

吉布斯采样在临界点前的快速收敛^{*}

陈小羽¹, 凤维明², 尹一通¹, 张昕渊¹

¹(计算机软件新技术全国重点实验室(南京大学), 新基石科学实验室, 江苏 南京 210023)

²(香港大学 计算与数据科学学院, 香港 999077)

通信作者: 尹一通, E-mail: yinyt@nju.edu.cn



摘 要: 吉布斯采样的临界行为是计算相变理论所关注的核心问题. 以硬核模型这一经典模型为例, 研究了吉布斯采样在临界点前的快速收敛. 在该模型中, 给定一个最大度为 $\Delta \geq 3$ 的 n 顶点图 G 以及参数 $\lambda \geq 0$, 则图 G 中的每个独立集 S 以正比于 $\lambda^{|S|}$ 的概率被采样. 研究了实现这一采样的经典吉布斯采样算法——Glauber dynamics, 在临界条件 $\lambda < (\Delta-1)^{\Delta-1}/(\Delta-2)^{\Delta}$ 下, 证明了该采样过程的马尔可夫链具有渐进最优的谱隙为 $\Omega(1/n)$, 因此这一经典采样算法在该临界点前始终快速收敛. 吉布斯采样过程在临界点前的快速收敛是马尔可夫链蒙特卡洛 (MCMC) 理论中的一类重要问题. 针对硬核模型上的这一问题, 此前已有若干依赖高等数学工具的证明. 为这个重要问题提供了一个简化的组合证明, 引入计算复杂性归约的思想来分析采样过程的收敛速率.

关键词: 计算相变; 马尔可夫链蒙特卡洛方法; 硬核模型

中图法分类号: TP301

中文引用格式: 陈小羽, 凤维明, 尹一通, 张昕渊. 吉布斯采样在临界点前的快速收敛. 软件学报, 2026, 37(4): 1615–1633. <http://www.jos.org.cn/1000-9825/7533.htm>

英文引用格式: Chen XY, Feng WM, Yin YT, Zhang XY. Rapid Convergence of Gibbs Sampling Before Critical Point. Ruan Jian Xue Bao/Journal of Software, 2026, 37(4): 1615–1633 (in Chinese). <http://www.jos.org.cn/1000-9825/7533.htm>

Rapid Convergence of Gibbs Sampling Before Critical Point

CHEN Xiao-Yu¹, FENG Wei-Ming², YIN Yi-Tong¹, ZHANG Xin-Yuan¹

¹(New Cornerstone Science Laboratory, State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210023, China)

²(School of Computing & Data Science, The University of Hong Kong, Hong Kong 999077, China)

Abstract: The critical behavior of Gibbs sampling is a central issue in the theory of computational phase transitions. This study takes the hard-core model, a classical model, as an example to study the rapid convergence of Gibbs sampling before the critical point. In this model, given an n -vertex graph G with a maximum degree of $\Delta \geq 3$ and a parameter $\lambda \geq 0$, each independent set S in graph G is sampled with a probability proportional to $\lambda^{|S|}$. This study investigates the canonical Gibbs sampling algorithm, Glauber dynamics, which implements this sampling. Under the critical condition $\lambda < (\Delta-1)^{\Delta-1}/(\Delta-2)^{\Delta}$, it is proven that the Glauber dynamics has an asymptotically optimal spectral gap of $\Omega(1/n)$, thus establishing that this classical sampling algorithm mixes rapidly up to the critical point. The rapid convergence of the Gibbs sampling process before the critical point is an important issue in Markov chain Monte Carlo (MCMC) theory. For this problem on the hard-core model, several proofs relying on advanced mathematical tools have been previously provided. This study offers a simplified combinatorial proof for this significant problem, introducing the idea of reductions from computational complexity to analyze the convergence rate of the sampling process.

Key words: computational phase transition; Markov chain Monte Carlo (MCMC) method; hardcore model

1 引言

吉布斯分布 (Gibbs distribution) 是一类由局部约束所描述的高维概率分布. 在计算机科学、概率论、统计物

* 收稿时间: 2025-03-01; 修改时间: 2025-08-23; 采用时间: 2025-08-27; jos 在线出版时间: 2025-12-10
CNKI 网络首发时间: 2025-12-11

理、数理统计与机器学习等多个领域中, 吉布斯分布都被广泛且深入地研究, 并有诸多重要应用. 在理论计算机科学当中, 人们关注的吉布斯分布通常和经典的图论问题紧密相关. 这些吉布斯分布的合法状态通常恰好就是某个图论问题的解空间, 例如硬核模型 (hardcore model): 状态空间为图的所有独立集 (independent set); 单体-双体模型 (monomer-dimer model): 状态空间为图的所有合法匹配 (matching); 玻茨模型 (Potts model): 参数取值恰当时状态空间为图的所有合法 q -染色 (q -vertex coloring).

吉布斯分布的采样问题是理论计算机科学的重要研究课题, 这是因为采样这种计算任务在吉布斯分布上常有计算相变现象——采样问题的计算复杂性会随着吉布斯分布的参数变化而发生骤然变化. 马尔可夫链蒙特卡洛 (Markov chain Monte Carlo, MCMC) 方法是针对这类问题研究最为深入的算法. 其中, 硬核模型是研究吉布斯采样计算相变所最经常用的基准 (benchmark) 模型, 也是各种采样算法以及分析工具的试金石^[1-17].

一个硬核模型实例由一个元组 $\mathcal{I} = (G, \lambda)$ 指定, 其中 $G = (V, E)$ 是一个简单无向图, $\lambda \in \mathbb{R}_{>0}$ 是一个正实数. 令 $\text{Ind}(G) \subseteq 2^V$ 表示图 G 中所有独立集的集合. 对于任意独立集 $S \subseteq \text{Ind}(G)$, 其权重定义为:

$$w(S) = w_{\mathcal{I}}(S) \triangleq \lambda^{|S|}.$$

硬核模型实例 $\mathcal{I}$ 指定了一个在 $\text{Ind}(G)$ 上的吉布斯分布 $\mu = \mu_{\mathcal{I}}$, 满足:

$$\forall S \in \text{Ind}(G), \quad \mu(S) \triangleq \frac{w(S)}{Z_G(\lambda)}.$$

在无向图 G 中, $Z_G(\lambda) \triangleq \sum_{S \in \text{Ind}(G)} w(S)$ 被称为配分函数. 计数问题, 即计算配分函数, 是硬核模型的核心计算任务. 精确计数是 #P 难问题^[18,19]. 更多研究集中在近似计数^[15,17,20,21]. 根据经典的规约算法^[22], 这一问题等价于吉布斯分布 μ 的采样问题.

硬核模型中有一个著名的相变现象. 在无穷 Δ -正则树中, 存在关键的唯一性阈值 (uniqueness threshold):

$$\lambda_c(\Delta) \triangleq \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta-2}.$$

如果 $\lambda < \lambda_c$, 根节点上的边缘分布与叶节点上的边界条件在深度为 ℓ 的情况下以指数方式衰减; 如果 $\lambda > \lambda_c$, 即使 $\ell \rightarrow \infty$, 长程相关性始终存在. 在相同的阈值处, 也发现了计算复杂性的相变. Weitz^[17]表明, 对于所有最大度数为常数 Δ 的图, 如果 $\lambda < \lambda_c(\Delta)$, 则存在一种完全多项式时间逼近方案 (FPTAS) 用于估计 $Z_G(\lambda)$. 另一方面, Sly^[15]证明除非 $\text{NP} = \text{RP}$, 否则当 $\lambda > \lambda_c(\Delta)$ 时, 不存在用于估计 $Z_G(\lambda)$ 的完全多项式随机化逼近方案. 对于硬核模型来说, 从其吉布斯分布中采样在计算上等同于近似计数^[22].

Glauber dynamics (也称吉布斯采样算法) 是从硬核模型当中采样的标准 (canonical) 马尔可夫链蒙特卡洛方法. 设 $\mathcal{I} = (G = (V, E), \lambda)$ 为一个硬核模型实例. 设 $\Gamma(v)$ 表示 G 中 $v \in V$ 的邻居. $\mathcal{I}$ 上的 Glauber dynamics 从任意一个独立集 $X_0 \in \text{Ind}(G)$ 开始. 在第 t 个转移步骤中, 它执行以下操作.

- 均匀随机选择一个顶点 $v \in V$.
- 如果 $X_{t-1} \cap \Gamma(v) \neq \emptyset$, 则 $X_t = X_{t-1}$; 否则, 令:

$$X_t = \begin{cases} X_{t-1} \setminus \{v\}, & \text{以概率 } \frac{1}{1+\lambda} \\ X_{t-1} \cup \{v\}, & \text{以概率 } \frac{\lambda}{1+\lambda} \end{cases}.$$

Glauber dynamics 最终会收敛到唯一的稳态分布 μ . 为了衡量收敛的速度, 我们定义 Glauber dynamics 的混合时间 (mixing time) 为:

$$\forall 0 < \varepsilon < 1, T_{\text{mix}}(\varepsilon) \triangleq \max_{X_0 \in \text{Ind}(G)} \min \{t \mid d_{\text{TV}}(X_t, \mu) \leq \varepsilon\},$$

其中, $d_{\text{TV}}(X_t, \mu)$ 表示 X_t 的分布与 μ 之间的全变差距离 (total variation distance).

对于硬核模型上的 Glauber dynamics, 之前有相当多的研究. 早期的工作证明了在一个亚临界区域 $\lambda < \frac{2}{\Delta-2}$ 内的 $O(n \log n)$ 混合时间, 其中 n 是顶点数量^[13,14,16]. 随后一些工作在特殊图上做到了任意的亚临界区域 (对任意

$\delta \in (0, 1)$, $\lambda \leq (1 - \delta)\lambda_c(\Delta)$ 内的 $O(n \log n)$ 混合时间^[10-12].

最近, 研究人员提出一种基于高维扩张器的马尔可夫链分析工具^[23-25], 也被称为谱独立性 (spectral independence) 技术. 借助这种技术, Anari 等人^[6]首次证明了一般图上任意亚临界区域以内的多项式混合时间. 这个结果之后被优化为 $n^{O(SI)}$ ^[8]. 其中, SI 是谱独立性参量, 只和目标分布相关. 对于硬核模型来说, 这个参量有如下的上界:

$$SI = O(1) \times \begin{cases} \min\{1/\delta, n\}, & \lambda \leq (1 - \delta)\lambda_c(\Delta) \\ n, & \lambda > \lambda_c(\Delta) \end{cases}.$$

所以, 当 δ 很小时, $n^{O(SI)}$ 是一个大多项式.

因为高维扩张器技术对一般的分布都适用 (甚至是唯一性区域以外的硬核模型), 所以在一般情形下 (不一定是硬核模型) 改进这个 $n^{O(SI)}$ 的混合时间界是不可能的 (不然会和 Sly 的下界矛盾^[15]). 想要改进这个混合时间界, 需要用到硬核模型独有的性质.

在 Chen 等人^[9]的工作中, 他们注意到在稀疏图上随机选取足够多比例的点删掉之后剩下的点大概率断开成若干规模很小的连通块. 利用这个额外性质, 他们证明了稀疏图 (最大度数为常数) 上硬核模型的 Glauber dynamics 的 $O(n \log n)$ 混合时间. 之后的若干工作引入了场动力系统 (field dynamics), 熵独立性 (entropic independence) 等新工具^[1,2], 它们能利用上硬核模型在 $\lambda \in (0, (1 - \delta)\lambda_c(\Delta))$ 这个区域内 SI 都是常数这一特殊性质. 最终成功证明了硬核模型上任意的亚临界区域内的 $O(n \log n)$ 混合时间^[3,7].

值得一提的是, 上面提到的这些证明通常都依赖一些高等数学工具例如随机微分方程. 本文通过引入计算复杂性规约的思想实现了类似的结果. 它的证明更简单并且完全是初等的.

1.1 主要结果

本文通过研究 Glauber dynamics 的谱隙 (spectral gap) 来研究其混合时间. 设 $P_{\mathcal{I}}$ 为 Glauber dynamics 在硬核模型实例 $\mathcal{I} = (G, \lambda)$ 上的转移矩阵. 众所周知, $P_{\mathcal{I}}$ 具有 $|\text{Ind}(G)|$ 个非负实特征值 $1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{|\text{Ind}(G)|} \geq 0$. 谱隙与混合时间密切相关, 定义 $P_{\mathcal{I}}$ 的谱隙如下:

$$\gamma_{\mathcal{I}} \triangleq 1 - \lambda_2.$$

我们的主要贡献是以下定理.

定理 1. 设 $0 < \delta < 1$ 为常数, 对于任何满足 $\Delta \geq 3$ 和 $\lambda \leq (1 - \delta)\lambda_c(\Delta)$ 的硬核模型实例 $\mathcal{I} = (G, \lambda)$, Glauber dynamics 在 $\mathcal{I}$ 上的谱隙满足:

$$\frac{1}{C(\delta)n} \leq \gamma_{\mathcal{I}} \leq \frac{2}{n},$$

并且 Glauber dynamics 的混合时间满足:

$$T_{\text{mix}}(\varepsilon) \leq 4C(\delta) \cdot n \cdot \left(n \log \Delta + \log \frac{1}{\varepsilon} \right),$$

其中, n 是 G 中顶点的数量, Δ 是 G 的最大度数, $C(\delta) = \left(\frac{1}{\delta} \right)^{O(1/\delta)}$ 是仅依赖于 δ 的常数.

我们的证明依赖于一种新的变换思想. 具体来说, 首先将硬核模型实例转换为具有足够大最大度数的新实例. 接下来, 分析变换后实例上的 Glauber dynamics. 这种分析利用了某些概率集中性质, 还使用了文献 [9,20] 中发展的“局部到全局”理论. 最后, 将原始实例上的 Glauber dynamics 与变换后实例上的 Glauber dynamics 进行比较. 这种转换提供了一定的优良特性, 以便每个证明步骤可以顺利进行. 请参见第 3 节查看证明概要.

