UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

DEPARTMENT OF PHYSICS AND ASTRONOMY

SENIOR HONORS THESIS

Unsupervised Learning of the Phase Transitions for 2D Ising and Potts Models

---

By

Yi HU

Advisor:
Joaquín E. DRUT

September 24, 2024

Abstract

Our study explores and evaluates the application of unsupervised learning for identifying phase transitions in different lattice spin systems, with a focus on the principal component analysis (PCA). The phase transitions of the Ising model and the Potts models with 2, 3, 5, and 10 states in 2D are studied in terms of lattice size and sample space for temperatures. I successfully identified the phase transition temperature using PCA by analyzing the explained variance ratios and lower-dimensional projections.

Contents

List of Figures

List of Tables

List of Abbreviations

QMC	**Q**uantum **M**onte **C**arlo
CNN	**C**onvolutional-**N**eural-**N**etwork
PCA	**P**rincipal-**C**omponent **A**nalysis
MCMC	**M**arkov **C**hain **M**onte **C**arlo

List of Symbols

σ_i	Individual Spin
$\{\sigma_i\}$	A Spin System Configuration
$\mathcal{H}$	Hamiltonian
T	System Temperature
k_B	Boltzmann Constant
J	Nearest-Neighbor Interaction Constant
h	Magnetic Field
E	Dimensionless Energy
M	Dimensionless Magnetization
Z	Partition Function
q	The Number of States for Potts Model
m	Order Parameter (Mean Spin per Site)
N_p	The Number of Nearest Interacting Spins
T_c	Critical Temperature
L	System Size in One Dimension
N	System Size in Two Dimensions (Lattice Size)
NS	Sample Size
n_{ij}	Interaction Variable
$\mathbf{X}$	Input Matrix of Configuration Samples
w_l	Vector of Weights (in PCA)
y_l	Projection Coordinate (in PCA)
β_T	$1/k_BT$
K	J/k_BT

Chapter 1

Introduction

1.1 Historical Overview

The critical behavior at an order-disorder phase transition has been a central interest in research of statistical physics and condensed matter physics. The research is facilitated by a large variety of mathematical models that have been introduced for real life experiments and theoretical modeling. Specifically, the classical Ising and Potts models, which belong to a smaller class of lattice models, have been found useful to model various phenomena and objects, including ferromagnetic materials and lattice gas.

The Ising model provides the archetypal example of the lattice spin models, for which spins take value of ± 1. As probably the single most commonly studied mathematical model of ferromagnetic in statistical mechanics, it has long played an important rule in the research of phase transitions and critical phenomena. The model allows the identification of the phase transition, which is one of the central topics of condensed matter physics research. Initially suggested by Wilhelm Lenz in 1920 to find the Curie's temperature, the model was solved in one dimension by his PhD student Ernst Ising in 1924. His analysis showed there was no phase transition to a ferromagnetic ordered state at any temperature in one dimension[6]. In 1941 Kramers and Wannier obtained the first exact quantitative result for the two-dimensional Ising model[4]. But Ising model in three dimension remains insoluble.

The Potts model is a generalization of the Ising model to more-than-two states, and was first proposed by Ranfrey Potts in 1951. Historically, a four-state version of the model was studied by Ashkin and Teller in 1943, but the model of general q states failed to catch much attention in its early years. In the past few decades however, there has been an increasing interest in the model, mainly because it has proven to be related to a variety of outstanding problems in lattice statistics; the critical behaviors of the model also have appeared to be more general and richer than that of Ising model[13].

In the ensuing efforts to explore the properties of these lattice spin models, many techniques have been borrowed from different fields of mathematics and statistics, such as Probability Theory and Combinatorics, and they have contributed to a more elaborated description of possible critical behaviours for the models. A group of Monte Carlo simulation methods has provided precise values for the critical temperature and critical exponents[6]. Meanwhile, new methods are required since the exponentially large space usually

makes calculations formidable for conventional approaches. There has been an effort to incorporate machine learning into the study of the many-body configurations. For example, convolutional-neural-network (CNN) in combination with quantum Monte Carlo (QMC) simulations has proven to be able to detect and characterize both continuous and discontinuous quantum phase transitions [14][5]. Because machine learning is adept at processing big data intelligently, it provides novel avenues to distill critical phase information and classify phases for condensed matter. This thesis focuses on an unsupervised learning method proposed by Lei Wang in [10] and the application of this method to study the system characteristics and phase transitions for Ising and Potts models.

1.2 Motivation and Research Objectives

As one of the central topics in the research of condensed matter physics, classifying phases of matter captures a lot of research attention in recent years. While the state of a system with large number of constituting particles is often sufficiently represented by only a few variables given the symmetry and degree of freedoms of the system, the identification of phases and phase transitions still becomes increasingly difficult as the number of new states increases. The difficulty comes from the fact that with more and more states, the order parameter used for labeling phases of matter may only be defined in an elusive way[11]. Therefore, new techniques to identify phases of matter and phase transitions are needed to meet the challenge.

Machine learning allows us to extract phase information directly from many-body configurations. As a burgeoning field, machine learning is widely adopted in different areas. While the application of supervised learning (e.g. regression and classification) to condensed matter physics has already successfully predicted crystal structures, approximated density functionals, and classified phases of statistic models[10], such methods by nature have a drawback: they require a set of existing training data and a training phase before making predictions. On the other hand, unsupervised learning requires no training phase. It is a group of self-taught methods that by themselves identify the relationships between data elements and classify the raw data. Even though unsupervised learning cannot be used to make predictions, the fact that it is not necessary for unsupervised learning to assume the phase transition exists makes it a preferred technique over the supervised learning for identifying phase transitions.

The research objective of the thesis is to explore and evaluate the application of the unsupervised learning technique, specifically principal component analysis, to phase transition identification for 2D Ising and Potts models.

1.3 Organization of the Thesis

The thesis is organized as follows. In Chapter 2 we introduce the Ising and Potts models, define the order parameter of the system and the spin correlation, and also discuss the phase transition temperature. In Chapter 3 we explain in detail the general procedure for using the unsupervised machine learning techniques to identify the phase transitions. Specifically, Wolff algorithm for sample generation and principal component analysis (PCA) for feature extraction are studied and explained. Chapter 4 offers the simulation results together with analyses that successfully identify the phase transitions. Chapter 5 concludes the thesis and provides our view for future work.

Chapter 2

Background: Theory

2.1 Spin Systems

Spin systems are random collections of spin variables assigned to the vertices of a lattice. The prototypical example of such a model is the Ising model, for which spins take value ± 1. The Potts model is a generalization of the Ising model to contain more than two possible states for each vertex.

2.1.1 Ising Model

The Ising model is a mathematical model in statistical mechanics originally used to study the behavior of magnetic particles in a magnetic field. The model consists of a collection of "spins" on lattice sites (an example of two-dimensional Ising model is shown in Fig.2.1). Each spin $\sigma_i$takes only the value of $+1$ (up) or -1 (down). The Hamiltonian $\mathcal{H}$ of a given spin configuration $\{\sigma_i\}$ is

$$\mathcal{H} = -J \sum_{<i,j>} \sigma_i \cdot \sigma_j - h \sum_i \sigma_i, \tag{2.1}$$

where $\langle i, j \rangle$ denotes the nearest-neighbor pairs, J is the interaction constant, and h is the external magnetic field. Note that the nearest-neighbor interaction is called ferromagnetic if $J > 0$, anti-ferromagnetic if $J < 0$, and the spins are non-interacting if $J = 0$. In other words, spins desire to be aligned in a ferromagnetic Ising model, and tend to have opposite signs in an anti-ferromagnetic model. Similarly, $h > 0$ if the spins tend to line up with the external field in the positive direction, $h < 0$ if the spins tend to line up in the negative direction, and $h = 0$ if there is no external field influence on the spin sites[8]. We define the dimensionless energy E as

$$E = \sum_{\langle i,j \rangle} \sigma_i \cdot \sigma_j, \tag{2.2}$$

and the dimensionless magnetization M as

$$M = \sum_i \sigma_i. \tag{2.3}$$

The Hamiltonian can thus be expressed as

$$\mathcal{H} = -JE - hM. \tag{2.4}$$

In most cases, we assume there is no external magnetic field and set $h = 0$. If we add the temperature dependence to the equation and define $\beta_T \equiv 1/k_B T$, $K \equiv J/k_B T$, where k_B is the Boltzmann constant and T is the temperature, the Hamiltonian thus becomes

$$-\beta_T \mathcal{H} = KE = K \sum_{\langle i,j \rangle} \sigma_i \cdot \sigma_j. \tag{2.5}$$

The partition function of the Ising model with only nearest-neighbour interactions can be written as

$$Z = \sum_{\{\sigma_i\}} \exp\left(-\beta_T \mathcal{H}\right) = \sum_{\{\sigma_i\}} \exp\left(K \sum_{\langle i,j \rangle} \sigma_i \cdot \sigma_j\right). \tag{2.6}$$

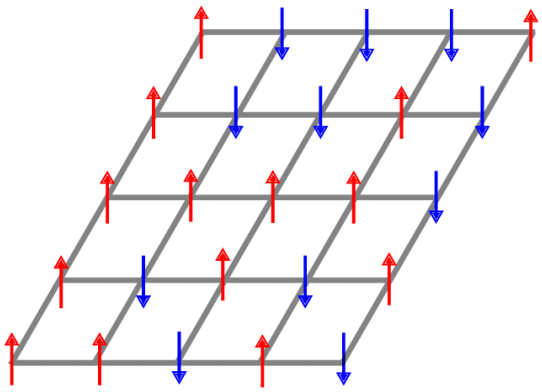


FIGURE 2.1: 2-D Ising Model

2.1.2 Potts Model

The Potts model is another lattice model of interacting spins derived from the Ising model. A q-state Potts model allows spins to take on a complex vector equally-spaced around the unit circle[13]

$$\sigma_i = \exp\left(i 2\pi n/q\right), \quad n = 0, 1, ..., q-1. \tag{2.7}$$

The standard Potts model was suggested by Potts with a Hamiltonian $\mathcal{H}$ for a given configuration $\{\sigma_i\}$ to be

$$-\beta_T \mathcal{H} = KE \tag{2.8}$$

where the dimensionless energy E is defined as

$$E = \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j). \tag{2.9}$$

$\delta(\sigma_i, \sigma_j)$ is the Kronecker delta; it equals one whenever $\sigma_i = \sigma_j$ and zero otherwise.

