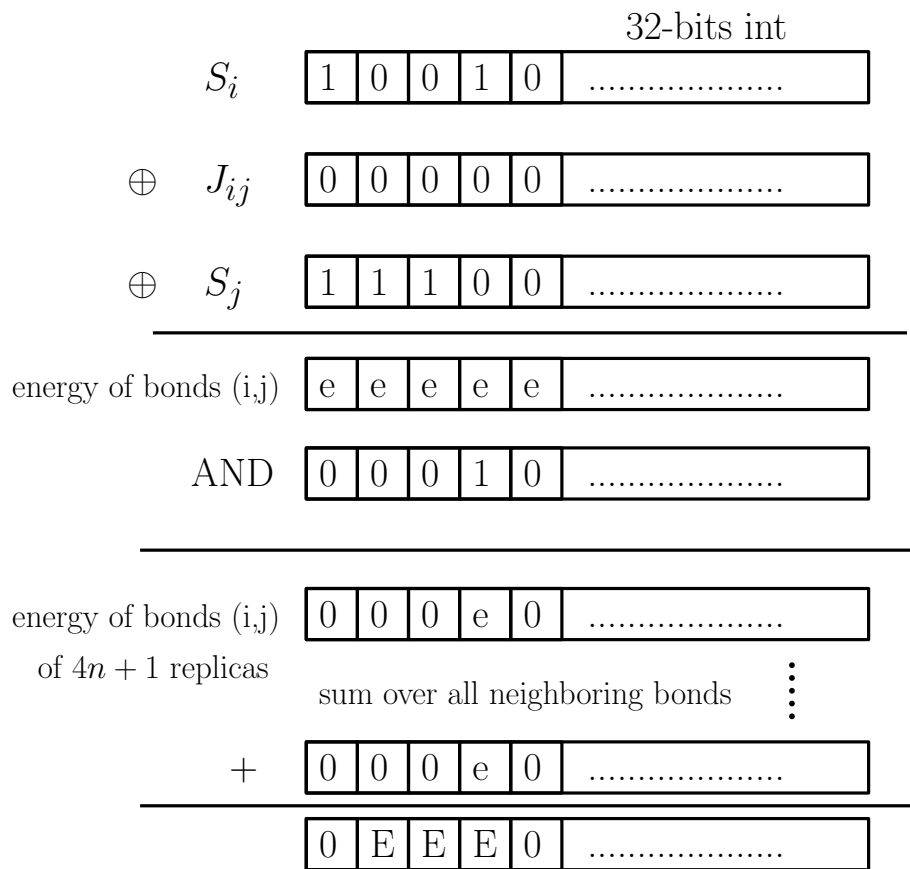


Figure 2.9: CAMSC scheme to calculate the energies of every single spin.



## Chapter 3

# Data Analysis and Finite-Size Scaling

In the previous chapter, we have mentioned the way to get the approximations of observables from the importance sampling. In this chapter, we present the details about data analysis of observables and how to estimate the critical exponents.

### 3.1 Data analysis

The expectation value of the observable  $O$  obtained from Monte Carlo simulation is given by

$$\langle O \rangle \approx \bar{O} = \frac{1}{N} \sum_{i=1}^N O(s_i), \quad (3.1)$$

where  $O(s_i)$  represents a measurement of quantity  $O$  of the configuration  $s_i$  and  $N$  is the number of measurement. However, since  $\bar{O}$  is just a approximation, we must estimate the statistical error. We define the standard deviation of the mean values as the error of the mean value

$$\sigma_{\bar{O}}^2 = \langle \bar{O}^2 \rangle - \langle \bar{O} \rangle^2. \quad (3.2)$$

According to the equation of variance, we have to repeat the sets of measurements for many times to get many mean values to estimate the error. It is neither realistic nor nec-

essary, because the variance of the mean values can be estimated by the variance of measurements.



We start from combining Eq. (3.1) and Eq. (3.2), getting

$$\sigma_O^2 = \frac{1}{N^2} \sum_{i,j} \langle O_i O_j \rangle - \frac{1}{N^2} \sum_{i,j} \langle O_i \rangle \langle O_j \rangle, \quad (3.3)$$

where  $O(\sigma_i)$  is represented by  $O_i$ . We rearrange (3.3) and get

$$\sigma_O^2 = \frac{1}{N^2} \sum_i (\langle O_i^2 \rangle - \langle O_i \rangle^2) + \frac{1}{N^2} \sum_{i \neq j} (\langle O_i O_j \rangle - \langle O_i \rangle \langle O_j \rangle). \quad (3.4)$$

We assume that all measurements are independent, so the second term should be zero, and the first term can be represented by the variance of individual measurements, so the variance of means can be written as

$$\sigma_O^2 = \frac{\sigma_O^2}{N}, \quad (3.5)$$

which is the relation between the variance of means and the variance of single measurements.

In the previous derivation, we assume that the measurements are independent, but the measurements between two consecutive configurations generated by Metropolis algorithm are strongly correlated. We can observe this phenomena in Fig. 3.1, the magnetization in  $x$  direction of the XY model. It takes thousands of MC sweeps to get statistical independent states. The feature of correlation between two measurements can be described by the autocorrelation function

$$A_O(t) = \frac{\langle O_{i+t} O_i \rangle - \langle O \rangle^2}{\langle O^2 \rangle - \langle O \rangle^2}, \quad (3.6)$$

where  $t$  is the simulation time between two measurements. The autocorrelation function asymptotically decay in exponential,  $A_O(t) \sim e^{-t/\tau_A}$ , where  $\tau_A$  is the autocorrelation time[4, 9]. Instead of Eq. (3.5), considering the correlation, the error must have a correc-

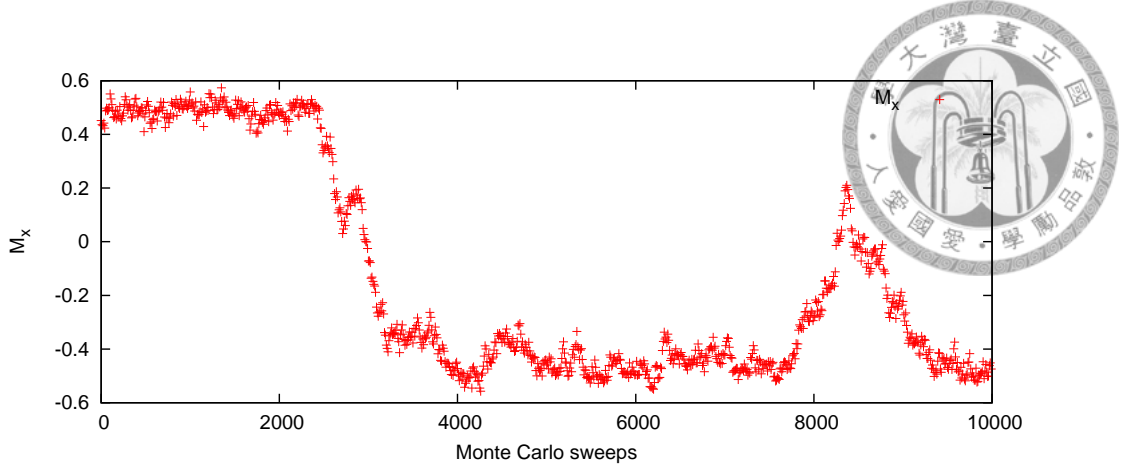


Figure 3.1: The magnetization in  $x$  direction versus Monte Carlo sweeps.

tion term contributed by the correlation. Starting from Eq. (3.4), we can get the error

$$\sigma_{\bar{O}}^2 = \frac{\sigma_O^2}{N}(1 + 2\tau_A). \quad (3.7)$$

However, estimating the autocorrelation time is complicated. In practical, we would not estimate the autocorrelation time and use Eq. (3.7) to get the correct errors. Instead, we perform the data-binning, i.e. we group data into bins with length  $M$ . If the length  $M$  is much longer than the autocorrelation time, the mean value and variance of the mean values can be written as

$$\bar{O} = \frac{1}{B} \sum_{b=1}^B O(\sigma_b), \quad \sigma_{\bar{O}}^2 = \sigma_{O_b}^2/B. \quad (3.8)$$

A simple way to understand why Eq. (3.8) works is starting from Eq. (3.7). Replacing  $\bar{O}$  with  $O_b$ , we can find that the factor contributed by the autocorrelation is fully counting in  $\sigma_{O_b}^2$  for  $M$  large enough. Therefore,  $O_b$ 's are statistical independent. The method to examine whether  $M$  is large enough is re-binning, combining  $M_b$  bins into a larger bin with size  $M_b M$ . If  $M$  is large enough, the error we get from Eq. (3.8) will not change. There are many other advantages to store all bin averages, For example, that make it easy to add new data and perform *Bootstrap* and *Jackknife* resampling on the data.



## 3.2 Finite-size scaling



Like Eq. (1.16), the scaling ansatz of quantity  $Q$  can be written as a homogeneous function

$$Q(t, l) = l^\sigma f(tl^{1/\nu}), \quad (3.9)$$

where  $x$  is an exponent depending on  $Q$ , and  $l$  is the length of coarse graining. In the thermodynamic limit, the length of coarse graining can be the correlation length,  $l = t^{1/\nu} \sim \xi$ . For finite sizes, the largest length of coarse graining is the size of the system. Eq. 3.9 for systems with infinite size and finite size can be written down

$$Q(t) = t^{\nu\sigma} f(1) \quad , \text{ for } L \rightarrow \infty, \quad (3.10)$$

$$Q(t) = L^\sigma f(tL^{1/\nu}) \quad , \text{ for finite } L, \quad (3.11)$$

where  $L$  here represents the size of system. Comparing Eq. 3.10 with the definition of critical exponents, we can write the finite-size scaling form of thermal quantities as

$$m = L^{-\beta/\nu} \mathcal{M}_0(tL^{1/\nu}), \quad (3.12)$$

$$\chi = L^{\gamma/\nu} \chi_0(tL^{1/\nu}), \quad (3.13)$$

$$C_v = L^{\alpha/\nu} \mathcal{C}_0(tL^{1/\nu}), \quad (3.14)$$

$$\xi = L\xi(tL^{1/\nu}). \quad (3.15)$$

With these finite-size scaling forms, the critical exponents can be extracted from simulation results of different sizes. Shown in Fig. 3.2, by tuning the critical exponents in the scaling form, the data of different sizes collapse to a universal curve of the scaling function  $f(x)$ . Using the data-collapse, the approximation of critical exponents are extracted. In this thesis, to perform data-collapse, we use the program by Harada [19] based on Bayesian inference.

According to the scaling form, if the scaling function of quantity  $Q$  has maximum near  $t = 0$ , such as  $\chi_0$  and  $\mathcal{C}_0$ ,  $Q_{max}$  is proportional to  $L^\sigma$ . Therefore, we can get the value

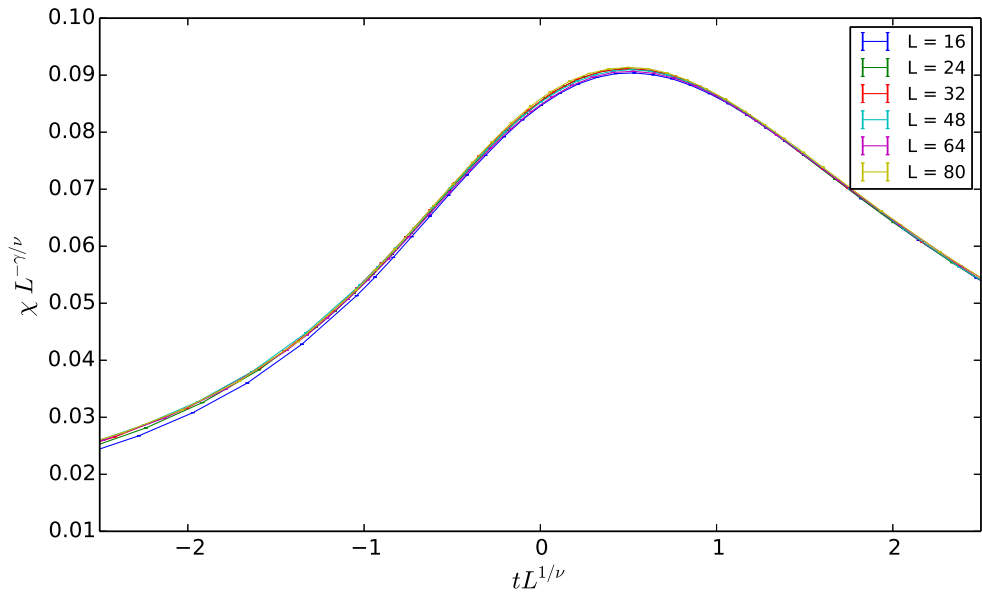


Figure 3.2: Data collapse of  $\chi$  of the 3D XY model.

of  $\gamma/\nu$  or  $\alpha/\nu$  by finding the maximum. With linear regression of the logarithm of the maximums and sizes, we can obtain the values. Besides, since for  $t = 0$ ,  $f(tL^{1/\nu})$  is a constant, we also can find critical exponents with  $Q_{t=0}$ .

