已提交至《应用概率年刊》

arXiv:2304.02109v2 [stat.CO] 9 Jul 2024

吉布斯采样器的协同性：谱隙

作者：IWONA CHLEBICKA [1,a] 和 KRZYSZTOF ŁATUSZY´NSKI [2,c] 和 BŁA´ZEJ MIASOJEDOW [1,b]

[1]华沙大学应用数学与力学研究所，Banacha 2, 02-097 华沙，波兰a.chlebicka@mimuw.edu.pl; b.miasojedow@mimuw.edu.pl

[2] 华威大学统计系，英国考文垂 CV4 7 ALc.K.G.Latuszynski@warwick.ac.uk

吉布斯采样器是计算物理和统计计算中广泛使用的马尔可夫链蒙特卡罗算法。然而，其最基本的性质，例如不同版本之间收敛特征的关系，尚未得到充分理解。

本文证明了吉布斯采样器谱隙的一致性原则：如果随机扫描或 $d!$ 确定性扫描中的任意一种具有谱隙，那么它们就都具有谱隙。我们的方法依赖于将吉布斯采样器几何地解释为交替投影算法，并分析冯·诺依曼–哈尔珀林循环交替投影方法的收敛速度。作为分析的副产品，我们还证明，尽管确定性扫描吉布斯采样器不可逆，但它们与可逆链一样具有许多稳健性，包括在相同条件下满足指数不等式和中心极限定理。

此外，我们还给出了一个定量结果：如果随机扫描吉布斯采样器的谱隙随维数以多项式速率 β 衰减，那么任意确定性扫描的速率都不劣于 $2\beta+2$。

1. 引言。 吉布斯采样器是源自统计力学的马尔可夫链蒙特卡罗（MCMC）算法。在统计力学中，它们以热浴动力学或 Glauber 动力学之名被引入，用于研究离散模型，尤其是其相变（Glauber (1963)）。继 Geman 和 Geman (1984) 以及 Gelfand 和 Smith (1990) 之后，吉布斯采样器也在主流计算科学领域广受欢迎。它们被用于从难以处理的高维概率分布中采样 π，方法是迭代地从 π 的条件分布中采样；在许多建模场景中，尤其是在贝叶斯范式下，这些条件分布通常更易于处理。

吉布斯采样器通常比其他方法更受青睐，因为它们往往易于实现和使用（Smith 和 Roberts (1993年)）。选择扫描策略——即按何种顺序从各自的条件分布中更新 d 维 π 的各个坐标——是建立吉布斯采样器的关键步骤。两种常见选择是随机扫描吉布斯采样器（$\mathrm{P_{RSG}}$）和确定性扫描吉布斯采样器（$\mathrm{P_{DSG}}$）。前者每次更新一个随机选取的坐标，后者则按确定顺序逐一更新各个坐标。乍看之下，选择特定的扫描方式似乎并非难事，而且 π 的任意一种$\mathrm{P_{RSG}}$ 或 $d!$ 种可能$\mathrm{P_{DSG}}$ 版本都会得到一个以 π 为不变分布的有效算法，但关于如何选择的具体指导却很少。事实上，没有任何扫描策略能够普遍最优，正如

MSC2020 主题分类：主要分类 60J22；次要分类 62F15、82B20、47B38。关键词和短语：马尔可夫链蒙特卡罗、吉布斯采样器、Glauber 动力学、谱隙、交替投影、迟缓混合、快速混合、中心极限定理。

	$\mathsf{P}_{\mathrm{RSG}}$		$\mathsf{P}_{\mathrm{DSG}}$
spectral gap in $L^2(\pi)$	+	$\iff$	+
	$\Updownarrow$		$\Downarrow$ $\Uparrow?$
geometric ergodicity	+	$\overset{\Longrightarrow}{?\Longleftarrow}$	+

图1.定理 *1.1*（上部 $\iff$）、推论 *1.2*（（i）为下部$\Longrightarrow$，（ii）位于左侧 $\Uparrow$）中吉布斯样本性质之间的关系，以及可能存在的带问号的等价推论（参见开放问题 *1.3* 关于$\Longleftarrow$）。所有 $\Downarrow$均出自（Kontoyiannis 和 Meyn，2012年，定理 1.3）。

Roberts 和 Rosenthal （2015年）以及 He 等人 （2016年）所示。此外，$\mathsf{P}_{\mathrm{RSG}}$ 的效率以非平凡的方式取决于选择概率向量（Roberts 和 Sahu （1997年）；Chimisov、Latuszynski 和 Roberts （2018年））。然而，关于其基本性质的理论研究仍然不足，难以为理解不同扫描方式之间可能存在的差异提供有用的指导。

马尔可夫转移算子的谱隙性质是马尔可夫链分析的核心，与马尔可夫链蒙特卡罗（MCMC）的收敛速率及其他稳定性特征密切相关（Meyn and Tweedie （2009年）；Kontoyiannis 和 Meyn （2003年）；Kipnis 和 Varadhan （1986年））。为拓展当前对一般状态空间吉布斯采样器的认识，本文通过证明以下二分性，对不同扫描策略下的谱隙性质给出了完备刻画：

定理 1.1. 设$\mathsf{P}_{\mathrm{DSG}}$ 和$\mathsf{P}_{\mathrm{RSG}}$ 是以 π为目标分布的吉布斯采样器。那么，以下情况恰有一种成立：
(i) 所有吉布斯采样器——$\mathsf{P}_{\mathrm{DSG}}$和$\mathsf{P}_{\mathrm{RSG}}$——都没有正的 $L^2(\pi)$ 谱隙；(ii) 所有吉布斯采样器——$\mathsf{P}_{\mathrm{DSG}}$和$\mathsf{P}_{\mathrm{RSG}}$——都有正的 $L^2(\pi)$ 谱隙。

因此，从定性角度看，所有确定性扫描和随机扫描吉布斯采样器的表现都同样好（另见图 1）。我们将这种二分性称为谱隙的一致性原则。此外，由于对于可逆马尔可夫链，谱隙等价于几何遍历性（简称 GE），Meyn and Tweedie (2009)，上述结果意味着：

推论 1.2. 假设任一以 π为目标分布的随机扫描吉布斯采样器具有几何遍历性。那么：
(i) 所有吉布斯采样器，$\mathsf{P}_{\mathrm{DSG}}$和$\mathsf{P}_{\mathrm{RSG}}$，都是几何遍历的；(ii) 所有吉布斯采样器，$\mathsf{P}_{\mathrm{DSG}}$和$\mathsf{P}_{\mathrm{RSG}}$，都有正的 $L^2(\pi)$ 谱隙。

上述推论的证明见第3 节，紧接在定理1.1的证明之后。然而，要为吉布斯采样器建立类似的几何遍历性一致性原则，还需要解决以下开放问题：

开放问题1.3. 几何遍历性是否能由$\mathsf{P}_{\mathrm{DSG}}$ 推出$\mathsf{P}_{\mathrm{RSG}}$的几何遍历性?

等价地，也可以问，$\mathsf{P}_{\mathrm{DSG}}$ 的几何遍历性是否意味着它具有谱隙。当$\mathsf{P}_{\mathrm{DSG}}$ 可逆时，也就是在2维情况下，答案是肯定的。否则，这是一个非平凡问题，将在第5节讨论。

我们在第3 节和定理3.1中给出了证明所需的主要工具。特别是，这意味着尽管$\mathsf{P}_{\mathrm{DSG}}$ 不可逆，但它满足可逆算子所具有的一项重要条件：其谱隙的存在等价于 $\|\mathsf{P}_{\mathrm{DSG}} - \Pi\| < 1$。我们

$$\begin{array}{ccccc} & \mathsf{P}_{\mathrm{RSG}} & & \mathsf{P}_{\mathrm{DSG}} \\ \text{rapid mixing} & + & \overset{\Longrightarrow}{?\Longleftarrow} & + \end{array}$$

图2.推论 *3.6*所揭示的吉布斯样本性质之间的关系，以及开放问题 *3.7* 提出的一种可能蕴含关系（以问号表示）。

将对此进行探讨，以证明确定性扫描吉布斯采样器在与随机扫描和其他可逆链相同的最小条件下，具有许多关键的遍历性质： Miasojedow (2014年)的霍夫丁不等式对$\mathsf{P}_{\mathrm{DSG}}$同样成立，函数仅属于 $L^2(\pi)$时，中心极限定理也同样成立4.1，参见定理；其中还给出了用谱隙表示渐近方差的显式界。

