

树上自旋系统的快速采样算法

白宗磊¹⁾ 王捍贫^{1),2)} 曹永知¹⁾ 王璐璐¹⁾

¹⁾(北京大学计算机学院高可信软件技术教育部重点实验室 北京 100871)

²⁾(广州大学计算机科学与网络工程学院 广州 510006)

摘要 自旋系统是统计物理学中用来描述微观粒子相互作用的重要框架,其可以描述伊辛模型,硬核模型,玻茨模型等统计物理学中的重要模型;通过求解自旋系统的配分函数可以得出物质的能量、磁矩等物理性质.作为一种重要的图模型,自旋系统在理论计算机、人工智能、概率论等领域中被称作马尔可夫随机场而广泛应用,其可以描述着色问题、图同态问题等图论中的重要问题.对图中的点和边赋予非负权重,自旋系统可以诱导出著名的吉布斯分布;配分函数的近似计算可以归结到对应的吉布斯采样问题,通过吉布斯采样可以求解系统的相关物理性质和统计规律.作为模型的简化,树上的自旋系统受到广泛研究;本文研究树上自旋系统的采样算法,并将其推广到树宽较小的图上.我们的主要工作可以列举如下:对于无外场的伊辛模型,基于节点的两种状态的对称性,可以直接计算出任意节点对应的边缘分布,然后通过简单变量的组合来模拟吉布斯分布.类似地,着色问题和玻茨模型也可以基于状态的对称性用简单变量来模拟吉布斯分布.对于一般的自旋系统,无法保证状态的对称性,我们先递归地计算出所有节点的边缘分布,然后基于这些边缘分布进行采样,并通过简单变量的组合来模拟吉布斯分布.对于普通图,我们引入树宽的概念来度量图与树的相似性,并且基于节点间的独立性将算法推广到树宽为2的伪森林和仙人掌图中.我们的算法仅需要线性时间来得到吉布斯分布中的一个样本,在时间复杂度上优于基于马尔可夫链蒙特卡洛模拟的采样算法.

关键词 着色问题;吉布斯分布;伊辛模型;采样算法;自旋系统

中图法分类号 TP301 DOI号 10.11897/SP.J.1016.2022.02093

Fast Sampling Algorithms for Spin Systems on Trees

BAI Zong-Lei¹⁾ WANG Han-Pin^{1),2)} CAO Yong-Zhi¹⁾ WANG Lu-Lu¹⁾

¹⁾(Key Laboratory of High Confidence Software Technologies (MOE), School of Computer Science, Peking University, Beijing 100871)

²⁾(School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006)

Abstract Spin systems model the interactions between microscopic particles in statistical physics, and provide a framework for many important statistical physical models, such as the Ising model, hard-core model, and Potts model. The partition function of spin systems captures the mean magnetic moment and the mean energy of the systems. The spin systems are also widely used as Markov random field in theoretical computer science, artificial intelligence, and probability theory. They provide a framework for many important models in graph theory, such as colorings, graph homomorphisms, etc. The spin systems induce the famous Gibbs measures when the vertices and edges are assigned with non-negative weights. The approximation of partition function can be reduced to Gibbs samplings, so the physical properties can be obtained from Gibbs samplings. The spin systems on trees are widely studied as a simplification of the systems. In this paper, we study the sampling algorithms for the Gibbs measures on trees, and our contributions can be

收稿日期:2021-09-20;在线发布日期:2022-03-22. 本课题得到国家自然科学基金(61972005,61932001,62172016)资助. 白宗磊,博士研究生,主要研究方向为随机过程、近似计数算法. E-mail: zongleibai@pku.edu.cn. 王捍贫(通信作者),博士,教授,中国计算机学会(CCF)会员,主要研究领域为计算机系统的形式语义和验证、算法和计算复杂性. E-mail: whpxhy@pku.edu.cn. 曹永知,博士,教授,中国计算机学会(CCF)会员,主要研究领域为形式化方法及其应用、隐私性与安全性、不确定性推理等. 王璐璐,硕士,主要研究方向为算法设计与分析.

enumerated as follows: for the Ising model without external fields, the two spins are symmetric, and we simulate the Gibbs measure via the combinations of simple Bernoulli variables. The same result is valid for colorings and the Potts model due to the symmetry of spins. For general spin systems, the symmetry of spins is not valid, so we need to calculate the marginal distributions of the vertices recursively at first. Then based on the distributions, we perform samplings on some simple variables, and simulate the Gibbs measure via delicate combinations of the variables. For general graph, we introduce the treewidth of graph to measure the distance of the graph from tree. Based on the independence of vertices, we extend our algorithms to some graphs with treewidth 2, including pseudo-forest and cactus graphs. Our algorithms only need linear time to obtain a sample from the Gibbs measures on trees, and exhibit better in running time than the algorithms based on Markov chain Monte Carlo method.

Keywords colorings; Gibbs measure; Ising model; sampling algorithm; spin systems

1 引言

自旋系统是来自于统计物理学的重要模型,其对微观粒子之间的相互作用进行建模.在统计物理学中,自旋系统是由点阵组成的,其中每个阵点代表一个微观粒子,每个粒子可能具有不同的自旋状态;自旋系统假设只有最近邻的自旋之间具有相互作用.自旋系统可以描述物质的相变(phase transition)现象,通过计算自旋系统的配分函数(partition function)可以推导出物质的各种物理性质,例如平均能量和平均磁矩等.统计物理学中很多重要的模型都可以描述为自旋系统,包括著名的伊辛模型、硬核模型、玻茨模型等.另外,自旋系统在理论计算机中也有广泛研究,图论中很多重要的问题可以描述为自旋系统,例如著名的着色问题、图同态问题等.在概率论、人工智能等领域中自旋系统也被称作马尔可夫随机场(Markov random field)而广泛研究.

自旋系统可以用图模型来描述,图中的每个节点被赋值为有限个状态之一,边表示系统中相邻节点之间的相互作用.我们用集合 \mathcal{S} 表示自旋系统的状态集,设集合 \mathcal{S} 中共有 q 种状态,节点的权重由其状态决定,用函数 $W: \mathcal{S} \rightarrow \mathbb{R}$ 来表示点权;边上的权重由其两个端点的状态决定,用 $q \times q$ 的对称矩阵 U 来表示边权.给定图 $G=(V, E)$ 和状态集 \mathcal{S} ,赋值 σ 定义为从节点集到状态集的映射,其对应的权重为所有边权与点权的乘积.容易看出,图 G 共有 $q^{|V|}$ 种可能的赋值,这些赋值构成了状态空间 Ω_G .如果所有赋值对应的权重均非负,则可以诱导出状态空间 Ω_G 上的概率分布函数,其中每个赋值出现的概率正

比于其权重,这一分布函数即为著名的吉布斯分布(Gibbs measure).将所有赋值对应的权重加和可得吉布斯分布的归一化因子,即为著名的配分函数(partition function).

自旋系统可以描述统计物理学和图论中的很多重要模型.例如,可以将对称矩阵 U 看作一个无向图 H 的邻接矩阵,当点权函数定义为 $W: \mathcal{S} \rightarrow \{1\}$ 时,赋值 $\sigma: V_G \rightarrow V_H$ 定义了一个从图 G 到 H 的同态函数.如果赋值 σ 的权重非零,那么图 G 中的任意边 (u, v) 经过映射后对应的 $(\sigma(u), \sigma(v))$ 必定也是图 H 中的边,由此可见自旋系统描述了从图 G 到 H 的图同态问题(counting graph homomorphisms).另外,若矩阵 U 主对角线上的取值均为 $\beta \geq 0$,其余元素均取值为1,此时自旋系统描述的是玻茨模型(Potts model).特别地,当参数 $\beta=0$ 时,每个权重非零的赋值对应于图 G 的一种合法着色,即任意边的两端节点状态均不同,此时自旋系统描述的是着色问题.

自旋系统中一个重要特例是状态集 \mathcal{S} 中只有两种状态的情况,称为双态自旋系统.其中大量研究集中于边权矩阵 $U = \begin{bmatrix} \beta & 1 \\ 1 & \gamma \end{bmatrix}$,即边的两端节点取值不同时边权为1的情况.此时点权函数可以表示为向量 $\mathbf{W} := (1, \lambda)$,其中参数 λ 在统计物理学中对应于系统受到的外场.双态自旋系统可以表示很多重要的物理模型,例如,当参数 $\beta=\gamma$ 时,即为著名的伊辛模型(Ising model),伊辛模型是统计物理学中用来描述物质铁磁性的数学模型,随着参数 β 的变化,物质呈现不同的铁磁性质, λ 则描述了物质受到的外场作用.当参数 $\gamma=0$ 时,双态自旋系统描述的是硬核模型(hard-core model),硬核模型是用于描述气

体状态的模型,图中的节点代表空间中的位置,节点取值为 0 表示其对应的位置不存在粒子,取值为 1 表示存在粒子,模型要求相邻位置不能同时存在粒子.在硬核模型中,权重非零的赋值 σ 对应于图 G 的一个独立集(赋值为 1 的节点在独立集中,赋值为 0 的节点不在独立集中).

另一方面,自旋系统是著名的计数约束可满足问题(Counting Constraint Satisfaction Problems, 简记为 #CSP)的特殊实例^[1-2],在 CSP 问题的术语中节点代表变量,变量的取值范围为状态集 S ,边可以看作二元约束函数,函数取值由矩阵 U 给出;另外,自旋系统也可以描述为 Holant 模型^[3-6]的重要实例,对于任意图 $G=(V, E)$,在 Holant 问题中将图 G 中的边抽象为节点,与原有的节点构成二部图 $G'=(V_V, V_E, E')$;图 G' 中,中边表示变量,并为节点关联约束函数,集合 V_V 中节点的关联函数由图 G 中的点权决定,集合 V_E 中节点的关联函数由图 G 中的边权决定.

自旋系统是计数问题和采样问题中的重要框架.自旋系统的采样即为从状态空间 Ω_G 中随机选取一个赋值,使得每个赋值被选中的概率正比于其权重.自旋系统的计数则为计算解的个数,即计算配分函数的值.由于赋值的个数为指数量级,除个别情况外,配分函数的计算都是 NP 难的,因此大量研究集中在配分函数的近似计算上^[7-8].

在统计物理学和人工智能领域,自旋系统上的吉布斯采样是重要的研究方向.另外,根据 Jerrum 等人^[9]给出的从计数问题到采样问题的归约算法,可以将配分函数的近似计算问题归约到吉布斯采样问题,因此基于吉布斯采样可以近似计算配分函数的值.在统计物理学中,对配分函数的计算是物理研究中的关键问题,通过计算配分函数,可以推算出该物理系统几乎所有的物理性质.例如,对配分函数取对数求导即可计算出系统的平均能量(mean energy)和平均磁矩(mean magnetic moment)^[10];配分函数的奇点(singularities)对应的是物理学中的相变(phase transition)现象,在这一点处微小的变动即可产生宏观上肉眼可见的变化.

在人工智能领域^[11-12],自旋系统框架被用来构建结构化概率模型,其使用图模型来表示随机变量之间的相互作用,其中节点代表随机变量,边代表变量间的直接相互作用.图模型为随机变量之间的直接相互作用提供了一个正式的建模框架,这种建模方式极大地减少了模型的参数个数,使得模型可以

使用较少的数据来对深度学习相关的问题进行有效的估计.在深度学习中,图模型上的采样算法受到广泛研究,在很多场景中我们需要从目标分布中快速得到样本;另一方面,可以通过对图模型采样来近似计算模型的配分函数,从而学习其相应的统计规律.

在统计物理学中研究的图一般是正则格 Z^d 的有穷子集,用于模拟真实的物理系统.其中一个重要的特例是贝特格(Bethe lattice) T^b ,即所有非叶节点的孩子数均为 $b \geq 2$ 的完全树.树上的自旋系统是对图上自旋系统的有效简化,作为采样问题中的经典实例,树上的采样算法得到广泛研究^[13-17].经典的采样方法是基于马尔可夫链蒙特卡洛模拟(Markov Chain Monte Carlo, 简记为 MCMC)的算法,马尔可夫链从初始状态出发,经过有限步随机游走,变量的分布逐渐趋近于平稳分布(stationary distribution),从而得到一个样本.通过构建合适的马尔可夫链,使其平稳分布等于吉布斯分布,即可实现吉布斯采样.对于树上的自旋系统,基于 MCMC 的采样算法需要 $O(n \log n)$ 的时间才能趋近于平稳分布,完成一次采样^[17].

本文中我们首先给出一种对树上自旋系统进行采样的新方法,其主要思想来源于统计推断,通过一系列简单变量的组合来模拟吉布斯分布.算法在采样过程中依赖于树上不同分支之间的独立性,可以对不同分支分别采样,从而得到符合整棵树上吉布斯分布的样本;但是对于普通图,这种独立性无法保证,不能直接使用上述方法进行采样.

为了将算法向普通图推广,我们引入树宽(treewidth)的概念来度量图与树的相似性.对于本文介绍的树上自旋系统的采样算法,其有效采样的基础是能够快速计算出节点的边缘分布函数,这一函数的计算主要是基于不同子树之间的相互独立性而求得的;但是普通图中存在圈,节点之间的相关性更加复杂,因此算法的有效性无法保证.我们引入树宽的概念来度量图中节点间的独立性,通俗地讲,普通图对应的树宽度量了该图与树的差距,树宽为 1 的图即为树或森林;树宽越大,图中节点之间的独立性越差,节点对应的分布函数的计算更加复杂.我们将树上的采样算法推广到部分树宽为 2 的特殊图中,包括伪森林(pseudo-forest)和仙人掌图(cactus graph).在算法的推广中最关键的问题是圈的引入,圈中的节点无法保证其两侧节点间的独立性.为此我们需要计算整个圈上节点的分布函数,并基于这一分布函数对圈中的节点进行采样.伪森林可以看

作从圈中的部分节点引申出树,而仙人掌图则是从圈中引申出树或圈。对于圈中节点引申出的树,可以使用树上的采样算法进行采样。对于引申出的圈,注意到仙人掌图中任意两个圈之间只有唯一公共节点,二者之间仍然可以保持较好的独立性;在计算第一个圈的分布函数时,可以将其他圈的配分函数赋值到公共节点上,基于圈中各个节点的分布函数对第一个圈采样后再对其他圈采样即可。如上所述,依次对图中所有圈和圈中引申出的树进行采样即可得到符合吉布斯分布的样本。

简单起见,我们首先研究双态自旋系统的采样算法,然后将算法扩展到多态自旋系统中。我们的主要结果可以列举如下:

(1) 对于无外场的伊辛模型,两种状态具有对称性,因此对树上的任意节点,我们可以基于参数 β 以及父节点的取值求出其边缘分布。基于这些边缘分布函数,可以通过 $O(n)$ 次伯努利采样来模拟吉布斯分布(定理 1)。对于圈中的情况,所有赋值可以按照圈中的同色边数目进行划分,基于状态的对称性可以快速计算出每个部分的概率,并按照此概率采样同色边数目;然后随机选择相应数目的边即可得出圈中节点的赋值。对于伪森林和仙人掌图,两种状态的对称性仍然成立,对于任意圈 C ,其每种赋值的概率不受其他圈和树的影响,可以递归地对图中的圈和树进行采样。

(2) 接下来将算法扩展到树上的双态自旋系统(定理 2),此时两种状态的对称性不再成立,因此我们先递归地计算树中所有节点的边缘分布函数,然后基于这些分布函数进行伯努利采样来模拟吉布斯分布。考虑圈中的情况,节点上的权重可以转移到其邻接边上;将边权看作二阶方阵,通过矩阵运算可以计算圈的配分函数,基于这一配分函数的比例依次对圈中的节点进行采样即可得到圈中节点的赋值。对于伪森林,圈中的部分节点可能引申出树结构,可以先递归地计算出树上的配分函数,并用根节点的外场来模拟两种状态的比值,从而计算圈中的分布函数,并以此为基础对圈中节点采样即可得到符合图中吉布斯分布的样本;对于仙人掌图,两个圈 C_1, C_2 可能有一个公共节点 v_{12} ,使用矩阵乘法可以计算圈 C_2 中节点 v_{12} 上两种状态的比值,用节点 v_{12} 的外场来模拟这一比值,然后计算圈 C_1 的分布函数并对圈 C_1 中的节点进行采样,最后从节点 v_{12} 出发对圈 C_2 中其他节点采样即可;按照上述方法,递归地对仙人掌图中的圈和树进行采样即可得到符合吉布斯

分布的样本。

(3) 对于玻茨模型,其状态具有对称性,同伊辛模型一样,对树中的任意节点,我们可以基于参数 β 和父节点的取值求得该节点的边缘分布函数,然后进行 $O(n)$ 次简单采样即可模拟吉布斯分布(定理 3)。对于圈中的情况,对所有赋值按照同色边数目进行划分,同样可以基于对称性计算出每个部分的概率并按照此概率选取同色边数目;然后随机地选择相应数目的边,被选中的边的两端节点可以收缩为一个节点(即在赋值时取值相同);最后依次随机地选取合法颜色为每个节点着色即可。对于伪森林和仙人掌图,各个状态的对称性仍然成立,递归地对图中的圈和树进行采样即可。

(4) 对于树上的多态自旋系统,我们首先递归地计算所有节点的边缘分布函数,然后通过 $O(n)$ 次简单采样即可模拟其吉布斯分布(定理 4)。对于圈上的情况,多态自旋系统与双态自旋系统的计算方法相同;由于状态数目不同,矩阵的大小会有差别,但使用矩阵乘法依然可以计算其配分函数从而得到各个节点的分布函数。在计算出分布函数后只需要按照该分布依次对节点进行采样并对节点着色即可得到符合吉布斯分布的赋值。对于伪森林和仙人掌图,只需要处理圈上引申出的树和圈即可,同样可以使用节点上的权重来模拟其引申出的结构,然后基于这些数据计算分布函数并依次采样即可。在算法中递归地计算边缘分布需要的时间是 $O(n)$,而模拟吉布斯分布需要 $O(n)$ 次简单采样,因此我们的算法只需要 $O(n)$ 时间即可从吉布斯分布中取得一个样本,要快于基于 MCMC 的采样算法。

如上所述,我们的采样算法适用于树和部分树宽为 2 的图;但是对于更加复杂的图,例如二部图,平面图,无爪图以及区间图等,其树宽更大,与树的相似性更低,这些图中节点间的相关性更加复杂,难以直接使用上述算法计算每个节点的边缘分布函数。另一方面,基于 Weitz^[18] 提出的自回避树模型(self-avoiding walk tree)可以近似计算图中节点的边缘分布函数;其适用范围更大,可以推广到有度数限制的普通图中,但是算法复杂度较高,需要 $O(n^c)$ 时间才能近似计算节点的概率分布,其中 c 是与输入图的最大度数 Δ 相关的常数。相比之下,我们的算法只需要线性时间即可完成一次采样,在时间复杂度上优于基于相关性衰减和基于 MCMC 的算法。另一方面,我们的方法适用的图较少,只对树和部分树宽较小的图有效,如果要推广到更广泛的图

中,需要引入新的工具来处理节点间的复杂关系.

