

分类号：
10112

学校代码：

太原理工大学

硕士学位论文

文

(学术学位)

基于二维 Ising-XY 耦合模型的相变研究

姓名：丁利利
学号：2020511050
培养单位：物理与光电工程学院
学科：物理学
研究方向：蒙特卡洛方法研究
指导教师：田彦婷 副教授
合作导师：张万舟 副教授

论文提交日期：2022 年 11 月

A Dissertation Submitted to
Taiyuan University of Technology
In partial fulfillment of the requirement
For the degree of Master

**Carlo Method and Machine Learning Study on
Phase Transition of Two Dimensional Ising XY
Coupling Model**

By
Lili Ding

College of Physics and Optoelectronics
2022

学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究所取得的成果。除文中已经注明引用的内容外，本论文不包含其他个人或集体已发表或撰写过的科研成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本声明的法律责任由本人承担。

论文作者签名： 签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者和指导教师完全了解太原理工大学有关保留、使用学位论文的规定：学校有权保留并向国家有关部门或机构送交学位论文的复印件和电子版；允许本学位论文被查阅和借阅；学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或其他复制手段保存和汇编本学位论文。

本学位论文属于：保密 ☐ 在____年解密后适用本授权书

不保密 ☐

论文作者签名： 导师签名：

签字日期： 年 月 日 签字日期： 年 月 日

学位论文答辩信息表

论文题目	蒙特卡罗和机器学习方法研究二维 Ising-XY 耦合模型的相变		
答辩日期		答辩秘书	
学位论文答辩委员会成员			
姓名	职称	工作单位	备注

摘 要

相变是自然界广泛存在的一类突变现象，相变理论是统计物理领域里一项长期且重要的研究内容，Ising 模型和 XY 模型是研究该类问题的重要载体。Ising 模型是经典统计力学的一个重要模型，XY 模型是一个更加普遍化的 Ising 模型，其自旋的朝向由简单的上和下解放到了 XY 平面的任意方向。本论文将二维 Ising 模型与二维 XY 模型结合起来得到一个二维双层 XY 耦合模型和一个二维 Ising-XY 耦合模型，分别采用蒙特卡罗方法和图神经网络方法研究了双层 XY 耦合模型和二维 Ising-XY 耦合模型的相变情况。

首先，采用蒙特卡罗方法中的 wolff 算法研究双层 XY 耦合模型的相变情况，固定 J_a ，通过调整 J_b 大小改变空间耦合系数；通过模拟计算不同耦合系数下的物理量，例如磁化强度、比热、Binder 累积量、超流密度等，再运用尺寸外推公式可以分析确定 Ising 型相变点和 XY 型相变点的位置。研究发现在空间各向同性极限、各向异性、各向异性极限三种情况下双层 XY 模型的 XY 型和 Ising 型同时发生相变，即只能观察到一次相变发生。

其次，采用 wolff 算法研究了二维 Ising-XY 耦合模型的相变情况，首先，通过计算比热大致确定相变点，分析了 $A = B$ 时的 AC 平面相图，研究发现在各向同性、各向异性、各向异性极限三种情况下 Ising 模型与 XY 模型发生的顺序各不相同。

最后，引入了颇具前景的图神经网络方法来模拟二维 Ising-XY 模型。相比较与之前的机器学习方法，图神经网络无疑是更适合物理模型，图神经网络的边特性可以较好的体现 Ising-XY 构型中不同自旋点之间的联系。经过实验可知，图神经网络具有的强大特性可以有效地模拟从传统蒙特卡罗方法中生成的构型，并且是全新的具有相同物理特性的构型。

关键词：蒙特卡罗算法;Ising 模型;XY 模型;各向异性;图神经网络

ABSTRACT

The Ising model and the XY model are very important models in the fields of thermodynamics and statistical physics, and have received extensive attention. In this paper, a two-dimensional Ising-XY coupled model is obtained by combining the two-dimensional Ising model with the two-dimensional XY model, and two possible phase diagrams of the model in the case of spatial anisotropy under periodic boundary conditions are given. There are four phases in the first case: (a) a disordered phase; (b) a Ising-ordered and XY-disordered; (c) a Ising-disordered and XY-ordered; (d) a fully ordered phase. In the second case, there are three phases in total: (a) Ising, XY are disordered; (b) Ising is ordered, XY is disordered; (c) Ising, XY are ordered. Therefore, there is an Ising-type phase transition and a KT phase transition in the Ising-XY model, and the sequence of these two phase transitions is not unique.

In this paper, we study the physical phenomena, sequence parameters and corresponding phase diagrams of the two-layer XY model and Ising-XY coupled model in spatial anisotropy under periodic boundary conditions. In this paper, we use the metropolis and wolff algorithms, which are commonly used in Monte Carlo, to analyze and compare the results with those obtained by previous theoretical derivations.

In the first chapter, Monte Carlo methods and their background are described; then the current research status of Monte Carlo methods in the field of statistical physics is briefly explained; then several commonly used Monte Carlo methods are introduced mainly including metropolis algorithm, wolff algorithm, Wang-Landau algorithm, and Swendsen-Wang algorithm.

In the second chapter, the physical model of this research is mainly introduced. Firstly, the common Ising model and XY model are introduced, then the double-layer XY model and Ising-XY coupling model are introduced in detail, and the concept of phase transition is briefly explained, and finally, the research content of this paper is introduced.

In Chapter 3, the phase transition of the two-layer XY coupled model is investigated by the wolff algorithm; fixing J (a) and changing the spatial coupling coefficient by adjusting the size of J (b); calculating the physical quantities under different coupling coefficients by simulation, such as

magnetization strength, specific heat, Binder accumulation, superfluid density, etc., and then applying the dimensional extrapolation formula can analytically determine the Ising-type phase transition point and XY-type phase transition point can be determined analytically by applying the dimensional extrapolation formula. It is found that the XY-type and Ising-type phase transitions of the two-layer XY model occur simultaneously in the three cases of spatial isotropy limit, anisotropy, and anisotropy limit, i.e., only one phase transition occurrence can be observed.

In Chapter 4, a study of the phase transition in the two-dimensional Ising-XY coupled model using the wolff algorithm is presented. First, the phase transition point is roughly determined by calculating the specific heat, and the AC plane phase diagram at $A=B$ is analyzed, and it is found that the Ising model occurs in a different order from the XY model in the three cases of isotropic, anisotropic, and anisotropic limits.

In Chapter 5, the current promising machine learning method is introduced to simulate the two-dimensional Ising-XY model. Compared with the previous machine learning methods, the graph neural network is undoubtedly more suitable for the physical model, and the edge properties of the graph neural network can better reflect the connection between different spin points in the Ising-XY configuration. After the experiments, it can be seen that the powerful properties of the graph neural network can effectively simulate the configurations generated from the traditional Monte Carlo methods, and they are new configurations with the same physical properties.

Keywords: Monte Carlo method; Ising model; XY model; anisotropic; Graph Neural Networks

目 录

摘 要	I
ABSTRACT	III
第 1 章 绪论	1
1.1 相变的分类	1
1.2 相变的研究方法	2
1.3 相变模型介绍	3
1.3.1 Ising 模型	3
1.3.2 XY 模型	4
1.3.3 Potts 模型	5
1.3.4 双层 XY 耦合模型	6
1.3.5 Ising-XY 耦合模型	6
第二章 蒙特卡罗方法	7
2.1 蒙特卡罗方法在统计物理领域的应用	9
2.2 蒙特卡罗的抽样方法	10
2.3 常用的蒙特卡罗方法	11
2.3.1 Metropolis 算法	11
2.3.2 Wolff 算法	14
2.3.3 王-朗道算法	15
2.3.4 Swendsen-Wang 算法	16
2.4 本论文的研究内容	17
第 3 章 蒙特卡罗方法研究空间各向异性双层 XY 耦合模型的相变	18
3.1 引言	18
3.2 模型与算法	18
3.2.1 模型	18

3.2.2 算法	19
3.2.3 物理量	21
3.3 结果与讨论	22
3.3.1 局部细节图	22
3.3.2 相图	25
第 4 章 蒙特卡罗方法研究二维 Ising-XY 耦合模型的相变	27
4.1 引言	27
4.2 模型及算法	27
4.2.1 模型	27
4.2.2 算法	29
4.2.3 物理量	30
4.3 结果与讨论	31
4.3.1 $A = B$ 相图	31
4.3.2 $A \neq B$ 相图	33
4.4 本章小结	35
第五章 图神经网络生成 IsingXY 模型	37
5.1 引言	37
5.2 模型与算法	37
5.2.1 Ising 模型	37
5.2.2 图神经网络	38
5.2.3 图卷积网络	39
5.2.4 变分自编码器	41
5.2.4 图变分自编码	43
5.3 模拟结果与分析	44
5.4 总结	44
第 6 章 结论与展望	46
展望:	47

参考文献	49
攻读学位期间取得的研究成果	52
致 谢	54

第 1 章 绪论

物质系统中化学物理性质完全相同，与其他部分具有明显分界面的均匀部分称为相。相变是指物质在一定条件下从一种物态转变为另一种物态的现象，相变现象在自然界中广泛存在，其中最常见有固-液相变、液-气相变等。本文主要研究统计物理中常见模型的相变，统计物理领域里一项长期且重要的任务是研究相变，相变理论既是统计物理学的基本科学问题，也可应用于若干实际物理体系。

在统计物理中，相变是指一个系统在一定的外部条件（如温度、压强、化学势等）下，随着某个内部参数（如粒子密度、自旋取向等）的改变，发生宏观物理性质的急剧变化的过程。相变的一个显著特征是在转变点附近系统的物理性质的变化呈现非连续性。例如，在固-液相变中，加热固体会使其温度上升，但当温度达到一定临界点时，固体会突然熔化成为液体，而在这个临界点上下，固体和液体的性质变化非常显著，比如说它们的热容和密度等性质的变化。物质内部粒子间的相互作用使系统倾向于有序，粒子自生的热运动使系统倾向于无序，粒子间相互作用和粒子自生热运动竞争产生了相变。二者缺一不可，如果没有相互作用不可能形成有序态，例如理想气体无法从气态相变为液态和固态；如果自旋磁矩间没有相互作用，就不会有铁磁相变；如无热运动干扰，系统只会是有序态，也不会有相变。

在相变研究领域，Ising 模型和 XY 模型是两个经典的理论模型，它们在理论和实践中都被广泛应用。尤其是 Ising 模型在各个领域都有应用，比如统计物理、量子场论、高能物理、材料科学等等。Ising-XY 耦合模型是 Ising 模型和 XY 模型的耦合，能够更好地描述真实物理系统的性质。研究 Ising-XY 模型的相变行为有助于更深入地理解物质相变机制，并为材料科学、能源材料、物理化学等领域提供理论支持。此外，Ising-XY 模型的相变研究也是统计物理领域的热点问题之一，吸引了众多学者的关注。当前，蒙特卡罗方法是研究相变问题的一种重要方法，已被广泛应用于相变的数值模拟和理论研究。尤其是在 Ising-XY 耦合模型的相变研究中，蒙特卡罗方法已经成为了主流的研究手段。通过模拟不同温度下 Ising-XY 模型的相变行为，可以研究相变的物理规律和性质，以及相变的临界现象等。研究 Ising-XY 模型的相变行为具有很大的意义和价值，有助于推进相变研究的进一步发展，也为物理学、化学、材料科学等应用领域提供理论指导。

1.1 相变的分类

第一种是著名的欧仑菲斯特相变理论，它的分类依据是考察热力学自由能作为热

力学变量的函数的性质。在这种分类下，如果自由能对某个热力学变量的导数在相变点的位置从 n 阶开始不连续，那么这样一个相变就称为 n 级相变。例如，如果自由能对某个热力学变量的一阶导数在相变位置不连续，那么该相变就被定义为一级相变；如果自由能对某个热力学变量的一阶导数在相变位置连续而二阶导数不连续，那么该相变就被定义为二级相变。二级及以上相变也称为连续相变。因此，气体液体固体的相互转变均被定义为一级相变，原因是密度即自由能对化学势的一阶导数在相变点存在不连续的变化。关于二级相变典型的例子是铁磁相变，在转变中磁化强度作为自由能对外加磁场的导数，在高温时为，同时会随着温度的降低连续增大。但是自由能对外加磁场的二阶导数，也就是磁化率，则在相变点的位置存在着不连续的变化。然而欧仑菲斯特分类法存在明显的缺陷，原因是它没有考虑到自由能的导数会在热力学极限下发散这种特殊情形。

第二类方法通常被称为现代分类法。在这种分类理论当中，如果相变过程中存在潜热，那就是该相变则为一级相变。例如冰的融化相变和水的沸腾都是属于一级相变。二级相变，也称作连续相变，其特征是无限大的关联长度、临界点附近按幂次律规则衰减的关联、以及发散的磁化率等。例如铁磁相变属于二级相变。另外，自然界中还存在无穷阶的相变，最著名的例子就是二维 XY 模型中的 KTB 相变。

序参量是刻画相变的有效参数，通常情况下它在相变过程中的一个相中的值为零而在另一个相中则为有限值。对于固液或汽液相变，可以选密度作为序参量；对于铁磁相变而言，我们可以选择净磁化强度作为序参量。前苏联物理学家朗道、金兹堡等提出“对称性破缺”理论来解释相变。该理论认为连续相变的特征是物质对称性和有序性的改变，在相变临界温度以下，系统的序参量非零、称性低、有序性高。在相变临界温度以上，系统的序参量为零、称性高、有序性低。随着温度的降低，序参量在临界点连续地从零变到非零。

除了常见的固液气三态之间的相变外，物质从铁磁变成顺磁、从金属变成超导或绝缘体等也都可以用相应物理量的对称性破缺来描述和解释。因此人们曾认为“对称性破缺”理论已经非常完美，以后凝聚态物理的研究仅仅是修修补补了。

1.2 相变的研究方法

相变是物理学研究中非常重要的一个主题，通常用 Ising 模型、XY 模型、Potts 模型、XY 耦合模型、Ising-XY 耦合模型等来研究相变现象。针对不同的模型特征物理学家们也设计了不同的算法，常见的有 metropolis 算法、wolff 算法、王-朗道算法、Swendsen-Wang 算法等。本文将主要对 XY 耦合模型、Ising-XY 耦合模型以及 metropolis 算法、wolff 算法进行介绍。

1.3 相变模型介绍

相变是自然界广泛存在的一类突变现象也是研究热力学与统计物理的问题，而 Ising 模型和 XY 模型正好是研究该类问题的重要载体。Ising 模型也是经典统计力学的一个重要模型，由德国物理学家威廉·楞次在 1920 年提出以描述铁磁性物质的内部的原子自旋状态及其与宏观磁矩的关系。伊辛模型可以近似描述单轴各向异性铁磁体的铁磁-顺磁相变，稍加改变还可以描述气-液相变、合金的有序无序相变等情形。20 世纪 30-40 年代，物理学家们使用平均场近似理论对二维伊辛点阵模型进行了研究。而 XY 模型是一个更加普遍化的 Ising 模型，其自旋的朝向由简单的上和下解放到了 XY 平面的任意方向。这些年来，Ising-XY 耦合模型、双层 XY 耦合模型、完全阻挫 XY 模型已经引起了人们的广泛关注。

1.3.1 Ising 模型

Ising 模型之所以具有如此广泛的应用并不仅仅在于它的模型机制的简单性，更重要的是它可以模拟出广泛存在于自然、社会、人工系统中的临界现象。Ising 模型是统计物理中迄今为止唯一一个同时具备：表述简单、内涵丰富、应用广泛这三种优点的模型。最初 Ising 模型的提出是为了用来描述单轴铁磁体的自旋现象、解释铁磁物质的相变，即铁磁在加热到一定临界温度以上会出现磁性消失的现象，而降温到临界温度以下又会表现出磁性。这种有磁性、无磁性两相之间的转变，是一种连续相变（也叫二级相变）。Ising 模型假设铁磁物质是由一堆规则排列的小磁针构成，每个磁针只有上下两个方向（自旋 $\sigma_i = \pm 1$ ）。相邻的小磁针之间通过能量约束发生相互作用，同时又会由于环境热噪声的干扰而发生磁性的随机转变（上变为下或反之）。存在某个临界温度 T_c ，当温度高于这个临界点时，随机涨落干扰越强，小磁针越容易发生无序而剧烈地状态转变，从而让上下两个方向的磁性相互抵消，材料的磁性消失；而当温度低于这个临界点时则小磁针相对安宁，系统处于能量约束高的状态，大量的小磁针方向一致，系统中又会出现稳定的自发磁化强度。这个临界温度又被称为居里温度。而当系统处于临界温度 T_c 的时候，Ising 模型表现出一系列幂律行为和自相似现象。所谓的临界现象，是指系统在相变临界点附近的时候表现出的一系列的标度现象，以及系统在不同尺寸之间的相似性。临界系统之中不同组成部分之间还会发生长程的关联，这种通过局部相互作用而导致长程联系的现象恰恰如同真实复杂系统，如社会、经济、认知神经网络等。因此，伊辛模型不仅仅是一个统计物理模型，它更是一个建模各种复杂系统模型的典范。分子排布为周期性点阵

Low T

High T

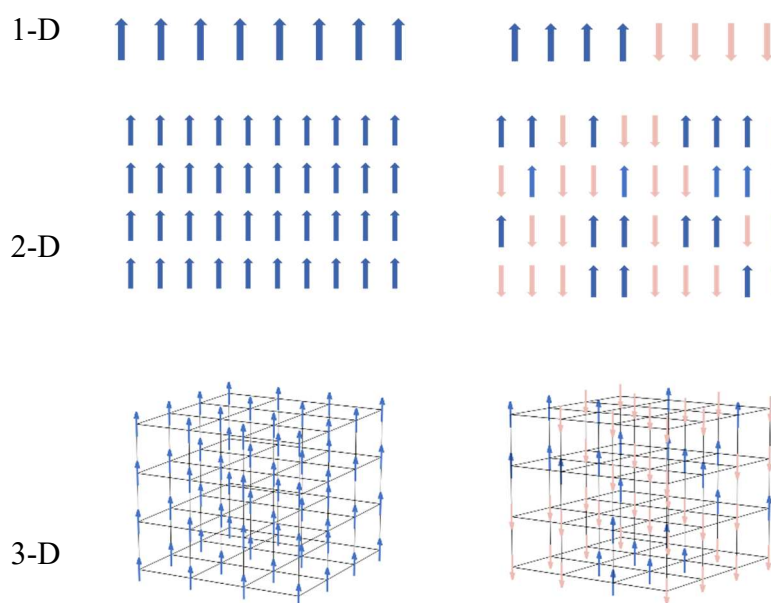


图 2-1 一维、二维、三维下低温、高温时 Ising 模型示意图

Fig.2-1 Schematic diagram of Ising model at low and high temperatures in 1D, 2D and 3D

Ising 模型的哈密顿量如下：

$$H = \pm J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \quad (1)$$

Ising 本人解析研究了 Ising 模型在一维情形下的性质,并确定其临界转变温度为 0,也就是不存在任何有限温度下的自发磁化强度。20 世纪 40 年代昂萨格对二维正方晶格上的伊辛模型采用解析法得到了严格解,作出了突出的成就。三维 Ising 模型至今没有严格解,但是通过数值模拟和理论分析手段可以较为精确地确定一个有限温度下的相变行为。