马尔可夫算子谱隙随维数增长而衰减的速率，是理解其计算效率如何随问题规模变化的关键。快速混合和迟缓混合分别对应多项式衰减和指数衰减；针对统计物理（Madras 和 Zheng (2003年)； Bhatnagar 和 Randall (2004年)）以及贝叶斯后验采样（Woodard、Schmidler 和 Huber (2009年a，b）等领域中的各种算法，相关研究已十分深入。对于吉布斯采样器，长期以来存在一个猜想（参见 Diaconis (2013年)；Levin、Peres 和 Wilmer (2009年)）认为，确定性扫描至多比随机扫描慢一个常数因子，也至多快一个对数因子。然而， Roberts 和 Rosenthal (2015年)）和He 等人 (2016年)）分别针对这两部分结论给出了反例。

基于这一思路，我们的推论 3.5 根据算子的几何性质，对$\mathsf{P}_{\mathrm{DSG}}$ 和 $\mathsf{P}_{\mathrm{RSG}}$ 的谱隙给出了定量估计。对于 $\mathsf{P}_{\mathrm{RSG}}$ 的估计是紧的，并且该结果意味着（通过推论 3.6），$\mathsf{P}_{\mathrm{RSG}}$ 的快速混合意味着$\mathsf{P}_{\mathrm{DSG}}$也快速混合。更确切地说，若$\mathsf{P}_{\mathrm{RSG}}$ 的谱隙随维数增长的多项式衰减速率为 β，那么它不会比 $2\beta+2$ 对 $\mathsf{P}_{\mathrm{DSG}}$的速率更差。这一事实由图 2加以说明。我们将相反的蕴含关系——即 $\mathsf{P}_{\mathrm{DSG}}$ 的快速混合意味着 $\mathsf{P}_{\mathrm{RSG}}$也快速混合——列为开放问题3.7。

我们的论文对大量相关研究作出了贡献。由于马尔可夫链蒙特卡罗的收敛速度在理论和应用中都至关重要，吉布斯采样器的交换策略效率已在许多背景下得到研究；然而，由于问题本身的困难，这些研究通常是在某些简化假设下进行的。高斯情形和近似高斯情形下的收敛速率见 Amit（1991年）； Amit 和 Grenander（1991年）； Roberts 和 Sahu（1997年）以及 Chimisov、Latuszynski 和 Roberts（2018年）。超越高斯情形，但针对双分量采样器，Qin 和 Jones（2022年）表明，无论随机扫描中的选择概率如何，确定性扫描在 $L^2(\pi)$ 中都比其随机扫描对应版本收敛得更快，并给出了两种收敛速率之间的精确关系。正如 Roberts 和 Rosenthal（2015年）所示，当 $d>2$ 时，这样的结论并不成立。He 等人（2016年）给出了更多例子，其中还在离散模型和进一步的特定假设下研究了随机扫描与适当修改的确定性扫描之间的关系。关于哪些条件能够保证双分量吉布斯采样器的所有版本同时具有几何遍历性，我们请参阅 Johnson、Jones 和 Neath（2013年）以及 Tan、Jones 和 Hobert（2013年）。Greenwood、McKeague 和 Wefelmeyer（1998年）从渐近方差的角度比较了二维扫描策略，Qin（2024年）进一步扩展了这一研究。在这种情况下，确定性扫描的渐近方差始终更小。随后，这一主题被扩展到更广泛的马尔可夫链蒙特卡罗算法族，见 Andrieu（2016年）；Maire、Douc 和 Olsson（2014年）； Andrieu 等人（2022年）。然而，根据 Roberts 和 Rosenthal（2015年）的例子以及 Greenwood、McKeague 和 Wefelmeyer（1998年）中的讨论，这些渐近方差结论无法推广到 $d=2$。这里需要指出

这并不令人意外：在研究 $L^2(\pi)$ 收敛时，可逆性发挥着主要作用，参见 Kontoyiannis 和 Meyn（2012年）。在双分量情形下，吉布斯采样器 $\mathrm{P}_{\mathrm{DSG}} = \mathrm{P}_1\mathrm{P}_2$ 和 $\mathrm{P}_{\mathrm{RSG}} = w_1\mathrm{P}_1 + w_2\mathrm{P}_2$ 都是可逆的。这是因为对于投影算子 P_1,P_2，我们有 $(\mathrm{P}_1\mathrm{P}_2)^n = (\mathrm{P}_1\mathrm{P}_2\mathrm{P}_1)^n$，其中 $\mathrm{P}_1\mathrm{P}_2\mathrm{P}_1$ 是自伴的（而 $\mathrm{P}_{\mathrm{RSG}}$ 在任意维数下始终是可逆的）。本文所分析的更高维 $\mathrm{P}_{\mathrm{DSG}}$ 已不再可逆；众所周知，转向多分量采样器会带来令人惊讶且违反直觉的现象（Roberts 和 Rosenthal（2015年））。

无需假设目标分布具有（近似）正态性，也无需限制于 $d = 2$，Liu、Wong 和 Kong 研究了吉布斯采样器的协方差结构与谱性质之间的关系（1994年），并证明 $\mathrm{P}_{\mathrm{RSG}}$ 是正的，即其谱包含于 $[0, 1]$。关于定性性质，Roberts 和 Rosenthal(1997年) 的命题 3.2 表明，若 $\mathrm{P}_{\mathrm{DSG}}$ 是一致遍历的，则 $\mathrm{P}_{\mathrm{RSG}}$也如此；另见 Łatuszy´nski、Roberts 和 Rosenthal (2013年) 的命题 4.8，其中给出了更直接的论证。

除一般抽象框架外，研究者还针对特定数据模型考察了吉布斯采样器的几何遍历性，以确立贝叶斯统计推断在各种统计框架中的可靠性，例如 Hobert 和 Geyer（1998年）；Johnson 和 Jones（2010年）；Jones 和 Hobert（2004年）；Marchev 和 Hobert（2004年）；Tan 和 Hobert（2009年）对一般层次模型的研究，以及 Papaspiliopoulos 和 Roberts（2008年）的研究；近年来，Zanella 和 Roberts（2021年）以及 Papaspiliopoulos、Roberts 和 Zanella（2020年）也进行了相关研究，并给出了多层次层级高斯模型的估计。在统计物理领域，自 De Santis（2002年）以来，伊辛模型的混合时间便通过谱理论得到研究。关于自旋玻璃 Glauber 动力学的更多结果，可参阅后续研究：Ben Arous 和 Jagannath（2018年）；Eldan、Koehler 和 Zeitouni（2022年）；Gheissari 和 Jagannath（2019年）；Jagannath（2019年）。需要指出的是，上述研究涵盖了广泛的问题，但一般理论仍有待形成统一的方法。

从正交投影算子的角度描述吉布斯采样器——这是本文采用的方法——始于 Amit（1991）的研究。他基于 Smith、Solmon 和 Wagner（1977年）给出的界，研究了高斯目标分布下吉布斯采样器收敛的一个特定情形。另一方面，Diaconis、Khare 和 Saloff-Coste（2010年）完整刻画了二维吉布斯采样器的 $L^2(\pi)$-收敛速率。为建立这一刻画，Diaconis、Khare 和 Saloff-Coste（2010年）应用了 Aronszajn（1950年）关于交替投影算法的经典结果。我们的方法则更精细地运用了 Badea、Grivaux 和 Müller（2011年）发展的算子理论几何方法，以分析冯·诺依曼–哈尔珀林方法中循环交替投影的收敛速率。根据 Badea、Grivaux 和 Müller（2011年）的命题 4.6，他们的方法得出的估计比 Smith、Solmon 和 Wagner（1977年）或 Deutsch 和 Hundal（1997年）的结果更精确。

本文其余部分安排如下。在第2 节中，我们介绍几何方法的相关符号和背景知识。第3 节讨论吉布斯采样器的谱性质。特别地，我们在此给出定理3.1 及其证明，随后证明定理1.1。第4 节给出其他结果，包括中心极限定理以及关于$\mathrm{P}_{\mathrm{DSG}}$的指数不等式；开放问题1.3 的讨论则见第5节。

2. 预备知识。记号。下文中，我们用 I 表示恒等算子，用 Im 表示算子的像（值域）。我们设 $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$——一个波兰空间