To extract the critical temperature and the critical exponent  $\nu$ , we use a set of parameters, which are called phenomenological coupling [8, 20]. A phenomenological coupling holds the property that it is size independent at  $T_c$ . A frequently-used one is the Binder ratio

$$R_4 = \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2}, \quad (3.16)$$

which is defined by the distribution of order parameter. We can also define ratios of other powers of  $m$ , such as  $R_2 = \frac{\langle m^2 \rangle}{\langle |m| \rangle}$ . For convenience, we can rescale the Binder ratio and get the Binder cumulant,

$$U_4 = \frac{n+2}{2} \left( 1 - \frac{n}{n+2} \right), \quad (3.17)$$

where  $n$  is the number of components of order parameter. For  $L \rightarrow \infty$ ,  $U_4 \rightarrow 1$  for  $T < T_c$ , while  $U_4 \rightarrow 0$  for  $T > T_c$ . According to Eq. (3.15),  $\xi/L$  is also a parameter

independent of size at  $T_c$ .

If we do not consider the correction terms, the phenomenological couplings will exactly cross at  $T_c$ . The derivatives of them near  $T_c$  is proportional to  $L^{1/\nu}$ , for example,

$$\frac{\partial U_4}{\partial T} \sim L^{1/\nu}. \quad (3.18)$$

The logarithmic derivatives of  $m^n$  holds the same property. In this thesis, using this property, we do fittings to find  $\nu$ .

However, there exist correction terms caused by the irrelevant scaling fields. The sub-leading correction term is proportional to  $L^{-w}$ , where  $w$  depends on the system. It is very hard to fitting many exponents simultaneously. Therefore, in this thesis, our strategy is to ignore the correction terms and simulate with larger size to diminish the correction terms.



## Chapter 4

# Three-dimensional XY model with $Z_q$

## Anisotropy

In this chapter, we study the three-dimensional XY model and the case with  $Z_q$  anisotropy, which is described with the Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j) - h \sum_i \cos(q\theta_i), \quad (4.1)$$

where  $h$  is the strength of anisotropy. We study the ferromagnetic case  $J > 0$ , and for convenience, we assume  $J = 1$  in the following.

### 4.1 3D XY model

The continuous phase transition in 3D XY model and the 3D XY universality class are widely studied, because there are physical systems with critical behaviors in the 3D XY universality class. The most famous one is the  $\lambda$ -transition of  $^4\text{He}$ . In Ref. 13, they carried out the measurements of the specific heat in the superfluid transition in zero gravity. They estimated the value of the specific heat exponent  $\alpha = -0.0127(3)$ , and by the hyperscaling relation, we know the value of the exponent  $\nu = 0.6709(1)$ .

Following this experimental result, Burovski *et al.* [21] obtained  $\nu = 0.6717(3)$  with the high-precision simulations, and Campostrini *et al.* [11] estimated  $\nu = 0.6717(1)$  with the high temperature expansion and the Monte Carlo simulations. They use finite-size scaling to analyze the results from Monte Carlo simulations. They considered the sub-leading correction term, since it will cause the systematic errors especially with small lattice sizes. And, they perform the simulations of the improved model of the  $\phi^4$  model to minimize the correction.

In our simulations, we perform simulations only up to  $L = 80$ , and do not consider the correction term. We perform fitting to find the log-log slope of the sizes and the maximums of logarithmic derivative of  $\langle |m| \rangle$ ,  $\langle m^2 \rangle$  and  $\langle m^4 \rangle$  with our simulation results of the sizes  $L = 48, 64, 80$ , obtaining  $\nu = 0.6708(1)$ . We also get  $\gamma = 1.324(4)$  by the maximum of susceptibility  $\chi$ .  $T_c = 2.2018(1)$  is estimated using the Binder cumulant and correlation length. Our results are consistent with previous results.

Our results of per site energy  $\langle e \rangle$ , magnetization  $\langle m \rangle$ , susceptibility  $\chi$ , specific heat  $C_v$ , Binder cumulant and correlation length are shown as follows.

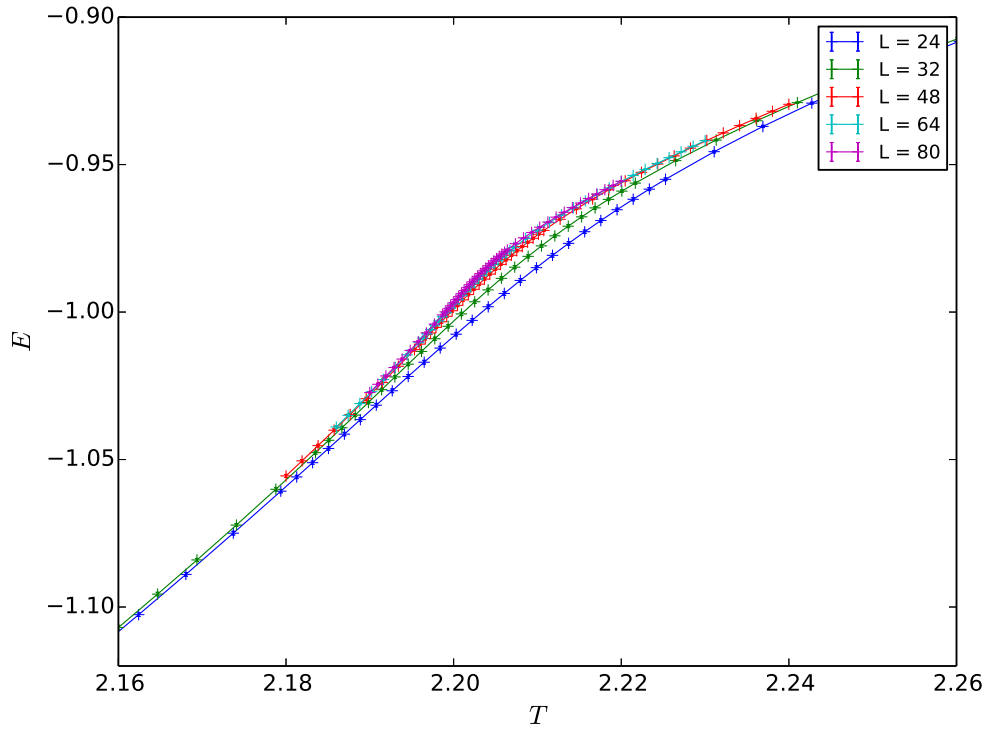
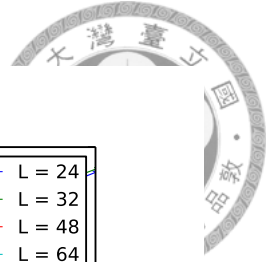


Figure 4.1: Energy per site v.s.  $T$  of the XY model.

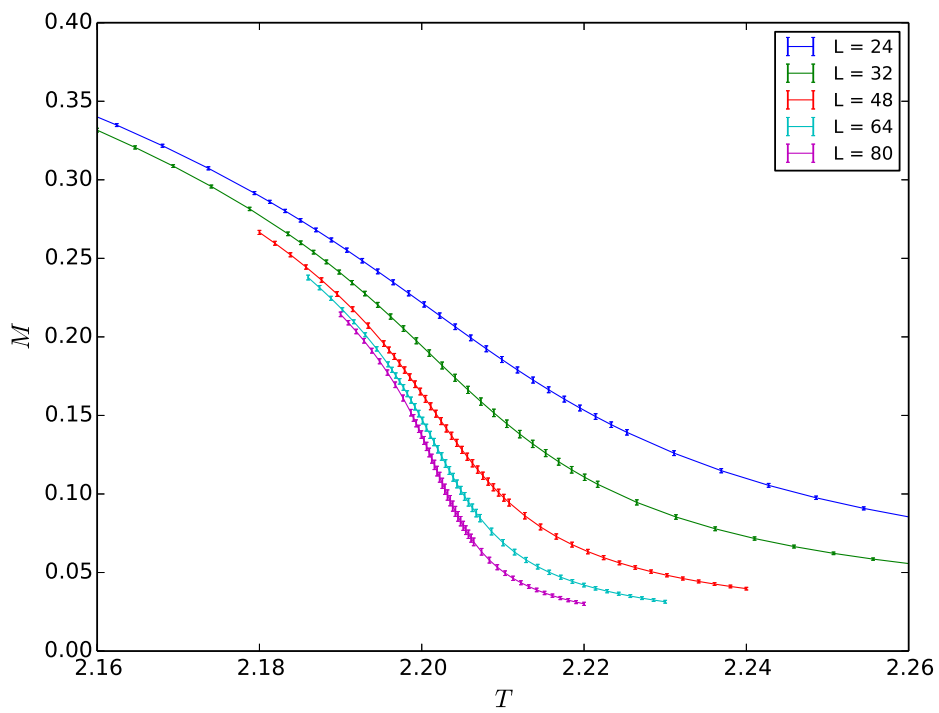


Figure 4.2: Magnetization v.s.  $T$  of the XY model.

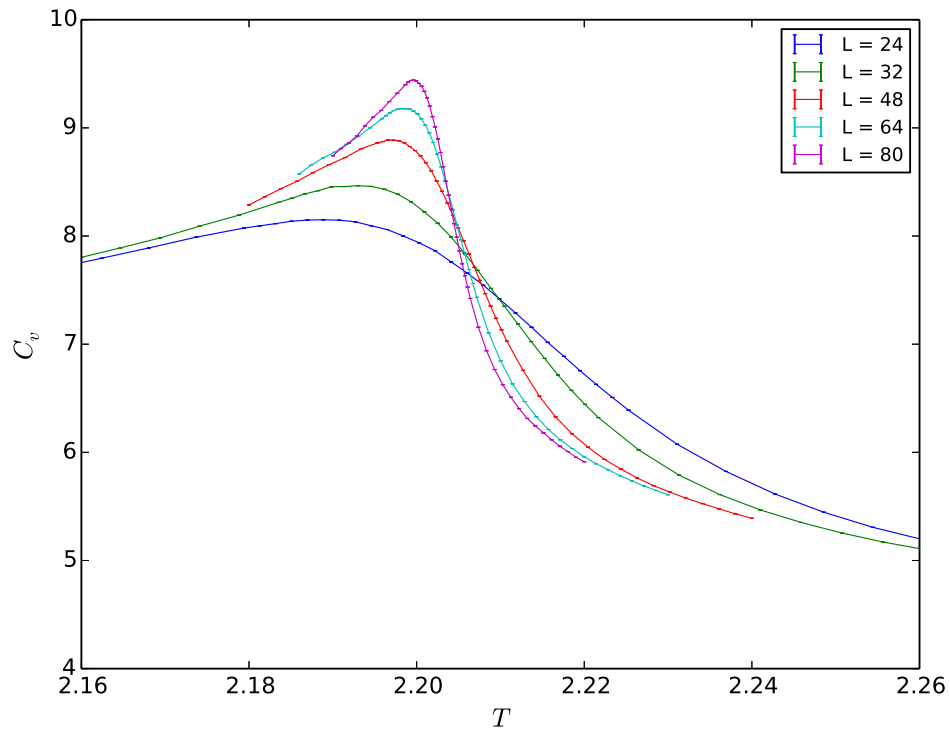
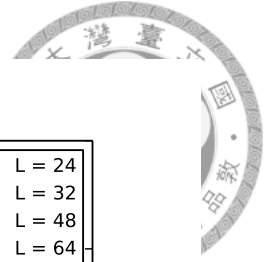


Figure 4.3: Specific heat of the XY model.

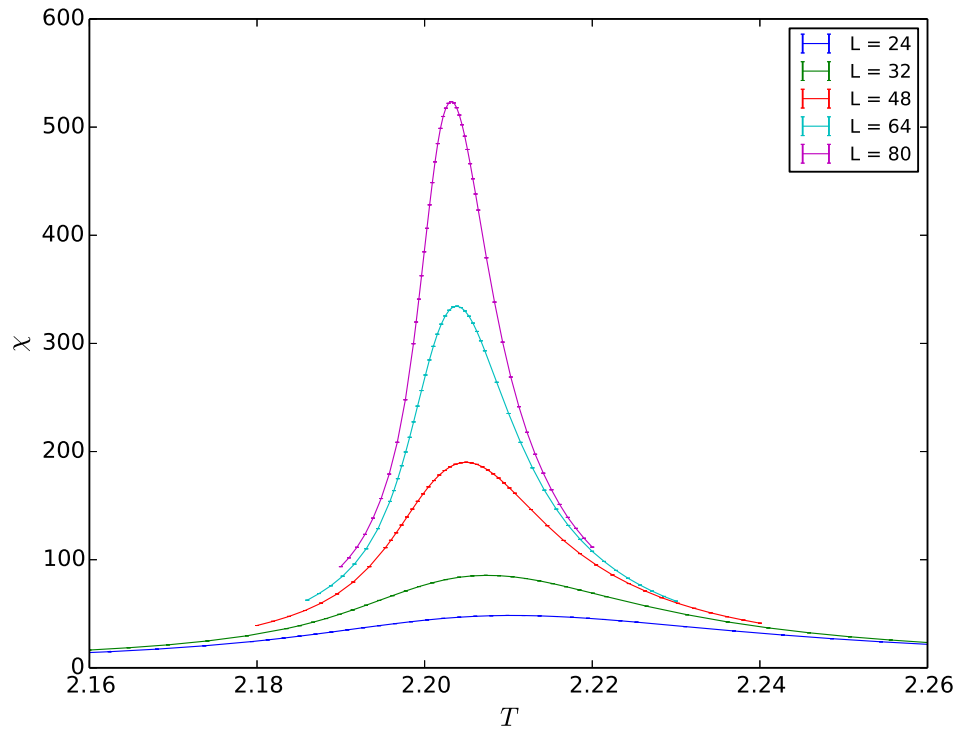


Figure 4.4: Susceptibility of the XY model.