2 相关工作

吉布斯采样在统计物理学与人工智能等领域中占有重要地位,实现吉布斯采样的经典方法是基于马尔可夫链蒙特卡洛模拟(MCMC)^[19-22]的采样算法.算法中马尔可夫链从初始状态出发,在状态空间 Ω_G 中进行随机游走直到趋近于平稳分布.MCMC方法的关键问题在于分析算法的运行时间,即马尔可夫链需要经过多少步才能趋近于平稳分布.对于自旋系统,MCMC算法的时间分析主要有基于正则轨线^[10](canonical path)和基于耦合性^[23](couplings)的分析方法.对于铁磁性伊辛模型($\beta = \gamma > 1$),Jerrum和Sinclair^[10]给出了基于正则轨线的时间分析,马尔可夫链可以在多项式时间内收敛于平稳分布.Goldberg等人^[24]将该方法扩展到了铁磁性双态自旋系统中($\beta\gamma > 1$),Liu等人^[25]将外场 λ 的上界提高到 $\lambda_{\text{MCMC}} := \beta/\gamma$,即对于给定的参数 $\beta \geq \gamma$,当外场满足 $0 < \lambda \leq \lambda_{\text{MCMC}}$ 时,马尔可夫链可以在多项式时间内收敛.对于硬核模型,Luby和Vigoda^[26]给出了基于MCMC的采样算法,并且基于路径耦合的方法(path coupling)证明了当图的最大度数 $\Delta \leq 4$ 时,马尔可夫链可以在多项式时间内收敛.

多态自旋系统的算法分析更加复杂.对于着色问题,Jerrum^[27]给出了马尔可夫链来对图中的合法着色进行采样,基于耦合技术可以证明当图的最大度数为 Δ 且色数 $q \geq 2\Delta + 1$ 时,马尔可夫链可以快速收敛.在此基础上,大量研究^[16,27-29]集中于设计新的马尔可夫链和分析方法,从而将色数限制中的系数减小,Vigoda^[30]使用了更复杂的马尔可夫链将色数的下界降低到 $q \geq 11/6\Delta$,最近Chen等人^[31]打破了 $11/6\Delta$ 的界,给出了一个常数 $\epsilon > 0$,使得当色数 $q \geq 11/6\Delta - \epsilon$ 时即可快速采样.如果对输入图加以限制,可以得到更低的界,对于周长不小于6且度数足够大的图,Dyer等人^[32]证明只要色数 $q \geq 1.49\Delta$ 即可实现快速采样.

玻茨模型可以看作着色问题的扩展,基于MCMC的方法同样广泛应用于玻茨模型的采样算法设计中^[33-34].对于不同的参数 β ,玻茨模型的物理性质以及计算复杂性相差巨大.当参数 $\beta < 1$ 时,相邻节点取值不同时边权更大,此时模型被称作反铁磁性的;当参数 $\beta > 1$ 时,相邻节点取值相同时权重更大,此时模型被称作铁磁性的.对于铁磁性的玻茨模型,其

对应的配分函数与随机聚类模型(Random cluster model)仅相差一个简单系数,基于这一等式,Swendsen和Wang^[35]给出了连接两个模型的马尔可夫链.经典的马尔可夫链在单次操作中只改变一个或几个节点的状态,而Swendsen-Wang算法交替地对节点和边进行操作,单次操作即可改变 $O(n)$ 个节点的状态.由于Swendsen-Wang算法的复杂性,研究主要集中在特殊图上的时间复杂度,比如完全图和二维晶格 \mathbb{Z}^2 .对于完全图,Swendsen-Wang算法的收敛速度^[36-37]已研究清楚,对于二维晶格 \mathbb{Z}^2 ,除少数特殊点外,Swendsen-Wang算法均可快速收敛^[38-39].对于普通图,Ullrich^[40]证明了Swendsen-Wang算法的收敛速度要快于随机聚类模型上的传统马尔可夫链.对于状态数 $q = 2$ 的情况,即无外场的伊辛模型,Guo和Jerrum^[41]证明了Swendsen-Wang算法可以在多项式时间内快速收敛.另一方面,图同态问题可以看作玻茨模型的扩展,图同态问题更加复杂,对其配分函数的研究主要有基于泰勒展开的近似计数算法^[42]和精确计数复杂度的二分性质^[43-44].

基于相关性衰减的方法被广泛用于近似计算自旋系统中节点的边缘分布函数.对于任意图 G ,选定图中任意节点 v ,其边缘分布函数可以由邻接点的分布函数求解出来,即节点 v 的概率可以记作邻接点概率的函数形式,由此可得边缘分布函数的递推关系;这一递推关系可以用树模型来描述,根据此递推关系构建的树即为自回避树.如果图中存在圈,则递推关系会沿着圈回到源节点,此时需要用带赋值的节点来代替源节点并停止这一分支(这一操作是“自回避”的由来).容易看出,自回避树的规模同图中节点的度数相关,度数越大,则自回避树的分支越多,随着递推深度的增加,自回避树的规模可能呈指数增长.要构造有效的算法来计算节点的概率分布,就要在合适的深度停止自回避树的构造,并通过节点间的相关性来度量对自回避树进行截断时引入的误差.当自旋系统符合相关性衰减性质时,即图中其他节点的赋值对节点 v 分布函数的影响随着距离增加而指数减小时,可以在一定深度截断自回避树,得到的模型仍然能有效地近似其概率分布.

基于相关性衰减的性质可以设计自旋系统配分函数的近似算法,也可以用于从吉布斯分布中采样.为了保证近似精度,自回避树的深度需要达到 $O(\log n)$,此时算法可以在多项式时间内实现近似采样,但其计算复杂性较高.对于硬核模型,我们用参数 λ 表示节点上的权重,Weitz^[18]证明了当图的最

大度数不超过 Δ 时,存在与度数 Δ 相关的阈值 λ_c ,对于任意正实数 $\lambda < \lambda_c$,硬核模型满足相关性衰减性质,基于自回避树模型可以在多项式时间内近似计算节点的边缘分布函数.特别地,对于参数 $\lambda = 1$ 的情况,硬核模型中每个赋值对应的是图中的独立集,当图的最大度数 $\Delta \leq 5$ 时,阈值 $\lambda_c > 1$,此时可以在多项式时间内近似计算图中任意节点出现在独立集中的概率.在硬核模型中,阈值 λ_c 与吉布斯分布的唯一性阈值相等,吉布斯分布的唯一性在统计物理学中占有重要地位,其对应的是物理系统的相变现象.对于 Δ -正则树,随着树的深度增加,叶节点的赋值对根节点概率分布的影响逐渐减小,但叶节点数目相应增加;如果树深趋于无穷时叶节点的不同赋值对根节点的影响趋于 0,即根节点的分布函数不受叶节点赋值的影响,那么该系统的吉布斯分布具有唯一性.在参数 $\lambda < \lambda_c$ 时,硬核模型上的吉布斯分布具有唯一性,在参数 $\lambda > \lambda_c$ 时,硬核模型上的吉布斯分布不具有唯一性,在阈值 λ_c 这一点存在相变现象.另一方面,当参数 $\lambda > \lambda_c$ 时,对应的硬核模型配分函数的近似计算是 NP-难的^[45-46],由此可见,硬核模型的可近似性与对应物理系统的相变现象是一致的,可以通过研究模型的可近似性推导出系统的各种物理性质,反之通过物理性质的研究可以推导出对应模型的近似复杂性.

同样的结论适用于反铁磁性双态自旋系统.当参数 $\beta\gamma < 1$ 且图的最大度数不超过 Δ 时,同样可以构造对应的自回避树模型,并构建相应的递推关系来近似计算图中任意节点的分布函数.通过对其递推函数进行分析,可得由参数 β, γ 和最大度数 Δ 决定的阈值 λ_c ,当外场 $\lambda < \lambda_c$ 时,对应的模型满足相关性衰减性质, Li, Lu 和 Yin^[47] 给出了对应配分函数的多项式时间近似算法;当外场 $\lambda > \lambda_c$ 时,系统对应的吉布斯分布唯一性不再成立,此时可以将最大割问题归约到双态自旋系统的近似计数问题, Sly 和 Sun^[48] 证明了其对应配分函数的近似计算是 NP-难的.对于铁磁性双态自旋系统($\beta\gamma > 1$), Guo 和 Lu^[49] 通过对系统的递推函数进行分析,得出了如下结论:

当图的最大度数 $\Delta \leq \Delta_c := \frac{2\sqrt{\beta\gamma}}{\sqrt{\beta\gamma}-1}$ 时,任意外场 $\lambda >$

0 的双态自旋系统均满足相关性衰减性质,对应的配分函数均存在多项式时间近似算法.另外, Guo 和 Lu^[49] 给出了如下猜想:当外场 $\lambda \leq \lambda_c := (\beta/\gamma)^{\Delta_c/2}$ 时,对于任意度数的图均可在多项式时间内近似计算对应的配分函数.对于参数 $\beta > 1 \geq \gamma$ 的情况,这一

猜想已被证实;对于参数 $\gamma > 1$ 的情况,其可近似性仍然是未解决的,其相关性衰减性质未被证实,也未能得到计算复杂性的相关结论.另一方面, Goldberg 和 Jerrum^[50] 证明了铁磁性双态自旋系统可以归约到二部图独立集的近似计数问题(counting Bipartite Independent Sets, 简记为 #BIS 问题);另外,给定参数 β, γ , Liu 等人^[25] 给出了外场下界 λ_0 , 对任意外场 $\lambda > \lambda_0$, 均可将 #BIS 问题近似归约到对应的铁磁性双态自旋系统. #BIS 问题是近似计数问题中的重要开放问题, Dyer 等人^[51-52] 猜想 #BIS 问题既不是 NP-难的,也不存在多项式时间近似算法(在 $RP \neq NP$ 条件下); #BIS 问题类被定义为与 #BIS 问题复杂度相同的问题集合,其为 #P 问题类的重要子类.如上所述,当外场 $\lambda > \lambda_0$ 时,对应的铁磁性双态自旋系统是 #BIS 完全的.

对于多态自旋系统,同样可以构建自回避树模型来建模各个节点之间的递推关系.基于对该递推关系的分析,可以证明相关性衰减的性质,从而得到多项式时间近似算法.基于对玻茨模型的分析, Gamarnik 和 Katz^[53] 给出了色数 $q \geq 2.78\Delta$ 时着色问题的多项式时间近似算法(要求输入图中不存在三角形);基于对多态自旋系统的分析, Lu 和 Yin^[54] 证明了当色数 $q \geq 2.58\Delta + 1$ 时,系统满足相关性衰减性质,可以近似计算图中合法着色的数目.对输入图加以限制,可以得到更好的结果:当图的度数不超过 3 时, Lu 等人^[55] 给出了色数为 4 时合法着色数的近似算法,这一结论对度数不超过 3 的图实现了色数 $q \geq \Delta + 1$ 的可近似性猜想^[27].

基于相关性衰减的方法可以有效地近似自旋系统的配分函数,但是其时间复杂度较高;另一方面,在自旋系统的吉布斯分布满足唯一性条件时,能否有效地进行吉布斯采样是一个著名的开放问题.近期 Anari 等人^[56] 基于马尔可夫链的高维扩展^[57-61] (high-dimensional expanders), 给出了通过分析对应系统的谱独立性(spectral independence)来证明马尔可夫链收敛速度的方法;基于此方法, Chen 等人^[62] 证明了当自旋系统满足相关性衰减性质时,对应的吉布斯采样可在 $O(n \log n)$ 时间内收敛.注意到在相关性衰减性质不满足时(吉布斯分布的唯一性不成立),对应的吉布斯采样是 NP-难的,因此该方法得出的吉布斯采样算法的有效区间上界是紧的.在对吉布斯采样的分析中马尔可夫链被看作在单纯复形(simplicial complex)上的高维随机游走,通过证明单纯复形的 LSE-性质(Local Spectral Expansion),

可以得到马尔可夫链的转移矩阵的最大特征值 (spectral gap) 上界^[57,60], 从而求得其收敛速度. Anari 等人^[56]证明了硬核模型上吉布斯采样算法的有效性, 即在节点权重 $\lambda < \lambda_c$ 时, 硬核模型上的吉布斯分布具有唯一性, 对应的吉布斯采样是快速收敛的; 对于反铁磁性双态自旋系统, 当系统满足相关性衰减性质时, 对应的吉布斯采样可以快速收敛^[63-64]. Feng 等人^[65]将该方法用于证明多态自旋系统上吉布斯采样的快速收敛, Chen 等人^[66]给出了着色问题上吉布斯采样的快速收敛性质. 另外, 基于 Alev 和 Lau^[57]以及 Cryan 等人^[67]对单纯复形上 LSE-性质的证明, Chen 等人^[62]将吉布斯采样的时间复杂度降低到 $O(n \log n)$, 这一结论适用于自旋系统框架下的多种模型, 包括硬核模型, 反铁磁性双态自旋系统, 无三角形图中的着色问题, 以及铁磁性双态自旋系统中相关性衰减算法的有效区域^[65,68]. 特别地, 通过对图上边着色问题的分析, Abdolazimi 等人^[14]证明其对应的单纯复形符合 LSE-性质, 在此基础可以得出: 对度数不超过 Δ 的树, 当色数 $q \geq (1+\epsilon)\Delta$ 时, 对应的吉布斯采样可以快速收敛, 这一结论证明了当色数 $q \geq \Delta$ 时即可近似采样的猜想^[27]在树上的有效性.

本文接下来的结构组织如下: 在第 3 节我们介绍相关的符号和定义, 包括伊辛模型、双态自旋系统、玻茨模型和图同态问题的定义; 然后引入树宽的概念并介绍树宽为常数的伪森林和仙人掌图; 在第 4 节我们研究双态自旋系统的采样算法, 其中在 4.1 节基于状态的对称性给出无外场伊辛模型的采样算法, 在 4.2 节将算法扩展到双态自旋系统中; 在第 5 节, 我们将算法推广到多态自旋系统中, 首先在 5.1 节基于状态的对称性给出玻茨模型的采样算法, 然后在 5.2 节给出多态自旋系统的采样算法, 其中图同态问题作为多态自旋系统的特例同样适用于此算法; 最后在第 6 节我们给出本文的总结.

3 预备知识

本节我们介绍相关符号和定义. 在 3.1 节我们介绍双态自旋系统的定义; 对于多态自旋系统, 我们先在 3.2 节给出玻茨模型的定义, 然后在 3.3 节给出多态自旋系统的定义, 最后在 3.4 节我们给出树宽的定义, 并介绍树宽为常数的伪森林和仙人掌图.

3.1 双态自旋系统

本节我们考虑状态集 $S = \{0, 1\}$ 的情况. 双态自

旋系统的一个实例是由无向图 $G = (V, E)$ 描述的, 其中赋值 $\sigma: V \rightarrow S$ 定义为从节点集 V 到状态集 S 的映射. 双态自旋系统是由参数矩阵 $U = \begin{bmatrix} \beta & 1 \\ 1 & \gamma \end{bmatrix}$ 和外场 λ 确定的, 其中矩阵 U 表示边的权重, 外场 λ 表示赋值为 1 的节点权重. 赋值 σ 的权重定义为所有节点权重和边权的乘积:

$$\omega(\sigma) = \beta^{m_0^\sigma} \gamma^{m_1^\sigma} \lambda^{n_1^\sigma} \quad (1)$$

其中 m_0^σ 和 m_1^σ 分别表示两端节点赋值为 $(0, 0)$ 和 $(1, 1)$ 的边的数目, n_1^σ 表示赋值为 1 的节点数目. 当参数 β, γ, λ 均非负且不全为 0 时, 吉布斯分布 π_G 定义为状态空间 $\Omega_G = S^V$ 上的概率分布函数, 其中赋值 σ 出现的概率为

$$\pi_G(\sigma) = \frac{\omega_G(\sigma)}{Z_G(\beta, \gamma, \lambda)} \quad (2)$$

归一化因子 $Z_G(\beta, \gamma, \lambda)$ 即为著名的配分函数 (partition function):

$$Z_G(\beta, \gamma, \lambda) := \sum_{\sigma \in \Omega_G} \omega(\sigma) = \sum_{\sigma \in \Omega_G} \beta^{m_0^\sigma} \gamma^{m_1^\sigma} \lambda^{n_1^\sigma} \quad (3)$$

双态自旋系统是采样问题和计数问题中的重要框架. 著名的伊辛模型对应的是参数 $\beta = \gamma$ 的情况, 硬核模型对应的是参数 $\gamma = 0$ 的情况.