（度量、可分且完备的）空间，以及 $X_n = (X_{n,1}, \ldots\ldots, X_{n,d}) \in \mathcal{X}$。我们将使用简写记号

$$X_{n,-i} := (X_{n,1}, \ldots, X_{n,i-1}, X_{n,i+1}, \ldots, X_{n,d}),$$

类似地 $\mathcal{X}_{-i} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_d$。考虑定义在 $\mathcal{X}$ 上的一个概率测度，其密度为 π，它将作为目标分布。我们将希尔伯特空间 $L^2(\pi)$ 记为函数空间 $f\colon \mathcal{X} \to \mathbb{R}$，其内积为 $\pi f^2 = \int_{\mathcal{X}} f^2(x)\pi(x)\, dx < \infty$，记作 $\langle f, g\rangle = \int_{\mathcal{X}} f(x)g(x)\pi(x)\, dx$。我们用 $\|\cdot\|$ 表示 $L^2(\pi)$ 中的范数。记 Π 为 $L^2(\pi)$ 上的一个线性算子，使得对于任意函数 $f \in L^2(\pi)$，都有

$$\Pi f = \pi(f). \tag{1}$$

以 π 为目标分布的吉布斯采样器。对于 $Z \sim \pi$，给定 $\pi(\cdot|x_{-i})$，我们记 $Z_i \,|\, Z_{-i} = x_{-i}$ 的条件分布。小步 P_i 指将 $X_n = x_n \in \mathcal{X}$ 替换为

$$X_{n+1} = (X_{n,1}, \ldots, X_{n,i-1}, Z_i, X_{n,i+1}, \ldots, X_{n,d}),$$

其中 $Z_i \sim \pi(\cdot|x_{-i})$。那么，确定性扫描吉布斯采样器是一种通过转移算子更新一个坐标的算法

$$\mathsf{P}_{\mathrm{DSG}} = \mathsf{P}_d \mathsf{P}_{d-1} \cdots \mathsf{P}_1 = \prod_{i=1}^{d} \mathsf{P}_i, \tag{2}$$

而随机扫描吉布斯采样器使用转移算子

$$\mathsf{P}_{\mathrm{RSG}} = \sum_{i=1}^{d} w_i \mathsf{P}_i, \tag{3}$$

其中权重 w_i 满足

$$w_i > 0 \text{ for every } i = 1, \ldots d \text{ and } \sum_{i=1}^{d} w_i = 1. \tag{4}$$

(3) 中权重的典型取法是$w_i = \frac{1}{d}$。我们回顾一下，与确定性扫描吉布斯采样器不同，随机扫描吉布斯采样器是可逆的。

谱与马尔可夫链。我们用 $Spec(T)$ 表示有界线性算子 T的谱，即数集 λ，其中 $(T - \lambda I)$ 不可逆，参见例如 Rudin(1987年。

关于马尔可夫链的基础信息，可参见Meyn andTweedie (2009年)。对于任意马尔可夫核 $P(x, dy)$，我们用P 表示其对应的线性算子，该算子作用于从 $L^2(\pi)$出发的函数，因此

$$\mathsf{P}f(x) = \int_{\mathcal{X}} f(y)P(x, dy).$$

本文中，我们始终假设所考虑的马尔可夫链是 π-不可约的。若存在 $C(x) < \infty$和 $\varrho < 1$，使得对于每个 n 以及几乎处处的 $x \in \mathcal{X}$，$\|P^n(x, \cdot) - \pi(\cdot)\|_{TV} \leq C\varrho^n$ 成立，则称具有转移核 $P(x, dy)$ 且平稳分布为 π 的马尔可夫链 $(X_n)_{n\geq 0}$ 几何遍历。

注意，由于$\mathsf{P}\Pi = \Pi\mathsf{P} = \Pi$ 以及 $\Pi^n = \Pi$，我们有以下事实。

引理EMMA 2.1. 如果P 是具有平稳分布 π的马尔可夫链的转移算子，那么对于每个 n，我们有

$$\|\mathsf{P}^n - \Pi\| = \|(\mathsf{P} - \Pi)^n\| \leq \|\mathsf{P} - \Pi\|^n .$$

我们称具有转移算子 $(X_n)_{n\geq 0}$ 和稳态分布 π 的马尔可夫链具有正谱隙$(1-\varrho)$，当且仅当

$$\sup\{|\lambda| : \lambda \in Spec(\mathsf{P} - \Pi)\} = \varrho < 1 ,$$

其中 $Spec$ 表示算子$(\mathsf{P} - \Pi)$ 在 $L^2(\pi)$中的谱。谱半径公式为 $\varrho = \limsup_{n\to\infty} \|(\mathsf{P} - \Pi)^n\|^{\frac{1}{n}}$，见（Rudin，1987，定理 18.9）。因此，由引理2.1，我们有

$$\varrho = \limsup_{n\to\infty} \|\mathsf{P}^n - \Pi\|^{\frac{1}{n}} .$$

反过来，马尔可夫链具有谱隙，当且仅当存在 n_0 使得

$$\|\mathsf{P}^{n_0} - \Pi\| < 1 . \tag{5}$$

此外，若马尔可夫链可逆，则 P 和$\mathsf{P} - \Pi$ 是自伴算子，并且对任意 n，有 $\|\mathsf{P}^n - \Pi\| = \|\mathsf{P} - \Pi\|^n$。因此，在可逆情形下 $\varrho = \|\mathsf{P} - \Pi\|$。

几何。对于 $i = 1, \ldots, d$，我们将由函数 $f \in L^2(\pi)$构成的、关于第 i 个坐标为常数的子空间定义为子空间 M_i。我们记作 $M := \bigcap_{i=1}^d M_i$。注意，M 仅包含常数函数。

根据吉布斯采样器的定义，很容易看出，对于每个 i，一个小步 P_i 是 M_i 上的正交投影。事实上，由于算子 $\mathsf{P}_i(x, \cdot)$ 关于 π 可逆，因此 P_i 是自伴算子，且 $\mathsf{P}_i^2 = \mathsf{P}_i$。此外，$\Pi$ 是 M 上的正交投影。我们将使用以下事实。

引理 2.2（Badea, Grivaux 和 Müller(2011年)，引理 4.2）。设 $f \in L^2(\pi)$。假设$\mathsf{P}_{\mathrm{DSG}}$ 以 为目标分布，由 (2) 给出，且 Π 由 (1)给出。那么对于每个 $j \in [1, d]$，有

$$\|\mathsf{P}_{j-1} \ldots \mathsf{P}_1 f - \mathsf{P}_j \ldots \mathsf{P}_1 f\|^2 \leq \|f - \Pi f\|^2 - \|\mathsf{P}_{\mathrm{DSG}} f - \Pi f\|^2 .$$

为了利用Badea、Grivaux 和 Müller （2011年）的一些结果，我们需要回顾其中引入的一些对象。我们将广义弗里德里希斯角c（M_1，.....，M_d）定义为

$$c(M_1, \ldots, M_d) := \sup \left\{ \frac{\sum_{i\neq j} \langle f_j, f_i \rangle}{(d-1)\sum_{i=1}^d \langle f_i, f_i \rangle} : \ f_i \in M_i \cap M^{\perp}, \ \sum_{i=1}^d \|f_i\| > 0 \right\} . \tag{6}$$

对于两个空间的情形，这一术语由 Friedrichs （1937年）引入。需要指出的是，巴拿赫空间之间的角这一概念至今仍受到关注。有关最新进展，参见 Oppenheim（2018年）。我们还注意到，与 c（M_1，.....，M_d）相关的一个量，即基于求和的相关系数，曾由 Qin 和 Wang （2024年）在吉布斯采样器的谱隙的望远镜性质背景下进行研究。基于求和的相关系数被 Badea、Grivaux 和 Müller （2011年）称为构型常数；他们证明，该常数与广义弗里德里希斯角等价（相差一个缩放和平移），参见（Badea、Grivaux 和 Müller，2011年，命题 3.6）。

请注意，根据范数的定义以及这一事实： M 只包含常值函数，我们有

$$c(M_1,\dots,M_d)=\sup\left\{\frac{\sum_{i\neq j}\langle f_j,f_i\rangle}{(d-1)\sum_{i=1}^d\langle f_i,f_i\rangle}:\ f_i\in M_i,\ \pi f_i=0,\ \sum_{i=1}^d\pi f_i^2>0\right\}.$$

我们还考虑倾斜度 ℓ（M_1，……，M_d），其

$$\ell(M_1,\dots,M_d):=\inf_{f:\ \mathrm{dist}\,(f,M)=1}\max_{1\le i\le d}\mathrm{dist}\,(f,M_i)\,. \tag{7}$$