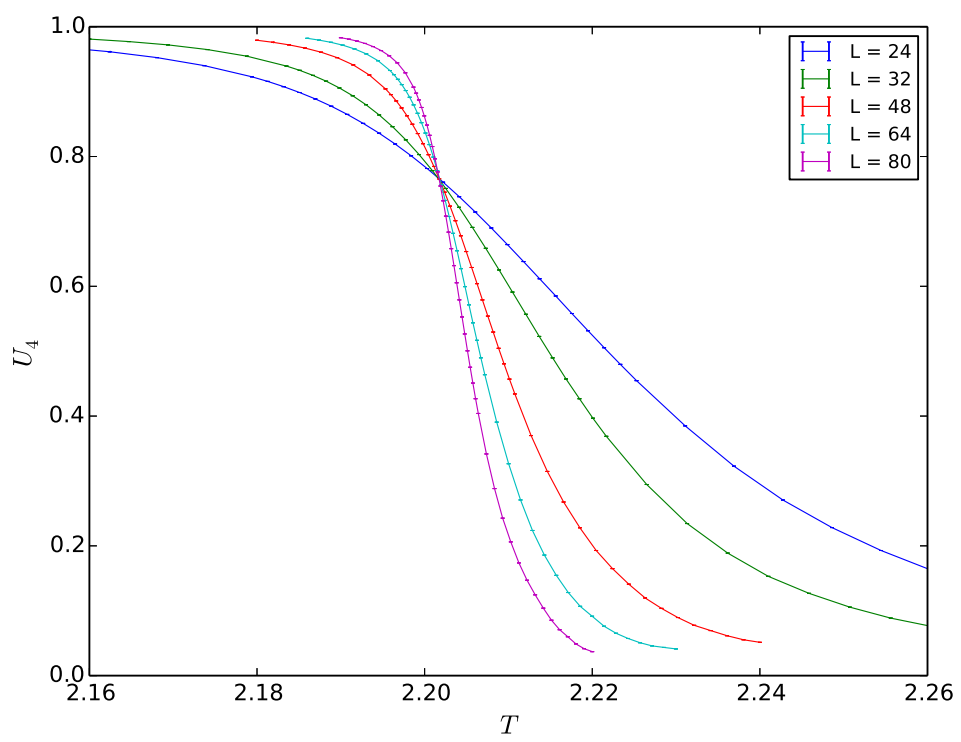
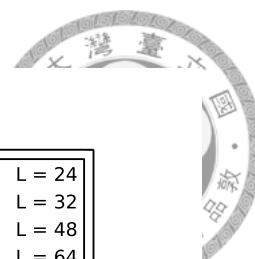


Figure 4.5: Binder cumulant of the XY model.

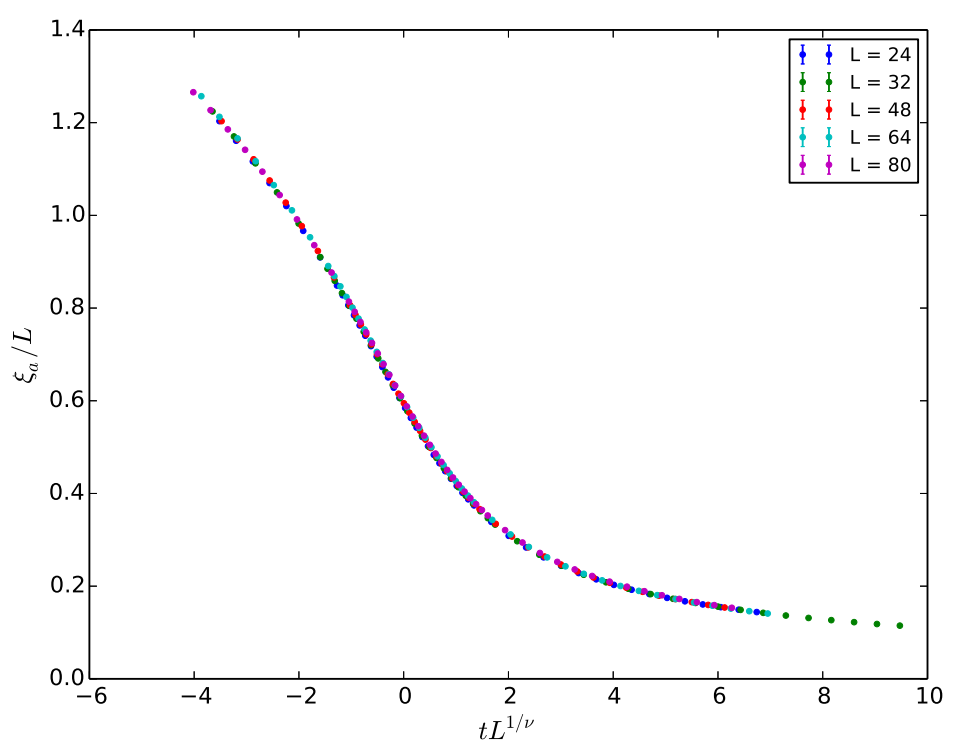


Figure 4.6: Finite size scaling plot of correlation length of the XY model.



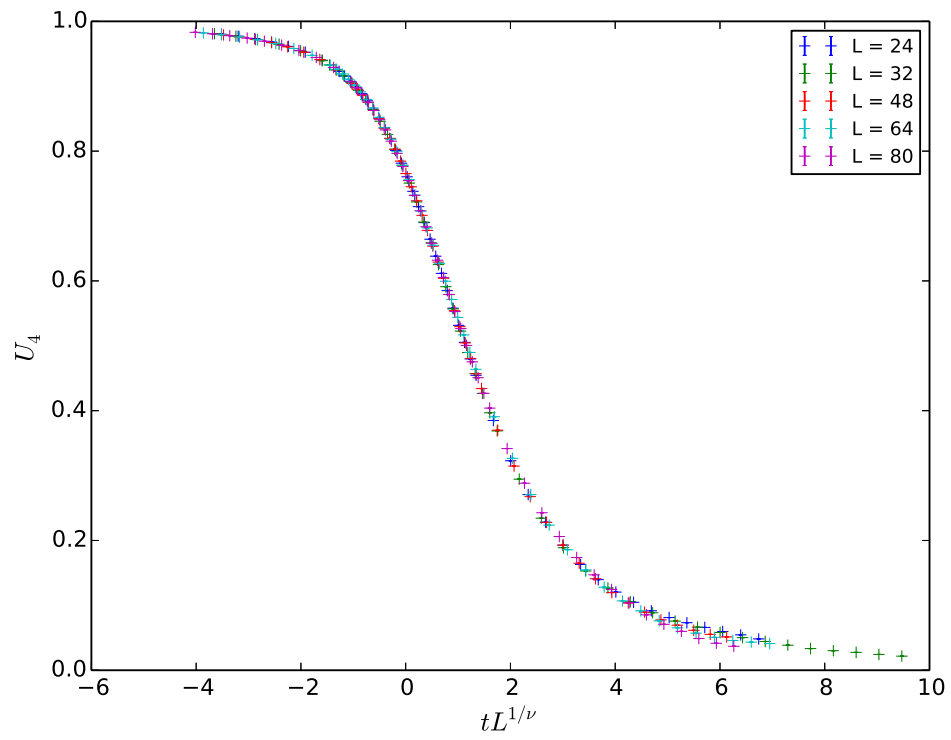


Figure 4.7: Finite size scaling plot of Binder cumulant of the XY model.

## 4.2 3D XY Model with $Z_q$ anisotropy and emergent U(1) symmetry



In this section, we follow Ref. [22] to discuss the XY model with anisotropy as a preparation for the next chapter.

Theoretic studies on the anisotropic XY model, such as the renormalization group theory, implies that the anisotropy is dangerously irrelevant for  $q \geq 4$ [23–26]; i.e. the anisotropic fields are irrelevant and the universality class maintain the same as the isotropic XY universality class. And it is also consistent with the studies on the  $q$ -state clock models, for which the anisotropy is dangerously irrelevant for  $q \geq 5$ . The clock models with the same Hamiltonian with the isotropic XY model, but the angle of a spin  $\theta$  is constraint at angles  $\theta_n = \frac{2\pi n}{q}$ , where  $n = 1 \dots q$ ; the clock model can be regarded as a case of the XY model with infinite anisotropy. And, the numerical results also support these predictions[27, 28].

In Ref. 22, while the anisotropy is irrelevant at  $T_c$ , for  $T < T_c$ , the perturbation becomes relevant with system size larger than a length scale  $\Lambda \sim \xi^a$ ,  $a > 1$ . It implies that the length scale  $\Lambda$  diverges faster than the correlation length as  $T$  close to  $T_c$ , and the anisotropy becomes irrelevant at  $T_c$ . However, as  $T$  being further from  $T_c$ ,  $\Lambda$  decrease and the anisotropy become relevant. To extract  $\Lambda$ , we observe the distribution of the order parameter  $P(m_x, m_y)$ , where  $m_x$  and  $m_y$  are defined in Eq. (1.23). We can also define an order parameter measuring the angular distribution and corresponding to the anisotropy

$$\langle m_q \rangle = \int_0^1 dr \int_0^{2\pi} d\theta r^2 P(r, \theta) \cos(q, \theta) \quad (4.2)$$

$$= \langle m \cos(q\theta) \rangle, \quad (4.3)$$

where  $\theta$  is the angle of the direction of  $(m_x, m_y)$ . Shown in Fig. 4.8, we can find that for  $L > \Lambda$ , the angular distribution occurs and in the case  $q = 8$ , the length scale  $8 < \Lambda < 32$ . However, in the  $Z_4$  case, the angular distributions can be found even for  $L = 8$  while close to  $T_c$ .

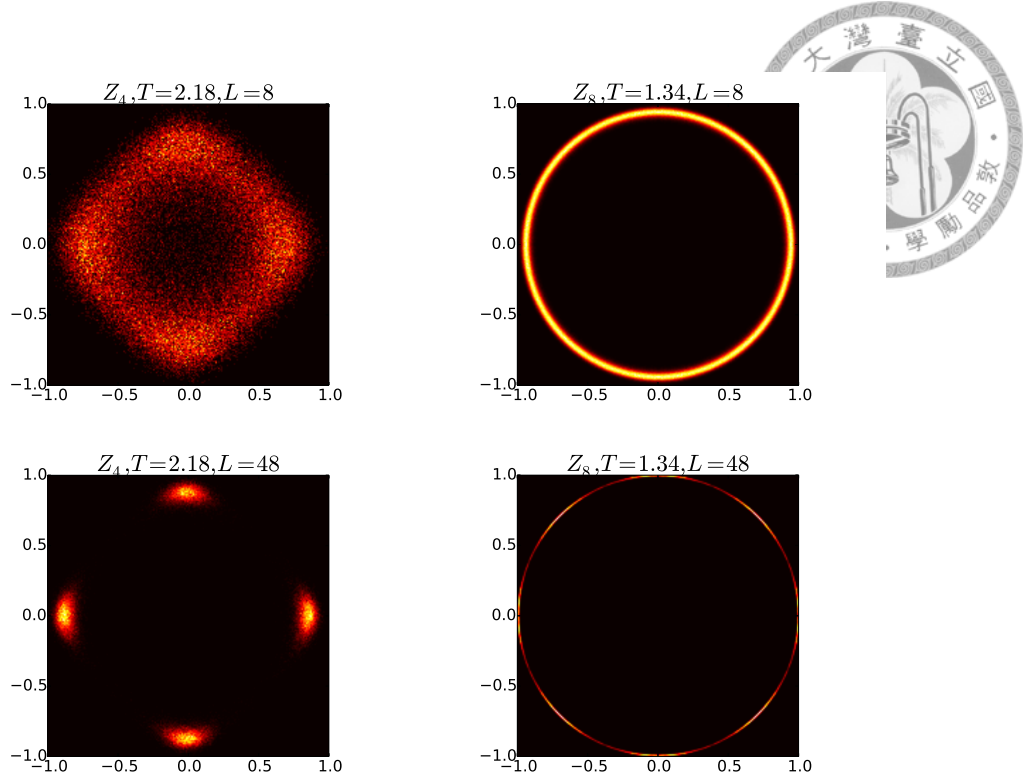


Figure 4.8:  $P(m_x, m_y)$  of  $Z_4$  and  $Z_8$  systems. For  $Z_4$  anisotropy, temperature is near  $T_c$ , we there is angular distribution even for  $L = 8$ . For  $Z_8$  anisotropy, at  $T = 1.34$ , which is much lower than  $T_c$ , the length scale is larger than 8 and smaller than 48.