更加直观地说, “局部到全局”技术的核心, 在于利用特定的一系列条件分布对目标分布 μ 进行分解. 在 Chen 等人^[9]以及更早的 Anari 等人^[6]的工作中, 通过固定一系列顶点的状态 (即 pinning) 来得到这些条件分布. 在我们的前期工作^[2]中, 为了改进 Chen 等人^[9]的工作而引入了场动力系统, 该技术在此前的基础上考虑了更复杂的一系列条件分布. 在场动力系统引入的条件分布当中, 除了一些特定顶点状态被固定之外, 硬核模型的参数也会有相应的偏移. 这些更复杂的条件分布虽然帮助我们解决了度数受限的问题, 但是处理它们需要更高等的数学工具以及更复杂的分析. 在本文的证明中, 首先通过一个组合变换将原分布 μ 转换为一个相似却又不同的新分布 μ_k . 之后本

文的证明仅在这个新分布上考虑通过固定变量状态得到的一系列条件分布,从而得到一个简单且初等的证明.

2 定义和背景知识

2.1 硬核模型和已知结果

设 $\mathcal{I} = (G = (V, E), \lambda)$ 是一个硬核模型, 其吉布斯分布为 $\mu = \mu_{\mathcal{I}}$. 在分析中, 我们经常把 μ 视为在 $\{0, 1\}^V$ 上的分布. 每个配置 $\sigma \in \{0, 1\}^V$ 代表一个唯一的子集 $S_{\sigma} \triangleq \{v \in V \mid \sigma_v = 1\}$. 我们使用 $\Omega = \Omega_{\mathcal{I}} = \{\sigma \in \{0, 1\}^V \mid S_{\sigma} \text{ 是一个独立集}\}$ 来表示 μ 的支撑集. 对于每个 $\sigma \in \Omega$, $\mu(\sigma) \propto \lambda^{|S_{\sigma}|}$.

对于硬核模型, Glauber dynamics 的谱隙和混合时间易于在 Dobrushin 条件下分析. 通过路径耦合^[26]及其对谱隙的推论^[27], 我们得到以下结果. 类似的结果也在文献^[28]中观察到.

命题 1^[26,27]. 设 $0 < \delta < 1$ 是一个常数. 设 $\mathcal{I} = (G, \lambda)$ 是具有最大度 $\Delta \geq 1$ 的 n 顶点硬核模型. 如果 $\mathcal{I}$ 满足 Dobrushin 条件 $\lambda \leq \frac{1-\delta}{\Delta-1+\delta}$, 那么对于 $\mathcal{I}$ 上的 Glauber dynamics, 其混合时间至多为 $\frac{n}{\delta} \log \frac{n}{\varepsilon}$, 谱间隙至少为 $\frac{\delta}{n}$.

基于局部到全局技术^[8,9,23], 可以证明在唯一性区域中硬核模型有以下谱隙.

命题 2^[8]. 令 $\delta \in (0, 1)$ 为常数. 令 $\mathcal{I} = (G, \lambda)$ 为一个具有最大度数 $\Delta \geq 3$ 的 n 顶点硬核模型. 如果 $\mathcal{I}$ 满足唯一性条件 $\lambda \leq (1-\delta)\lambda_c(\Delta)$, 那么 $\mathcal{I}$ 上 Glauber dynamics 的谱隙至少为 $(5n)^{-66/\delta}$.

注意到在文献^[8]中的结果意味着 Glauber dynamics 有 $n^{-1-32/\delta} e^{-O(1/\delta^2)}$ 的谱隙下界. 结合文献^[8]和文献^[29]的 Theorem 1.1 和 Theorem 3.1, 可以得到命题 2 中的结果.

2.2 马尔可夫链和混合时间

设 Ω 为状态空间. 设 $(X_t)_{t \geq 0}$ 是 Ω 上的马尔可夫链, 其转移矩阵 (transition matrix) 为 $P: \mathcal{R}_{\geq 0}^{\Omega \times \Omega}$. 当上下文清楚时, 我们通常使用矩阵 P 来表示对应的马尔可夫链. 当对任意的 $X, Y \in \Omega$, 存在整数 t 使得 $P^t(X, Y) > 0$ 时, 马尔可夫链是不可约 (irreducible) 的. 当对任意的 $X \in \Omega$, $\gcd\{t \mid P^t(X, X) > 0\} = 1$ 时, 马尔可夫链是非周期性 (aperiodic) 的. 如果一个分布 μ 满足 $\mu = \mu P$, 则称 μ 是 P 的一个稳态分布 (stationary distribution). 如果一个马尔可夫链既是不可约的又是非周期性的, 那么它有一个唯一的稳态分布. 如果对于分布 μ , 马尔可夫链 P 满足下面的详细平衡方程 (detailed balance equation):

$$\mu(X)P(X, Y) = \mu(Y)P(Y, X), \forall X, Y \in \Omega.$$

这意味着 μ 是 P 的一个稳态分布. 我们也说这个马尔可夫链是时间可逆的 (time reversible). 在本文中, 我们考虑的所有马尔可夫链都是可逆的.

令 μ 是集合 Ω 上的分布. 令 P 是一个关于 Ω 的不可约、非周期、时间可逆的马尔可夫链, 具有唯一的稳态分布 μ . P 的混合时间的定义为:

$$T_{\text{mix}}(\varepsilon) \triangleq \max_{X \in \Omega} \min \{t \mid d_{\text{TV}}(P^t(X, \cdot), \mu) \leq \varepsilon\}, \forall 0 < \varepsilon < 1,$$

其中, $P^t(X, \cdot)$ 是从 X 开始经过 t 个转移步骤后生成的分布的马尔可夫链, $d_{\text{TV}}(P^t(X, \cdot), \mu)$ 表示 $P^t(X, \cdot)$ 和 μ 之间的全变差距离 (total variation distance), 形式上为:

$$d_{\text{TV}}(P^t(X, \cdot), \mu) \triangleq \frac{1}{2} \sum_{Y \in \Omega} |P^t(X, Y) - \mu(Y)|.$$

2.3 Block dynamics

在本文中, 我们研究布尔随机变量上的 block dynamics. 设 V 是一个基底集合. 设 μ 是定义在 $\{0, 1\}^V$ 上的分布, 支撑集为 Ω . 对于任意 $\Lambda \subseteq V$, 我们用 μ_{Λ} 表示从 μ 投影到 Λ 上的边缘分布. 对于某些集合 $\Lambda \subseteq V$ 上的配置 $\sigma \in \{0, 1\}^{\Lambda}$, 如果 $\mu_{\Lambda}(\sigma) > 0$, 则我们称 σ 是可行的. 对于任意可行的 $\tau \in \{0, 1\}^{V \setminus \Lambda}$, 令 μ_{Λ}^{τ} 表示在给定 τ 条件下 Λ 上的边缘分布. 我们用 $\Omega_{\Lambda}^{\tau} \subseteq \{0, 1\}^{\Lambda}$ 来表示 μ_{Λ}^{τ} 的支撑集. 对于任意整数 $1 \leq \ell \leq |V|$, 令 $\binom{V}{\ell} \triangleq \{S \subseteq V \mid |S| = \ell\}$ 表示所有大小为 ℓ 的子集的集族. μ 上的 block dynamics 定义如下.

定义 1. ℓ -block dynamics. 对于任意正整数 $1 \leq \ell \leq |V|$, μ 上的 ℓ -block dynamics 是一个马尔可夫链 $(X_t)_{t \geq 0}$, 定义如下. 它从任意配置 $X_0 \in \Omega$ 开始, 在第 t 个转移步骤中, 执行如下操作.

- 均匀随机选择一个集合 $S \in \binom{V}{\ell}$, 并使 $X_t(V \setminus S) = X_{t-1}(V \setminus S)$.
- 从分布 $\mu_S^{X_t(V \setminus S)}$ 中采样 $X_t(S) \in \Omega_S^{X_t(V \setminus S)}$.

特别地, 1-block dynamics 其实就是 μ 上的 Glauber dynamics. 我们考虑在条件分布 μ_Λ^σ 上的 block dynamics, 其中 $\Lambda \subseteq V$ 是一个子集, $\sigma \in \{0, 1\}^{V \setminus \Lambda}$ 是一个可行配置. 特别地, 如果 $\Lambda = V$, 则 $\mu_\Lambda^\sigma = \mu$. 以下命题很容易验证.

命题 3. 对于从硬核模型诱导出的任何条件分布 μ_Λ^σ , 对于任何 $1 \leq \ell \leq |\Lambda|$, 在 μ_Λ^σ 上的 ℓ -block dynamics 是不可约的、非周期的, 并且关于 μ_Λ^σ 是时间可逆的.

2.4 分析工具

设 μ 是具有支撑集 Ω 的分布. 令 $\langle \cdot, \cdot \rangle_\mu$ 表示由 μ 加权的内积, 即, 对于任何两个函数 $f, g: \Omega \rightarrow \mathbb{R}$,

$$\langle f, g \rangle_\mu \triangleq \sum_{X \in \Omega} \mu(X) f(X) g(X).$$

设 P 是一个不可约的马尔可夫链, 对于 μ 是时间可逆的. 我们通常将转移矩阵 P 视为一个马尔可夫算子, 将函数映射到函数. 对于任何函数 $f: \Omega \rightarrow \mathbb{R}$, P 作用在 f 上产生一个新的函数 $Pf: \Omega \rightarrow \mathbb{R}$, 使得:

$$(Pf)(X) \triangleq \sum_{Y \in \Omega} P(X, Y) f(Y), \quad \forall X \in \Omega.$$

如果将 f 视为一个列向量, 那么 Pf 就是简单的矩阵乘法. 由于 P 是不可约且时间可逆的, P 有 $|\Omega|$ 个实特征值 $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq -1$ (因为 P 是不可约的, 所以 $\lambda_2 < 1$)^[30]; 每个 λ_i 对应一个实特征函数 $f_i: \Omega \rightarrow \mathbb{R}$, 使得 $Pf_i = \lambda_i f_i$, 其中 $f_1 = 1$ 为全 1 向量, 而 $f_1, f_2, \dots, f_{|\Omega|}$ 形成一个内积空间 $(\mathbb{R}^\Omega, \langle \cdot, \cdot \rangle_\mu)$ 的正交基. 假设 $|\Omega| \geq 2$. P 的绝对谱隙定义为 $1 - \lambda_\star \triangleq 1 - \max\{|\lambda_i| \mid 2 \leq i \leq |\Omega|\}$. P 的谱隙 (spectral gap) 定义为 $1 - \lambda_2$. P 的弛豫时间 (relaxation time) 定义为:

$$t_{\text{rel}}(P) \triangleq \begin{cases} \frac{1}{1 - \lambda_\star}, & \text{if } |\Omega| \geq 2 \\ 0, & \text{if } |\Omega| = 1 \end{cases} \quad (1)$$

在退化情况 $|\Omega| = 1$ 中, 我们简单假设 $t_{\text{rel}}(P) = 0$. 关于混合时间和弛豫时间之间的关系是众所周知的^[30]:

$$T_{\text{mix}}(\varepsilon) \leq t_{\text{rel}} \log \left(\frac{1}{\varepsilon \mu_{\min}} \right), \quad \text{其中 } \mu_{\min} = \min_{X \in \Omega} \mu(X) \quad (2)$$

此外, 如果 P 是在 μ 上的一个不可约的 block dynamics (定义 1), 则 P 具有 $|\Omega|$ 个非负实特征值 $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq 0$ ^[23,30,31]. 因此, 有以下关系成立:

$$1 - \lambda_\star = 1 - \lambda_2.$$

在本文中, 我们将使用以下工具来分析谱隙.

2.4.1 Poincaré 不等式

对于任意 $f: \Omega \rightarrow \mathbb{R}$, 关于 μ 的期望值被定义为:

$$E_\mu[f] \triangleq \sum_{X \in \Omega} \mu(X) f(X).$$

关于 μ 的方差被定义为:

$$\text{Var}_\mu[f] \triangleq E_\mu[f^2] - (E_\mu[f])^2 = \frac{1}{2} \sum_{X, Y \in \Omega} \mu(X) \mu(Y) (f(X) - f(Y))^2 \quad (3)$$

其中, f^2 是一个函数, 使得 $f^2(x) = (f(x))^2$. 设 P 是关于 μ 时间可逆的马尔可夫链. 关于 P 和 μ 的狄利克雷形定义为:

$$\mathcal{E}_P(f, f) \triangleq \langle f, (I - P)f \rangle_\mu = \frac{1}{2} \sum_{X, Y \in \Omega} \mu(X) P(X, Y) (f(X) - f(Y))^2 \quad (4)$$

谱隙可以通过以下 Poincaré 不等式来表征^[30]:

$$1 - \lambda_2 = \inf \left\{ \frac{\mathcal{E}_P(f, f)}{\text{Var}_\mu[f]} \mid f: \Omega \rightarrow \mathbb{R}, \text{Var}_\mu[f] \neq 0 \right\} \quad (5)$$

此外, 如果 P 是 μ 上的 block dynamics (定义 1), 我们有以下性质.

引理 1. 设 V 是一个基底集合. 设 μ 是在 $\{0, 1\}^V$ 上的支撑集为 Ω 的分布. 令 $1 \leq \ell \leq |V|$ 是一个整数. 令 P 是 μ 上的 ℓ -block dynamics. 如果 P 是不可约的, 那么:

$$\text{Var}_\mu[f] \leq t_{\text{rel}}(P) \mathcal{E}_P(f, f), \quad \forall f: \Omega \rightarrow \mathbb{R}.$$

证明: 如果 $|\Omega| \geq 2$, 由于 P 是不可约的, $\lambda_2 < 1$, 因此根据公式 (5), 命题成立. 如果 $|\Omega| = 1$, 那么根据公式 (3) 和公式 (4) 中的定义, 我们有 $\text{Var}_\mu[f] = \mathcal{E}_P(f, f) = 0$, 因此命题显然成立.