1.3.2 XY 模型

在统计场论中,XY 模型是一种 $O(n)$ 模型 ($n=2$), $O(n)$ 模型是统计物理和凝聚态物理领域的基本模型,可用于描述自规避随机行走、量子磁体、冷原子系统等体系的相变行为。同时也可以把 XY 模型看成一种更普遍的伊辛模型。他和伊辛模型的区别在于自旋方向由向上或向下转变为任何方向,Ising 模型的自旋变量只能取为离散值;而 XY 模型的自旋变量则是一个可以在二维平面内旋转的单位矢量,自旋等效为一个相角。XY 模型的哈密顿量如下:

$$H = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j) \quad (2)$$

这里 $\langle i,j \rangle$ 表示最近邻格点上的求和, θ_i 表示格点上矢量自旋的相位,它可以在

$(0, 2\pi)$ 之间连续取值。

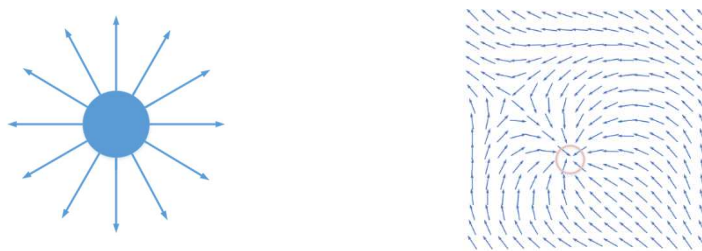


图 2-2 单个 XY 模型自旋取向、二维 XY 模型存在涡旋时的自旋分布图

Fig. 2-2 Spin orientation of a single XY model, spin distribution in the presence of vortices in the two-dimensional XY model

XY 模型在空间二维的情形下具有独特的相变性质。它会随着温度的变化展现出 KTB 相变的特性。不难看出上述公式所定义的自旋系统的能量大小在全局自旋的连续改变下不发生任何变化，也就是说该模型展现出‘ $O(2)$ ’对称性。然而，根据 Memin-Wegner 的理论，这样的对称性在任何有限温度下都不可能自发地破缺，换句话说，不可能像二维模型那样在有限温度下存在一个有序态。有趣的是，二维模型中存在一个罕见的无穷阶相变，即 KTB 相变。当温度很高时，该系统处于无序态。在低温下系统则处于一个准长程有序的态。系统中可能出现涡旋。低温下，系统中只有少量涡旋，并且涡旋和反涡旋会成对。KTB 相变发生时伴随着涡旋和反涡旋的拆对现象。目前，通过高温展开方法计算得到的最精确的相变临界点为 $J_c=1.1200$ 。^[14] 值得注意的是，KTB 相变不仅在 XY 模型以及超导超流系统中产生，同时也可以在两分量离散自旋系统 Ashkin-Teller 模型中得以呈现。

1.3.3 Potts 模型

Potts 模型是由十九世纪五十年代的物理学家 Potts 的导师 Domb 所提出来的一个新的模型，并于 1952 年由 Potts 发表了相关的研究成果。Potts 模型近年来受到了许多科研工作者的关注，它同 XY 模型一样作为 Ising 模型的一种推广，Potts 模型将 Ising 模型中自旋的取向自由度放宽到 q 个，每一个取向自由度代表一个空间取向。表现出了比 Ising 模型更加普遍、丰富、有趣的临界行为和现象。模型的复杂性使问题的求解带来了比 Ising 模型更大的困难；不过，经过多年不懈的研究，最终还是取得了很多有意义的结果。

q 个取向可以定义为角度： $\theta_n = 2\pi n/q, n = 0, 1, \dots, q-1$ ；

哈密顿量为： $H = -\sum_{\langle i,j \rangle} J(\theta_{ij})$ ；

其中， i, j 为相邻的位点， $\theta_{ij} = \theta_{n_i} - \theta_{n_j}$ 。

Potts 模型一般有两种定义方法：平面 Potts 模型和标准 Potts 模型，我们通常选用标准 Potts 模型，其哈密顿量如下：

$$H = -\sum_{\langle i,j \rangle} J \delta_{\sigma_i \sigma_j} \quad (3)$$

其中 σ_i 为格点 i 上的自旋，它可以取 $\sigma_i = 1, 2, \dots, q$ 中的任何值；与 Ising 模型一样， J 是自旋相互作用系数， $J > 0$ 代表铁磁性系统， $J < 0$ 代表反铁磁系统；函数 $\delta_{\sigma_i \sigma_j}$ 在 $\sigma_i = \sigma_j$ 时等于 1，否则为 0； $\langle i, j \rangle$ 代表求和只对最近邻的自旋对进行。在 q 为 2 时，式（3）转变为 Ising 模型的哈密顿量。

1.3.4 双层 XY 耦合模型

众所周知，哈尔伯德型哈密顿量公式如下：

$$\begin{aligned} H = & \sum_i t_a b_{1,i}^+ b_{1,i+\hat{x}} - t_b b_{1,i}^+ b_{1,i+\hat{y}} + H.c. \\ & + t_a b_{2,i}^+ b_{2,i+\hat{y}} - t_b b_{2,i}^+ b_{2,i+\hat{x}} + H.c. \\ & + g b_{1,i}^+ b_{1,i} b_{2,i} b_{2,i} + H.c. + \dots \end{aligned} \quad (4)$$

这里 $b_{1,i}^+$ 和 $b_{2,i}^+$ 是位点 i 上的 p_x 和 p_y 原子相对应的湮灭算子。上式包括从产生的密度 - 密度排斥相互作用原子之间的 s 波散射。在式(4)中， $t_a > t_b > 0$ ，如果玻色子算子被转子算子通过 $b_{x,i} \sim (-1)^{i_x} \exp(-i\theta_x)$ 和 $b_{y,i} \sim (-1)^{i_y} \exp(-i\theta_y)$ 替代，哈密顿量变为各向异性二维耦合 XY 模型^[15]。

$$\begin{aligned} H_c = & \sum_i -J_a \cos(\nabla_x \theta_{i,1}) - J_b \cos(\nabla_y \theta_{i,1}) \\ & - J_b \cos(\nabla_x \theta_{i,2}) - J_a \cos(\nabla_y \theta_{i,2}) + \gamma \cos(2\theta_{i,1} - 2\theta_{i,2}) \end{aligned} \quad (5)$$

1.3.5 Ising-XY 耦合模型

近年来，具有连续 $U(1)$ 和离散 Z_2 对称性的统计力学模型引起了广泛关注，因为它们可以显示相变，而且可以产生有趣且不寻常的临界行为。一般来说，随着温度或其他一些参数的变化，可能会发生两个连续的相变，在强耦合激发的情况下，也可能发生一个相变。大多数具备这些特征的简单模型由耦合 XY 模型和伊辛模型组成，即 Ising-XY 模型，其哈密顿量 Eq (1)

$$\frac{H}{kT} = -\sum_{\langle i,j \rangle} [(A + B\sigma_i \sigma_j) \cos(\theta_i - \theta_j) + C\sigma_i \sigma_j] \quad (6)$$

上式中， i 代表当前格点， j 代表最近邻格点； θ_i (θ_j) 代表二维 XY 模型上第 i (j) 格点的自旋值，取值为 $\theta = [0, 2\pi]$ ； σ_i (σ_j) 代表二维 Ising 模型上第 i (j) 格点的自旋值，取值为 $\sigma = \pm 1$ ； A ， B ， C 为空间耦合参数。在这个系统中有两种类型的拓扑缺陷：畴壁和旋涡。

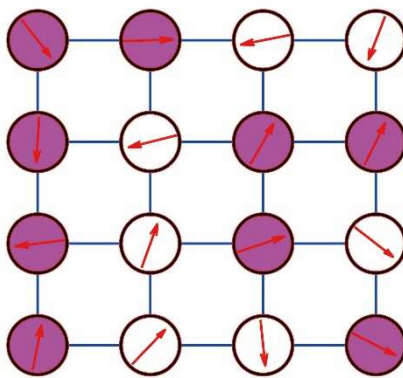


图 2-3 模型的典型配置，箭头表示 θ_i ，颜色代表 σ_i 的值

Fig2-3 A typical configuration of the CIXY model, the arrows denote θ_i the colors represent the values of σ_i

各向同性极限 $A = B$:此时在水平方向和竖直方向上与最近邻自旋相互作用的强度相同，系统此时成为各向同性的 Ising-XY 耦合模型，其哈密顿量形式如下

$$\frac{H}{kT} = -\sum_{\langle i,j \rangle} [A(1 + \sigma_i \sigma_j) \cos(\theta_i - \theta_j) + C \sigma_i \sigma_j] \quad (7)$$

对于 $A > 0$ 和 $C + A/2 > 0$ ，它与 FFXY 系统具有相同的对称破缺模式。因此，它代表了这类模型的另一个例子。许多物理系统可以用 Eq (2) 来描述。例如，在正方晶格上的具有最近邻相互作用 J_1 和次近邻相互作用 J_2 的反铁磁 $J_1 - J_2$ XY 模型，在其基态中最多分成两个正方形子晶格，只要 $2J_2 > J_1$ 每个子晶格就按反铁磁排序。二阶自旋波微扰理论产生的有效相互作用有利于两个子晶格上自旋之间的平行排列，从而打破了经典基态的 $U(1) \times U(1)$ 对称性下降到 $U(1) \times Z_2$ 对称。另一个例子是正方形（或三角形）晶格上的完全受挫 XY 自旋模型。该系统的经典基态破坏了自旋 $U(1)$ 对称性和与自旋围绕每个单位平方旋转有关的手性 Z_2 对称性。

Ising-XY 模型将这两个模型耦合在一起，描述了晶格上的自旋和相位的耦合。在这种模型中，自旋和相位之间存在相互作用，因此它可以用于描述量子磁性系统和量子相变。Ising-XY 模型具有许多有趣的性质，例如在某些条件下可以出现拓扑序和拓扑激发。这些性质使得 Ising-XY 模型在凝聚态物理、量子信息等领域具有广泛的应用。

第二章 蒙特卡罗方法

蒙特卡罗方法（Monte Carlo method），也称统计模拟方法，是 1940 年代中期由于科学技术的发展和电子计算机的发明，由美国的冯·诺伊曼提出的一种以概率统计理论为指导的数值计算方法。蒙特卡罗方法通过抽样调查法得到统计值来推定未知特性量，使用随机数（或更常见的伪随机数）来解决很多计算问题的方法，因此适用于对离散

系统进行计算仿真试验^[1]。蒙特卡罗方法在计算物理学^[3]（如统计物理计算、粒子输运计算^[4]、量子热力学计算、空气动力学计算）、金融工程学、宏观经济学、生物医学以及机器学习等领域都有着广泛的应用。

蒙特卡罗方法两类：一类是所求解的问题本身具有内在的随机性，借助计算机的运算能力可以直接模拟这种随机的过程。例如在核物理的研究中，用其分析中子在反应堆中的传输过程。中子与原子核的作用受到量子力学规律的制约，人们只能知道它们相互作用发生的概率，却无法准确获得中子与原子核作用时的位置以及裂变产生的新中子的行进速率和方向。科学家依据其概率进行随机抽样得到裂变位置、速度和方向，这样模拟大量中子的行为后，经过统计就能获得中子传输的范围，作为反应堆设计的依据。另一种类型是所求解问题可以转化为某种随机分布的特征数，比如随机事件出现的概率，或者随机变量的期望值。通过随机抽样的方法，以随机事件出现的频率估计其概率，或者以抽样的数字特征估算随机变量的数字特征，并将其作为问题的解，这种方法常用于求解复杂的多维积分问题^[5]。

蒙特卡罗方法的经典案例之一是用撒点法求圆的面积，假设在不知道直径时我们需要计算一个圆的面积，蒙特卡罗方法则基于这样的思想：假想你有一袋豆子，把豆子均匀地朝这个图形上撒，然后数这个图形之中有多少颗豆子，这个豆子的数目就是图形的面积。当你的豆子越小，撒的越多的时候，结果就越精确。借助计算机程序可以生成大量均匀分布坐标点，然后统计出图形内的点数，通过它们占总点数的比例和坐标点生成范围的面积就可以求出图形面积。这种方法也可以求其他不规则图形，我们知道图形的不规则程度和分析计算的复杂程度是成正比的，使用蒙特卡罗方法便可以避免复杂的分析计算过程，大大提高我们的效率。

$$S_0 = \frac{N_0}{N} \times S \quad (1)$$

上式中， N_0 代表圆内点数， N 代表总点数， S_0 代表圆面积， S 总面积。

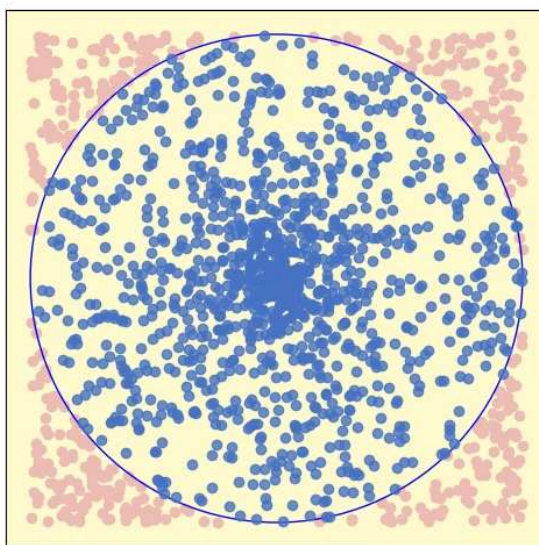


图 1-1 用撒点法求圆的面积：红色点为圆内的点，蓝色点为圆外的点

Fig.1-1 Using the scattering method to find the area of a circle: red dots are points inside the circle, blue dots are points outside the circle

使用蒙特卡罗方法进行分子模拟计算是按照如下步骤进行的：

- (1) 使用随机数生成器产生一个随机的分子构型；
- (2) 对此分子构型的其中粒子坐标做无规则的改变，产生一个新的分子构型；
- (3) 计算新的分子构型的能量；
- (4) 比较新的分子构型与改变前的分子构型的能量变化，判断是否接受该构型；

若这个随机数大于所计算出的玻尔兹曼因子，则放弃这个构型，重新计算；

若这个随机数小于所计算出的玻尔兹曼因子，则接受这个构型，使用这个构型重复再做下一次迭代；

若新的分子构型能量低于原分子构型的能量，则接受新的构型，使用这个构型重复再做下一次迭代；

若新的分子构型能量高于原分子构型的能量，则计算玻尔兹曼因子，并产生一个随机数；

- (5) 如此进行迭代计算，直至最后搜索出低于所给能量条件的分子构型结束。

2.1 蒙特卡罗方法在统计物理领域的应用

随着当代计算机技术的发展，蒙特卡罗方法已经成为人们解决复杂统计模型和高维问题的主要工具。特别是经过最近几十年，蒙特卡罗极大地拓展了人们的科学视野，无论是早期计算物理学的发展，还是现代的计算生物学，都帮助人们研究许多极其复

针对不同的算法选择合适的抽样方法能大大提高计算效率，为我们节省大量的时间和成本，蒙特卡罗中最常用的抽样方法有三种，分别是直接采样、接受-拒绝采样、重要性采样。**直接采样依赖于累积分布函数**；接受拒绝采样虽然完美的解决了累积分布函数不可求时的采样问题，但是非常依赖于提议分布的选择，如果提议分布选择的不好，可能采样时间很长却获得很少满足分布的粒子；而重要性采样的提出就解决了这一难题。

$$\langle Q \rangle = \frac{\sum_{\mu} Q_{\mu} e^{-\beta E_{\mu}}}{\sum_{\mu} e^{-\beta E_{\mu}}} \quad (2)$$
$$\langle Q \rangle = \frac{\sum_{i=1}^M P_{\mu_i}^{-1} Q_{\mu_i} e^{-\beta E_{\mu_i}}}{\sum_{i=1}^M P_{\mu_i}^{-1} e^{-\beta E_{\mu_i}}} \quad (5)$$
$$\langle Q \rangle = \frac{\sum_{i=1}^M e^{\beta E_{\mu_i}} Q_{\mu_i} e^{-\beta E_{\mu_i}}}{\sum_{i=1}^M e^{\beta E_{\mu_i}} e^{-\beta E_{\mu_i}}} = \frac{\sum_{i=1}^M Q_{\mu_i}}{M} \quad (6)$$

The graph consists of 9 nodes, μ_1 through μ_9 , arranged in a circular pattern. Directed edges are labeled with probabilities: P_{12} , P_{13} , P_{14} , P_{15} , P_{21} , P_{31} , P_{32} , P_{41} , P_{51} , P_{61} , P_{71} , P_{81} , P_{91} . The graph shows a complex network of dependencies between the nodes.

图 1-2 马尔科夫链 (Markov) 过程示意图

Fig.1-2 Schematic diagram of Markov chain (Markov) process

Markov 过程的每一步都是一次状态转移，如上图所示，在一步中可能转移为 μ_1 , μ_2 , μ_3 , μ_4 , μ_5 ，转移概率分别为： $P_1 \rightarrow 2$, $P_1 \rightarrow 3$, $P_1 \rightarrow 4$, $P_1 \rightarrow 5$ ，对一个“好的” Markov 过程，无论初始状态是什么，只要经历足够多的步数，最终都会按固定的概率输出末状态，这个末状态叫稳态。

一个“好的” Markov 过程最好应满足以下两个条件：

- (1) 遍历性(Ergodicity);
- (2) 细致平衡(Detailed Balance)。

遍历性 (Ergodicity) 要求无论从哪个状态出发，都能找到一条概率不为 0 的路径到达任意的另一个状态。

细致平衡 (Detailed Balance) 要求每一步中由状态 μ 转移为 ν 的概率流 $J_{u \rightarrow v}$ 等于由 ν 转移为 u 概率流 $J_{v \rightarrow u}$ 。概率流 $J_{u \rightarrow v}$ 定义为 u 的稳态概率乘以由 u 转移为 ν 的概率，即

$$J_{u \rightarrow v} = P_u P_{u \rightarrow v} = P_v P_{v \rightarrow u} = J_{v \rightarrow u} \quad (7)$$

可以把 Markov 过程理解为一套水管，每个状态就是一个结点。当 Markov 过程达到稳态时，由结点 u 流向 ν 的流量应等于由结点 ν 流向 u 的流量，这就是细致平衡。

2.3 常用的蒙特卡罗方法

2.3.1 Metropolis 算法

马尔可夫链蒙特卡洛 (Markov Chain Monte Carlo, 简称 MCMC) 是一类从复杂概率分布中进行采样的方法。Metropolis^[10]算法是严格意义上第一个 MCMC 算法，它于 1953 年由物理学家 Metropolis 等人正式发表，本小节将介绍 MCMC 方法的先驱—Metropolis 算法。

Metropolis 算法是 MCMC 的核心，MCMC 的基本思想是构造一个遍历的马尔科夫链，使得其不变分布成为人们所需要的抽样分布^[1]。做到这点似乎相当复杂，但实际上由于人们可以非常灵活地选择简单的转移概率，所以构造该算法并不困难。

Metropolis 算法的动机是接受-拒绝抽样方法。在接受-拒绝抽样方法中，有布标概

率密度函数 f ，建议概率密度函数 g 和一个接受准则 $h(x,y)$ 。同样地，假设 f 是全空间 Ω 上的目标概率密度函数，人们需要在 Ω 上产生样本马尔科夫链 x_1, x_2, \dots, x_t ，使得它的稳定分布恰好是这个目标概率密度函数 f 的分布。类似于接受-拒绝方法，在当前状态 x_t 下产生下一个状态分两个步骤进行：

(1) 从建议概率密度函数 $g(\cdot|x_t)$ 产生一个随机数 y 作为建议的下一个状态（注意，这是依赖当前状态的条件概率）；

(2) 然后按均匀分布生成一个随机数 r ，如果 $r < h(x,y)$ 则接受该建议的随机数，即 $x_{t+1} = y$ ，否则放弃 y 而采用原状态 $x_{t+1} = x_t$ 。重复这个过程产生随机序列。很显然，这个过程所产生的随机序列做到了下一随机数仅依赖于当前的随机数，而和以前产生的随机数无关。