我们将使用以下辅助事实，该事实可直接由（Badea、Grivaux 和 Müller，2011年，命题 3.7 和 3.9）得出。

命题2.3. 设 $d\ge 2$。假设 M_1，……，M_d是 L^2（π）的闭子空间，且其交集为 $M=\bigcap_{i=1}^d M_d$。进一步假设，对于每个 i，P_i 是 M_i 上的正交投影，而 Π 是 M 上的正交投影。那么

$$\left\|\frac{1}{d}\sum_{i=1}^d\mathsf{P}_i-\Pi\right\|=\frac{d-1}{d}\left[c(M_1,\dots,M_d)+\frac{1}{d-1}\right] \tag{8}$$

and

$$1-\frac{2d}{d-1}\ell(M_1,\dots,M_d)\le c(M_1,\dots,M_d)\le 1-\frac{1}{d-1}\ell^2(M_1,\dots,M_d)\,. \tag{9}$$

3. 吉布斯采样器的谱隙。我们给出存在 $\mathsf{P}_{\mathrm{DSG}}$ 和 $\mathsf{P}_{\mathrm{RSG}}$ 的谱隙的精确充分条件。

定理3.1. 对于波兰空间 $\mathcal{X}=\mathcal{X}_1\times\cdots\times\mathcal{X}_d$ 上的任意平稳分布 π，以下条件等价。

(i) 存在如 (4) 中所示的正权重 $w_1,\dots,w_d$，使得随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}}=\sum_{i=1}^d w_i\mathsf{P}_i$ 具有 $L^2(\pi)$ 谱隙。(ii) 对于满足 (4) 的任意权重 $w_1,\dots,w_d$，随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}}=\sum_{i=1}^d w_i\mathsf{P}_i$ 具有 $L^2(\pi)$ 谱隙。(iii) 存在一个排列 σ，使得确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}}=\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}$ 具有 $L^2(\pi)$ 谱隙。(iv) 对于每个排列 σ，确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}}=\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}$ 都具有 $L^2(\pi)$ 谱隙。(v) 存在一个排列 σ，使得 $\left\|\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}-\Pi\right\|<1$。(vi) 对于每个排列 σ，都有 $\left\|\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}-\Pi\right\|<1$。

条件 (v) 和 (vi) 通常分别强于条件 (iii) 和 (iv)。不过，我们证明反向蕴含也成立。此外，由于对于任意算子 T，$\|T^*T\|=\|T^2\|$ 条件 (v)、(vi) 等价于

$$\left\|\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}\mathsf{P}_{\sigma(d-1)}\cdots\mathsf{P}_{\sigma(2)}\mathsf{P}_{\sigma(1)}-\Pi\right\|<1$$

对于某个排列 σ。

请注意，条件 *(v)* 和 *(vi)* 用于推导马尔可夫链的 Hoeffding不等式（Miasojedow，2014年），参见第 4.2节中的定理 4。

在定理的证明中3.1 我们使用以下引理。

引理 3.2. 对于任意具有转移算子 P 和稳态分布 π 的马尔可夫链，我们有

$$\|\mathsf{P}-\Pi\|\leq 1\,.$$

P证明。记作 $f_0=f-\pi f$。简单计算可得 $\|f_0\|\leq\|f\|$。事实上 ，

$$\|f_0\|=\pi f_0^2=\pi((f-\pi(f))^2)=\pi(f^2)-(\pi(f))^2\leq\pi(f^2)=\|f\|\,.$$

对于任意 $f\in L^2(\pi)$，根据詹森不等式和平稳分布的定义，我们有

$$\|\mathsf{P}f-\Pi f\|=\|\mathsf{P}f_0\|=\int_{\mathcal{X}}(\mathsf{P}f_0(x))^2\pi(x)dx\leq\int_{\mathcal{X}}\mathsf{P}f_0^2(x)\pi(x)dx=\pi(f_0^2)\leq\|f\|\,.$$

□

下一个引理也可以由（Jones、Roberts 和 Rosenthal，2014年，命题 2）推出，但为求简洁，我们给出一个简单的证明。

引理 3.3。如果存在 w_1、……、w_d 满足（4）且使得

$$\left\|\sum_{i=1}^{d}w_i\mathsf{P}_i-\Pi\right\|<1\,,$$

那么对于所有满足 $\tilde{w}_1$、……、$\tilde{w}_d$ 的权重（4），我们有

$$\left\|\sum_{i=1}^{d}\tilde{w}_i\mathsf{P}_i-\Pi\right\|<1\,.$$

ˆ证明。对于任意选取的 $\tilde{w}_1,\ldots,\tilde{w}_d$，令 $\alpha=$最小值$_{1\leq i\leq d}\frac{w_i}{\tilde{w}_i}$。显然 $\alpha\leq 1$。对于 $i=1,\ldots,d$，我们定义 $\hat{}w_i=\frac{\tilde{w}_i-\alpha w_i}{1-\alpha}$。根据定义，$\hat{}w_i\geq 0$ 且 $\sum_{i=1}^{d}\hat{}w_i=1$。令 $\hat{}\mathsf{P}=\sum_{i=1}^{d}\hat{}w_i\mathsf{P}_i$。显然，$\hat{}\mathsf{P}$ 是一个平稳分布为 π 的马尔可夫转移算子。我们有

$$\left\|\sum_{i=1}^{d}\tilde{w}_i\mathsf{P}_i-\Pi\right\|=\left\|\alpha\left[\sum_{i=1}^{d}w_i\mathsf{P}_i-\Pi\right]+(1-\alpha)\left[\hat{\mathsf{P}}-\Pi\right]\right\|\,.$$

应用三角不等式可得

$$\left\|\sum_{i=1}^{d}\tilde{w}_i\mathsf{P}_i-\Pi\right\|\leq\alpha\left\|\sum_{i=1}^{d}w_i\mathsf{P}_i-\Pi\right\|+(1-\alpha)\left\|\hat{\mathsf{P}}-\Pi\right\|\,. \tag{10}$$

最后，根据假设和引理 3.2 可得

$$\left\|\sum_{i=1}^{d}\tilde{w}_i\mathsf{P}_i-\Pi\right\|<\alpha+(1-\alpha)=1\,.$$

□

引理 3.4。设 $d\geq 2$。假设 $M_1,\ldots,M_d$ 是 $L^2(\pi)$ 的闭子空间，且它们的交为 $M=\bigcap_{i=1}^{d}M_d$。进一步假设，对于每个 i，P_i 是 M_i 上的正交投影，而 Π 是 M 上的正交投影。那么对于每个排列 σ，我们有

$$\left\|\mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}-\Pi\right\|<\sqrt{1-\frac{\ell^2}{d^2}}\,, \tag{11}$$

其中 $\ell:=\ell$（M_1，……，M_d）定义见（7）。

证明。我们遵循（Badea、Grivaux 和 Müller，2011年，定理 4.1）的证明思路。根据引理 2.2，对于每个 $j \in [1,d]$ 我们可以估计

$$\begin{aligned}(\operatorname{dist}(f, M_j))^2 = \|f - \mathsf{P}_j f\|^2 &\leq \|f - \mathsf{P}_j \ldots \mathsf{P}_1 f\|^2 \\ &\leq (\|f - \mathsf{P}_1 f\| + \|\mathsf{P}_1 f - \mathsf{P}_2 \mathsf{P}_1 f\| + \cdots + \|\mathsf{P}_{j-1} \ldots \mathsf{P}_1 f - \mathsf{P}_j \ldots \mathsf{P}_1 f\|)^2 \\ &\leq j \left(\|f - \mathsf{P}_1 f\|^2 + \|\mathsf{P}_1 f - \mathsf{P}_2 \mathsf{P}_1 f\|^2 + \cdots + \|\mathsf{P}_{j-1} \ldots \mathsf{P}_1 f - \mathsf{P}_j \ldots \mathsf{P}_1 f\|^2\right) \\ &\leq j^2 \left(\|f - \Pi f\|^2 - \|\mathsf{P}_{\mathrm{DSG}} f - \Pi f\|^2\right) \\ &\leq d^2 \left(\|f - \Pi f\|^2 - \|\mathsf{P}_{\mathrm{DSG}} f - \Pi f\|^2\right).\end{aligned}$$

因此，由于 $\|f - \Pi f\|^2 = \operatorname{dist}(f,\ M)$ 根据 ℓ 的定义，我们看到

$$\ell^2 \|f - \Pi f\|^2 \leq \left(\max_{j \in [1,d]} \operatorname{dist}(f, M_j)\right)^2 \leq d^2 (\|f - \Pi f\|^2 - \|\mathsf{P}_{\mathrm{DSG}} f - \Pi f\|^2).$$