双态自旋系统可以扩展到带局部变量的情况, 此时不同的边和节点上的权重可以不同. 赋值 σ 的权重 $\omega(\sigma)$ 定义为

$$\prod_{e=(u,v) \in E} \beta_e^{(1-\sigma(u))(1-\sigma(v))} \gamma_e^{\sigma(u)\sigma(v)} \prod_{v \in V} \lambda_v^{\sigma(v)},$$

其中矩阵 $\begin{bmatrix} \beta_e & 1 \\ 1 & \gamma_e \end{bmatrix}$ 表示边 e 的权重, 参数 λ_v 表示节点 v 的外场. 由此诱导出的吉布斯分布 π_G 同样满足对任意赋值 $\sigma \in \Omega_G$ 都有 $\pi_G(\sigma) \propto \omega(\sigma)$. 容易看出, 当边权、点权分别取相同值时, 即对任意边 $e \in E$, 都有 $\beta_e = \beta, \gamma_e = \gamma$; 对任意节点 $v \in V$ 都有 $\lambda_v = \lambda$ 时, 即表示普通的双态自旋系统.

3.2 玻茨模型

考虑状态集 $S = \{0, 1, \dots, q-1\}$ 的情况, 我们用正整数 $q \geq 2$ 表示状态个数. 参数 $\beta \geq 0$ 的玻茨模型定义如下: 给定图 $G = (V, E)$, 赋值 $\sigma: V \rightarrow S$ 定义为从节点集 V 到状态集 S 的函数. 赋值 σ 的权重定义为 $\omega_G(\sigma) = \beta^{m(\sigma)}$, 其中 $m(\sigma)$ 表示赋值 σ 中同色边的数目, 即满足 $\sigma(i) = \sigma(j)$ 的边 (i, j) 的数目. 玻茨模型的状态空间为 $\Omega_G := S^V$, 配分函数:

$$Z_G(\beta) := \sum_{\sigma \in \Omega_G} \omega_G(\sigma)$$

定义为所有权重 $\omega(\sigma)$ 的和. 其对应的吉布斯分布 π_G 满足对任意赋值 $\sigma \in \Omega_G$ 都有:

$$\pi_G(\sigma) := \frac{\omega_G(\sigma)}{Z_G(\beta)}.$$

注意到当任意边 $(i, j) \in E$ 都满足 $\sigma(i) \neq \sigma(j)$ 时, 赋值 σ 对应于图 G 的一种合法着色. 因此当参数 $\beta=0$ 时, 配分函数 $Z_G(\beta)$ 即为图 G 中合法着色的数目. 因此可以将着色问题看作玻茨模型的特殊情况.

3.3 多态自旋系统

本节我们给出多态自旋系统的定义. 对于图 $G=(V, E)$ 和状态集 $\mathcal{S}=\{0, 1, \dots, q-1\}$, 赋值 $\sigma: V \rightarrow \mathcal{S}$ 定义为从节点集到状态集的函数. 赋值 σ 的权重定义为

$$\omega(\sigma) := \prod_{(u, v) \in E} U(\sigma(u), \sigma(v)) \prod_{v \in V} W(\sigma(v)) \quad (4)$$

其中边 $(u, v) \in E$ 的权重由对称矩阵 $U: \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ 给出, 节点 $v \in V$ 的权重由函数 $W: \mathcal{S} \rightarrow \mathbb{R}$ 给出. 如果矩阵 U 和函数 W 中的取值均非负且存在非零值, 那么由自旋系统可以诱导出吉布斯分布 π_G , 其状态空间为 $\Omega_G := \mathcal{S}^V$, 赋值 σ 出现的概率正比于其权重 $\omega(\sigma)$, 归一化因子即为配分函数:

$$Z_G(U, W) := \sum_{\sigma \in \Omega_G} \omega(\sigma) \quad (5)$$

图同态问题可以看作多态自旋系统的特殊情况. 给定图 $G=(V_G, E_G)$ 和图 $H=(V_H, E_H)$, 我们用 $q \times q$ 的对称矩阵 U 来表示图 H 的邻接矩阵. 定义从节点集 V_G 到 V_H 的函数 $\sigma: V_G \rightarrow V_H$, 如果函数 σ 满足对任意边 $(u, v) \in E_G$ 都有 $(\sigma(u), \sigma(v)) \in E_H$, 那么函数 σ 表示从图 G 到图 H 的同态函数, 其权重定义为 $\omega(\sigma) := \prod_{(i, j) \in E_G} U_{\sigma(i), \sigma(j)}$. 令 Ω_G 表示从图 G 到 H 的所有同态函数构成的集合, 图同态问题的配分函数即为

$$Z_G(U) = \sum_{\sigma \in \Omega_G} \omega(\sigma) \quad (6)$$

图论中的很多模型可以表示为图同态问题. 例如, 在 3.2 节定义的玻茨模型对应的是邻接矩阵 U 满足

$$U_{ij} = \begin{cases} \beta, & \text{如果 } i=j \\ 1, & \text{其他情况} \end{cases}$$

的图同态问题. 图 G 上的着色问题可以表示为从图 G 到 q 阶团的图同态问题. 另外, 定义图 $H=(V_H, E_H)$ 如下: 节点集 $V_H = \{u, v\}$, 边集 $E_H = \{(u, v), (v, v)\}$, 则从图 G 到 H 的同态函数对应的是图 G 上的独立集(节点 u 的原象在独立集中, 节点 v 的原象不在独立集中). 另一方面, 图同态问题可以看作计数约束可满足问题(Counting Constraint Satisfaction

Problems, 简记为 #CSP) 的特例, 其对应的是约束函数均为对称二元关系的情况.

3.4 从树到图的扩展

本节我们介绍图论中树宽(treewidth)的概念来度量图与树的相似性, 并基于树宽的概念将算法从树向普通图进行推广. 通俗地讲, 普通图对应的树宽度量了该图与树的差距, 树宽为 1 的图即为树或森林; 树宽越大, 图中节点之间的独立性越差, 节点对应的分布函数的计算更加困难.

首先引入树分解(tree decomposition)的概念来建模图的连通度:

定义 1^[69]. 树分解(tree decomposition). 给定无向图 $G=(V, E)$, 其树分解由树 T 以及 T 中每个节点 z 所关联的集合 $S_z \subseteq V$ 组成. 其中树 T 和 S_z 满足如下条件:

- (1) $\bigcup_{z \in T} S_z = V$, 即对于图 G 中的任意节点 $v \in V$, 树 T 中均存在节点 z , 其关联的集合 S_z 包含节点 v ;
- (2) 对于图 G 中任意一条边 $(u, v) \in E$, 存在树中节点 $z \in T$ 满足 $\{u, v\} \subseteq S_z$;
- (3) 对于图 G 中的任意节点 $v \in V$, 树 T 中所有满足 $v \in S_z$ 的树节点 z 构成一个连通子图(连通子树).

基于树分解可以定义树宽(treewidth)的概念:

定义 2^[69]. 树宽(treewidth). 给定无向图 $G=(V, E)$ 和一个树分解 T , 树 T 的宽(width)定义为 $\max_{z \in V_T} |S_z| - 1$. 图 G 的树宽定义为其所有可能的树分解中树宽最小的树对应的宽.

在定义中使用 $|S_z| - 1$ 作为树宽是为了保证树的树宽为 1. 对于完全图 K_n , 其树宽为 $n-1$; 对于节点个数 ≥ 2 的连通图, 其树宽为 1 当且仅当该图为树; 另一方面, 若图中存在圈, 则其树宽至少为 2. 圈的树宽即为 2, 其树分解可以构造如下: 任意选定圈中节点 v , 将圈中每条边的两个端点加节点 v 构成的集合作为树中的节点, 由此得到的树分解满足定义 1 的性质且树宽为 2. 给定常数 k , 树宽不超过 k 的图称作偏 k -树(partial k -tree), 很多重要的图模型为偏 k -树, 包括仙人掌图(cactus graph)、伪森林(pseudo-forest)、Halin 图、Apollonian 网等, 本文中我们将采样算法应用到伪森林和仙人掌图中.

如图 1 所示, 满足如下性质的图即为伪森林: 图中每个连通分量至多有一个圈, 即任意两个圈之间均不存在公共节点, 且不存在连接两个圈的路径. 特

别地,连通的伪森林称为伪树(pseudo-tree),即为至多有一个圈的连通图. 容易看出,伪森林为稀疏图,由于每个连通分量中至多有一个圈,图中边的数目不会超过节点数目. 根据树宽的定义,圈 C 的树宽为 2,对于从圈 C 中的节点 v 引申出的树,可以将树中每条边的两个端点构成的集合作为节点,即可得到对应的树分解,由此可得伪森林的树宽仍然为 2,与树的相似性较高,可以利用节点间的独立性计算节点的边缘分布函数.

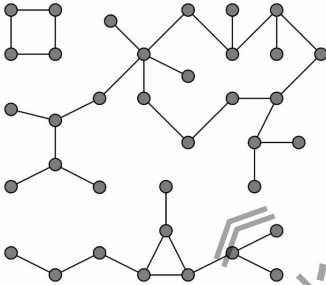


图 1 伪森林(pseudoforest)示例

伪森林在组合中有着广泛应用,可以用于定义拟阵(matroid). 给定图 $G=(V,E)$,可以在图 G 的边集 E 上定义拟阵,边集 E 的任意子集 A 是独立的当且仅当边集 A 的诱导子图构成一个伪森林. 在拟阵中最小的非独立集即为图 G 中包含两个圈的极小的连通子图,这些连通子图也被称作脚踏车(bicycle). 另外,脚踏车可以作为判定图是否为伪森林的标准,给定的图为伪森林当且仅当其所有子图均不是脚踏车.

如图 2 所示,仙人掌图(cactus graph)为连通图,且其中任意两

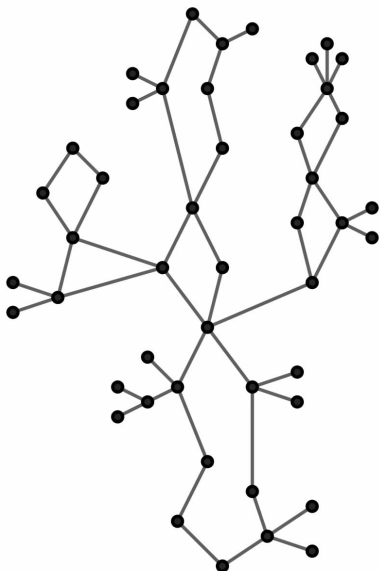


图 2 仙人掌图(cactus graph)示例

个简单圈之间至多只有一个公共节点. 显然伪树符合仙人掌图的定义. 考虑仙人掌图的树宽,注意到圈的树宽为 2,简单起见,设仙人掌图 G 由两个圈 C_1, C_2 组成,两个圈有公共节点 v_{12} ,先按照前述方法分别对圈 C_1 和 C_2 进行树分解,得到的树 T_1, T_2 中存在节点 z_1, z_2 ,其对应的集合 S_{z_1}, S_{z_2} 均包含节点 v_{12} ,连接节点 z_1, z_2 即可得到图 G 的树分解,其对应的树宽仍然为 2. 对于更复杂的仙人掌图,按照上述方法即可递归地构造对应的树分解,且保证其树宽依然为 2. 由此可见,仙人掌图中节点之间的独立性较高,可以将我们的算法推广到仙人掌图中.

对于更加复杂的图,例如二部图、无爪图以及区间图等,其树宽更大,节点间的相关性更复杂,难以直接计算节点对应的分布函数. 对于这一类图,如果图中任意节点 v 受到其他节点的影响随着距离的增加而指数减小,则可以基于 Weitz^[18] 提出的自回避树模型(self-avoiding walk tree)来近似计算节点 v 的分布函数并进行采样;如果系统满足相关性衰减性质,那么可以基于第 2 节中介绍的方法进行采样或近似计算配分函数.

最后,我们定义本文中用到的一些符号. 给定图 $G=(V,E)$,我们用 $n=|V|$ 表示图中节点的个数,用 $m=|E|$ 表示边的条数. 对于正整数 $n>0$,令 $[n]$ 表示集合 $\{1, \dots, n\}$;对于随机变量 X ,其状态空间表示为 Ω_X ,概率分布函数表示为函数 $\mu_X: \Omega_X \rightarrow [0, 1]$,即对任意状态 $a \in \Omega_X$ 都有 $\mu_X(a) := Pr[X=a]$. 为叙述方便,我们同样将函数 $\mu_X = (\mu_X(a))_{a \in \Omega_X}$ 看作由状态 $a \in \Omega_X$ 索引的向量.

4 双态自旋系统

本节我们研究双态自旋系统的采样算法. 在 4.1 节我们基于状态的对称性给出无外场伊辛模型的采样算法,然后在 4.2 节将该算法推广到整个双态自旋系统中.

4.1 无外场伊辛模型

本节我们研究无外场伊辛模型的采样算法. 首先考虑树上的情况,对于树 $T=(V_T, E_T)$,我们用 $\Omega_T := \{0, 1\}^{V_T}$ 表示树上伊辛模型的状态空间. 根据式(1)中给出的定义,对任意赋值 $\sigma \in \Omega_T$,其权重为

$$\omega(\sigma) = \beta^{m(\sigma)} \quad (7)$$

其中 $m(\sigma)$ 表示同色边的条数. 我们用 π_T 表示状态

空间 Ω_T 上的吉布斯分布, $Z_T(\beta) = \sum_{\sigma \in \Omega_T} \omega_T(\sigma)$ 表示配分函数. 对树上伊辛模型的采样如算法 1 所示.

算法 1. 伊辛模型的采样算法.

输入: 树 $T=(V_T, E_T)$; 参数 $\beta > 0$

输出: 赋值 X

将节点集 V_T 标记为 $\{v_1, \dots, v_n\}$;

对变量 Y_1, \dots, Y_n 独立地进行伯努利采样;

FOR $i \leftarrow 1$ to n

$X(v_i) = Y_{i_1} \oplus \dots \oplus Y_{i_{k_i}}$;

ENDFOR

RETURN X ;

首先我们定义函数 $l: V_T \rightarrow [n]$ 给树中的节点加标记, 要求函数 l 满足如下性质: 对任意节点 $v \in V_T \setminus \{r\}$, 其父节点为 f_v , 函数 l 满足 $l(v) > l(f_v)$ (可以使用广度优先搜索算法从根节点到叶节点依次加标记即可满足条件). 注意到根节点标记为 v_1 , 对于任意节点 $v_i \in V_T$, 存在一条从根节点 v_1 到 v_i 的路径 $\mathcal{P}_i = (v_{i_1}, \dots, v_{i_{k_i}})$, 其中 $v_{i_1} = v_1$ 且 $v_{i_{k_i}} = v_i$.

对于图中任意节点 v_i , 定义对应的随机变量 $Y_i \sim \text{Bernoulli}(p)$, 其中 p 表示该节点取值为 0 的概率, 由系统参数决定. 令 $\mathcal{Y} := (Y_1, \dots, Y_n)$ 表示节点集 V_T 对应的变量序列, 其中变量 Y_1 的分布函数 μ_1 满足 $\mu_1(0) = \mu_1(1) = 1/2$, 对于任意 $i \in [n] \setminus \{1\}$, 变量 Y_i 的分布函数 μ_i 满足:

$$\mu_i(0) = \frac{\beta}{1+\beta}; \mu_i(1) = \frac{1}{\beta+1}.$$

在算法中我们对序列 \mathcal{Y} 中的变量独立地进行伯努利采样. 对于概率分布为 $\mu_Y = (p, 1-p)$ 的随机变量 Y , 其伯努利采样可以实现如下: 从 $[0, 1]$ 区间内随机选取实数 u , 并令

$$Y = \begin{cases} 0, & \text{如果 } u \leq p \\ 1, & \text{其他情况} \end{cases} \quad (8)$$

对任意节点 v_i , 存在从根节点 r 到 v_i 的路径 \mathcal{P}_i , 在算法 1 所示的 FOR 循环中, 我们对路径 \mathcal{P}_i 中的节点对应的变量依次进行异或操作来求得节点 v_i 的赋值. 其中 \oplus 表示异或操作, 其形式化定义如下:

$$\begin{aligned} i \oplus j &= 0, \text{ 如果 } i = j; \\ i \oplus j &= 1, \text{ 如果 } i \neq j. \end{aligned}$$

注意到算法 1 的输出 X 满足 $X(v_1) = Y_1$, 而变量 Y_1 的概率分布为 $\mu_1 = (1/2, 1/2)$. 要证明赋值的正确性, 只需要证明吉布斯分布 π_T 满足 $\pi_T^0 = \pi_T^1 = 1/2$ 即可. 我们直接证明这一结论对普通图均成立, 给定图 G 和任意节点 $v \in V$, 定义变量 $\pi_{G,v}^i := \pi_G(\sigma(v) = i)$

表示节点 v 取值为 i 的概率, 则有如下引理.

引理 1. 在无外场伊辛模型中, 给定图 G 和任意节点 $v \in V$, 对应的吉布斯分布 π_G 满足 $\pi_{G,v}^0 = \pi_{G,v}^1 = 1/2$.

证明. 对于图 $G=(V, E)$ 和任意节点 $v \in V$, 定义集合 $\Omega_i := \{\sigma \in \Omega_G \mid \sigma(v) = i\}$ 表示所有将节点 v 映射为 i 的赋值构成的集合, 其中 $i=0$ 或 1 . 定义从 Ω_0 到 Ω_1 的函数 $g := \Omega_0 \rightarrow \Omega_1$ 如下: 对于任意赋值 $\sigma_0 \in \Omega_0$, 对应的赋值 $\sigma_1 := g(\sigma_0)$ 定义为赋值 σ_0 的反转, 即对于任意的节点 $u \in V$ 都有 $\sigma_1(u) = 1 - \sigma_0(u)$. 容易看出函数 g 是从 Ω_0 到 Ω_1 的一一映射. 根据式 (7), 赋值 σ 的权重为 $\omega(\sigma) = \beta^{m(\sigma)}$, 其中 $m(\sigma)$ 表示单色边的数目. 由于 σ_1 是对赋值 σ_0 的反转, 其单色边数目 $m(\sigma_1) = m(\sigma_0)$. 因此, 对于任意赋值 $\sigma_0 \in \Omega_0$, 均有 $\pi_G(\sigma_0) = \pi_G(g(\sigma_0))$, 对集合 Ω_0 中所有赋值求和, 可得 $\pi_G(\Omega_0) = \pi_G(\Omega_1) = 1/2$. 证毕.