显然，在算法中建议的概率密度函数 g 主要目的是用来产生状态转移，即为每个状态构造出一个邻域，并使邻域中的某个邻居倍选中。要使得它产生的序列是马尔科夫链，这要求建议概率密度函数 g 必须在整个 Ω 上都有定义，这通常是容易做到的。接受-拒绝抽样方法是最简单的情况：对所有的 $x, g(\cdot|x) = g(\cdot)$ ，单 MCMC 推广了它，使它变成了一个条件概率密度函数。当然必须要求在 x 的某个邻域 N_x 里有 $g(y|x) > 0$ ，其中任意 $y \in N_x$ 这样才能够使状态产生遍历性的转移。一种特殊的 Metropolis 算法需要 g 满足对称性，即满足条件 $g(y|x) = g(x|y)$ ，这种对称性在大多数情况下可以很自然地做到，当然它也可以稍微减弱成下面的条件 $g(y|x) > 0 \Leftrightarrow g(x|y) > 0$ ，即要求状态转移是可逆的。

算法接下来要做的是，将选中的状态和当前状态进行比较，以一定的概率接受其中之一为随机序列链的下一状态。这里与接受-拒绝方法的另一不同之处在于，接受概率不但取决于下一状态 y ，而且还与当前状态 x 有关。具体的方法是所谓的 Metropolis-Hastings 算法，该算法的接受概率定义为

$$h(x, y) = \min\left(1, \frac{f(y)g(x|y)}{f(x)g(y|x)}\right) \quad (8)$$

当对称性条件满足时，上式简化成

$$h(x, y) = \min\left(1, \frac{f(y)}{f(x)}\right) \quad (9)$$

就是说，如果建议的下一状态 y 具有比当前状态 x 的概率较大，即 $f(y) \geq f(x)$ 则肯定接受它；否则，接受概率是 $\frac{f(y)}{f(x)}$ 。这里需要指出：MCMC 算法只需知道 f 的相对值，即只要给出一个正比于 f 的函数即可，这种方便也是 MCMC 的优势之一，因为有些应用问题难以将 f 归一化。虽然这个 Metropolis-Hastings 算法看起来简单，但它却非常有用，连同所有的改进算法一起，它们现已在许多学科领域里广泛应用。

下面给出 Metropolis-Hastings 算法的伪代码：

Algorithm 1

```

1 Select  $x_0$  // 选择一个初始状态，一般选择  $f$  大的状态
2   for  $t = 1$  to  $N$  do // 重复迭代  $N$  次
3      $y \sim g(\cdot | x_{t-1})$  // 生成服从概率密度  $g$  的随机数
4      $h(x_{t-1}, y) \leftarrow \min \{1, \frac{f(y)g(x_{t-1}|y)}{f(x_{t-1})g(y|x_{t-1})}\}$  // 计算接受概率
5     If  $r \sim U(0,1) \leq h(x_{t-1}, y)$  then
6        $x_t \leftarrow y$  // 接受新状态
7     else
8        $x_t \leftarrow x_{t-1}$  // 状态不变
9     End if
10  End for
11  Draw histogram of  $\{x_0, x_1, \dots, x_N\}$  输出抽样序列 //
```

如果建议概率密度函数 g 满足对称性条件，则概率函数 h 简化为 $h(x, y) = \min(1, \frac{f(y)}{f(x)})$ ，此时算法被称为对称 Metropolis-Hastings 算法。细致平衡在 MCMC 中是非常重要的，不满足细致平衡条件一般无法得到正确的平衡态结果。

Metropolis 算法细致平衡推导：

如果 Metropolis 算法达到稳态，也就是状态转移矩阵它的任意两个状态是连通的，所以我们所设计的算法要符合细致平衡原则，这样我们的算法才能在 MCMC 步数下达到稳态。

细致平衡原则：当一个系统达到平衡态时，我们有 $P(A)$ 表示系统在平衡态下处于 A 态的概率 $P(A \rightarrow B)$ 表示系统从微观态 A 变为微观态 B 的可能性

$$P(A)P(A \rightarrow B) = P(B)P(B \rightarrow A) \quad (10)$$

Metropolis 算法需要构造一个合理的接受率来满足细致平衡原则，为了得到这个接受率，我们需要将状态转移矩阵 $P(A \rightarrow B)$ 应当写为：

$$P(A \rightarrow B) = Q(A \rightarrow B)\alpha(A \rightarrow B) \quad (11)$$

则细致平衡公式 $P(A)P(A \rightarrow B) = P(B)P(B \rightarrow A)$ 就改写成为：

$$P(A)Q(A \rightarrow B)\alpha(A \rightarrow B) = P(B)Q(B \rightarrow A)\alpha(B \rightarrow A) \quad (12)$$

$Q(A \rightarrow B)$ 表示由状态 A 开始，选取状态 B 的概率。由于一步只能翻转一个分子，所以只有在 A, B 的差别不大于一个分子的时候， $Q(A \rightarrow B)$ 才不为 0，由于选取分子是随机的，所以：

$$Q(A \rightarrow B) = 1/N \quad (N \text{ 为分子总数}) \quad (13)$$

同理 $Q(B \rightarrow A) = \frac{1}{N}$ 所以 metropolis 的接受率为：

$$\alpha(A \rightarrow B) = \min\{1, \frac{P(B)}{P(A)}\} \quad (14)$$

由 $P(A) = e^{-\beta E}$ 可知, 接受率可以写为 $\alpha(A \rightarrow B) = \min\{1, e^{-\beta \Delta E}\}$

2.3.2 Wolff 算法

一般的算法, 比如 Metropolis, 每次随机选择一个自旋来考虑是否翻转, 这样在远离相变点的时候, 效果比较理想, 但是在相变点附近, Metropolis 效果往往难以达到预期。简单概括就是: Metropolis 算法的局限性在于一次翻转的分子太少, 翻转多次时难以使两次的样本达到统计独立, 而 Wolff 算法是一种高效的集团翻转的方式^[11]。Wolff 算法可以看做 Metropolis 算法的便捷形式, 常用于 Ising 模型和 Potts 模型, 它的具体步骤是:

1. 在系统中随机抽取一个分子作为种子
2. $p_{add}(1 - e^{-2\beta J})$ 的概率依次决定是否添加种子周围的同向分子到 cluster 中 (即 cluster 中的分子要么全为 1 要么全为 -1)。
3. 将原来的种子标记为 operated; 将新加入的分子作为新的种子。
4. 重复步骤 2, 3, 直到 cluster 中所有的分子都被标记为 operated, 即不再有种子。
5. 以 $\alpha(A \rightarrow B)$ 的概率决定是否翻转 cluster。

Wolff 算法细致平衡的细节推导:

Wolff 算法与 metropolis 算法相比, 不同的是选取概率 $Q(A \rightarrow B)$ 不再是 $1/N$, 不过在已经得到的细致平衡公式中我们更关心的是 $\frac{Q(A \rightarrow B)}{Q(B \rightarrow A)}$ 。

一个 cluster 的停止, 取决于这个 cluster 相邻的同向分子的数目。所以在 $A \rightarrow B$ 的过程中, 存在着一个中间 cluster 的状态

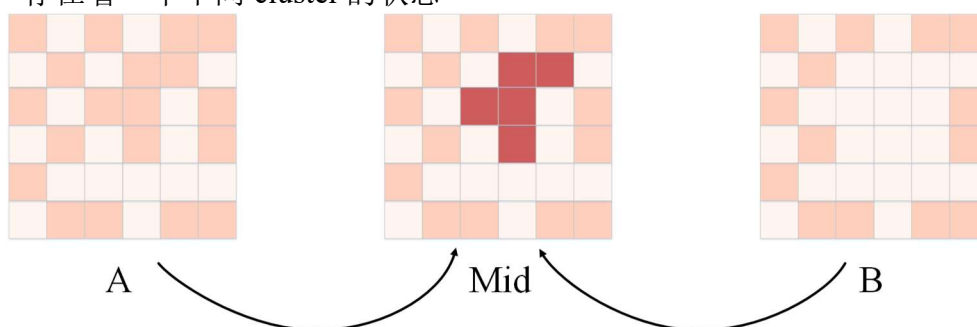


图 1-3 Wolff 算法翻转过程示意图

Figure 1-3 Schematic diagram of the Wolff algorithm flipping process

如果 $A \rightarrow mid$ 需要断 m 个键, $B \rightarrow mid$ 需要断 n 个键那么:

$$\frac{Q(A \rightarrow B) \cdot (1 - P_{add})^m}{Q(B \rightarrow A) \cdot (1 - P_{add})^n} = (1 - P_{add})^{m-n} \quad (15)$$

由细致平衡公式: $P(A)P(A \rightarrow B) = P(B)P(B \rightarrow A)$ 可得:

$$\frac{P(A \rightarrow B)}{P(B \rightarrow A)} = \frac{Q(A \rightarrow B)\alpha(A \rightarrow B)}{Q(B \rightarrow A)\alpha(B \rightarrow A)} = (1 - P_{add})^{m-n} \frac{\alpha(A \rightarrow B)}{\alpha(B \rightarrow A)} = e^{-\beta(E_B - E_A)} \quad (16)$$

又因为 $E_B - E_A = 2J(m - n)$ ，于是我们得到：

$$\frac{(1 - P_{add})^{m-n}\alpha(A \rightarrow B)}{\alpha(B \rightarrow A)} = e^{-2\beta J(m-n)} \quad (17)$$

最后简化得到：

$$\frac{\alpha(A \rightarrow B)}{\alpha(B \rightarrow A)} = e^{2\beta J}(1 - P_{add})^{n-m} \quad (18)$$

当 $P_{add} = 1 - e^{-2\beta}$ 时：

$$\frac{\alpha(A \rightarrow B)}{\alpha(B \rightarrow A)} = 1 \quad (19)$$

由此可知概率为该值时，只要构造出 cluster，就一定可翻转。

$$P_\beta(\omega) = \frac{1}{Z_\beta} e^{-\beta H(\omega)} \quad (20)$$

2.3.3 王-朗道算法

王-朗道算法是一种计算机模拟方法，于 2001 年由王建民和陆朝阳提出，用于计算统计物理模型中的自由能或密度状态。该算法通过调整势能函数的权重来实现对模型状态的平衡采样，还可以用于计算相变和研究复杂多体系统的能级结构。

在统计物理的正则分布中，大家知道微观态的密度函数是 $P_\beta(\omega) = 1/Z_\beta * e^{-\beta H(\omega)}$ ，我们回忆一下在 Ising 模型的模拟中用到过的 Metropolis 算法^[12]，在改变当前系统的位形时，需要考虑接受率 $\min\{e^{-\beta \Delta H}, 1\}$ ，当系统从能量较低的区域想要跳跃到能量较高的区域时将会很容易被拒绝位形变换，而往往系统不同的低能区之间有一些高能的区域隔开，这一区域的接受率比较低，因此，这种稳态的束缚将大大限制 Monte-Carlo 算法对位形空间的遍历和概率权重估计的效率。

当考虑能量（或者其他物理量）时，就会因为能级简并出现一个权重因子， $P_\beta(\varepsilon) = \frac{1}{Z_\beta} e^{-\beta \varepsilon} \int_{\{\omega \in \Gamma: H(\omega) = \varepsilon\}} d\omega = \frac{1}{Z_\beta} g(\varepsilon) e^{-\beta \varepsilon}$ ，因为 Wang 和 Landau 的论文里使用不同能级之间的转移来定算法的，所以概率密度用的是上面的形式，也就要出现一个权重因子 $g(\varepsilon) = \int_{\{\omega \in \Gamma: H(\omega) = \varepsilon\}} d\omega = e^{S(\varepsilon)}$ ，实际上就是微观态的数量，也就是熵的函数，按照正常的 Metropolis 算法，并且不考虑温度的影响时，接受率应该是 $\alpha(\varepsilon_1 \rightarrow \varepsilon_2) = \min\{\frac{g(\varepsilon_2)}{g(\varepsilon_1)}, 1\}$ ，同样也无法避免从态密度很高的区域到态密度很低的区域将导致的拒绝率过高的问题。此时，Wang、Landau 想到的引导采样系统进入这些低态密度区的办法就是把这个接受率倒过来，变成 $\alpha(\varepsilon_1 \rightarrow \varepsilon_2) = \min\{\frac{g(\varepsilon_1)}{g(\varepsilon_2)}, 1\}$ ，这样，当目前所处的状态是高态密度区时，下一步将会以比较高的接受率保持在高态密度区或者进入低态密度区。通过这样来使得采样系统尽可能遍历所有位形足够多次，提高 MC 算法的效率和精度。

怎么实现 Wang-Landau 算法？

假设用到的提议分布是对称的 $P(x,y)$ ，这样可以保证接受率是上面的形式：

1. 首先对能谱进行剖分，分为 N 个区域，那么间隔 $\Delta = \frac{E_{\max}-E_{\min}}{N}$ ，区间 $E_i = [E_{\min} + (i-1)\Delta, E_{\min} + i\Delta)$, $i = 1, 2, \dots, N-1$, $E_N = [E_{\max} - \Delta, E_{\max}]$ ，映射 I 把能量映射到对应的区间下标

2. 初始化

(1) 由于没有先验信息，将各个区域的初始态密度函数设置成 1，为了避免溢出，取对数用熵来计算， $S_i = 0, i = 1, 2, \dots, N$

(2) 步长 $f=1$

(3) 初始化位形 $\omega_0 \in$

3. while($f > \epsilon$ && $t < \max_iteration$) 进入迭代

(1) 提议新的位形 $\omega_{t+1} \sim P(\omega_t, \cdot)$

(2) 计算接受率 $\alpha = \min\{\exp(S_{I \circ H(\omega_t)} - S_{I \circ H(\omega_{t+1})}), 1\}$

(3) 是否接受 if($\text{rand} < \alpha$) ~ accept;

(4) 修改当前位形和数据, $\text{Hist}_{I \circ H(\omega_{t+1})}++; \sim S_{I \circ H(\omega_{t+1})} += f; \sim t++; \sim \omega_t = \omega_{t+1};$

(5) 若直方图平坦，则修改步长 则修改步长 if($\text{Hist.is_flat}()$) $\text{Hist}=0; f*=0.5;$

2.3.4 Swendsen-Wang 算法

Swendsen-Wang 算法是一种用于计算统计物理模型中的自由能和相关物理量的计算机模拟方法。Swendsen-Wang 算法^[13]最初用于伊辛模型与 Potts 模型，后来被推广到其他模型之中。该算法的关键是按照 Fortuin 与 Kasteleyn 的理论将 Potts 模型变换为逾渗模型，相邻自旋间按概率成键。之后再通过霍森-科佩尔曼算法标识联键的集团，并将每个集团内的所有自旋赋以相同的随机值。该算法通过将相邻的同一自旋聚类成簇，然后将这些簇随机翻转，以实现模型状态的平衡采样。由于该算法可以一次改变整个集团的自旋，因而在临界点附近能够显著提高效率，以解决临界慢化问题。

Swendsen-Wang 算法最初由 Robert Swendsen 和 Joon Ha Kim 在 1987 年提出，并且被广泛应用于计算相变点和研究磁性系统的性质。该算法在计算效率和准确性方面的表现非常优秀，因此在统计物理学和计算物理学领域受到了广泛的关注和应用。

Swendsen-Wang 算法的基本流程如下：

(1) 初始化系统的自旋状态，即给定每个自旋的初始方向。

(2) 将所有相邻的同一自旋聚类成簇。可以使用深度优先搜索算法 (DFS) 或广度优先搜索算法 (BFS) 来实现这一步骤。具体而言，以某个自旋为起点，使用 DFS 或 BFS 遍历所有与之相邻的自旋，并将它们标记为同一簇。如果两个相邻的自旋具有不同的自旋方向，则它们不属于同一簇。

(3) 对每个簇进行随机翻转, 即随机选择一个自旋方向, 将簇中所有自旋的方向都设为该方向。这相当于将整个簇视为一个单元, 然后随机地将它们翻转。

(4) 重复第 2 和第 3 步, 直到系统达到平衡态。可以使用 Metropolis-Hastings 算法等随机漫步算法来实现对系统状态的采样。

(5) 计算系统的物理量, 如自由能、磁化强度等。

2.4 本论文的研究内容

本文主要的研究内容分为三个章节, 分别是第三章、第四章、第五章。第三章通过 wolff 算法对空间各向异性下双层耦合 XY 模型进行了详细的探究, 为后人研究 XY 耦合模型或者关于 XY 模型的衍生模型提供了更多的参数和数据支撑。第四章研究了更为广泛的 Ising-XY 耦合模型, 通过蒙特卡罗计算得到了具体的相图验证了前人的理论推导, 对目前关于 Ising-XY 耦合模型的相图以及相变情况的研究提供了坚实有力的支撑。第五章采用图神经网络研究 Ising-XY 耦合模型, 有效地模拟从传统蒙特卡罗方法中生成的构型, 并且生成全新的具有相同物理特性的构型, 相比较与之前的机器学习方法效率大大提升。

第 3 章 蒙特卡罗方法研究空间各向异性双层 XY 耦合模型的相变

3.1 引言

相变是自然界广泛存在的一类突变现象，也是研究热力学与统计物理的问题，而 Ising 模型和 XY 模型正好是研究该类问题的重要载体。Ising 模型是经典统计力学的一个重要模型，由德国物理学家威廉·楞次在 1920 年提出以描述铁磁性物质的内部的原子自旋状态及其与宏观磁矩的关系。伊辛模型可以近似描述单轴各向异性铁磁体的铁磁-顺磁相变，稍加改变还可以描述气-液相变、合金的有序无序相变等情形。20 世纪 30-40 年代，物理学家们使用平均场近似理论对二维伊辛点阵模型进行了研究。而 XY 模型是一个更加普遍化的 Ising 模型，其自旋的朝向由简单的上和下解放到了 XY 平面的任意方向。

XY 模型是一个重要的统计物理模型，它可以用于描述许多不同的物理系统，如超导体、自旋玻璃和凝聚态物质中的磁性自旋系统。而 XY 耦合模型更是由于其丰富有趣的物理现象引起了广大学者的关注与研究，为我们留下了众多的理论和数值计算作为支撑。其中较为经典的有三维强耦合 XY 模型的蒙特卡罗模拟，在强耦合下发现存在三个不同的相并且有三临界点的存在。

在本章中，采取马尔科夫链蒙特卡罗中的 Metropolis 算法和 Wolff 算法研究正方晶格下不同尺寸的二维耦合 XY 模型。通过测量能量、磁化强度、比热、宾得累积量、超流密度，探究了各向同性极限、各向异性极限、各向异性等不同情况下的相变情况并且作出两层 XY 自旋之间无耦合时 XY 模型的相图。本章研究表明，在 J_a 、 J_b 取值不同时有着相同的物理现象，固定 J_a 值，任取 $J_b \in [0,1]$ ，空间各向异性情况下 Ising 型相变与 XY 型相变同时发生，不同 J_a 值 J_b 相变点不同。相图呈双曲线趋势变化。

3.2 模型与算法

3.2.1 模型

众所周知，哈尔伯德型哈密顿量公式如下：

$$\begin{aligned} H = & \sum_i t_a b_{1,i}^+ b_{1,i+\hat{x}} - t_b b_{1,i}^+ b_{1,i+\hat{y}} + H.c. \\ & + t_a b_{2,i}^+ b_{2,i+\hat{y}} - t_b b_{2,i}^+ b_{2,i+\hat{x}} + H.c. \\ & + g b_{1,i}^+ b_{1,i} b_{2,i} b_{2,i} + H.c. + \dots \end{aligned} \quad (1)$$