The partition function of an ordinary Potts model with q-valued spins considering only the nearest-neighbour interactions can be written as

$$Z = \sum_{\{\sigma_i\}} \exp\left(-\beta_T \mathcal{H}\right) = \sum_{\{\sigma_i\}} \exp\Big(K \sum_{\langle i,j\rangle} \delta(\sigma_i, \sigma_j)\Big). \tag{2.10}$$

2.1.3 Order Parameter and Correlation Function

The order parameter describes the average spin state in the system and is defined as the vector mean of all individual spins[7]

$$m = \frac{\langle M\rangle}{N} = \frac{\langle \sum_i \sigma_i\rangle}{N} = \langle \sigma\rangle \tag{2.11}$$

where N is the total number of sites, or the lattice size. For completely uncorrelated systems, the order parameter will evaluate to 0 since the phase factors from the uncorrelated states will cancel.

From the coupling in the Hamiltonian, the correlation $\langle \sigma_i \sigma_j\rangle$ between neighboring spins for a q-state Potts model is[13]:

$$\langle \sigma_i \sigma_j\rangle = \frac{q}{q-1}\frac{1}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}(\delta(\sigma_i,\sigma_j) - \frac{1}{q}), \tag{2.12}$$

where the sum is over neighboring spins. N_p denotes the number of nearest interacting spins for a single site. For a completely correlated system, the probability of having the same spin value with the nearest spins is $1/q$. Therefore, the expectation value of the delta function is $1/q$ and the correlation will evaluate to zero. For a completely correlated system the expectation of the delta function will be 1, and the value in the sum will evaluate to $\frac{N_p(q-1)}{q}$, giving the correlation a value of 1.

We expect Eq.(2.12) works properly for Ising model if $q = 2$. In that case, $\sigma_i = \{-1, 1\}$ for both the Potts and the Ising models, and we have $\sigma_i \cdot \sigma_j = 2\delta(\sigma_i, \sigma_j) - 1$. Eq.(2.12) becomes

$$\langle \sigma_i \sigma_j\rangle = \frac{2}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}(\delta(\sigma_i,\sigma_j) - \frac{1}{2}) = \frac{1}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}(2\delta(\sigma_i,\sigma_j) - 1) = \frac{1}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}\sigma_i \cdot \sigma_j, \tag{2.13}$$

which is the spin correlation for the Ising model. Therefore, Eq.(2.12) holds in the Ising limit.

2.2 Phase Transition

In two dimensions, the critical temperature for Ising model has been found to be[2]

$$\frac{kT_c^I}{J} = \frac{1}{K_I} = \frac{2}{\ln\left(1+\sqrt{2}\right)} \tag{2.14}$$

where T_c^I is the critical temperature for Ising model and $K_I = \frac{J}{kT_c^I}$. And for a q-state Potts model the critical temperature T_c^P satisfies[13]

$$\frac{kT_c^P}{J} = \frac{1}{K_P} = \frac{1}{\ln(1+\sqrt{q})}, \tag{2.15}$$

where $K_P = \frac{J}{kT_c^P}$.

The order parameter and correlation of the system will drop sharply as the temperature approaches from below the critical temperature.

Note that although the Potts model is supposed to be a generalized version of Ising model, the two models are not exactly equivalent when $q = 2$. There is a small caveat: since anti-parallel spin pairs give different values of interaction energy for the two models ($\sigma_i\sigma_j = -1$ for Ising model but $\delta(\sigma_i, \sigma_j) = 0$ for 2-state Potts model), the critical temperature scales correspondingly and gives the relation

$$T_P(q=2) = \frac{1}{2}T_I. \tag{2.16}$$

2.3 Summary

Model	Ising	Potts
Partition function	$\sum_{\{\sigma_i\}} \exp\left(K\sum_{\langle i,j\rangle}\sigma_i\cdot\sigma_j\right)$	$\sum_{\{\sigma_i\}} \exp\left(K\sum_{\langle i,j\rangle}\delta(\sigma_i,\sigma_j)\right)$
Order parameter	$\langle\sigma\rangle$	
Correlation function	$\frac{q}{q-1}\frac{1}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}(\delta(\sigma_i,\sigma_j)-\frac{1}{q})$	$\frac{1}{N_p}\sum_{\sigma_i,\sigma_j}^{N_p}\sigma_i\cdot\sigma_j$
Phase transition	$K_I = \ln(1+\sqrt{2})/2$	$K_P = \ln(1+\sqrt{q})$

TABLE 2.1: Summary

Chapter 3

Background: Methodology

The general procedure of applying unsupervised learning to find the phase transitions is to (1) generate a number of uncorrelated spin configuration samples at different temperatures as input data using Monte Carlo simulation, (2) use dimension reduction techniques to extract essential features from the data matrix and find a low-dimensional representation, and (3) apply cluster analysis algorithms to the dimension-reduced data to find clusters of data representing different phases[10].

In this chapter, Section 3.1 explains the Monte Carlo simulation techniques used for configuration generation, with special focus on Wolf algorithm. Section 3.2 introduces the unsupervised learning, the principal component analysis (PCA), for dimensionality reduction.

3.1 Monte Carlo Simulation

Monte Carlo simulation is a computerized mathematical technique that produces a series of random sampling. There are two basic sampling techniques: direct sampling and Markov Chain sampling. In the direct sampling, the samples are taken independently for each step by choosing a random site on the space; therefore, later samples are independent of the previous ones. However, in the Markov Chain Monte Carlo (MCMC), the next sample depends on the current one. The sampling starts at a given initial state or where the last simulation ends. Then, from the initial state, we move in a random direction to a new state with a distance limited to a predetermined value ∂. Finally, we decide whether to remain in the new state based on the acceptance rate calculated at the new site.

Direct sampling is practically impossible for large configuration spaces. For MCMC, the most general implementation of this procedure is Metropolis algorithm, also known as Metropolis-Hastings algorithm. To apply this method to generating random 2-D spin configurations, a vertex is selected randomly to change its spin. The new system configuration will be accepted if the change reduces the total free energy of the system, or conditionally accepted to a higher energy by a Boltzmann factor. In other words, if the current system configuration μ and the new one ν satisfy $\mathcal{H}_\nu < \mathcal{H}_\mu$, the new system configuration is always accepted and replaces the current configuration; if $\mathcal{H}_\nu \geq \mathcal{H}_\mu$, the new configuration is accepted with an acceptance rate $A(\mu \rightarrow \nu) = e^{-\beta_T(\mathcal{H}_\nu - \mathcal{H}_\mu)}$[3][8]. However, the performance of Metropolis

algorithm suffers as the temperature approaches the critical temperature. A lot of Markov Chain steps are required to obtain two uncorrelated configuration samples because the correlation length diverges and there are critical fluctuations near phase transition.

Wolf Algorithm

The Wolff algorithm is a cluster algorithm that involves flipping large groups of spins simultaneously. Proposed by Wolff[12], this method gets rid of the critical slowing down of the Metropolis algorithm by flipping the whole cluster quickly, as opposed to relying on random reassignment of adjacent sites with large correlation. Baillie in [1] compared the performance of different cluster algorithms for two-dimensional Potts model and found Wolf algorithm converge efficiently. Therefore, Wolf cluster algorithm is chosen for the Monte Carlo simulation of sample configurations for this project. See Appendix B for explanation and a detailed implementation scheme of the cluster Wolf algorithm for Ising and Potts models.

3.2 Unsupervised Learning

3.2.1 Feature Extraction

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical technique used for dimensionality reduction. It uses orthogonal transformation to convert vectors into linear combinations of a set of mutually orthogonal vectors called principal components. Along the directions of these components the variances of the data decrease monotonically. The orthogonal transformation is a linear transformation of the original coordinates $\mathbf{Y} = \mathbf{XW}$. $\mathbf{X}$ with dimensions $n \times m$ is the original data matrix where there are n data entries as row vectors. $\mathbf{W} = (w_1, w_2, ..., w_m)$ is called vectors of weights composed of column vectors w_l that represent the weights of the principal components in the configuration space. Each vector of weights can be obtained by solving the eigenproblem

$$\mathbf{X^T X} w_l = \lambda_l w_l \tag{3.1}$$

where λ_l is the l^{th} eigenvalue of $\mathbf{X^T X}$ sorted in a descending order and w_l is the corresponding eigenvector. The explained variance ratio is the normalized eigenvalue defined as $\tilde{\lambda_l} = \lambda_l / \sum_i \lambda_i$.

The projected coordinates y_l of the l^{th} dimension in the configuration space can be obtained:

$$y_l = \mathbf{X} w_l. \tag{3.2}$$

Chapter 4

Simulation and Results

4.1 Sample generation and procedures

We implement the unsupervised learning technique described in Chapter 3 and conduct five experiments respectively for the Ising model and the Potts model with $q = 2, 3, 5$ and 10.

During each experiment, a number of temperatures near the theoretical critical temperature are chosen and calculated for β_T and $1/K$. For each of the five models, three sets of temperatures shown in Table 4.1-4.3 are used to identify how the choice of temperature sets influences the unsupervised learning results. Temperature set 2 adds four higher temperatures for each model to those in temperature set 1, and temperature set 3 adds four lower temperatures to those in temperature set 1.

100 uncorrelated spin configurations are generated at each temperature and collected into a $M \times N$ matrix $\mathbf{X}$, where M equals $100\times$number of temperatures and N equals the total number of spins. Every 100 rows of the data use the same temperature. Values of 40^2, 80^2, 120^2, and 160^2 are used for N to see how the system size will influence the simulation results.