由引理 1 可得, 我们可以用变量 Y_1 来模拟节点 v_1 的概率分布. 对于树 T 中的任意节点 $v_i \in V_T$, 令 T_i 表示根节点为 v_i 的子树. 注意到对任意子树 T_i , 两种状态的对称性依然成立, 因此我们可以基于参数 β 和父节点 v_{f_i} 的取值来推算出节点 v_i 的边缘分布. 如算法 1 中所示, 基于这些边缘分布, 可以通过伯努利采样来模拟吉布斯分布 π_T .

定理 1. 给定树 $T=(V_T, E_T)$ 和参数 $\beta > 0$, 算法 1 的输出 X 符合吉布斯分布, 即对任意 $\sigma \in \Omega_T$ 都有 $Pr[X = \sigma] = \pi_T(\sigma)$.

证明. 首先考虑树 T 只包含一个节点的情况, 此时吉布斯分布 $\pi_T = (1/2, 1/2)$, 因此我们有 $\mu_X = \mu_1 = \pi_T$.

考虑有 n 个节点的树 $T=(V_T, E_T)$. 我们用 v_1 表示根节点, 用 $\mathcal{P}_i = (v_{i_1}, \dots, v_{i_{k_i}})$ 表示从根节点 v_1 到节点 v_i 的路径. 令序列 $\mathcal{Y} = (Y_1, \dots, Y_n)$ 表示节点集 V_T 对应的变量, $\Omega_{\mathcal{Y}} := \{0, 1\}^{V_T}$ 表示序列 \mathcal{Y} 的状态空间. 定义从 $\Omega_{\mathcal{Y}}$ 到 Ω_T 的函数 $h: \Omega_{\mathcal{Y}} \rightarrow \Omega_T$ 如下: 对于 $\Omega_{\mathcal{Y}}$ 中的任意状态 \mathbf{y} , 若赋值 $\sigma_{\mathbf{y}} = h(\mathbf{y})$, 则有

$$\begin{aligned} \sigma_{\mathbf{y}}(v_1) &= y_1; \\ \sigma_{\mathbf{y}}(v_i) &= y_{i_1} \oplus \dots \oplus y_{i_{k_i}}, \text{ 对任意 } i \in [n], \end{aligned}$$

其中 $(y_{i_1}, \dots, y_{i_{k_i}})$ 表示路径 $\mathcal{P}_i = (v_{i_1}, \dots, v_{i_{k_i}})$ 中节点对应的变量取值.

对于任意 $i \in [n] \setminus \{1\}$, 令 v_{f_i} 表示节点 v_i 的父节点. 注意到赋值 σ 满足

$$\sigma_{\mathbf{y}}(v_i) = y_{i_1} \oplus \dots \oplus y_{i_{k_i}} = \sigma_{\mathbf{y}}(v_{f_i}) \oplus y_i,$$

因此函数 h 是良定义的. 另一方面, 我们可以由赋值 σ_y 唯一地构造出取值 \mathbf{y} : 令 $y_1 = \sigma_y(v_1)$, 对所有 $i \in [n] \setminus \{1\}$, 令 $y_i = \sigma_y(v_i) \oplus \sigma_y(v_{f_i})$. 因此函数 h 是从 Ω_y 到 Ω_T 的一一映射.

接下来证明吉布斯分布 $\pi_T = \mu_y$. 只需证明对任意的 $\mathbf{y} \in \Omega_y$, 都有 $Pr[\mathcal{Y} = \mathbf{y}] = \pi_T(\sigma_y)$ 即可. 注意到对任意的 $i \in [n] \setminus \{1\}$, 变量 $y_i = \sigma_y(v_i) \oplus \sigma_y(v_{f_i})$, 因此等式 $\sigma_y(v_i) = \sigma_y(v_{f_i})$ 成立当且仅当 $y_i = 0$, 即变量 $y_i = 0$ 时对应的边为同色边. 对边集 E_T 中的所有边求和, 可得 $m(\sigma_y) = m(\mathbf{y})$, 其中 $m(\mathbf{y})$ 表示序列 (y_2, \dots, y_n) 中取值为 0 的变量个数, $m(\sigma_y)$ 表示赋值 σ_y 中单色边的数目. 由于函数 h 是一一映射, 因此我们有

$$Z_{Be}(\beta) := \sum_{\mathbf{y} \in \Omega_y} \beta^{m(\mathbf{y})} = \sum_{\sigma_y \in \Omega_T} \beta^{m(\sigma_y)} = Z_T(\beta).$$

另外 \mathcal{Y} 中变量对应的分布函数 μ_i 满足

$$\mu_1 = (1/2, 1/2),$$

$$\mu_i \propto (\beta, 1), \text{ 对任意 } i \in [n] \setminus \{1\}.$$

因此对任意取值 $\mathbf{y} \in \Omega_y$, 我们有

$$Pr[\mathcal{Y} = \mathbf{y}] = \frac{\beta^{m(\mathbf{y})}}{Z_{Be}(\beta)} = \frac{\beta^{m(\sigma_y)}}{Z_T(\beta)} = \pi_T(\sigma_y) \quad (9)$$

由于函数 h 是一一映射的, 因此式(9)对所有赋值 $\sigma \in \Omega_T$ 均成立, 即对任意赋值 $\sigma_y \in \Omega_T$, 算法 1 的输出 X 均满足 $Pr[X = \sigma_y] = Pr[\mathcal{Y} = \mathbf{y}] = \pi_T(\sigma_y)$. 证毕.

注意到算法 1 中描述的过程本质上是从根节点开始, 依次为树中所有节点赋值的过程, 每个节点的概率分布取决于父节点的取值以及系统的参数 β , 即节点的取值仅由父节点的取值和采样结果决定. 这一过程类似于马尔可夫链的随机游走过程, 马尔可夫链的操作过程可以看作是路径上的赋值过程, 树上的赋值过程可以看作是对马尔可夫链的一种扩展. 从根节点出发依次为子节点赋值, 当赋值过程延伸到叶节点时, 对应路径上的节点均被赋值; 当所有节点均被赋值时, 即完成了一次采样操作. 树上节点的边缘分布容易计算出来, 可以使用上述过程直接赋值, 但是普通图上的分布更加复杂, 难以直接计算出边缘分布函数并进行采样; 普通图上的吉布斯分布对应的是马尔可夫随机场, 可以使用第 2 节所述的相关技术进行分析, 当系统满足相关性衰减性质时, 可以构造马尔可夫链来进行吉布斯采样.

我们简单分析算法 1 的时间复杂度. 注意到算法中需要进行 $O(n)$ 次伯努利采样; 另外在 FOR 循环中, 可以对算法进行修改, 从根节点到叶节点依次进行异或操作: 对于节点 v_i 可以用公式 $\sigma_y(v_i) =$

$\sigma_y(v_{f_i}) \oplus y_i$ 得到赋值 $\sigma_y(v_i)$, 这样只需要进行 $O(n)$ 次异或操作即可. 因此算法 1 的时间复杂度为 $O(n)$. 相比于第 2 节中所述的吉布斯采样算法, 我们的算法在时间复杂度上要优于基于 MCMC 的采样算法, 其对应的马尔可夫链需要 $O(n \log n)$ 的时间才能收敛.

上述基于对称性的算法可以直接推广到 3.4 节中定义的伪森林和仙人掌图中. 首先我们介绍圈上的情况. 对于有 k 个节点的简单圈 $C = (V_C, E_C)$, 其边数和点数均为 k , 可以直接计算出圈上节点每个赋值的权重. 从节点 $v_0 \in V_C$ 出发遍历圈 C 上的节点回到 v_0 , 在遍历过程中同色边的数目可能为 $k, k-2, k-4, \dots, 1(0)$, 即圈上节点的取值只能出现偶数次反转; 每种同色边数目对应的权重分别为 $\beta^k, \beta^{k-2}, \dots, \beta^1(\beta^0)$, 最后的取值取决于 k 的奇偶性. 考虑权重为 β^{k-2} 的所有可能赋值, 只需要从 k 条边中任意选择两条反转边即可, 因此有 $\binom{k}{2}$ 种可能的赋值. 对于其他同色边数的推理相同, 由此可以算出对应的配分函数为

$$Z_C(\beta) = 2 \left(\beta^k + \binom{k}{2} \beta^{k-2} + \binom{k}{4} \beta^{k-4} + \dots \right) \\ = (\beta + 1)^k + (\beta - 1)^k,$$

其中系数 2 来自于节点 v_0 的两种可能取值. 基于此公式可以快速计算出圈 C 对应的配分函数, 并且计算出反转边数量为 $k, k-2, k-4, \dots, 1(0)$ 的赋值对应的权重和, 基于这些数据可以设计对应的采样算法.

推论 1. 给定图 $G = (V, E)$ 和伊辛模型的参数 $\beta > 0$, 如果图 G 为伪森林或仙人掌图, 则存在算法在 $O(n)$ 时间内从吉布斯分布中取得样本 X , 满足对任意 $\sigma \in \Omega_G$ 都有 $Pr[X = \sigma] = \pi_G(\sigma)$.

证明. 首先考虑圈 C 上的采样, 已知其配分函数为 $Z_C := (\beta + 1)^k + (\beta - 1)^k$, 使用如下策略进行采样: 根据引理 1 中所述的状态的对称性, 圈中节点 v_0 的分布函数为 $(1/2, 1/2)$, 基于此为节点 v_0 采样并赋值; 定义权重的前 j 项和为

$$S_j = \beta^k + \binom{k}{2} \beta^{k-2} + \dots + \binom{k}{2j} \beta^{k-2j},$$

从 $[0, 1]$ 区间内随机选取实数 u , 根据 u 的值得到反转边的数目 i ,

$$i = \begin{cases} 0, & \text{如果 } u \leq \frac{2S_0}{Z_C} \\ j, & \text{如果 } \frac{2S_{j-1}}{Z_C} < u \leq \frac{2S_j}{Z_C} \end{cases}.$$

然后从圈 C 上的 k 条边中随机选取 i 条边作为反转边, 从节点 v_0 出发, 遇到非反转边则节点取值不变, 遇到反转边则将对应节点的取值取反即可得到对应的赋值. 容易看出, 反转边数目的选取是依据其分布函数确定的, 在确定数目后随机选择反转边的位置, 这使得所有满足该反转边数目的赋值以均等概率随机出现, 因此得到的赋值符合圈上的吉布斯分布.

对于伪森林, 其可以看作圈的简单推广, 由于每个连通分量是相互独立的, 因此可以依次对每个分量分别采样. 根据 3.4 节中给出的定义, 每个连通分量中至多有一个圈 C . 按照上述方法对圈 C 中的节点进行赋值, 此时赋值 σ_C 出现的概率正比于 $\omega(\sigma_C) = \beta^{m(\sigma_C)}$, 即由 σ_C 中的同色边数目决定. 对于从圈中任意节点 $v \in V_C$ 延伸出的树, 将节点 v 看作其树根, 树上的分布函数符合引理 1 中所述的状态对称性, 根节点取值两种状态的概率相同, 因此树的加入只是对节点 v 的两种取值的权重乘以一个相同的常数, 不影响圈 C 的分布函数; 依然可以按照上述方法对圈 C 中的节点进行赋值, 然后按照算法 1 中的方法对圈 C 上所有节点引申出的树上的节点依次进行赋值. 根据定理 1 中的证明, 该样本符合吉布斯分布.

对于仙人掌图, 其结构更加复杂, 可能有多个圈连通; 另一方面, 仙人掌图要求任意两个简单圈之间至多有一个公共节点, 这保证了各个圈之间的独立性, 可以依次对圈进行采样. 给定仙人掌图 $G=(V, E)$, 为叙述简单起见, 假设图 G 由两个圈 C_1, C_2 组成, 圈 C_1 与 C_2 存在一个公共节点 v_{12} . 首先考虑圈 C_1 的情况, 其配分函数可以按照公式 $Z_C(\beta) = (\beta+1)^k + (\beta-1)^k$ 求得, 根据伊辛模型中两种状态的对称性, 加入圈 C_2 后, 节点 v_{12} 的两种取值的权重比例不变, 对于圈 C_1 的每种赋值, 圈 C_2 的引入相当于为每种赋值的权重乘以一个相同的系数, 因此可以按照反转边数目对圈 C_1 中的节点进行采样, 结果依然符合吉布斯分布. 然后使用同样的方法对圈 C_2 中的节点进行采样, 此时赋值的出发节点为 v_{12} , 在根据配分函数的值随机采样出反转边个数后, 按照顺序对圈 C_2 中节点进行赋值即可得到整个图的样本, 且样本符合吉布斯分布.

对于更加复杂的仙人掌图, 可以按照上述方法从任意圈开始依次对图中的圈进行采样, 最后对圈中节点引申出的树进行采样即可得整个图的赋值; 基于伊辛模型的对称性可以依次证明该样本符合仙人掌图的吉布斯分布. 证毕.

4.2 双态自旋系统的采样算法

本节我们介绍双态自旋系统的采样算法, 根据

3.1 节中给出的定义, 其边权由矩阵 $U = \begin{bmatrix} \beta & 1 \\ 1 & \gamma \end{bmatrix}$ 给出, 点权由外场 λ 确定.

首先考虑树上的情况, 给定树 $T=(V_T, E_T)$, 对于一般的双态自旋系统, 引理 1 中所述的状态对称性不再成立, 无法直接计算节点的边缘分布函数, 因此我们需要先递归地计算出所有节点的边缘分布. 定义 Z_T^i 表示树 T 的根节点 v_1 赋值为 i 的条件下树上的配分函数. 定义比值 $R_T := \frac{Z_T^1}{Z_T^0}$; 概率 $p := Pr[\sigma(v_1) = 1]$ 表示吉布斯分布中根节点 v_1 取值为 1 的概率. 注意到概率 $p = \frac{Z_T^1}{Z_T^1 + Z_T^0}$, 因此有 $R_T = \frac{p}{1-p}$ 以及 $p = \frac{R_T}{1+R_T}$. 由此可得树 T 上的递推函数.

引理 2. 给定树 T 和根节点 r , 令 $\{c_1, \dots, c_d\}$ 表示根节点 r 的子节点, 令 T_i 表示根节点为 c_i 的子树. 则比值 R_T 满足递归式 $R_T = \lambda \prod_{i=1}^d \frac{\gamma R_{T_i} + 1}{\beta + R_{T_i}}$.

证明. 根据式 (3) 中给出的配分函数定义, 可将函数 Z_T 写作子树的配分函数乘积的形式, 即:

$$Z_T^1 = \lambda \prod_{i=1}^d (\gamma Z_{T_i}^1 + Z_{T_i}^0);$$

$$Z_T^0 = \prod_{i=1}^d (Z_{T_i}^1 + \beta Z_{T_i}^0).$$

因此通过简单计算可得:

$$R_T = \frac{Z_T^1}{Z_T^0} = \frac{\lambda \prod_{i=1}^d (\gamma Z_{T_i}^1 + Z_{T_i}^0)}{\prod_{i=1}^d (Z_{T_i}^1 + \beta Z_{T_i}^0)} = \lambda \prod_{i=1}^d \frac{\gamma R_{T_i} + 1}{\beta + R_{T_i}}.$$

证毕.

基于引理 2 给出的递归方程, 我们给出算法 2 来求解比值 R_T . 然后基于等式 $p = \frac{R_T}{1+R_T}$ 可以计算出概率 p .

算法 2. Ratio.

输入: 树 $T=(V_T, E_T)$

输出: 比值 R_T

IF $|V_T| = 1$

 RETURN $R_T = \lambda$;

ENDIF

令 $\{T_1, \dots, T_d\}$ 表示根节点 v_1 的子树;

令 $temp = 1$;

FOR $i \leftarrow 1$ to d :

$x = Ratio(T_i)$;

$h(x) = \frac{\gamma x + 1}{\beta + x}$;

$temp = temp * h(x)$;

ENDFOR

RETURN $R_T = temp * \lambda$;

接下来证明算法 2 给出的结果 \mathcal{R}_T 等于树中根节点对应的比值 R_T .

引理 3. 给定树 T , 算法 2 给出的结果 \mathcal{R}_T 满足等式 $\mathcal{R}_T = R_T = \frac{Z_T^1}{Z_T^0}$.

证明. 我们递归地证明 $\mathcal{R}_T = \frac{Z_T^1}{Z_T^0}$. 如果树 T 仅包含一个节点 r , 则有 $\mathcal{R}_T = \lambda = \frac{Z_T^1}{Z_T^0}$.

考虑包含 n 个节点的树 $T = (V_T, E_T)$, 其根节点为 r , 假设对于任意子树 T_i , 调用算法 2 得到的比值 $\mathcal{R}_{T_i} = \text{Ratio}(T_i)$ 满足 $\mathcal{R}_{T_i} = \frac{Z_{T_i}^1}{Z_{T_i}^0} = R_{T_i}$. 往证对于树 T , 等式 $\mathcal{R}_T = R_T$ 同样成立. 根据引理 2, 我们有:

$$R_T = \frac{Z_T^1}{Z_T^0} = \lambda \prod_{i=1}^d \frac{\gamma R_{T_i} + 1}{\beta + R_{T_i}} = \mathcal{R}_T.$$

因此可以递归地证明算法 2 计算出的比值 \mathcal{R}_T 等于树 T 的比值 R_T . 证毕.

对算法 2 进行简单修改, 在递归地计算比值的过程中将所有子树的比值记录在一个数据结构中, 这样只需调用一次算法 2 即可在 $O(n)$ 时间内计算出所有子树的比值. 基于这些比值数据, 我们可以通过伯努利采样来模拟吉布斯分布 π_T .