此外, 块动态的狄利克雷形可以用引理 2^[32,33]来重新表述. 对于任意不相交集 $A, B \subseteq V$, 任意 $\sigma_A \in \{0, 1\}^A$, $\sigma_B \in \{0, 1\}^B$, 我们用 $\sigma_A \uplus \sigma_B$ 表示一个在 $A \cup B$ 上的配置, 该配置在 A 中与 σ_A 一致, 在 B 中与 σ_B 一致.

引理 2. 设 V 是一个基底集合. 设 μ 是一个在 $\{0, 1\}^V$ 上支撑集为 Ω 的分布. 设 $1 \leq \ell \leq |V|$ 为一个整数. 设 P 是 μ 上的 ℓ -block dynamics. 对于任意 $f: \Omega \rightarrow \mathbb{R}$, 有:

$$\mathcal{E}_P(f, f) = \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{\rho \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\rho) \cdot \text{Var}_{\mu_S^\rho}[f_\rho],$$

其中, $\Omega_{V \setminus S}$ 表示边缘分布 $\mu_{V \setminus S}$ 的支撑集, $f_\rho: \Omega_S^\rho \rightarrow \mathbb{R}$ 定义为 $f_\rho(\tau) \triangleq f(\rho \uplus \tau)$, 而 Ω_S^ρ 表示条件分布 μ_S^ρ 的支撑集.

证明: 通过 ℓ -block dynamics 的定义, 我们知道对于所有的 $\sigma, \tau \in \Omega$, 转移概率可以写成:

$$P(\sigma, \tau) = \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu_S^{\sigma_{V \setminus S}}(\tau_S) \cdot \mathbf{1}[\sigma_{V \setminus S} = \tau_{V \setminus S}].$$

于是, 根据公式 (4) 中狄利克雷形的定义, 我们有:

$$\begin{aligned} \mathcal{E}_P(f, f) &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} \mu(\sigma) \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu_S^{\sigma_{V \setminus S}}(\tau_S) \cdot \mathbf{1}[\sigma_{V \setminus S} = \tau_{V \setminus S}] \cdot (f(\sigma) - f(\tau))^2 \\ &= \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{\rho \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\rho) \cdot \frac{1}{2} \sum_{\alpha, \beta \in \Omega_S^\rho} \mu_S^\rho(\alpha) \mu_S^\rho(\beta) (f_\rho(\alpha) - f_\rho(\beta))^2 \\ &= \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{\rho \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\rho) \cdot \text{Var}_{\mu_S^\rho}[f_\rho]. \end{aligned}$$

2.4.2 Cheeger 不等式

设 P 是相对于 μ 可逆的马尔可夫链. 对于状态空间的任意子集 $S \subseteq \Omega$, 我们使用 $\mu(S)$ 表示 $\sum_{X \in S} \mu(X)$. 马尔可夫链 P 的导通率 (conductance) 定义为:

$$\Phi \triangleq \min_{S \subseteq \Omega: \mu(S) < 1/2} \frac{\sum_{X \in S, Y \in \Omega \setminus S} \mu(X) P(X, Y)}{\mu(S)}.$$

命题 4. Cheeger 不等式^[34,35]. 导通率和谱隙满足如下关系:

$$\frac{\Phi^2}{2} \leq 1 - \lambda_2 \leq 2\Phi.$$

2.5 集中度

我们将使用以下超几何分布的概率集中不等式.

引理 3^[36]. 设 V 为大小为 n 的基底集合. 设 $\Lambda \subseteq V$ 为一个子集. 给定一个整数 $1 \leq \ell \leq n$, 令 $S \in \binom{V}{\ell}$ 为从 $\binom{V}{\ell}$ 中

均匀随机抽取的随机子集. 对于每个 $i \in \Lambda$, 定义指示随机变量 $X_i = \mathbf{1}_{i \in S}$. 令 $X = \sum_{i \in \Lambda} X_i$, 则对于任意 $t > \frac{2e\ell|\Lambda|}{n} \mu_k$ $= 2eE[X]$, 有 $\Pr[X \geq t] \leq 2^{-t}$.

3 证明大纲

本节证明主定理 (定理 1). 首先, 构造一个变换, 将原本的目标硬核模型分布 μ 变换为另外一个硬核模型分布 μ_k (定义 3). 这两个分布紧密相关然而 μ_k 对应的硬核模型总是具有 $\Omega(\log n)$ 量级的最大度数. 我们会首先证明 μ_k 上的 Glauber dynamics 具有最优的谱隙 (引理 6). 然后, 通过比较 μ 和 μ_k 上的 Glauber dynamics, 来证明这两个分布上的 Glauber dynamics 的谱隙具有相同量级 (引理 7). 最后结合引理 6 和引理 7, 我们就能证明定理 1.

我们考虑以下一系列硬核模型实例.

定义 2. 硬核模型实例族 $\mathcal{F}(\delta)$. 令 $0 < \delta < 1$ 为一个常数. 定义 $\mathcal{F}(\delta)$ 为满足 $\Delta \geq 3$ 且 $\lambda \leq (1-\delta)\lambda_c(\Delta)$ 的所有硬核模型实例 $\mathcal{I} = (G, \lambda)$ 的集合, 其中 Δ 是图 G 的最大度数.

3.1 谱隙的下界

我们首先证明了谱间隙的下界.

引理 4. 设 $0 < \delta < 1$ 是一个常数. 存在一个常数 $C(\delta) = \left(\frac{1}{\delta}\right)^{O(1/\delta)}$, 对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, Glauber dynamics 在 $\mathcal{I}$ 上的谱隙满足:

$$\gamma_{\mathcal{I}} \geq \frac{1}{C(\delta)n},$$

其中, n 是 G 中顶点的数量, Δ 是 G 的最大度数.

固定一个常数 $0 < \delta < 1$. 为了证明引理 4, 我们使用 $8/\delta$ 作为阈值, 将 $\mathcal{F}(\delta)$ 中的所有实例分为两部分: 最大度数小于 $8/\delta$ 的实例和最大度数大于等于 $8/\delta$ 的实例. 注意对于第 1 部分, 我们已经有了文献 [9] 中的结果如下.

引理 5^[9]. 设 $0 < \delta < 1$ 为常数. 存在一个常数 $C_0(\delta) = \left(\frac{1}{\delta}\right)^{O(1/\delta)}$, 使得对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 如果 $\Delta < 8/\delta$, 则 $\mathcal{I}$ 上 Glauber dynamics 的谱隙满足以下关系:

$$\gamma_{\mathcal{I}} \geq \frac{1}{C_0(\delta)n},$$

其中, n 是 G 中顶点的数目, Δ 是 G 的最大度数.

引理 5 是通过结合文献 [9] 中的几个结果证明的. 为了完整起见, 我们在附录 B 中给出了证明. 我们现在只关注最大度数至少为 $8/\delta$ 的核心实例. 下面的变换是本文证明中的一个关键工具.

定义 3. k -变换. 设 $k \in \mathbb{N}^+$ 为一个整数. 给定一个硬核模型实例 $\mathcal{I} = (G = (V, E), \lambda)$, 定义一个变换实例 $\mathcal{I}' = (G', \lambda') \triangleq \text{Trans}(\mathcal{I}, k)$, 使得 $\lambda' \triangleq \lambda/k$, 图 $G' = (V', E')$ 构造如下.

- 对于任意顶点 $v \in V$, 存在 k 个顶点 $v_1, v_2, \dots, v_k \in V'$ 在 G' 中形成一个 k -团.
- 对于任意边 $\{u, v\} \in E$, 任意 $1 \leq i, j \leq k$, 在 G' 中 u_i 和 v_j 通过一条边相连.

图 1 是一个 k -变换 ($k=3$) 的示例. 首先, 对于 G 中的每个顶点, 在 G' 中有一个 3-团. 以顶点 a 为例, 我们在 G' 中有顶点 a_1, a_2, a_3 组成的 3-团. 接下来, 我们在团之间添加边. 以 G 中的边 $\{a, b\}$ 为例, 对于任意的 $i, j \in [3]$, 在 G' 中有一条边 $\{a_i, b_j\}$.

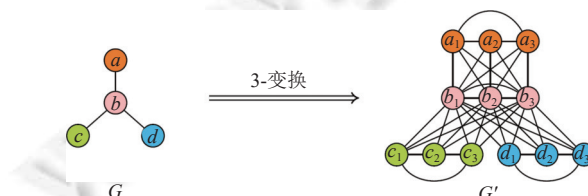


图 1 3-变换的示例

我们使用以下特定的变换来证明引理 4. 设 $0 < \delta < 1$ 为常数. 对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 定义参数:

$$\kappa(\mathcal{I}) \triangleq \left\lceil \frac{10^6}{\delta^2} \log n \right\rceil \quad (6)$$

其中, n 为 G 中的顶点数. 定义转换后的实例:

$$\mathcal{I}' = (G' = (V', E'), \lambda') \triangleq \text{Trans}(\mathcal{I}, \kappa(\mathcal{I})) \quad (7)$$

我们有以下两个引理.

引理 6. Glauber dynamics 在 $\mathcal{I}'$ 上的谱隙. 设 $0 < \delta < 1$ 为一个常数. 存在一个常数 $C_1(\delta) = \exp(O(1/\delta))$ 使得对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 如果 $\Delta \geq 8/\delta$, 那么对于转换后实例 $\mathcal{I}' = (G', \lambda')$ 上 Glauber dynamics 的谱隙 (参照公式 (7) 中的定义) 满足:

$$\gamma_{\mathcal{I}'} \geq \frac{1}{C_1(\delta)n'},$$

其中, Δ 是图 G 的最大度数, n' 是图 G' 的顶点数.

引理 7. 在 $\mathcal{I}$ 和 $\mathcal{I}'$ 上 Glauber dynamics 的比较. 设 $0 < \delta < 1$ 为常数. 对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 有:

$$\gamma_{\mathcal{I}} \geq \frac{\kappa(\mathcal{I})}{5} \gamma_{\mathcal{I}'},$$

其中, $\gamma_{\mathcal{I}}$ 是 $\mathcal{I}$ 上 Glauber dynamics 的谱隙, $\gamma_{\mathcal{I}'}$ 是变换后实例 $\mathcal{I}'$ 上 Glauber dynamics 的谱隙 (见公式 (7) 中定义), $\kappa(\mathcal{I})$ 是在公式 (6) 中定义参数.

引理 6 的证明在第 4 节给出. 变换后, 新的实例 $\mathcal{I}'$ 仍然处于唯一性区域, 其最大度数足够大. 最大度数的下界保证了某种概率集中性质, 对我们的证明起着重要作用. 引理 7 的证明在第 5 节中给出. k -变换保证了存在一个函数 $g: \Omega_{\mathcal{I}'} \rightarrow \Omega_{\mathcal{I}}$, 如果 $X \in \Omega_{\mathcal{I}'}$ 服从 Gibbs 分布 $\mu_{\mathcal{I}'}$, 那么 $Y = g(X)$ 服从 Gibbs 分布 $\mu_{\mathcal{I}}$. 使用此函数, 我们将两个马尔可夫链联系起来, 以证明引理 7 中的关系. 有了引理 6 和引理 7, 我们现在来证明引理 4.

证明: 令 $0 < \delta < 1$ 为常数. 考虑一个实例 $\mathcal{I} = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$. 如果 G 的最大度数 $\Delta < 8/\delta$, 那么根据引理 5, 谱隙至少为 $\frac{1}{C_0(\delta)n}$, 其中 $C_0(\delta) = \left(\frac{1}{\delta}\right)^{O(1/\delta)}$ 是引理 5 中的常数. 假设 $\Delta \geq 8/\delta$. 设 $\mathcal{I}'$ 表示公式 (7) 中的变换实例. 结合引理 6 和引理 7, 谱隙满足:

$$\gamma_{\mathcal{I}} \geq \frac{\kappa(\mathcal{I})}{5} \gamma_{\mathcal{I}'} \geq \frac{\kappa(\mathcal{I})}{5} \cdot \frac{1}{C_1(\delta)n'} = \frac{1}{5C_1(\delta)n},$$

其中, $C_1(\delta) = \exp(O(1/\delta))$ 是引理 6 中的常数, n' 是 $\mathcal{I}'$ 中顶点的数量, 最后一个等式成立是因为 $n' = \kappa(\mathcal{I})n$. 将两种情况结合起来, 谱隙 $\gamma_{\mathcal{I}}$ 至少为 $\frac{1}{C(\delta)n}$, 其中:

$$C(\delta) = \max\{C_0(\delta), C_1(\delta)\} = \left(\frac{1}{\delta}\right)^{O(1/\delta)}. \quad \square$$

3.2 定理 1 的证明

固定一个常数 $0 < \delta < 1$. 固定一个具有 Gibbs 分布 μ 的实例 $\mathcal{I} = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$, 其中 $|V| = n$. 设 Ω 表示 μ 的支撑集. 我们可以假设 $\lambda \geq \frac{1}{2\Delta}$. 否则, 由于 Dobrushin 条件成立, 根据命题 1, 主要结果成立. 令 P 表示 $\mathcal{I}$ 上的 Glauber dynamics. 根据引理 4, P 的谱隙至少为 $\frac{1}{C(\delta)n}$, 其中 $C(\delta) = (1/\delta)^{O(1/\delta)}$ 是引理 4 中的常数. 注意, Glauber dynamics 是定义 1 中 block dynamics 的特例. 结合引理 4、公式 (1) 和公式 (2), 混合时间满足:

$$T_{\text{mix}}(\varepsilon) \leq C(\delta)n \log\left(\frac{1}{\varepsilon \mu_{\min}}\right), \mu_{\min} = \min_{X \in \Omega} \mu(X).$$

给出配置 $\sigma \in \{0, 1\}^V$, 它的权重定义为 $w(\sigma) = \lambda^{\sum_{v \in V} \sigma_v}$. 设 $Z = \sum_{\sigma \in \Omega} w(\sigma)$ 表示配分函数. 吉布斯分布由 $\mu(\sigma) = \frac{w(\sigma)}{Z}$ 定义.