We can explain the result for  $Z_q$  anisotropy by coarse graining. Considering a single spin, the configuration must have  $q$ -preferred directions. As performing coarse graining, after combining  $l^3$  spins into a block spin, the effect of  $q$ -preferred directions is diminishes, and the angular distribution become more flat. Within  $\Lambda$ , the anisotropy diminished with increasing size. As shown in Fig. 4.9 and Fig. 4.10, there are crossing points and  $m_q$  decreases with size near  $T_c$ . However, for  $L \gtrsim \Lambda$ , the angular distribution become pronounced. We can write the finite-size scaling form of  $m_q$  as

$$\langle m_q \rangle = L^{-\sigma} g(tL^{1/\nu_q}), \quad (4.4)$$

which is generalized from Eq. (3.12) with  $\sigma = \beta/\nu$ .

In Ref. 22, they analyze the scaling of the order parameter distributions, and finally get the relation that  $\nu_q = \Delta_q \nu / 3 = a_q \nu$ , where  $a_q > 1$  is the exponent in the relation  $\Lambda = \xi^{a_q}$  and  $\Delta_q > 3$  is the scaling dimension of the irrelevant anisotropy. For  $a_q > 1$ ,  $\nu_q$

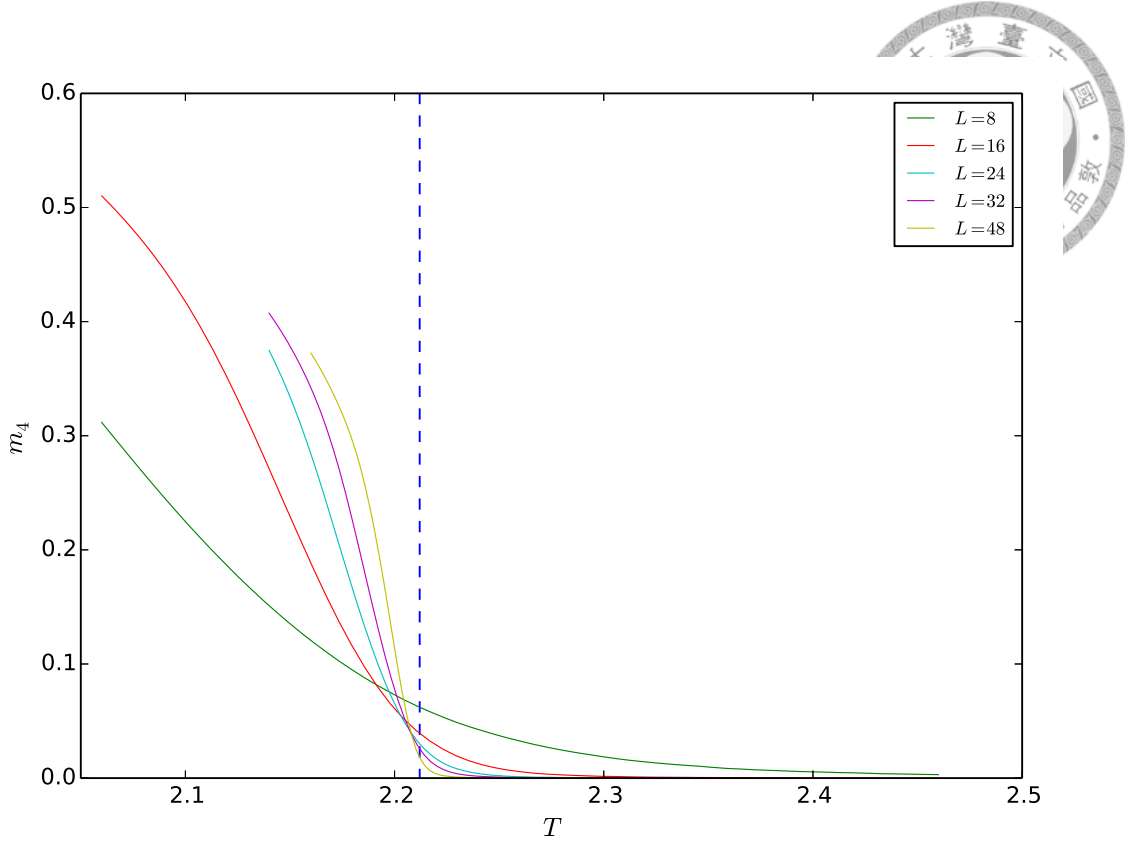


Figure 4.9:  $m_q$  order parameter for  $q = 4$  and  $h = 2.0$

must be larger than  $\nu$ . The function form of  $g(X)$  is given by

$$g(X) \propto \frac{I_1(\tilde{h}X^{3\nu_q})}{I_0(\tilde{h}X^{3\nu_q})}, \quad (4.5)$$

where  $\tilde{h}$  is a nonuniversal scale factor and  $I_n$  represents the modified Bessel function of order  $n$ [22, 25].

With the finite size scaling form and the relations, we can extract the length scale  $\Lambda$  by analyzing the behavior of  $m_q$ . Fig. 4.11 shows  $\langle m \rangle$  and  $\langle m_q \rangle$  for  $q = 5, 6$  and Fig. 4.12 shows the scaling of  $\langle m_4 \rangle$ ,  $\langle m_5 \rangle$  and  $\langle m_6 \rangle$ . We obtain  $a_4 = 1.06(4)$ ,  $a_5 = 1.6(1)$ ,  $a_6 = 2.4(1)$ ,  $a_6 = 2.4(1)$ ,  $a_7 = 3.0(1)$ ,  $a_8 = 4.4(3)$ , which are consistent with the results of Lou *et al.* and the form  $a_q = a_4(q/4)^2$ .

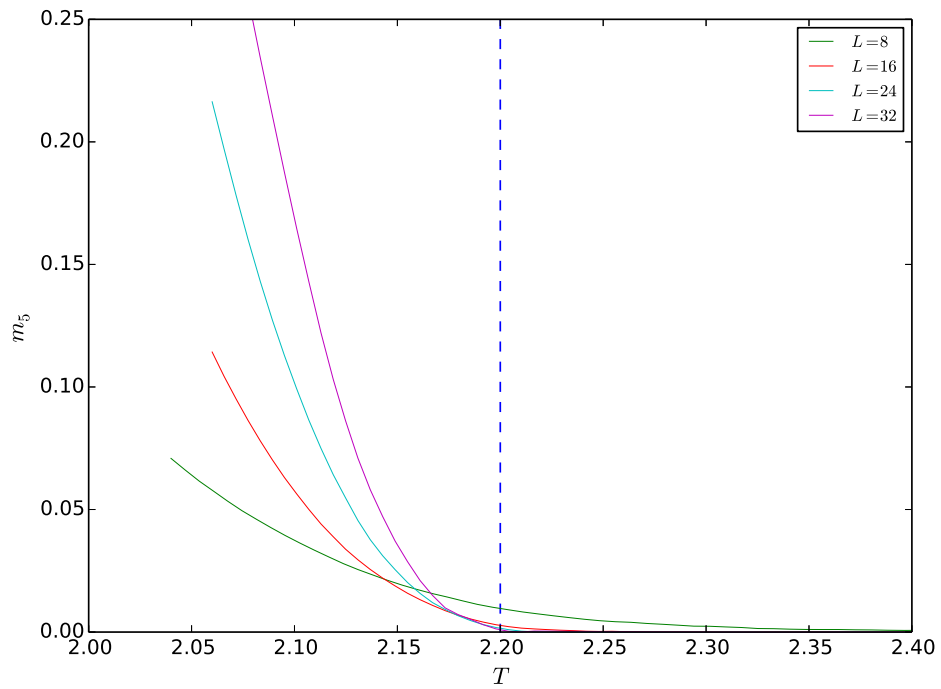


Figure 4.10:  $m_5$  order parameter for  $Z_5$  clock model

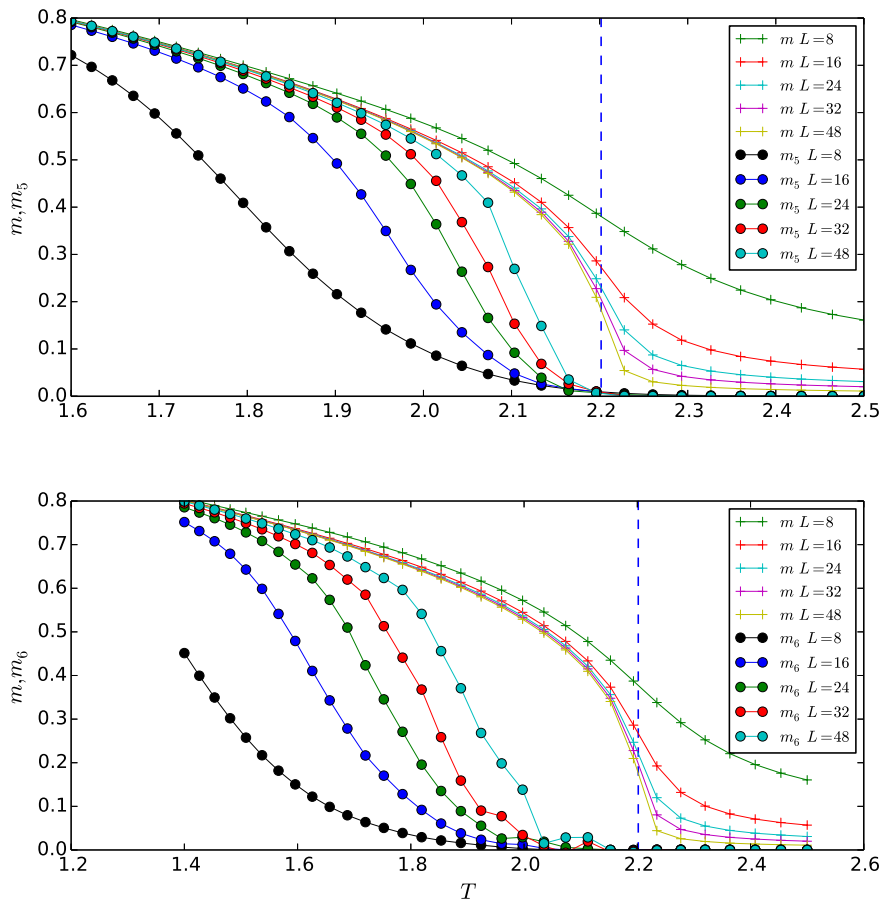


Figure 4.11:  $m_q$  and  $m$  versus  $T$  for  $q = 5, 6$ .

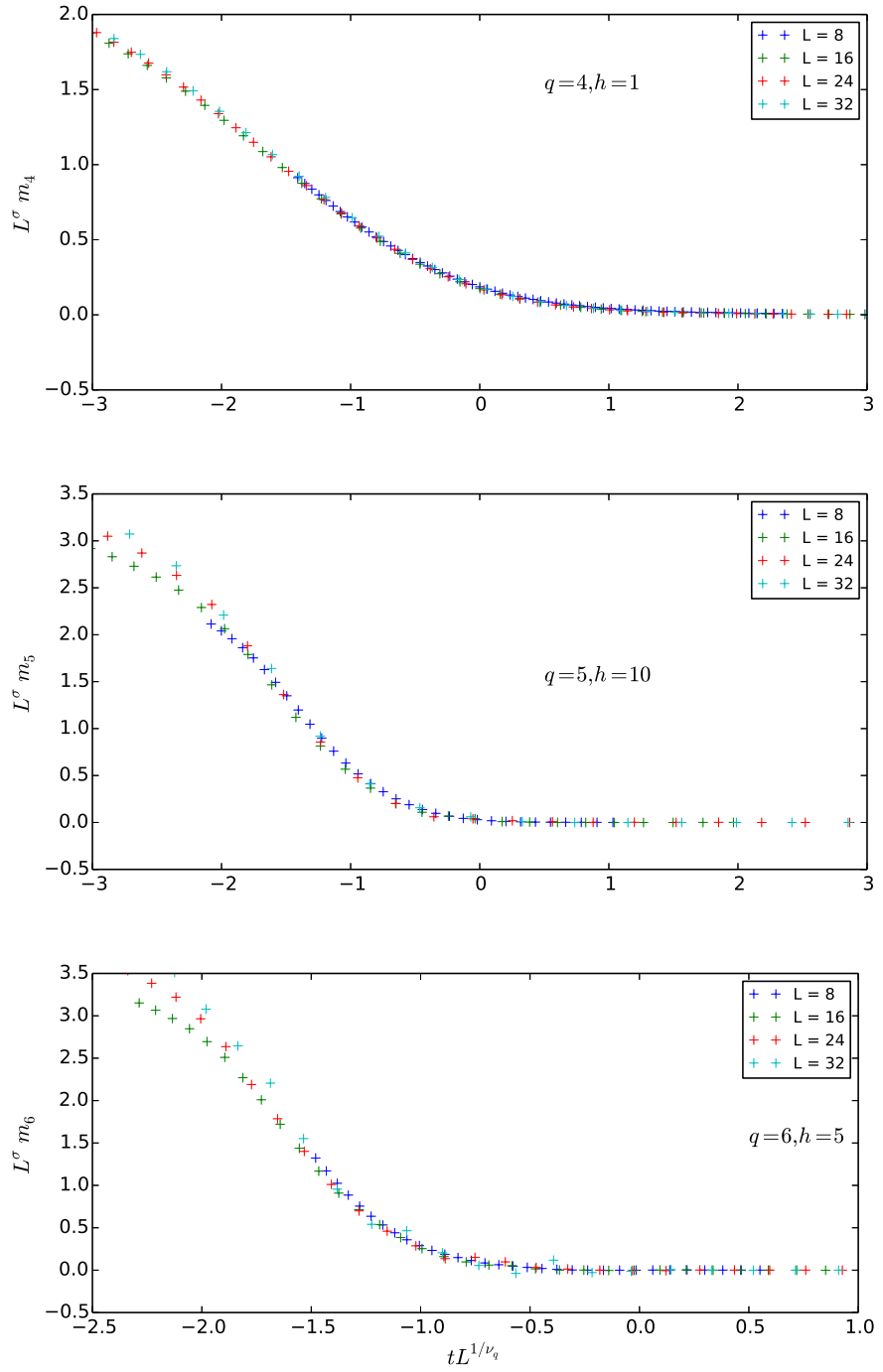
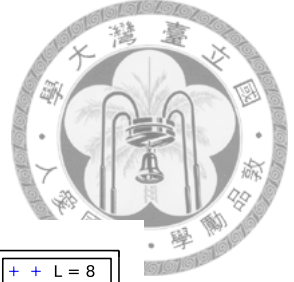