接下来介绍采样算法, 首先使用集合 $[n]$ 来标记 V_T 中的节点, 其中根节点标记为 v_1 ; 用 v_{f_i} 表示节点 v_i 的父节点, 且满足 $f_i < i$. 对于子树 T_i , 定义 $p_i := \frac{R_{T_i}}{1 + R_{T_i}}$, 其中 $R_{T_i} = \frac{Z_{T_i}^1}{Z_{T_i}^0}$ 表示由算法 2 求得的子树 T_i 的比值. 注意到子树 T_i 的根节点 v_i 对应的边缘分布取决于树 T_i 的结构和父节点 v_{f_i} 的取值. 因此我们需要定义两个分布函数:

$$\mu_i^0 \propto (\beta, R_{T_i}) \text{ 和 } \mu_i^1 \propto (1, \gamma R_{T_i}) \quad (6)$$

分别对应于父节点 v_{f_i} 取值为 0 和 1 的情况. 定义 n 个随机变量 $\mathcal{Y} := (Y_1, \dots, Y_n)$ 对应于 V_T 中的节点. 在算法 3 中, 我们通过对 \mathcal{Y} 中的变量独立地进行伯努利采样来模拟吉布斯分布 π_T . 伯努利采样可用式(8)中所述的方法来实现.

算法 3. *Samp_2_spin*.

输入: 树 $T = (V_T, E_T)$; 取值 y

输出: 赋值 X

IF $|V_T| = 1$

RETURN $\mathcal{R}_T = \lambda$;

ENDIF

令 $\{T_1, \dots, T_d\}$ 为根节点 v_1 的子树;

令 $X(r) = y$;

FOR $i \leftarrow 1$ to d ;

IF $y = 0$;

对 Y_i 进行伯努利采样, 其分布函数为 μ_i^0 ;

$X_i = \text{Samp}_2_spin(T_i, Y_i)$;

ELSE;

对 Y_i 进行伯努利采样, 其分布函数为 μ_i^1 ;

$X_i = \text{Samp}_2_spin(T_i, Y_i)$;

ENDIF

ENDFOR

RETURN $X = X(r) \cup \bigcup_{i=1}^d X_i$;

为了实现从吉布斯分布 π_T 中采样, 我们首先调用算法 2 计算所有子树的比值 R_{T_i} . 然后对变量 Y_i 进行一次伯努利采样, 其分布函数正比于树 T 的比值:

$$\mu_1 \propto (1, R_{T_1}) \quad (11)$$

最后使用参数 (T, Y_1) 调用算法 3 得到赋值 X .

注意到算法 3 所示的代码实现了树 T 上的递归函数. 我们通过分析第 j 轮递归的情况来证明算法的正确性. 其参数包含子树 T_{v_j} 以及根节点 v_j 的取值 y . 可以看出, 在进入这一轮递归函数之前, 部分节点已被赋值, 令 $S_j \subseteq V_T$ 表示已被赋值的节点, $s_j \in \{0, 1\}^{S_j}$ 表示 S_j 中节点的取值. 令 $C_j := \{c_1, \dots, c_d\}$ 表示根节点 v_j 的子节点集合. 在本轮调用中, 我们对 C_j 中节点对应的变量进行伯努利采样. 则 $S'_j := S_j \cup C_j$ 为本轮调用结束后被赋值的节点集合, 令 $s'_j \in \{0, 1\}^{S'_j}$ 表示其对应的取值.

对集合 $S \subseteq V_T$ 和赋值 σ , 令 $\sigma(S)$ 表示在集合 S 上的部分赋值. 对于任意取值 $s \in \{0, 1\}^S$, 我们用 $\pi_T(\sigma(S) = s)$ 表示在吉布斯分布中, 集合 S 取值为 s 的边缘概率. 假设在第 j 轮调用前, 部分赋值的边缘分布符合吉布斯分布, 如果在调用结束后, 其部分赋值仍然符合吉布斯分布, 那么可以递归地证明在所有调用结束后, 对任意的赋值 $\sigma \in \{0, 1\}^{V_T}$ 都有 $\text{Pr}[X = \sigma] = \pi_T(\sigma)$, 即可证明算法 3 的正确性.

定理 2. 给定树 T 和参数 (β, γ, λ) , 算法 3 的输出 X 符合吉布斯分布, 即对于任意的 $\sigma \in \{0, 1\}^{V_T}$, 都有 $\text{Pr}[X = \sigma] = \pi_T(\sigma)$.

证明. 根据式(2)给出的定义, 对于任意赋值 $\sigma \in \{0, 1\}^{V_T}$, 吉布斯分布 π_T 满足:

$$\pi_T(\sigma) = \frac{\omega(\sigma)}{Z_T(\beta, \gamma, \lambda)} = \frac{\beta^{m_0(\sigma)} \gamma^{m_1(\sigma)} \lambda^{n_1(\sigma)}}{Z_T(\beta, \gamma, \lambda)}.$$

令 μ_X 表示变量 X 的分布函数, 我们递归地证明 $\mu_X = \pi_T$. 根据引理 3, 调用算法 2 得到的比值 \mathcal{R}_T 满足 $\mathcal{R}_T = R_T = \frac{Z_T^1}{Z_T^0}$. 当树 T 只有一个节点时, 分布函数 $\mu_X \propto (1, \lambda)$, 可得 $\mu_X = \pi_T$.

考虑包含 n 个节点的树 $T = (V_T, E_T)$, 设其根节点为 r . 根据引理 3, 对任意 $i \in \{0, 1\}$, 均有:

$$Pr[\sigma(r) = i] = \mu_1(i) = \pi_T(\sigma(r) = i),$$

其中 μ_1 表示式(11)中给出的变量 Y_1 的分布函数. 假设在算法 3 的第 j 轮调用前, 对任意取值 $s_j \in \{0, 1\}^{S_j}$, 其赋值 $\sigma(S_j)$ 满足:

$$Pr[\sigma(S_j) = s_j] = \pi_T(\sigma(S_j) = s_j) \quad (12)$$

往证在第 j 轮调用结束后, 对任意取值 $s'_j \in \{0, 1\}^{S'_j}$, 赋值 $\sigma(S'_j)$ 满足:

$$Pr[\sigma(S'_j) = s'_j] = \pi_T(\sigma(S'_j) = s'_j) \quad (13)$$

在算法 3 的第 j 轮调用中, 我们用 $C_j = \{c_1, \dots, c_d\}$ 表示根节点 v_j 的子节点. 注意到根节点 v_j 被赋值为参数 y , 因此对于根节点为 c_i 的子树 T_i , 我们有:

$$\begin{aligned} Pr[\sigma(c_i) = k \mid \sigma(S_j) = s_j] &= Pr[\sigma(c_i) = k \mid \sigma(v_j) = y] \\ &= \mu_{c_i}^y(k) \end{aligned} \quad (14)$$

其中 $\mu_{c_i}^0 \propto (\beta, R_{T_i})$ 和 $\mu_{c_i}^1 \propto (1, \gamma R_{T_i})$ 表示式(10)中给出的变量 Y_i 的分布函数. 对于部分赋值 s_j , 定义集合 $\Omega_j := \{\sigma \in \{0, 1\}^{V_T} \mid \sigma(S_j) = s_j\}$ 表示满足该部分赋值的所有可能赋值, 令

$$Z_{T, S_j}^j := \sum_{\sigma \in \Omega_j} \omega(\sigma) \quad (15)$$

表示在集合 S_j 赋值为 s_j 的条件下树 T 的配分函数. 定义 $S_{j,i} := S_j \cup \{c_1, \dots, c_i\}$, 令 $s_{j,i}$ 表示 $S_{j,i}$ 的取值. 不失一般性, 我们研究 $s_{j,1}(c_1) = 0, s_{j,1}(v_j) = y = 0$ 的情况, 其他情况的推导相同.

注意到在第 j 轮调用开始前, 子树 T_1 中的节点均未被赋值. 对于集合 Ω_j 中的任意赋值 τ , 定义集合 $[\tau] := \{\sigma \in \Omega_j \mid \text{任意 } v \in V_T \setminus V_{T_1} \text{ 有 } \sigma(v) = \tau(v)\}$. 容易看出集合 $[\tau]$ 中的元素构成等价关系, 等价类 $[\tau]$ 诱导出集合 Ω_j 的等价划分. 选择一组代表元 $\gamma = \{\tau_1, \dots, \tau_t\}$ 使得 $\bigcup_{i=1}^t [\tau_i] = \Omega_j$ 并且对任意的 $a \neq b$ 都有 $[\tau_a] \cap [\tau_b] = \emptyset$. 定义配分函数

$$Z_{T_1, \tau}^i = \sum_{\sigma \in [\tau], \sigma(c_1) = i} \omega(\sigma).$$

则有如下关系:

$$\frac{Z_{T_1, \tau}^1}{Z_{T_1, \tau}^0} = \frac{\sum_{\sigma \in \{0,1\}^{T_1}, \sigma(c_1)=1} \omega(\sigma)}{\beta \cdot \sum_{\sigma \in \{0,1\}^{T_1}, \sigma(c_1)=0} \omega(\sigma)} = \frac{R_{T_1}}{\beta},$$

等式中的系数 β 对应的是边 (v_j, c_1) 上的权重. 根据式(10)和(15)可得:

$$\frac{Z_{T, S_j}^{j,1}}{Z_{T, S_j}^j} = \frac{\sum_{[\tau] \in \gamma} Z_{T_1, \tau}^0}{\sum_{[\tau] \in \gamma} (Z_{T_1, \tau}^1 + Z_{T_1, \tau}^0)} = \frac{\beta}{\beta + R_{T_1}} = \mu_{c_1}^0(0) \quad (16)$$

结合式(12)、(14)和式(16), 我们有:

$$\begin{aligned} Pr[\sigma(S_{j,1}) = s_{j,1}] &= Pr[\sigma(S_j) = s_j] \cdot \mu_{c_1}^y(0) \\ &= \pi_T(\sigma(S_j) = s_j) \cdot \mu_{c_1}^y(0) \\ &= \frac{Z_{T, S_j}^j}{Z_T} \cdot \frac{Z_{T, S_{j,1}}^{j,1}}{Z_{T, S_j}^j} = \frac{Z_{T, S_{j,1}}^{j,1}}{Z_T} \\ &= \pi_T(\sigma(S_{j,1}) = s_{j,1}) \end{aligned} \quad (17)$$

其中第一个等式是由贝叶斯公式得出的.

使用同样的推导过程, 可以证明对节点 c_i 和 v_j 的其他取值均可得到相同结论. 另外, 我们可以递归地对集合 C_j 中的所有节点进行赋值, 从而得到等式(13)中的结论. 因此如果在进入第 j 轮调用前, 等式(12)中的条件 $Pr[\sigma(S_j) = s_j] = \pi_T(\sigma(S_j) = s_j)$ 成立, 那么经过本轮调用后, 等式(13)中的结论 $Pr[\sigma(S'_j) = s'_j] = \pi_T(\sigma(S'_j) = s'_j)$ 也成立. 当算法 3 终止时, 递归调用结束, 此时集合 $S'_n = V_T$, 对于任意的赋值 $s'_n \in \{0, 1\}^{V_T}$, 变量 X 满足:

$$Pr[X = s'_n] = Pr[\sigma(S'_n) = s'_n] = \pi_T(s'_n). \quad \text{证毕.}$$

注意到在算法 2 中计算递归函数需要 $O(n)$ 时间, 在算法 3 中需要执行 n 次伯努利采样. 将所有时间相加可得, 我们需要 $O(n)$ 时间来实现一次吉布斯采样.

接下来我们将上述算法推广到 3.4 节中定义的伪森林和仙人掌图中. 首先考虑圈上的情况. 对于有 k 个节点的简单圈 $C = (V_C, E_C)$, 其边数和点数均为 k . 由于圈上的节点度数均为 2, 因此可以将节点上的权重 λ 分配到其相邻的两条边中, 即节点权重函数变为 $W = (1, \lambda)$, 边上的权重变为矩阵 $U = \begin{bmatrix} \beta & \sqrt{\lambda} \\ \sqrt{\lambda} & \gamma \lambda \end{bmatrix}$; 容易验证, 经过变换后每个赋值的权重不变, 对应的配分函数也不变.

接下来我们介绍经典的使用矩阵乘法来计算圈上配分函数的方法, 首先考虑长度为 2 的路径, 其权重可以由简单的矩阵运算来求得, 即 $U^2 = U \cdot U$, 其中 U^2 表示长度为 2 的路径对应的权重函数, 容易看出矩阵 U^2 中的项 U_{ij}^2 刚好对应于路径两端取值为 (i, j) 时的权重; 对于长度为 k 的路径, 递归地进行上述运算, 可得其对应的权重为 k 个矩阵的积 $U^k = U \cdot \dots \cdot U$. 将该路径上的节点依次标记为 v_0, v_1, \dots, v_k , 然后将两个端点 v_0, v_k 合并为一个节点 v , 得到有 k 个节点的圈, 容易看出, 在节点 v 取值为 0, 1 的条件下圈上的配分函数分别对应于矩阵 U^k 中 U_{00}^k, U_{11}^k 项的值.

推论 2. 给定图 $G = (V, E)$ 和参数 (β, γ, λ) , 如

果图 G 为伪森林或仙人掌图, 则存在算法在 $O(n)$ 时间内从双态自旋系统对应的吉布斯分布中取得样本 X , 对任意 $\sigma \in \Omega_G$ 都有 $Pr[X = \sigma] = \pi_G(\sigma)$.

证明. 首先考虑圈 C 上的采样, 按照 U_{00}^k, U_{11}^k 的比例进行采样即可得到节点 v 的取值. 然后对节点 v 的邻接点 v_1, v_{k-1} 进行赋值, 设 $U^{k-1} = U \cdot \dots \cdot U$ 表示从 v_1 到 v_{k-1} 的路径对应的权重矩阵, 此时节点 v 已赋值, 节点 v_1, v_{k-1} 的分布函数可以由矩阵 U^{k-1} 中的项和边 $(v, v_1), (v, v_{k-1})$ 的权重求得, 将矩阵中的项 U_{ij}^{k-1} 同两条边的权重相乘即可得节点 v_1, v_{k-1} 的四种取值对应的权重比例, 按照这一比例进行采样即可为节点 v_1, v_{k-1} 赋值. 注意到对每个节点的采样均符合圈上的吉布斯分布, 按照上述方法递归地对圈中节点进行采样赋值, 即可得到满足圈上吉布斯分布的样本. 注意到算法中主要的运算集中在矩阵乘法中, 我们可以选择从节点 $v_{\lfloor k/2 \rfloor}$ 出发向两边进行矩阵运算, 每次计算出的矩阵存储在数据结构中, 即可直接用于上述采样, 因此在线性时间内即可完成对圈上节点的采样.

上述方法可以推广到节点权重不同的情况, 此时圈中任意节点 v 的权重 λ_v 依然可以平均分配到其邻边中, 即两条边的权重分别乘以 $\sqrt{\lambda_v}$ 即可 (注意此时边权重矩阵不再保证是对称阵), 此时的采样算法依然有效.

对于伪森林的情况, 其圈上的节点 v 引申出的树用 T_v 表示, 通过算法 2 计算出树 T_v 中根节点 v 的比值 \mathcal{R}_{T_v} 并以此作为节点 v 的外场, 根据引理 2, 比值 \mathcal{R}_{T_v} 等于节点 v 上两种取值的权重比, 因此使用这些数据计算出的矩阵符合整个图的配分函数比值, 得到的样本符合吉布斯分布. 对于伪森林, 先使用算法 2 中的方法对圈中所有引申出的树进行计算并用其比值替换对应节点的外场, 然后使用上述方法对圈中的节点进行采样得到圈中节点的赋值, 在此基础上调用算法 3 对树上的节点采样赋值即可得到伪森林的样本, 且样本符合吉布斯分布.

对于仙人掌图, 同样可以使用矩阵乘法来求解圈中节点的权重. 简单起见, 考虑图 G 中只有两个圈 C_1, C_2 的情况, 两个圈存在一个公共节点 v_{12} . 首先计算圈 C_1 对应的矩阵, 在计算过程中, 遇到节点 v_{12} 时, 则需要先对圈 C_2 进行矩阵运算来求得节点 v_{12} 的比值, 并以此作为节点 v_{12} 的外场来参与圈 C_1 的运算, 由此求得的比值符合图 G 的配分函数, 可以基于这一数据对圈 C_1 中的节点进行采样. 然后考虑圈 C_2 的情况, 从 v_{12} 出发向两边展开依次对节点

进行采样即可. 对于更复杂的仙人掌图, 可以先对圈中引申出的树进行运算, 修改树根节点的外场值, 然后从某个圈出发, 递归地对圈中节点进行矩阵运算并依次进行采样, 最后依据树中节点的分布函数对节点进行采样, 即可得到满足吉布斯分布的样本.

经过简单计算可得, 递归地对圈中节点进行矩阵运算需要的时间是 $O(n)$, 对整个图进行采样需要的时间同样是 $O(n)$, 因此在线性时间内即可取得符合吉布斯分布的样本. 证毕.

5 多态自旋系统

本节我们研究多态自旋系统的采样算法. 在 5.1 节我们基于状态的对称性给出玻茨模型的采样算法. 然后在 5.2 节将采样算法推广到一般的多态自旋系统中.

5.1 玻茨模型的采样算法

本节我们给出玻茨模型的采样算法. 同 4.1 节中的算法相同, 我们基于状态的对称性求得分布函数, 从而使用简单变量的组合来模拟吉布斯分布. 注意到着色问题可以看作玻茨模型的特殊情况, 因此我们的算法可以对树上的合法着色进行采样.