We mainly focus on analyzing the explained variance ratios and the matrix projection on the first two principal components, as will be discussed in Section 4.2.1 and 4.2.2 respectively. The supporting results of simulation under temperature set 2 and 3 for the five models that are not included in this chapter can be found in Appendix A.

$$\mathbf{X} = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} & \cdots & \sigma_{1,N} \\ \sigma_{2,1} & \sigma_{2,2} & \sigma_{2,3} & \cdots & \sigma_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{100,1} & \sigma_{100,2} & \sigma_{100,3} & \cdots & \sigma_{100,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{M,1} & \sigma_{M,2} & \sigma_{M,3} & \cdots & \sigma_{M,N} \end{pmatrix}_{M\times N} \tag{4.1}$$

4.2 PCA

After sample generation, we need to extract features from the matrix $\mathbf{X}$ by applying PCA. As mentioned in Chapter 3, this step does not assume the existence of the phase transition as most supervised learning methods do, and

Model	$k_BT/J(=1/K)$
Ising	2.1, 2.15, 2.2, 2.25, ..., 2.5, 2.55
2-state Potts	1.0, 1.05, 1.1, 1.15, ..., 1.4, 1.45
3-state Potts	0.7, 0.75, 0.8, 0.85, ..., 1.1, 1.15
5-state Potts	0.6, 0.65, 0.7, 0.75, ..., 1.05, 1.1
10-state Potts	0.6, 0.65, 0.7, 0.75, ..., 1.05, 1.1

TABLE 4.1: Temperature set 1

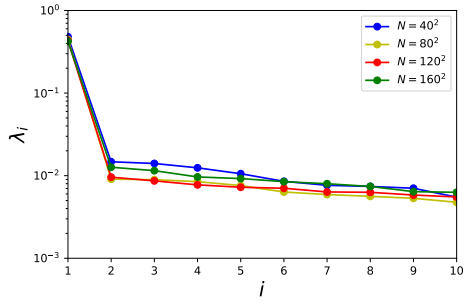
Model	$k_BT/J(=1/K)$
Ising	2.1, 2.15, 2.2, 2.25, ..., 2.7, 2.75
2-state Potts	1.0, 1.05, 1.1, 1.15, ..., 1.6, 1.65
3-state Potts	0.7, 0.75, 0.8, 0.85, ..., 1.3, 1.35
5-state Potts	0.6, 0.65, 0.7, 0.75, ..., 1.25, 1.3
10-state Potts	0.6, 0.65, 0.7, 0.75, ..., 1.25, 1.3

TABLE 4.2: Temperature set 2

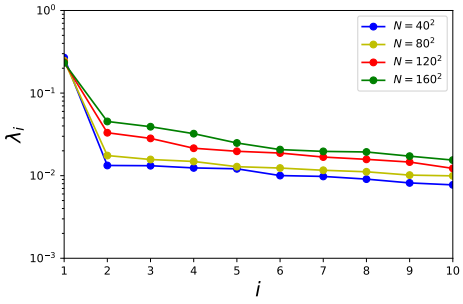
Model	$k_BT/J(=1/K)$
Ising	1.9, 1.95, 2.0, 2.05, ..., 2.5, 2.55
2-state Potts	0.8, 0.85, 0.9, 0.95, ..., 1.4, 1.45
3-state Potts	0.5, 0.55, 0.6, 0.65, ..., 1.1, 1.15
5-state Potts	0.4, 0.45, 0.5, 0.55, ..., 1.05, 1.1
10-state Potts	0.4, 0.45, 0.5, 0.55, ..., 1.05, 1.1

TABLE 4.3: Temperature set 3

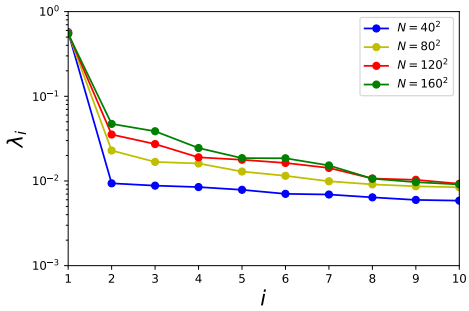
the input matrix $\mathbf{X}$ is the only data we need for performing the unsupervised learning. Our knowledge about the critical temperature of the models is used for verification only.



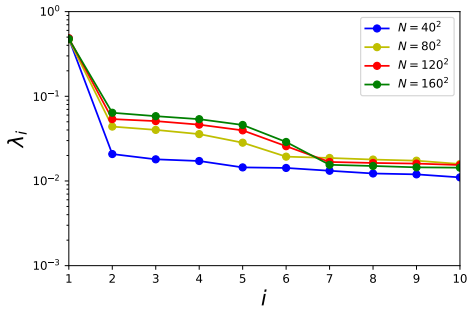
(a) Ising Model



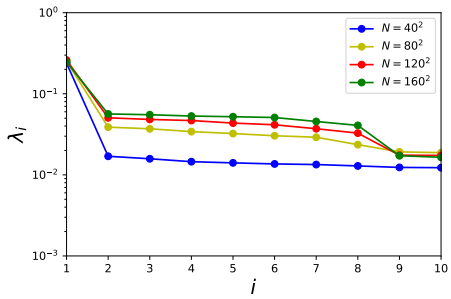
(b) 2-state Potts Model



(c) 3-state Potts Model

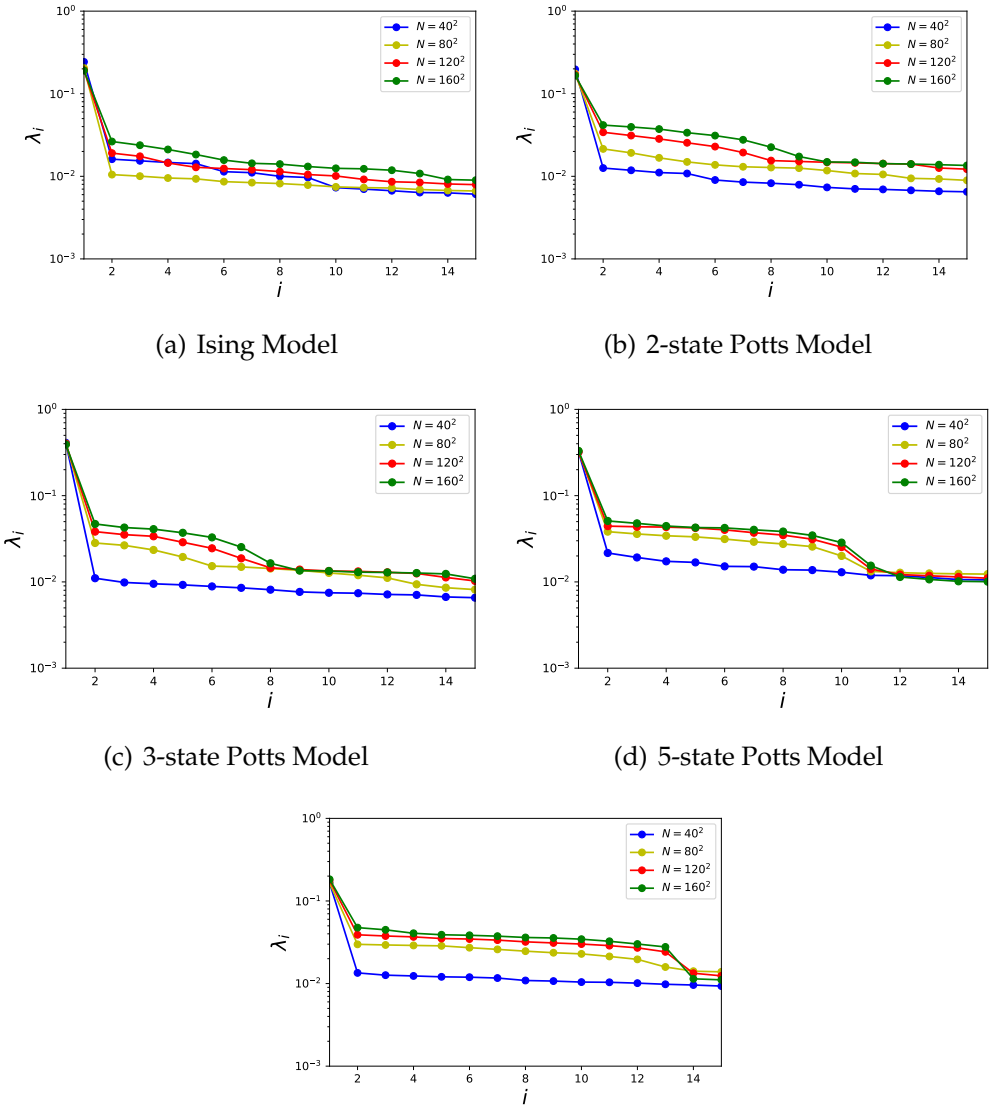


(d) 5-state Potts Model



(e) 10-state Potts Model

FIGURE 4.1: The first ten explained variance ratios under temperature set 1 for the Ising and Potts models.



(a) Ising Model

(b) 2-state Potts Model

(c) 3-state Potts Model

(d) 5-state Potts Model

(e) 10-state Potts Model

FIGURE 4.2: The first fifteen explained variance ratios under temperature set 2 for the Ising and Potts models.

4.2.1 Explained variance ratio

When keeping only the first few principal components, PCA is an effective dimension reduction approach which captures most variations of the original data. The first ten explained variance ratios under temperature set 1 and the first fifteen explained variance ratios under temperature set 2 for the Ising model and the four Potts models of interest are shown in Fig.4.1 and 4.2 respectively. All the plots show that there is only one dominant principal component and it is the case for all five models. Since the temperature is the only variable when generating configuration samples, the first principal component indicates the direction along which the spin systems vary most significantly as the temperature changes.

We can also see from Fig.4.1 and Fig.4.2 that for Potts model, as q increases, there is also an increase in the explained variance ratios of the principal components after the first one. Notably, when the lattice size is large enough, the plots show a nearly flat region with similar variance ratios for a number of components. This region has variance ratios much lower than the variance ratio of the first component, but significantly greater than those of the later ones. If we compare Fig.4.1 and Fig.4.2, the flat region is found to be influenced by the temperature set used for generating the input matrix. Under temperature set 2 with four higher temperature values added to temperature set 1, we get a longer flat region for 5- and 10-state Potts models, and also a tendency of forming the flat region for 2- and 3-state Potts models as shown in Fig.4.2(b) and 4.2(c). The length of the flat region is positively related to the number of high temperatures used.

4.2.2 Low-Dimensional Projection

Following Eq.(3.2) we project the samples onto the space spanned by the first two principal components. The projections for the Ising model and 2- and 3-state Potts models using temperature set 1 are shown in Fig.4.3-4.5 respectively, where the panels (a-d) are for $N = 40^2, 80^2, 120^2$, and 160^2 sites. As expected, the plots show the variation along the first principal axis y_1 is much stronger than that along the second principal axis y_2.

As the lattice size increases, we can see all the data points have the tendency to split into three clusters along the first principal axis. At low temperatures, the samples lie symmetrically at finite y_1, while at high temperatures, they mainly locate around the origin. During the phase transition when the temperature is around the critical one, the samples spread broadly along y_1 axis because of large fluctuations. One also observes that the samples also tend to split into clusters along the second principal axis y_2 as the system size grows. After the sample points converge to the origin horizontally, they move away vertically from the origin as temperature increases.

We calculate the simulated expectation value of $|y_1|$ and y_2 with variance as a function of temperature in temperature set 1, 2 and 3. The results are shown in Fig.4.6-4.10, Fig.A.4-A.8 and Fig.A.13-A.17 respectively, together with the theoretical critical temperature for each system noted. During the phase transition, we expect $|y_1|$ to drop drastically to zero with large fluctuations. We can estimate the phase transition temperatures by looking for a

Model	Identified phase transition temperature	Theoretical T_c
Ising	$2.25 - 2.3$	2.27
2-state Potts	$1.1 - 1.15$	1.13
3-state Potts	$0.95 - 1.0$	0.99
5-state Potts	$0.85 - 0.9$	0.85
10-state Potts	$0.7 - 0.75$	0.7

TABLE 4.4: Identified phase transition temperature with PCA.