这里 $b_{1,i}^+$ 和 $b_{2,i}^+$ 是 p_x 和 p_y 的湮没算子位点 i 上的原子。在这个哈密顿量式 (4) 中 $t_a > t_b > 0$ 。如果玻色子算子被转子算子通过 $b_{x,i} \sim (-1)^{i_x} \exp(-i\theta_x)$ 和 $b_{y,i} \sim (-1)^{i_y} \exp(-i\theta_y)$ 替代，哈密顿量变为各向异性二维耦合 XY 模型。由于其基态同时表现出连续和离散简并性，故而引起了人们的注意。

$$H_c = \sum_i -J_a \cos(\nabla_x \theta_{i,1}) - J_b \cos(\nabla_y \theta_{i,1}) - J_b \cos(\nabla_x \theta_{i,2}) - J_a \cos(\nabla_y \theta_{i,2}) + \gamma \cos(2\theta_{i,1} - 2\theta_{i,2}) \quad (2)$$

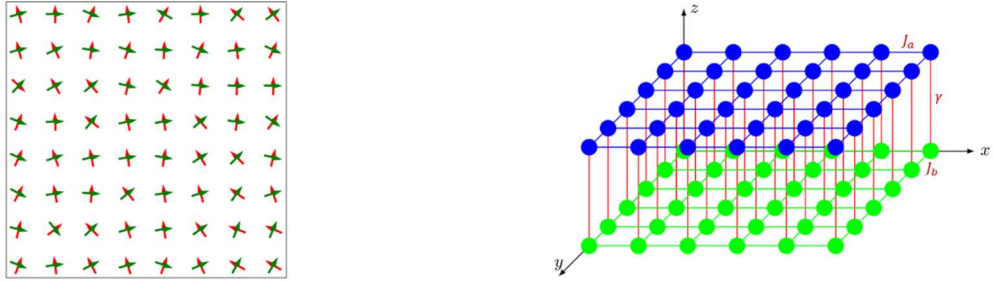


图 3-1 (a)二维耦合 XY 模型平面图；(b)二维耦合 XY 模型空间图

Fig.3-1 (a) Plan view of the 2D coupled XY model; (b) Spatial view of the 2D coupled XY model

本节内容主要研究模型（2）中有限温度下的相变，两个 XY 变量之间的耦合在整个代数阶段总是相关的，这意味着无论 γ 有多弱，在大尺度上 γ 总是重整化的强，因此本章中我们研究时将在 $\gamma=0$ ，且在足够长的尺度上，用 Ising 变量 $\sigma = \pm 1$ 替换 $\theta_2 = \theta_1 + \frac{\sigma\pi}{2}$ 。在 $J_a = J_b$ 的各向同性极限中，系统成为各向同性耦合 Ising-XY 模型；在 $J_b = 0$ 的各向异性极限中，系统成为各向异性极限下 Ising-XY 耦合模型；在 $J_a > J_b > 0$ 的各向异性下，系统成为各向异性 Ising-XY 耦合模型。以上三种情况下关于 Ising-XY 模型的哈密顿量都可简化如下：

$$H_1 = -J \sum_{\langle i,j \rangle} (1 + \sigma_i \sigma_j) \cos(\theta_i - \theta_j) \quad (3)$$

$$H_2 = -J \sum_{\langle i,j \rangle} \cos(\nabla_x \theta_i) - J \sigma_i \sigma_{i+\bar{y}} \cos(\nabla_y \theta_j) \quad (4)$$

$$H_3 = \sum_{\langle i,j \rangle} -J(1 + \beta \sigma_i \sigma_{i+\bar{x}}) \cos(\theta_i - \theta_{i+\bar{x}}) - J(\beta/4 + \sigma_i \sigma_{i+\bar{y}}) \cos(\theta_i - \theta_{i+\bar{y}}) \quad (5)$$

3.2.2 算法

Metropolis 算法：最简单的模拟 Metropolis 算法如下。一个蒙特卡洛迭代步骤给所有的 Ising 和 XY 自旋一次更新的机会，如图 1 所示。当更新 Ising 自旋值时，XY 自旋的相互作用可以被看作是一个伊辛自旋的耦合系数，即 $J_{eff} = \cos(\theta_i - \theta_j)$ 。一个简单的方法去模拟 XY 模型是选择一个轴 α ，让自旋以该轴为对称轴翻转， $\theta_k^* \leftarrow 2\alpha - \theta_k$ 。伪代码如下。

Algorithm 2

Input:

构型 $\{\theta_1, \theta_2, \dots, \theta_N\}, \{S_1, S_2, \dots, S_N\}$

1 | For $i = 1 \rightarrow N$ do

```

2      Update XY spins
3       $k \leftarrow N * rand() + 1, \alpha \leftarrow \pi * rand()$ 
4       $\theta_k^* \leftarrow 2\alpha - \theta_k, \Delta E \leftarrow E(\theta_k^*) - E(\theta_k)$ 
5       $Pacc \leftarrow \min \{1, -\frac{\Delta E}{T, \tau}\}$ 
6      If  $rand() \leq Pacc$  then
7           $\theta_k \leftarrow \theta_k^*$ 
8      End if
9  End for
10     更新伊辛的旋转方式与 2-10 号线类似
11  $S_k \rightarrow -S_k$ 
12 Return
    
```

Wolff 算法：通常情况下，集群更新算法对于纯 XY 模型需要将自旋矢量分解为两个不同的相互垂直的分量，即 $\vec{S} = S_x \vec{i} + S_y \vec{j}$ ，此时 $H_{XY} = H_{xx} + H_{yy}$ ，然后再分别更新 H_{xx} 和 H_{yy} 。与更新 Ising 自旋相同，当更新 XY 自旋值时，Ising 自旋的相互作用可以被看作是一个 XY 自旋的耦合系数。

在大量参考文献中，点 i, j 形成集群的概率为

$$P(k, i) = 1 - \min \{1, e^{\frac{2H_{xx}}{k_B T}}\} \quad (6)$$

这里，我们给出了一个更普遍的表达式

$$P(k, i) = 1 - \min \{1, \exp [\frac{E(\theta_k - \theta_i) - E(\theta_k^* - \theta_i)}{k_B T}]\} \quad (7)$$

其中 $E(\theta_k - \theta_i)$ 是旧的键能，新的键能是 $E(\theta_k^* - \theta_i)$ 。伪代码如下。

Algorithm 3 Wolff 算法的一个 MC 步骤

Input:

构型 $\{\theta_1, \theta_2, \dots, \theta_N\}, \{S_1, S_2, \dots, S_N\}$

```

1      Update XY spins
2       $k \leftarrow N * rand() + 1, \alpha \leftarrow \pi * rand(), id = 1$ 
3       $\theta_k^* \leftarrow 2\alpha - \theta_k, stack[1] \leftarrow k$ 
4      while  $stack$  not null or  $id \geq 1$  do
5           $k \leftarrow stack[id]$ 
6          for  $i$  in all neighboring sites of site  $k$  do
7              Case 1: 点  $i$  没有被访问
8              Case 2:  $rand < P(k, i)$ 
9              If case 1 and case 2 then
10                 Split  $\theta_i \leftarrow 2\alpha - \theta_i$ 
    
```

```

11         stack.push(i)
12         id ← id + 1, stack[id] = i
13     End if
14 End for
15 id ← id - 1
16 End while
17 End for
18 更新伊辛自旋的方式与 2-17 号线类似
19 Return
    
```

3.2.3 物理量

采用 Metropolis 算法和 Wolff 算法对正方晶格上的 Ising-XY 模型进行蒙特卡洛模拟并取周期性边界条件。众所周知，模拟无法直接在热力学极限下(无限大尺寸)进行。因此，我们取有限大小的系统进行模拟，通过系统尺寸的变化和有限尺寸标度，我们可推断出系统在热力学极限下的行为。用 L 表示系统的线性尺寸，对于正方晶格而言，格点总数为 $N=L \times L$ 。接下来将利用上述哈密顿量定义蒙特卡洛模拟中的观测量。

序参量 M 为铁磁系统的关于 Ising 模型的平均磁化强度：

$$M = \sum_i S_i / L^2 \quad (8)$$

比热 C_v 可被定义为：

$$C_v = \frac{1}{k_B T^2} [\langle E^2 \rangle - \langle E \rangle^2] \quad (9)$$

这表明能量的波动，比热的峰值又称为“比热尖峰”，“比热尖峰”的个数代表着相变的次数。

采用宾得累积量 B_r 来定义伊辛型的相变点，宾得累积量可被定义为：

$$B_r = \langle M^2 \rangle^2 / \langle M^4 \rangle \quad (10)$$

采用超流密度 ρ_s 来确定 XY 型相变对应的临界点，超流密度的公式定义如下：

$$\rho_s = \frac{1}{Nd} \frac{\partial^2 F(\phi)}{\partial \phi^2} \quad (11)$$

根据 H_1 、 H_2 、 H_3 和式 (6)，各向同性极限的情况下超流密度可被描述为：

$$\rho_s = \frac{1}{Nd} (\langle H \rangle - (1 + \sigma_i \sigma_j) \beta \sum_{a=1}^d \langle I_a^2 \rangle) \quad (12)$$

其中， $I_x = J \sum_{\langle i,j \rangle_x} \sin(\theta_j - \theta_i)$ 。

同样，根据 H_2 和式 (6)，各向异性的情况下超流密度可被描述为：

$$\rho_s = \rho_{s_x} + \rho_{s_y} \quad (13)$$

$$\rho_{s_x} = \frac{1}{Nd} (\langle H \rangle - \beta \sum_{a=1}^d \langle I_x^2 \rangle) \quad (9)$$

$$\rho_{sy} = \frac{1}{Nd} (\langle H \rangle - \sigma_i \sigma_j \beta \sum_{a=1}^d \langle I_y^2 \rangle) \quad (10)$$

因此推断，各向异性极限的情况下超流密度可以写成如下：

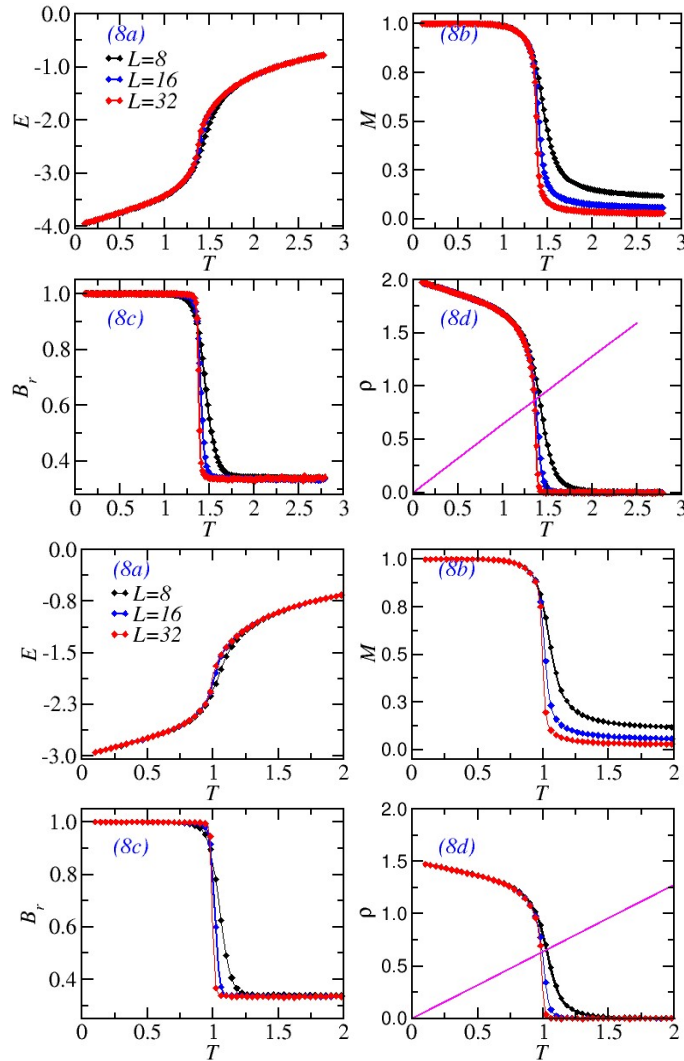
$$\rho_s = \rho_{sx} + \rho_{sy} \quad (11)$$

$$\rho_{sx} = \frac{1}{Nd} (\langle H \rangle - (1 + \beta \sigma_i \sigma_{i+\vec{x}}) \sum_{a=1}^d \langle I_x^2 \rangle) \quad (12)$$

$$\rho_{sy} = \frac{1}{Nd} (\langle H \rangle - (\beta + \sigma_i \sigma_{i+\vec{y}}) \sum_{a=1}^d \langle I_y^2 \rangle) \quad (13)$$

3.3 结果与讨论

3.3.1 局部细节图



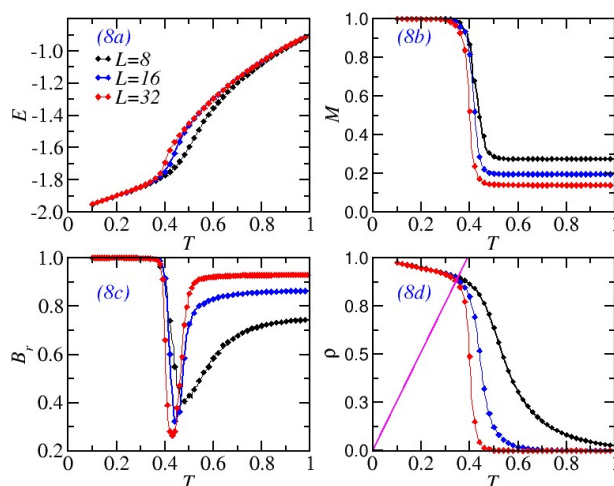


图 3-2 (a) $J_a=J_b$ 时, 双层 XY 耦合模型中能量 E 、磁化强度 M 、宾得累积量 B_r 、超流密度 ρ 随温度 T 的变化情况; (b) $J_b=0.5J_a$ 时, 双层 XY 耦合模型中能量 E 、磁化强度 M 、宾得累积量 B_r 、超流密度 ρ 随温度 T 的变化情况; (c) $J_b=0$ 时, 双层 XY 耦合模型中能量 E 、磁化强度 M 、宾得累积量 B_r 、超流密度 ρ 随温度 T 的变化情况。

Fig. 3-2 (a) Variation of energy E , magnetization intensity M , Binde accumulation B_r , and supercurrent density ρ with temperature T for $J_a=J_b$; (b) Variation of energy E , magnetization intensity M , Binde accumulation B_r , and supercurrent density ρ with temperature T for $J_b=0.5J_a$; (c) Variation of energy E , magnetization intensity M , Binde accumulation B_r , and supercurrent density ρ with temperature T for $J_b=0$ with temperature T .

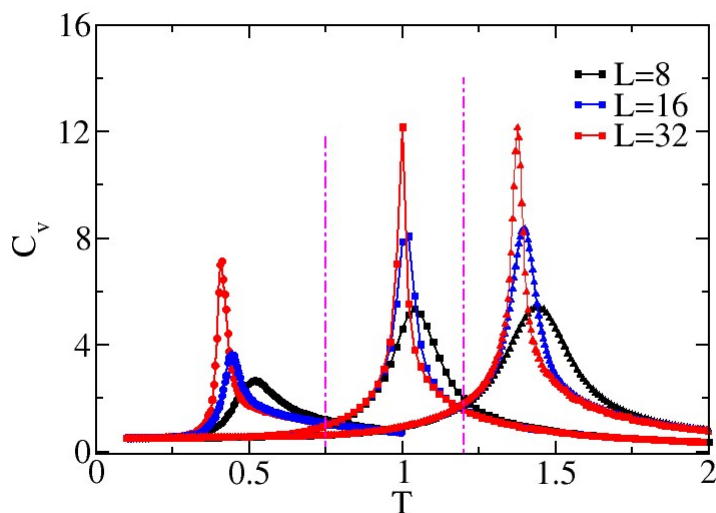


图 3-3 从左到右, 依次是 $J_b = 0$ 、 $J_b = 0.5J_a$ 、 $J_b = J_a$ 三种比率情况下, 双层 XY 耦合模型中比热 C_v 随温度 T 的变化情况

Fig. 3-3 From left to right, the variation of specific heat C_v with temperature T for three ratios of $J_b = 0$, $J_b = 0.5J_a$, $J_b = J_a$.

图 3-2 中描述了不同情况下该模型的物理量变化规律; 其中图 3-2 (a)、图 3-2 (b)、图 3-3 (c) 分别代表 $J_b = 0$ 、 $J_b = 0.5J_a$ 、 $J_b = J_a$ 三种比率情况下, 所观测物

理量随温度 T 的变化情况，其中观测物理量包含能量 E 、磁化强度 m 、宾得累积量 B_T 、超流密度 ρ 四种。从图 3-2 中我们可以观察到各物理量的值在相变点附近发生突变；在同一比率下，物理量的计算值呈现明显的尺寸效应，即随着尺寸增大相变点提前且突变现象更加明显。根据图 3-2 (a)、图 3-2 (b)、图 3-3 (c) 可知在 $J_b = J_a$ 、 $J_b = 0.5J_a$ 、 $J_b = 0$ 三种比率情况下，随温度的增加能量都由最低态 -4eV 、 -3eV 、 -2eV 开始逐步增加直至接近 0eV ，尺寸越大增加越快。同样在这三种情况下随温度的增加，自旋取向由有序态转向为无序态，其磁化强度取值由 1 逐渐降低至 0，在相变点附近尺寸越大磁化强度值降低的越快。一般而言，我们使用宾得累积量在不同尺寸下的交点来确定 Ising 型相变点，由图 3-2 可知在 $J_b = J_a$ 、 $J_b = 0.5J_a$ 、 $J_b = 0$ 三种比率情况下发生 Ising 型相变的温度分别为 1.3623°C 、 0.9825°C 、 0.4298°C ，即随 J_b 所占比率的减小发生 Ising 型相变的温度不断变低。普遍情况下，我们将超流密度与 $\frac{\Delta\rho_s}{T} \sim \frac{2}{\pi}$ 的交点记为相变点，在 $J_b = 0$ 时相变点变为超流密度与 $\frac{\Delta\rho_s}{T} \sim \frac{8}{\pi}$ 的交点；尺寸越大超流密度降低的越快最终减小至 0。图 3-3 从左到右，依次是 $J_b = 0$ 、 $J_b = 0.5J_a$ 、 $J_b = J_a$ 三种比率情况下，比热 C_v 随温度 T 的变化情况，上图只有一个比热尖峰，因此我们推测可能该模型只存在一个相变点。然而由比热尖峰判断的结果是否准确，仍然还需要我们进一步精确推断。已知我们可以用宾得累积量的交点判断 Ising 型相变点的位置，接下来只需要判断 XY 型相变点的位置与 Ising 型相变是否重合便可判断该 Ising-XY 耦合模型相变点的个数。若二者重合则与比热尖峰所传达的信息一致，也就是说 Ising 型相变与 XY 型相变同时发生，即该模型只存在一个相变点；反之，若二者不重合则与比热尖峰所传达的信息相悖，即该模型其实存在两个相变点。

前面我们已经简要说明了判定 xy 型相变点位置的方法，下面我们来详细介绍一下具体方法。如图 3-2 所示，不同的耦合系数对应不同的情况，在 $J_b = J_a$ 以及 J_b 存在但不非常小时，可以采用超流密度与直线 $\Delta\rho_s/Tc \sim 2/\pi$ 的交点确定某一尺寸下相变发生的位置；在 $J_b = 0$ 以及 J_b 存在但非常小时，需要采用超流密度与直线 $\Delta\rho_s/Tc \sim 8/\pi$ 的交点确定某一尺寸下相变发生的位置。

