arXiv:2302.00678v1 [math.NA] 1 Feb 2023

采用 Besov 随机树先验求解贝叶斯椭圆反问题的多层马尔可夫链蒙特卡罗

Andreas Stein *†‡ Viet Ha Hoang §

摘要

我们提出一种马尔可夫链蒙特卡罗有限元算法，用于求解采用“Besov 随机树先验”的椭圆型贝叶斯反问题。这些先验由带有随机系数的小波级数给出，其中展开中的某些项按照所谓 *Galton-Watson trees* 的规律随机消失。这使得模型能够纳入随机分形结构和对数扩散中的大偏差；这些现象在地球物理学或医学成像的许多应用中自然存在。该框架面临两大难点：首先，相关扩散系数不满足一致椭圆性条件，从而产生不可积项，导致标准多层估计量发散。其次，相关参数空间是波兰空间，但不是赋范线性空间。针对第一个问题，我们在估计量中引入截断函数，以补偿不可积项；针对第二个问题，则采用独立性 Metropolis-Hastings 采样器予以解决。所得算法在均方意义下收敛，具有近乎最优的渐近复杂度，且接受概率与维数无关。

1 引言

自然科学和工程学中的无数现象都可以用偏微分方程（PDE）建模。由于信息不完整、测量误差等原因，相应模型中的参数通常具有不确定性。因此，偏微分方程参数常被视为随机变量，或（可能是）无限维随机场。一个经过深入研究的例子是二阶椭圆方程，其中随机扩散系数被用作给定物理区域内不确定渗透率/电导率的统计模型。在许多应用中，人们需要求解相应的逆问题，即根据偏微分方程模型解的离散观测，推断模型参数的实现。重要应用包括地球物理工程中的电阻率层析成像（[10]）以及医学应用中的肌电图（[23]）。无论哪种情况，逆问题都是不适定的，需要采用适当的正则化技术。

一种常见方法是从统计或贝叶斯 视角考察逆问题（[16, 3, 30, 7]），并将其解表示为定义在适当参数空间上的概率测度。这种所谓的后验分布 测度，是通过将先验选定的先验 测度以观测数据为条件推断得到的。在温和假设下，贝叶斯逆问题（BIP）具有适定性；此外，还可以通过为参数空间选择适当的先验 模型来纳入先验模型信息。自 Stuart [30], 的开创性工作以来，

*应用数学研讨班，数学系，ETH 苏黎世¨
†通讯作者。电子邮箱：andreas.stein@sam.math.ethz.ch
‡本研究期间生成和/或分析的数据集，可在提出合理请求后向通讯作者获取。§新加坡南洋理工大学物理与数学科学学院数学科学系

过去十年间，人们对贝叶斯逆问题和逆不确定性量化的兴趣激增，参见例如 [4, 11, 6, 27, 15, 18, 22]。

从计算角度来看，求解逆问题相当于从一个条件概率测度中采样，而该测度仅在归一化常数意义下已知。因此，贝叶斯推断中广泛采用马尔可夫链蒙特卡洛（MCMC）方法，参见例如 [5, 7, 24, 19]。这类接受-拒绝算法依赖于求解相应的（偏微分方程）模型正问题，其中会产生离散化误差（例如由有限元近似导致），且计算成本可能很高。针对这些问题，人们开发了用于贝叶斯逆问题的多层蒙特卡洛算法，相关工作不胜枚举，部分例子包括[14, 26, 20, 12, 9, 21]。与“标准”MCMC方法相比，多层马尔可夫链蒙特卡罗（ML-MCMC）方法能够将计算贝叶斯后验分布下感兴趣量的复杂度降低数个数量级。然而，许多用于椭圆型贝叶斯逆问题的ML-MCMC方法有一个缺点：它们要求随机扩散系数满足一致椭圆性条件。这一要求排除了重要的对数高斯先验，更不用说具有更重尾部的模型，例如Besov先验 [25]。据我们所知，只有 [12, 13] 全面解决了这一问题，分别针对采用高斯先验的椭圆型和抛物型贝叶斯逆问题。

遗憾的是，由于尾部衰减过快，高斯先验模型无法捕捉大偏差。此外，高斯先验或 Besov先验无法在后验模型中纳入分形（空间）结构，而这类结构自然存在于地下流动或医学成像应用中。为此，近期已有研究在 [17] 中针对线性逆问题引入了***Besov*** 随机树先验，并在[28]中将其用作随机椭圆型偏微分方程模型中的对数扩散系数。这些先验由带有随机系数的小波级数给出，其中展开中的某些项会根据所谓***Galton-Watson*** 树的规律随机消失。相应随机场的样本具有分形几何结构，因此，在基于高斯随机场的模型灵活性不足的应用中，Besov 随机树先验或许是一个可行的候选方案。分形结构的程度和豪斯多夫维数由一个调控参数 $\beta \in [0,1]$控制，该参数称为小波密度。

1.1 主要贡献

我们针对采用 Besov 随机树先验的椭圆型贝叶斯逆问题，开发了一种多层马尔可夫链蒙特卡洛有限元采样算法。本文的结果建立在相关配套论文 [28]对采用 Besov 随机树系数的椭圆正问题所作分析的基础上，并对其进行了补充。算法的超参数根据相应正问题的正则性进行调节，我们还对多层马尔可夫链蒙特卡洛算法进行了误差-计算量分析。计算复杂度估计表明，所提出的方法与 [28] 中的正问题多层蒙特卡洛方法具有基本相同的计算复杂度（相差至多对数项），因此在渐近意义下是最优的。特别地，这些结果适用于环面上具有小波密度 $\beta = 1$的“标准”Besov先验。我们强调，正问题模型无需满足一致椭圆性假设，因为我们的多层马尔可夫链蒙特卡洛估计量能够补偿贝叶斯势中的不可积项，且不会引入额外偏差。相反，如果不考虑正问题方程解和贝叶斯势的无界性，就会导致结果极不准确（例如，参见 [12]中的数值实验）。我们还使用独立性 Metropolis-Hastings 采样器，因此该算法可用于一般的（非线性）参数空间，例如 GW 树构成的波兰空间。尽管本文将分析限定于 Besov 随机树先验，但很容易将所提出的算法应用于与非赋范参数空间相关的其他先验模型，且无需一致椭圆性条件。

1.2 本文结构

我们在第1.3节中规定本文的基本符号。第 2 节介绍一般的椭圆型贝叶斯逆问题，并给出后验测度适定性和数据依赖性的相关结果。第3节介绍 Besov 随机树先验，同时为方便读者，回顾 [28] 中关于相应椭圆正问题的适定性和逐路径近似结果

。第4节介绍采用 Besov 随机树先验的贝叶斯逆问题，以及后验测度的维数截断与有限元近似的结合。我们进一步证明后验近似在海林格距离下的先验误差估计。第5节介绍多层马尔可夫链蒙特卡洛算法，证明均方根误差的收敛性，并针对先验模型中全部正则性参数范围，给出相应的误差-计算量分析。第6节通过若干数值实验验证我们的理论结果。

1.3 符号

对于任意向量空间 $\mathcal{V}$，我们用 $\mathcal{V}'$ 表示其拓扑对偶，用 ${}_{\mathcal{V}'}\langle\cdot,\cdot\rangle_{\mathcal{V}}$ 表示相应的对偶配对。对于任意度量空间 $(\mathcal{X}, d_{\mathcal{X}})$，我们用 $B_\lambda(y) := \{x \in \mathcal{X} | \, d_{\mathcal{X}}(x,y) \leq \lambda\} \subset \mathcal{X}$表示以 $y \in \mathcal{X}$ 为中心、半径为 $\lambda > 0$ 的闭球。如果 $d_{\mathcal{X}}$ 由范数 $\|\cdot\|_{\mathcal{X}}$ 通过 $d_{\mathcal{X}}(x,y) = \|x-y\|_{\mathcal{X}}$ 诱导，我们用 $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ 表示相应的赋范空间。对于 $k \in \mathbb{N}$，如果 $\mathcal{X} = \mathbb{R}^k$，则除非另有说明，我们使用欧几里得度量。

任意度量空间 $\mathcal{X}$ 的博雷尔 σ-代数由 $\mathcal{X}$ 中的开集生成，记作 $\mathcal{B}(\mathcal{X})$。对于任意 σ-有限且完备的测度空间 $(E, \mathcal{E}, \mu)$、巴拿赫空间 $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ 以及可积指数 $p \in [1, \infty]$，我们定义勒贝格-博赫纳空间

$$L^p(E, \mu; \mathcal{X}) := \{\varphi : E \to \mathcal{X} | \, \varphi \text{ is strongly measurable and } \|\varphi\|_{L^p(E,\mu;\mathcal{X})} < \infty\},$$