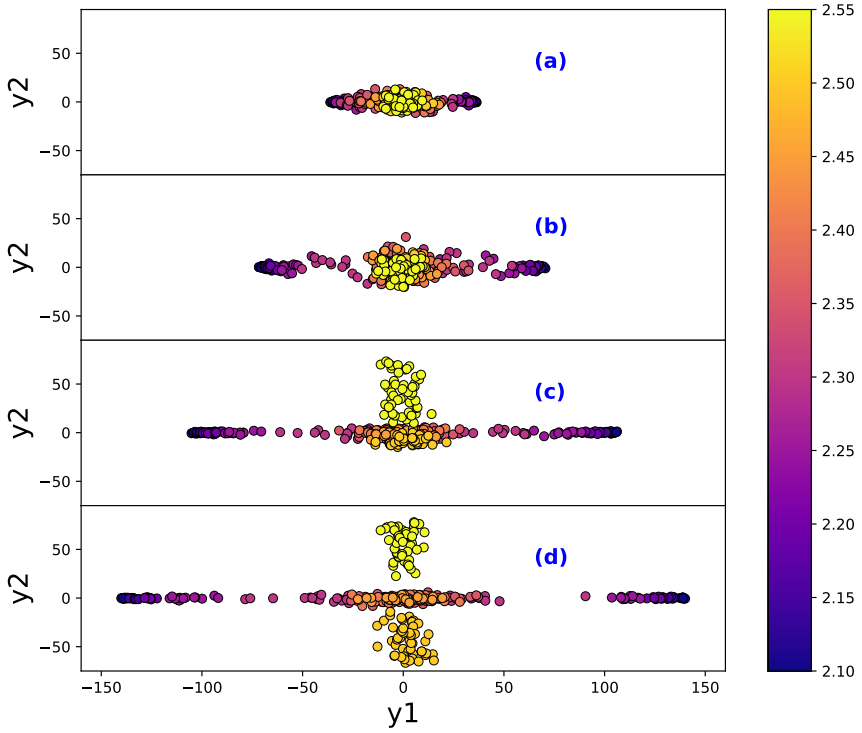


FIGURE 4.3: Projection of the samples onto the plane of the leading two principal components for the Ising model with temperature set 1.

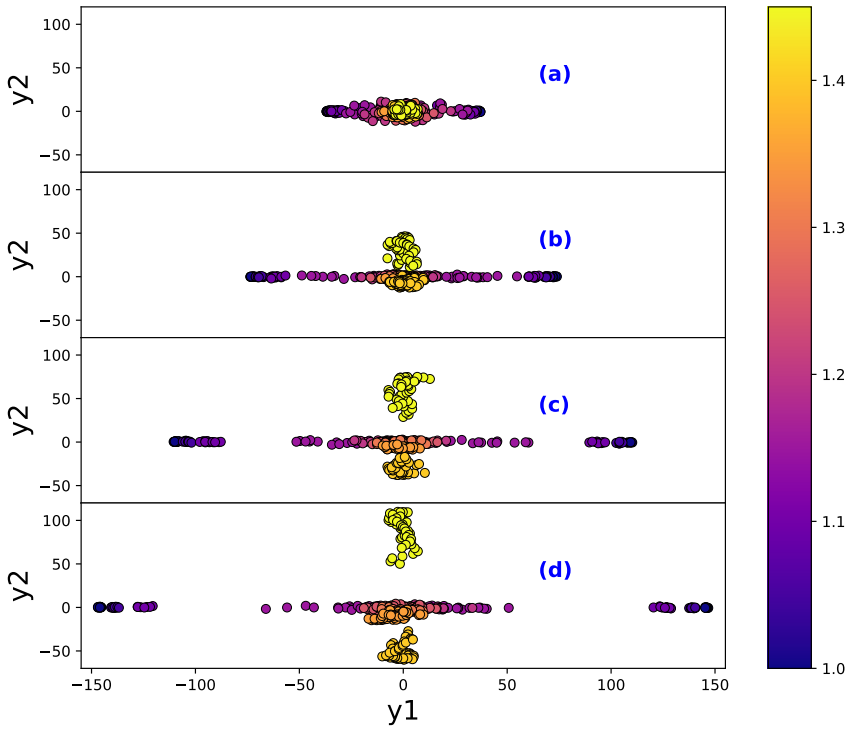


FIGURE 4.4: Projection of the samples onto the plane of the leading two principal components for the 2-state Potts model with temperature set 1.

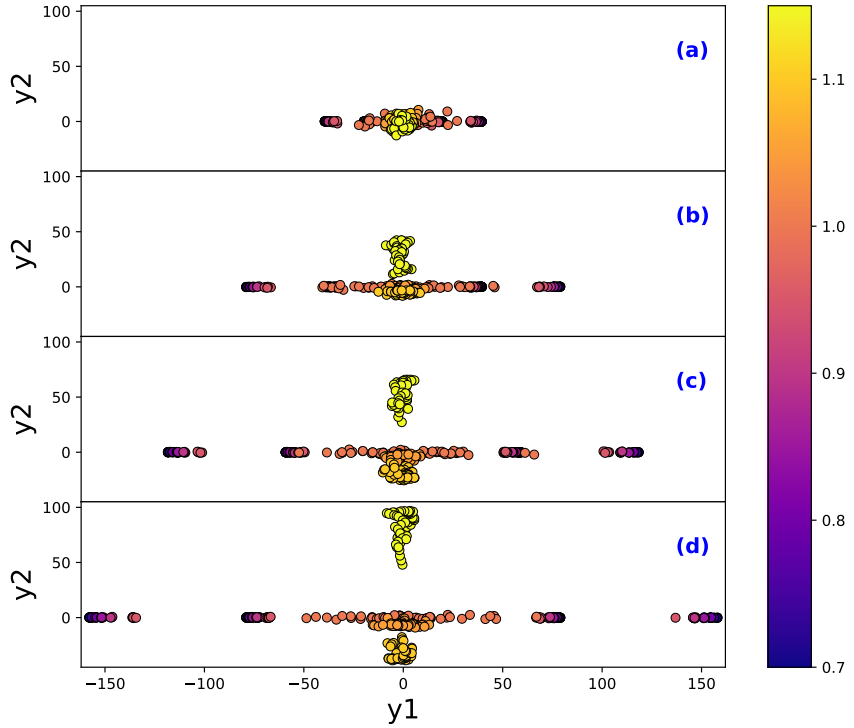
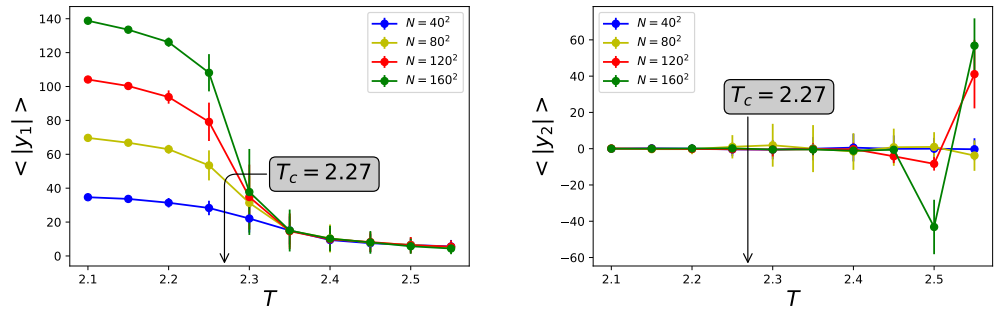


FIGURE 4.5: Projection of the samples onto the plane of the leading two principal components for the 3-state Potts model with temperature set 1.

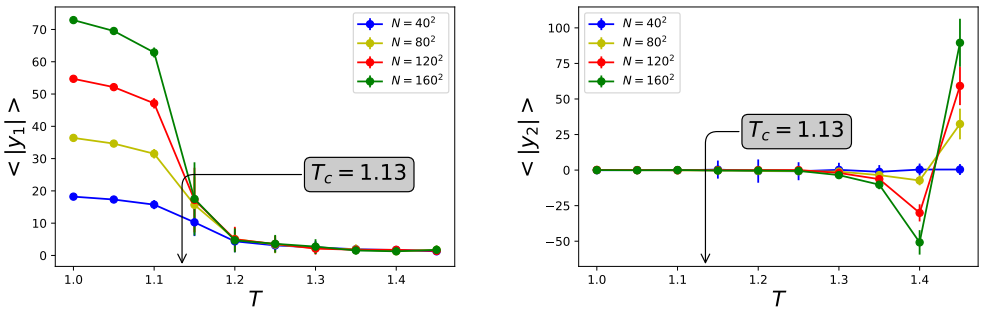
temperature range where there is a sudden drop in $|y_1|$ with large data variance. All plots of Fig.4.6(a)-4.10(a), Fig.A.4(a)-A.8(a) and Fig.A.13(a)-A.17(a) are indeed as expected, and even with different temperature sets, the plots give similar estimation of the temperature of phase transitions for the five models. Table 4.4 summarizes the simulated phase transition temperatures and theoretical true values for different systems. Our identified possible ranges for phase transitions nicely agree with theoretical values.

When looking at y_2 axis from Fig.4.6(b)-4.10(b), Fig.A.4(b)-A.8(b) and Fig. A.13(b)-A.17(b), one observes that the five systems uniformly remain indifferent to the temperature change along the second principal axis before the phase transitions. However, as the temperature increases, samples start reacting to the temperature change. While the starting points of the reaction vary from case to case depending on the model type, the temperature set and the system size, the plots all show oscillation in y_2 at high temperatures after the phase transition predicted by $|y_1|$. Such oscillation in the second principal axis y_2 can be best explained as the statistical compensation for the near-zero values of y_1 and for the intrinsic asymmetry of the input matrix. However, the starting point of reaction of y_2 axis alone does give a good estimation of the upper limit for the critical temperature.



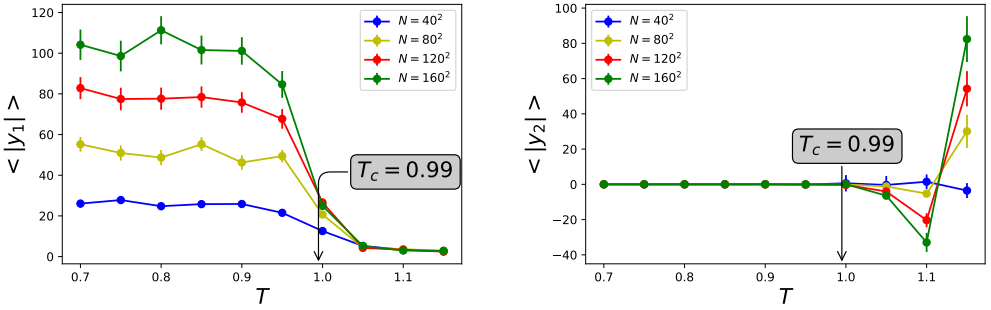
(a) Ising Model, $\langle |y_1| \rangle$ (b) Ising Model, $\langle y_2 \rangle$

FIGURE 4.6: The simulated expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 1 for Ising model.