注意到 $Z \leq \sum_{\sigma \in \{0, 1\}^V} \lambda^{\sum_{v \in V} \sigma_v} = (1 + \lambda)^n$, 而 $w(\sigma) \geq \min\{1, \lambda^n\}$. 这意味着:

$$\mu_{\min} \geq \min \left\{ \frac{1}{(1+\lambda)^n}, \frac{\lambda^n}{(1+\lambda)^n} \right\}.$$

注意到 $\frac{1}{2\Delta} \leq \lambda \leq \lambda_c(\Delta) \leq \frac{2e}{\Delta-2}$. 这意味着 $\log\left(\frac{1}{\mu_{\min}}\right) \leq n \log(8\Delta)$. 我们有:

$$T_{\text{mix}}(\varepsilon) \leq C(\delta)n \log\left(\frac{1}{\varepsilon\mu_{\min}}\right) \leq C(\delta)n \left(n \log(8\Delta) + \log \frac{1}{\varepsilon}\right) \leq 4C(\delta) \cdot n \cdot \left(n \log \Delta + \log \frac{1}{\varepsilon}\right).$$

谱隙的上界的证明是对 Cheeger 不等式的简单应用. 固定任意顶点 $u \in V$. 定义一个集合 $S \triangleq \{\sigma \in \Omega_{\mathcal{I}} \mid \sigma_u = 1\}$. 换句话说, S 表示所有独立集 I 的集合, 其中 $u \in I$. 注意, 对于任意的 $\sigma \in S$, 存在唯一的 $\tau \in \Omega \setminus S$ 使得 $P(\sigma, \tau) > 0$, 其中 $\tau_u = 0$, 且对于所有 $v \in V \setminus \{u\}$, $\tau_v = \sigma_v$. 根据 Glauber dynamics 的转移规则, 有 $P(\sigma, \tau) \leq \frac{1}{n}$. 因此, 集合 S 的导通率可以被限制为:

$$\Phi(S) = \frac{\sum_{\sigma \in S, \tau \in \Omega \setminus S} \mu(\sigma)P(\sigma, \tau)}{\sum_{\sigma \in S} \mu(\sigma)} \leq \frac{\sum_{\sigma \in S} \mu(\sigma) \frac{1}{n}}{\sum_{\sigma \in S} \mu(\sigma)} = \frac{1}{n}.$$

类似地, 集合 $\Omega \setminus S$ 的导通率可以被限制为 $\Phi(\Omega \setminus S) = \frac{\sum_{\sigma \in \Omega \setminus S, \tau \in S} \mu(\sigma)P(\sigma, \tau)}{\sum_{\sigma \in \Omega \setminus S} \mu(\sigma)} \leq \frac{1}{n}$. 于是, Glauber dynamics 的导通率 Φ 满足:

$$\Phi = \min_{H \subseteq \Omega: \mu(H) \leq 1/2} \Phi(H) \leq \max\{\Phi(S), \Phi(\Omega \setminus S)\} \leq \frac{1}{n}.$$

根据 Cheeger 不等式 (命题 4), 我们有 $\gamma_{\mathcal{I}} \leq 2\Phi = \frac{2}{n}$.

4 Glauber dynamics 在 k -变换实例上的谱隙

本节使用以下两个结果证明引理 6. 第 1 个结果 (命题 5) 表明, 在 k -变换之后, 新实例仍然满足唯一性条件. 第 2 个结果 (引理 8) 表明, 如果一个硬核模型实例满足唯一性条件并且具有足够大的最大度数, 则相应 Glauber dynamics 的谱隙有下界.

命题 5. 设 $0 < \delta < 1$ 为常数, $k \in \mathbb{N}^+$ 为正整数. 对于任意实例 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 其中图 G 的最大度数 $\Delta_G \geq \frac{8}{\delta}$, 新实例 $\mathcal{I}' = (G', \lambda') \triangleq \text{Trans}(\mathcal{I}, k)$ 满足 $\mathcal{I}' \in \mathcal{F}(\delta/2)$.

证明: 设 Δ 和 Δ' 分别表示图 G 和 G' 的最大度数. 可以验证 $\Delta' = k(\Delta + 1) - 1$. 注意到 $\frac{e}{\Delta-2} \leq \lambda_c(\Delta) \leq \frac{e(\Delta-1)}{(\Delta-2)^2}$. 因此:

$$\frac{\lambda_c(\Delta)}{k\lambda_c(\Delta')} \leq \frac{(\Delta-1)(\Delta'-2)}{k(\Delta-2)^2} \leq \left(1 - \frac{2}{\Delta}\right)^{-2} \stackrel{(*)}{\leq} \frac{1}{1 - \frac{4}{\Delta}} \leq \frac{1}{1 - \frac{4}{\frac{8}{\delta} - 4}} = \frac{1 - \frac{\delta}{2}}{1 - \delta},$$

其中, (*) 成立是由于伯努利不等式和 $\Delta \geq 8/\delta \geq 8$ 这一事实. 因此:

$$\lambda' = \frac{\lambda}{k} \leq \frac{(1-\delta)\lambda_c(\Delta)}{k} \leq \left(1 - \frac{\delta}{2}\right)\lambda_c(\Delta'). \quad \square$$

引理 8. 令 $0 < \delta < 1$ 为常数. 存在常数 $C_2(\delta) = \exp(O(1/\delta))$, 使得对于任意 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$, 如果 $\Delta \geq \frac{12000}{\delta} \log n$, 则 $\mathcal{I}$ 上 Glauber dynamics 的谱隙至少为 $\frac{1}{C_2(\delta)n}$, 其中 Δ 是图 G 的最大度数, n 是图 G 的顶点数.

我们将引理 8 的证明推迟到第 4.1 节中进行, 本文下面证明引理 6.

证明: 令 $\mathcal{I} = (G, \lambda) \in \mathcal{F}(\delta)$ 以及 $\mathcal{I}' = (G', \lambda') \triangleq \text{Trans}\left(\mathcal{I}, \left\lceil \frac{10^6}{\delta^2} \log n \right\rceil\right)$ 是通过 $\left\lceil \frac{10^6}{\delta^2} \log n \right\rceil$ -变换得到的实例, 其中 n

是 G 中顶点的数目. 通过变换, G' 中的顶点数为 $n' = \left\lceil \frac{10^6}{\delta^2} \log n \right\rceil \cdot n$. 通过命题 5, 我们知道如果 $\Delta \geq \frac{8}{\delta}$, 那么 $\mathcal{I}' = (G', \lambda') \in \mathcal{F}(\delta/2)$. 记 $\delta' = \delta/2$, Δ' 表示 G' 的最大度数. 我们声明以下结论:

$$\Delta' \geq \left\lceil \frac{12000}{\delta'} \log n' \right\rceil \quad (8)$$

通过引理 8, 可以知道 $\mathcal{I}'$ 上 Glauber dynamics 的谱隙至少为 $\frac{1}{C_2(\delta')n'}$, 其中 $C_2(\delta') > 0$ 是某个仅依赖于 $\delta' = \delta/2$ 的常数. 最后, 我们设置 $C_1(\delta) = C_2(\delta') = C_2(\delta/2) = \exp(O(1/\delta))$.

现在证明公式 (8). 根据 k -变换的定义, 有 $n' = \left\lceil \frac{10^6}{\delta^2} \log n \right\rceil \cdot n \leq \frac{2^{20}}{\delta^2} n \log n$. 由于 $0 < \delta < 1$, 以下不等式很容易验证:

$$\frac{40}{\delta} \log n \geq \log \left(n \cdot \frac{2^{20}}{\delta^2} \cdot \log n \right) \geq \log n'.$$

另外, 根据 k -变换的定义, 我们有 $\Delta_{G'} \geq \left\lceil \frac{10^6}{\delta^2} \log n \right\rceil$. 结合上述不等式, 我们有:

$$\Delta' \geq \frac{10^6}{\delta^2} \log n = \left(\frac{25000}{\delta} \right) \left(\frac{40}{\delta} \log n \right) \geq \frac{25000}{\delta} \cdot \log n' \stackrel{(*)}{=} \frac{12500}{\delta'} \log n' \geq \left\lceil \frac{12000}{\delta'} \log n' \right\rceil,$$

其中, 标 $(*)$ 等式成立是因为 $\delta = 2\delta'$. □

4.1 引理 8 的证明

固定一个常数 $0 < \delta < 1$ 及一个最大度数为 $\Delta \geq \frac{12000}{\delta} \log n$ 的硬核模型实例 $\mathcal{I} = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$, 其中 $n = |V|$ 是 G 中顶点的数量. 我们注意到 $n \geq \Delta \geq \frac{12000}{\delta}$. 令 $\mu = \mu_{\mathcal{I}}$ 表示其吉布斯分布. 令 $\Omega = \Omega_{\mathcal{I}} \subseteq \{0, 1\}^V$ 表示 μ 的支撑集. 根据定义 1, 对于任意整数 $1 \leq \ell \leq n$, 在分布 μ 上定义 ℓ -block dynamics P_{ℓ} . 具体地, P_{ℓ} 从任意可行配置 $X \in \Omega$ 开始; 在每个更新步骤中, 它执行以下操作:

- 从大小为 ℓ 的 $\binom{V}{\ell}$ 中均匀随机选择一个集合 S .
- 从条件分布 $\mu_S^{X_{V \setminus S}}$ 中重新采样 X_S .

我们证明的起点是引理 9.

引理 9. 如果 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$, 那么 P_{ℓ} 的谱隙至少为 $350^{-67/\delta}$, 这意味着其弛豫时间为:

$$t_{\text{rel}}(P_{\ell}) \leq 350^{\frac{67}{\delta}}.$$

通过参考文献 [9,37] 中开发的局部到全局技术, 可以证明引理 9. 为了完整起见, 我们在附录 A 中证明引理 9. 使用 $P = P_1$ 表示吉布斯分布 μ 的 Glauber 动力学 (1-block dynamics). 引理 9 给出了 block dynamics P_{ℓ} 的谱隙, 其中 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$. 我们通过比较 P_{ℓ} 和 P 的弛豫时间来证明引理 8. 严格来说, 可以有引理 10.

引理 10. 如果 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$, 那么弛豫时间 $t_{\text{rel}}(P_{\ell})$ 和 $t_{\text{rel}}(P)$ 满足:

$$t_{\text{rel}}(P) \leq \frac{3n^2}{\ell} t_{\text{rel}}(P_{\ell}).$$

引理 10 的证明将推迟到第 4.2 节. 我们现在证明引理 8. 结合引理 9 和引理 10, Glauber dynamics 的弛豫时间可以被界定为:

$$t_{\text{rel}}(P) \leq \frac{3n^2}{\ell} t_{\text{rel}}(P_{\ell}) \leq 300n \cdot t_{\text{rel}}(P_{\ell}) \leq 350^{\frac{70}{\delta}} n.$$

这意味着 Glauber dynamics 的谱隙至少为 $\frac{1}{C_4(\delta)n}$, 其中 $C_4(\delta) = 350^{\frac{70}{\delta}} = \exp(O(1/\delta))$.

4.2 Block dynamics 和 Glauber dynamics 之间的比较 (引理 10 的证明)

令 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$, 根据命题 3, block dynamics P_{ℓ} 是不可约的. 对于任何 $f: \Omega \rightarrow \mathbb{R}$, 结合引理 1 和引理 2, 可以得到以下不等式:

$$\text{Var}_\mu[f] \leq t_{\text{rel}}(P_\ell) \cdot \mathcal{E}_{P_\ell}(f, f) \leq t_{\text{rel}}(P_\ell) \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{\rho \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\rho) \cdot \text{Var}_{\mu_S^\rho}[f_\rho] \quad (9)$$

其中, $\Omega_{V \setminus S}$ 表示边缘分布 $\mu_{V \setminus S}$ 的支撑集, $f_\rho: \Omega_S^\rho \rightarrow \mathbb{R}$ 的定义为 $f_\rho(\tau) \triangleq f(\rho \uplus \tau)$, Ω_S^ρ 表示条件分布 μ_S^ρ 的支撑集. 对于任何 $S \in \binom{V}{\ell}$ 和 $\rho \in \Omega_{V \setminus S}$, 令 P_S^ρ 表示条件分布 μ_S^ρ 上的 Glauber dynamics, 其中 Glauber dynamics 是定义 1 中的 1-block dynamics. 根据命题 3, Glauber 动力学 P_S^ρ 是不可约的. 注意到 $|S| = \ell$. 再次根据引理 1 和引理 2, 对于任何 $f_\rho: \Omega_S^\rho \rightarrow \mathbb{R}$, 有:

$$\text{Var}_{\mu_S^\rho}[f_\rho] \leq t_{\text{rel}}(P_S^\rho) \cdot \mathcal{E}_{P_S^\rho}(f_\rho, f_\rho) \leq t_{\text{rel}}(P_S^\rho) \cdot \frac{1}{\ell} \sum_{v \in S} \sum_{\sigma \in \Omega_{S \setminus \{v\}}^\rho} \mu_{S \setminus \{v\}}^\rho(\sigma) \cdot \text{Var}_{\mu_v^{\rho \uplus \sigma}}[f_{\rho \uplus \sigma}] \quad (10)$$