其中

$$\|\varphi\|_{L^p(E,\mu;\mathcal{X})} := \begin{cases} \left(\int_E \|\varphi(x)\|_{\mathcal{X}}^p \mu(dx)\right)^{1/p}, & p \in [1,\infty) \\ \operatorname{ess\,sup}_{x \in E} \|\varphi(x)\|_{\mathcal{X}}, & p = \infty. \end{cases}$$

若 $\mathcal{X} = \mathbb{R}$，我们使用简写记号 $L^p(E, \mu) := L^p(E, \mu; \mathbb{R})$。若 $E \subset \mathbb{R}^d$ 是欧几里得空间的子集，我们假设 $\mathcal{E} = \mathcal{B}(E)$ 且 μ 是勒贝格测度，并记作 $L^p(E) := L^p(E, \mu; \mathbb{R})$，除非另有说明。

对于概率空间 $(\Omega, \mathcal{A}, \mathbb{P})$ 和巴拿赫空间值随机变量 $X : \Omega \to \mathcal{X}$，我们将 $\mathbb{E}_{\mathbb{P}}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ 记为 X 关于 $\mathbb{P}$ 的期望。对于定义在 $(\Omega, \mathcal{A})$ 上且均关于参考测度 $\mathbb{Q}_0$（定义在 $(\Omega, \mathcal{A})$ 上）绝对连续的任意两个测度$\mathbb{Q}_1$，$\mathbb{Q}_2$，$\mathbb{Q}_1$ 和 $\mathbb{Q}_2$ 的海林格距离 定义为

$$d_{\mathrm{Hell}}(\mathbb{Q}_1, \mathbb{Q}_2) := \left(\frac{1}{2}\int_\Omega \left(\sqrt{\frac{d\mathbb{Q}_1}{d\mathbb{Q}_0}(\omega)} - \sqrt{\frac{d\mathbb{Q}_2}{d\mathbb{Q}_0}(\omega)}\right)^2 d\mathbb{Q}_0(\omega)\right)^{1/2}.$$

对于任意有界连通空间区域 $\mathcal{D} \subset \mathbb{R}^d$，对于 $k \in \mathbb{N}$ 和 $p \in [1, \infty]$，我们将具有 $L^p(\mathcal{D})$ 中 k 阶弱导数的标准索伯列夫空间记为 $W^{k,\,p}(\mathcal{D})$。分数阶 $s \geq 0$ 的索伯列夫-斯洛博杰茨基空间记为 $W^{s,\,p}(\mathcal{D})$。此外，对于任意 $s \geq 0$，$H^s(\mathcal{D}) := W^{s,\,2}(\mathcal{D})$，并且我们采用恒等 $H^0(\mathcal{D}) = L^2(\mathcal{D})$。若 $\mathcal{D}$ 是利普希茨区域，我们定义任意 $s > 1/2$

$$H_0^s(\mathcal{D}) := \ker(\gamma_0) = \{\varphi \in H^s(\mathcal{D}) | \, \gamma_0(\varphi) = 0 \text{ on } \partial\mathcal{D}\}, \tag{1}$$

此处，$\gamma_0 \in \mathcal{L}(H^s(\mathcal{D}) \; H^{s-1/2}(\partial\mathcal{D}))$ 表示迹算子。

令 $\mathsf{C}(\overline{\mathcal{D}})$ 表示所有连续函数 $\varphi : \mathcal{D} \to \mathbb{R}$ 构成的空间。对于任意 $\alpha \in \mathbb{N}$，$\mathsf{C}^\alpha(\mathcal{D})$ 是所有具有 α 阶连续偏导数的函数 $\varphi \in \mathsf{C}(\mathcal{D})$ 构成的空间。对于非整数 $\alpha > 0$，我们用 $\mathsf{C}^\alpha(\mathcal{D})$ 表示所有具有 $\alpha - \lfloor\alpha\rfloor$-Hölder 连续的 $\lfloor\alpha\rfloor$ 阶偏导数的 $\varphi \in \mathsf{C}^{\lfloor\alpha\rfloor}(\mathcal{D})$ 构成的空间。对于任意正实数 $\alpha > 0$，我们还用 $\mathcal{C}^\alpha(\mathcal{D})$ 表示光滑度为 α 的 *Hölder-Zygmund* 空间。例如，定义见 [31, 第 1.2 节。2] 我们用 $\mathsf{S}(\mathbb{R}^d)$ 表示施瓦茨空间，即所有光滑且快速衰减的函数构成的空间；用 $\mathsf{S}'(\mathbb{R}^d)$ 表示其对偶空间，即缓增分布空间。此外，对于任意开集 $O \subseteq \mathbb{R}^d$，$\mathsf{D}(O)$ 表示所有支撑紧含于 O 的光滑函数 $\varphi \in \mathsf{C}^\infty(O)$ 构成的空间。

对于有限元误差分析，我们引入一个可数集合 $\mathfrak{H} \subset (0, \infty)$，并用$h \in \mathfrak{H}$ 表示一般的有限元细化参数。我们进一步假设存在一个严格递减序列 $(h_\ell, \ell \in \mathbb{N}) \subset \mathfrak{H}$，使得 $\lim_{\ell\to\infty} h_\ell = 0$。

2 贝叶斯椭圆反问题

2.1 正向偏微分方程模型

设 $(\Omega, \mathcal{A}, \mathbb{P})$ 是一个完备的概率空间（参数为 $\omega \in \Omega$），且 $\mathcal{D} \subset \mathbb{R}^d$, $d \in \{1,2,3\}$是一个凸多边形区域，其边界 $\partial\mathcal{D}$ 由有限条线段或平面片段构成。对于给定的 $\omega \in \Omega$，我们考虑随机（或"参数化"）椭圆问题，求 $u(\omega) : \mathcal{D} \to \mathbb{R}$，使得

$$-\nabla \cdot (a(\omega)\nabla u(\omega)) = f \quad \text{in } \mathcal{D}, \quad u(\omega) = 0 \quad \text{on } \partial\mathcal{D}. \tag{2}$$

问题 (2) 中的扩散系数 $a : \Omega \to L^\infty(\mathcal{D})$ 是一个适当的随机场；为简单起见，假设源项 $f : \mathcal{D} \to \mathbb{R}$ 是一个确定性函数。对于问题 (2) 的变分形式，我们定义 $H := L^2(\mathcal{D})$、$V := H_0^1(\mathcal{D})$，并回顾一下，根据庞加莱不等式，$\|\cdot\|_V : V \to \mathbb{R}_{\geq 0}$, $v \mapsto \|\nabla v\|_H$ 在 V 上定义了一个范数。对于固定的 $\omega \in \Omega$，如果对任意 $v \in V$ 均有下式成立，则称 $u(\omega) \in V$ 是问题 (2) 的逐路径弱解。

$$\int_{\mathcal{D}} a(\omega)\nabla u(\omega) \cdot \nabla v dx = {}_{V'}\langle f, v\rangle_V. \tag{3}$$

为保证逐路径弱解的存在性与唯一性，我们假设 $f \in V'$，并且 $a : \Omega \to L^\infty(\mathcal{D})$ 是强 $\mathcal{A}/\mathcal{B}(L^\infty(\mathcal{D}))$-可测的，使得

$$a_-(\omega) := \operatorname*{ess\,inf}_{x \in \mathcal{D}} a(x, \omega) > 0, \quad P\text{-a.s.} \tag{4}$$

这是一个标准结果（例如参见 [28, 定理 3.2]），用于说明参数到解的映射 $u : \Omega \to V$ 是良定义的，并且是（强）$\mathcal{A}/\mathcal{B}(V)$-可测的。

2.2 贝叶斯逆问题

为了介绍逆问题，我们考虑参数到观测的映射

$$\mathcal{G} : \Omega \to \mathbb{R}^k, \quad \omega \mapsto [\mathcal{O} \circ u](\omega) \tag{5}$$

其线性观测泛函 $\mathcal{O} \in ((H_0^\theta \mathcal{O}(\mathcal{D}))')^k$ 在 $\theta_{\mathcal{O}} \in (1/2, 1]$ 和 $k \in \mathbb{N}$上有界。我们假设含噪观测 δ 具有如下形式

$$\delta = \mathcal{G}(\omega) + \vartheta, \tag{6}$$

其中 ϑ 是定义在 $\mathbb{R}^k$上的中心高斯观测噪声。因此，ϑ 服从如下密度

$$\rho(x) = (2\pi \det(\Sigma))^{-\frac{k}{2}} \exp\left(-\frac{1}{2} x^\top \Sigma^{-1} x\right), \quad x \in \mathbb{R}^k, \tag{7}$$