Figure 4.12: Scaling  $m_q$  for  $q = 4, 5, 6$ , where  $\nu_4 = 0.709, \nu_5 = 1.08, \nu_6 = 1.59$ .



## Chapter 5

# Nonlinear Three-dimensional XY

## Model with $Z_q$ Anisotropy

In this chapter, we generalize the simple XY model into the nonlinear XY model with different nearest neighbor interaction

$$H = J \sum_{\langle i,j \rangle} V(\theta_i - \theta_j) - h \sum_i \cos(q\theta_i), \quad (5.1)$$

where potential function  $V(\theta)$  is written as

$$V(\theta) = -2 \left( \frac{1 + \cos(\theta)}{2} \right)^P. \quad (5.2)$$

This model is first studied in two dimensions by Domany *et al.* [29]. Shown in Fig. 5.1, the potential function gets narrow as  $P$  increase. This implies that the model is getting close to the Potts model with infinitely many states, which has a first-order transition. Therefore, in Ref. 29, they try to make the transition change from the Kosterlitz-Thouless transition to a first-order phase transition, and they indeed found first-order transitions for  $P > 50$ .

We report our results of the critical behaviors of the 3D nonlinear XY models with and without anisotropy term in this chapter.



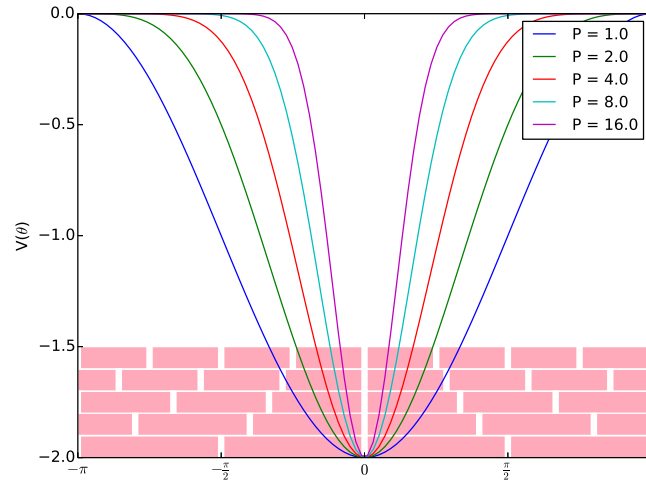


Figure 5.1:  $V(\theta)$  vs.  $\theta$  and the pink blocks indicates the angles of  $Z_4, Z_5, Z_6,$  and  $Z_8$

## 5.1 Nonlinear 3D XY model

First, we study the nonlinear XY model without anisotropy and examine that whether there exist first-order transitions. There are theoretical proofs and discussions predicting the existence of the first-order phase transition as  $P$  is large enough[30, 31]. For  $P = 8$ , the negative peaks of binder cumulant indicate the transition is first-order, shown in Fig. 5.2. Though the energy histograms do not show clear double peaks at any temperature in Fig. 5.3, there is a clear energy shift between the temperatures near the critical temperature. From these indicators, the transition should be first-order, and this conclusion is consistent with the previous micro-canonical Monte Carlo study[32]. We also show the energy and order parameter in Fig. 5.5 and 5.4. The result of micro-canonical Monte Carlo study[32] also shown in Fig. 5.6, the van der Waals-like loop is discovered. The Binder cumulant for  $P = 7.5$ , shown in Fig. 5.7, reveals that there is a continuous transition, and for  $P > 8$ , for the width of potential  $V(\theta)$  getting smaller we predict there must be a first-order transition. Therefore, there exists a  $7.5 < P_c < 8.0$  and the transition turns to first-order for  $P > P_c$ .

For  $1 < P < P_c$ , there are continuous transitions. With the finite-size scaling analysis, we estimate the correlation length exponent  $\nu$  by the maximum of the derivatives of Binder

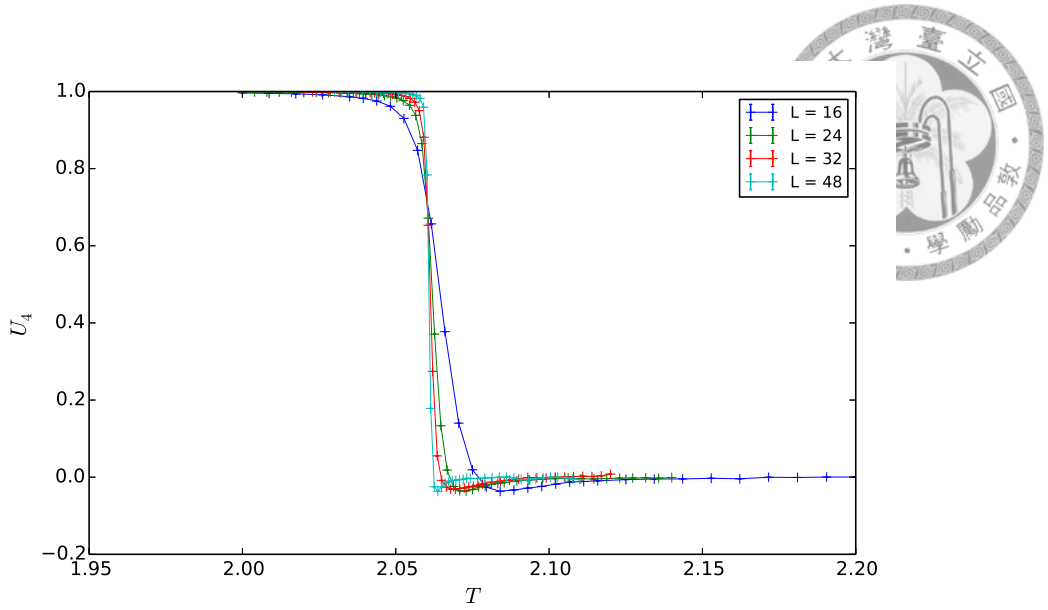


Figure 5.2: Binder cumulants for  $P = 8$

cumulants and logarithmic derivative of  $\langle m^4 \rangle$ . Since our simulation size is just up to  $L = 48$  and the systematic errors are very large, it is hard to judge whether  $\nu$  varies with  $P$ . However, we also measure  $\gamma$  by the maximum of susceptibility  $\chi$ , and it reveals a trend that  $\gamma$  decrease with increasing  $P$ . Our results are shown in Table. 5.1

Table 5.1: Critical exponent  $\nu$  and  $\gamma$  versus  $P$

$P$	1.0	2.0	3.0	4.0	5.0	6.0	7.0	7.5
$\nu$	0.6709(1)	0.665(2)	0.674(5)	0.660(4)	0.659(2)	0.668(11)	0.658(12)	0.59(4)
$\gamma$	1.256(5)	1.277(30)	1.214(20)	1.215(8)	1.160(4)	1.110(70)	1.081(78)	

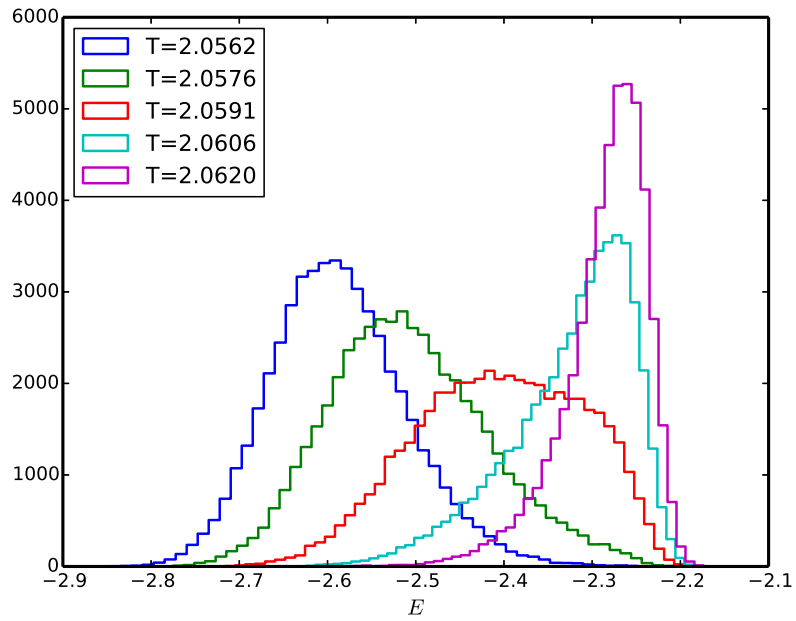
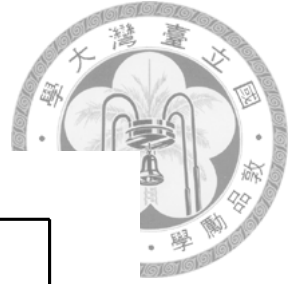


Figure 5.3: Energy histogram for  $P = 8, L = 32$

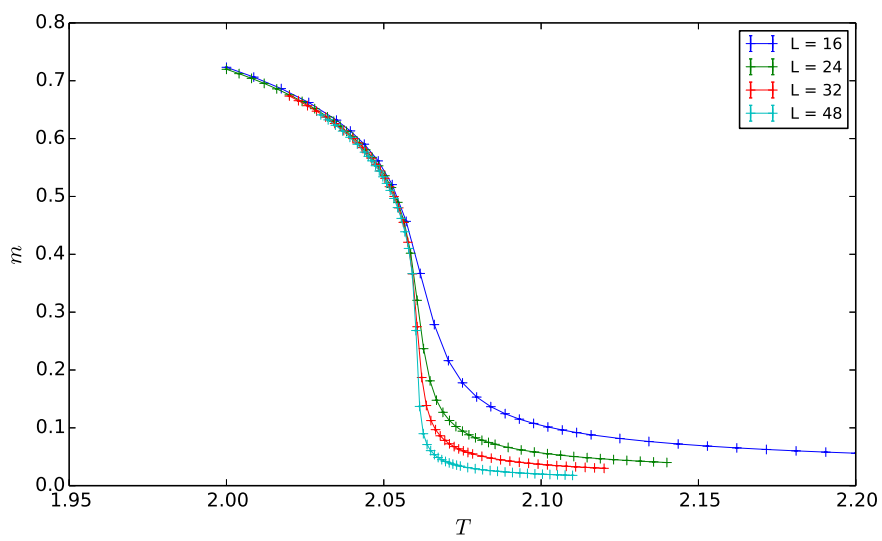


Figure 5.4: Order parameter for  $P = 8$

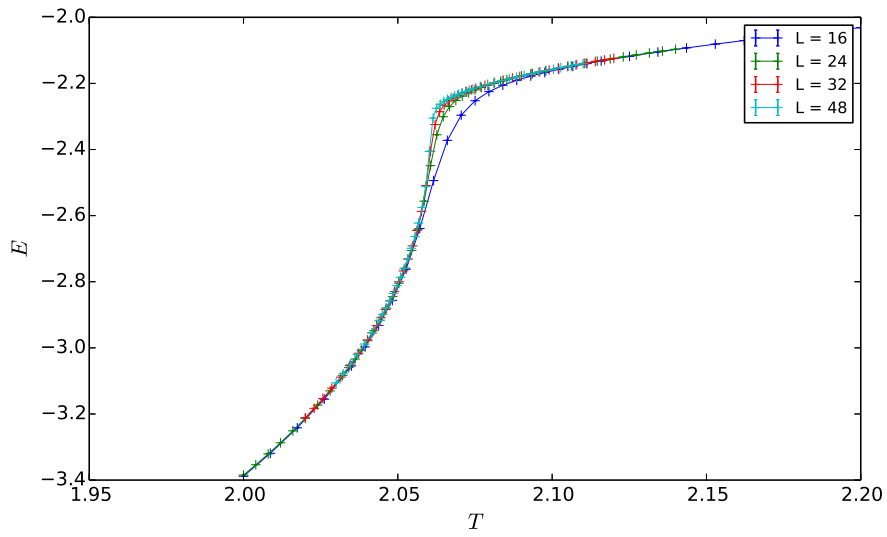
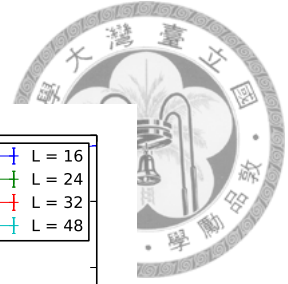