其中, $\mu_{S \setminus \{v\}}^\rho$ 是从 μ_S^ρ 投影到 $S \setminus \{v\}$ 上的边缘分布; $\Omega_{S \setminus \{v\}}^\rho \subseteq \{0, 1\}^{S \setminus \{v\}}$ 是 $\mu_{S \setminus \{v\}}^\rho$ 的支撑集. 函数 $f_{\rho \uplus \sigma}: \Omega_v^{\rho \uplus \sigma} \rightarrow \mathbb{R}$ 的定义为 $f_{\rho \uplus \sigma}(c_v) = f_\rho(\sigma \uplus c_v)$, 其中 $\Omega_v^{\rho \uplus \sigma} \subseteq \{0, 1\}$ 是 $\mu_v^{\rho \uplus \sigma}$ 的支撑集; $\sigma \uplus c_v$ 表示 S 上的配置, 其中 v 的值为 $c_v \in \{0, 1\}$, $S \setminus \{v\}$ 上的配置为 σ . 对于任何 $S \subseteq \binom{V}{\ell}$, 定义最大弛豫时间为:

$$t_{\text{rel}}^{\max}(S) = \max_{\rho \in \Omega_{V \setminus S}} t_{\text{rel}}(P_S^\rho) \quad (11)$$

注意到, 方差是非负的. 结合公式 (9)–(11), 对于任意 $f: \Omega \rightarrow \mathbb{R}$, 有:

$$\begin{aligned} \text{Var}_\mu[f] &\leq t_{\text{rel}}(P_\ell) \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{\rho \in \Omega_{V \setminus S}} \mu_{V \setminus S}(\rho) \cdot t_{\text{rel}}^{\max}(S) \cdot \frac{1}{\ell} \sum_{v \in S} \sum_{\sigma \in \Omega_{S \setminus \{v\}}^\rho} \mu_{S \setminus \{v\}}^\rho(\sigma) \cdot \text{Var}_{\mu_v^{\rho \uplus \sigma}}[f_{\rho \uplus \sigma}] \\ (*) &= t_{\text{rel}}(P_\ell) \cdot \frac{1}{\binom{n}{\ell}} \cdot \frac{1}{\ell} \sum_{S \in \binom{V}{\ell}} t_{\text{rel}}^{\max}(S) \sum_{v \in S} \sum_{\substack{\rho \in \Omega_{V \setminus S} \\ \sigma \in \Omega_{S \setminus \{v\}}^\rho}} \mu_{V \setminus \{v\}}(\rho \uplus \sigma) \cdot \text{Var}_{\mu_v^{\rho \uplus \sigma}}[f_{\rho \uplus \sigma}] \\ (\text{令 } \tau \triangleq \rho \uplus \sigma) &\leq t_{\text{rel}}(P_\ell) \cdot \frac{1}{\binom{n}{\ell}} \cdot \frac{1}{\ell} \sum_{S \in \binom{V}{\ell}} t_{\text{rel}}^{\max}(S) \sum_{v \in V} \sum_{\tau \in \Omega_{V \setminus \{v\}}} \mu_{V \setminus \{v\}}(\tau) \cdot \text{Var}_{\mu_v^\tau}[f_\tau]. \end{aligned}$$

注意, 在 (*) 中, 我们列举了集合 S 中的所有顶点 v , 但在最后一行中, 我们列举了集合 V 中的所有顶点 v . 由于 $S \subseteq V$ 且所有变量都是非负的, 最后一个不等式成立. 设 $S \in \binom{V}{\ell}$ 是从 $\binom{V}{\ell}$ 中均匀随机抽取的随机集合. 令 P 为 Glauber dynamics (1-block dynamics) 在 μ 上的转移矩阵, $t_{\text{rel}}(P)$ 表示其弛豫时间. 通过上述不等式, 对于任意函数 $f: \Omega \rightarrow \mathbb{R}$, 有:

$$\begin{aligned} \text{Var}_\mu[f] &\leq t_{\text{rel}}(P_\ell) \cdot \frac{n}{\ell} \cdot \mathbb{E}_S[t_{\text{rel}}^{\max}(S)] \cdot \frac{1}{n} \sum_{v \in V} \sum_{\tau \in \Omega_{V \setminus \{v\}}} \mu_{V \setminus \{v\}}(\tau) \cdot \text{Var}_{\mu_v^\tau}[f_\tau] \\ (\text{根据引理 2}) &= t_{\text{rel}}(P_\ell) \cdot \frac{n}{\ell} \cdot \mathbb{E}_S[t_{\text{rel}}^{\max}(S)] \cdot \mathcal{E}_P(f, f). \end{aligned}$$

其中, $\mathcal{E}_P(f, f)$ 是 Glauber dynamics P 的狄利克雷形. 通过公式 (5) 中 P 的 Poincaré 不等式和公式 (1) 中弛豫时间的定义, Glauber dynamics 的弛豫时间 $t_{\text{rel}}(P)$ 满足:

$$t_{\text{rel}}(P) \leq t_{\text{rel}}(P_\ell) \cdot \frac{n}{\ell} \cdot \mathbb{E}_S[t_{\text{rel}}^{\max}(S)].$$

为了证明引理 10, 只需证明:

$$\mathbb{E}_S[t_{\text{rel}}^{\max}(S)] \leq 3n \quad (12)$$

为了给出 $\mathbb{E}_S[t_{\text{rel}}^{\max}(S)]$ 的上界, 我们将所有 $S \in \binom{V}{\ell}$ 分类为好和坏情况. 固定一个 $S \in \binom{V}{\ell}$, 让 $G[S]$ 表示 G 在 S 上的导出子图. 令 $\Delta_{G[S]}$ 表示 $G[S]$ 的最大度. 令 λ 为硬核模型实例的参数. 我们通过以下方式判断集合 S 是好还是坏.

- 好集合: 如果 $\lambda < \frac{1}{2(\Delta_{G[S]} + 2)}$, 则 S 是好的.

- 坏集合: 如果 $\lambda \geq \frac{1}{2(\Delta_{G[S]} + 2)}$, 则 S 是坏的.

通过这个定义, 我们可以定义以下两个集合:

$$\mathcal{G} \triangleq \left\{ S \in \binom{V}{\ell} \mid S \text{ 是好的} \right\}, \mathcal{B} \triangleq \left\{ S \in \binom{V}{\ell} \mid S \text{ 是坏的} \right\}.$$

我们断言对于好和坏的集合有以下结果.

引理 11. 对于 $\mathcal{G}$ 和 $\mathcal{B}$, 以下结果成立.

- 对于任何 $S \in \mathcal{G}$, $t_{\text{rel}}^{\max}(S) \leq 2\ell$.
- 对于任何 $S \in \mathcal{B}$, $t_{\text{rel}}^{\max}(S) \leq \ell^{140/\delta}$.

引理 11 的证明被推迟至第 4.3 节. 我们现在使用引理 11 来证明公式 (12), 有:

$$E_S[t_{\text{rel}}^{\max}(S)] \leq \Pr_S[S \in \mathcal{G}] \cdot 2\ell + \Pr_S[S \in \mathcal{B}] \cdot \ell^{140/\delta} \quad (13)$$

其中, S 是从 $\binom{V}{\ell}$ 中均匀随机抽样的. 为了证明公式 (12), 需要界定 $S \in \mathcal{B}$ 的概率. 令 $\mathcal{I} = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$ 表示原始的硬核实例; Δ_G 表示图 G 的最大度数. 在引理 8 的假设下, 有 $\Delta_G \geq \frac{12000}{\delta} \cdot \log n \geq 12000$. 注意到 $\lambda \leq (1 - \delta)\lambda_c(\Delta_G) \leq \frac{4}{\Delta_G - 2}$, 很容易验证如果 $\lambda \geq \frac{1}{2(\Delta_{G[S]} + 2)}$, 那么必然有 $\Delta_{G[S]} \geq \frac{1}{16}\Delta_G$. 对于每个顶点 $v \in V$, 定义随机变量 $D_v = |\{S \cap \Gamma_G(v)\}|$, 其中 $\Gamma_G(v)$ 是图 G 中顶点 v 的邻域, 而 D_v 计算了被随机集合 S 选择的顶点 v 的邻居数量. 我们有:

$$\Pr_S[S \in \mathcal{B}] = \Pr_S\left[\lambda \geq \frac{1}{2(\Delta_{G[S]} + 1)}\right] \leq \Pr_S\left[\Delta_{G[S]} \geq \frac{\Delta_G}{16}\right] \leq \sum_{v \in V} \Pr_S\left[D_v \geq \frac{\Delta_G}{16}\right].$$

因为 $\ell = \lfloor \frac{n}{32e} \rfloor$, 所以 $E_S[D_v] \leq \frac{\ell \Delta_G}{n} \leq \frac{1}{2e} \frac{\Delta_G}{16}$. 又因为 $\Delta_G \geq \frac{12000}{\delta} \cdot \log n$, 所以根据引理 3, 有:

$$\Pr_S[S \in \mathcal{B}] = \sum_{v \in V} 2^{-\Delta_G/16} \leq n^{-300/\delta} \quad (14)$$

结合公式 (13)、公式 (14) 和 $\ell = \lfloor \frac{n}{32e} \rfloor$, 我们可以证明公式 (12) 如下:

$$E[t_{\text{rel}}^{\max}(S)] \leq 2\ell + n^{-300/\delta} \cdot \ell^{140/\delta} \leq 3n.$$

这就证明了引理 10.

4.3 好情况和坏情况的弛豫时间 (引理 11 的证明)

回忆在第 4.1 节开始时固定的硬核模型实例 $\mathcal{I} = (G = (V, E), \lambda)$. 固定一个集合 $S \in \binom{V}{\ell}$. 我们使用 $\sigma \in \Omega_{V \setminus S}$ 表示在公式 (11) 中达到最大值的配置, 即:

$$t_{\text{rel}}^{\max}(S) = \max_{\rho \in \Omega_{V \setminus S}} t_{\text{rel}}(P_S^\rho) = t_{\text{rel}}(P_S^\sigma).$$

我们将界定弛豫时间 $t_{\text{rel}}(P_S^\sigma)$, 将集合 S 划分为两部分:

$$\begin{cases} A \triangleq \{v \in S \mid \forall u \in \Gamma_G(v) \setminus S, \sigma_u = 0\} \\ B \triangleq \{v \in S \mid \exists u \in \Gamma_G(v) \setminus S, \sigma_u = 1\} \end{cases}.$$

其中, $\Gamma_G(v)$ 是原始图 G 中节点 v 的邻域. 很容易看出, 在条件分布 μ_S^σ 中, $\tau = 0 \in \{0, 1\}^B$ 是 B 上唯一可行的配置. 如果 $|A| = 0$, 那么 μ_S^σ 的支撑集只有一个状态. 根据公式 (1) 中的定义, $t_{\text{rel}}(P_S^\sigma) = 0$, 引理 11 是显然成立的. 现在假设 $|A| > 0$. 考虑分布 $\mu_A^{\sigma \cup \tau}$. 让 $P_A^{\sigma \cup \tau}$ 表示条件分布 $\mu_A^{\sigma \cup \tau}$ 上的 Glauber dynamics. 虽然 P_S^σ 的状态空间是 $\{0, 1\}^S$, 但 B 上的配置始终固定为 $\tau = 0$. 我们可以将 P_S^σ 视为状态空间 $\{0, 1\}^A$ 上的马尔可夫链. 因此, P_S^σ 是 $P_A^{\sigma \cup \tau}$ 的懒惰版本. 形式上, 在每一步中, 以 $\frac{|B|}{|S|}$ 的概率, P_S^σ 保持在当前状态; 以 $\frac{|A|}{|S|}$ 的概率, P_S^σ 以与链 $P_A^{\sigma \cup \tau}$ 相同的方式演变. 以下关系容易验证:

$$\lambda_2(P_S^\sigma) = \frac{|A|}{|S|} \cdot \lambda_2(P_A^{\sigma \cup \tau}) + \left(1 - \frac{|A|}{|S|}\right).$$

注意到 P_S^σ 和 $P_A^{\sigma \cup \tau}$ 都是不可约的, 因此 $\lambda_2(P_S^\sigma), \lambda_2(P_A^{\sigma \cup \tau}) < 1$. 我们有:

$$t_{\text{rel}}(P_S^\sigma) = \frac{|S|}{|A|} \cdot t_{\text{rel}}(P_A^{\sigma\omega\tau}) = \frac{\ell}{|A|} \cdot t_{\text{rel}}(P_A^{\sigma\omega\tau}) \quad (15)$$

我们现在要界定 $t_{\text{rel}}(P_A^{\sigma\omega\tau})$. 注意到 $P_A^{\sigma\omega\tau}$ 恰好是硬核模型实例 $\mathcal{I} = (G[A], \lambda)$ 上的 Glauber dynamics, 其中 $G[A]$ 是 G 在 A 上的导出子图. 注意到 $\Delta_{G[A]} \leq \Delta_{G[S]} \leq \Delta_G$.

如果 $S \in \mathcal{G}$. 则有 $\lambda \leq \frac{1}{2(\Delta_{G[S]} + 1)}$, 这意味着 $\lambda \leq \frac{1}{2(\Delta_{G[A]} + 1)}$. 在这种情况下, 我们要证明:

$$t_{\text{rel}}(P_A^{\sigma\omega\tau}) \leq 2|A| \quad (16)$$

然后, 结合公式 (15) 和公式 (16) 证明引理. 如果 $\Delta_{G[A]} = 0$, 那么 A 中的所有变量是独立的, 公式 (16) 显而易见. 如果 $\Delta_{G[A]} \geq 1$, 那么根据命题 1, 实例 $\mathcal{I}$ 满足 Dobrushin 条件, $\delta = 1/2$. 因此, 依然有弛豫时间 $t_{\text{rel}}(P_A^{\sigma\omega\tau}) \leq 2|A|$.