其中协方差矩阵 $\Sigma \in \mathbb{R}^{k \times k}$对称且正定。

给定观测 δ，我们的目标是推导贝叶斯后验概率测度 $\mathbb{P}_\delta := \mathbb{P}(\cdot \mid \delta)$，其定义在 $(\Omega, \mathcal{A})$ 上。注意，对于给定的 $\omega \in \Omega$，$\delta = \mathcal{G}(\omega) + \vartheta$（在给定实现 $\mathcal{G}(\omega)$ 的条件下）的分布几乎处处相对于 $\mathcal{N}(0, \Sigma)$ 绝对连续，因此 $\mathbb{P}_\delta$ 由贝叶斯定理给出：

命题 2.1。后验测度 $\mathbb{P}_\delta$ 关于先验测度 $\mathbb{P}$绝对连续，其拉东–尼科迪姆导数为

$$\frac{d\mathbb{P}_\delta}{d\mathbb{P}}(\omega) = \frac{\rho(\delta - \mathcal{G}(\omega))}{\int_\Omega \rho(\delta - \mathcal{G}(\omega)) d\mathbb{P}(\omega)} = \frac{\exp(-\Phi(\omega; \delta))}{Z(\delta)}. \tag{8}$$

在(8)中，我们定义了贝叶斯势 $\Phi : \Omega \times \mathbb{R}^k \to \mathbb{R}$ 为

$$\Phi(\omega; \delta) := -\log(\rho(\delta - \mathcal{G}(\omega))) = \frac{k \log(2\pi \det(\Sigma))}{2} + \frac{1}{2}(\delta - \mathcal{G}(\omega))^\top \Sigma^{-1}(\delta - \mathcal{G}(\omega)), \tag{9}$$

归一化常数 $Z(\delta) > 0$ 由下式给出

$$Z(\delta) := \int_{\Omega} \exp\left(-\Phi(\omega;\delta)\right) d\mathbb{P}(\omega). \tag{10}$$

证明。式（5）中的观测映射 $\mathcal{G} : \Omega \to \mathbb{R}^k$ 继承了 $u : \Omega \to V$ 的可测性，因此由 [7,定理 14] 可得结论。 □

我们对贝叶斯势 Φ 作出一些假设，以推导映射 $\delta \to \mathbb{P}_\delta$ 关于海林格距离的 Lipschitz 连续性。

假设 2.2。

1) 对每个 $\lambda > 0$，都存在一个常数 $\kappa_1(\lambda) > 0$ 以及一个集合 $\Omega_\lambda \in \mathcal{A}$，使得 $\mathbb{P}(\Omega_\lambda) > 0$，并且

$$\Phi(\omega;\delta) \le \kappa_1(\lambda), \quad \text{for all } \omega \in \Omega_\lambda \text{ and } \delta \in B_\lambda(0).$$

2) 对每个 $\lambda > 0$，都存在 $\kappa_2(\lambda, \cdot) \in L^2(\Omega\ \mathbb{P})$，使得

$$|\Phi(\omega;\delta) - \Phi(\omega;\delta')| \le \kappa_2(\lambda,\omega)\|\delta - \delta'\|_2, \quad \text{for all } \omega \in \Omega \text{ and } \delta, \delta' \in B_\lambda(0).$$

命题 2.3。 *[11, 定理 2.4]* 在假设 *2.2* 下，对任意 $\lambda > 0$，都存在一个常数$C(\lambda) > 0$，使得

$$d_{\mathrm{Hell}}(\mathbb{P}_\delta, \mathbb{P}_{\delta'}) \le C(\lambda)\|\delta - \delta'\|_2, \quad \textit{for all } \delta, \delta' \in B_\lambda(0). \tag{11}$$

[11, 定理 2.4]的证明还给出了式 (10) 中归一化常数 $Z(\delta)$ 的一个下界，该下界只依赖于数据 δ 的范数:

推论 2.4。 在假设 *2.2* 下，对于任意 $\lambda > 0$，存在一个常数 $c(\lambda) > 0$，使得

$$Z(\delta) \ge c(\lambda) > 0 \quad \textit{for all } \delta \in B_\lambda(0). \tag{12}$$

证明。 固定 $\lambda > 0$，对于任意 $\delta \in B_\lambda(0)$，由假设 2.2 的条目 (1) 可得

$$Z(\delta) = \int_{\Omega} \exp\left(-\Phi(\omega;\delta)\right) d\mathbb{P}(\omega) \ge \int_{\Omega_\lambda} d\mathbb{P}(\omega) \exp\left(-\kappa_1(\lambda)\right) =: c(\lambda) > 0.$$

□

我们考虑本文 [17, 28] 中的 *Besov* 随机树先验。这一特定先验已在 [28] 中用于对椭圆正问题 (2) 中的对数扩散系数 $b := \log(a)$ 进行建模。下一节将回顾 Besov 随机树先验的构造，并汇总关于式 (3) 的适定性以及逐路径弱解正则性的一些结果。

3 Besov 随机树先验

本节将介绍 Besov 随机树先验，以及相应的椭圆正问题，以求得 (2) 中的 u，其中 $\log(a)$ 由 Besov 随机树先验给出。首先回顾多分辨率分析（MRA）的一些工具，以及 Besov 空间的小波表示。随后构造 Besov 随机树先验，并总结相应椭圆正问题适定性和正则性的若干结果。本节最后一部分将讨论通过先验的维数截断和有限元实现的逐路径近似。后者对于从第 2 节中的后验测度 $\mathbb{P}_\delta$ 进行（近似）采样是必要的。

3.1 Besov 空间的多分辨率分析与小波表示

令 $\mathbb{T}^d := [0,1]^d$ 表示 $d \in \mathbb{N}$ 的 d 维环面。我们简要回顾在 $L^2(\mathbb{R}^d)$ 和 $L^2(\mathbb{T}^d)$ 上构造正交小波基的方法，以及相关Besov空间的小波表示。有关更详细的介绍，请参阅 [32, 第 1], [33, 章第 1 节。2], 以及关于多分辨率分析中正交小波的 [8,第 5] 章。

设 ϕ 和 ψ 是紧支撑的尺度函数和小波函数，属于 $C^\alpha(\mathbb{R})$，$\alpha \geq 1$，且适用于 $L^2(\mathbb{R})$上的多分辨率分析。此外，我们假设 ψ 满足消失矩条件

$$\int_{\mathbb{R}} \psi(x) x^m dx = 0, \quad m \in \mathbb{N}_0,\ m < \alpha. \tag{13}$$

一个例子是具有 $M := \lfloor \alpha \rfloor \in \mathbb{N}$ 个消失矩的 Daubechies小波（也称为 DB($\lfloor \alpha \rfloor$)-小波），其支撑为 $[-M+1, M]$，且属于 $C^1(\mathbb{R})$，其中 $M \geq 5$（参见例如 [8,第 7.1] 节）。对于任意 $j \in \mathbb{N}_0$ 和 $k \in \mathbb{Z}$，定义缩放和平移函数

$$\psi_{j,k,0}(x) := \phi(2^j x - k), \quad \text{and} \quad \psi_{j,k,1}(x) := \psi(2^j x - k), \quad x \in \mathbb{R}. \tag{14}$$

由于 $\|\phi\|_{L^2(\mathbb{R})} = \|\psi\|_{L^2(\mathbb{R})} = 1$，可知 $((\psi_{0,k,0})\ k \in \mathbb{Z}) \cup ((2^{j/2}\psi_{j,k,1}), (j,k) \in \mathbb{N}_0 \times \mathbb{Z})$ 是 $L^2(\mathbb{R})$ 的一组标准正交基。

相应的各向同性[1] 小波基可通过单变量多分辨率分析的张量化构造，并且在 $L^2(\mathbb{R}^d)$ 中是正交的，$d \geq 2$。我们定义指标集 $\mathcal{L}_0 := \{0,1\}^d$ 和 $\mathcal{L}_j := \mathcal{L}_0 \backslash \{(0,\ldots,0)\}$，其中 $j \in \mathbb{N}$。注意，若 $j = 0$，则 $\mathcal{L}_j$ 的基数为 $|\mathcal{L}_j| = 2^d$；否则为 $|\mathcal{L}_j| = 2^d - 1$。对于任意 $l \in \mathcal{L}_0$，我们进一步定义

$$\psi_{j,k,l}(x) := 2^{dj/2} \prod_{i=1}^{d} \psi_{j,k_i,l(i)}(x_i), \quad j \in \mathbb{N}_0,\ k \in \mathbb{Z}^d,\ x \in \mathbb{R}^d, \tag{15}$$