图 3-2 中斜线与曲线的交点对应的横坐标代表在特定耦合常数和尺寸下的相变温度，从上图我们发现超流密度也有着明显的尺寸效应，即随着尺寸增大其值降低越快。除此之外，在 $J_b = J_a$ 时，尺寸 $L=16$ 、 $L=32$ 、 $L=48$ 时相变点分别为 1.368°C 、 1.354°C 、 1.353°C ，说明同一耦合参数下尺寸越大由温度对应的相变点越小。尺寸 $L=32$ 时，当 $J_a = 1$ 时，在 $J_b = J_a$ 、 $J_b = 0.5J_a$ 、 $J_b = 0$ 三种情况下相变点分别为 1.354°C 、 0.9831°C 、 0.3482°C ；可以看出相同尺寸时，相变点随空间各向异性情况的加剧而减小。

尺寸越大得出的相变点越精确，但也可以想象如果我们设置超级大的尺寸时给计算机带来的工作量是巨大的，实际操作起来非常困难，为了简化我们的计算量，可以借助超流密度外推公式由几个小尺寸的相变点外推出尺寸无穷大时该模型对应的相变

点。超流密度尺寸外推公式如下所示：

$$T_c(L) = T_{BKT}(\infty) + \frac{a}{\ln^2(bL)} \quad (14)$$

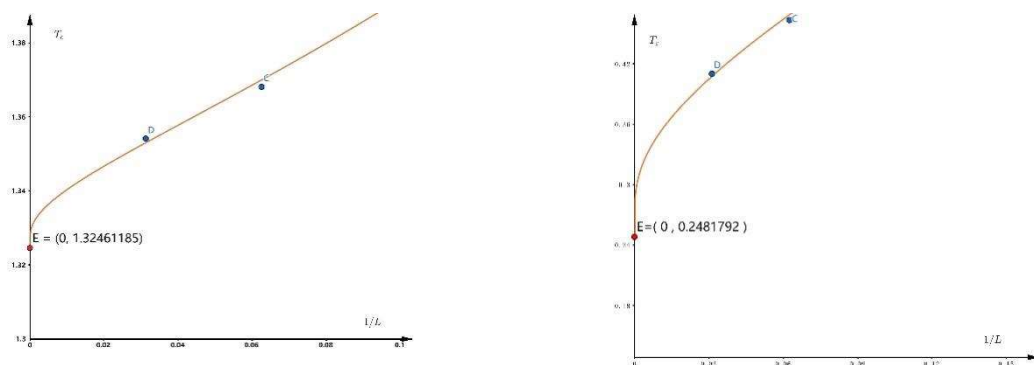


图 3-4 各向同性极限和各向异性极限情况下，由小尺寸相变点外推无限大尺寸的相变点的拟合图

Fig.3-4 Fitted plots of phase transition points of infinite size extrapolated from small size phase transition points in the isotropic and anisotropic limits

由图 3-4 所示，在各向同性极限情况下，该双层 XY 耦合模型在尺寸无限大时的 XY 型相变点是 1.3246°C。同时对比 Ising 型相变在各向同性极限情况下的相变点是 1.3623°C，我们发生二者相差很小，在误差允许的范围内可以看成两种相变同时发生，即该模型只存在一个相变点。结论与只存在一个比热尖峰的现象相符。我们也可以利用式 (13) 拟合其他耦合常数下该模型在尺寸无限大或其他任意尺寸下时的相变点，大大简化我们的计算难度。

3.3.2 相图

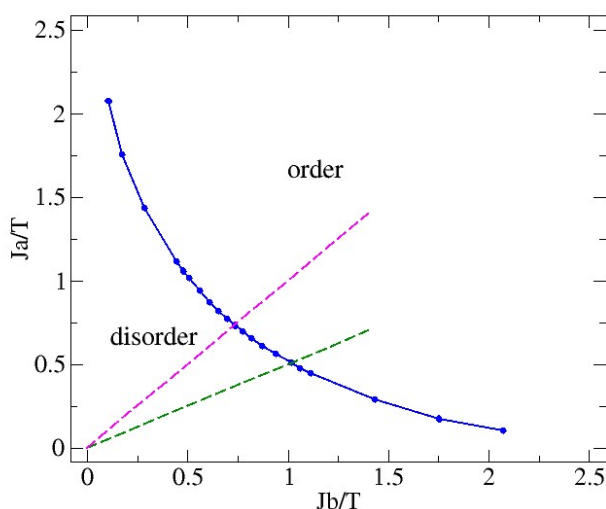


图 3-5 双层 XY 耦合模型的相图

Fig.3-5 Phase diagram of the two-layer XY coupling model

通过不断调整 J_b 和 J_a 间的比率，我们得到了在不同比率下双层 XY 耦合模型的相变点。图 10 中蓝线上的每一点都代表着一个特定比率下发生相变时参数 J_a 和 J_b 与发生相变时的温度之间的关系。在蓝线的右边是双层 XY 模型的低温有序态，在蓝线的左边是双层 XY 模型的高温无序态；品红线的斜率为 1、绿线的斜率为 0.5，它们与相变线的交点坐标代表比率为 1、0.5 时的相变信息。

3.4 本章小节

第 4 章 蒙特卡罗方法研究二维 Ising-XY 耦合模型的相变

4.1 引言

Ising-XY 耦合模型是 Ising 模型和 XY 模型的耦合, 能够更好地描述真实物理系统的性质。研究 Ising-XY 模型的相变行为有助于更深入地理解物质相变机制, 并为材料科学、能源材料、物理化学等领域提供理论支持。此外, Ising-XY 模型的相变研究也是统计物理领域的热点问题之一, 吸引了众多学者的关注。

这些年来, Ising-XY 耦合模型、双层 XY 耦合模型、完全阻挫 XY 模型已经引起了人们的广泛关注。1990 年, Kosterlitz 教授等人利用实空间重整化群和蒙特卡罗的方法研究了混合二维 Ising-XY 耦合模型在各向同性极限下的相图, 测量了一些常见的临界指数, 发现不同情况下有着一级相变和连续相变的区别。2005 年, 伊斯兰自由大学 Rasool Ghanbari 等人研究了三维下耦合 XY 模型的相图。仿真结果表明该模型存在三临界点, 在三临界点两种不同的序同时存在。2007 年, Cenke Xu 教授从理论上介绍了具有空间各向异性和 $U(1) \times Z_2$ 对称性的二维耦合 XY 系统的相变, 系统的相变可分为 Ising 型相变和 XY 型相变, 相变的顺序取决于各向异性耦合的比率并画出了经典相变和量子相变下的相图。遗憾的是, 各位学者都仅仅从理论上给出了 Ising-XY 耦合模型的相图, 没有具体的数据支撑^[4]。我们本次课题研究的目的是在数值计算层面得出详细的数据从而画出 Ising-XY 模型的相图。

本章中采取了马尔科夫链蒙特卡罗中的 Metropolis 算法和 Wolff 算法研究正方晶格下不同尺寸的 Ising-XY 耦合模型, 从数值上验证 Kosterlitz 的结论。通过测量能量、磁化强度、比热、熵得累积量、超流密度, 研究了各向同性极限、各向异性极限、各向异性等不同情况下的相变情况并且作出 Ising-XY 模型的相图。根据现有的研究表明, 在 C 取值不同时有着不同的现象; 在 $C > C^*$ 时, 固定温度不变, 空间各向同性情况下 Ising 模型先于 XY 模型发生相变, XY 模型在 A 为 0.321 时发生相变, Ising 模型在 A 为 0.461 时发生相变。固定 T 不变, 随着 B/A 逐渐减小至 0, XY 相变点与 Ising 相变点逐渐接近并相交, 最后 XY 相变先于 Ising 相变发生与各向异性极限情况重合; (除此之外, 各向异性极限情况下 XY 模型在 A 为 0.941 时发生相变。) 在 $C < C^*$ 时, 固定温度不变, 空间各向同性情况下 Ising 模型与 XY 模型同时发生相变, 耦合模型在 A 为 0.731 时发生相变。固定 T 不变, 随着 B/A 逐渐减小至 0, XY 相变点与 Ising 相变点逐渐分开, 最后 XY 相变先于 Ising 相变发生与各向异性极限情况重合。

4.2 模型及算法

4.2.1 模型

近年来, 具有连续 $U(1)$ 和离散 Z_2 对称性的统计力学模型引起了广泛关注, 因

为它们不仅可以显示相变，而且可以产生有趣且不同寻常的临界行为。一般来说，随着温度或其他一些参数的变化，可能会发生两个连续的相变，在强耦合激发的情况下，也可能发生一个相变。大多数具备这些特征的简单模型由耦合 XY 模型和伊辛模型组成，即 Ising-XY 耦合模型，其哈密顿量 Eq (1)

$$\frac{H}{kT} = -\sum_{\langle i,j \rangle} [(A + B\sigma_i\sigma_j) \cos(\theta_i - \theta_j) + C\sigma_i\sigma_j] \quad (1)$$

上式中， i 代表当前格点， j 代表最近邻格点； θ_i (θ_j) 代表二维 XY 模型上第 i (j) 格点的自旋值，取值范围为 $\theta = [0, 2\pi]$ ； σ_i (σ_j) 代表二维 Ising 模型上第 i (j) 格点的自旋值，取值为 $\sigma = \pm 1$ ； A ， B ， C 为空间耦合系数。在这个系统中有两种类型的拓扑缺陷：畴壁和旋涡。

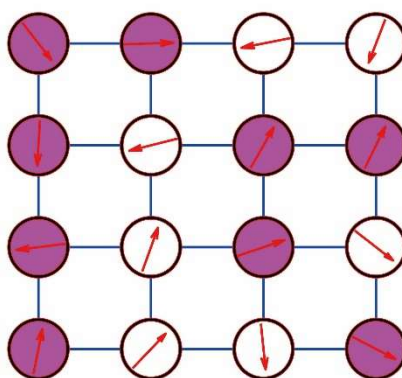


图 4-1 模型的典型配置，箭头表示 θ_i ，颜色代表 σ_i 的值

Fig.4-1 A typical configuration of the CIXY model, the arrows denote θ_i , the colors represent the values of σ_i

当空间耦合参数 $A = B$ 时，系统处于各向同性极限状态：此时在水平方向上和竖直方向上当前格点自旋与最近邻自旋相互作用的强度相同，系统此时成为各向同性的 Ising-XY 耦合模型，其哈密顿量形式如下：

$$\frac{H}{kT} = -\sum_{\langle i,j \rangle} [A(1 + \sigma_i\sigma_j) \cos(\theta_i - \theta_j) + C\sigma_i\sigma_j] \quad (2)$$

当 $A > 0$ 和 $C + A/2 > 0$ 时，该模型与完全受挫 XY 模型具有相同的对称破缺模式；因此，它代表了这类模型的另一个例子。许多物理系统可以用 Eq (2) 来描述。例如，在正方晶格上的具有最近邻相互作用 J_1 和次近邻相互作用 J_2 的反铁磁 $J_1 - J_2$ XY 模型在其基态中最多分成两个正方形子晶格，只要 $2J_2 > J_1$ 每个子晶格就按反铁磁排序。二阶自旋波微扰理论产生的有效相互作用有利于两个子晶格上自旋之间的平行排列，从而打破了经典基态的 $U(1) \times U(1)$ 对称性下降到 $U(1) \times Z_2$ 对称。另一个例子是正方形（或三角形）晶格上的完全受挫 XY 自旋模型。该系统的经典基态破坏了自旋 $U(1)$ 对称性和与自旋围绕每个单位平方旋转有关的手性 Z_2 对称性。

4.2.2 算法

Metropolis 算法：最简单的模拟 Metropolis 算法如下。一个蒙特卡洛迭代步骤给所有的 Ising 和 XY 自旋一次更新的机会，如图 1 所示。当更新 Ising 自旋值时，XY 自旋的相互作用可以被看作是一个伊辛自旋的耦合系数，即 $J_{eff} = \cos(\theta_i - \theta_j)$ 。同样，当更新 XY 自旋值时，Ising 自旋的相互作用可以被看作是一个 XY 自旋的耦合系数，即 $J_{eff} = A(1 + \sigma_i \sigma_j)$ 。一个简单的方法去模拟 XY 模型是选择一个轴 α ，让自旋以该轴为对称轴翻转， $\theta_k^* \leftarrow 2\alpha - \theta_k$ 。伪代码如下。

Algorithm 4

Input:

输入构型 $\{\theta_1, \theta_2, \dots, \theta_N\}, \{s_1, s_2, \dots, s_N\}$

```

1   For  $i = 1 \rightarrow N$  do
2       Update XY spins
3        $k \leftarrow N * rand() + 1, \alpha \leftarrow \pi * rand()$ 
4        $\theta_k^* \leftarrow 2\alpha - \theta_k, \Delta E \leftarrow E(\theta_k^*) - E(\theta_k)$ 
5        $Pacc \leftarrow \min \{1, \exp(-\frac{\Delta E}{k_B T})\}$ 
6       If  $rand() \leq Pacc$  then
7            $\theta_k \leftarrow \theta_k^*$ 
8       End if
9   End for
10  更新伊辛的旋转方式与 2-10 号线类似
11   $s_k \rightarrow -s_k$ 
12 Return
```

Wolff 算法：通常情况下，Wolff 算法对于纯 XY 模型需要将自旋矢量分解为两个不同的相互垂直的分量，即 $\vec{\theta} = \theta_x \vec{i} + \theta_y \vec{j}$ ，此时 $H_{XY} = H_{xx} + H_{yy}$ ，然后再分别更新 H_{xx} 和 H_{yy} 。与更新 Ising 自旋相同，当更新 XY 自旋值时，Ising 自旋的相互作用可以被看作是一个 XY 自旋的耦合系数。

在参考文献中，点 i, j 形成集群的概率为：

$$P(k, i) = 1 - \min \{1, e^{\frac{2H_{xx}}{k_B T}}\} \quad (3)$$

这里，给出了一个更普遍的表达式：

$$P(k, i) = 1 - \min \{1, \exp \left[\frac{E(\theta_k - \theta_i) - E(\theta_k^* - \theta_i)}{k_B T} \right]\} \quad (4)$$

其中 $E(\theta_k - \theta_i)$ 是旧的键能，新的键能是 $E(\theta_k^* - \theta_i)$ ；使用这种翻转概率可以之间对 XY 模型的自旋 θ 进行翻转，无需分解。伪代码如下。

Algorithm 5 Wolff 算法的一个 MC 步骤

Input:

 构型 $\{\theta_1, \theta_2, \dots, \theta_N\}, \{S_1, S_2, \dots, S_N\}$

```

1      Update XY spins
2       $k \leftarrow N * rand() + 1, \alpha \leftarrow \pi * rand(), id = 1$ 
3       $\theta_k^* \leftarrow 2\alpha - \theta_k, stack[1] \leftarrow k$ 
4      while stack not null or id  $\geq 1$  do
5           $k \leftarrow stack[id]$ 
6          for i in all neighboring sites of site k do
7              Case 1: 点 i 没有被访问
8              Case 2:  $rand < P(k, i)$ 
9              If case 1 and case 2 then
10                 Split  $\theta_i \leftarrow 2\alpha - \theta_i$ 
11                  $stack.push(i)$ 
12                  $id \leftarrow id + 1, stack[id] = i$ 
13             End if
14         End for
15      $id \leftarrow id - 1$ 
16     End while
17 End for
18 更新伊辛自旋的方式与 2-17 号线类似
19 Return
```

4.2.3 物理量

本研究采用了 Metropolis 算法和 Wolff 算法对正方晶格上的 Ising-XY 模型进行了蒙特卡洛模拟，并选用了周期性边界条件。由于热力学极限下(无限大尺寸)的模拟是不现实的，因此我们选择了有限大小的系统进行模拟，并利用系统尺寸的变化和有限尺寸标度来推断系统在热力学极限下的行为。在本研究中，我们使用 L 来表示系统的线性尺寸，对于正方晶格而言，其格点总数为 $N=L \times L$ 。我们采用的尺寸为 $L=16/32/48$ 。接下来，我们将利用上述哈密顿量来定义 Ising-XY 耦合模型，并计算在蒙特卡洛模拟中的能量。

序参量 M 为铁磁系统的关于 Ising 模型的平均磁化强度：

$$M = \sum_i S_i / L^2 \quad (3)$$

序参量 M_{anti} 为反铁磁系统的关于 Ising 模型的平均磁化强度：

$$M_{anti} = \sum_{x=2n}(\sum_{y=2n} S_{x,y}) + \sum_{x=2n+1}(\sum_{y=2n+1} S_{x,y}) \quad (4)$$

比热 C_v 可被定义为:

$$C_v = \frac{1}{k_B T^2} [\langle E^2 \rangle - \langle E \rangle^2] \quad (5)$$

其中, 能量 E 即可通过模型的哈密顿量 H 计算得到。

用宾得累积量来精确判断伊辛型的相变点的位置, 宾得累积量 B_r 可被定义为:

$$B_r = \langle M^2 \rangle^2 / \langle M^4 \rangle \quad (6)$$

序参量 M_{xy} 为系统的关于 XY 模型的平均磁化强度:

$$M_{xy} = \sum_j \cos \theta_j / L^2 \quad (7)$$

采用 $M_{xy} * L^{1/8}$ 来确定 XY 型的相变点的具体位置。

4.3 结果与讨论

4.3.1 $A = B$ 相图

在各向同性极限 $A=B$ 的情况下, 在温度 T 固定为 1°C 时, 通过改变参数 A 和 C 可以观察到相变的变化情况:

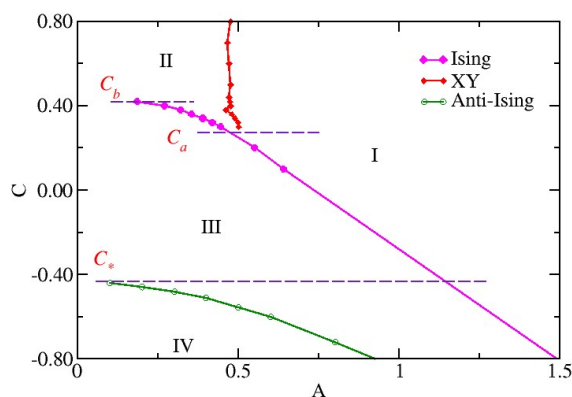


图 4-2 在各向同性极限的情况下通过改变参数 A 和 C 进行蒙特卡罗模拟得到的相图

Figure 4-2 Phase diagram obtained by Monte Carlo simulation in the anisotropic limit

针对 Ising-XY 耦合模型中 $A = B$ 的相位图, 本文通过 Metropolis 算法进行了模拟。我们研究这个 $A=B$ 的相图的动机如下: 首先, 我们需要一个相图, 从这个相图中得出的结果是仿真的出发点, 以确保我们的代码和结果是可靠的^[16]。第二, $A \neq B$ 相图的结构需要由 $A = B$ 相图来验证。另外, 还需要对尚未报道的相的存在进行测试。最重要的是, 这一结果有助于检查有空间各向异性 Ising-XY 耦合模型的相图是否正确。通过比热尖峰的个数粗略判断相变的个数及位置, 并在图 2 中, 全局相位图在 C - A 参数平面中展示。值得注意的是, 当 $C_a < C < C_b$ 时, 系统发生两次相变, 由 Ising 自旋磁化强度和 XY 自旋磁化强度可判定先发生 Ising 型相变再发生 XY 型相变; 当 $C < C_a$ 或 $C >$

C_b 时，系统只发生一次相变，Ising 型相变和 XY 型相变同时发生。此外，在 C-A 参数平面中还可以观察到反铁磁相，在 A 值较小时，若 C 值从-0.44 以下开始增加，可观察到 Ising 自旋由反铁磁相-无序相-铁磁相的转变。