Figure 5.5: Internal energy for  $P = 8$

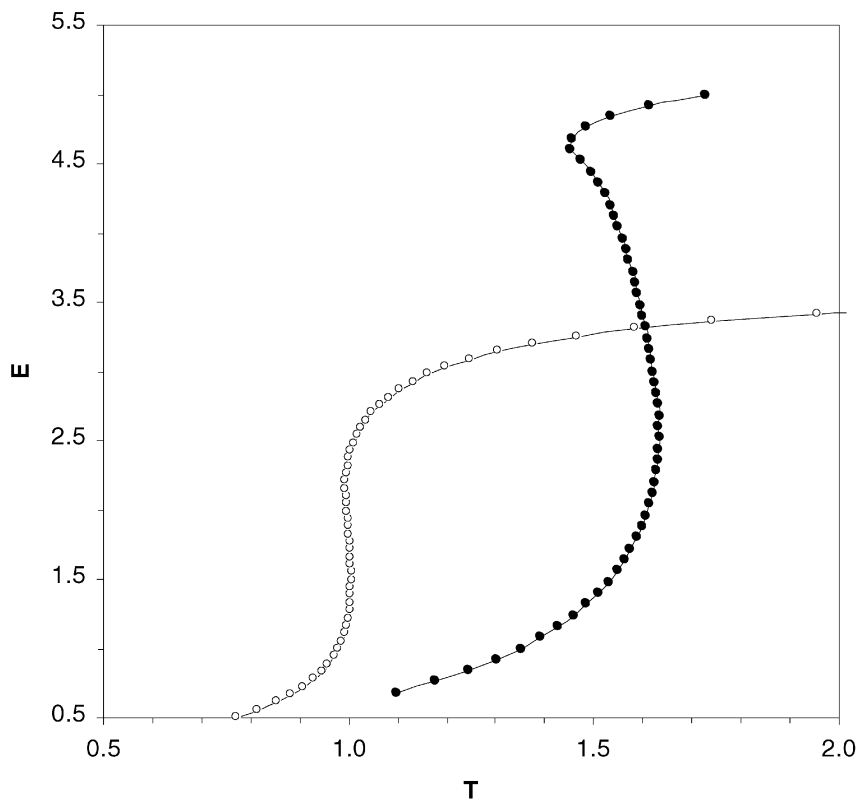


Figure 5.6: Micro-canonical Monte Carlo simulation result for  $P = 55$  2D system( $\circ$ ) and 3D system( $\bullet$ ). There is a van der Waals-like loop implying that the transition is first order.[32]

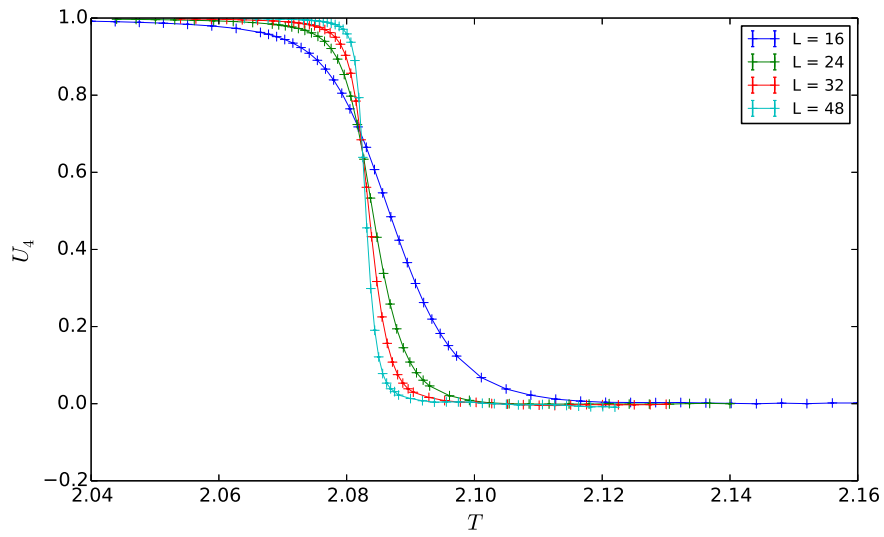


Figure 5.7: Binder cumulants for  $P = 7.5$

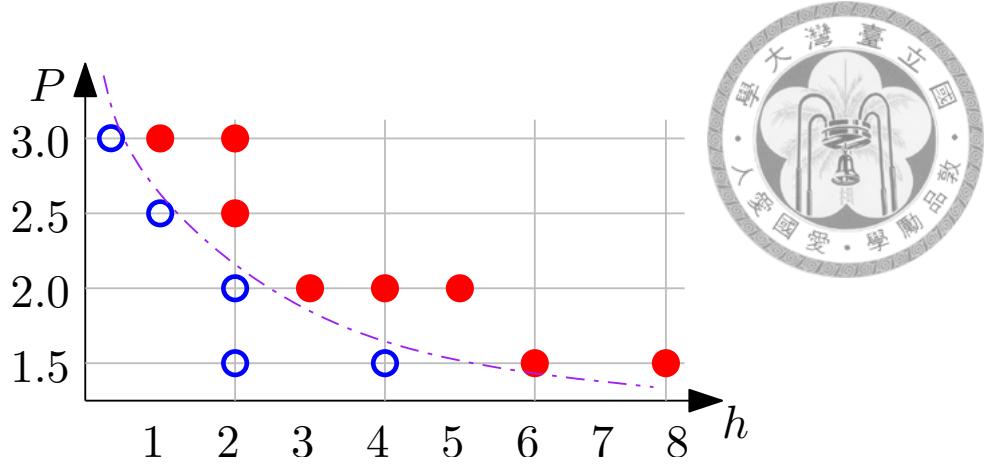


Figure 5.8: First-order and continuous phase transitions for  $q = 4$  with different  $P$  and  $h$ , the first-order phase transitions are labeled by red dashes while the continuous phase transitions are by blue circles

## 5.2 Nonlinear 3D XY model with $Z_q$ anisotropy

### 5.2.1 From continuous to first-order

After considering the anisotropy, the situation becomes much more complicated. Unlike the simple XY model with anisotropy, the anisotropy term may mightily influence the phase transition and even change the order of the transitions. Since there are many parameters in the Hamiltonian, to explain the results clearly, we start from the  $Z_4$  case. For  $q = 4$ , as  $h$  is large than some value  $h_c$ , the transition become first-order for  $P \geq 1.5$ . Shown in Fig. 5.8,  $h_c$  decrease as  $P$  increase, and at  $P = 1$ , the transition of simple XY model with  $Z_4$  anisotropy is absolutely a continuous phase transition. The relation of  $h_c$  and  $P$  is also found for  $q > 4$ . And, we claim that for  $q = 4$  and  $P$  slightly larger than 1, the first-order transitions exist for  $h > h_c$ .

As an example of the first-order phase transitions, the energy histogram and the magnetization distribution for  $P = 2.5$ ,  $h = 2.0$  and  $L = 48$  are shown in Fig. 5.9 and Fig. 5.10. From the magnetization distribution, we find that there is no emergent U(1) symmetry. The double-peaks of energy and order parameter are evidence for the first-order transition, and, of course, negative peaks of the Binder cumulant are found.

For  $q > 4$ , we perform some simulations for  $q = 5, 6, 7$ . We find that for each  $q > 4$ ,

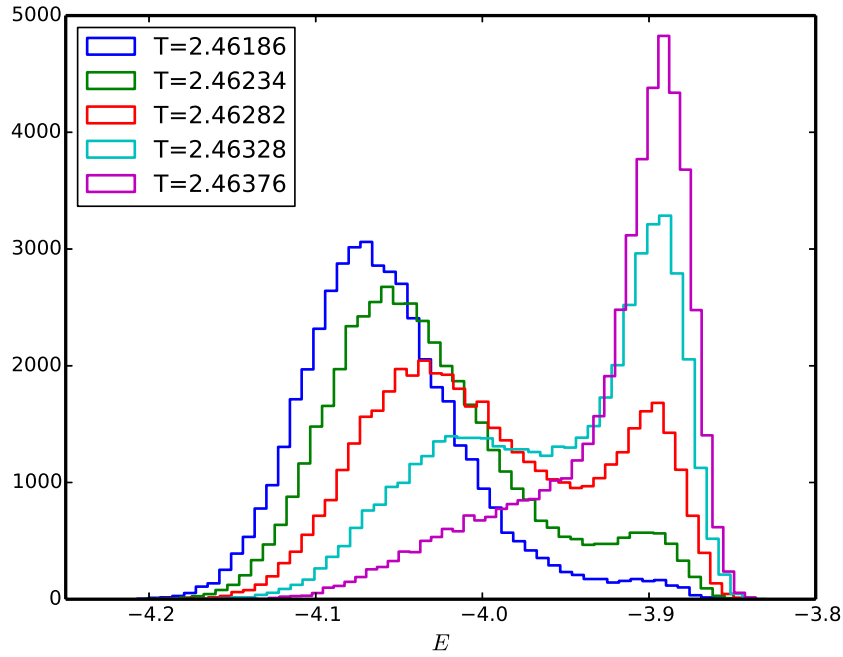
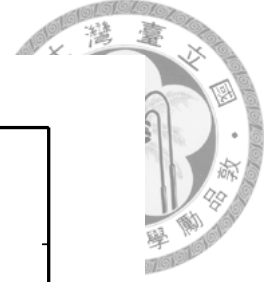


Figure 5.9: The energy histogram for  $P = 2.5$ ,  $q = 4$  and  $h = 2.0$ . The double-peak is very clear at  $T = 2.46282$ .

there exists a region of  $P$  where the transitions maintain continuous for any  $h$ , and when  $P$  is larger than a value  $P_q$  for each  $q$ , first-order transitions can be found for  $h$  large enough. The results are shown in Fig. 5.11. As an example, for  $q = 5$  and  $P = 2.0$ , the transition is always continuous. We also check this diagram with the  $Z_q$  clock model with the potential in Eq. (5.2) as an extreme case that the strength of anisotropy goes to infinity. We find that the exponent  $P_q$  increase with  $q$ , and for  $q \rightarrow \infty$ , i.e. the case with no anisotropy,  $P_\infty = P_c$  is near 8.0.

This result is reasonable under the framework of  $q$ -state clock model. Observing Fig. 5.1, we estimate that the width of potential  $V(\theta)$  is proportional to  $\sqrt{P}$ . As  $P$  increasing, the width decreases and the internal energy between two spins with the smallest nonzero angle  $V(2\pi/q)$  goes to zero. Therefore, as  $P$  increases, the model becomes more similar to the  $q$ -state Potts model; for smaller  $q$ ,  $V(2\pi/q)$  decreases faster with  $P$ , so  $P_q$  is smaller.