从而得到 $((\psi_{j,\ k,\ l})\ \ j \in \mathbb{N}_0, \quad k \in \mathbb{Z}^d, \quad l \in \mathcal{L}_j)$ 是 $L^2(\mathbb{R}^d)$ 的正交基。

环面上由局部支撑的周期函数组成的正交基 $\mathbb{T}^d$ 可通过张量化引入，例如 [32, 第 1 节3]中的构造。我们采用 [28, 第 2 节1]中的构造：给定 ϕ 和 ψ，我们选取缩放因子 $w \in \mathbb{N}$，使得

$$\operatorname{supp}(\psi_{w,0,l}) \subset \left\{ x \in \mathbb{R}^d \middle|\ \|x\|_2 < \frac{1}{2} \right\}, \quad l \in \mathcal{L}_0.$$

采用这种 w 选择，可得 $j \in \mathbb{N}_0$ th at

$$\operatorname{supp}(\psi_{j+w,0,l}) \subset \left\{ x \in \mathbb{R}^d \middle|\ \|x\|_2 < 2^{-j-1} \right\}.$$

现在令 $K_j := \{k \in \mathbb{Z}^d | \ 0 \leq k_1, \ldots, k_d < 2^j\} \subset 2^j \mathbb{T}^d$，并注意到 $|K_{j+w}| = 2^{d(j+w)}$。定义一周期小波函数

$$\psi^{per}_{j,k,l}(x) := \sum_{n \in \mathbb{Z}^d} \psi_{j,k,l}(x-n), \quad j \in \mathbb{N}_0,\ k \in K_j,\ l \in \mathcal{L}_0,\ x \in \mathbb{R}^d,$$

及其在 $\mathbb{T}^d$ 上的限制如下：

$$\psi^l_{j,k}(x) := \psi^{per}_{j,k,l}(x), \quad j \in \mathbb{N}_0,\ k \in K_j,\ l \in \mathcal{L}_0,\ x \in \mathbb{T}^d. \tag{16}$$

现在，对于指标集 $\mathcal{I}_w := \{j \in \mathbb{N}_0,\ k \in K_{j+w},\ l \in \mathcal{L}_j\}$，我们得到

$$\mathbf{\Psi}_w := \left((\psi^l_{j+w,k}),\ (j,k,l) \in \mathcal{I}_w \right) \tag{17}$$

[1]可以考虑各向异性张量化，截断后会得到所谓的“双曲交叉近似”。由于这类构造往往会在近似中沿笛卡尔坐标轴引入优先方向，因此本文不予考虑。

是一个 $L^2(\mathbb{T}^d)$-正交归一基，参见 [32, 命题 1.34]。我们进一步定义子空间 $V_{w+1} := \text{span}\{\psi^l_{w,k} |\ k \in K_w,\ l \in \mathcal{L}_0\} \subset L^2(\mathbb{T}^d)$，并注意到 $\dim(V_{w+1}) = 2^{d(w+1)}$。根据 [8, 第 9 章.3]，中针对单周期单变量函数的多重分辨率分析可知，$((\psi^l_{j,\ k})\ \ j \le w,\ \ k \in K_j,\ \ l \in \mathcal{L}_j)$ 是 V_{w+1} 的另一个正交归一基。因此，我们可以替换式 (16) 中前 $2^{d(w+1)}$ 个基函数，从而得到（在计算上更方便的）$L^2(\mathbb{T}^d)$-正交归一基

$$\mathbf{\Psi} := \left((\psi^l_{j,k}),\ (j,k,l) \in \mathcal{I}_{\mathbf{\Psi}}\right), \quad \mathcal{I}_{\mathbf{\Psi}} := \{j \in \mathbb{N}_0,\ k \in K_j,\ l \in \mathcal{L}_j\}. \tag{18}$$

定义 3.1. 设 $s > 0, p \in [1,\infty]$ 和 $\varphi \in L^2(\mathbb{T}^d)$。我们定义*Besov* 范数

$$\|\varphi\|_{B^s_{p,p}(\mathbb{T}^d)} := \left(\sum_{(j,k,l)\in\mathcal{I}_{\mathbf{\Psi}}} 2^{jp(s+\frac{d}{2}-\frac{d}{p})} |(\varphi,\psi^l_{j,k})_{L^2(\mathbb{T}^d)}|^p\right)^{1/p}, \quad p \in [1,\infty), \tag{19}$$

and

$$\|\varphi\|_{B^s_{\infty,\infty}(\mathbb{T}^d)} := \sup_{(j,k,l)\in\mathcal{I}_{\mathbf{\Psi}}} 2^{j(s+\frac{d}{2})} |(\varphi,\psi^l_{j,k})_{L^2(\mathbb{T}^d)}| < \infty. \tag{20}$$

相应的*Besov* 空间 定义在 $\mathbb{T}^d$ 上，具体为

$$B^s_{p,p}(\mathbb{T}^d) := \{\varphi \in L^2(\mathbb{T}^d) |\ \|\varphi\|_{B^s_{p,p}(\mathbb{T}^d)} < \infty\}. \tag{21}$$

我们先介绍一些记号，供本文后续部分使用，涉及 Besov 空间、H¨older 空间和 Zygmund 空间。由于后续分析中（周期）区域 $\mathbb{T}^d$ 保持不变，为方便起见，以下采用简写 $B^s_p := B^s_{p,\ p}(\mathbb{T}^d)$、$\mathsf{C}^\alpha := \mathsf{C}^\alpha(\mathbb{T}^d)$ 和 $\mathcal{C}^\alpha := \mathcal{C}^\alpha(\mathbb{T}^d)$。

3.2 Besov 随机树先验

我们首先将 Besov 随机树先验引入为相对于 $\mathbf{\Psi}$的小波展开，其中将 $L^2(\mathbb{T}^d)$-正交投影系数替换为 p- 指数随机变量。为此，设 $p \in [1,\infty)$，并考虑一个独立同分布（ i.i.d. ）序列 $X = ((X^l_{j,\ k}),\ (j,\ k,\ l) \in \mathcal{I}_{\mathbf{\Psi}})$，其元素为 p-指数随机变量。也就是说，每个 $X^l_{j,\ k}$ 的密度为

$$\phi_p(x) := \frac{1}{c_p}\exp\left(-\frac{|x|^p}{\kappa}\right), \quad x \in \mathbb{R}, \qquad c_p := \int_{\mathbb{R}} \exp\left(-\frac{|x|^p}{\kappa}\right) dx, \tag{22}$$

其中，$\kappa > 0$ 是一个固定的尺度参数。令 $\mathbb{Q}_0$ 表示定义在 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上的相应一维 p-指数测度。当 $p = 2$时，我们得到方差为 $\frac{\kappa}{2}$ 的正态分布；当 $p = 1$时，则得到尺度为 κ 的拉普拉斯分布。

我们先验构造中的随机树结构基于某些集合值随机变量，即所谓的 *Galton-Watson* 树。离散树、Galton-Watson（GW）树的定义及其基本性质，见 [28], 中附录 A；该文讨论了椭圆正问题。

定义 3.2。 [17, 定义 3] 令 $s > 0, p \in [1,\infty)$，以及 $X = ((X^l_{j,k}),\ (j,k,l) \in \mathcal{I}_{\mathbf{\Psi}})$ 为一列 p-指数分布的随机变量。令 $\mathfrak{T}$ 表示所有无无限节点的树的集合（参见 [28, 定义 A.1], ），并令 $T: \Omega \to \mathfrak{T}$ 为一棵 GW 树（参见 [28, 定义 A.3]），其后代分布为 $\mathcal{P} = \text{Bin}(2^d, \beta)$，其中 $\beta \in [0,1]$，且独立于 X。此外，令$\mathfrak{I}_T$ 表示与 T 相关联的小波指标集合，该集合来自 [28, 公式（79）]。

定义随机树索引集 $\mathcal{I}_T(\omega) := \{(j,k,l) |\ (j,k) \in \mathfrak{I}_T(\omega)\ \ l \in \mathcal{L}_j\}$ 和

$$b_T(\omega) := \sum_{(j,k,l)\in\mathcal{I}_T(\omega)} \eta_j X^l_{j,k}(\omega)\psi^l_{j,k}, \quad \omega \in \Omega, \quad \text{where} \quad \eta_j := 2^{-j(s+\frac{d}{2}-\frac{d}{p})}, \quad j \in \mathbb{N}_0. \tag{23}$$