假设 $S \in \mathcal{B}$. 由于 $\mathcal{I}$ 在唯一性区域内, 我们有 $\lambda \leq (1-\delta)\lambda_c(\Delta_G) \leq (1-\delta)\lambda_c(\Delta_{G[A]})$. 因此, 实例 $\mathcal{J} \in \mathcal{F}(\delta)$ 也处于唯一性区域内. 我们考虑以下两种情况.

• 情况 1: $\Delta_{G[A]} \geq 3$. 根据命题 2, 有 $t_{\text{rel}}(P_A^{\sigma\omega\tau}) \leq (5|A|)^{66/\delta}$.

• 情况 2: $\Delta_{G[A]} \leq 2$. 注意到 $\lambda < \lambda_c(\Delta_G)$. 由于 $\Delta_G \geq 20$, 有 $\lambda \leq \lambda_c(20) \leq \frac{1}{6}$. 由于 $\Delta_{G[A]} \leq 2$, 有 $\lambda \leq \frac{1}{6} \leq \frac{1}{2(\Delta_{G[A]} + 1)}$.

根据该情况的证明, 我们有 $t_{\text{rel}}(P_A^{\sigma\omega\tau}) \leq 2|A| \leq (5|A|)^{66/\delta}$.

根据公式 (15) 和 $|A| \leq \ell$, 我们有:

$$t_{\text{rel}}(P_S^\sigma) \leq \frac{\ell}{|A|} \cdot (5|A|)^{66/\delta} \leq (5\ell)^{70/\delta}.$$

因为 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$ 和 $n \geq 12000$, 这意味着 $\ell \geq 100$. 所以我们有 $t_{\text{rel}}(P_{\mu_S^\sigma}) \leq \ell^{140/\delta}$.

5 Glauber dynamics 之间的比较

在本节中, 我们证明引理 7. 设 $\mathcal{I} = (G = (V, E), \lambda)$ 是图 G 上的硬核模型实例. 令 $k \in \mathbb{N}^+$ 是一个正整数. 在实例 $\mathcal{I}$ 上使用 k -变换 (定义 3), 我们定义以下转换后的实例:

$$\mathcal{I}' = (G', \lambda') \triangleq \text{Trans}(\mathcal{I}, k), \quad G' = (V', E'), \quad \lambda' = \frac{\lambda}{k}.$$

令 μ 和 μ' 分别表示由 $\mathcal{I}$ 和 $\mathcal{I}'$ 诱导的吉布斯分布, 其中 μ 的支撑集是 Ω , μ' 的支撑集是 Ω' . 令 P 和 P' 分别表示在 $\mathcal{I}$ 和 $\mathcal{I}'$ 上的 Glauber dynamics 的转移矩阵. 我们证明以下一般情形下的比较引理.

引理 12. 比较引理. 对于任何 $k \in \mathbb{N}^+$, 任何硬核模型实例 $\mathcal{I} = (G, \lambda)$, 让 $\mathcal{I}' = \text{Trans}(\mathcal{I}, k)$ 表示变换后的实例, P 和 P' 的谱隙满足:

$$1 - \lambda_2(P) \geq \frac{k + \lambda}{1 + \lambda} (1 - \lambda_2(P')).$$

注意, 上述引理对于任何 $k \in \mathbb{N}^+$ 和任何硬核模型实例 $\mathcal{I}$ 都成立, 即使 $\mathcal{I}$ 超出了唯一性阈值. 很容易验证引理 7 是引理 12 的推论. 让 Δ 表示 G 的最大度数. 在引理 7 中, 对于 $\Delta \geq 3$, 有 $\lambda < \lambda_c(\Delta_G) \leq 4$. 根据公式 (6), 我们有 $\kappa(\mathcal{I}) = k \geq 5$, 因此 $\frac{k + \lambda}{1 + \lambda} \geq \frac{k}{5}$. 本节的其余部分将专注于证明引理 12. 为了证明引理 12, 引入一个新的链 P_{lazy} , 该链位于状态空间 Ω 上, 是 Glauber dynamics P 的一个懒化版本. 定义 $r \triangleq 1 - \frac{1 + \lambda}{k + \lambda}$, 可以很容易验证 $0 \leq r < 1$. P_{lazy} 的转移规则定义如公式 (17). 在每个转移步骤中, 以概率 r , 链 P_{lazy} 停留在当前状态; 否则, 它将执行与链 P 相同的更新. 转移矩阵 P_{lazy} 和 P 有以下关系:

$$P_{\text{lazy}} = rI + (1 - r)P \quad (17)$$

其中, $I \in \mathbb{R}^{\Omega \times \Omega}$ 为单位矩阵. 很容易验证 P_{new} 是不可约的、非周期性的, 并且关于 μ 是时间可逆的. 我们首先比较 P 和 P_{lazy} , 然后比较 P_{lazy} 和 P' . 根据公式 (17) 中 P_{lazy} 的定义, 容易验证引理 13.

引理 13. $1 - \lambda_2(P) = \frac{1}{1 - r} (1 - \lambda_2(P_{\text{lazy}})) = \frac{k + \lambda}{1 + \lambda} (1 - \lambda_2(P_{\text{lazy}}))$.

然后, 我们有引理 14.

引理 14. $1 - \lambda_2(P_{\text{lasy}}) \geq 1 - \lambda_2(P')$.

引理 12 是通过结合引理 13 和引理 14 来证明的. 第 5.1 节专门用来证明引理 14.

5.1 引理 14 的证明

我们使用引理 15 来证明引理 14. 回想一下, P_{lasy} 的状态空间是 Ω , P_{lasy} 的稳态分布是 μ ; P' 的状态空间是 Ω' , P' 的稳态分布是 μ' ; 狄利克雷形 $\mathcal{E}(\cdot, \cdot)$ 和方差 $\text{Var}[\cdot]$ 在公式 (4) 和公式 (3) 中定义.

引理 15. 对于任何满足 $\text{Var}_\mu[f] \neq 0$ 的函数 $f: \Omega \rightarrow \mathbb{R}$, 存在一个函数 $f': \Omega' \rightarrow \mathbb{R}$, 使得 $\text{Var}_{\mu'}[f'] \neq 0$, 并且:

$$\frac{\mathcal{E}_{P_{\text{lasy}}}(f, f)}{\text{Var}_\mu(f)} = \frac{\mathcal{E}_{P'}(f', f')}{\text{Var}_{\mu'}(f')}.$$

引理 15 的证明见第 5.2 节. 现在我们来证明引理 14.

证明: 定义 $W_\Omega = \{f: \Omega \rightarrow \mathbb{R} \mid \text{Var}_\mu[f] \neq 0\}$ 以及 $W_{\Omega'} = \{f: \Omega' \rightarrow \mathbb{R} \mid \text{Var}_{\mu'}[f] \neq 0\}$. 根据引理 15, 存在一个从 W_Ω 到 $W_{\Omega'}$ 的映射 h , 对于任何 $f \in W_\Omega$, 我们有:

$$\frac{\mathcal{E}_{P_{\text{lasy}}}(f, f)}{\text{Var}_\mu(f)} = \frac{\mathcal{E}_{P'}(h(f), h(f))}{\text{Var}_{\mu'}(h(f))},$$

其中, $h(f) = f'$, f' 是引理 15 中的函数. 我们使用 $h(W_\Omega)$ 来表示 h 的像空间. 由于 $h(W_\Omega) \subseteq W_{\Omega'}$, 有以下关系:

$$1 - \lambda_2(P_{\text{lasy}}) = \inf_{f \in W_\Omega} \frac{\mathcal{E}_{P_{\text{lasy}}}(f, f)}{\text{Var}_\mu[f]} = \inf_{f' \in h(W_\Omega)} \frac{\mathcal{E}_{P'}(f', f')}{\text{Var}_{\mu'}(f')} \geq \inf_{f' \in W_{\Omega'}} \frac{\mathcal{E}_{P'}(f', f')}{\text{Var}_{\mu'}(f')} = 1 - \lambda_2(P'). \quad \square$$

5.2 证明引理 15

为了证明引理 15, 我们定义函数 g . 回想一下 $\mathcal{I} = (G = (V, E), \lambda)$ 是原始的硬核模型实例, $\mathcal{I}' = (G' = (V', E'), \lambda') = \text{Trans}(\mathcal{I}, k)$ 是转化后的实例. 对于任意顶点 $v \in V$, 图 G' 中有一个包含顶点 $v_1, v_2, \dots, v_k$ 的团.

定义 4. 设 $g: \Omega' \rightarrow \Omega$ 是一个函数, 使得对于任何 $\sigma' \in \Omega'$, $\sigma = g(\sigma')$ 构造如下:

$$\forall v \in V, \sigma_v \triangleq \begin{cases} 1, & \text{如果存在 } 1 \leq i \leq k \text{ 使得 } \sigma'(v_i) = 1 \\ 0, & \text{否则} \end{cases}.$$

对于任意 $\tau \in \Omega$, 我们使用 $g^{-1}(\tau) \subseteq \Omega'$ 来表示 τ 的原像, 则以下性质成立.

引理 16. 函数 g 是一个从 Ω 到 Ω' 的映射, 并且满足如下条件.

- 对于任意 $\sigma \in \Omega$, 都有:

$$\sum_{\sigma' \in g^{-1}(\sigma)} \mu'(\sigma') = \mu(\sigma) \quad (18)$$

这意味着如果 $X' \sim \mu'$, 那么 $g(X') \sim \mu$.

- 对于任意 $\sigma, \tau \in \Omega$, 任意 $\sigma' \in g^{-1}(\sigma)$, 都有:

$$\sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') = P_{\text{lasy}}(\sigma, \tau) \quad (19)$$

这意味着如果 $(X'_t)_{t \geq 0}$ 是马尔可夫链 P' , 那么 $(g(X'_t))_{t \geq 0}$ 是马尔可夫链 P_{lasy} .

证明: 根据定义 3, 对于图 $G = (V, E)$ 中的任意顶点 v , $G' = (V', E')$ 中都存在一个包含顶点 $v_1, v_2, \dots, v_k$ 的 k -团 C_v . 对于图 G 中的任意边 $\{u, v\}$, 在 G' 中, C_u 和 C_v 中的所有顶点对都是相邻的. 对于任意配置 $\sigma' \in \Omega' \subseteq \{0, 1\}^{V'}$, 集合 $S'(\sigma') \triangleq \{v \in V' \mid \sigma'_v = 1\}$ 是 G' 中的一个独立集, 这意味着如果 C_v 中的某个顶点在 $S'(\sigma')$ 中, 那么对于 G 中的所有边 $\{u, v\}$, C_u 中的任意顶点都不在 $S'(\sigma')$ 中. 根据 g 的定义, 很容易看出 $g(\sigma')$ 表示 G 中的一个独立集. 这意味着 g 是一个从 Ω' 到 Ω 的映射. 对于任意配置 $\sigma \in \Omega$, 我们用 $|\sigma|$ 表示 $\sum_{v \in V} \sigma_v$, 它的权重定义为 $w(\sigma) \triangleq \lambda^{|\sigma|}$. 让 $Z \triangleq \sum_{\sigma \in \Omega} w(\sigma)$ 表示硬核模型实例 $\mathcal{I} = (G = (V, E), \lambda)$ 的配分函数. 对于任意配置 $\sigma' \in \Omega'$, 我们用 $|\sigma'|$ 表示 $\sum_{v \in V'} \sigma'_v$, 它的权重定义为 $w'(\sigma') \triangleq (\lambda')^{|\sigma'|} = (\lambda/k)^{|\sigma'|}$. 让 $Z' \triangleq \sum_{\sigma' \in \Omega'} w'(\sigma')$ 表示硬核模型实例 $\mathcal{I}' = (G' = (V', E'), \lambda' = \lambda/k)$ 的配分函数. 固定配置 $\sigma \in \Omega$. 考虑集合 $g^{-1}(\sigma) \subseteq \Omega'$. 假设 $\sigma' \in g^{-1}(\sigma)$. 如果 $\sigma_v = 1$, 那么团 C_v 中有且只有一个顶点 v_i 满足 $\sigma'_{v_i} = 1$; 如果 $\sigma_v = 0$, 那么对于 C_v 中的所有顶点 v_i , 都有 $\sigma'_{v_i} = 0$. 因此, 我们有:

$$\begin{cases} |g^{-1}(\sigma)| = k^{|\sigma|} \\ \forall \sigma' \in g^{-1}(\sigma), w'(\sigma') = \left(\frac{\lambda}{k}\right)^{|\sigma|}. \end{cases}$$

对于任意 $\sigma \in \Omega$, 有:

$$w(\sigma) = \lambda^{|\sigma|} = k^{|\sigma|} \cdot \left(\frac{\lambda}{k}\right)^{|\sigma|} = \sum_{\sigma' \in g^{-1}(\sigma)} w'(\sigma').$$

由于 g 是从 Ω' 到 Ω 的映射, 我们有:

$$Z' = \sum_{\sigma' \in \Omega'} w'(\sigma') = \sum_{\sigma \in \Omega} \sum_{\sigma' \in g^{-1}(\sigma)} w'(\sigma') = \sum_{\sigma \in \Omega} w(\sigma) = Z.$$

这意味着, 对于任何 $\sigma \in \Omega$, 有:

$$\mu(\sigma) = \frac{w(\sigma)}{Z} = \frac{\sum_{\sigma' \in g^{-1}(\sigma)} w'(\sigma')}{Z'} = \sum_{\sigma' \in g^{-1}(\sigma)} \mu'(\sigma').$$

然后, 证明公式 (19). 固定两个配置 $\sigma, \tau \in \Omega$. 令 $\sigma \oplus \tau \triangleq \{v \in V \mid \sigma_v \neq \tau_v\}$ 表示 σ 和 τ 之间的不一致顶点集. 固定一个配置 $\sigma' \in g^{-1}(\sigma)$. 我们考虑以下 3 种情况 (以 $|\sigma \oplus \tau|$ 分类).