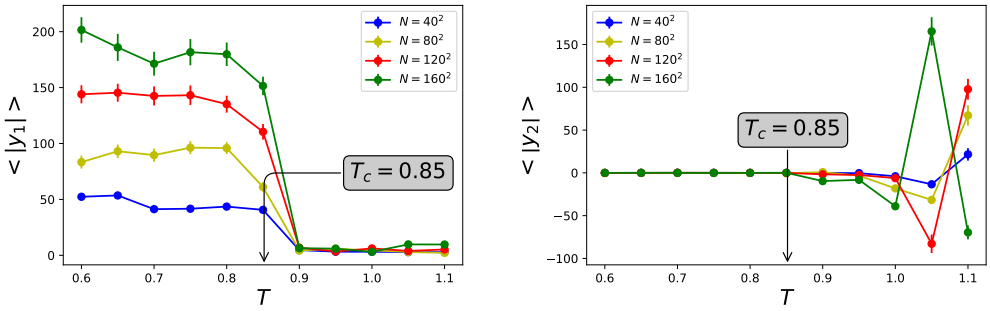
(a) 2-state Potts Model, $\langle |y_1| \rangle$ (b) 2-state Potts Model, $\langle y_2 \rangle$

FIGURE 4.7: The simulated expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 1 for 2-state Potts model.



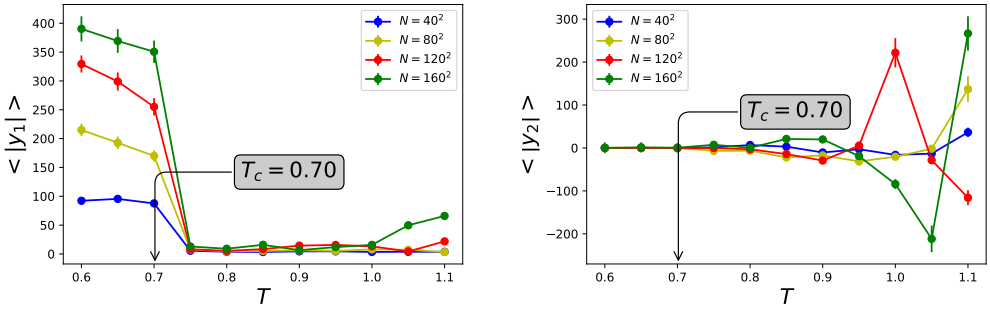
(a) 3-state Potts Model, $\langle |y_1| \rangle$ (b) 3-state Potts Model, $\langle y_2 \rangle$

FIGURE 4.8: The simulated expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 1 for 3-state Potts model.



(a) 5-state Potts Model, $\langle |y_1| \rangle$ (b) 5-state Potts Model, $\langle y_2 \rangle$

FIGURE 4.9: The simulated expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 1 for 5-state Potts model.



(a) 10-state Potts Model, $\langle |y_1| \rangle$ (b) 10-state Potts Model, $\langle y_2 \rangle$

FIGURE 4.10: The simulated expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 1 for 10-state Potts model.

Chapter 5

Conclusion

5.1 Summary and Conclusion

Both the Ising model and the Potts model have wide applicability to many systems of statistical physics and condensed mass physics, and have close connection to natural phase transitions. In this project, we successfully identify the phase transitions of the Ising model and the Potts model with $q = 2, 3, 5$ and 10 using unsupervised learning. The general procedure is to first generate configuration samples under a set of temperatures and put them into a matrix, and then apply PCA to the sample matrix. PCA is able to extract features of the input and reduce it to lower dimensions without losing salient information. By looking at the matrix's projection along the first principal axis, we can discover the phase transitions.

We successfully use this unsupervised learning scheme to pinpoint a small temperature range where the phase transition happens for each model. The range can be further narrowed down by using temperatures that are closer together for sample generation. A qualitative value can also be obtained by doing cluster analysis of the sample projections, but it is outside of the scope of this project.

Along the first principal axis, larger systems show not only a more sudden value shift but also greater fluctuations during the phase transition. Meanwhile, there is also symmetry about the origin, suggestive of a deeper mathematical framework behind those models. On the other hand, the asymmetry along the second principal axis is found to be related to the temperature set for sample generation and the intrinsic asymmetry in samples.

5.2 Future Work

Although with the unsupervised learning scheme we are able to identify and analysis the dominant collective phases of the system related to the phase transition in two dimensions for Ising and Potts model, simple PCA can only identify relations that are linear. Therefore, future work can focus on the identification of nonlinear phase transitions. Some interesting and promising directions are the application of kernel techniques or neural-network based analysis.

While the successful identification of phase transitions using PCA is rather encouraging given how simple the method is, it still lacks the level of accuracy that can be obtained by conventional Monte Carlo simulation techniques. Improvement can be done by incorporating cluster analysis that is able to identify those overlapping phases.

Appendix A

Supporting Simulation Results

A.1 Temperature Set 2

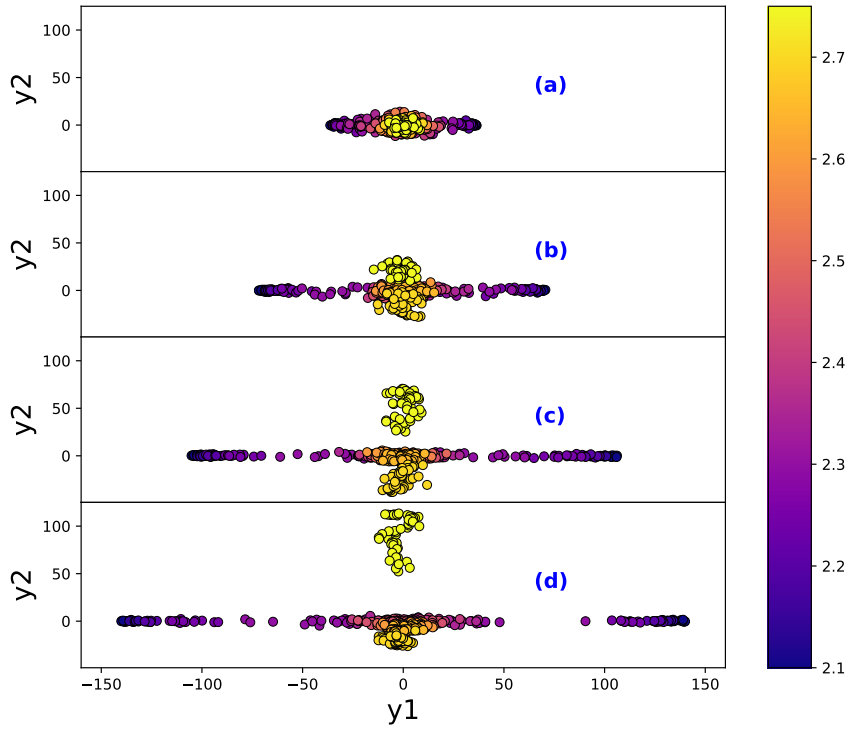


FIGURE A.1: Projection of the samples onto the plane of the leading two principal components with temperature set 2 for the Ising model.

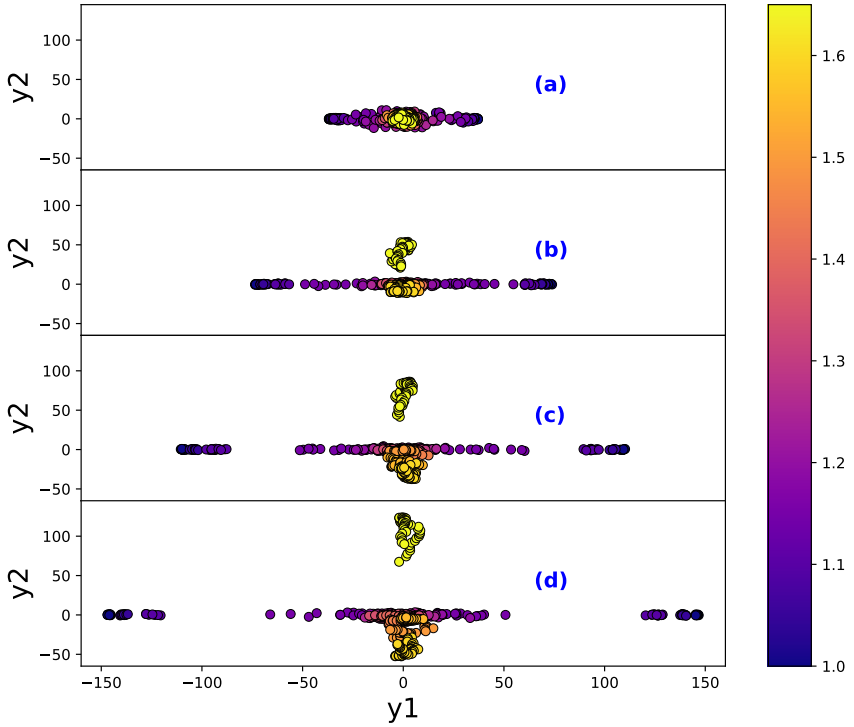


FIGURE A.2: Projection of the samples onto the plane of the leading two principal components with temperature set 2 for the 2-state Potts model.