我们将 b_T 称为具有小波密度 的 B^s_p-随机变量。

注 3.3。我们立即得到经典的*Besov*先验，它在 [25] 中引入，是 $\beta = 1$的特例，其中 $\mathcal{I}_T(\omega) = \mathcal{I}_{\mathbf{\Psi}}$ 几乎处处成立。级数 (23) 可自然地解释为随机函数关于（确定的、固定的）基$\mathbf{\Psi}$的正交展开。 b_T 的树结构在 $\mathbb{T}^d$上产生随机分形；当定义 3.2 中的树 T在有限个节点之后仍未终止时，就会出现这种情况。由 [28, 引理 A.4],可知，当 $\beta \in (2^{-d},\ 1]$时，后一事件以正概率发生。在这种情况下，分形的豪斯多夫维数为 $d +$ 对数$_2(\beta) \in (0,\ d]$；详情见 [17, 第 3] 节。

为处理以 b_T 作为先验模型的椭圆型反问题，我们描述相应的参数概率空间。令 $\mathbb{Q}_0$ 表示定义在 $(\mathbb{R},\ \mathcal{B}(\mathbb{R}))$ 上的随机变量 $X^l_{j,\ k}$ 的一元 p 指数测度，其勒贝格密度如式 (22) 所示。p 指数 X 的乘积概率空间为 $(\Omega_p,\ \mathcal{A}_p,\ \mathbb{Q}_p)$，其中

$$\Omega_p := \mathbb{R}^{\mathbb{N}}, \quad \mathcal{A}_p := \bigotimes_{n\in\mathbb{N}} \mathcal{B}(\mathbb{R}), \quad \text{and} \quad \mathbb{Q}_p := \bigotimes_{n\in\mathbb{N}} \mathbb{Q}_0. \tag{24}$$

现在令 $s > 0$ 和 $p \in [1, \infty)$ 固定，并满足 $s > \frac{d}{p}$。我们定义加权 ℓ^p 空间

$$\ell^p_s := \left\{x = \left(x^l_{j,k}, (j,k,l) \in \mathcal{I}_{\mathbf{\Psi}}\right) \in \mathbb{R}^{\mathbb{N}} |\ \|x\|_{s,p} < \infty\right\},$$

其中

$$\|x\|_{s,p} := \left(\sum_{(j,k,l)\in\mathcal{I}_{\mathbf{\Psi}}} 2^{-jps}|x^l_{j,k}|^p\right)^{1/p}.$$

由于 $1 \le p < \infty$，$(\ell^p_s,\ \|\cdot\|_{s,\ p})$ 是一个可分巴拿赫空间。我们注意到，对于 $X \sim \mathbb{Q}_p$，有 s

$$\mathbb{E}(\|X\|^p_{s,p}) \le \sum_{(j,k,l)\in\mathcal{I}_{\mathbf{\Psi}}} 2^{-jps}\mathbb{E}(|X^l_{j,k}|^p) \le C\sum_{j=0}^{\infty} 2^{-jps}2^{dj}(2^d - 1) \le C\sum_{j=0}^{\infty} 2^{-jp(s-\frac{d}{p})} < \infty,$$

由于 $s > \frac{d}{p}$，因此 $\mathbb{Q}_p$ 集中于 ℓ^p_s 上。因此，我们可以将 $(\ell^p_s,\ \mathcal{B}(\ell^p_s)\ \mathbb{Q}_p)$ 视为展开式 (23) 中随机系数序列 X 的概率空间。

集值随机变量 T 是一个 GW 树，因此取值于由所有无无限节点的树构成的波兰空间 $(\mathfrak{T},\ \delta_{\mathfrak{T}})$。关于 $\mathfrak{T}$ 的度量 $\delta_{\mathfrak{T}}$ 及其相关的 Borel σ-代数 $\mathcal{B}(\mathfrak{T})$ 可在 [28, 定义 A2], 或 [1, 第 2 节1]中明确表示。于是，GW 树 T 在 $(\mathfrak{T},\ \mathcal{B}(\mathfrak{T}))$ 上的像测度$\mathbb{Q}_T$ 仅取决于后代分布 $\mathcal{P} = \mathrm{Bin}(2^d,\ \beta)$ 的参数 β 和 d，其表达式见 [28, 公式（77）]。因此，GW 树的参数概率空间为 $(\mathfrak{T},\ \mathcal{B}(\mathfrak{T})\ \mathbb{Q}_T)$。

为将随机系数 X 与 GW 树 T结合起来，我们定义笛卡尔积 $\Omega := \ell^p_s \times \mathfrak{T}$，并在 Ω 上赋予度量

$$d_\Omega((x_1, \mathbf{t_1}), (x_2, \mathbf{t_2})) := \|x_1 - x_2\|_{s,p} + \delta_{\mathfrak{T}}(\mathbf{t_1}, \mathbf{t_2}).$$

命题 3.4。空间 $(\Omega\ d_\Omega)$ 是波兰空间，其 *Borel* σ-代数为 $\mathcal{B}(\Omega) = \mathcal{B}(\ell^p_s \times \mathfrak{T}) = \mathcal{B}(\ell^p_s) \otimes \mathcal{B}(\mathfrak{T})$。

证明。根据 [1, 引理 2，1] 度量空间 $(\mathfrak{T}, \delta_{\mathfrak{T}})$（其中 $\delta_{\mathfrak{T}}$ 由 [28, 定义 A.2] 给出）是完备且可分的。因此，根据 [2, 推论 3，39]（$\Omega\ d_\Omega$）具有可分性和完备性。此外，由 [2, 定理 4.44]可知，$\mathcal{B}(\Omega) = \mathcal{B}(\ell^p_s \times \mathfrak{T}) = \mathcal{B}(\ell^p_s) \otimes \mathcal{B}(\mathfrak{T})$成立。 □

现在我们可以定义与 $\ell^p_s \times \mathfrak{T}$值随机变量 (X, T)相关联的先验概率空间：令 $(\Omega\ \mathcal{A}, \mathbb{P})$ 表示由下式给出的乘积概率空间

$$\Omega := \ell^p_s \times \mathfrak{T}, \quad \mathcal{A} := \mathcal{B}(\ell^p_s) \otimes \mathcal{B}(\mathfrak{T}), \quad \text{and} \quad \mathbb{P} := \mathbb{Q}_p \otimes \mathbb{Q}_T. \tag{25}$$

我们指出，测度 $\mathbb{P} = \mathbb{Q}_p \otimes \mathbb{Q}_T$ 的乘积结构等价于 X 和 T的随机独立性。

我们仍需确定随机变量 $(X,\ T)$ 的一个实现，使其与相应的随机树先验 b_T 相对应。为此，我们考虑典范映射

$$b_T : \Omega \to L^2(\mathbb{T}^d), \quad \omega \mapsto \sum_{(j,k,l)\in\mathcal{I}_T(\omega)} \eta_j X_{j,k}^l(\omega)\psi_{j,k}^l. \tag{26}$$

映射 b_T：$\Omega \to L^2(\mathbb{T}^d)$ 确实定义良好，因为 $\|b_T\|_{L^2(\mathbb{T}^d)} < \infty$ 成立，这是由于 $s > \frac{d}{p}$。此外，正如下面的命题 3.5 所示，b_T 是 $\mathcal{A}/\mathcal{B}(L^2(\mathbb{T}^d))$-可测的。因此，$b_T$ 在先验测度 $\mathbb{P}$ 下的推前概率测度为

$$b_T\#\mathbb{P}(B) := \mathbb{P}(b_T^{-1}(B)), \quad B \in \mathcal{B}(L^2(\mathbb{T}^d)). \tag{27}$$

B_p^s-随机变量 b_T 及其小波密度 β 所对应的概率空间为

$$(L^2(\mathbb{T}^d),\ \mathcal{B}(L^2(\mathbb{T}^d)),\ b_T\#\mathbb{P}).$$

根据 [28, 备注 2.9]，我们知道对于任意 $t \in (0, s-\frac{d}{p})$，$b_T\#\mathbb{P}$ 集中于 B_p^t。下面的定理 3.6 给出了一个更精细的结果：对于满足光滑度指数 $t = t(s,d,p,\beta,q)$ 的 $q \geq 1$，$b_T\#\mathbb{P}$ 集中于 Besov 空间 B_q^t。本节最后给出关于 b_T的可测性和逐路径正则性的两个结果。