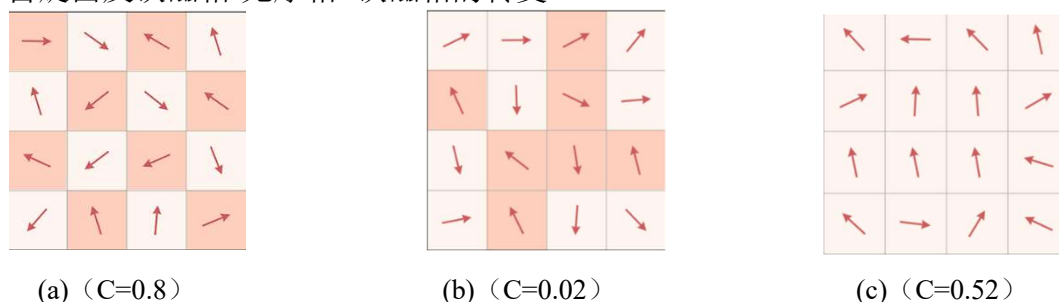


图 4-3 在 $A=B=0.4$ 的情况下，C 值不同时 Ising-XY 耦合模型的构型图

Fig.4-3 Conformation diagram of Ising-XY coupling model with different values of C for $A=B=0.4$

图 4-3 的三种构型图所描述的都是 $A=B=0.4$ 的情况，图中方格颜色代表 Ising 自旋值 σ_i —深色代表 1、浅色代表-1，箭头角度代表 XY 自旋值 θ_i 。忽略 XY 的自旋只考虑 Ising 的自旋时，我们发现在 $C=0.8$ 时，自旋取值为 1 或-1 交替出现，任意格点与最近邻自旋的乘积为-1； $C=0.52$ 时，所有自旋取值相同，任意格点与最近邻自旋的乘积为-1； $C=0.02$ 时，自旋取值看起来比较随机、无明显规律。

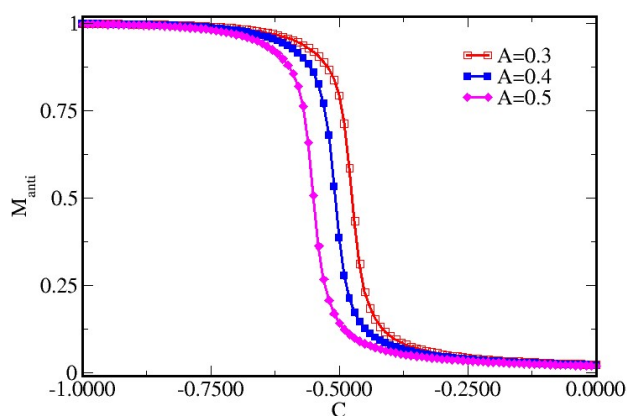


图 4-4 $A=B$ 时，不同 A 值下序参量 M_{anti} 随 C 值的变化情况

Fig. 4-4 Variation of the sequential parameter M_{anti} with the value of C for different values of $A=B$

从图 4-2 观察到当 C 取值较小时，随着 A 值增加 Ising 模型自旋分布从反铁磁序变成无序态最后至铁磁序。图 4 给出了 $A=B$ 的情况下，随着 C 值的增加，反铁磁磁化强度 M_{anti} 由 1 逐渐减小至 0，并且 $A=0.3 \setminus 0.4 \setminus 0.5$ 时发生突变的位置随 A 值的增大而减小。

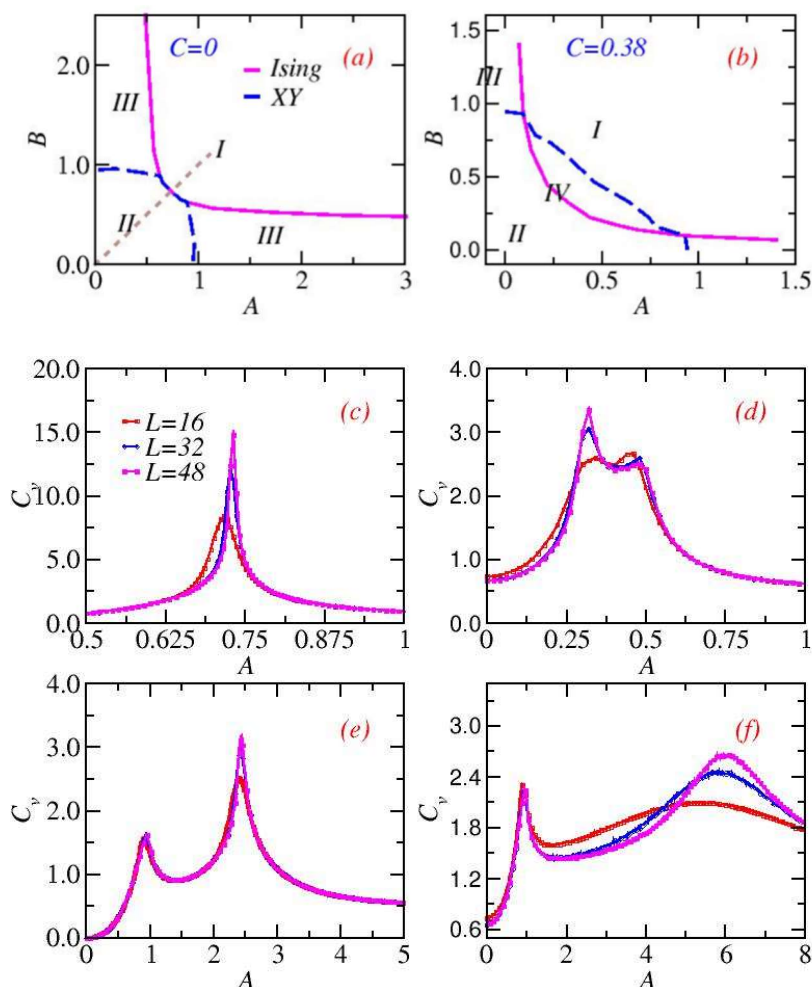
4.3.2 $A \neq B$ 相图


图 4-5 (a) (b) 在 $C=0$ 、 $C=0.38$ 时, Ising_XY 模型在 A , B 参数平面上的经典相图。蓝线是 XY 相变, 红线是 Ising 相变; (c) (d) (e) (f) 分别为不同参数下比热尖峰 C_v 随参数 A 的变化, (c) $C=0.0$, $B/A=1$; (d) $C=0.38$, $B/A=1$; (e) $C=0.0$, $B/A=0.2$; (f) $C=0.38$, $B/A=0.01$;

Fig. 4-5 (a) (b) Classical phase diagram of the Ising_XY model in the A , B parameter plane at $C = 0$ and $C = 0.38$. The blue line is the XY phase transition and the red line is the Ising phase transition; (c) (d) (e) (f) Variation of specific heat spike C_v with parameter A for different parameters, (c) $C=0.0$, $B/A=1$; (d) $C=0.38$, $B/A=1$; (e) $C=0.0$, $B/A=0.2$; (f) $C=0.38$, $B/A=0.01$, respectively.

图 4-5 (a) (b) 中蓝色虚线表示 XY 相变的边界, 另一条线表示 Ising 相变, 分别描述了 $C=0$ 和 $C=0.38$ 情况下关于参数 AB 平面的相图。两条线在 A 、 B 比较接近且 $C=0$ 的区域重叠, $C = 0$ 时, 沿切线 $A = B$ (棕色虚线) 扫描, 两种类型的相变同时发生, 在图 3(c)中, 比热尖峰 C_v 只有一个峰值证实了这一点。然而, 当 $C=0.38$ 时, 沿切线 $A = B$ (棕色虚线) 扫描, 在图 3(d)中, 对于比热尖峰 C_v 除了一个主峰外, 还有一个较小的扭结, 意味着这里可能还存在一个相变。在图 3(e)和图 3(f)中, 即 $C=0.0$,

$B/A=0.2$ 和 $C=0.38$, $B/A=0.01$ 两种情况下对于比热尖峰 C_v 很明显有两个峰存在, 意味着这里存在两个相变。就如前面所讲的一样, 通过比热尖峰我们只能判断相变的个数无法判断相变的类型及先后顺序, 接下来我们需要计算宾得累积量 B_r 和 M_{xy} 随参数 A 变化的关系才能判断 Ising 型相变以及 XY 型相变的位置并确定同一系统中两个相变发生的先后关系。

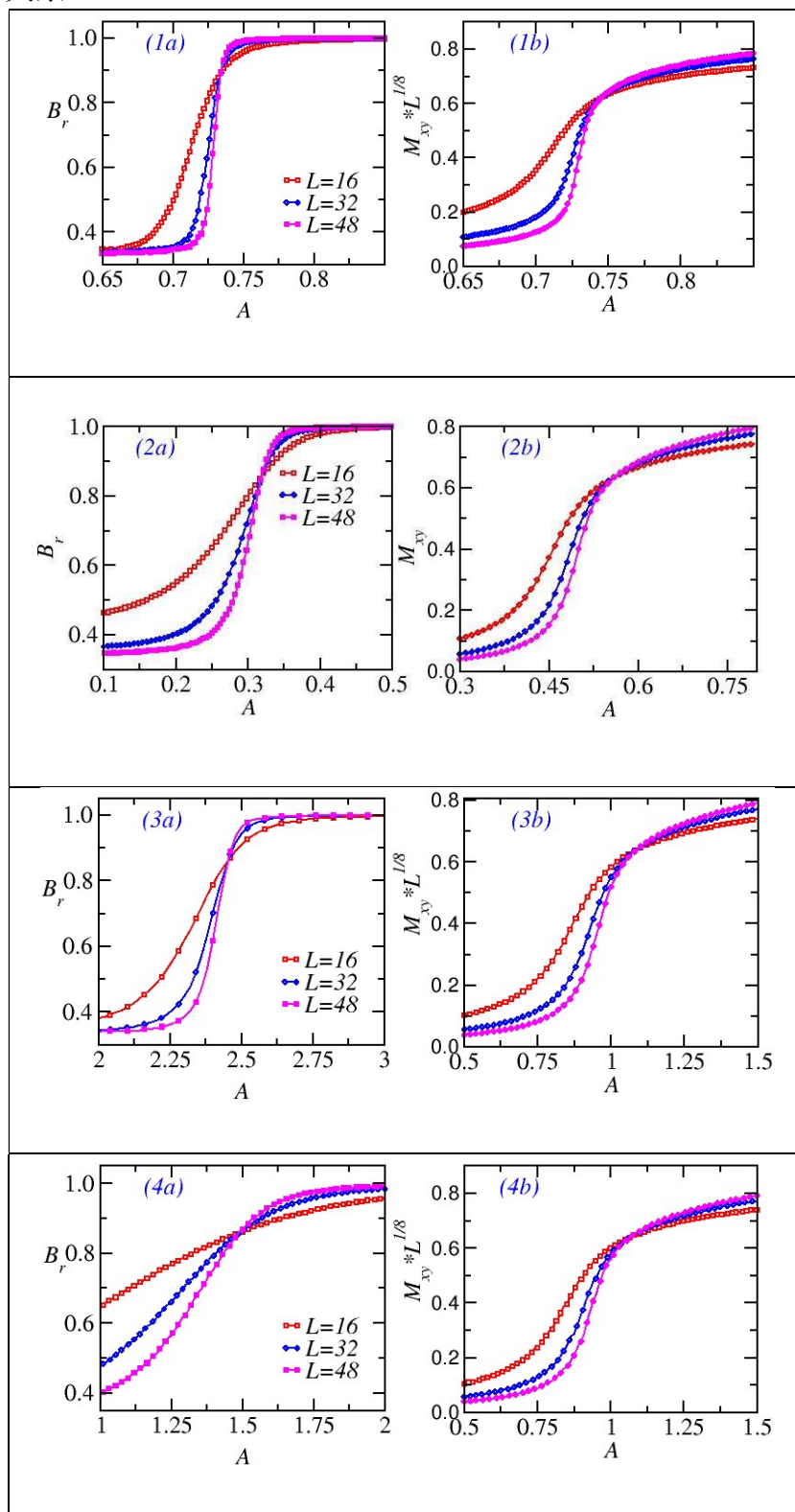


图 4-6 (1a)(1b) - (4a)(4b) 分别为不同参数下宾得累积量 B_r 与物理量 $M_{xy} * L^{\frac{1}{8}}$ 随参数 A 的变化, (1a)、(1b) $C=0.0$, $B/A=1$; (2a)、(2b) $C=0.38$, $B/A=1$; (3a)、(3b) $C=0.0$, $B/A=0.2$; (4a)、(4b) $C=0.38$, $B/A=0.01$;

Figure.4-6 (1a)(1b)-(4a)(4b) shows the variation of the cumulative quantity B_r and the physical quantity $M_{xy} * L^{\frac{1}{8}}$ with parameter A for different parameters, (1a), (1b) $C=0.0$, $B/A=1$; (2a), (2b) $C=0.38$, $B/A=1$; (3a), (3b) $C=0.0$, $B/A=0.2$; (4a), (4b) $C=0.38$, $B/A=0.01$.

图 6 中宾得累积量 B_r 在不同尺寸下的交点可用来相对精确判断 Ising 型相变点的位置, 而 XY 的磁化强度与尺寸的八分之一次方的乘积 $M_{xy} * L^{\frac{1}{8}}$ 在不同尺寸下的交点可用来相对精确判断 XY 型相变点的位置。从上图 (1a)(1b) 观察到在 $C=0.0$, $B/A=1$ 时, Ising 型相变和 XY 型相变都在 $A=0.731$ 时刻发生; 从上图 (2a)(2b) 观察到在 $C=0.38$, $B/A=1$ 时, Ising 型相变在 $A=0.331$ 时刻发生、XY 型相变都在 $A=0.561$ 时刻发生; 从上图 (3a)(3b) 观察到在 $C=0.0$, $B/A=0.2$ 时, Ising 型相变在 $A=2.441$ 时刻发生、XY 型相变在 $A=2.441$ 时刻发生 $A=0.4882$; 从上图 (4a)(4b) 观察到在 $C=0.38$, $B/A=0.05$ 时, Ising 型相变在 $A=0.32$ 时刻发生、XY 型相变在 $A=0.461$ 时刻发生。也就是说除了 $C=0.0$, $B/A=1$ 的情况之外, 其他三种情况下都能观察到两个相变, 该结果与比热尖峰 C_v 相吻合。

4.4 本章小结

本研究使用了蒙特卡罗方法中最为常用的 Metropolis 算法以及效率更高的 Wolff 算法, 对空间各向异性的 Ising-XY 耦合模型的相变情况进行了计算研究。在研究过程中, 我们发现了许多有趣的现象, 并成功绘制出了该模型在空间各向异性情况下的相图。在研究各向同性极限 ($A=B$) 下的情况时, 我们得到了图 2, 并发现不同的 C 值对应的相变情况也不同。当 $C > C_b$ 时, 只有 XY 型相变, Ising 自旋一直有序分布; 当 $C_a < C < C_b$ 时, 先发生 XY 型相变再发生 Ising 型相变; 当 $C_* < C < C_a$ 时, Ising 型相变和 XY 型相变同时发生; 当 $C < C_*$ 时, 先发生相变, Ising 自旋和 XY 自旋同时由有序变无序然后 Ising 自旋再由无序变成反铁磁序。根据图 2 我们选定 $C=0.38$ 分析 $A \neq B$ 时该模型的相图; 在各向异性极限下的情况, XY 相变先于 Ising 相变发生; 随着 B/A 的增大 Ising 相变与 XY 相变对应的 A 值逐渐近并相交, 此时两种相变同时发生; 此后随着 B/A 的增大, Ising 相变要先于 XY 相变发生, 待 $A=B$ 时与各向同性极限的情况相重合。当 $C = 0$ 时, 在各向异性极限下的情况, XY 相变先于 Ising 相变发生; 随着 B/A 的增大到某一值时 Ising 相变与 XY 相变对应的 A 值逐渐近并相交, 此时两种相变同时发生; 此后随着 B/A 的增大, 两种相变仍一同发生, 待 $A=B$ 时与各向同性极限的情况相重合。交换 A 和 B 相图是对称的。本章计算结果与前人的理论推导符合, 可作为有力的支撑。

第五章 图神经网络生成 IsingXY 模型

5.1 引言

近年来,物理学家已经开始采用机器学习技术来研究相变^[17]。大多数任务都是通过监督学习算法来完成的^[18]。在有监督学习中,算法是对带标签的数据进行训练,将标签分配给数据点。经过成功的训练,可以对以前未见数据的标签进行高精度的预测。除了有监督学习外,还有一些无监督学习算法可以在无标签数据中找到结构。已经有可能使用无监督的学习技术来重现伊辛模型的蒙特卡罗采样状态^[19],以无监督的方式发现相变已经成为主流。全连接神经网络(Fully Connected Neural Network)^[20]、卷积神经网络(Convolution Neural Networks, CNN)^[21]、变分自编码器(Variational Auto Encoder, VAE)^[22]、生成对抗网络(Generative Adversarial Network, GAN)^[23]等都在 Ising 模型的模拟上广泛应用^[24],并取得了不错的效果。但是这些模型在训练时都无法兼顾物理模型分子与分子之间的相互作用和分子之间边的影响,同时很多物理模型的微观结构是不规则的,生成模型的泛用性差一般只能适应一种温度下的一种尺寸。这就需要一种能考虑非结构化数据,并且能考虑节点与节点、节点与边的关系的深度神经网络。

幸运的是近年来,深度学习领域关于图神经网络(Graph Neural Networks, GNN)的研究热情日益高涨^[25],GNN 处理非结构化数据的出色能力使得其在物理、生物、化学方面的应用尤为突出^[26]。IsingXY 模型本质上也是一种图结构,图卷积网络(Graph Convolution Networks, GCN)^[25]可以聚合邻居节点的信息,相比于传统卷积,GCN 更善于编码图的结构信息,能够学习到更好的表示特征。变分图自编码(Variational Graph Auto Encoder, VGAE)是一种无监督图生成算法^[27],是图神经网络与生成网络相结合的产物,它能够有效利用数据的图结构去模拟训练数据的分布。**本章中主要研究什么?**

5.2 模型与算法

5.2.1 Ising 模型

在统计物理中,Ising 模型被描述为某些晶格中具有耦合相互作用的二元自旋的集合。考虑 N 个自旋 $s=\{s_i\}$ 可以取值 ± 1 ,其中索引 i 标记自旋 s_i 的位置,Ising 系统的标准哈密顿量仅包括最近邻相互作用,并且每个自旋方向可能是"向上"(+1)或"向下"(-1),尽管广义模型可能包括长程相互作用和更多的自旋选择方向.标准哈密顿量^[28]是:

$$H = -J \sum_{neighbors} S_i * S_j \quad (1)$$

5.2.2 图神经网络

图是一种数据结构,它对一组对象(节点)及其关系(边)进行建模。今年来,由于图结构的强大表现力,用机器学习方法分析图的研究越来越受到重视^[29]。图神经网络是指使用神经网络学习图结构数据,提取和发掘图结构数据中的特征和模式,满足聚类,分类,预测,分割,生成等图学习任务需求的算法总称。图神经网络发展的一个动机源于卷积神经网络。卷积神经网络的广泛应用带来了机器学习领域的突破并开启了深度学习时代^[21]。然而卷积神经网络只能在规则的欧式空间数据中提取潜在特征,不能很好的对应现实中复杂多变的图数据,如何将卷积神经网络应用于图结构这一非欧式空间成为图神经网络模型重点解决的问题^[30]。以下给出图神经网络中相关符号的说明和定义^[31]。