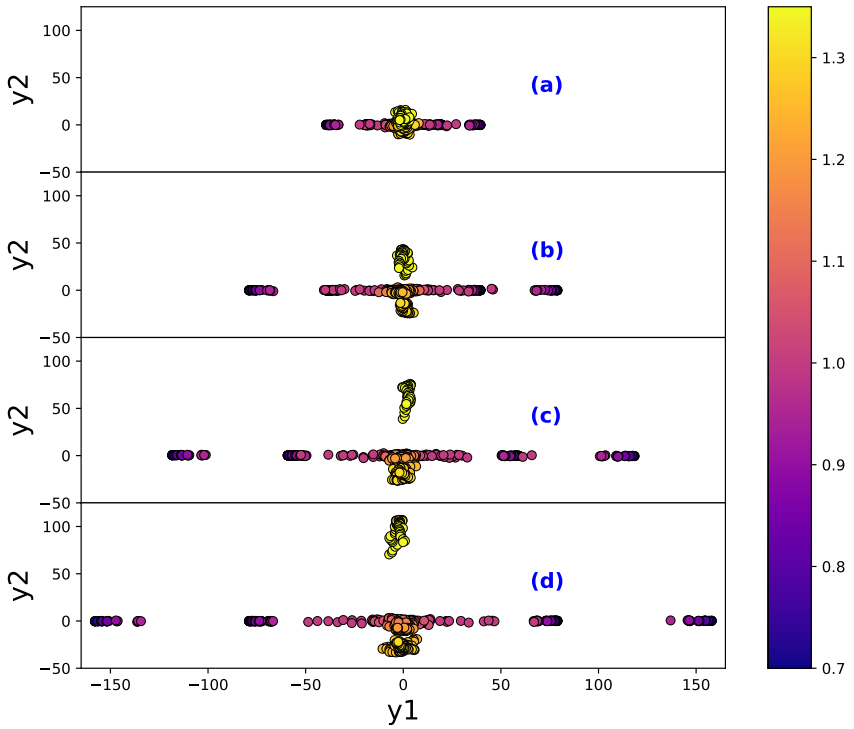
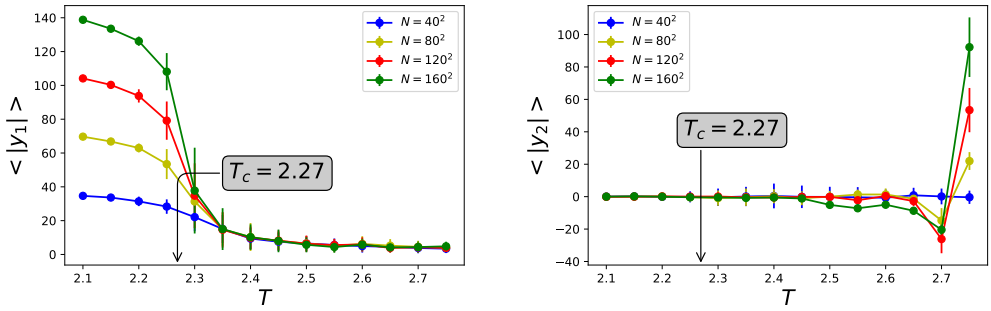
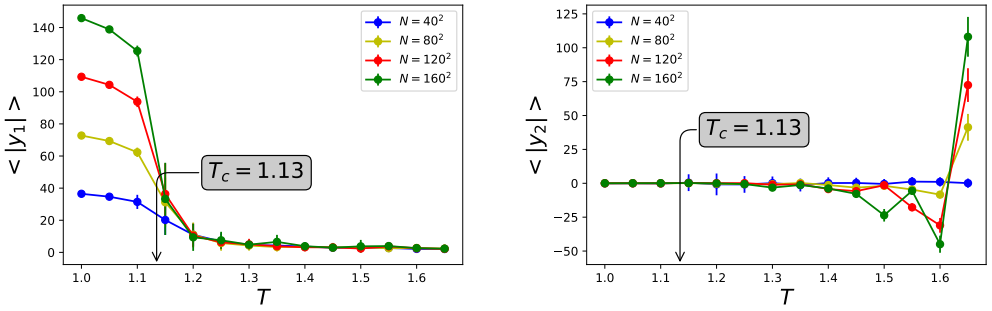


FIGURE A.3: Projection of the samples onto the plane of the leading two principal components with temperature set 2 for the 3-state Potts model.



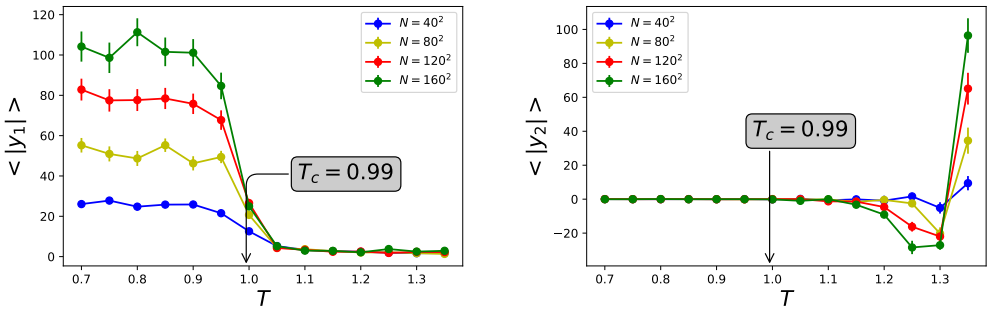
(a) Ising Model, $\langle |y_1| \rangle$ (b) Ising Model, $\langle y_2 \rangle$

FIGURE A.4: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 2 for Ising model.



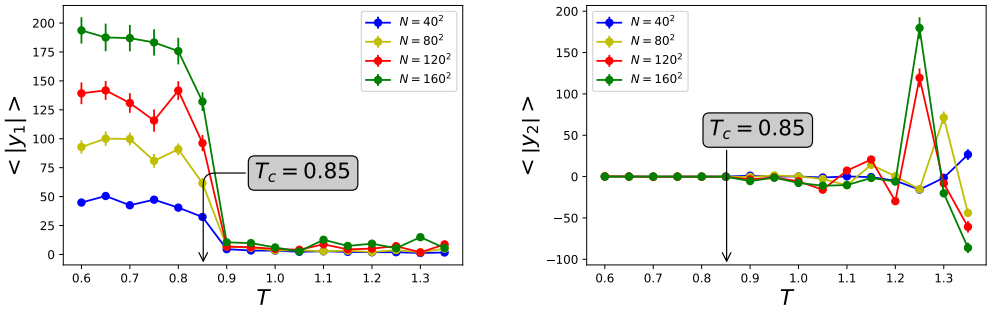
(a) 2-state Potts Model, $\langle |y_1| \rangle$ (b) 2-state Potts Model, $\langle y_2 \rangle$

FIGURE A.5: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 2 for 2-state Potts model.



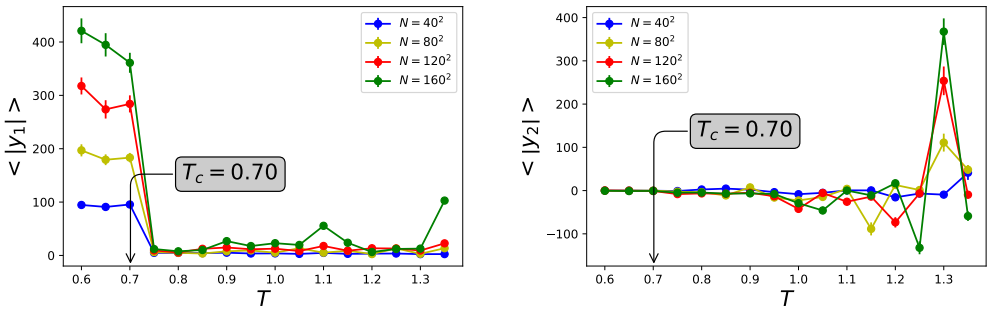
(a) 3-state Potts Model, $\langle |y_1| \rangle$ (b) 3-state Potts Model, $\langle y_2 \rangle$

FIGURE A.6: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 2 for 3-state Potts model.



(a) 5-state Potts Model, $\langle |y_1| \rangle$ (b) 5-state Potts Model, $\langle y_2 \rangle$

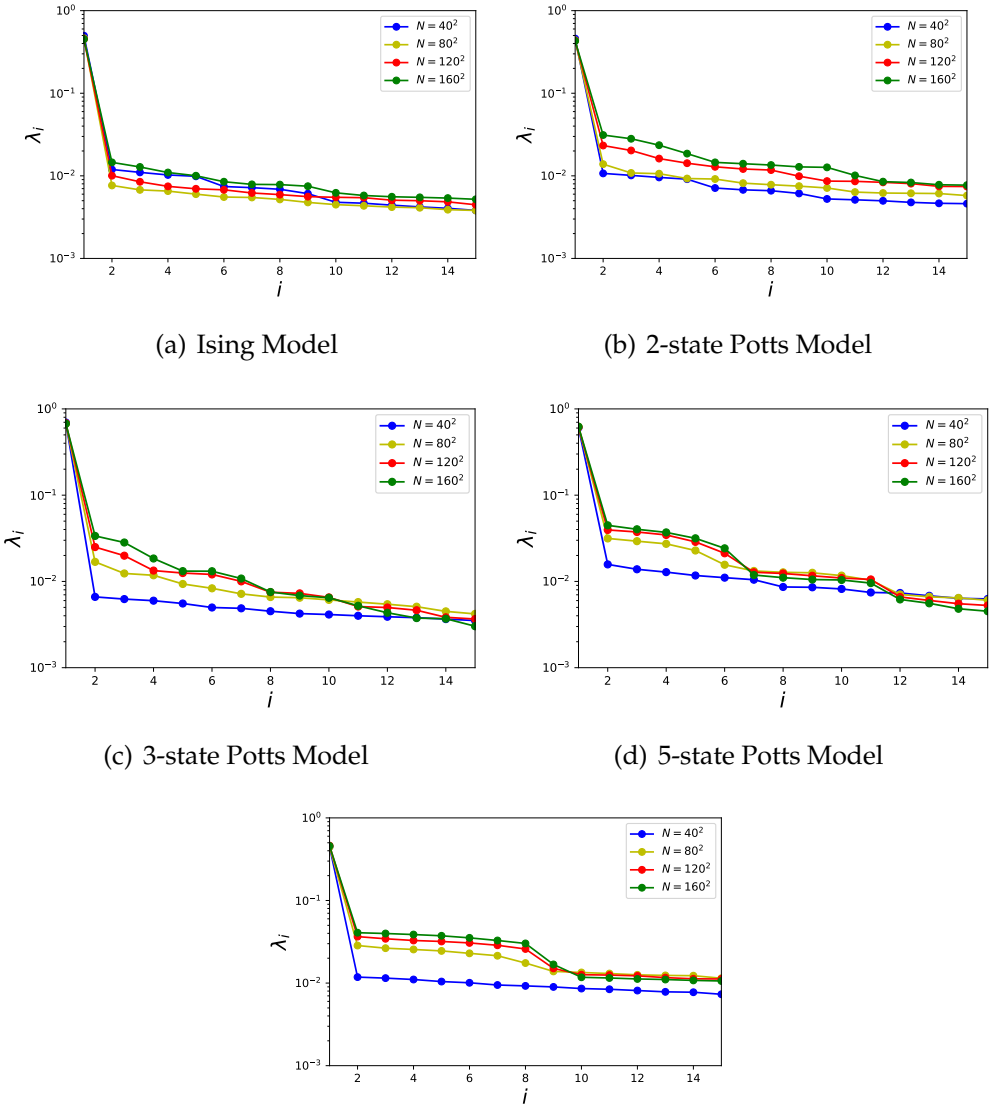
FIGURE A.7: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 2 for 5-state Potts model.



(a) 10-state Potts Model, $\langle |y_1| \rangle$ (b) 10-state Potts Model, $\langle y_2 \rangle$

FIGURE A.8: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 2 for 10-state Potts model.