命题 3.5。 *[28, 命题 2.10]* 设 $s > \frac{d}{p}$，$\beta \in [0,1]$，并设 b_T是一个具有小波密度 β的 B_p^s-随机变量。那么 b_T：$\Omega \to \mathrm{C}(\mathbb{T}^d)$，且 b_T 是（强） $\mathcal{A}/\mathcal{B}(\mathrm{C}(\mathbb{T}^d))$-可测的。

定理 3.6。 *[28, 定理 2.11]* 设 b_T 是一个具有小波密度 $\beta = 2^{\gamma-d}$的 B_p^s-随机变量，如定义 *3.2* 所述，其中 $\gamma \in (-\infty, d]$。

1.）成立：$b_T \in L^q(\Omega;\ \mathbb{P};\ B_q^t)$，因此对所有满足$t < s + \frac{d-\gamma}{q} - \frac{d}{p}$ 的 $t > 0$和 $q \geq 1$，该结论几乎处处成立。

2.）设 $s - \frac{d}{p} > 0$和 $t \in (0,\ s-\frac{d}{p})$。那么存在一个 $\varepsilon_p > 0$，使得

$$\mathbb{E}_{\mathbb{P}}\left(\exp\left(\varepsilon\|b\|_{\mathcal{C}^t}^p\right)\right) < \infty, \quad \varepsilon \in (0, \varepsilon_p),$$

特别地，对于任意 $q \geq 1$，都有 $b_T \in L^q(\Omega;\ \mathbb{P};\ \mathcal{C}^t)$。

3.）令 $q \geq 1$ 且 $s - \frac{d}{p} - \frac{\min(\gamma,\ 0)}{q} > 0$。对于任意 $t \in (0,\ s - \frac{d}{p} - \frac{\min(\gamma,\ 0)}{q})$，都有 $b_T \in L^q(\Omega;\ \mathbb{P};\ \mathcal{C}^t)$。

3.3 正问题的适定性与正则性

令 $\mathcal{D} \subset \mathbb{R}^d$，$d \in \{1,2,3\}$ 为凸多边形区域，其边界 $\partial\mathcal{D}$ 由有限条线段或平面片段构成。我们进一步假设 $\mathcal{D} \subseteq \mathbb{T}^d$。令 $\varphi|_{\mathcal{D}}$ 表示任意 $\varphi \in \mathrm{S}'(\mathbb{R}^d)$ 在 $\mathcal{D}$ 上的限制，而它又由元素 $\varphi|_{\mathcal{D}} \in \mathrm{D}'(\mathcal{D})$ 给出，使得

$${}_{\mathrm{D}'(\mathcal{D})}\langle \varphi|_{\mathcal{D}}, v\rangle_{\mathrm{D}(\mathcal{D})} = {}_{\mathrm{S}'(\mathbb{R}^d)}\langle \varphi, v_0\rangle_{\mathrm{S}(\mathbb{R}^d)}, \quad v \in \mathrm{D}(\mathcal{D}),$$

其中 $v_0 \in \mathrm{D}(\mathbb{R}^d) \subset \mathrm{S}(\mathbb{R}^d)$ 表示任意 $v \in \mathrm{D}(\mathcal{D})$ 的零延拓（参见 [32, 第 2] 节）。根据 [32, 定理 1.29]，对于任意 $b \in B_p^s$，存在唯一的一周期延拓 $\mathrm{ext}^{per}(b) : \mathbb{R}^d \to \mathbb{R}$，使得 $b = \mathrm{ext}^{per}(b)|_{\mathcal{D}}$；进一步细节也请参见 [28, 第 4节。

因此，定义 3.2 中给出的 b_T 对 $\mathcal{D} \subseteq \mathbb{T}^d$ 的限制为

$$b_{T,\mathcal{D}}(\omega) := (\mathrm{ext}^{per} b_T(\omega))|_{\mathcal{D}}. \tag{28}$$

我们称 $b_{T,\ \mathcal{D}}$ 一个 $B_p^s(\mathcal{D})$值随机变量。

注 3.7。 请注意，如果 $\mathcal{D} \subsetneq \mathbb{T}^d$，$b_{T,\ \mathcal{D}}$ 不一定是周期函数，而只是一个周期函数的限制，该函数定义在环面 $\mathbb{T}^d$上。为简洁起见，假设 $\mathcal{D} \subseteq \mathbb{T}^d$，这对以下结果没有实质影响：若 $\mathcal{D} \not\subset \mathbb{T}^d$ 是有界区域，我们可以将定义（3.2）从环面 $\mathbb{T}^d$ 推广到足够大的（周期）区域 $[-L, L]^d$，其中 $L > 1$ 使得 $\mathcal{D} \subset [-L, L]^d$。然后，我们只需将 $b_{T,\ \mathcal{D}}$ 定义为该扩展区域上一个L-周期函数的限制。

现在，我们在 (2) 中令 $a = \exp(b_{T,\mathcal{D}})$，得到带有 Besov 随机树先验的椭圆正问题，即对给定的 $\omega \in \Omega$，求满足下式的 $u(\omega) : \mathcal{D} \to \mathbb{R}$

$$-\nabla \cdot (\exp(b_{T,\mathcal{D}}(\omega))\nabla u(\omega)) = f \quad \text{in } \mathcal{D}, \quad u(\omega) = 0 \quad \text{on } \partial\mathcal{D}. \tag{29}$$

定理 3.8。 *[28,* 定理 *3.9]* 设 $b_{T,\mathcal{D}}$ 在(28) 中给定，其中 $p \in [1, \infty)$、 $s > 0$ 和 $\beta \in [0, 1]$，使得 $sp > d$。此外，设 $f \in V'$。则以下断言成立。

1.) 几乎必然存在唯一的弱解 $u(\omega) \in V$，它满足(29)，且 $u : \Omega \to V$ 是强可测的。

2.) 对于 *(22)* 中充分小的 $\kappa > 0$，存在常数 $q \in (1, \infty)$ 和 $C^* > 0$，使得

$$\|u\|_{L^q(\Omega,\mathbb{P};V)} \leq C\|f\|_{V'} < \infty \quad \begin{cases} \text{for } q \in [1, \overline{q}) \text{ if } p = 1, \text{ and} \\ \text{for any } q \in [1, \infty) \text{ if } p > 1. \end{cases}$$

3.) 令 $f \in H$, $r \in (0, s - \frac{d}{p}) \cap (0, 1]$, 若 $r < 1$，则令 $r_0 \in (0, r)$；若 $r = 1$，则令 $r_0 = 1$。对于 *(22)* 中充分小的 $\kappa > 0$，存在常数 $q \in (1, \overline{\infty})$ 和 $C > 0$，使得。

$$\|u\|_{L^q(\Omega,\mathbb{P};H^{1+r_0}(\mathcal{D}))} \leq C\|f\|_H < \infty \quad \begin{cases} \text{for } q \in [1, \overline{q}) \text{ if } p = 1 \text{ and} \\ \text{for any } q \in [1, \infty) \text{ if } p > 1. \end{cases}$$

我们注意到，正向偏微分方程 (29) 中的非负扩散系数不具有严格正的下界。我们还指出，定理 3.8 的第 2.) 和第 3.) 部分中“对于充分小的 $\kappa > 0$”这一条件仅适用于 $p = 1$，并保证对于某个（依赖于 κ 的） $q \geq 1$，有 $\exp(\|b_{T,\ \mathcal{D}}\|_{L^\infty(\mathcal{D})}) \in L^q(\Omega)$，或相应地有 $\exp(\|b_{T,\ \mathcal{D}}\|_{C^r(\mathcal{D})}) \in L^q(\Omega)$。

3.4 正向问题的路径逼近

为获得对 (23) 中 b_T 的可处理近似，我们在 $N \in \mathbb{N}$ 个尺度后截断小波级数展开，从而得到截断随机树 *Besov* 先验

$$b_{T,N}(\omega) := \sum_{\substack{(j,k,l) \in \mathcal{I}_T(\omega) \\ j \leq N}} \eta_j X^l_{j,k}(\omega)\psi^l_{j,k}, \quad \omega \in \Omega. \tag{30}$$

对于固定的 $\omega \in \Omega$，具有截断系数的相应扩散问题的弱形式为：求 $u_N(\omega) \in V$，使得对所有 $v \in V$

$$\int_{\mathcal{D}} a_N(\omega)\nabla u_N(\omega) \cdot \nabla v dx = {}_{V'}\langle f, v\rangle_V, \tag{31}$$

其中

$$a_N : \Omega \to L^\infty(\mathcal{D}), \quad \omega \mapsto \exp(b_{T,N}(\omega)|_{\mathcal{D}}). \tag{32}$$