• 假设 $|\sigma \oplus \tau| > 1$. 在这种情况下, $P_{\text{lazy}}(\sigma, \tau) = 0$, 因为 P_{lazy} 每一步只更新一个顶点. 固定一个配置 $\tau' \in g^{-1}(\tau)$. 很容易验证 $|\sigma' \oplus \tau'| \geq |\sigma \oplus \tau| > 1$. 因此 $P'(\sigma', \tau') = 0$, 最后性质成立.

• 假设 $|\sigma \oplus \tau| = 1$. 这里有两种情况. 首先假设存在一个顶点 $v \in V$, 使得 $\sigma_v = 1, \tau_v = 0$, 且对所有 $u \in V \setminus \{v\}$, $\sigma_u = \tau_u$. 在这种情况下, 根据公式 (17), 有 $P_{\text{lazy}}(\sigma, \tau) = \frac{(1-r)}{n(1+\lambda)}$, 其中 $n = |V|$. 考虑链 P' , 回顾此前, 我们有 $\sigma' \in g^{-1}(\sigma)$. 在团 C_v 中恰有一个顶点 v_i , 使得 $\sigma'_{v_i} = 1$. 在链 P' 中, 从 σ' 移动到 $\tau' \in g^{-1}(\tau)$ 的唯一可能性是 P' 选择顶点 v_i 并将 v_i 的值更新为 0. 记为 $r \triangleq 1 - \frac{1+\lambda}{k+\lambda}$, $|V'| = kn$. 因此, 我们有:

$$\sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') = \frac{1}{kn(1+\lambda')} = \frac{1}{n(k+\lambda)} = \frac{1-r}{n(1+\lambda)} = P_{\text{lazy}}(\sigma, \tau).$$

现在假设存在一个顶点 $v \in V$, 使得 $\sigma_v = 0, \tau_v = 1$, 且对所有 $u \in V \setminus \{v\}$, $\sigma_u = \tau_u$. 在这种情况下, 根据公式 (17), 有 $P_{\text{lazy}}(\sigma, \tau) = \frac{(1-r)\lambda}{n(1+\lambda)}$. 考虑链 P' . 回顾此前, 我们有 $\sigma' \in g^{-1}(\sigma)$. 所有团 C_v 中的顶点 v_i 都满足 $\sigma'_{v_i} = 0$. 在链 P' 中, 从 σ' 移动到 $\tau' \in g^{-1}(\tau)$ 的唯一可能性是 P' 选择团 C_v 中的一个顶点 v_i 并将 v_i 的值更新为 1. 因此, 我们有:

$$\sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') = \frac{k\lambda'}{kn(1+\lambda')} = \frac{\lambda}{n(k+\lambda)} = \frac{(1-r)\lambda}{n(1+\lambda)} = P_{\text{lazy}}(\sigma, \tau).$$

• 假设 $|\sigma \oplus \tau| = 0$, 这意味着 $\sigma = \tau$. 注意马尔可夫链 P' 每一步只更新一个顶点. 因此, $P'(\sigma', \tau') > 0$ 当且仅当 $|\sigma' \oplus \tau'| \leq 1$. 这意味着 $P'(\sigma', \tau') > 0$, 当且仅当 $|g(\sigma') \oplus g(\tau')| = |\sigma \oplus \tau| \leq 1$. 由于 P' 是随机矩阵, 所以:

$$\sum_{\tau' \in \Omega'} P'(\sigma', \tau') = \sum_{\substack{\tau' \in \Omega' \\ |\sigma \oplus g(\tau')| \leq 1}} P'(\sigma', \tau') = \sum_{\tau' \in g^{-1}(\sigma)} P'(\sigma', \tau') + \sum_{\substack{\tau \in \Omega \\ |\sigma \oplus \tau| = 1}} \sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') = 1.$$

同样, P_{lazy} 也每一步只更新一个顶点, 我们有:

$$\sum_{\tau \in \Omega} P_{\text{lazy}}(\sigma, \tau) = P_{\text{lazy}}(\sigma, \sigma) + \sum_{\substack{\tau \in \Omega \\ |\sigma \oplus \tau| = 1}} P_{\text{lazy}}(\sigma, \tau) = 1.$$

因此, 可以证明:

$$\sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') = P_{\text{lazy}}(\sigma, \tau), \quad \forall \tau \in \Omega \text{ with } |\sigma \oplus \tau| = 1.$$

通过对比, 我们得出 $\sum_{\tau' \in g^{-1}(\sigma)} P'(\sigma', \tau') = P_{\text{lazy}}(\sigma, \sigma)$.

现在证明引理 15. 给定一个满足 $\text{Var}_{\mu}[f] \neq 0$ 的函数 $f: \Omega \rightarrow \mathbb{R}$, 定义 $f': \Omega' \rightarrow \mathbb{R}$ 为:

$$\forall \sigma' \in \Omega', f'(\sigma') = f(g(\sigma')) \quad (20)$$

于是, $f' = f \circ g$. 我们将证明以下两个性质成立, 这意味着引理 15 成立.

- $\mathcal{E}_{P'}(f', f') = \mathcal{E}_{P_{\text{lazy}}}(f, f)$.
- $\text{Var}_{\mu'}(f') = \text{Var}_{\mu}(f)$.

根据公式 (4) 中的定义, 我们有:

$$\begin{aligned} \mathcal{E}_{P'}(f', f') &= \frac{1}{2} \sum_{\sigma', \tau' \in \Omega'} \mu'(\sigma') P'(\sigma', \tau') (f'(\sigma') - f'(\tau'))^2 \\ (g \text{ 将 } \Omega' \text{ 映射到 } \Omega) &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} \sum_{\substack{\sigma' \in g^{-1}(\sigma) \\ \tau' \in g^{-1}(\tau)}} \mu'(\sigma') P'(\sigma', \tau') (f'(\sigma') - f'(\tau'))^2 \\ (\text{根据 } f' \text{ 的定义}) &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} \sum_{\substack{\sigma' \in g^{-1}(\sigma) \\ \tau' \in g^{-1}(\tau)}} \mu'(\sigma') P'(\sigma', \tau') (f(g(\sigma')) - f(g(\tau')))^2 \\ &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} (f(\sigma) - f(\tau))^2 \sum_{\sigma' \in g^{-1}(\sigma)} \mu'(\sigma') \sum_{\tau' \in g^{-1}(\tau)} P'(\sigma', \tau') \\ (\text{根据公式(18)和(19)}) &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} \mu(\sigma) P_{\text{lazy}}(\sigma, \tau) (f(\sigma) - f(\tau))^2 = \mathcal{E}_{P_{\text{lazy}}}(f, f). \end{aligned}$$

类似地, 对于方差, 根据公式 (3), 我们有:

$$\begin{aligned} \text{Var}_{\mu'}[f'] &= \frac{1}{2} \sum_{\sigma', \tau' \in \Omega'} \mu'(\sigma') \mu'(\tau') (f'(\sigma') - f'(\tau'))^2 \\ &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} (f(\sigma) - f(\tau))^2 \sum_{\sigma' \in g^{-1}(\sigma)} \mu'(\sigma') \sum_{\tau' \in g^{-1}(\tau)} \mu'(\tau') \\ &= \frac{1}{2} \sum_{\sigma, \tau \in \Omega} \mu(\sigma) \mu(\tau) (f(\sigma) - f(\tau))^2 = \text{Var}_{\mu}[f]. \end{aligned} \quad \square$$

6 总 结

将点替换成 k -团是传统计算复杂性规约证明中常见的 gadget. 本文通过在图上考虑这种 gadget 证明了 Glauber dynamics (吉布斯采样算法) 在硬核模型临界点前的快速收敛. 类似的结果此前已知的证明需要用到高等数学工具且都很复杂. 本文给出了一个简单且初等的证明, 为硬核模型上 Glauber dynamics 的收敛现象提供了一种新的理解方式.

目前, 我们的方法暂时无法推广至 Ising 模型等其他吉布斯分布, 因为该证明使用了一个关键性质, 即硬核模型进行 k -变换 (见定义 3) 之后仍然是硬核模型. 对于 Ising 模型, 这一性质并不成立. 如何推广当前的方法, 为一般的吉布斯分布的高效采样的临界行为给出组合证明, 是一个未来值得探索的方向.

References

- [1] Chen XY, Feng WM. Rapid mixing via coupling independence for spin systems with unbounded degree. arXiv:2407.04672, 2024.
- [2] Chen XY, Feng WM, Yin YT, Zhang XY. Rapid mixing of Glauber dynamics via spectral independence for all degrees. In: Proc. of the 62nd IEEE Annual Symp. on Foundations of Computer Science (FOCS). Denver: IEEE, 2022. 137–148. [doi: 10.1109/FOCS52979.2021.00022]
- [3] Chen YS, Eldan R. Localization schemes: A framework for proving mixing bounds for Markov chains (extended abstract). In: Proc. of the 63rd IEEE Annual Symp. on Foundations of Computer Science (FOCS). Denver: IEEE, 2022. 110–122. [doi: 10.1109/FOCS54457.2022.00018]
- [4] Mézard M, Montanari A. Information, Physics, and Computation. Oxford: Oxford University Press, 2009.
- [5] Anari N, Jain V, Koehler F, Pham HT, Vuong TD. Entropic independence: Optimal mixing of down-up random walks. In: Proc. of the

- 54th Annual ACM SIGACT Symp. on Theory of Computing. Rome: ACM, 2022. 1418–1430. [doi: [10.1145/3519935.3520048](https://doi.org/10.1145/3519935.3520048)]
- [6] Anari N, Liu KK, Gharan SO. Spectral independence in high-dimensional expanders and applications to the hardcore model. In: Proc. of the 61st IEEE Annual Symp. on Foundations of Computer Science (FOCS). Durham: IEEE, 2020. 1319–1330. [doi: [10.1109/FOCS46700.2020.00125](https://doi.org/10.1109/FOCS46700.2020.00125)]
- [7] Chen XY, Feng WM, Yin YT, Zhang XY. Optimal mixing for two-state anti-ferromagnetic spin systems. In: Proc. of the 63rd IEEE Annual Symp. on Foundations of Computer Science (FOCS). Denver: IEEE, 2022. 588–599. [doi: [10.1109/focs54457.2022.00062](https://doi.org/10.1109/focs54457.2022.00062)]
- [8] Chen ZC, Liu KK, Vigoda E. Rapid mixing of Glauber dynamics up to uniqueness via contraction. In: Proc. of the 61st IEEE Annual Symp. on Foundations of Computer Science (FOCS). Durham: IEEE, 2020. 1307–1318. [doi: [10.1109/FOCS46700.2020.00124](https://doi.org/10.1109/FOCS46700.2020.00124)]
- [9] Chen ZC, Liu KK, Vigoda E. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion. In: Proc. of the 53rd Annual ACM SIGACT Symp. on Theory of Computing. ACM, 2021. 1537–1550. [doi: [10.1145/3406325.3451035](https://doi.org/10.1145/3406325.3451035)]
- [10] Dyer M, Sinclair A, Vigoda E, Weitz D. Mixing in time and space for lattice spin systems: A combinatorial view. *Random Structures & Algorithms*, 2004, 24(4): 461–479. [doi: [10.1002/rsa.20004](https://doi.org/10.1002/rsa.20004)]
- [11] Efthymiou C, Hayes TP, Štefankovič D, Vigoda E, Yin YT. Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. *SIAM Journal on Computing*, 2019, 48(2): 581–643. [doi: [10.1137/17M1127144](https://doi.org/10.1137/17M1127144)]
- [12] Hayes TP, Vigoda E. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *The Annals of Applied Probability*, 2006, 16(3): 1297–1318. [doi: [10.1214/105051606000000330](https://doi.org/10.1214/105051606000000330)]
- [13] Luby M, Vigoda E. Approximately counting up to four (extended abstract). In: Proc. of the 29th Annual ACM Symp. on Theory of Computing. El Paso: ACM, 1997. 682–687. [doi: [10.1145/258533.258663](https://doi.org/10.1145/258533.258663)]
- [14] Luby M, Vigoda E. Fast convergence of the Glauber dynamics for sampling independent sets. *Random Structures & Algorithms*, 1999, 15(3-4): 229–241. [doi: [10.1002/\(SICI\)1098-2418\(199910/12\)15:3/4<229::AID-RSA3>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1098-2418(199910/12)15:3/4<229::AID-RSA3>3.0.CO;2-X)]
- [15] Sly A. Computational transition at the uniqueness threshold. In: Proc. of the 51st IEEE Annual Symp. on Foundations of Computer Science. Las Vegas: IEEE, 2010. 287–296. [doi: [10.1109/FOCS.2010.34](https://doi.org/10.1109/FOCS.2010.34)]
- [16] Vigoda E. A note on the Glauber dynamics for sampling independent sets. *The Electronic Journal of Combinatorics*, 2001, 8(1): R8. [doi: [10.37236/1552](https://doi.org/10.37236/1552)]
- [17] Weitz D. Counting independent sets up to the tree threshold. In: Proc. of the 38th Annual ACM Symp. on Theory of Computing. Seattle: ACM, 2006. 140–149. [doi: [10.1145/1132516.1132538](https://doi.org/10.1145/1132516.1132538)]
- [18] Greenhill C. The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Computational Complexity*, 2000, 9(1): 52–72. [doi: [10.1007/PL00001601](https://doi.org/10.1007/PL00001601)]
- [19] Valiant LG. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 1979, 8(3): 410–421. [doi: [10.1137/0208032](https://doi.org/10.1137/0208032)]
- [20] Patel V, Regts G. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 2017, 46(6): 1893–1919. [doi: [10.1137/16M1101003](https://doi.org/10.1137/16M1101003)]
- [21] Peters H, Regts G. On a conjecture of Sokal concerning roots of the independence polynomial. *Michigan Mathematical Journal*, 2019, 68(1): 33–55. [doi: [10.1307/mmj/1541667626](https://doi.org/10.1307/mmj/1541667626)]
- [22] Jerrum MR, Valiant LG, Vazirani VV. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 1986, 43: 169–188. [doi: [10.1016/0304-3975\(86\)90174-X](https://doi.org/10.1016/0304-3975(86)90174-X)]
- [23] Alev VL, Lau LC. Improved analysis of higher order random walks and applications. In: Proc. of the 52nd Annual ACM SIGACT Symp. on Theory of Computing. Chicago: ACM, 2020. 1198–1211. [doi: [10.1145/3357713.3384317](https://doi.org/10.1145/3357713.3384317)]
- [24] Anari N, Liu KK, Gharan SO, Vinzant C. Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid. In: Proc. of the 51st Annual ACM SIGACT Symp. on Theory of Computing. Phoenix: ACM, 2019. 1–12. [doi: [10.1145/3313276.3316385](https://doi.org/10.1145/3313276.3316385)]
- [25] Cryan M, Guo H, Mousa G. Modified log-Sobolev inequalities for strongly log-concave distributions. In: Proc. of the 60th IEEE Annual Symp. on Foundations of Computer Science (FOCS). Baltimore: IEEE, 2019. 1358–1370. [doi: [10.1109/FOCS.2019.00083](https://doi.org/10.1109/FOCS.2019.00083)]
- [26] Bubley R, Dyer M. Path coupling: A technique for proving rapid mixing in Markov chains. In: Proc. of the 38th Annual Symp. on Foundations of Computer Science. Miami Beach: IEEE, 1997. 223–231. [doi: [10.1109/sfcs.1997.646111](https://doi.org/10.1109/sfcs.1997.646111)]
- [27] Chen MF. Trilogy of couplings and general formulas for lower bound of spectral gap. In: Accardi L, Heyde CC, eds. *Probability Towards 2000*. New York: Springer, 1998. 123–136. [doi: [10.1007/978-1-4612-2224-8_7](https://doi.org/10.1007/978-1-4612-2224-8_7)]
- [28] Weitz D. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures & Algorithms*, 2005, 27(4): 445–475. [doi: [10.1002/rsa.20073](https://doi.org/10.1002/rsa.20073)]
- [29] Feng WM, Guo H, Yin YT, Zhang CH. Rapid mixing from spectral independence beyond the boolean domain. In: Proc. of the 2021