定义 1. 图是由节点和连接节点的边所构成的,通常记为 $G=(V,E)$ 。其中 $V=\{v_1, v_2, \dots, v_n\}$ 代表节点集合, $E=\{e_1, e_2, \dots, e_n\}$ 代表边集合,边也可以用 (v_1, v_2) 的方式来表示.通常节点也被称为顶点或者交点.边也成为链接或者弧.通用的图表示是一个五元组: $G(V,E,A,X,D)$.其中 A 代表图的邻接矩阵, X 代表节点的特征矩阵, D 代表度矩阵。

定义 2. 邻接矩阵:图的邻接矩阵指用于表示图中节点的连接情况的矩阵.该矩阵可以是二值的,也可以是带权的.对于有 N 个节点的无向图来说,邻接矩阵是一个 $N*N$ 的实对称矩阵。



图 5-1 (a)图结构;(b)图的邻接矩阵

Figure.5-1 (a) graph structure; (b) adjacency matrix of the graph

定义 3. 度矩阵:节点的度表示与该节点相连的边的数量.图的度矩阵即用于描述图中每个节点的度的矩阵.度矩阵是一个对角矩阵,对于无向图来说,一般只使用入度矩阵或者出度矩阵。

定义 4. 组合拉普拉斯矩阵,又称标准拉普拉斯矩阵,由对角矩阵和邻接矩阵组合而成:

$$L = D - A \quad (2)$$

该矩阵只在中心节点和一阶相连的节点上有非零元素,其余之处均为零.拉普拉斯矩阵也是图的一种表现形式。

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

图 5-2 (a)度矩阵;(b)拉普拉斯矩阵

Fig.5-2 (a) Degree matrix; (b) Laplace matrix

定义 5. 归一化拉普拉斯矩阵^[32]

$$L^{sym} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \quad (3)$$

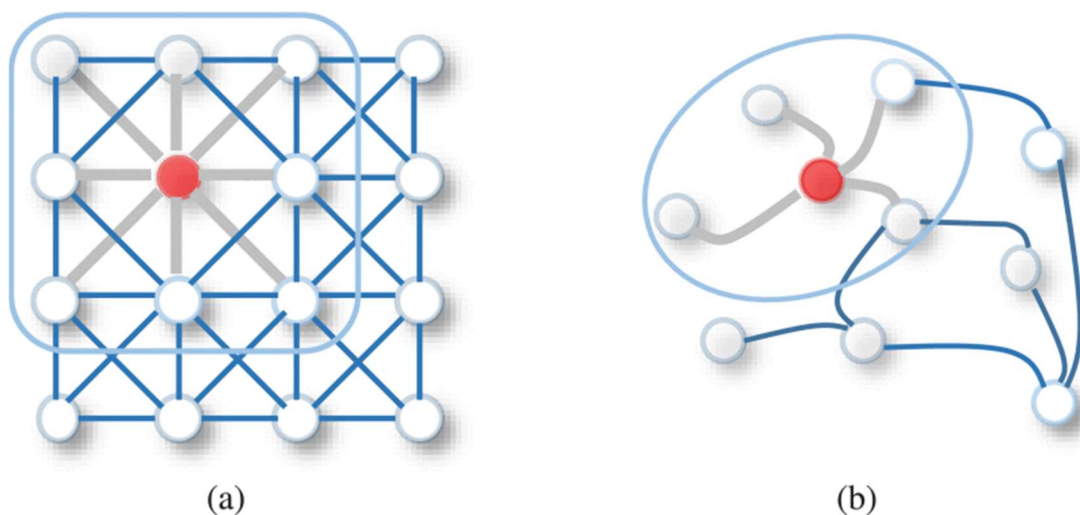
其元素值为:

$$L_{i,j}^{sym} = \begin{cases} 1 & i=j, \text{且 } \deg(v_i) \neq 0 \\ -\frac{1}{\sqrt{\deg(v_i)\deg(v_j)}}, & i \neq j, \text{并且 } v_i \text{与 } v_j \text{相连} \\ 0 & \text{其他} \end{cases} \quad (4)$$

其中 $\deg(v)$ 表示为节点 v 的度

5.2.3 图卷积网络

图是一种数据结构，它对一组对象（节点）及其关系（边）进行建模。今年来，由于图结构的强大表现力，用机器学习方法分析图的研究越来越受到重视。图神经网络是指使用神经网络学习图结构数据，提取和发掘图结构数据中的特征和模式，满足聚类，分类，预测，分割，生成等图学习任务需求的算法总称^[33]。图神经网络发展的一个动机源于卷积神经网络。卷积神经网络的广泛应用带来了机器学习领域的突破并开启了深度学习时代。然而卷积神经网络只能在规则的欧式空间数据中提取潜在特征，不能很好的对应现实中复杂多变的图数据，如何将卷积神经网络应用于图结构这一非欧式空间成为图神经网络模型重点解决的问题。以下给出图神经网络中相关符号的说明和定义。



(a) 二维卷积. 类似于图, 图像中的每个像素都被视为节点, 其中相邻节点由过滤器大小决定。二维卷积取红色节点及其相邻节点的像素值的加权平均值。相邻节点是有序的, 并且具有固定的大小。

(b) 图形卷积. 为了获得红色节点的隐藏表示, 图卷积运算的一个简单解决方案是获取红色节点及其相邻节点的节点特征的平均值。与图像数据不同, 相邻节点是无序的, 大小可变。

图 5-3 二维卷积与图卷积的比较^[34]

Figure.5-3 Comparison of two-dimensional convolution and graph convolution

对于图卷积网络来说, 想要学习到它对于每个节点的特征表示, 任意的图卷积都可以写成这样的一个非线性函数:

$$H^{l+1} = f(H^l, A) \quad (5)$$

$H^0 = X$ 为第一层的输入, l 代表神经网络的层数, A 代表邻接矩阵。

传统的图卷积网络做法就是将邻居节点的信息加入到本节点当中:

$$f(H^l, A) = \sigma(AH^lW^l) \quad (6)$$

W_l 代表第 l 层的参数矩阵, σ 代表激活函数。根据矩阵乘法可以看出上式中的每个节点都结合了相邻节点的信息, 但由于仅仅使用了邻接矩阵, 邻接矩阵对角线为 0, 无法体现节点自身的信息, 所以我们使用拉普拉斯矩阵来代替邻接矩阵

$$f(H^l, A) = \sigma(LH^lW^l) \quad (7)$$

上式中引入了拉普拉斯矩阵, 从而解决了没有考虑自身节点信息自传递的问题, 但是由于没有被规范化, 我们将拉普拉斯矩阵规则化得到图卷积

$$f(H^l, A) = \sigma(L^{sym}H^lW^l) \quad (8)$$

上面讲述的都是以矩阵形式计算, 我们也可以从单个节点角度来观察公式, 对于第 $l+1$ 层的节点特征 h^{l+1} , 对于他的邻接结点 $j \in N$, N 是结点 i 的所用邻居结点的集合, 所以图神经网络的更新公式同样可以描写为

$$h_i^{l+1} = f(h_i^l) = \sigma(h_i^l \cdot W_1^{l+1} + \sum_j h_j^l \cdot W_2^{l+1}) \quad (9)$$

5.2.4 变分自编码器

变分自编码器是一种变分推理的生成模型，由编码器和解码器两部分网络组成，它是包含隐变量的一种模型。在变分自编码器当中，我们基于这样的假设：我们的样本 x 是某些隐变量 z (latent variable) 通过某种映射产生出的，而 z 也不是一个固定值，而是服从一个分布： $z \sim P_\theta(z)$ ，则 $x \sim p_\theta(x|z)$ ，这里的 P 是由参数 θ 决定的分布族，而 θ 就是我们需要找到的真实分布。所以我们最根本的目标就是确定 z 和 θ 的值，这样就能够得到样本 x 。为了推断 $p_\theta(x|z)$ ，我们需要最大化边际对数似然 $\log(p_\theta(x))$ 我们可以重写该式^[35]：

$$\begin{aligned} \log(p_\theta(x)) &= \log \int p_\theta(x|z) dz \\ &= \log \int \frac{p_\theta(x|z)p(z)q_\phi(z|x)}{q_\phi(z|x)} dz \\ &\geq E_{z \sim q_\phi(z|x)} [\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x) \| p(z)) \\ &:= \varepsilon(x, \theta, \phi) \end{aligned} \quad (10)$$

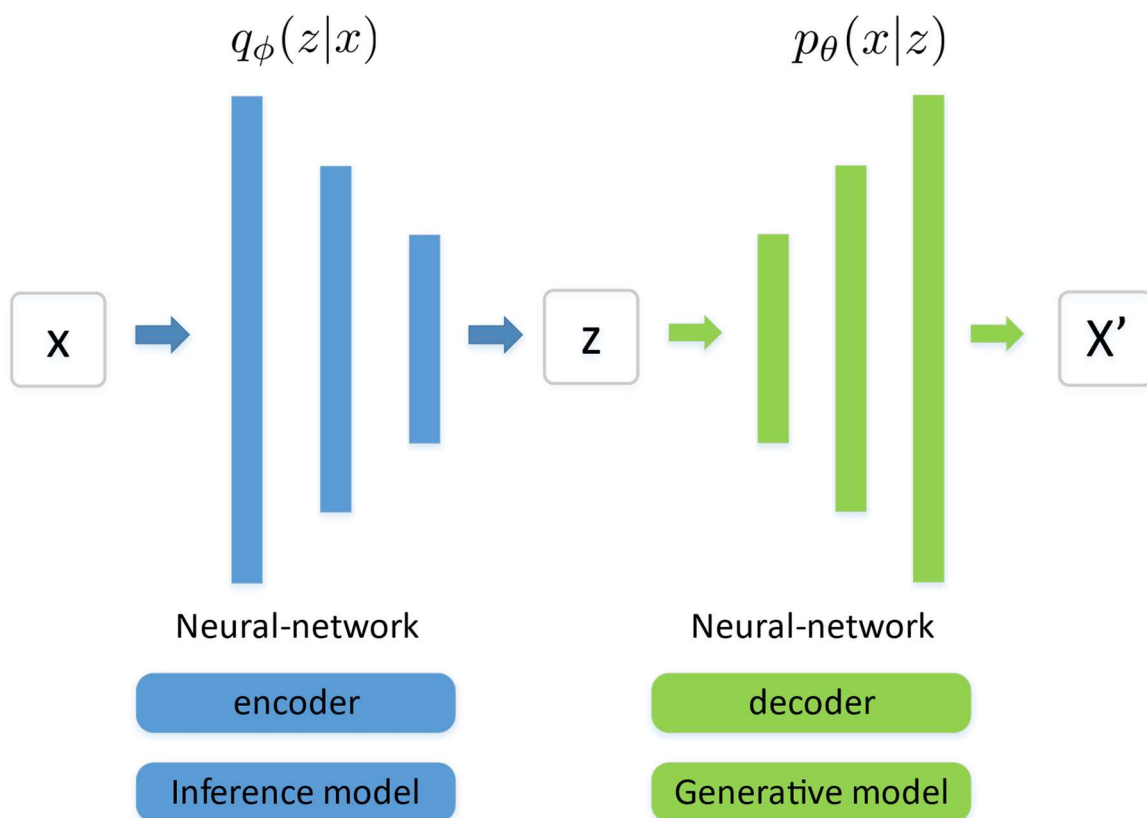


图 5-4 变分自编码模型

Figure.5-4 Variational AutoEncode model

在第四步应用了 Jensen 不等式，在第二步中对潜在变量 z 的积分通常是棘手的，因此引入了带有参数集 ϕ 的近似后验分布 $q_\phi(z|x)$ 并使用变分推理原理来获得边缘对数似

然的易处理界限,即证据下界(ELBO)^[22]。使用 $p(z)=N(1,0)$ 作为潜在变量先验.ELBO 是对数似然的易处理下界,因此可以最大化推断 $p_\phi(x|z)$ 。 $Ez \sim q_\phi(z|x)[\log(p_\phi(x|z))]$ 可以理解为重构误差,因为最大化它会使解码器的输出类似于编码器的输入, $DKL(q_\phi(z|x)|p(z))$ 是 KL 散度,一种衡量两个分布相似性的数值,它可以确保潜在表示是高斯的,使得具有相似特征的数据点具有相似的高斯表示。

我们已经概述了潜在变量模型背后的一般思想,但我们仍然需要指定近似(变分)后验 $q_\phi(z|x)$ 和模型似然 $\log_\phi(p(x|z))$ 。近似后验的常见选择是分解高斯分布

$$q_\phi(z|x) = \prod_{j=1}^d N(z_j|\mu_j, \sigma_j^2) \quad (11)$$

其中 μ 和 σ 表示模型的均值和标准差。

总而言之,VAE 是一种基于提取出 $p(z|x)$ 的潜在高斯表示 $q_\phi(z|x)$ 的均值与方差的编码器网络。VAE 中的解码器,它使用高斯 $q_\phi(z|x)$ 的样本作为输入,根据分布 $p_\theta(x|z)$ 生成新样本。我们通过使用反向传播最大化 ELBO 来计算所有的参数,但是高斯分布是随机的,不可微的。因此需要在编码器的输出和解码器的输入之间建立一个确定性和可微分的映射,需要将随机变量 z 表示为另一个辅助随机变量 ϵ 的可微可逆变化 g (即 $z = g(\mu, \sigma, \epsilon)$),我们采用重参数化技巧^[22],使 $g(\mu, \sigma, \epsilon) = \mu + \sigma \odot \epsilon$ 即

$$z = \mu + \sigma \odot \epsilon \quad (12)$$

其中 $\epsilon \sim N(0,1)$ 和 \odot 是元素乘积

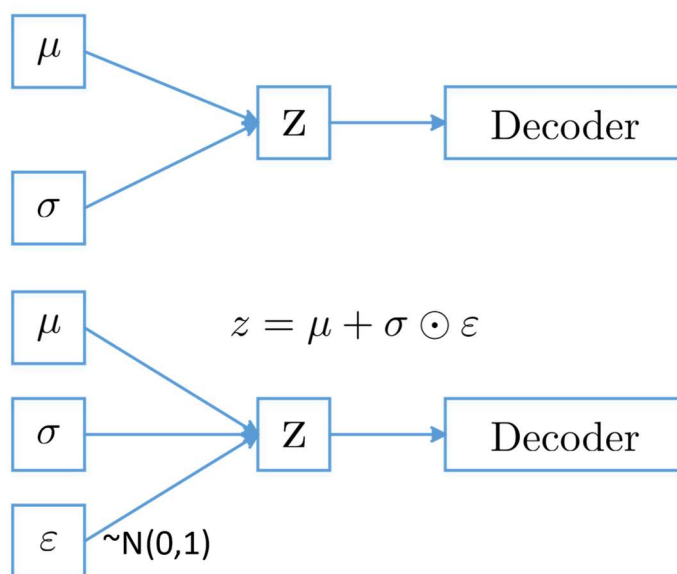


图 5-5 重参数技巧

Fig.5-5 Reparameter trick

5.2.4 图变分自编码

编码器部分:

我们的编码器由三个卷积层构成,分别具有 32,64,64 个通道和一个整流线性单元 (ReLU) 作为激活函数.为了归一化神经网络我们使用了 `batchNorm`^[36],`GraphNorm`^[37]和 0.5 的 `dropout`^[38]率.前两个 GCN^[39]层生成一个高维特征矩阵,同时我们考虑图结构中的边信息,将边的特征参与到卷积聚合当中,参考之前图卷积的更新公式可以得到我们新的更新公式:

$$x'_i = W_1 x_i + W_2 \sum_{j \in N(i)} e_{j,i} \cdot x_j \quad (13)$$

其中 x_i 表示节点特征, $e_{j,i}$ 表示从源节点 j 到目标节点 i 的边权重.

卷积操作中上层卷积的结果作为下层卷积的输入来进行卷积的堆叠

$$\hat{x} = GCN(GCN(X, A), A) \quad (14)$$

第三个卷积层我们采用两个不同的卷积网络来分别获取 $\mu, \log \sigma^2$

$$\mu = GCN_u(\hat{x}, A) \quad (15)$$

$$\log \sigma^2 = GCN_{\log \sigma^2}(\hat{x}, A) \quad (16)$$

然后我们就可以根据 Eq (12) 从分布中进行采样 Z 。

解码器部分:

我们的采样结果 z 是节点的高维表示,在解码器部分,由于 IsingXY 模型的节点之间的连接方式是固定的,所以我们在解码器部分可以忽略 IsingXY 模型的边属性,对潜在变量 z 中的每个节点使用全连接神经网络^[20],使其维度降维至 1.此时我们可以将重构问题认为是对每个节点进行分类,使用 `sigmoid` 激活函数,再将其按值进行分类,之后对所有节点进行重构就可以得到我们生成的 IsingXY 构型。

$$\bar{x}_i = MLP(z_i) \quad (17)$$

z_i 代表节点的高维表示,通过全连接神经网络使其变为我们需要的重构节点.由于我们得到的结果为节点信息,实现的方式为对每个节点进行卷积和降维,这样做的好处不仅仅是可以处理不同节点数量的图而且避免了传统矩阵形式的卷积核提取特征所带来的噪音,我们的模型在适应各种尺寸的构型的同时精度也得到了保证。

通过后续的实验信息可以看出,不同尺寸的构型一起进行训练,可以让模型提取出不同尺寸下稳态构型的共性特征,从而使生成的构型更加符合真实构型。

损失函数:

$$L = MSE(x, \bar{x}) + KL(q(Z|X, A) || p(Z)) \quad (18)$$

第一项 `MSE` 是均方误差,用于测量网络重构数据的程度,防止重构数据过度偏离原始数据;第二项 `KL` 散度,其中 $p(Z) = N(0, 1)$.它衡量我们的 $q(Z | X, A)$ 与 $p(Z)$ 的匹配程度

在经过多次迭代以后,损失函数的值小于阈值后,我们就可以认为模型训练成功,这时候我们可以保存模型和训练出的 μ 与 σ 并使用解码器来生成我们想要的构型.我们的模

型可以生成多种尺寸的构型,将不同尺寸的构型一起训练,并在训练完成后,将想要尺寸的构型再进行 1~5 步的尺寸固定训练,这种训练几乎不耗费时间,但能将网络中的生成节点数量变换成我们想要的构型节点数量,这样就可以生成不同尺寸的构型。

5.3 模拟结果与分析

为了验证上述网络的有效性,我们建立了一个 IsingXY 模拟器,在临界温度附近的 32×32 晶格与 16×16 晶格中产生 IsingXY 状态。我们所模拟的模型由 Eq(1)的哈密顿描述。