因此，考虑到 $\|f - \Pi f\|^2 \leq \|f\|^2$，我们得到

$$\|\mathsf{P}_{\mathrm{DSG}} f - \Pi f\|^2 \leq \left(1 - \frac{\ell^2}{d^2}\right) \|f - \Pi f\|^2 \leq \left(1 - \frac{\ell^2}{d^2}\right) \|f\|^2.$$

上文中，投影的顺序无关紧要，因此我们推断，对于集合 σ 的$\{1, \ldots, d\}$ 任意排列，均有（11）。 □

我们来证明两类吉布斯采样器存在谱隙的充分条件之间的等价性。

定理 3.1 的证明。由于蕴含关系 $(ii) \Rightarrow (i)$、$(iv) \Rightarrow (iii)$ 和 $(vi) \Rightarrow (v)$ 显然成立，我们将重点放在其余关系上。在此证明中，我们记 $\ell := \ell(M_1, \ldots, M_d)$。

$(i) \Rightarrow (ii)$ 对于任意权重 $w_1, \ldots, w_d$ 随机扫描吉布斯采样器都是可逆的。因此，谱隙性质等价于 $\left\|\sum_{i=1}^d w_i \mathsf{P}_i - \Pi\right\| < 1$ 以及 (ii) 引理 3.3 的结果。

$(i) \Rightarrow (vi)$ 由 (i) 和引理 3.3 可得 $\left\|\frac{1}{d}\sum_{i=1}^d \mathsf{P}_i - \Pi\right\| < 1$。另一方面，命题 2.3 的假设均满足，且我们有（8）。因此$c(M_1, \ldots, M_d) < 1$。根据（9）可得 $\ell > 0$。由引理 3.4 可得

$$\left\|\mathsf{P}_{\sigma(1)} \mathsf{P}_{\sigma(2)} \cdots \mathsf{P}_{\sigma(d)} - \Pi\right\| < \sqrt{1 - \frac{\ell^2}{d^2}} < 1$$

并(vi) 成立。

$(vi) \Rightarrow (iv)$ 我们任取一个固定的排列 σ。于是有

$$\left\|\mathsf{P}_{\sigma(1)} \mathsf{P}_{\sigma(2)} \cdots \mathsf{P}_{\sigma(d)} - \Pi\right\| < 1$$

以及(iv) 由（5）得到。

$(v) \Rightarrow (iii)$ 这是$(vi) \Rightarrow (iv)$的一个特例。

$(iii) \Rightarrow (i)$ 直接由（Roberts 和 Rosenthal，1997年，命题 3.2）得出

□

现在，我们可以据此推导出主要结果。

证明定理 1.1。我们注意到，所需的二分性已在定理3.1的证明中给出。事实上，如果任一吉布斯采样器具有正的 $L^2(\pi)$ 谱隙，那么定理3.1 中的性质 (i) 或 (iii) 成立。于是，根据 (ii) 和 (iv)，所有吉布斯采样器都具有正的 $L^2(\pi)$ 谱隙。否则，没有任何吉布斯采样器具有正的 $L^2(\pi)$ 谱隙。 □

证明 推论 1.2。对于任意 π-不可约马尔可夫链，谱隙性质意味着几何遍历性（Kontoyiannis 和 Meyn，2012年，定理 1.3）。此外，对于可逆链，几何遍历性等价于谱隙性质（Roberts 和 Rosenthal，1997年，定理 2.1）或（Kontoyiannis 和 Meyn，2012年，命题 1.2）。由于随机扫描吉布斯采样器是可逆的，根据定理 3.1 中的等价条件，我们推知所有吉布斯采样器都有正谱隙。 □

利用定理 3.1 的证明，我们可以估计吉布斯采样器的谱隙。需要指出的是，随机扫描吉布斯采样器的估计是精确的。

推论 3.5。 设 $c := c(M_1, \ldots\ldots, M_d)$ 由 (6)给出。谱隙满足以下估计。

(i) 若 $\alpha := d$最小值$_{1\le i\le d} w_i$，对于具有任意权重 $w_1, \ldots, w_d$的随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}} = \sum_{i=1}^{d} w_i \mathsf{P}_i$，满足 (4)，则以下界成立

$$\|\mathsf{P}_{\mathrm{RSG}} - \Pi\| \le \tfrac{d-1}{d}\alpha\left(c + \tfrac{1}{d-1}\right) + 1 - \alpha\,. \tag{12}$$

此外，当对每个 $i = 1, \ldots, d$， $w_i = \frac{1}{d}$ 都成立时，等号成立，并且

$$\tfrac{1}{d} \le \|\mathsf{P}_{\mathrm{RSG}} - \Pi\|\,. \tag{13}$$

(ii) 对于集合 $\{1, \ldots, d\}$ 的任意排列 σ，确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}} = \mathsf{P}^{\sigma(1)}\mathsf{P}^{\sigma(2)}\cdots\mathsf{P}^{\sigma(d)}$ 满足

$$\|\mathsf{P}_{\mathrm{DSG}} - \Pi\| \le \sqrt{1 - \frac{(d-1)^2}{4d^4}(1-c)^2}\,. \tag{14}$$

证明。(i) 我们注意到

$$\left\|\sum_{i=1}^{d} w_i \mathsf{P}_i - \Pi\right\| = \left\|\alpha\left[\sum_{i=1}^{d}\frac{1}{d}\mathsf{P}_i - \Pi\right] + (1-\alpha)\left[\sum_{i=1}^{d}\frac{w_i - \frac{\alpha}{d}}{1-\alpha}\mathsf{P}_i - \Pi\right]\right\|$$
$$\le \alpha\left\|\sum_{i=1}^{d}\frac{1}{d}\mathsf{P}_i - \Pi\right\| + 1 - \alpha\,,$$

这里我们使用了三角不等式和引理 3.2。应用（8）可得所需的不等式。

若$w_i = \frac{1}{d}$对每个 $i = 1, \ldots, d$都成立，则（8）意味着我们处理的是（12）中的等号情形。为说明（13），只需注意，根据定义 $c = c(M_1, \ldots, M_d) \ge 0$。*(ii)* 不等式（14）由（11）结合（9）得到。

□

推论 3.6。 根据快速混合与Woodard、Schmidler 和 Huber（2009年b）的迟缓混合概念，如果随机扫描快速混合，则任何确定性扫描也会快速混合。等价地，如果确定性扫描迟缓混合，则任何随机扫描也会迟缓混合。

证明。如果随机扫描快速混合，其谱隙 $1-\varrho_{\mathrm{RSG}}(d)$ 随维数 d多项式衰减，即 极限 $\lim_{d\to\infty} \frac{1-\varrho_{\mathrm{RSG}}(d)}{d^{-\beta}} > 0$ 对于某个 ~~$\beta>0$~~成立。因此，对于足够大的d，有 $\varrho_{\mathrm{RSG}}(d) \le 1-\gamma d^{-\beta}$ 对于某个 $\gamma>0$成立。我们取这样的 $\gamma>0$。因此，根据（8）可得

$$c(M_1,\dots,M_d) \le 1-\gamma d^{-\beta}.$$

结合 (14) 上述结果可推出

$$\varrho^2_{\mathrm{DSG}}(d) \le 1-\tfrac{(d-1)^2}{4d^2}\gamma^2 d^{-2\beta-2} \le 1-\tfrac{\gamma^2}{16}d^{-2\beta-2}.$$

因此

$$1-\varrho_{\mathrm{DSG}}(d) \ge \tfrac{1}{2}(1-\varrho^2_{\mathrm{DSG}}(d)) \ge \tfrac{\gamma^2}{32}d^{-2\beta-2},$$

这意味着确定性扫描也能快速混合。 □

我们提出一个问题：反向蕴含是否也成立?

开放问题 3.7. 已知 $\varrho_{\mathrm{DSG}}(d)$ 以维数 d的多项式速率衰减，是否可以证明 $\varrho_{\mathrm{RSG}}(d)$ 的衰减是维数 d的多项式函数?