A.2 Temperature Set 3



(a) Ising Model

(b) 2-state Potts Model

(c) 3-state Potts Model

(d) 5-state Potts Model

(e) 10-state Potts Model

FIGURE A.9: The first fifteen explained variance ratios under temperature set 3 for the Ising and Potts models.

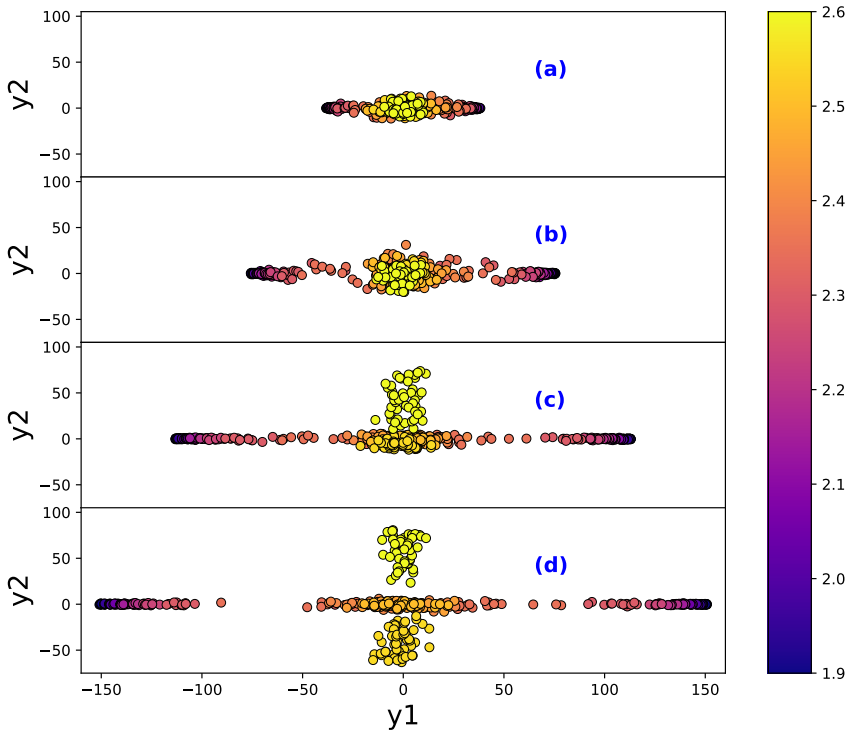


FIGURE A.10: Projection of the samples onto the plane of the leading two principal components with temperature set 3 for the Ising model.

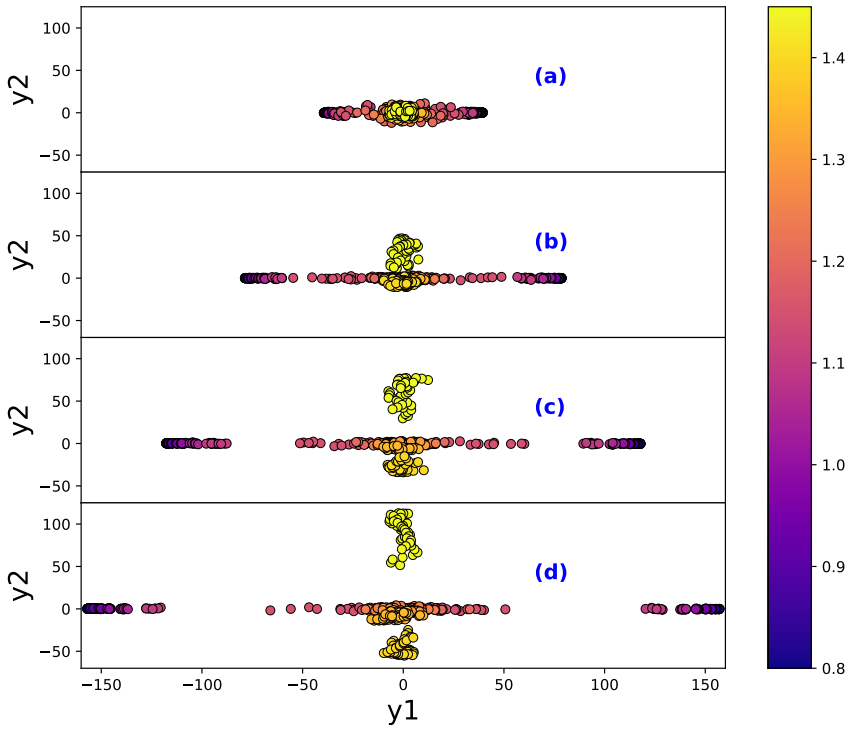


FIGURE A.11: Projection of the samples onto the plane of the leading two principal components with temperature set 3 for the 2-state Potts model.

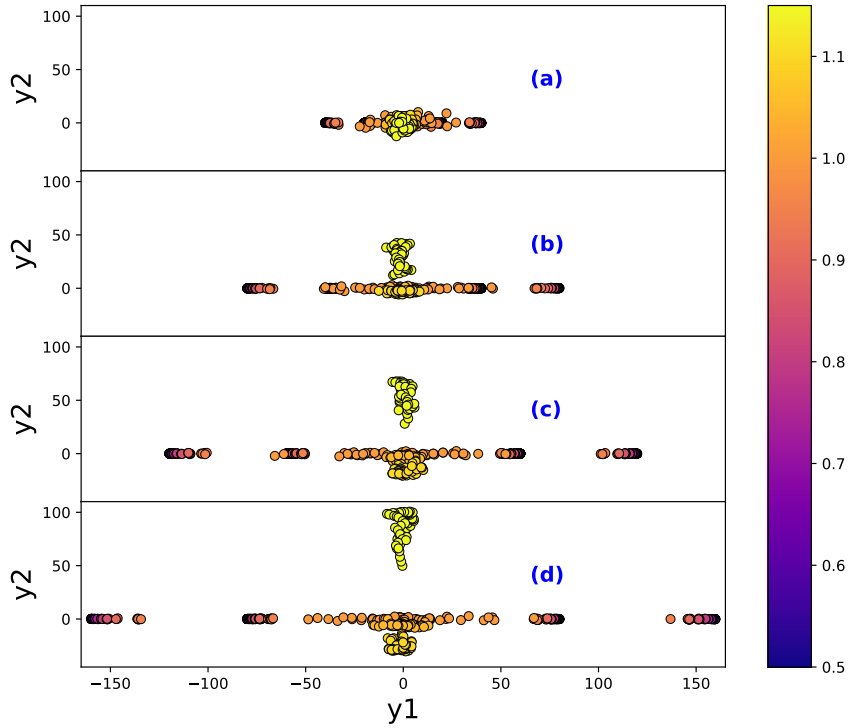
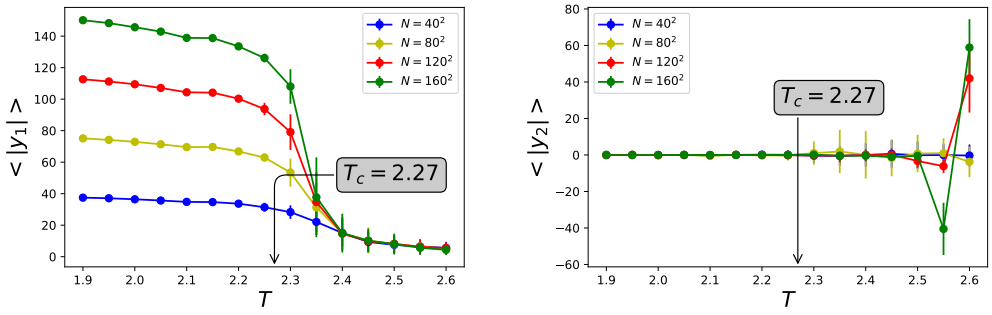
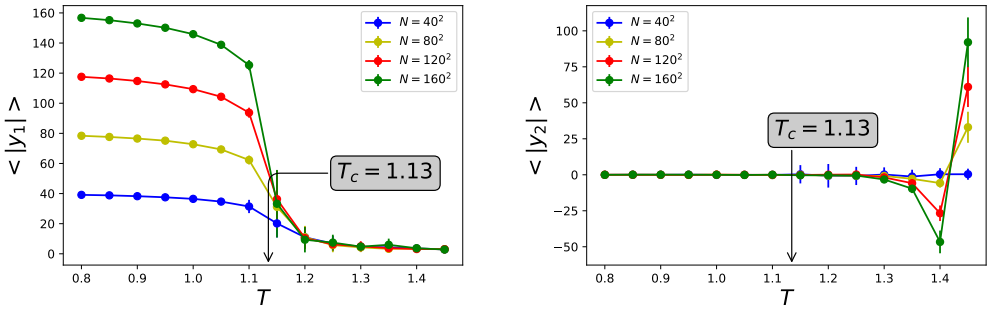


FIGURE A.12: Projection of the samples onto the plane of the leading two principal components with temperature set 3 for the 3-state Potts model.



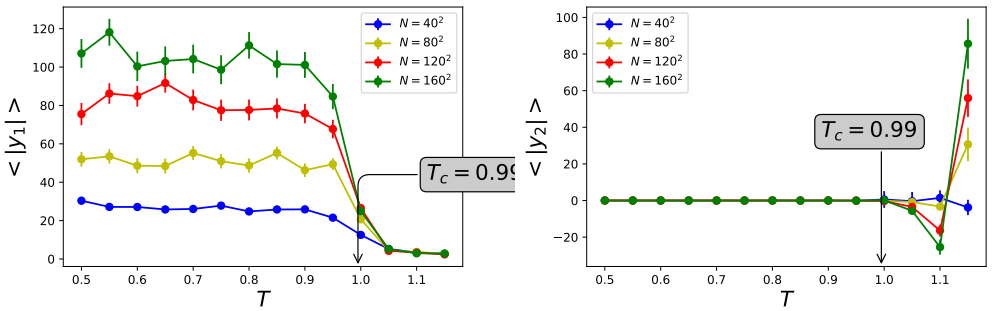
(a) Ising Model, $\langle |y_1| \rangle$ (b) Ising Model, $\langle y_2 \rangle$

FIGURE A.13: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 3 for Ising model.



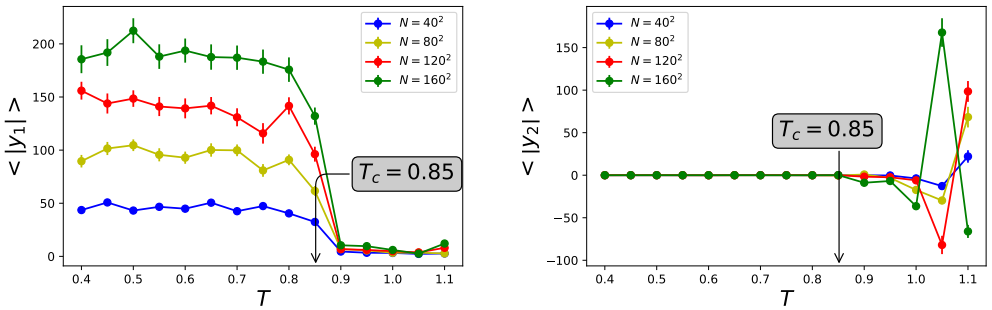
(a) 2-state Potts Model, $\langle |y_1| \rangle$ (b) 2-state Potts Model, $\langle y_2 \rangle$

FIGURE A.14: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 3 for 2-state Potts model.