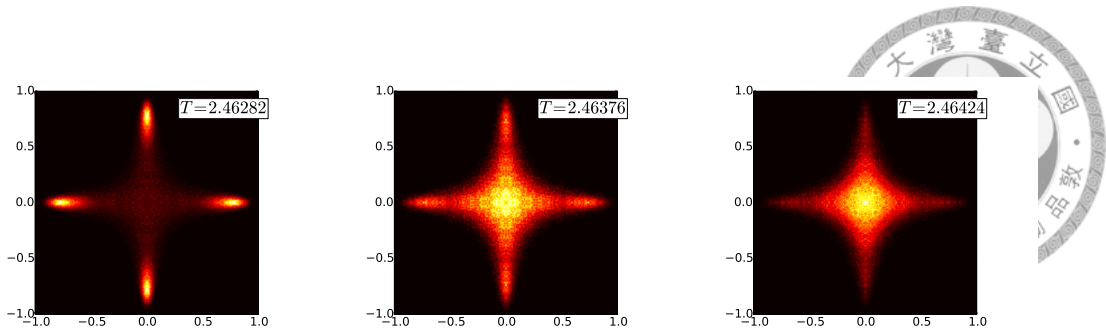


Figure 5.10: The The order parameter distribution for  $P = 2.5$  and  $h = 2.0$ . There are symmetric peaks and a central peak. They coexist at  $T = 2.46376$

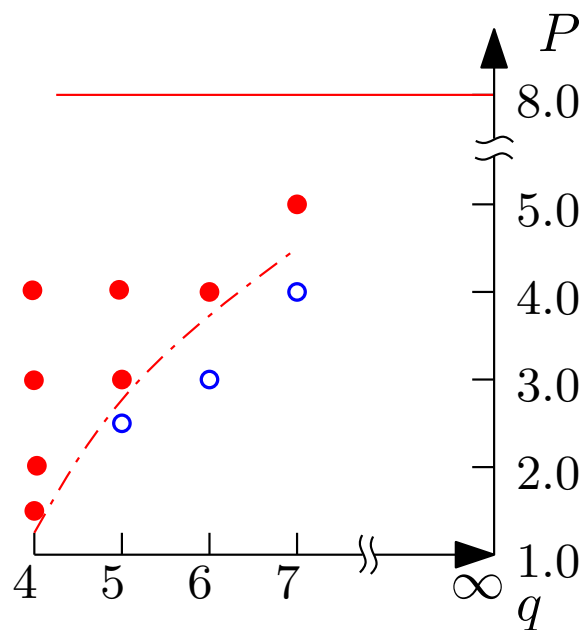


Figure 5.11: Existence of first-order phase transition for different  $P$  and  $q$ . The red dashes means that the transitions turn into first-order transitions for  $h$  large enough while the blue circles means that there is no first-order transition no matter what value  $h$  is.

## 5.2.2 From irrelevance to relevance

In this and next subsection, we show some quantitative results of critical exponent and discuss about the anisotropy relevance and irrelevance. In the simple XY model, for  $q \geq 4$ , the anisotropy is dangerous irrelevant and the critical exponents do not change. The exponent  $\nu_q$  must be larger than  $\nu_{XY} = 0.671$ , ensuring that the anisotropy is irrelevant at critical temperature.

However, in the nonlinear XY model,  $\nu_q$  changes with  $h$  and it may be smaller than  $\nu$  for  $h = 0$ . In addition, the critical exponents may vary as  $h$  varies. Caused by the



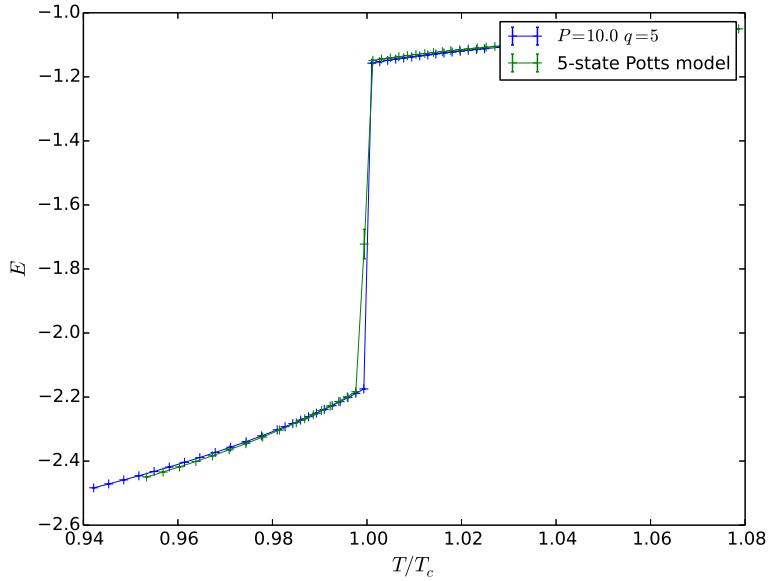


Figure 5.12: Energy versus the reduced temperature  $T/T_c$  for 5-state Potts model and nonlinear XY model for  $P = 10.0$  and  $h = 40.0$ . For the nonlinear XY model, the energy is only the energy of interactions between spins. We can see that the behaviors are similar and the latent heat is almost the same.

systematic errors, there exists some uncertainty in our results. Nevertheless, for  $q = 4$ , the changes of  $\nu_q$  and the critical exponents are obvious. In addition, since there is no need to consider the region  $1 < P < P_q$  for  $q = 4$ , the behaviors are easier to understand. Therefore, we start from the results for  $q = 4$ , again.

By Eq. (4.4), we extract  $\nu_4$  with data collapse using the program by Kenji Harada[19]. For  $P = 2.0$ , we find that the  $\nu_q(h)$  depends on  $h$ ,  $\nu_4(0.1) = 0.68(2)$ ,  $\nu_4(0.5) = 0.63(1)$ ,  $\nu_4(0.7) = 0.62(1)$ ,  $\nu_4(1.0) = 0.56(1)$ , and the correlation length exponent  $\nu$  is also decreasing from 0.671 to 0.56(2), shown in Fig. 5.13. Unlike in simple XY model,  $\nu_q(h)$  and  $\nu$  vary with  $h$ , and for  $h \geq 1.0$ ,  $\nu_4$  is smaller than  $\nu$  of isotropic nonlinear XY model with  $P = 2.0$ . For the region where  $\nu_4 < \nu_{P=2.0}$ , by the relation of length scale  $\Lambda$  and correlation length in chapter 4, we can inference that the anisotropy is relevant and there is no emergent U(1) symmetry, and the varying  $\nu$  also give us the same conclusion.

For  $P = 1.5$  and  $P = 3.0$ , the behaviors are the same. Therefore, we claim that for  $P \geq 1.5$ , for  $h$  large enough the anisotropy is relevant. There exists a non-universal

region, where the critical exponents vary with the symmetry-breaking field  $h$ , before the transition become first-order. We also find that before  $h$  is larger enough for the transition to be a first-order transition,  $\nu_4(h)$  should be smaller than  $\nu$  for  $h = 0$  first.

There exist two open questions here. One is how to describe this system with  $\nu_q$  varying with  $h$ ; the other is whether the exponents are universal for  $\nu_4 > \nu_{h=0}$ , i.e. by the theoretical description of length scale  $\Lambda$ , the anisotropy should be irrelevant for  $\nu_4 > \nu_{h=0}$  and the exponents should be the same. Since the systematic errors are not small, the region of the non-universal behavior is not clear enough to confirm the relation between the non-universal behavior and  $\nu_q$ .

For  $q > 4$ , we study  $q = 5, 6$ . For  $P > P_q$ , the properties of transitions and  $\nu_q$  are nearly the same as the case for  $q = 4$ . However, shown in Fig. 5.15, due to the systematic error, the region and existence of non-universal behavior are uncertain. It is hard to judge whether the difference of  $\nu$  is caused by errors and the finite-size correlation terms.

For  $P < P_q$ , shown in Fig. 5.14,  $\nu_q(h)$  is also decreasing as  $h$  increasing, but for any  $h$ , even for the clock model, the extreme case,  $\nu_q(h)$  is large than  $\nu_{h=0}$ . The existence or absence of non-universal behavior is also hard to be confirmed.

For  $P < P_q$ , since  $\nu_q(h)$  is always larger than  $\nu_{h=0}$ , anisotropy is expected to be irrelevant. However, the critical exponents change slightly as  $h$  increases, and we are not sure whether there exists the non-universal behavior.

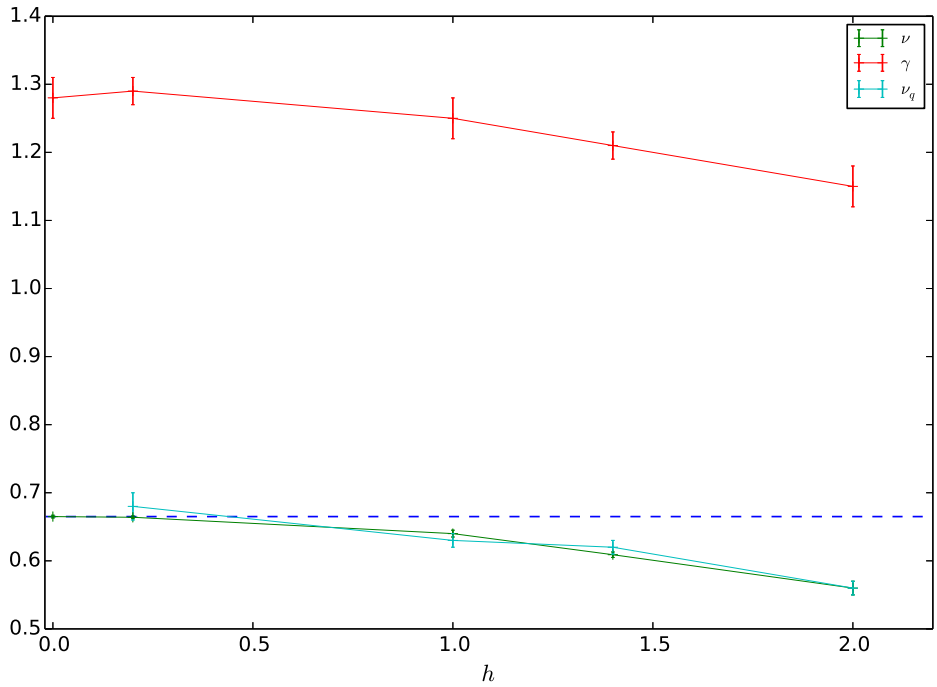
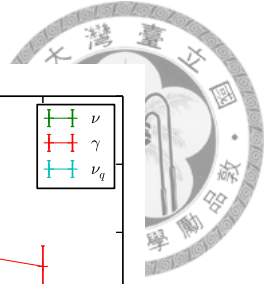


Figure 5.13: Critical exponent versus  $h$  for  $P = 2$  and  $q = 4$ .  $\nu_4$  decreases with  $h$ . The non-universal behaviors are obvious, and  $\nu$  is near to  $\nu_4$  for  $\nu_4 < \nu_{h=0}$ .

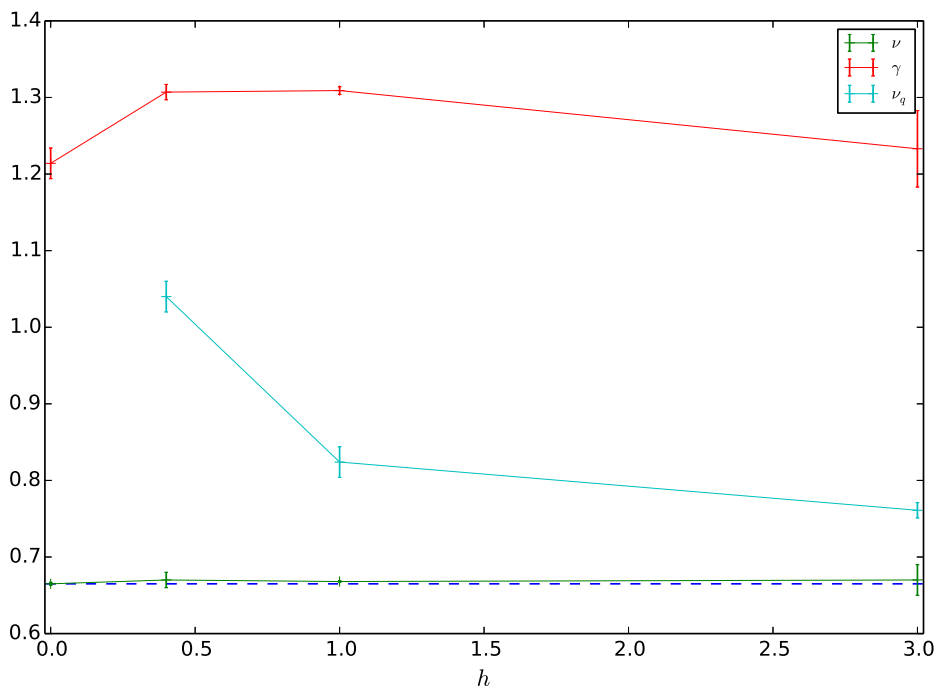


Figure 5.14: Critical exponents versus  $h$  for  $P = 2$  and  $q = 5$ .  $\nu_5$  decreases with  $h$  but always larger than  $\nu_{h=0}$ . The change of  $\gamma$  may be caused by the finite-size effect and errors. We also measure the exponents of the 5-state clock model,  $\nu = 0.668(5)$ ,  $\nu_q = 0.76(3)$  and  $\gamma = 1.26(1)$ . The extreme case is consistent with the finite- $h$  case.