首先考虑树 $T = (V_T, E_T)$ 上的情况, 令 $\mathcal{S} = \{0, 1, \dots, q-1\}$ 表示状态集合, 我们用 π_T 表示树上玻茨模型的吉布斯分布, 用 $\Omega_T := \mathcal{S}^{V_T}$ 表示状态空间. 我们使用 4.1 节中给出的方法将 V_T 中的节点标记为 $\{v_1, \dots, v_n\}$. 定义一组变量 $\mathcal{Y} := (Y_1, \dots, Y_n)$, 序列 \mathcal{Y} 中所有变量的状态空间均为集合 \mathcal{S} . 变量 Y_1 的分布函数定义为 $\mu_1 = (1/q, \dots, 1/q)$; 对于 $i \in [n] \setminus \{1\}$, 变量 Y_i 的分布函数定义为 $\mu_i \propto (\beta, 1, \dots, 1)$. 对变量 Y 的采样可以如下实现: 随机地从区间 $[0, 1]$ 中选择实数 U , 且令

$$Y = \begin{cases} 0, & \text{如果 } U \leq \frac{\beta}{\beta + q - 1} \\ j, & \text{如果 } \frac{\beta + j - 1}{\beta + q - 1} < U \leq \frac{\beta + j}{\beta + q - 1} \end{cases} \quad (18)$$

对于变量 Y_1 , 只需将上式中的参数 β 换成 1 即可. 对树上玻茨模型的采样如算法 4 所示.

算法 4. 玻茨模型的采样算法.

输入: 树 $T = (V_T, E_T)$; 参数 $\beta \geq 0$; 正整数 $q > 0$

输出: 赋值 X

将节点集 V_T 标记为 $\{v_1, \dots, v_n\}$;

对变量 Y_1, \dots, Y_n 独立地进行采样;

令 $X(v_1) = Y_1$;

```

FOR  $i \leftarrow 2$  to  $n$ ;
   $X(v_i) = (X(v_{f_i}) + Y_i) \bmod q$ ;
ENDFOR
RETURN  $X$ ;

```

首先考虑根节点的情况,在算法 4 中我们对变量 Y_1 采样来给根节点赋值,即 $X(v_1) = Y_1$. 令 $\pi_T^i := \pi_T(\sigma(v_1) = i)$ 表示在吉布斯分布中根节点赋值为 i 的概率. 接下来我们基于对称性证明吉布斯分布 π_T 满足 $\pi_T^i = 1/q$. 同引理 1 类似,我们直接证明对普通图的任意节点,上述等式均成立. 对于图 $G=(V, E)$ 和任意给定节点 $v \in V$, 定义 $\pi_{G,v}^i := \pi_G(\sigma(v) = i)$ 表示节点 v 赋值为 i 的概率, 则有如下结论:

引理 4. 对于玻茨模型,其各个状态满足对称性,即对图 $G=(V, E)$ 中的任意节点 $v \in V$ 和状态 $i \in \mathcal{S}$, 都有 $\pi_{G,v}^i = 1/q$.

证明. 对于图 $G=(V, E)$ 和给定的节点 $v \in V$, 定义集合 $\Omega_i := \{\sigma \in \Omega_G \mid \sigma(v) = i\}$ 表示状态空间 Ω_G 中节点 v 取值为 i 的所有赋值. 令 S_q 表示状态集 \mathcal{S} 中元素的所有置换函数组成的集合. 对于任意赋值 σ 和置换函数 $\tau \in S_q$, 定义赋值 σ_τ , 对任意 $v \in V$, 都有 $\sigma_\tau(v) = \tau(\sigma(v))$. 可以看出两个赋值 σ_τ 和 σ 中同色边的数目是相同的. 根据 3.2 节中玻茨模型的定义, 对应的权重 $\omega(\sigma) = \omega(\sigma_\tau)$. 对于赋值 σ , 定义集合 $[\sigma] := \{\sigma_\tau \mid \tau \in S_q\}$ 表示通过 S_q 中的置换函数得到的所有赋值. 可以看出集合 $[\sigma]$ 中的元素满足等价关系, 因此集合 $[\sigma]$ 诱导出集合 Ω_G 的一个划分.

注意到在集合 S_q 中有 $q!$ 个置换函数, 由置换群的对称性可得, 对任意的 $i, j \in \mathcal{S}$, 集合 S_q 中存在 $(q-1)!$ 个置换函数将 i 映射到 j . 因此对任意赋值 $\sigma \in \Omega_G$, 其对应的集合 $[\sigma]$ 满足 $|\Omega_i \cap [\sigma]| = 1/q \cdot |[\sigma]|$. 对集合 Ω_G 的各个划分求和可得, 对任意状态 $i \in \mathcal{S}$ 都有 $\pi_G(\Omega_i) = 1/q$. 证毕.

根据引理 4, 我们可以使用变量 Y_1 来模拟根节点 v_1 上的边缘分布. 注意到所有子树中状态的对称性都是成立的, 因此对任意节点 v_i , 可以通过参数 β 和父节点 v_{f_i} 的取值来计算 v_i 上的边缘分布. 基于这些边缘分布, 我们可以使用简单变量来模拟树上的吉布斯分布 π_T , 如算法 4 所示. 接下来我们证明算法 4 的输出 X 满足 $\mu_X = \pi_T$.

定理 3. 给定树 $T=(V_T, E_T)$ 和参数 $\beta \geq 0$, 算法 4 的输出 X 符合吉布斯分布, 即对任意的赋值 $\sigma \in \Omega_T$ 都有 $Pr[X = \sigma] = \pi_T(\sigma)$.

证明. 如果图 T 中仅包含一个节点, 则其吉布斯分布为 $\pi_T = (1/q, \dots, 1/q)$, 因此可以使用分布

函数为 $\mu_1 = (1/q, \dots, 1/q)$ 的变量 Y_1 来模拟吉布斯分布.

对于有 n 个节点的树 $T=(V_T, E_T)$, 我们使用 n 个独立地随机变量 Y_1, \dots, Y_n 来模拟吉布斯分布. 令 Ω_Y 表示序列 $\mathcal{Y} := (Y_1, \dots, Y_n)$ 的状态空间. 定义从 Ω_Y 到 Ω_T 的函数 $g: \Omega_Y \rightarrow \Omega_T$ 如下: 对变量序列 \mathcal{Y} 的任意取值 \mathbf{y} , 赋值 $\sigma_y := g(\mathbf{y})$ 定义为

$$\sigma_y(v_1) = y_1;$$

$$\sigma_y(v_i) = (\sigma_y(v_{f_i}) + y_i) \bmod q, \text{ 所有 } i \in [n] \setminus \{1\}.$$

接下来我们证明函数 g 是从 Ω_Y 到 Ω_T 的一一映射. 由赋值 $\sigma_y = g(\mathbf{y})$ 可以唯一地构造出对应的取值 \mathbf{y} : 首先令 $y_1 = \sigma_y(v_1)$, 然后对于任意的 $i \in [n] \setminus \{1\}$, 令 $y_i = (\sigma_y(v_i) - \sigma_y(v_{f_i})) \bmod q$. 因此函数 g 是双射函数.

注意到对任意 $i \in [n] \setminus \{1\}$, 其对应的节点 v_i 满足 $\sigma_y(v_i) = \sigma_y(v_{f_i})$ 当且仅当对应变量的取值 $y_i = 0$. 对 E_T 中所有边求和, 可得 $m_0(\mathbf{y}) = m(\sigma_y)$, 其中 $m_0(\mathbf{y})$ 表示序列 (y_2, \dots, y_n) 中取值为 0 的变量数目, $m(\sigma)$ 表示在赋值 σ 中单色边的数目. 因此可得:

$$Z_Y(\beta) := \sum_{\mathbf{y} \in \Omega_Y} \beta^{m_0(\mathbf{y})} = \sum_{\sigma_y \in \Omega_T} \beta^{m(\sigma_y)} = Z_T(\beta).$$

算法 4 的输出 X 有如下性质: 对任意的取值 $\mathbf{y} \in \Omega_Y$, 都有:

$$Pr[X = \sigma_y] = Pr[\mathcal{Y} = \mathbf{y}] = \frac{\beta^{m_0(\mathbf{y})}}{Z_Y(\beta)} = \frac{\beta^{m(\sigma_y)}}{Z_T(\beta)} = \pi_T(\sigma_y) \quad (19)$$

注意到函数 g 是双射的, 因此对于任意的赋值 $\sigma \in \Omega_T$, 同样有 $Pr[X = \sigma] = \pi_T(\sigma)$. 结论得证. 证毕.

同伊辛模型相同, 上述基于对称性的算法可以直接推广到 3.4 节中定义的伪森林和仙人掌图中. 首先考虑圈上的采样, 通过计算圈上的配分函数, 我们使用同色边数目对所有赋值进行划分, 然后通过随机选取同色边和随机选取颜色来保证每种赋值着色概率相同.

对于有 k 个节点的简单圈 $C=(V_C, E_C)$, 其边数和点数均为 k . 从圈中节点 $v_0 \in V_C$ 出发遍历圈上的节点回到 v_0 , 形成序列 $v_0, v_1, \dots, v_{k-1}, v_0$. 首先考虑圈上合法着色的数目, 设长度为 k 的圈合法着色数目为 A_k , 按照上述序列依次对圈中节点着色, 节点 v_0 有 q 种颜色可选, v_1, \dots, v_{k-2} 均各有 $q-1$ 种颜色可选. 考虑 v_{k-1} 的着色, 如果节点 v_0 与 v_{k-2} 颜色相同, 那么 v_{k-1} 有 $q-1$ 种颜色可选, 此时可以把节点 v_0 到 v_{k-2} 的着色看成对长度为 $k-2$ 的圈进行着色; 否则 v_{k-1} 有 $q-2$ 种颜色可选, 此时可以把节点 v_0 到 v_{k-2} 的着色看成对长度为 $k-1$ 的圈进行着色.

因此有如下递推关系:

$$A_k = (q-1)A_{k-2} + (q-2)A_{k-1},$$

注意到 $A_2 = q(q-1)$, $A_3 = q(q-1)(q-2)$, 经过简单计算可得:

$$A_k + A_{k-1} = q(q-1)^{k-1},$$

对上述递推式进行交叉加减操作可得: 当 k 为奇数时有:

$$A_k + A_2 = (q-1)^k + (q-1)^2,$$

此时有 $A_k = (q-1)^k - (q-1)$; 当 k 为偶数时有:

$$A_k - A_2 = (q-1)^k - (q-1)^2,$$

此时有 $A_k = (q-1)^k + (q-1)$.

现在考虑玻茨模型, 在对圈上的节点进行遍历过程中同色边的数目可能是 $0, 1, \dots, k-2, k$. 如果圈上存在 i 条同色边, 那么可以把同色边收缩为一个点, 满足条件的赋值可以看作长度为 $k-i$ 的圈的合法着色, 因此共有 A_{k-i} 种可能的赋值, 每个赋值的权重为 β^i , 由此可以直接计算圈上的配分函数:

$$Z_C(\beta) = q\beta^k + \sum_{i=0}^{k-2} A_{k-i}\beta^i.$$

基于此公式可以快速计算出圈 C 对应的配分函数, 并且计算出同色边数量为 i 的赋值所占比重.

在计算出配分函数后, 可以使用如下策略进行采样: 将区间 $[0, 1]$ 按照 k 种同色边数目对应的权重划分成 k 份, 从 $[0, 1]$ 区间内随机选取实数 U , 根据 U 所在的区间得到对应的整数 i , 然后从圈 C 上的 k 条边中随机选取 i 条边作为同色边. 根据引理 4 所述的对称性, 可以从集合 S 中为节点 v_0 随机选取一种颜色 (若 v_0 的邻边被选中为同色边, 则将其邻接点赋值为相同颜色), 从节点 v_0 出发按顺序依次为圈上节点着色: 对于节点 v_i , 如果其前邻边为同色边, 那么节点 v_i 着色与 v_{i-1} 相同; 否则从节点 v_i 所有合法着色中随机选取一种并着色. 容易看出, 在选定同色边数目时, 选定的数目 i 符合对应赋值的权重比例; 然后随机地选取 i 条同色边, 这保证了满足条件的每种赋值出现的概率相同; 在对节点着色过程中, 每次都从合法着色中随机选取一种, 保证了每种赋值出现的概率相同, 因此该着色方案得到的样本满足圈上的吉布斯分布.

推论 3. 给定图 $G=(V, E)$ 和参数 β , 如果图 G 为伪森林或仙人掌图, 则存在算法在 $O(n)$ 时间内从玻茨模型对应的吉布斯分布中取得样本 X , 满足对任意 $\sigma \in \Omega_G$ 都有 $Pr[X=\sigma] = \pi_G(\sigma)$.

证明. 上述方法可以实现圈上玻茨模型的采样算法, 对于伪森林, 可以对每个连通分量依次进行

采样. 由于每个连通分量中至多有一个圈 C . 按照上述方法对圈 C 中的节点进行赋值, 此时赋值 σ_C 出现的概率正比于 $\omega(\sigma_C) = \beta^{m(\sigma_C)}$. 对于从圈中任意节点 $v \in V_C$ 引申出的树, 将节点 v 看作树根, 其子树的赋值均符合引理 4 所述的对称性, 按照算法 4 中的操作对树上的节点进行赋值, 即可得到符合吉布斯分布的样本; 对圈 C 上所有节点引申出的树依次采样, 即可得到整个图的样本. 按照定理 3 中的证明, 该样本符合吉布斯分布.

对于仙人掌图, 其中任意两个简单圈之间至多有一个公共节点, 同 4.1 节中伊辛模型的情况相同, 可以依次对圈进行采样. 简单起见, 假设图 G 由两个圈 C_1, C_2 组成, 圈 C_1 与圈 C_2 存在一个公共节点 v_{12} . 首先考虑圈 C_1 的情况, 其配分函数可以按照公式 $Z_C(\beta)$ 快速计算出来. 根据玻茨模型中 q 种状态的对称性, 加入圈 C_2 后, 节点 v_{12} 各种状态的权重比例不变, 即对于圈 C_1 的每种着色, 圈 C_2 的引入相当于为其对应的权重乘以一个相同的系数, 因此按照公式 $Z_C(\beta)$ 计算出的配分函数直接对圈 C_1 中的节点进行采样依然符合吉布斯分布. 在对圈 C_1 中节点进行采样后, 使用同样的方法对圈 C_2 中的节点进行采样, 此时赋值的出发节点为 v_{12} , 在根据配分函数的值随机得出同色边个数后, 按照顺序对圈 C_2 中节点进行着色即可得到整个图的样本, 且样本符合吉布斯分布.

对于一般的仙人掌图, 按照上述方法依次对图中的圈进行采样赋值即可, 最后对圈中节点引申出的树进行着色即可得整个图的赋值. 基于对称性可以证明新圈的引入不影响圈中的分布函数, 依次对圈采样依然符合吉布斯分布. 证毕.

5.2 多态自旋系统的采样算法

本节我们给出多态自旋系统的采样算法. 本节的算法为 4.2 节中双态自旋系统采样算法的扩展. 在 3.3 节中定义的图同态问题可以看作本节中算法的特殊情况.

根据式 (4) 的定义, 我们使用 $q \times q$ 阶对称矩阵 $U: S \times S \rightarrow \mathbb{R}$ 表示边的权重, 函数 $W: S \rightarrow \mathbb{R}^+$ 表示节点的权重. 首先考虑树 $T=(V_T, E_T)$ 上的情况, 可以递归地计算根节点的边缘分布函数. 如算法 5 所示, 我们递归地计算出序列 R_T 使得对任意状态 $i \in S$, 都有:

$$R_T(i) = Z_T^i := Z_T(\sigma(r) = i),$$

其中 Z_T^i 表示根节点 r 赋值为 i 的条件下树 T 的配分函数. 对于叶节点, 令序列 $R_T = W$ 表示其权重函

数. 对于 $|V_T|=n$ 的情况, 令 $C=\{c_1, \dots, c_d\}$ 表示根节点 r 的子节点集合, T_j 表示根节点为 c_j 的子树. 对于子树 T_j , 令 $R_{T_j}(i):=Z_{T_j}^i$ 表示根节点 v_j 赋值为 i 的条件下子树 T_j 的配分函数. 通过递归地调用函数 $\text{RatioMP}(T_j)$ 可以计算序列 R_{T_j} 的值.

算法 5. RatioMP .

输入: 树 $T=(V_T, E_T)$

输出: 序列 \mathcal{R}_T

IF $|V_T|=1$

RETURN $\mathcal{R}_T=W$;

ENDIF

令 $\{T_1, \dots, T_d\}$ 为根节点 r 的子树;

FOR $j \leftarrow 1$ to d :

$x_j = \text{RatioMP}(T_j)$;

ENDFOR

FOR $i \leftarrow 0$ to $q-1$:

$temp=1$;

FOR $i \leftarrow 1$ to d :

$h = \sum_{t=0}^{q-1} U_{it} x_j(t)$;

$temp = temp * h$;

ENDFOR

$\mathcal{R}_T(i) = temp * W(i)$

ENDFOR

RETURN \mathcal{R}_T ;

接下来我们证明基于算法 5 的输出序列 \mathcal{R}_T 可以计算出吉布斯分布中根节点 r 的边缘分布.

引理 5. 算法 5 的输出序列 \mathcal{R}_T 正比于根节点的边缘分布, 即对任意状态 $i \in \mathcal{S}$, 都有 $\mathcal{R}_T(i) = Z_T^i$.

证明. 我们递归地证明本引理. 对于只有一个节点的树 T , 易得对任意状态 $i \in \mathcal{S}$, 都有 $\mathcal{R}_T(i) = W(i) = Z_T^i$.

对于有 n 个节点的树 $T=(V_T, E_T)$, 令 $\mathcal{R}_{T_j} = \text{RatioMP}(T_j)$ 表示算法 5 对其子树 T_j 运算得到的序列. 假设对任意的子树 T_j , 算法 5 的输出 \mathcal{R}_{T_j} 均满足: 对任意 $i \in \mathcal{S}$, 都有:

$$\mathcal{R}_{T_j}(i) = R_{T_j}(i) = Z_{T_j}^i \quad (20)$$

往证树 T 对应的序列 \mathcal{R}_T 同样满足条件.