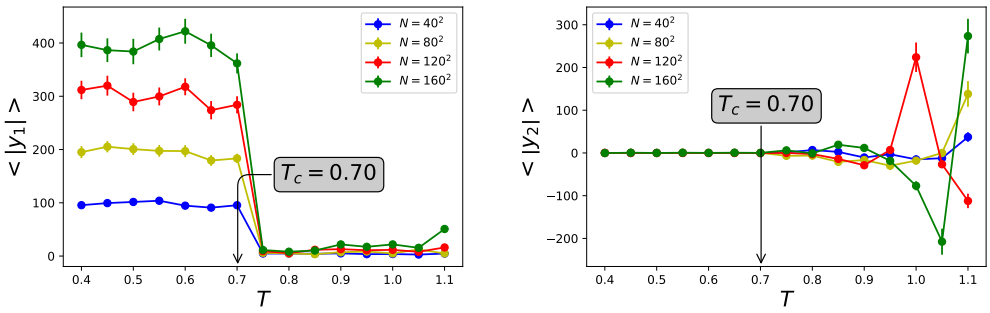
(a) 3-state Potts Model, $\langle |y_1| \rangle$ (b) 3-state Potts Model, $\langle y_2 \rangle$

FIGURE A.15: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 3 for 3-state Potts model.



(a) 5-state Potts Model, $\langle |y_1| \rangle$ (b) 5-state Potts Model, $\langle y_2 \rangle$

FIGURE A.16: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 3 for 5-state Potts model.



(a) 10-state Potts Model, $\langle |y_1| \rangle$

(b) 10-state Potts Model, $\langle y_2 \rangle$

FIGURE A.17: The expectation value of $|y_1|$ and y_2 as a function of temperature in temperature set 3 for 10-state Potts model.

Appendix B

Wolf Cluster Algorithm

B.1 For Ising Model

From Fortuin-Kasteleyn representation, the partition function for the Ising model can be written as[7]

$$\begin{aligned} Z &= \sum_{\{\sigma_i\}} \exp\left(K \sum_{\langle i,j \rangle} \sigma_i \sigma_j\right) \\ &= \sum_{\{\sigma_i\}} \prod_{\langle i,j \rangle} e^K [(1-p) + p\delta(\sigma_i, \sigma_j)] \\ &= \sum_{\{\sigma_i\}} \sum_{\{n_{ij}\}} \prod_{\langle i,j \rangle} e^K [(1-p)\delta(n_{ij}, 0) + p\delta(\sigma_i, \sigma_j)\delta(n_{ij}, 1)] \end{aligned} \tag{B.1}$$

with

$$p = 1 - e^{-2K}, \tag{B.2}$$

where n_{ij} are assigned a value of 0 for inactive bonds and 1 for active bonds.

Equ.(B.1) gives an effective parameterization of the partition function in terms of clusters, and it leads to the cluster simulation algorithm[9]:

1. Choose a random seed site and flip the spin σ_i to $-\sigma_i$.
2. Find all neighboring spins that match the original seed spin and add them to cluster with probability p.
3. If a spin is added to cluster, flip the spin to new spin. Check that spin's neighbors and repeat procedure 2.
4. One step is complete once all spins have been checked.
5. Repeat until convergence.

```
import numpy as np
from numpy.random import rand

def wolffStep(S, beta):
    N = len(S)
    padd = 1-np.exp(-2*beta)
    stack = []
    a = np.random.randint(N)
```

```
    b = np.random.randint(N)
    oldS = S[a,b]
    newS = -oldS
    stack.append(a*N+b)
    S[a,b] = newS

    while (len(stack) != 0):
        cur = stack.pop()
        a = cur//N
        b = cur%N

        a1 = a+1 if a<N-1 else 0
        a2 = a-1 if a>0 else N-1
        b1 = b+1 if b<N-1 else 0
        b2 = b-1 if b>0 else N-1

        if(S[a1,b]==oldS and rand()<padd):
            stack.append(a1*N+b)
            S[a1,b] = newS
        if(S[a2,b]==oldS and rand()<padd):
            stack.append(a2*N+b)
            S[a2,b] = newS
        if(S[a,b1]==oldS and rand()<padd):
            stack.append(a*N+b1)
            S[a,b1] = newS
        if(S[a,b2]==oldS and rand()<padd):
            stack.append(a*N+b2)
            S[a,b2] = newS
    return S

def wolff(S, beta, stepN):
    for i in range (stepN):
        S = wolffStep(S, beta)
    return S


def genSample(L, NS, T):
    r = np.zeros((NS*len(T), L*L))
    for i in range(len(T)):
        t = T[i]
        beta = 1.0/t
        S = np.ones((L,L))
        S = wolff(S, beta, 500)
        for j in range(NS):
            S = wolff(S, beta, 30)
            r[i*NS+j, :] = S.flatten()
    return r
```

B.2 For Potts Model

From Fortuin-Kasteleyn representation, the partition function for Potts model can be written as[7]

$$\begin{aligned} Z &= \sum_{\{\sigma_i\}} \exp\left(K \sum_{\langle i,j\rangle} \delta(\sigma_i, \sigma_j)\right) \\ &= \sum_{\{\sigma_i\}} \prod_{\langle i,j\rangle} e^K[(1-p) + p\delta(\sigma_i, \sigma_j)] \\ &= \sum_{\{\sigma_i\}} \sum_{\{n_{ij}\}} \prod_{\langle i,j\rangle} e^K[(1-p)\delta(n_{ij}, 0) + p\delta(\sigma_i, \sigma_j)\delta(n_{ij}, 1)]. \end{aligned} \tag{B.3}$$

This time, $p = 1 - e^{-K}$. The cluster simulation algorithm is[1][9]:

1. Choose a random seed site and change the spin to a random new spin.
2. Find all neighboring spins that match the original seed spin and add them to cluster with probability p.
3. If a spin is added to cluster, flip the spin to the new spin. Check that spin's neighbors and repeat procedure 2.
4. One step is complete once all spins have been checked.
5. Repeat until convergence.

```
import numpy as np
from numpy.random import rand

def wolffStep(S, beta, q):
    N = len(S)
    padd = 1-np.exp(-beta)
    stack = []
    a = np.random.randint(N)
    b = np.random.randint(N)
    oldS = S[a,b]
    newS = (oldS+np.random.randint(1,q))
    if (newS>=q):
        newS -=q;
    stack.append(a*N+b)
    S[a,b] = newS

    while (len(stack) != 0):
        cur = stack.pop()
        a = cur//N
        b = cur%N

        a1 = a+1 if a<N-1 else 0
        a2 = a-1 if a>0 else N-1
```

```
        b1 = b+1 if b<N-1 else 0
        b2 = b-1 if b>0 else N-1

        if(S[a1,b]==oldS and rand()<padd):
            stack.append(a1*N+b)
            S[a1,b] = newS
        if(S[a2,b]==oldS and rand()<padd):
            stack.append(a2*N+b)
            S[a2,b] = newS
        if(S[a,b1]==oldS and rand()<padd):
            stack.append(a*N+b1)
            S[a,b1] = newS
        if(S[a,b2]==oldS and rand()<padd):
            stack.append(a*N+b2)
            S[a,b2] = newS
    return S

def wolff(S, beta, stepN, q):
    for i in range (stepN):
        S = wolffStep(S, beta, q)
    return S

def genSample(L, NS, T, q):
    r = np.zeros((NS*len(T), L*L))
    for i in range(len(T)):
        t = T[i]
        beta = 1.0/t
        S = np.zeros((L,L))
        S = wolff(S, beta, 500, q)
        for j in range(NS):
            S = wolff(S, beta, 30,q)
            r[i*NS+j, :] = S.flatten()
    return r
```

Bibliography

1. Baillie, C. F. & Coddington, P. D. Comparison of cluster algorithms for two-dimensional Potts models. *Phys. Rev. B* **43,** 10617–10621 (13 May 1991).
2. Baxter, R. *Exactly solved models in statistical mechanics* ISBN: 9780120831807 (Academic Press, London New York, 1982).
3. Binder, K. & Heermann, D. *Monte Carlo Simulation in Statistical Physics: An Introduction* ISBN: 9783642031632. https://books.google.com/books?id=y6oDME582TEC (Springer Berlin Heidelberg, 2010).
4. BRUSH, S. G. History of the Lenz-Ising Model. *Rev. Mod. Phys.* **39,** 883–893 (4 Oct. 1967).
5. Dong, X.-Y., Pollmann, F. & Zhang, X.-F. *Machine learning of quantum phase transitions* 2018. eprint: arXiv:1806.00829.
6. Ising, T., Folk, R., Kenna, R., Berche, B. & Holovatch, Y. *The Fate of Ernst Ising and the Fate of his Model* 2017. eprint: arXiv:1706.01764.
7. Janke, W. in *Computational Physics: Selected Methods Simple Exercises Serious Applications* (eds Hoffmann, K. H. & Schreiber, M.) 10–43 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996). ISBN: 978-3-642-85238-1. doi:10.1007/978-3-642-85238-1_3. https://doi.org/10.1007/978-3-642-85238-1_3.
8. Kardar, M. *Statistical Physics of Fields* ISBN: 9780521873413. https://books.google.com/books?id=nTxBhGX01P4C (Cambridge University Press, 2007).
9. Landau, D. P. & Binder, K. *A Guide to Monte Carlo Simulations in Statistical Physics* 2nd ed. doi:10.1017/CBO9780511614460 (Cambridge University Press, 2005).
10. Wang, L. Discovering Phase Transitions with Unsupervised Learning. doi:10.1103/PhysRevB.94.195105. eprint: arXiv:1606.00318 (2016).
11. Wen, X.-g. *Quantum Field Theory of Many-Body Systems* 2004.
12. Wolff, U. Collective Monte Carlo Updating for Spin Systems. *Phys. Rev. Lett.* **62,** 361–364 (4 Jan. 1989).
13. Wu, F. Y. The Potts model. *Rev. Mod. Phys.* **54,** 235–268 (1 Jan. 1982).
14. Zhao, X. L. & Fu, L. B. *Machine Learning Phase Transition: An Iterative Proposal* 2018. eprint: arXiv:1808.01731.