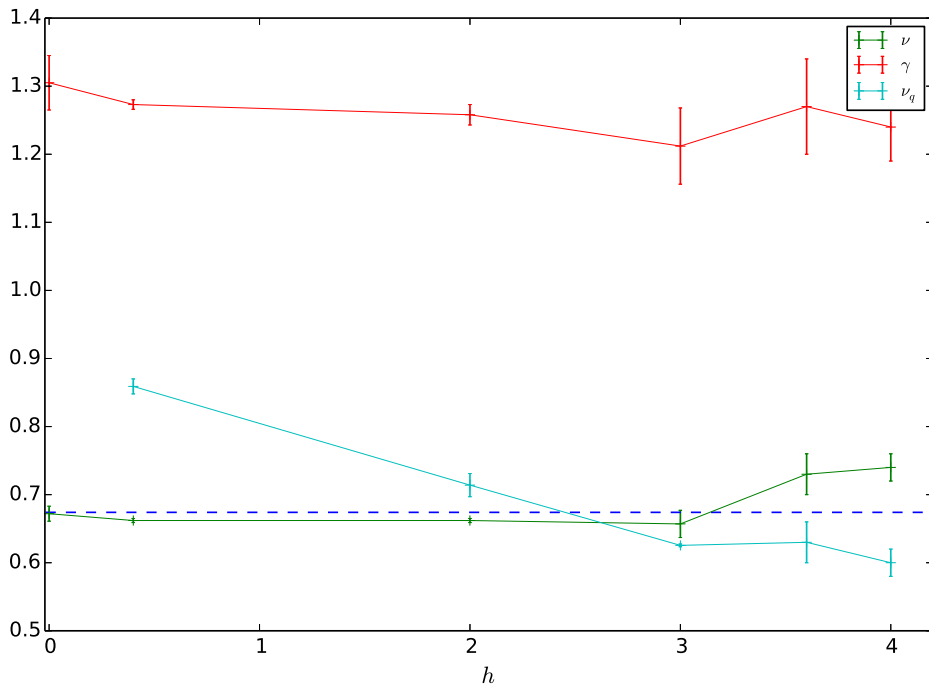
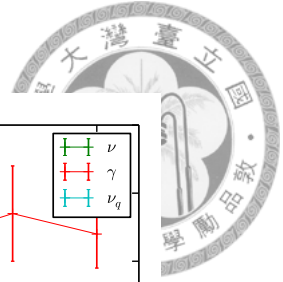


Figure 5.15: Critical exponents versus  $h$  for  $P = 3$  and  $q = 5$ .  $\nu_5$  is decreasing. However, the changes of  $\nu$  and  $\gamma$  are not clear. The non-universal behavior can not be confirmed.

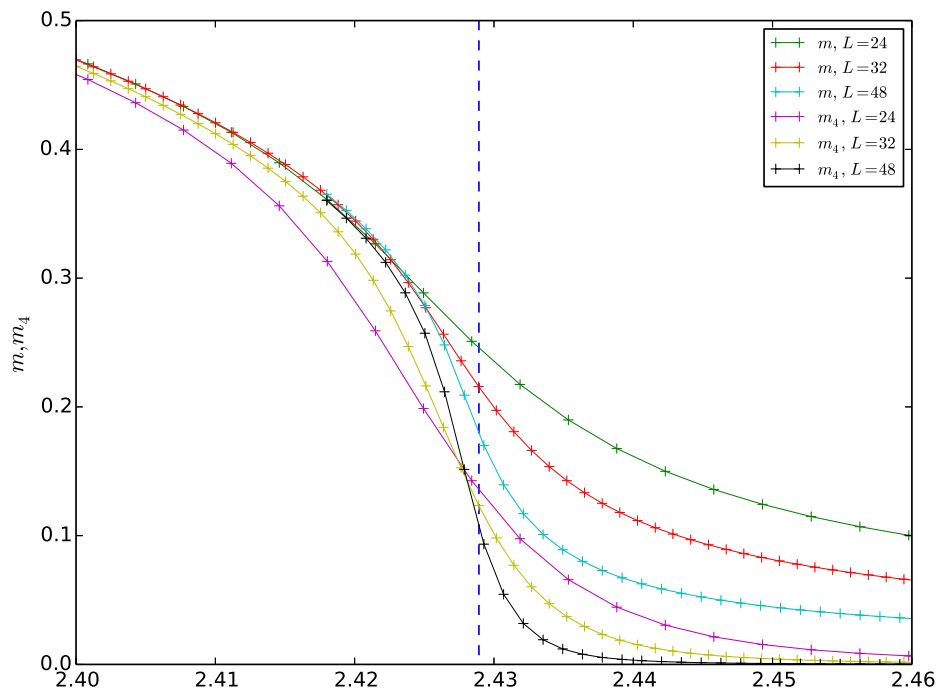


Figure 5.16:  $\langle m \rangle$  and  $\langle m_4 \rangle$  versus  $T$  for  $P = 2.0$ ,  $h = 2.0$  and  $q = 4$ .

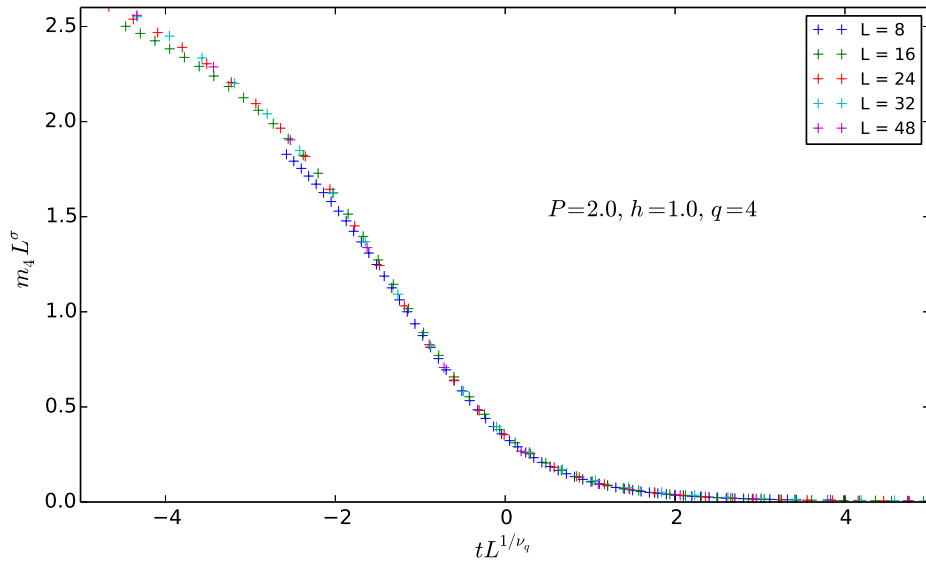


Figure 5.17: Scaling plot of  $\langle m_4 \rangle$  for  $P = 2.0, h = 1.0$  and  $q = 4$ .



# Chapter 6

## Summary

To study phase transitions and critical behaviors, Monte Carlo methods and finite-size scaling are necessary. Caused by the irrelevant field of finite size, there is some correction depending on size. To get precise results, large-size simulations should be performed.

To simulate large size system in less time, we implement the Monte Carlo method on GPUs. According to the checkerboard algorithm, We can parallelize the single-spin flips and measurements into threads on hundreds of GPU cores. It does reduce the time for Monte Carlo simulations.

We study the simple XY model. With no anisotropy, we obtain  $\nu = 0.6708(1)$  and  $\gamma = 1.324(4)$ , consistent with previous results. The emergent U(1) symmetry is a characteristic in the transitions of anisotropic XY model. Following the work of Lou *et al.*, we study the the length scale  $\Lambda$  of U(1) symmetry with the Monte Carlo method on GPU. For the dangerous irrelevance,  $\Lambda$  diverge faster than the correlation length,  $\Lambda \sim \xi^{a_q}$ ,  $a = \nu_q/\nu > 1$ . For the simple XY model with  $Z_q$  anisotropy, we find  $a_4 = 1.06(4)$ ,  $a_5 = 1.6(1)$ ,  $a_6 = 2.4(1)$ ,  $a_6 = 2.4(1)$ ,  $a_7 = 3.0(1)$ ,  $a_8 = 4.4(3)$ , which are consistent with previous works.

For the XY model with nonlinear potential, the transitions are first-order for  $P$  large enough. For  $P > 1$ , with  $Z_q$  anisotropy,  $\nu_q$  and  $a_q$  vary with  $h$ , the strength of anisotropy. And,  $\nu_q$  may be smaller than  $\nu$  for  $h = 0$ . In the case with  $Z_4$  anisotropy, for  $P \geq 1.5$ ,

the transitions turn into first-order transition for  $h$  large enough, and as  $h$  increasing, the exponents become non-universal. For  $q > 4$ , there are two different cases depending on  $P$ . For  $P$  large enough, the critical behaviors and the behavior of  $\Lambda$  are almost the same as the  $Z_4$  case, but the non-universal behaviors can not be confirmed clearly because of errors caused by finite-size effects. In the case that  $P$  is not so large, the transition is always continuous. Though  $\nu_q$  also vary with  $h$ , it is always larger than  $\nu$  for  $h = 0$ . The anisotropy should be irrelevant, but it is not sure that the critical exponents are constant with different  $h$ .

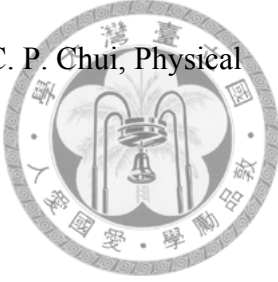
An important question left is the regions of non-universal behavior. To answer this question, simulations on systems with larger sizes should be carried out.

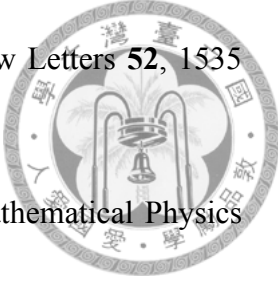


# Bibliography

- [1] K. Huang, *Statistical Mechanics* (Wiley, 1987).
- [2] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, 1996).
- [3] B. Linder, in *Thermodynamics and Introductory Statistical Mechanics* (John Wiley & Sons, Inc., 2004) pp. 119–126.
- [4] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2005).
- [5] K. G. Wilson, *Physical Review B* **4**, 3174 (1971).
- [6] H. G. Katzgraber, arXiv preprint arXiv:0905.1629 (2009).
- [7] L. Onsager, *Physical Review* **65**, 117 (1944).
- [8] A. Pelissetto and E. Vicari, *Physics Reports* **368**, 549 (2002).
- [9] A. W. Sandvik, arXiv preprint arXiv:1101.3281 (2011).
- [10] J. M. Kosterlitz and D. J. Thouless, *Journal of Physics C Solid State Physics* **6**, 1181 (1973).
- [11] M. Campostrini, M. Hasenbusch, A. Pelissetto, and E. Vicari, *Physical Review B* **74**, 144506 (2006).
- [12] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, *Physical Review B* **63**, 214503 (2001).



- 
- [13] J. A. Lipa, J. A. Nissen, D. A. Stricker, D. R. Swanson, and T. C. P. Chui, *Physical Review B* **68**, 174518 (2003).
- [14] F. Y. Wu, *Reviews of Modern Physics* **54**, 235 (1982).
- [15] R. H. Swendsen and J.-S. Wang, *Physical Review Letters* **57**, 2607 (1986).
- [16] <http://docs.nvidia.com/cuda/cuda-c-programming-guide/#axzz36f11V4UC>.
- [17] T. Preis, P. Virnau, W. Paul, and J. J. Schneider, *Journal of Computational Physics* **228**, 4468 (2009).
- [18] Y. Fang, S. Feng, K.-M. Tam, Z. Yun, J. Moreno, J. Ramanujam, and M. Jarrell, arXiv preprint arXiv:1311.5582 (2013).
- [19] K. Harada, *Physical Review E* **84**, 056704 (2011).
- [20] P. H. Lundow and I. A. Campbell, *Physical Review B* **82**, 024414 (2010).
- [21] E. Burovski, J. Machta, N. Prokof'ev, and B. Svistunov, *Physical Review B* **74**, 132502 (2006).
- [22] J. Lou, A. Sandvik, and L. Balents, *Physical Review Letters* **99**, 207203 (2007).
- [23] J. V. José, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, *Physical Review B* **16**, 1217 (1977).
- [24] M. Caselle and M. Hasenbusch, *Journal of Physics A: Mathematical and General* **31**, 4603 (1998).
- [25] M. Oshikawa, *Physical Review B* **61**, 3430 (2000).
- [26] J. Manuel Carmona, A. Pelissetto, and E. Vicari, *Physical Review B* **61**, 15136 (2000).
- [27] P. D. Scholten and L. J. Irakliotis, *Physical Review B* **48**, 1291 (1993).
- [28] J. Hove and A. Sudbø, *Physical Review E* **68**, 046107 (2003).

- 
- [29] E. Domany, M. Schick, and R. H. Swendsen, *Physical Review Letters* **52**, 1535 (1984).
- [30] A. C. D. v. Enter and S. B. Shlosman, *Communications in Mathematical Physics* **255**, 21 (2005).
- [31] A. C. D. van Enter and S. B. Shlosman, *Physical Review Letters* **89**, 285702 (2002), arXiv: cond-mat/0205455.
- [32] S. Ota and S. B. Ota, *Physics Letters A* **367**, 35 (2007).