根据等式(4)和(5)中定义的配分函数, 我们有:

$$R_T(i) = Z_T^i = \sum_{\substack{\sigma \in \Omega_T \\ \sigma(r)=i}} \omega(\sigma) = W(i) \cdot \prod_{j \in [d]} \sum_{t=0}^{q-1} U_{it} Z_{T_j}^t \quad (21)$$

其中 U_{it} 表示边 (r, c_j) 上的权重, $Z_{T_j}^t$ 表示在节点 c_j 赋值为 t 的条件下子树 T_j 的配分函数. 根据式(20)和(21)可得, 经过算法 5 中的运算, 其输出序列 \mathcal{R}_T 满足对任意的 $i \in \mathcal{S}$ 都有 $\mathcal{R}_T(i) = Z_T^i$. 证毕.

通过对算法 5 简单修改, 可以在递归调用的过程中将任意子树 T_i 对应的序列 R_{T_i} 保存在一个数据结构中, 这样只需要遍历一次就可以得到所有子树的权重序列, 算法的运行时间为 $O(n)$.

根据引理 5, 我们可以用算法 5 求出的序列 R_T 来计算根节点 r 取值为 i 的概率:

$$p_r^i := \Pr[\sigma(r)=i] = \frac{R_T(i)}{\sum_{j=0}^{q-1} R_T(j)} \quad (22)$$

定义 n 个简单变量 $\mathcal{Y} := (Y_1, \dots, Y_n)$ 来模拟吉布斯分布 π_T . 注意到在吉布斯分布中, 节点 v_j 的取值对应的边缘分布函数是由子树 T_j 的结构和其父节点 v_{f_j} 的取值决定的. 因此对于父节点 v_{f_j} 取值为 $i \in \mathcal{S}$ 的情况, 我们定义 μ_j^i 来表示变量 Y_j 的分布函数. 对任意的 $j \in [n] \setminus \{1\}$, 分布函数 μ_j^i 定义为

$$\mu_j^i(k) := \frac{U_{ik} R_{T_j}(k)}{\sum_{t=0}^{q-1} U_{it} R_{T_j}(t)} \quad (23)$$

对于根节点 r , 只需定义分布函数 $\mu_1 \propto R_T$ 即可.

对于分布函数为 μ 的变量 Y , 其采样实现如下: 从区间 $[0, 1]$ 中随机采样正实数 U , 然后令

$$Y = \begin{cases} 0, & \text{如果 } \frac{\mu(0)}{\sum_{i=0}^{q-1} \mu(i)} \\ \dots \\ j, & \text{如果 } \frac{\sum_{i=0}^{j-1} \mu(i)}{\sum_{i=0}^{q-1} \mu(i)} < U \leq \frac{\sum_{i=0}^j \mu(i)}{\sum_{i=0}^{q-1} \mu(i)} \end{cases} \quad (24)$$

为了实现吉布斯采样, 我们首先调用算法 5 来计算所有节点 $v \in V_T$ 对应子树的权重序列 R_{T_v} . 然后对随机变量 Y_1 进行采样, 其分布函数为 μ_1 . 最后将树 T 和变量 Y_1 作为参数递归地调用算法 6 得到赋值 X .

算法 6. SampMP

输入: 树 $T=(V_T, E_T)$; 取值 y

输出: 赋值 X

令 c_1, \dots, c_d 表示根节点 r 的子节点, T_i 表示根节点为 c_i 的子树;

令 $X(r) = y$;

FOR $i \leftarrow 1$ to d :

对变量 Y_i 进行采样, 其分布函数为 μ_i^y ;

$X(v_i) = \text{SampMP}(T_i, Y_i)$;

ENDFOR

RETURN $X = X(r) \cup \bigcup_{i=1}^d X_i$;

对于赋值 σ 和任意节点集 $S \subseteq V_T$, 令 $\sigma(S)$ 表示

在集合 S 上的部分赋值. 对于状态空间 \mathcal{S}^S 中任意取值 s , 令 $\pi_T(\sigma(S)=s)$ 表示在吉布斯分布中将 S 中的节点赋值为 s 的概率. 注意到算法 6 中的伪代码描述了树 T 上的递归算法. 对于第 j 轮递归调用, 设其参数为子树 T_{v_j} 和赋值 y . 在第 j 轮递归调用之前, 存在节点集 $S_j \subseteq V_T$ 已被赋值, 令 $s_j \in \mathcal{S}^{S_j}$ 表示 S_j 中节点的取值. 在子树 T_{v_j} 中, 令 $C_j := \{c_1, \dots, c_d\}$ 表示根节点 v_j 的子节点. 在本轮调用中, 我们对 C_j 中节点对应的变量进行采样. 令 $S'_j := S_j \cup C_j$ 表示本轮调用结束后被赋值的节点集, 其对应的取值为 s'_j . 当算法 6 终止时, 若对任意的赋值 $\sigma \in \mathcal{S}^{V_T}$ 都有 $Pr[X=\sigma] = \pi_T(\sigma)$, 则算法 5 和 6 实现了吉布斯采样.

定理 4. 给定树 T 、边权重矩阵 U 和点权函数 W , 算法 6 的输出 X 符合吉布斯分布, 即对于任意的 $\sigma \in \mathcal{S}^{V_T}$, 都有 $Pr[X=\sigma] = \pi_T(\sigma)$.

证明. 我们通过对算法 6 中函数的递归深度进行归纳来证明 $\mu_X = \pi_T$. 对于 $|V_T|=1$ 的情况, 注意到分布函数 $\mu_X \propto W$, 立即可得 $\mu_X = \pi_T$.

考虑包含 n 个节点的树 $T = (V_T, E_T)$, 设其根节点为 r . 根据引理 5, 对所有 $i \in \mathcal{S}$, 均有:

$$Pr[\sigma(r)=i] = \mu_1(i) = \pi_T(\sigma(r)=i) \quad (25)$$

其中 $\mu_1 \propto R_T$ 表示变量 Y_1 的分布函数. 假设在算法 6 的第 j 轮调用之前, 其在 S_j 上的部分赋值 $\sigma(S_j)$ 满足对任意取值 $s_j \in \mathcal{S}^{S_j}$, 都有:

$$Pr[\sigma(S_j)=s_j] = \pi_T(\sigma(S_j)=s_j) \quad (26)$$

往证在本轮调用结束后, 在集合 S'_j 上的部分赋值 $\sigma(S'_j)$ 满足对任意的取值 $s'_j \in \mathcal{S}^{S'_j}$, 都有:

$$Pr[\sigma(S'_j)=s'_j] = \pi_T(\sigma(S'_j)=s'_j) \quad (27)$$

对于取值 s_j , 定义集合

$$\Omega_j := \{\sigma \in \mathcal{S}^{V_T} \mid \sigma(S_j) = s_j\},$$

表示所有满足 $\sigma(S_j) = s_j$ 的赋值集合, 令

$$Z_{T,S_j}^j := \sum_{\sigma \in \Omega_j} \prod_{v \in V} W(\sigma(v)) \prod_{(u,v) \in E_T} U_{\sigma(u)\sigma(v)} \quad (28)$$

表示在 S_j 赋值为 s_j 的条件下树 T 对应的配分函数.

对于以 c_i 为根节点的子树 T_i , 由式(23)可得:

$$\begin{aligned} Pr[\sigma(c_i)=k \mid \sigma(S_j)=s_j] &= Pr[\sigma(c_i)=k \mid \sigma(v_j)=y] \\ &= \frac{U_{yk} R_{T_i}(k)}{\sum_{t=0}^{q-1} U_{yt} R_{T_i}(t)} = \mu_{c_i}^y(k) \quad (29) \end{aligned}$$

定义 $S_{j,i} := S_j \cup \{c_1, \dots, c_i\}$, 令 $s_{j,i}$ 表示 $S_{j,i}$ 的取值. 不失一般性, 我们研究 $s_{j,1}(c_1) = 0, s_{j,1}(v_j) = y = 0$ 的情况, 其他情况的推导相同.

注意到在第 j 轮调用开始前, 子树 T_1 中的节点均未被赋值. 对于集合 Ω_j 中的任意赋值 τ , 定义集合

$$[\tau] := \{\sigma \in \Omega_j \mid \sigma(v) = \tau(v) \text{ 对所有 } v \in V_T \setminus V_{T_1}\}.$$

容易看出同一集合 $[\tau]$ 中的元素构成等价关系, 集合类 $[\tau]$ 诱导出集合 Ω_j 的等价划分. 选择一组代表元 $\gamma = \{\tau_1, \dots, \tau_r\}$ 使得 $\bigcup_{i=1}^r [\tau_i] = \Omega_j$ 并且对任意的 $a \neq b$ 都有 $[\tau_a] \cap [\tau_b] = \emptyset$. 定义函数

$$Z_{T_1, \tau}^i = \sum_{\sigma \in [\tau], \sigma(c_1)=i} \omega(\sigma).$$

则对于任意的 $\tau \in \Omega_j$, 对应的函数 $Z_{T_1, \tau}^i$ 有如下关系:

$$\frac{Z_{T_1, \tau}^0}{\sum_{i \in \mathcal{S}} Z_{T_1, \tau}^i} = \frac{\sum_{\sigma \in [\tau], \sigma(c_1)=0} \omega(\sigma)}{\sum_{\sigma \in [\tau]} \omega(\sigma)} = \frac{U_{00} R_{T_1}(0)}{\sum_{i=0}^{q-1} U_{0i} R_{T_1}(i)},$$

其中系数 U_{0i} 表示的是边 (v_j, c_1) 上的权重.

根据式(28), 可得:

$$\begin{aligned} \frac{Z_{T,S_{j,1}}^{s_{j,1}}}{Z_{T,S_j}^j} &= \frac{\sum_{[\tau] \in \Gamma} Z_{T_1, \tau}^0}{\sum_{[\tau] \in \Gamma} \sum_{i \in \mathcal{S}} Z_{T_1, \tau}^i} = \frac{U_{00} R_{T_1}(0)}{\sum_{i=0}^{q-1} U_{0i} R_{T_1}(i)} = \mu_{c_1}^0(0) \quad (30) \end{aligned}$$

其中 $\mu_{c_1}^0$ 表示的是节点 c_1 对应的分布函数, 其定义如式(23)所示, 注意此时节点 v_j 的取值为 0, 因此对应的分布函数为 $\mu_{c_1}^0$. 结合式(26)、(29)和式(30), 我们有:

$$\begin{aligned} Pr[\sigma(S_{j,1})=s_{j,1}] &= Pr[\sigma(S_j)=s_j] \cdot \mu_{c_1}^0(0) \\ &= \pi_T(\sigma(S_j)=s_j) \cdot \mu_{c_1}^0(0) \\ &= \frac{Z_{T,S_j}^j}{Z_T} \cdot \frac{Z_{T,S_{j,1}}^{s_{j,1}}}{Z_{T,S_j}^j} = \frac{Z_{T,S_{j,1}}^{s_{j,1}}}{Z_T} \\ &= \pi_T(\sigma(S_{j,1})=s_{j,1}) \quad (31) \end{aligned}$$

其中第一个等式是由贝叶斯公式求得的.

对于节点 c_1 和 v_j 的不同取值, 使用上述推导过程, 均可得到相同的结论. 另外, 我们可以递归地对集合 C_j 中的所有节点进行赋值, 从而得到等式(27)中的结论. 因此如果在进入第 j 轮调用前, 等式(26)中的条件 $Pr[\sigma(S_j)=s_j] = \pi_T(\sigma(S_j)=s_j)$ 成立, 则经过本轮调用后, 等式(27)中的结论 $Pr[\sigma(S'_j)=s'_j] = \pi_T(\sigma(S'_j)=s'_j)$ 也成立. 当算法 6 终止时, 递归函数结束, 此时集合 $S'_n = V_T$, 对应的输出 X 满足: 对任意赋值 $s'_n \in \mathcal{S}^{V_T}$, 都有:

$$Pr[X=s'_n] = Pr[\sigma(S'_n)=s'_n] = \pi_T(s'_n). \text{ 证毕.}$$

注意到在算法 5 中我们需要 $O(n)$ 时间来递归地计算权重序列, 在算法 6 中需要进行 $O(n)$ 次简单采样. 因此我们需要 $O(n)$ 时间来实现树上多态自旋系统的吉布斯采样.

接下来将采样算法推广到 3.4 节中定义的伪森林和仙人掌图中. 在 4.2 节中我们通过矩阵运算的

方法计算出双态自旋系统的分布函数,同样的方法可以推广到多态自旋系统中.首先考虑圈上的情况.对于有 k 个节点的简单圈 $C=(V_C, E_C)$,其边数和节点数均为 k .由于圈上的节点度数均为 2,因此可以将节点上的权重 $W(i)$ 分配到你相邻的两条边中,使得节点权重变为 $W(i)=1$,边上的权重矩阵 U 中的 i 行 j 列变为 $U'_{ij}=U_{ij}\sqrt{W(i)}\sqrt{W(j)}$;为表述简单起见,我们仍然用 U 表示转换后的边权矩阵.容易验证,经过上述变换后每个赋值对应的权重不变,圈 C 对应的配分函数也不变.对于长度为 k 的路径,同样可以使用矩阵乘法来计算对应的配分函数.首先考虑长度为 2 的路径,其权重可以由简单的矩阵运算来求得,即 $U^2=U \cdot U$,其中 U^2_{ij} 表示路径两端节点取值为 i, j 时对应的权重;对于长度为 k 的路径,递归地进行上述运算,可得其对应的权重为 k 个矩阵的积 $U^k=U \cdots U$.将路径上的节点依次标记为 v_0, v_1, \dots, v_k ,将两个端点 v_0, v_k 合并为一个节点 v ,得到有 k 个节点的圈,容易看出,节点 v 取值为 i 时圈 C 的配分函数对应于矩阵 U^k_{ii} 项的值.

同双态自旋系统中一样,可以基于矩阵的取值对圈上的节点依次进行采样.按照 U^k_{ii} 的权重比例进行采样即可得到节点 v 的取值,然后对节点 v_1, v_{k-1} 进行赋值,设 $U^{k-1}=U \cdots U$ 表示从 v_1 到 v_{k-1} 的路径对应的权重矩阵,此时节点 v 已根据采样结果赋值,节点 v_1, v_{k-1} 的分布函数可以通过将矩阵 U^{k-1} 中的项与两条边 $(v, v_1), (v, v_{k-1})$ 对应的权重相乘来求得,然后按照这个分布函数进行采样即可为节点 v_1, v_{k-1} 赋值.按照上述方法依次对圈中节点进行采样赋值,从节点 v 开始,对每个节点的采样均符合系统的吉布斯分布,通过简单的递归证明可得最后的样本是符合吉布斯分布的.算法中的计算量主要集中在矩阵运算,可以从节点 $v_{\lfloor k/2 \rfloor}$ 出发向两边进行矩阵运算,每次计算出的矩阵存储在数据结构中用于计算对应路径的分布函数,即可在线性时间内完成采样.

推论 4. 给定图 $G=(V, E)$ 和参数矩阵 U 、点权函数 W ,如果图 G 为伪森林或仙人掌图,则存在算法在 $O(n)$ 时间内从多态自旋系统对应的吉布斯分布中取得样本 X ,满足对任意 $\sigma \in \Omega_G$ 都有 $Pr[X=\sigma]=\pi_G(\sigma)$.

证明. 上述方法给出了圈上多态自旋系统的采样算法,注意到矩阵运算中每个矩阵的取值可以是不同的,因此算法可以推广到各个节点的权重函数 W 取不同值的情况,我们用 W_v 表示节点 v 的权

重函数,此时节点 v 的权重依然可以平均分配到你邻边中,即两条边的权重分别乘以 $\sqrt{W_v(i)}$ 即可.对于伪森林,圈上的节点 v 引申出的树用 T_v 表示,通过算法 5 可以计算出树 T_v 中根节点 v 的 q 种取值对应的比值 \mathcal{R}_{T_v} ,并以此作为节点 v 的权重函数 W_v .根据引理 5,函数 W_v 符合树上的吉布斯分布,根节点 v 取值的概率分布正比于序列 W_v .因此基于函数 W_v 对圈上的配分函数进行计算,求得的分布函数符合整个伪森林上的吉布斯分布.基于矩阵运算的值对圈中的节点进行采样可得对应节点的赋值,然后从根节点 v 出发,通过调用算法 6 对树 T_v 中的其他节点进行采样即可得伪森林的样本.

对于仙人掌图,其任意两个圈之间至多有一个公共节点,使用矩阵相乘的方法同样可以求解圈中节点对应的分布函数.首先考虑图 G 中只有两个圈 C_1, C_2 的情况,两个圈存在一个公共节点 v_{12} .首先将圈 C_1 从某个节点分拆,然后使用矩阵运算来计算其配分函数,遇到节点 v_{12} 时,则先对圈 C_2 进行矩阵运算来求得在圈 C_2 中节点 v_{12} 的分布函数,并以此作为节点 v_{12} 的权重函数来参与圈 C_1 的运算,由此求得的分布函数符合图 G 的配分函数,可以基于这些分布函数来对圈 C_1 中的节点进行采样,在对节点 v_{12} 采样赋值后,基于此赋值,从节点 v_{12} 开始对圈 C_2 中的节点进行采样赋值即可.在求解圈 C_2 中节点 v_{12} 权重函数的过程中计算出的矩阵可以用来计算采样过程中需要的分布函数.

对于更复杂的仙人掌图,先调用算法 5 对引申出的树进行计算,得到的比值作为树根节点的权重函数;然后任选一个圈出发,递归地进行上述过程,由于任意两个圈之间至多有一个公共节点,因此在遇到该节点时递归地进行计算可得其对应的权重函数,从而对所有圈中的节点进行采样;最后依据树中的运算结果对节点进行采样,即可得到满足吉布斯分布的样本.在递归计算过程中需要的矩阵乘法次数跟圈中节点个数相同,采样过程中简单采样的次数同图中节点个数相同,二者相加即可计算出整个过程需要的时间是 $O(n)$,在线性时间内即可取得一个符合吉布斯分布的样本. 证毕.