带有截断系数的 Problem (31) 的解 $u_N : \Omega \to V$ 仍然无法完全处理，因为它的取值属于无限维希尔伯特空间 V。因此，在本节余下部分，我们针对固定的截断指标 N，考虑 u_N 的伽辽金有限元近似。

首先，我们通过一系列单形（区间/三角形/四面体）或平行多面体（区间/平行四边形/平行六面体）对凸域 $\mathcal{D} \subset \mathbb{T}^d$, $d \in \{1, 2, 3\}$ 进行离散化，记为 $(\mathcal{K}_h)_{h\in\mathfrak{H}}$。细化参数 $h > 0$ 取自可数指标集$\mathfrak{H} \subset (0, \infty)$，对应于单形/平行多面体 $K \in \mathcal{K}_h$的最长边。为获得一系列“性质良好”的三角剖分，我们对 $(\mathcal{K}_h)_{h\in\mathfrak{H}}$ 作出以下假设。

假设 3.9。 序列 $(\mathcal{K}_h)_{h\in\mathfrak{H}}$ 满足：

1. 容许性： 对于每个 $h \in \mathfrak{H}$， $\mathcal{K}_h$ 由开且非空的单形/平行多面体 K组成，使得

   - $\mathcal{D} = \bigcup_{K\in\mathcal{K}_h} K$，
   - $K_1 \cap K_2 = \emptyset$ 对于任意两个 K_1， $K_2 \in \mathcal{K}_h$，满足 $K_1 \neq K_2$，并且
   - 交集 $K_1 \cap K_2$ 对于 $K_1 \neq K_2$ 而言，要么为空，要么是公共边、公共顶点，或者（在空间维数 $d = 3$ 中）是 K_1 和 K_2的公共面。

2. 形状正则性： 设 $\rho_{K,\ in}$ 和 $\rho_{K,\ out}$ 分别表示给定 的内接圆和外接圆半径， re-
   分别地，对于给定的 $K \in \mathcal{K}_h$。 存在一个常数 $\rho > 0$，使得

$$\rho := \sup_{h\in\mathfrak{H}} \sup_{K\in\mathcal{K}_h} \frac{\rho_{K,out}}{\rho_{K,in}} < \infty.$$

基于给定的剖分 $\mathcal{K}_h$，我们定义分片（多）线性有限元空间

$$V_h := \begin{cases} \{v \in V |\ v|_T \text{ is linear for all } K \in \mathcal{K}_h\}, & \text{if } \mathcal{K}_h \text{ consists of simplices,} \\ \{v \in V |\ v|_T \text{ is } d\text{-linear for all } K \in \mathcal{K}_h\}, & \text{if } \mathcal{K}_h \text{ consists of parallelotopes.} \end{cases}$$

Cle因此， $V_h \subset V$ 是一个有限维空间，并定义 $n_h := \dim(V_h) \in \mathbb{N}$。由此得到 for
fixed $\omega \in \Omega$ 上的全离散问题 为求解 $u_{N,h}(\omega) \in V_h$，使得对所有 $v_h \in V_h$

$$\int_{\mathcal{D}} a_N(\omega)\nabla u_{N,h}(\omega) \cdot \nabla v_h dx = {}_{V'}\langle f, v_h\rangle_V. \tag{33}$$

截断误差与有限元近似误差之和由下一个结果给出界限。

定理 3.10。 *[28, 定理 4.4、4.7 和 4.8]* 设$(\mathcal{K}_h)_{h\in\mathfrak{H}}$ 是满足假设 *3.9* 的三角剖分序列，并设 u、 u_N 和 $u_{N,h}$ 分别为给定 $N \in \mathbb{N}$ 和 $h \in \mathfrak{H}$时方程(29)、(30)和(33) 的逐路径弱解。此外，设 $p \in [1, \infty)$ 且 $s > 0$ 满足 $sp > d$。

对于任意的 $f \in H$，充分小的 $\kappa > 0$ 属于(22)，任意的 $r \in (0, s - \frac{d}{p}) \cap (0, 1]$ 以及 $t \in (0, s - \frac{d}{p})$，存在常数 $q \in (1, \infty)$ 和 $C > 0$，使得对于任意的 $N \in \mathbb{N}$ 和 $h \in \mathfrak{H}$，均有

$$\|u - u_N\|_{L^q(\Omega,\mathbb{P};V)} \leq C2^{-Nt} \quad \begin{cases} \text{for } q \in [1, \overline{q}) \text{ if } p = 1, \\ \text{for any } q \in [1, \infty) \text{ if } p > 1, \end{cases}$$

$$\|u_N - u_{N,h}\|_{L^q(\Omega,\mathbb{P};V)} \leq Ch^r \quad \begin{cases} \text{for } q \in [1, \overline{q}) \text{ if } p = 1, \\ \text{for any } q \in [1, \infty) \text{ if } p > 1, \end{cases}$$

$$\|u_N - u_{N,h}\|_{L^q(\Omega,\mathbb{P};H)} \leq Ch^{2r} \quad \begin{cases} \text{for } q \in [1, \overline{q}) \text{ if } p = 1, \\ \text{for any } q \in [1, \infty) \text{ if } p > 1. \end{cases}$$

4 具有 Besov 随机树先验的逆问题

再次考虑第 2 节中的贝叶斯逆问题设定，其中假设式（2）中的 $\log(a)$由 Besov 随机树先验给出，即 $a = \exp(b_{T,\mathcal{D}})$。我们在此设定下验证假设 2.2，以确保命题 2.3 和推论 2.4 成立。

引理 4.1。 设定理 *3.8* 的假设在 $q \geq 2$ 下成立，且令 对数(a) 在 (2) 中由 *Besov* 随机树先验给出，即为 $a = \exp(b_{T,\mathcal{D}})$，并令 ρ 由 (7) 给出。则对于 Φ 和 ρ，假设 *2.2* 成立。

证明。 由定理 3.8 可得 $u \in L^q(\Omega\ \mathbb{P}; V)$。这意味着存在一个常数$C_u > 0$ 和一个集合 $\Omega_u \in \mathcal{A}$，满足 $\mathbb{P}(\Omega_u) > 0$，使得对所有 $\omega \in \Omega_u$， $\|u(\omega)\|_V \le C_u$。现在固定 $\lambda > 0$，并定义 $\chi :$ $= \lambda + \|\mathcal{O}\|_{((H_0^\theta \mathcal{O})')_k} C_u$。于是，对所有 $\omega \in \Omega_u$ 及 $\delta \in B_\lambda(0)$，我们得到

$$\Phi(\omega;\delta) \le |\log(\rho(\delta - [\mathcal{O}\circ u](\omega)))| \le \sup_{x\in B_\chi(0)} |\log(\rho(x))| \le \frac{k}{2}|\log(2\pi\det(\Sigma))| + \frac{1}{2}\|\Sigma^{-1}\|_2\chi^2 < \infty.$$

因此， Φ 满足假设 2.2 的第 1 项，其中 $\Omega_\lambda := \Omega_u$，且

$$\kappa_1(\lambda) := \frac{k}{2}|\log(2\pi\det(\Sigma))| + \frac{1}{2}\|\Sigma^{-1}\|_2(\lambda + \|\mathcal{O}\|_{((H_0^\theta\mathcal{O})')^k} C_u)^2.$$

为证明第 2 项，我们固定 $\omega \in \Omega$，并令 $\delta, \delta' \in B_\lambda(0)$。由 (7) 可得

$$\begin{aligned}
|\Phi(\omega;\delta) - \Phi(\omega;\delta')| &= \left|\log\left(\frac{\rho(\delta - [\mathcal{O}\circ u](\omega))}{\rho(\delta' - [\mathcal{O}\circ u](\omega))}\right)\right| \\
&= \frac{1}{2}\left|-\delta^\top\Sigma^{-1}\delta + (\delta')^\top\Sigma^{-1}(\delta') + 2[\mathcal{O}\circ u](\omega)^\top\Sigma^{-1}(\delta - \delta')\right| \\
&\le \frac{1}{2}\left|-\delta^\top\Sigma^{-1}(\delta - \delta') + (\delta')^\top\Sigma^{-1}(\delta' - \delta)\right| + \left|[\mathcal{O}\circ u](\omega)^\top\Sigma^{-1}(\delta' - \delta)\right| \\
&\le \left(\max(\|\delta\|_2, \|\delta'\|_2) + \|\mathcal{O}\|_{((H_0^\theta\mathcal{O})')^k}\|u(\omega)\|_V\right)\|\Sigma^{-1}\|_2\|\delta - \delta'\|_2 \\
&\le \left(\lambda + \|\mathcal{O}\|_{((H_0^\theta\mathcal{O})')^k}\|u(\omega)\|_V\right)\|\Sigma^{-1}\|_2\|\delta - \delta'\|_2.
\end{aligned}$$