- ACM-SIAM Symp. on Discrete Algorithms (SODA). Virtual: Society for Industrial and Applied Mathematics, 2021. 1558–1577. [doi: 10.1137/1.9781611976465.95]
- [30] Levin DA, Peres Y. Markov Chains and Mixing Times. 2nd ed., Providence: American Mathematical Society, 2017.
- [31] Dyer M, Greenhill C, Ullrich M. Structure and eigenvalues of heat-bath Markov chains. Linear Algebra and Its Applications, 2014, 454: 57–71. [doi: 10.1016/j.laa.2014.04.018]
- [32] Caputo P, Menz G, Tetali P. Approximate tensorization of entropy at high temperature. Annales de la Faculté des Sciences de Toulouse: Mathématiques, 2015, 24(4): 691–716. [doi: 10.5802/afst.1460]
- [33] Martinelli F, Sinclair A, Weitz D. The Ising model on trees: Boundary conditions and mixing time. In: Proc. of the 44th Annual IEEE Symp. on Foundations of Computer Science, 2003. Cambridge: IEEE, 2003. 628–639. [doi: 10.1109/SFCS.2003.1238235]
- [34] Lawler GF, Sokal AD. Bounds on the L^2 spectrum for Markov chains and Markov processes: A generalization of Cheeger's inequality. Trans. of the American Mathematical Society, 1988, 309(2): 557–580. [doi: 10.2307/2000925]
- [35] Sinclair A, Jerrum M. Approximate counting, uniform generation and rapidly mixing Markov chains. Information and Computation, 1989, 82(1): 93–133. [doi: 10.1016/0890-5401(89)90067-9]
- [36] Joag-Dev K, Proschan F. Negative association of random variables with applications. The Annals of Statistics, 1983, 11(1): 286–295. [doi: 10.1214/AOS/1176346079]
- [37] Guo H, Mousa G. Local-to-global contraction in simplicial complexes. arXiv:2012.14317, 2021.

附录 A. Block dynamics 分析

在本附录中, 我们证明引理 9. 我们将在一般情形下证明以下引理.

引理 A1. 令 $0 < \delta < 1$ 是一个常数. 令 $\mathcal{I} = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$ 是一个 n -顶点的硬核模型实例. 对于任何整数 $\frac{66}{\delta} \leq \ell \leq n$, ℓ -block dynamics P_ℓ 的谱隙至少为:

$$1 - \lambda_2(P_\ell) \geq \left(\frac{\ell}{2n}\right)^{\frac{67}{\delta}}.$$

引用引理 9 是引理 A1 的一个推论. 在引理 9 中, 实例 $I \in \mathcal{F}(\delta)$ 满足 $n \geq \Delta \geq \frac{12000}{\delta}$ 且 $\ell = \left\lfloor \frac{n}{32e} \right\rfloor$. 可以验证 $\ell \geq \frac{66}{\delta}$. 因此, 弛豫时间最多为 $\left(\frac{2n}{\ell}\right)^{67/\delta} \leq 350^{67/\delta}$. 现在我们来证明引理 A1. 首先引入一些定义. 设 μ 为硬核模型实例 $\mathcal{I}$ 的吉布斯分布. 设 Ω 为 μ 的支持. 固定任意的 $\Lambda \subset V$ 和任意的 Λ 上的可行配置 $\sigma_\Lambda \in \Omega_\Lambda$. 对于任意不同的 $u, v \in V \setminus \Lambda$, 定义 u 对 v 的(两两)影响为:

$$\Psi_\mu^{\sigma_\Lambda}(u, v) \triangleq \Pr_{X \sim \mu}[X_v = 1 \mid X_u = 1, X_\Lambda = \sigma_\Lambda] - \Pr_{X \sim \mu}[X_v = 1 \mid X_u = 0, X_\Lambda = \sigma_\Lambda] \quad (\text{A.1})$$

对于任意的 $u \in V \setminus \Lambda$, 令 $\Psi_\mu^{\sigma_\Lambda}(u, u) = 0$. 矩阵 $\Psi_\mu^{\sigma_\Lambda}$ 是文献 [6] 中的影响矩阵. 令 $\zeta_0, \zeta_1, \dots, \zeta_{n-2} \in [0, 1]$ 为一个实数序列. 如果对于所有整数 $0 \leq k \leq n-2$, 对于所有属于 $\binom{V}{k}$ 的子集 Λ , 对于在 Λ 上的任意可行配置 $\sigma_\Lambda \in \Omega_\Lambda$, 都有 $\frac{\lambda_{\max}(\Psi_\mu^{\sigma_\Lambda})}{n-k-1} \leq \zeta_k$, 那么分布 μ 被称为 $(\zeta_0, \zeta_1, \dots, \zeta_{n-2})$ -谱独立^[6].

定理 A1^[8]. 令 $0 < \delta < 1$ 为一个常数. 令 $\mathcal{I} \in \mathcal{F}(\delta)$ 为一个有 n 个顶点的硬核模型实例, 其吉布斯分布为 μ . 分布 μ 是 $(\zeta_0, \zeta_1, \dots, \zeta_{n-2})$ -谱独立的, 其中对于 $0 \leq i \leq n-2$, $\zeta_i = \min \left\{ \left\lceil \frac{32}{\delta} \right\rceil \frac{1}{n-i-1}, 1 \right\}$.

定理 A2^[9]. 令 V 为一个大小为 n 的基底集合. 令 $\zeta_0, \zeta_1, \dots, \zeta_{n-2} \in [0, 1]$ 为一个实数序列. 对于任何在 $\{0, 1\}^V$ 上的分布 μ , 只要该分布是 $(\zeta_0, \zeta_1, \dots, \zeta_{n-2})$ -谱独立的, 那么对于任何 $1 \leq \ell \leq n$, ℓ -block dynamics P_ℓ 在 μ 上的谱隙至少为:

$$1 - \lambda_2(P_\ell) \geq \frac{\sum_{k=n-\ell}^{n-1} \Gamma_k}{\sum_{k=0}^{n-1} \Gamma_k},$$

其中, 对于 $k > 0$, $\Gamma_k \triangleq \prod_{j=0}^{k-1} \frac{1 - \zeta_j}{1 + \zeta_j}$, $\Gamma_0 \triangleq 1$.

证明: 设 $C \triangleq \left\lceil \frac{32}{\delta} \right\rceil$. 固定整数 ℓ , 使得 $\frac{66}{\delta} \leq \ell \leq n$. 我们假设 $n \geq \ell \geq \frac{66}{\delta} \geq 2C$, 否则, 如果不存在这样的 ℓ , 引理显然成立. 通过一些计算, 可以验证对于任意 $1 \leq k \leq n-1$, 总是有 $\Gamma_k = \frac{(n-k-C)(n-k-C+1)\dots(n-k+C-1)}{(n-C)(n+1-C)\dots(n+1+C)} = \frac{\binom{(n-1)-k+C}{2C}}{\binom{(n-1)+C}{2C}}$. 因此, 结合定理 A1 和定理 A2, 有:

$$1 - \lambda_2(P_\ell) \geq \frac{\sum_{k=0}^{n-1} \Gamma_k}{\sum_{k=0}^{n-1} \Gamma_k} \stackrel{(\star)}{=} \prod_{k=-C}^C \frac{\ell+k}{n+k} \geq \left(\frac{\ell-C}{n-C}\right)^{2C+1} \geq \left(\frac{\ell}{2n}\right)^{2C+1} \geq \left(\frac{\ell}{2n}\right)^{2\lceil \frac{n}{\delta} \rceil + 1} \geq \left(\frac{\ell}{2n}\right)^{\frac{66}{\delta}},$$

其中, 我们利用了 $n \geq \ell \geq \frac{66}{\delta} \geq 2C$ 这一事实, 并且标 $(\star)$ 等式可以从方程 $\sum_{j=0}^{N-1} \binom{j}{k} = \binom{N}{k+1}$ 中获得. $\square$

附录 B. 小度数实例的分析

在本节中, 我们证明引理 5.

证明: 令 $I = (G = (V, E), \lambda) \in \mathcal{F}(\delta)$ 为一个硬核模型实例, 其吉布斯分布为 μ . 令 $\Delta \leq \frac{8}{\delta}$ 表示图 G 的最大度数. 不失一般性, 假设 $\frac{1}{2\Delta} \leq \lambda \leq \lambda_c(\Delta)$. 如果 $\lambda < \frac{1}{2\Delta}$, 那么 $\mathcal{I}$ 满足 Dobrushin 条件, 根据命题 1, 引理成立. 如果对于任意 $\Lambda \subseteq V$, 任意合法的 $\sigma \in \{0, 1\}^\Lambda$, 任意 $v \in V \setminus \Lambda$, 都有:

$$\forall c \in \Omega_{V \setminus \Lambda}^\sigma, \mu_v^\sigma(c) \geq b,$$

那么我们就说 $\mathcal{I}$ 满足 b -边缘有界. 注意到 $\lambda \geq \frac{1}{2\Delta}$, 由于硬核模型实例的自规约性, 可以取顶点 v 两跳邻域上的最坏情况, 从而说明 $\mathcal{I}$ 在 $b = \Theta\left(\frac{1}{\Delta}\right)$ 时是 b -边缘有界的. 令 $\theta = \frac{b^2}{12\Delta} = \Theta\left(\frac{1}{\Delta^3}\right) = \Omega(\delta^3)$. 不失一般性, 我们假设 $n \geq \frac{66}{\theta\delta}$. 否则 $n < \frac{66}{\theta\delta} = O\left(\frac{\Delta^3}{\delta}\right) \leq O\left(\frac{1}{\delta^4}\right)$. 在这种情况下, 我们使用命题 2, 在 $\mathcal{I}$ 上 Glauber 动力学的谱间隙至少为:

$$(5n)^{-\frac{66}{\delta}} = \frac{1}{C'_0(\delta)n}, \quad C'_0 = \left(\frac{1}{\delta}\right)^{O(1/\delta)}.$$

考虑 μ 上的 ℓ -block dynamics, 其中 $\ell = \lceil \theta n \rceil$. 因为 $\theta < 1$, $n \geq \ell \geq \theta n \geq \frac{66}{\delta}$. 根据引理 A1, μ 上 ℓ -block dynamics 的谱隙至少为 $\left(\frac{\ell}{2n}\right)^{67/\delta} \geq \left(\frac{\theta}{2}\right)^{67/\delta}$. 这意味着 Gibbs 分布 μ 满足方差的均匀块分解 (uniform block factorization)^[9], 其中常数 $C_{UB} = \frac{\ell}{n} \left(\frac{2}{\theta}\right)^{67/\delta} = (1/\delta)^{O(1/\delta)}$. 结合文献 [9] 中的引理和事实, 可以证明 Glauber dynamics 的谱隙至少为:

$$\frac{b^4}{18 \log(1/b) C_{UB} n} = \frac{1}{C''_0(\delta)n}, \quad \text{其中 } C''_0 = \left(\frac{1}{\delta}\right)^{O(1/\delta)}.$$

这样就证明了引理 5. $\square$

作者简介

陈小羽, 博士, 主要研究领域为理论计算机, 随机算法, 近似计数与采样.

凤维明, 博士, 助理教授, CCF 专业会员, 主要研究领域为理论计算机, 随机算法, 近似计数与采样.

尹一通, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为随机算法, 数据结构, 并行算法.

张昕渊, 博士生, 主要研究领域为理论计算机, 随机算法, 近似计数与采样.