6 结 论

本文中我们研究了树上自旋系统的采样算法.对于伊辛模型和玻茨模型,各个状态具有对称性,我们基于状态的对称性计算出节点的边缘分布函数,

从而构造出采样算法. 对于一般的自旋系统, 对称性不再成立, 因此我们首先递归地计算所有节点对应的边缘分布, 然后通过简单变量的组合来模拟吉布斯分布. 我们的算法对树上的自旋系统均成立, 为了将算法推广到更加广泛的图中, 我们引入了树宽的概念来度量图与树的相似性, 当给定图的树宽为 2 时, 节点间具有一定的独立性, 可以直接计算其概率分布并将我们的采样算法应用到对应的图中, 包括伪森林和仙人掌图. 需要注意的是, 并非所有树宽为 2 的图都可以直接使用我们的算法, 若图中存在两个圈 C_1, C_2 , 且两个圈之间有多条公共边(公共边是连续的, 我们称之为公共路径), 此时可以选定公共路径的某一端节点, 与图中所有边的两端节点组合作为节点, 来构造该图的树分解(公共路径的另一个端点会引出树的两个分支), 因此该图的树宽仍然为 2, 但是其结构更加复杂, 难以直接计算图中节点的概率分布函数.

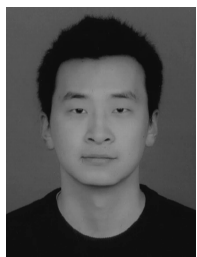
我们的采样算法仅需要 $O(n)$ 时间即可从吉布斯分布中取得一个样本. 基于 MCMC 的采样算法^[17]需要 $O(n \log n)$ 时间来取得一个样本, 我们的算法在时间复杂度上表现更好. 另一方面, 我们的算法依赖于节点之间的独立性, 对于普通图, 节点的概率分布函数会受到图中结构的影响, 很难通过简单计算来得到边缘分布. 目前已有很多技术用来对普通图的吉布斯分布进行采样以及对配分函数进行近似计算, 例如我们在第 2 节中介绍的基于 MCMC 的采样方法、基于相关性衰减的近似方法, 另外还有基于局部重采样和多项式泰勒展开的算法等. 这些方法可以在多项式时间内对结构较复杂的图进行采样或者近似计算出配分函数, 但是相比于我们的算法, 这些方法的时间复杂度较高. 若能将我们的算法推广到普通图中, 从而降低采样的时间复杂度, 这是很有意义的工作, 这需要新的技术来快速计算节点的概率分布.

参 考 文 献

- [1] Cai J, Lu P, Xia M. The complexity of complex weighted Boolean #CSP. *Journal of Computer and System Sciences*, 2014, 80(1): 217-236
- [2] Cai J, Chen X. Complexity of counting CSP with complex weights. *Journal of the ACM*, 2017, 64(3): 19:1-19:39
- [3] Cai J, Guo H, Williams T. A complete dichotomy rises from the capture of vanishing signatures. *SIAM Journal on Computing*, 2016, 45(5): 1671-1728
- [4] Lin J, Wang H. The complexity of Holant problems over Boolean domain with non-negative weights//*Proceedings of the 44th International Colloquium on Automata, Languages, and Programming*. Warsaw, Poland, 2017, 80(29): 1-14
- [5] Lin J, Wang H. The complexity of Boolean Holant problems with nonnegative weights. *SIAM Journal on Computing*, 2018, 47(3): 798-828
- [6] Cai J, Lu P, Xia M. Dichotomy for real Holant^c problems//*Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms*. New Orleans, USA, 2018: 1802-1821
- [7] Zhang Chi-Hao. *Approximation Algorithms for Counting Problems*[Ph. D. dissertation]. Shanghai Jiao Tong University, Shanghai, 2016(in Chinese)
(张驰豪. 计数问题的近似算法[博士学位论文]. 上海交通大学, 上海, 2016)
- [8] Qiu Guo-Liang, Zhang Chi-Hao. Approximability of partition functions of ferromagnetic two-state spin systems. *Computer Science*, 2020, 47(5): 22-26(in Chinese)
(邱国良, 张驰豪. 铁磁性双态自旋系统配分函数的可近似性. *计算机科学*, 2020, 47(5): 22-26)
- [9] Jerrum M, Valiant L G, Vazirani V V. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 1986, 43: 169-188
- [10] Jerrum M, Sinclair A. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 1993, 22(5): 1087-1116
- [11] Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press, 2016
- [12] Zhang Jian, Ding Shi-Fei, Zhang Nan, et al. Restricted Boltzmann machines: A review. *Journal of Software*, 2019, 30(7): 2073-2090(in Chinese)
(张健, 丁世飞, 张楠等. 受限玻尔兹曼机研究综述. *软件学报*, 2019, 30(7): 2073-2090)
- [13] Delcourt M, Heinrich M, Perarnau G. The Glauber dynamics for edge-colorings of trees. *Random Structures & Algorithms*, 2020, 57(4): 1050-1076
- [14] Abdolazimi D, Liu K, Gharan S O. A matrix Trickle-Down theorem on simplicial complexes and applications to sampling colorings//*Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science*. Denver, USA, 2022: 161-172
- [15] Kenyon C, Mossel E, Peres Y. Glauber dynamics on trees and hyperbolic graphs//*Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*. Las Vegas, USA, 2001: 568-578
- [16] Hayes T P, Vigoda E. A non-markovian coupling for randomly sampling colorings//*Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*. MA, USA, 2003: 618-627
- [17] Martinelli F, Sinclair A, Weitz D. Fast mixing for independent sets, colorings and other models on trees//*Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*. Louisiana, USA, 2004: 456-465

- [18] Weitz D. Counting independent sets up to the tree threshold// Proceedings of the 38th Annual ACM Symposium on Theory of Computing. Seattle, USA, 2006: 140-149
- [19] Xu J, Zhao Y, Chen J, Han C. A structure learning algorithm for Bayesian network using prior knowledge. *Journal of Computer Science and Technology*, 2015, 30(4): 713-724
- [20] Liu L, Hasan O, Tahar S. Formal reasoning about finite-state discrete-time Markov chains in HOL. *Journal of Computer Science and Technology*, 2013, 28(2): 217-231
- [21] Garcia V, Chen C S, Zhou Y, Shi J. Gibbs sampling based distributed OFDMA resource allocation. *Science China Information Sciences*, 2014, 57(4): 1-12
- [22] Zou B, Peng Z, Xu Z. The learning performance of support vector machine classification based on Markov sampling. *Science China Information Sciences*, 2013, 56(3): 1-16
- [23] Bubley R, Dyer M E. Path coupling: A technique for proving rapid mixing in Markov chains//Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science. Florida, USA, 1997: 223-231
- [24] Goldberg L A, Jerrum M, Paterson M. The computational complexity of two-state spin systems. *Random Structures & Algorithms*, 2003, 23(2): 133-154
- [25] Liu J, Lu P, Zhang C. The complexity of ferromagnetic two-spin systems with external fields//Proceedings of the Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM). Barcelona, Spain, 2014: 843-856
- [26] Luby M, Vigoda E. Approximately counting up to four (extended abstract)//Proceedings of the 29th Annual ACM Symposium on the Theory of Computing. Texas, USA, 1997: 682-687
- [27] Jerrum M. A very simple algorithm for estimating the number of k -colorings of a low-degree graph. *Random Structures & Algorithms*, 1995, 7(2): 157-166
- [28] Hayes T P, Vigoda E. Coupling with the stationary distribution and improved sampling for colorings and independent sets// Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms. British Columbia, Canada, 2005: 971-979
- [29] Molloy M. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM Journal on Computing*, 2004, 33(3): 721-737
- [30] Vigoda E. Improved bounds for sampling colorings// Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science. New York, USA, 1999: 51-59
- [31] Chen S, Delcourt M, Moitra A, et al. Improved bounds for randomly sampling colorings via linear programming// Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms. California, USA, 2019: 2216-2234
- [32] Dyer M E, Frieze A M, Hayes T P, et al. Randomly coloring constant degree graphs//Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science. Rome, Italy, 2004: 582-589
- [33] Galanis A, Stefankovic D, Vigoda E, et al. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 2016, 45(6): 2004-2065
- [34] Goldberg L A, Jerrum M. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM*, 2012, 59(5): 25:1-25:31
- [35] Swendsen R H, Wang J. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letter*, 1987, 58: 86-88
- [36] Blanca A, Sinclair A. Dynamics for the mean-field random-cluster model//Proceedings of the Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM). NJ, USA, 2015: 528-543
- [37] Galanis A, Stefankovic D, Vigoda E. Swendsen-Wang algorithm on the mean-field Potts model. *Random Structures & Algorithms*, 2019, 54(1): 82-147
- [38] Ullrich M. Comparison of Swendsen-Wang and heat-bath dynamics. *Random Structures & Algorithms*, 2013, 42(4): 520-535
- [39] Ullrich M. Rapid mixing of Swendsen-Wang dynamics in two dimensions. *Dissertationes Mathematicae*, 2014, 502: 3-64
- [40] Ullrich M. Swendsen-Wang is faster than single-bond dynamics. *SIAM Journal on Discrete Mathematics*, 2014, 28(1): 37-48
- [41] Guo H, Jerrum M. Random cluster dynamics for the Ising model is rapidly mixing//Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms. Spain, Hotel Porta Fira, 2017: 1818-1827
- [42] Barvinok A, Soberón P. Computing the partition function for graph homomorphisms. *Combinatorica*, 2017, 37(4): 633-650
- [43] Bulatov A, Grohe M. The complexity of partition functions. *Theoretical Computer Science*, 2005, 348(2): 148-186
- [44] Cai J, Chen X, Lu P. Graph homomorphisms with complex values: A dichotomy theorem. *SIAM Journal on Computing*, 2013, 42(3): 924-1029
- [45] Sly A. Computational transition at the uniqueness threshold //Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science. Las Vegas, USA, 2010: 287-296
- [46] Galanis A, Štefanković D, Vigoda E. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combinatorics, Probability and Computing*, 2016, 25(4): 500-559
- [47] Li L, Lu P, Yin Y. Correlation decay up to uniqueness in spin systems//Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms. New Orleans, Louisiana, USA, 2013: 67-84
- [48] Sly A, Sun N. The computational hardness of counting in two-spin models on d -regular graphs//Proceedings of the IEEE 53rd Annual Symposium on Foundations of Computer Science. New Brunswick, NJ, USA, 2012: 361-369

- [49] Guo H, Lu P. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. *ACM Transactions on Computation Theory*, 2018, 10(4): 1-25
- [50] Goldberg L A, Jerrum M. The complexity of ferromagnetic Ising with local fields. *Combinatorics, Probability and Computing*, 2007, 16(1): 43-61
- [51] Dyer M, Goldberg L A, Greenhill C, et al. On the relative complexity of approximate counting problems//Proceedings of the 3rd International Workshop on Approximation Algorithms for Combinatorial Optimization. Saarbrücken, Germany, 2000: 108-119
- [52] Dyer M, Goldberg L A, Greenhill C, et al. The relative complexity of approximate counting problems. *Algorithmica*, 2004, 38(3): 471-500
- [53] Gamarnik D, Katz D. Correlation decay and deterministic FPTAS for counting colorings of a graph. *Journal of Discrete Algorithms*, 2012, 12: 29-47
- [54] Lu P, Yin Y. Improved FPTAS for multi-spin systems//Proceedings of the Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. Springer, Berlin, Heidelberg, 2013: 639-654
- [55] Lu P, Yang K, Zhang C, et al. An FPTAS for counting proper four-colorings on cubic graphs//Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms. Barcelona, Spain, 2017: 1798-1817
- [56] Anari N, Liu K, Gharan S O. Spectral independence in high-dimensional expanders and applications to the hardcore model//Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science. Durham, USA, 2020: 1319-1330
- [57] Alev V L, Lau L C. Improved analysis of higher order random walks and applications//Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing. Chicago, USA, 2020: 1198-1211
- [58] Dinur I, Kaufman T. High dimensional expanders imply agreement expanders//Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science. Berkeley, CA, USA, 2017: 974-985
- [59] Kaufman T, Mass D. High dimensional random walks and colorful expansion//Proceedings of the 8th Innovations in Theoretical Computer Science Conference. Berkeley, USA, 2017: 4:1-4:27
- [60] Kaufman T, Oppenheim I. High order random walks: Beyond spectral gap. *Combinatorica*, 2020, 40: 245-281
- [61] Oppenheim I. Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps. *Discrete and Computational Geometry*, 2018, 59(2): 293-330
- [62] Chen Z, Liu K, Vigoda E. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion//Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. Italy, 2021: 1537-1550
- [63] Chen Z, Liu K, Vigoda E. Rapid mixing of glauber dynamics up to uniqueness via contraction//Proceedings of the 61st IEEE Annual Symposium on Foundations of Computer Science. Durham, USA, 2020: 1307-1318
- [64] Chen X, Feng W, Yin Y, et al. Rapid mixing of Glauber dynamics via spectral independence for all degrees//Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science. Denver, USA, 2022: 137-148
- [65] Feng W, Guo H, Yin Y, et al. Rapid mixing from spectral independence beyond the Boolean domain//Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms. 2021: 1558-1577
- [66] Chen Z, Galanis A, Štĕfanković D, et al. Rapid Mixing for Colorings via Spectral Independence//Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms. Virtual Conference, 2021: 1548-1557
- [67] Cryan M, Guo H, Mousa G. Modified log-Sobolev inequalities for strongly log-concave distributions//Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science. Baltimore, USA, 2019: 1358-1370
- [68] Shao S, Sun Y. Contraction: a unified perspective of correlation decay and zero-freeness of 2-spin systems. *Journal of Statistical Physics*, 2021, 185(2): 1-25
- [69] Robertson N, Seymour P D. Graph minors. III. Planar tree-width. *Journal of Combinatorial Theory Series B*, 1984, 36(1): 49-64



BAI Zong-Lei, Ph.D. candidate. His research interests include random process and approximate counting algorithms.

WANG Han-Pin, Ph.D., professor. His research interests include formal semantics and verification of computer systems, algorithms and computational complexity.

CAO Yong-Zhi, Ph.D., professor. His research interests include formal methods, privacy and security, and reasoning about uncertainty.

WANG Lu-Lu, M.S. His research interests include algorithm design and analysis.

Background

Spin systems capture a wide range of probabilistic models studied in statistical physics, probability theory, artificial intelligence, etc. Spin system is usually described by a graph, in which each vertex is assigned with one of a fixed number of states, and edges specify the neighborhood relation of the system. In statistical physics, the graph G is usually a finite portion of some regular lattice, such as \mathbb{Z}^d . An important special case is the Bethe lattice \mathbb{T}^b , i. e. G is a complete rooted tree in which each interior vertex has $b \geq 2$ children. Spin systems on trees are not only a useful simplification of the classical counterparts on \mathbb{Z}^d , but have attracted a lot of attention in their own right as the canonical example of sampling algorithms.

Spin systems are widely studied as a framework for sampling and counting problems, which encompasses many important models, such as the Potts model, coloring problem, and graph homomorphisms, etc. An important special case is when there are only 2-spins, which encompasses the famous Ising model and hard-core model. Spin systems can also be viewed as basic cases of Counting Constraint Satisfaction Problems (#CSP), which is a widely studied framework for counting problems.

Sampling from the Gibbs measures plays important roles in statistical physics and statistical inference. The classical technique to establish sampling algorithms is Markov Chain Monte Carlo (MCMC) method, in which the procedure starts from an initial state, and performs a random process among the state space Ω_G . The challenge of MCMC method is to analyze the mixing time, that is, the number of steps it takes to approach the stationary distribution. For spin systems, two main techniques, one based on canonical path and the other on couplings, were designed to bound the mixing time. For the ferromagnetic Ising model ($\beta = \gamma > 1$), Jerrum and Sinclair showed the rapid mixing of MCMC through analyzing the canonical path of the Markov chains. Then the method was generalized to ferromagnetic 2-spin systems ($\beta\gamma > 1$), and the bound of λ is improved to $\lambda_{\text{MCMC}} := \beta/\gamma$, that is, for fixed $\beta \geq \gamma$, the Markov chain mixes rapidly for 2-spin systems when $0 < \lambda \leq \lambda_{\text{MCMC}}$. For the hard-core model, Luby and Vigoda showed that the Markov chain mixes rapidly for graphs with maximum degree $\Delta \leq 4$, based on the path

coupling method.

The multi-spin systems are more sophisticated. For colorings and Potts model, a series of works were established based on MCMC method, in which the efficiency of Markov chains is also proved by the canonical paths and coupling techniques. In addition, the Potts model is equivalent to the Random cluster model in the sense of having the same partition function up to an easily computable factor. Based on the relation, Swendsen and Wang proposed a Markov chain for the Potts model, and the mixing time is well understood for complete graphs and \mathbb{Z}^2 . On the other hand, for general graphs, Ullrich showed that the mixing time of the Swendsen-Wang dynamics is always no more than that of the edge-flip dynamics. For the case $q = 2$, Guo and Jerrum provided a polynomial upper bound on the mixing time of the Swendsen-Wang algorithm. For graph homomorphisms, much attention is focused on the approximate counting, and the complexity dichotomy of exact counting.

In the paper, we studied the sampling algorithms for spin systems on trees. We presented efficient algorithms to simulate the Gibbs measures. For the Ising model and Potts model, the sampling algorithms were established based on the symmetry property of spins. For general spin systems, the symmetry property fails, so we must calculate the marginal distributions of vertices recursively, and simulate the Gibbs measures via delicate combinations of simple variables. Our algorithms only need $O(n)$ time to obtain a sample from the Gibbs measures, and exhibit better than the $O(n \log n)$ algorithms based on MCMC method. However, the results rely on the independence of sub-trees. For general graphs, the marginal distribution of fixed vertex may be influenced by vertices in the distance, so it is difficult to capture the distributions via simple calculations. Many interesting techniques have been designed to establish sampling and counting algorithms for general graphs, such as the MCMC, correlation decay, Taylor expansion method. It is valuable to extend our algorithms to general graphs, but may need new techniques to capture the marginal distributions of vertices.

This work was supported by the National Natural Science Foundation of China under Grant Nos. (61972005, 61932001, and 62172016).