4. 吉布斯采样器的中心极限定理与霍夫丁不等式。作为定理3.1 的直接推论，我们证明以下版本的中心极限定理与霍夫丁不等式对两种吉布斯采样器均成立。

定理4.1. 设马尔可夫链$(X_n)_{n\ge0}$ 具有平稳分布 π，由具有正谱隙的确定性扫描吉布斯采样器$\mathsf{P}_{\mathrm{DSG}}$ 或具有正谱隙的随机扫描吉布斯采样器$\mathsf{P}_{\mathrm{RSG}}$ 生成。进一步假设$(X_n)_{n\ge0}$是哈里斯常返的。那么，对于任意满足 $\pi(f^2)<\infty$ 的函数 f 以及任意初始分布，我们有

$$\frac{1}{\sqrt{n}}\left[\sum_{i=0}^{n-1}(f(X_i)-\pi f)\right] \xrightarrow{D} \mathcal{N}(0,\sigma^2(f)),$$

其中渐近方差 $\sigma^2(f)$ 有界于

$$\sigma^2(f) \le \frac{1+\varrho}{1-\varrho}\pi(f-\pi f)^2,$$

其中$(1-\varrho)$ 是所考虑的吉布斯采样器的 $L^2(\pi)$ 谱隙。
此外，对于满足（4）的任意权重$w_1,\dots,w_d$，随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}} = \sum_{i=1}^d w_i \mathsf{P}_i$ 满足 $\varrho = \|\mathsf{P}_{\mathrm{RSG}} - \Pi\|$；而对于集合 $\{1,\dots,d\}$ 的任意排列 σ，确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}} = \mathsf{P}_{\sigma(1)}\mathsf{P}_{\sigma(2)}\cdots\mathsf{P}_{\sigma(d)}$ 满足$\varrho \le \|\mathsf{P}_{\mathrm{DSG}} - \Pi\|$。在这两种情况下，$\varrho$ 和 $\sigma^2(f)$ 都具有关于维数 d和广义弗里德里希斯角 $c(M_1,\dots,M_d)$的显式界。

证明。令 P 为 $(X_n)_{n\ge0}$的转移算子。我们将采用几何遍历马尔可夫链的一般结果，即（Meyn andTweedie，2009年，定理 17.4.4）。根据定理 3.1，存在 $\varrho<1$ 使得 $\|\mathsf{P}-\Pi\| \le \varrho$ 和 $\|\mathsf{P}^i-\Pi\| \le \varrho^i$。我们定义$\hat{f} = \sum_{i=0}^{\infty}(\mathsf{P}^i f - \Pi f)$。由三角不等式可推得

$$\|\hat{f}\| \le \sum_{i=0}^{\infty}\|\mathsf{P}^i f - \Pi f\| \le \sqrt{\Pi f^2}\sum_{i=0}^{\infty}\|\mathsf{P}^i-\Pi\| \le \sqrt{\Pi f^2}\sum_{i=0}^{\infty}\varrho^i < \infty.$$

ˆ此外， f 满足所谓的泊松方程

$$\hat{f} - \mathsf{P}\hat{f} = f - \Pi f\,.$$

因此，(Meyn andTweedie，2009年，定理 17.4.4) 的假设均已得到验证，因此我们可得

$$\frac{1}{\sqrt{n}}\left[\sum_{i=0}^{n-1}(f(X_i)-\pi f)\right] \xrightarrow{D} \mathcal{N}(0,\sigma^2(f))\,,$$

其中 $\sigma^2(f) = \pi(\hat{f}^2 - (\mathsf{P}\hat{f})^2)$。此外，我们有

$$\sigma^2(f) = \pi(f-\pi f)^2 + 2\sum_{i=1}^{\infty}\langle f-\pi f, (\mathsf{P}-\Pi)(f-\pi f)\rangle\,.$$

根据柯西–施瓦茨不等式，且由于 $\|\mathsf{P}^i - \Pi\| \le \varrho^i$，我们得到

$$\sigma^2(f) \le \pi(f-\pi f)^2\left[1+2\sum_{i=1}^{\infty}\varrho^i\right] = \frac{1+\varrho}{1-\varrho}\pi(f-\pi f)^2\,.$$

对于 ϱ，关于$\mathsf{P}_{\mathrm{RSG}}$ 和$\mathsf{P}_{\mathrm{DSG}}$ 的界分别由 (12) 和 (14) 得出。 □

我们的分析和定理3.1 的另一个推论是，吉布斯采样器的两个版本都满足以下指数不等式。

定理 4.2. 设马尔可夫链 $(X_n)_{n\ge 0}$ 及平稳分布 π 由具有正谱隙的确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}}$，或由具有正谱隙的随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}}$ 生成。设 f 为取值于 $[0,1]$的函数，并记 $\mu := \pi(f)$。则对于满足 $\mu+\varepsilon \le 1$的任意 ε、任意 $n \ge 1$ 以及满足 $\frac{d\nu}{d\pi} \in L^2(\pi)$ 的任意初始测度 $\nu \ll \pi$，均有

$$\mathbb{P}\left(\sum_{i=1}^{n} f(X_i) \ge n(\mu+\varepsilon)\right) \le \left\|\tfrac{d\nu}{d\pi}\right\| \exp\left(-\tfrac{1-\varrho}{1+\varrho}n\varepsilon^2\right)$$

其中 $(1-\varrho)$ L^2是所考虑吉布斯采样器的 π（L^2）-谱隙。
此外，对于权重为 w_1，……，w_d且满足（4）的随机扫描吉布斯采样器 $\mathsf{P}_{\mathrm{RSG}} = \sum_{i=1}^{d} w_i \mathsf{P}_i$，我们有 $\varrho = \|\mathsf{P}_{\mathrm{RSG}} - \Pi\|$；而对于由集合 $\{1,\dots,d\}$ 的任意排列 σ 确定的确定性扫描吉布斯采样器 $\mathsf{P}_{\mathrm{DSG}} = \mathsf{P}^{\sigma(1)}\mathsf{P}^{\sigma(2)}\cdots\mathsf{P}^{\sigma(d)}$，我们有 $\varrho \le \|\mathsf{P}_{\mathrm{DSG}} - \Pi\|$。在这两种情况下，$\varrho$均可根据维数 d和广义弗里德里希斯角 $c(M_1, \dots, M_d)$得到显式界。

证明。结合定理3.1 与（Miasojedow，2014年，定理 1.1），即可得到所需的不等式。对于 ϱ 的$\mathsf{P}_{\mathrm{RSG}}$ 和$\mathsf{P}_{\mathrm{DSG}}$ 的界分别由（12）和（14）得出。 □

5. 关于开放问题**1.3**。本节给出与开放问题1.3 得到肯定答案等价的条件，并举例说明答案可能是否定的。

注记 5.1。为了证明由 $\mathsf{P}_{\mathrm{DSG}}$ 的几何遍历性可推出定理 1.1的假设，只需证明 $\mathsf{P}_{\mathrm{DSG}}$ 的几何遍历性蕴含 $\mathsf{P}_{\mathrm{RSG}}$（或者等价地，定理 3.1中的任一谱隙性质）。根据（$\mathsf{P}_{\mathrm{DSG}}$ 的几何遍历性等价于以下各项断言，参见（Roberts 和 Rosenthal，1997年，命题 2.1）和（Meyn and Tweedie，2009年，定理 15.0.2）。

(*i*)（返回时间的指数矩）存在某个小可测集 C，以及 τ——返回集合 C 的返回时间，并且 $b>1$，使得 $\sup_{x\in C}\mathbb{E}(b^{\tau}|X_0=x)<\infty$。(*ii*)（漂移条件）存在一个可测函数 $V\geq 1$、一个小集 C、$\lambda<1$ 和 $K<\infty$，使得 $\mathrm{P}_{\mathrm{DSG}}V(x)\leq\lambda V(x)+K\mathbb{1}_C(x)$。(*iii*)（$L_V$中的谱隙）对于可测函数 $V\geq 1$ 和赋范数 $\|f\|_{L_V}$ 的空间 $=:$ $\sup_{x\in\mathcal{X}}\frac{|f(x)|}{V(x)}$（DSG $-\Pi$）（ ）算子（P L_V）在 中具有谱隙。

如果存在 $\pi(C)>0$ 以及 n、$\beta>0$ 和概率测度 μ，使得对于每个可测集 A 都有 $P^n_{\mathrm{DSG}}(x,A)\geq\mathbb{1}_C(x)\beta\mu(A)$，则称集合 C 为小集。

让我们通过一个例子来说明，为什么我们怀疑可能无法从 $\mathsf{P}_{\mathrm{DSG}}$ 的几何遍历性推断出其正谱隙的存在。注意，对于转移算子为 P 的马尔可夫链，其谱隙为正，当且仅当其加性可逆化 $\frac{1}{2}(\mathsf{P}+\mathsf{P}^*)$ 具有正谱隙。如果算子 $\frac{1}{2}(\mathsf{P}+\mathsf{P}^*)$ 是几何遍历的，就可能出现这种情况。事实上，我们构造了一个马尔可夫链，使得 P 和 P^* 是几何遍历的，但 $\frac{1}{2}(\mathsf{P}+\mathsf{P}^*)$ 却不是。因此，要对开放问题 1.3给出可能的肯定答案，就必须依赖吉布斯采样器的特殊性质。