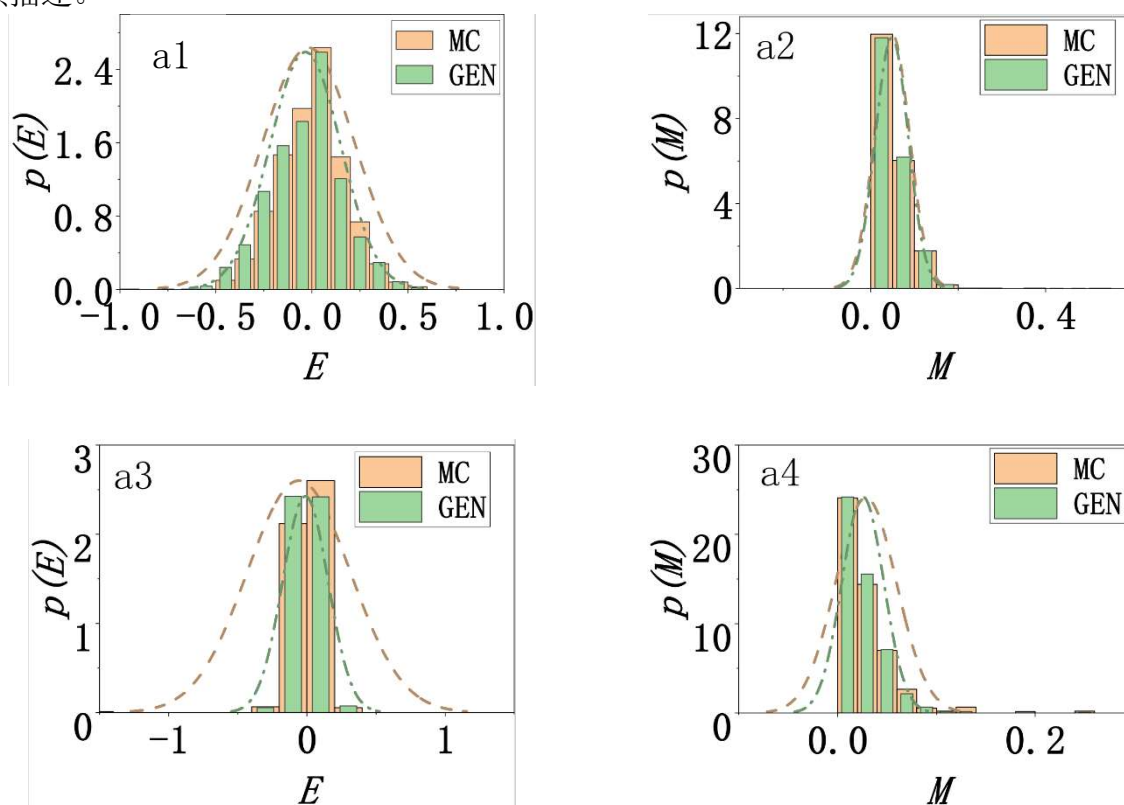


图 5-1 (a1)为 16×16 尺寸下, IsingXY 构型的能量分布;(a2)为 32×32 尺寸下, IsingXY 构型的磁化强度分布;(a3)为 32×32 尺寸下 IsingXY 构型的能量分布;(a4)为 32×32 尺寸下, IsingXY 构型的磁化强度分布

Fig.5-1 (a1) shows the energy distribution of IsingXY configuration at 16×16 size; (a2) shows the magnetization intensity distribution of IsingXY configuration at 32×32 size; (a3) shows the energy distribution of IsingXY configuration at 32×32 size; (a4) shows the magnetization intensity distribution of IsingXY configuration at 32×32 size

5.4 总结

本章采用了机器学习中目前较为火热的图神经网络来模拟 IsingXY 构型,实验证明,图神经网络能寻找到构型中自旋点之间的联系,并能较为精确的模型 IsingXY 模型。实验中,通过使用带有边属性的图卷积神经网络来训练 IsingXY 模拟器,我们生

成了与传统蒙特卡洛模拟类似的 IsingXY 构型。由于 VGAE 的性质，我们生成的构型不是现有构型的重建，而是重新生成了独一无二的全新构型。本文中描述的模型代表了一个用图神经网络模拟物理构型的一个论证，在未来，通过采用更复杂的训练方法与更加强大的计算能力，可以进一步提高其精度。我们设想这种生成网络在模拟一般多体系统中具有广泛的应用。如对传统采样方法的加速，根据图的特性可以在一种模型中生成多种尺寸不同的构型以及基于实验观察的多体系统的精确表示。

第 6 章 结论与展望

本论文主要利用蒙特卡罗方法中的 metropolis 算法和 wolff 算法，对 XY 耦合模型、ISing-XY 耦合模型在不同尺寸下进行大规模模拟运算，对在空间各向异性下两个模型的物理量如：能量、磁化强度、比热、宾得累积量、超流密度进行了运算，并画出 XY 耦合模型、ISing-XY 耦合模型的相图，在第五章中，我们使用机器学习中的图神经网络来模拟 Ising-XY 模型，并将模拟生成的构型与传统蒙特卡罗方法生成的构型进行物理性质的比对。具体内容如下：

(1)通过耦合 XY 模型的哈密顿量计算其能量作为蒙特卡罗模拟时自旋是否翻转的判断条件，经过足够的弛豫时间达到平衡后计算大量样本的物理量求取平均值，利用物理量的取值可判断自旋分布的情况从而判断模型是有序态还是无序态以及是否发生相变。接下来我们以 $J_a = J_b$ 时的 XY 耦合模型为例，在任何尺寸下能量在温度为 0°C 时对应的能量值大约为 -4eV ，与理论推导吻合；随着温度的增高能量逐渐增加至 0eV ，在温度为 $1.3-1.4^\circ\text{C}$ 左右时增速最快发生跃变，尺寸越大跃变现象越明显，因此我们猜测相变点应该在这附近。磁化强度的值在初始时为 1，自旋排列有序，随着温度的增加热运动越来越剧烈，自旋排列逐渐趋于无序故而磁化强度逐渐降低，与能量类似在温度为 1.4 时降低速度最快。宾得累积量也是随着温度升高从 1 至 0，区别于上述物理量之处是它在不同尺寸下会交于一点，通常我们将该点对应的位置记为相变点，通过计算宾得累积量可精确计算出相变点的位置为 1.3246°C 。超流密度通常用于计算 XY 型相变点的位置，在 $J_b \neq 0$ 的情况下，超流密度随温度变化的曲线与直线 $\frac{\Delta\rho_s}{T_c} \sim 2/\pi$ 的交点对应的横坐标记为相变点，可得到不同尺寸下相变点的位置，通过尺寸外推公式可得尺寸为无限大时相变点为 1.3246°C 。比较 ISing 型相变点与 XY 型相变点，可得出误差在允许范围内，可视为两种相变同时发生。以上是 $J_a = J_b$ 时的情况，保持 J_a 不变、随着 J_b 减小，发生相变时的温度逐渐减小，但两种相变仍在误差允许的范围内可视为同时发生。 $J_b = 0.5J_a$ 时，相变点为 0.9718°C 或 0.2482°C 。

(2) 同样，通过 Ising-XY 耦合模型的哈密顿量计算其能量作为蒙特卡罗模拟时自旋是否翻转的判断条件，经过足够的弛豫时间达到平衡后计算大量样本的物理量求取平均值，利用物理量的取值可判断自旋分布的情况从而判断模型是有序态还是无序态以及是否发生相变。接下来我们以 $A = B$ 时的 Ising-XY 耦合模型为例，通过调节 C 值可以观察到在不同取值时对应的相变个数不同，由此可选择恰当的 C 值观察研究 $A \neq B$ 时 Ising-XY 耦合模型的相图。比热是用来反映能量涨落的物理量，比热越大说明涨落现象越明显，比热尖峰是在一段时间内涨落最明显的位置，通常在相变点附近，一般情况下比热尖峰的个数代表相变点的个数。当 $C=0.0$ ， $B/A=1$ 时只存在一个比热尖峰即只有一个相变点；其他三种情况下都存在两个相变点。比热尖峰只能判断相变点的个

数无法判断相变点的类型，因此我们借助宾得累积量和 $M_{xy} * L^{\frac{1}{8}}$ 分别判定 Ising 型相变和 XY 型相变的位置。通过调整 A 与 B 的比例关系可算出各向异性下 Ising-XY 耦合模型的相图。

(3) 构建了一个生成 Ising-XY 构型的图卷积网络生成器，并对 16, 32 两种不同尺寸的 Ising-XY 模型在相变点附近的构型进行模拟。从实验结果可知，我们确定了使用图神经网络来模拟凝聚态物理模型的可行性。通过使用图卷积神经网络来训练 Ising-XY 生成器，我们成功生成了想要的全新的 IsingXY 模。对于不同卷积层的图变分自编码器来讲，三层为生成精度最高的卷积结构，由此可以得出卷积层数的多少决定了生成精度，对于卷积来说就是将周围节点的信息聚合到中心节点，层数过多会考虑更远程的影响从而影响不同距离节点之间影响的比重，同时更多层数的卷积层也对实验设备有着更高的要求。

综上，本文所研究的双层 XY 耦合模型和 Ising-XY 耦合模型都存在着许多有趣的物理现象，在超导系统中有着广泛的应用前景，值得进一步的探究讨论。对于图神经网络来说，我们可以通过不同的网络模型以及权重处理来解决上述问题，从而使精度更加提高，我们猜想生成网络在凝聚态物理中具有更加广泛的应用以及我们尚未发现的潜在特性，如对传统采样算法的加速以及对复杂系统的精确表示。

展望：

本论文虽已成功计算出双层耦合 XY 模型以及 Ising-XY 模型的物理量和相图，但依然有些地方需要进一步的改进和完善，对此我们进行如下展望：

(1) 本文中通过蒙特卡罗计算双层 XY 耦合模型只发现了一个相变，与先前学者们的理论推导存在两种相变的情况不符，需要进一步分析造成这种矛盾的原因。此外，还需进一步探究双层 XY 模型在由低温至高温过程中自旋构型图，寻找是否存在涡旋或者半涡旋，进一步探究该模型中可能存在的其他的有趣现象并与其他类似模型相比较。

(2) 第四章中虽然已经比较全面准确的论证了 Ising-XY 模型相变的具体内容，但是对于 $C > C_b$ 和 $C < C_*$ 情况下各向异性时相变的研究缺乏具体全面的研究，值得进一步补充。

(3) 由于仪器设备和实验时间的有限，本文所有计算的尺寸大都小于 48，不够精确，并且对于其他参数（例如临界指数）的测定以及普适类的划分还需进一步完善。

(4) XY 耦合模型、ISing-XY 耦合模型、完全阻挫 XY 模型三者都存在有趣的临界行为，但是关于它们的数值计算和三者之间结果的比较相对其他模型来说是非常匮乏的，值得进一步研究。

(5) 图神经网络从理论上来说可以将不同尺寸的构型一起进行训练，这样生成器

就可以生成不同尺寸下的稳态构型，但是本模型并没有解决该想法所遇到的欠拟合问题，并且 VAE 受限于 σ 与 μ 并且模型的损失比较单位并不是物理性质而是被重构的原构型，所以导致 VAE 具有精度上限，在未来我们可以引入更加强大的图神经网络模型以及更加多样化的权重处理方式来解决卷积层数过多带来的欠拟合与过拟合问题。

参考文献

- [1] 雷桂媛. 关于蒙特卡罗及拟蒙特卡罗方法的若干研究[D]. 浙江大学.
- [2] 陈晓怀, 薄晓静, 王宏涛. 基于蒙特卡罗方法的测量不确定度合成[J]. 仪器仪表学报, 2005(z1): 759-761.
- [3] 何延才. 蒙特卡罗方法在物理学中的应用[J]. 自然杂志, 1982, 5(6): 453-457.
- [4] 邓力李刚, DENG Li L G. 粒子输运蒙特卡罗模拟现状概述[J]. 计算物理, 2010, 27(6): 791.
- [5] 王岩, 尹海丽, 窦在祥. 蒙特卡罗方法应用研究[J]. 青岛理工大学学报, 2006, 27(2): 111-113.
- [6] 王俊峰. 逾渗模型的蒙特卡罗研究[D]. 合肥: 中国科学技术大学, 2013.
- [7] 应和平, 董绍静, 杨清健. 二维 Spin Glass 系统——随机格点 SO (3) 自旋模型的 Monte Carlo 模拟[J]. 计算物理, 1988, 5(1): 107.
- [8] 朱新玲. 马尔科夫链蒙特卡罗方法研究综述[J]. 统计与决策, 2009(21): 151-153.
- [9] 李长隆, 钮可. 马尔科夫链蒙特卡罗方法浅析[J]. 黑龙江科技信息, 2016(26): 48-48.
- [10] 郭茂祖, 王亚东, 孙华梅, 等. 基于 Metropolis 准则的 Q-学习算法研究[J]. 计算机研究与发展, 2002, 39(6): 684-688.
- [11] 李晓寒, 王宗笠, 宁旭. 二维 Ising 模型临界相变的 Monte-Carlo 数值模拟[J]. 安徽大学学报: 自然科学版, 2008, 32(3): 56-59.
- [12] 王松松. Monte Carlo 方法研究点随机集团模型的渗流[D]. 太原理工大学, 2016.
- [13] 刘会平. 低维受限系统热力学, 相变和临界现象的研究[D]. 华中科技大学, 2006.
- [14] 冯乐. 蒙特卡罗方法研究含集团权重伊辛模型和玻色—哈伯德模型: 相变和超固体[D]. 太原理工大学, 2017.
- [15] Hellmann M, Deng Y, Weiss M, et al. Phase transition of a two-dimensional, multiplicatively coupled XY –Potts model[J]. Journal of Physics A: Mathematical and Theoretical, 2009, 42(22): 225001.
- [16] Henley C L. Ordering due to disorder in a frustrated vector antiferromagnet[J]. Physical Review Letters, 1989, 62(17): 2056-2059.

- [17] Nguyen H C, Zecchina R, Berg J. Inverse statistical problems: from the inverse Ising problem to data science[J]. *Advances in Physics*, 2017, 66(3): 197-261.
- [18] Efthymiou S, Beach M J S, Melko R G. Super-resolving the Ising model with convolutional neural networks[J]. *Physical Review B*, 2019, 99(7): 075113.
- [19] Wetzel S J. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders[J]. *Physical Review E*, 2017, 96(2): 022140.
- [20] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015: 3431-3440.
- [21] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [22] Kingma D P, Welling M. Auto-Encoding Variational Bayes[J]. 2013.
- [23] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks[M]. *arXiv*, 2014.
- [24] D'Angelo F, Böttcher L. Learning the Ising model with generative neural networks[J]. *Physical Review Research*, 2020, 2(2): 023266.
- [25] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[J]. *arXiv:1609.02907 [cs, stat]*, 2017.
- [26] De Cao N, Kipf T. MolGAN: An implicit generative model for small molecular graphs[J]. *arXiv:1805.11973 [cs, stat]*, 2018.
- [27] Kipf T N, Welling M. Variational Graph Auto-Encoders[J]. *arXiv:1611.07308 [cs, stat]*, 2016.
- [28] Chandra P, Coleman P, Larkin A I. Ising transition in frustrated Heisenberg models[J]. *Physical Review Letters*, 1990, 64(1): 88-91.
- [29] Zhou J, Cui G, Hu S, et al. Graph neural networks: A review of methods and applications[J]. *AI Open*, 2020, 1: 57-81.
- [30] Waikhom L, Patgiri R. Graph Neural Networks: Methods, Applications, and Opportunities[J]. *ArXiv*, 2021, abs/2108.10733.
- [31] 数据结构综述（严蔚敏版）[EB]//知乎专栏.
- [32] GNN 入门之路: 01.图拉普拉斯矩阵的定义、推导、性质、应用 - 知乎[EB].
- [33] Song Z, Yang X, Xu Z, 等. Graph-based Semi-supervised Learning: A Comprehensive Review[J]. *IEEE transactions on neural networks and learning systems*,

2021, PP.

[34] Wu Z, Pan S, Chen F, 等. A Comprehensive Survey on Graph Neural Networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 32: 4-24.

[35] Blei D M, Kucukelbir A, McAuliffe J D. Variational Inference: A Review for Statisticians[J]. Journal of the American Statistical Association, 2017, 112(518): 859-877.

[36] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[J]. arXiv:1502.03167 [cs], 2015.

[37] Cai T, Luo S, Xu K, et al. GraphNorm: A Principled Approach to Accelerating Graph Neural Network Training[J]. arXiv:2009.03294 [cs, math, stat], 2021.

[38] Luo D, Cheng W, Yu W, et al. Learning to Drop: Robust Graph Neural Network via Topological Denoising[J]. arXiv:2011.07057 [cs], 2020.

[39] Morris C, Ritzert M, Fey M, et al. Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 4602-4609.

攻读学位期间取得的研究成果

[1]丁利利, 张旭, 吴学英, 李媛, 田彦婷。基于 $Y_{0.08}Zr_{0.92}O_{2-\delta}/Er_{0.4}Bi_{1.6}O_3$ 复合电解质的 $Ca_3Co_4O_{9+\delta}$ 氧电极电化学性能。《复合材料学报》, 2022, 002, 13801。(EI)

[2]张旭, 吴萍萍, 丁利利, 田彦婷。基于交替浸渍法对 $La_{0.65}Sr_{0.35}MnO_3$ 氧电极的性能优化。《复合材料学报》, 2021, 002, 13801。(EI)

[3]张旭, 田彦婷, 吴萍萍, 丁利利。溶液浸渍法优化可逆固体氧化物电池氧电极研究进展。《化工新型材料》, 2021.

致 谢

即将落笔，竟有些许不舍。仍记得考研面试时的紧张和初入理工大的欣喜，现将要离开生活三年的地方，我舍不得。三年的学习生涯给予我很多感动，深深留恋这片育人的土地！在各位老师的帮助下，我顺利完成了了解专业、深入专业、实验摸索、分析数据、整理数据与撰写论文的学习过程。在毕业论文收官之时，我在这儿向所有帮助、呵护过我的人们表示诚挚感谢。

特别感谢我的导师曹铃老师，三年的相处时光太过短暂，但他尽全力给予我无微不至的关心。可能我不是您优秀的学生，但您始终是我心里最为尊敬的老师。学习方面，时常听到来自曹老师的鼓励，这也是我前进的动力。科研方面，我深深感受到了曹老师认真严谨的科学态度，从理论计算到药品称量、溶液配制，他从不马虎。生活方面，曹老师时常找我谈话，询问并帮我解决最近遇到的问题和生活中的困难，他用丰富的人生阅历教导我正确的处事原则。日常生活中的他高大帅气，却没有距离感。感谢曹老师三年的关爱和付出，祝老师青春常驻、事事顺遂。

非常感谢课题组的范明明老师。范老师为人严谨，事无巨细，对学生真诚相待、对实验一丝不苟、认真负责的态度深深令我折服。办公室里的他兢兢业业、勤勤恳恳，实验室里的他早出晚归、精益求精。在实验探索中，范老师始终不厌其烦的解决我遇到的问题，亲自带我做实验摸索条件，在我气馁想要放弃的时候，他总是鼓励我，还给予我很多人文关怀，让我深深感动；在数据分析中，范老师认真教我软件作图的细则和注意事项，从坐标刻度的粗细到字号大小的选择再到图片排版，他总是耐心指导；在论文撰写中，范老师与我每日一谈，每一章、每一句、每个标点、句式逻辑他都会细细斟酌，每章节历经十余次的修改，我知道这是他牺牲了无数休息时间和节假日换来的定稿。衷心感谢范老师三年间对我的指导和鼓励，希望您日后工作顺利、家庭幸福。同时我也很感谢物电学院对困难低谷中的我伸出援助之手，感谢课题组的杨致老师、李秀燕老师、刘瑞萍老师、徐利春老师在科研上给我的帮助，老师们的谆谆教导让我铭记终身。

感谢 409 课题组的一切。感谢王亚通师兄、郝宇师兄、赵婉彤师姐、申嘉乐师姐给我的关怀和照顾。感谢同窗李荣对我的照顾有加。感谢 409 的孙庆、霍亦琦、赵文阳、郭宇宏、王岳师妹和梁建新师弟在科研上和生活中的帮助。

感谢舍友张梦、耿鑫、沈洁莲、王佩佩、潘红三年的悉心陪伴，也感谢李雪梅、尚蕾、马晓燕、韩雪婷等小伙伴在我日常学习生活上的相伴，缘分让我们相遇，今后的我们也许各奔前程，但这份友谊永在。

最后，感谢一直默默藏在我身后的避风港和坚强后盾，他们是陪伴着我的爸爸、

妈妈、哥哥、嫂子，共同为我筑起了一道道安心学习的屏障，圆了我的求学梦。

此刻的我，心里也许放空了，但终放不下。毕业，是结束，亦是开始。感恩出现在我生命中的所有人！