例 5.2. 设 $\mathbb{Z}^+=\{0,1,2,\dots\}$。我们将状态空间 $\mathbb{Z}^+\times\mathbb{Z}^+$ 上的马尔可夫链 $\{X_n\}_{n\geq 0}$ 定义为，其转移矩阵 P 如下：

$$\begin{aligned}
P((0,0),(n,n)) &= p(n) \quad \text{for all } n\in\mathbb{Z}^+,\\
P((n,k),(n,k-1)) &= 1 \quad \text{for all } n\in\mathbb{Z}^+,\ k=2,\dots n,\\
P((n,1),(0,0)) &= 1 \quad \text{for all } n\in\mathbb{Z}^+,\\
P((k,l),(i,j)) &= 0 \quad \text{otherwise},
\end{aligned}$$

其中，$p(n)$ 是任意选取的分布，满足 $p(n)>0$ 对所有 n 均成立，且具有有限的指数矩。显然，马尔可夫链 $\{X_n\}_{n\geq 0}$ 是非周期且不可约的。我们定义停时 $\tau=\inf\{n:\ n>0, X_n=(0,0)\}$。根据（Meyn and Tweedie，2009年，定理 15.0.2），要证明几何遍历性，只需证明对某个 $b>1$，$\mathbb{E}(b^{\tau}|X_0=(0,0))<\infty$ 的指数矩存在。但若链从 (0, 0) 开始，则 τ 的分布为 $\mathbb{P}(\tau=n)=p(n-1)$，其中 $n\geq 1$。因此，根据关于 $p(n)$ 的假设，我们可得 τ 的指数矩有限，进而链 $\{X_n\}_{n\geq 0}$ 是几何遍历的。由（Meyn and Tweedie，2009年，定理 10.0.1）可推得，平稳分布π 与第二个变量无关。因此，由方程定义的伴随算子 P^*

$$\pi(k,l)P^*((k,l),(i,j))=\pi(i,j)P((i,j),(k,l))$$

由.....给出

$$\begin{aligned}
P^*((0,0),(n,1)) &= p(n) \quad \text{for all } n\in\mathbb{Z}^+,\\
P^*((n,k-1),(n,k)) &= 1 \quad \text{for all } n\in\mathbb{Z}^+,\ k=2,\dots n,\\
P^*((n,n),(0,0)) &= 1 \quad \text{for all } n\in\mathbb{Z}^+,\\
P^*((k,l),(i,j)) &= 0 \quad \text{otherwise},
\end{aligned}$$

按照 P^* 生成的链，通过类似的论证也具有几何遍历性。不过，我们将证明其加性可逆化 $\mathsf{K}:=\frac{1}{2}(\mathsf{P}+\mathsf{P}^*)$ 不具有几何遍历性。只需证明该链的导通率为 0，因为根据 Cheeger 的

不等式（Lawler 和 Sokal，1988，定理 2.1），我们将得到该链不存在谱隙。因此，根据 K 的可逆性，我们得到 K 不具有几何遍历性。K 的导通率定义为

$$\varkappa := \inf_{A:\pi(A)>0} \frac{\sum_{x\in A}\pi(x)K(x,A^c)}{\pi(A)\pi(A^c)},$$

其中 A^c 是 A 的补集。对于任意 $n>0$，我们定义集合 $A_n=\{(n,k):\ k=1,\dots,n\}$。注意，对于 $k=2,\dots,n-1$，$K((n,k)\ A_n^c)=0$；而对于 $k\in\{1,n\}$，$K((n,k)\ A_n^c)=\frac{1}{2}$。由于 $\pi(A_n^c)>\pi(0,0)>0$，且 $\pi(n,k)=\pi(n,j)=\delta$，因此对于某个 $\delta\geq 0$，我们得到

$$\varkappa \leq \frac{\sum_{x\in A_n}\pi(x)K(x,A_n^c)}{\pi(A_n)\pi(A_n^c)} \leq \frac{\delta}{n\delta\pi(0,0)} = \frac{1}{n\pi(0,0)} \xrightarrow[n\to\infty]{} 0.$$

由于 n 可以任意大，我们推知导通率 $\varkappa=0$，进而链 K 不具有几何遍历性。

6. 致谢。我们感谢匿名审稿人提出非常有用的意见，使本文得到了改进。IC 的研究受到 NCN 资助项目 2019年/34/E/ST1/00120 的支持，BM 的研究受到 NCN 资助项目 2018年/31/B/ST1/00253 的支持。KŁ 获得了英国皇家学会通过英国皇家学会大学研究奖学金提供的支持。

参考文献

AMIT, Y.（1991年）。关于高斯分布和非高斯分布随机松弛的收敛速率。《多元分析杂志》38 82–99。https://doi.org/10.1016/0047-259X(91)90033-X MR1128938 AMIT, Y. 和 GRENANDER, U.（1991年）。比较随机松弛的扫描策略。《多元分析杂志》37 197–222。https://doi.org/10.1016/0047-259X(91)90080-L MR1114473 ANDRIEU, C.（2016年）。关于随机扫描和系统扫描采样器。《生物计量学》103 719–726。MR3551794 ANDRIEU, C.、LEE, A.、POWER, S. 和 WANG, A. Q.（2022年）。利用弱庞加莱不等式比较马尔可夫链，并应用于伪边际马尔可夫链蒙特卡罗。《统计年刊》50 3592–3618。https://doi.org/10.1214/22-aos2241 MR4524509 ARONSZAJN, N.（1950年）。再生核理论。《美国数学会汇刊》68 337–404。MR51437 BADEA, C.、GRIVAUX, S. 和 MÜLLER, V.（2011年）。交替投影方法的收敛速率。《代数与分析》23 1–30。https://doi.org/10.1090/S1061-0022-2012-01202-1 MR2896163 BEN AROUS, G. 和 JAGANNATH, A.（2018年）。平均场自旋玻璃的谱隙估计。《数学物理通讯》361 1–52。https://doi.org/10.1007/s00220-018-3152-6 MR3825934 BHATNAGAR, N. 和 RANDALL, D.（2004年）。Potts 模型模拟回火的迟缓混合。载于《第十五届 ACM-SIAM 离散算法研讨会论文集》478–487。ACM，纽约。MR2291087CHIMISOV, C.、LATUSZYNSKI, K. 和 ROBERTS, G.（2018年）。调整吉布斯采样器。DE SANTIS, E.（2002年）。自旋玻璃在低温和高温下的 Glauber 动力学。《亨利·庞加莱研究所概率与统计年刊》38 681–710。https://doi.org/10.1016/S0246-0203(02)01106-8 MR1931583 DEUTSCH, F. 和 HUNDAL, H.（1997年）。交替投影方法的收敛速率。II。《数学分析与应用杂志》205 381–405。MR1428355 DIACONIS, P.（2013年）。我们学到的一些东西（关于马尔可夫链蒙特卡罗）。《伯努利》19 1294–1305。https://doi.org/10.3150/12-BEJSP09 MR3102552 DIACONIS, P.、KHARE, K. 和 SALOFF-COSTE, L.（2010年）。随机交替投影。《伊利诺伊数学杂志》54 963–979。MR2928343ELDAN, R.、KOEHLER, F. 和 ZEITOUNI, O.（2022年）。谱隙的谱条件：高温伊辛模型中的快速混合。《概率论及相关领域》182 1035–1051。https://doi.org/10.1007/s00440-021-01085-x MR4408509 FRIEDRICHS, K.（1937年）。关于解析函数和二元函数的某些不等式及特征值问题。《美国数学会汇刊》41 321–364。MR1501907GELFAND, A. E. 和 SMITH, A. F. M.（1990年）。基于抽样的边缘密度计算方法。《美国统计学会杂志》85 398–409。MR1141740GEMAN, S. 和 GEMAN, D.（1984年）。随机松弛、吉布斯分布与图像的贝叶斯复原。《IEEE 模式分析与机器智能汇刊》PAMI-6 721-741。

GHEISSARI, R. 和 JAGANNATH, A.（2019年）。球形自旋玻璃动力学的谱隙。《亨利·庞加莱研究所概率年刊》55 756–776。https://doi.org/10.1214/18-aihp897 MR3949952 GLAUBER, R. J.（1963年）。伊辛模型的时变统计。《数学物理杂志》4 294–307。https://doi.org/10.1063/1.1703954 MR148410 GREENWOOD, P. E.、MCKEAGUE, I. W. 和 WEFELMEYER, W.（1998年）。吉布斯采样器的信息界。《统计年刊》26 2128–2156。MR1700224 HE, B. D.、DE SA, C. M.、MITLIAGKAS, I. 和 RÉ, C.（2016年）。吉布斯采样中的扫描顺序：哪些模型会受其影响，以及影响程度的界。载于《神经信息处理系统进展》第29卷。HOBERT, J. P. 和 GEYER, C. J.（1998年）。层次随机效应模型的吉布斯采样器和分块吉布斯采样器的几何遍历性。《多元分析杂志》67 414–430。JAGANNATH, A.（2019 60 083305, 9 10 1063 1 5094173 3993758）。平均场自旋玻璃在短期和长期时间尺度上的动力学。《数学物理杂志》。https://doi.org/ . / . MR JOHNSON, A. A. 和 JONES, G. L.（2010年）。贝叶斯层次广义线性模型的吉布斯采样。《电子统计学杂志》4 313–333。https://doi.org/10.1214/09-EJS515 MR2645487 JOHNSON, A. A.、JONES, G. L. 和 NEATH, R. C.（2013年）。逐分量马尔可夫链蒙特卡罗：混合与复合条件下的一致遍历性和几何遍历性。《统计科学》28 360–375。https://doi.org/10.1214/13-STS423 MR3135537 JONES, G. L. 和 HOBERT, J. P.（2004年）。层次随机效应模型吉布斯采样器的充分预热期。《统计年刊》32 784–817。https://doi.org/10.1214/009053604000000184 MR2060178 JONES, G. L.、ROBERTS, G. O. 和 ROSENTHAL, J. S.（2014年）。条件 Metropolis-Hastings 采样器的收敛。《应用概率进展》46 422–445。https://doi.org/10.1239/aap/1401369701 MR3215540 KIPNIS, C. 和 VARADHAN, S.（1986年）。可逆马尔可夫过程加性泛函的中心极限定理及其在简单排斥过程中的应用。《数学物理通讯》104 1–19。KONTOYIANNIS, I. 和 MEYN, S. P.（2003年）。几何遍历马尔可夫过程的谱理论与极限定理。《应用概率年刊》13 304–362。https://doi.org/10.1214/aoap/1042765670 MR1952001 (2003m:60187) KONTOYIANNIS, I. 和 MEYN, S. P.（2012年）。不可逆马尔可夫链的几何遍历性与谱隙。《概率论及相关领域》154 327–339。https://doi.org/10.1007/s00440-011-0373-4 MR2981426 ŁATUSZYŃSKI, K.、ROBERTS, G. O. 和 ROSENTHAL, J. S.（2013年）。自适应吉布斯采样器及相关马尔可夫链蒙特卡罗方法。《应用概率年刊》23 66–98。https://doi.org/10.1214/11-AAP806 MR3059204 LAWLER, G. F. 和 SOKAL, A. D.（1988年）。马尔可夫链与马尔可夫过程的 L^2 谱界：Cheeger 不等式的推广。《美国数学会汇刊》309 557–580。LEVIN, D. A.、PERES, Y. 和 WILMER, E. L.（2009年）。马尔可夫链与混合时间。美国数学会，普罗维登斯，罗得岛州。含 James G. Propp 和 David B. Wilson 撰写的一章。https://doi.org/10.1090/mbk/058 MR2466937 LIU, J. S.、WONG, W. H. 和 KONG, A.（1994年）。吉布斯采样器的协方差结构及其在估计量和增广方案比较中的应用。《生物计量学》81 27–40。https://doi.org/10.1093/biomet/81.1.27 MR1279653 MADRAS, N. 和 ZHENG, Z.（2003年）。关于交换算法。《随机结构与算法》22 66–97。https://doi.org/10.1002/rsa.10066 MR1943860 MAIRE, F.、DOUC, R. 和 OLSSON, J.（2014年）。非齐次马尔可夫链渐近方差的比较及其在马尔可夫链蒙特卡罗方法中的应用。《统计年刊》42 1483–1510。MR3262458 MARCHEV, D. 和 HOBERT, J. P.（2004年）。van Dyk 和 Meng 多元 Student t 模型算法的几何遍历性。《美国统计学会杂志》99 228–238。MARTINELLI, F.、SINCLAIR, A. 和 WEITZ, D.（2004年）。树上的 Glauber 动力学：边界条件与混合时间。《数学物理通讯》250 301–334。https://doi.org/10.1007/s00220-004-1147-y MR2094519 MEYN, S. P. 和 TWEEDIE, R. L.（2009年）。马尔可夫链与随机稳定性。剑桥大学出版社。MIASOJEDOW, B.（2014年）。一般状态空间上几何遍历马尔可夫链的 Hoeffding 不等式。《统计与概率快报》87 115–120。MR3168944 OPPENHEIM, I.（2018年）。平均投影一致收敛以及投影的循环乘积或随机乘积的角度判据。《以色列数学杂志》223 343–362。MR3773065 PAPASPILIOPOULOS, O. 和 ROBERTS, G.（2008年）。贝叶斯层次模型中吉布斯采样器的稳定性。《统计年刊》36 95–117。https://doi.org/10.1214/009053607000000749 MR2387965 PAPASPILIOPOULOS, O.、ROBERTS, G. O. 和 ZANELLA, G.（2020年）。交叉随机效应模型的可扩展推断。《生物计量学》107 25–40。MR4064138 QIN, Q.（2024年）。利用双投影理论分析双分量吉布斯采样器。QIN, Q. 和 JONES, G. L.（2022年）。双分量马尔可夫链蒙特卡罗采样器的收敛速率。《伯努利》28 859–885。https://doi.org/10.3150/21-bej1369 MR4388922

QIN, Q. and WANG, G. (2024年). 谱望远镜：基于层次结构的随机扫描吉布斯采样器收敛速率界。Ann. Appl. Probab. 34 1319–1349. https://doi.org/10.1214/23-aap1992 MR4700260 ROBERTS, G. O. and ROSENTHAL, J. S. (1997年). 几何遍历性与混合马尔可夫链。Electron. Comm. Probab. 2 no. 2, 13–25. https://doi.org/10.1214/ECP.v2-981 MR1448322 ROBERTS, G. O. and ROSENTHAL, J. S. (2015年). 多种扫描方式下某些简单吉布斯采样器令人惊讶的收敛性质。International Journal of Statistics and Probability 5 51. ROBERTS, G. O. and SAHU, S. K. (1997年). 吉布斯采样器的更新方案、相关结构、分块与参数化。J. Roy. Statist. Soc. Ser. B 59 291–317. MR1440584 RUDIN, W. (1987年). 《实分析与复分析》，第三版。McGraw-Hill Book Co., 纽约。MR924157 SMITH, A. F. M. and ROBERTS, G. O. (1993年). 通过吉布斯采样器及相关马尔可夫链蒙特卡罗方法进行贝叶斯计算。J. Roy. Statist. Soc. Ser. B 55 3–23. MR1210421 SMITH, K. T., SOLMON, D. C. and WAGNER, S. L. (1977年). 从射线影像重建物体问题的实践与数学方面。Bull. Amer. Math. Soc. 83 1227–1270. MR490032 TAN, A. and HOBERT, J. P. (2009年). 具有非正规先验的贝叶斯随机效应模型的分块吉布斯采样：收敛与再生。Journal of Computational and Graphical Statistics 18 861–878. TAN, A., JONES, G. L. and HOBERT, J. P. (2013年). 关于双变量吉布斯采样器的几何遍历性。载于《现代统计理论与应用进展：献给 Morris L. Eaton 的纪念文集》。Inst. Math. Stat. (IMS) Collect. 10 25–42. Inst. Math. Statist., 俄亥俄州比奇伍德。MR3586937 WOODARD, D. B., SCHMIDLER, S. C. and HUBER, M. (2009年a). 多峰分布上并行回火与模拟回火快速混合的条件。Ann. Appl. Probab. 19 617–640. https://doi.org/10.1214/08-AAP555 MR2521882 WOODARD, D. B., SCHMIDLER, S. C. and HUBER, M. (2009年b). 并行回火迟缓混合的充分条件。Electron. J. Probab. 14 no. 29, 780–804. https://doi.org/10.1214/EJP.v14-638 MR2495560 ZANELLA, G. and ROBERTS, G. (2021年). 多层线性模型、吉布斯采样器与多重网格分解（附讨论）。Bayesian Anal. 16 1308–1390. https://doi.org/10.1214/20-BA1242 MR4381136