由于根据定理 3.8， $u \in L^2(\Omega\ \mathbb{P}; V)$，因此假设 2.2 的第 2 项成立，其中

$$\kappa_2(\lambda,\omega) := \left(\lambda + \|\mathcal{O}\|_{((H_0^\theta\mathcal{O})')^k}\|u(\omega)\|_V\right)\|\Sigma^{-1}\|_2.$$

□

注释 4.2. 令 $\lambda_{\min} > 0$ 表示 Σ的最小特征值。从引理 4.1 的证明可明显看出，对于固定的 ω 和 λ， $\kappa_1(\lambda) = \mathcal{O}(\lambda^{-1}$ 最小值$)$， $\kappa_2(\lambda,\omega) = \mathcal{O}(\lambda^{-1}$ 最小值$)$。因此，当 $\lambda_{\min} \to 0$时，假设 2.2 中的界在“小噪声极限”下会变差。相应地，对于命题 2.3 中固定的 $\underline{\lambda} > 0$， $C(\lambda) = \mathcal{O}(\lambda^{-1}$ 最小值$)$；而在推论 2.4 中，取 κ_1 如引理 4.1 的证明所示时， $c(\lambda) = \mathcal{O}(\lambda^{\frac{k}{2}}$ 最小值$)$。

由（8）得到的（精确）后验分布 $\mathbb{P}_\delta$ 通常无法获得，因为只能获得正问题的有偏样本 $u_{N,\ h} \approx u$，如（33）所示。因此，我们考虑近似后验

$$d\mathbb{P}_{\delta,N,h}(\omega) := \frac{\exp\left(-\Phi_{N,h}(\omega;\delta)\right)d\mathbb{P}(\omega)}{Z_{N,h}(\delta)}, \tag{34}$$

其离散贝叶斯势和归一化常数定义如下

$$\Phi_{N,h}(\omega;\delta) := -\log\left(\rho(\delta - [\mathcal{O}\circ u_{N,h}](\omega))\right), \quad \text{and} \quad Z_{N,h}(\delta) := \int_\Omega \exp\left(-\Phi_{N,h}(\omega;\delta)\right)d\mathbb{P}(\omega) > 0.$$

命题 4.3。 设定理 *3.10* 的假设成立，使得 $q \ge 4$ 在 $p = 1$ 的情况下成立。那么，对任意 $\lambda > 0$，存在一个 $C(\lambda) > 0$，它与 N 和 h 无关，使得

$$d_{\text{Hell}}(\mathbb{P}_\delta, \mathbb{P}_{\delta,N,h}) \le C(\lambda)(2^{-Nt} + h^{(2-\theta_\mathcal{O})r}) \quad \textit{for all } \delta \in B_\lambda(0).$$

为了证明命题 4.3，我们需要 $Z_{N,\ h}(\delta)$ 的一致下界：

引理 4.4. 设定理 *3.10* 的假设成立。那么，对于任意 $\lambda > 0$，存在一个常数$c(\lambda) > 0$，该常数与 N 和 h 无关，使得

$$Z_{N,h}(\delta) \ge c(\lambda) > 0 \quad \textit{for all } \delta \in B_\lambda(0). \tag{35}$$

~证明. 根据定理 3.10，存在一个常数 $C>0$，使得对于所有 $N\in\mathbb{N}$ 和 $h\in\mathfrak{H}$，均有

$$\|u-u_{N,h}\|_{L^1(\Omega,\mathbb{P};V)}\le C(2^{-Nt}+h^r)\le\widetilde{C}<\infty. \tag{36}$$

由于 $u\in L^1(\Omega\ \mathbb{P};V)$，由反三角不等式可得

$$\|u_{N,h}\|_{L^1(\Omega,\mathbb{P};V)}\le\widetilde{C}+\|u\|_{L^1(\Omega,\mathbb{P};V)}<\infty. \tag{37}$$

~ ~现在定义 $C_u:=2(C+\|u\|_{L^1(\Omega\ \mathbb{P};V)})>0$，以及集合

$$\widetilde{\Omega}_u:=\{\omega\in\Omega\,\big|\,\|u_{N,h}\|_V\le\widetilde{C}_u\}\in\mathcal{A}.$$

~集合 Ω_u 通常取决于 N 和 h。由马尔可夫不等式、不等式 (37) 以及~ C_u 的定义可得

$$\mathbb{P}(\widetilde{\Omega}_u)=1-\mathbb{P}(\|u_{N,h}\|_V>\widetilde{C}_u)\ge1-\frac{\|u_{N,h}\|_{L^1(\Omega,\mathbb{P};V)}}{\widetilde{C}_u}\ge\frac{1}{2}, \tag{38}$$

~对所有 $N\in\mathbb{N}$ 和 $h\in\mathfrak{H}$均成立。现在固定 $\lambda>0$。对所有 $\omega\in\Omega_u$ 和 $\delta\in B_\lambda(0)$，我们有

$$\Phi_{N,h}(\omega;\delta)\le|\log(\rho(\delta-[\mathcal{O}\circ u_{N,h}](\omega)))|\le\sup_{x\in B_\chi(0)}|\log(\rho(x))|<\infty,$$

其中 $\chi:=\lambda+\|\mathcal{O}\|_{((H_0^\theta\mathcal{O})')_k}\widetilde{C}_u$，最后一个估计由 $\log\circ\rho:\ \mathbb{R}^k\to\mathbb{R}$.~的连续性可知是有限的。因此，$\Phi_{N,\ h}$ 满足假设 2.2 的第一部分，其中 $\Omega_\lambda:=\Omega_u$，且

$$\kappa_1(\lambda):=\sup_{x\in B_\chi(0)}|\log(\rho(x))|<\infty,$$

~由（38）可知。由于 $\mathbb{P}(\Omega_u)$ 和$\kappa_1(\lambda)$ 关于 N 和 h一致有界，因此该结论可仿照推论 2.4 的证明得出。 □

命题 *4.3* 的证明：证明基本遵循 [14, 命题 10], 的证明，只需在适当位置代入定理 3.10 中的估计。由于 $\mathbb{P}_\delta$ 和 $\mathbb{P}_{\delta,\ N,\ h}$关于 $\mathbb{P}$都绝对连续，因此对于固定的 $\lambda>0$ 以及任意 $\delta\in B_\lambda(0)$，有

$$\begin{aligned}
&2d_{\mathrm{Hell}}(\mathbb{P}_\delta,\mathbb{P}_{\delta,N,h})^2\\
&=\int_\Omega\left(Z(\delta)^{-1/2}\exp\left(-\frac{1}{2}\Phi(\omega;\delta)\right)-Z_{N,h}(\delta)^{-1/2}\exp\left(-\frac{1}{2}\Phi_{N,h}(\omega;\delta)\right)\right)^2d\mathbb{P}(\omega)\\
&\le2\int_\Omega Z(\delta)^{-1}\left(\exp\left(-\frac{1}{2}\Phi(\omega;\delta)\right)-\exp\left(-\frac{1}{2}\Phi_{N,h}(\omega;\delta)\right)\right)^2d\mathbb{P}(\omega)\\
&\quad+2\int_\Omega\left(Z(\delta)^{-1/2}-Z_{N,h}(\delta)^{-1/2}\right)^2\exp\left(-\Phi_{N,h}(\omega;\delta)\right)\mathbb{P}(\omega)\\
&:=2(I+II).
\end{aligned}\tag{39}$$

为了对 I 进行界定，我们使用泰勒展开，并利用式（9）中的 Σ 是正定的，得到

$$\begin{aligned}
&\left|\exp\left(-\frac{1}{2}\Phi(\omega;\delta)\right)-\exp\left(-\frac{1}{2}\Phi_{N,h}(\omega;\delta)\right)\right|\\
&\quad\le\frac{1}{2}\exp\left(-\frac{1}{2}\frac{k\log\left(2\pi\det(\Sigma)\right)}{2}\right)|\Phi(\omega;\delta)-\Phi_{N,h}(\omega;\delta)|\\
&\quad\le C\left|\Phi(\omega;\delta)-\Phi_{N,h}(\omega;\delta)\right|,
\end{aligned}$$

其中 $C>0$ 是一个确定性常数。现在令 $C_\rho \in (0,\infty)$为式（7）中 ρ 的Lipschitz 常数。我们结合式（39）中的界和H¨older不等式，对第一项进行界定，得到

$$
\begin{aligned}
I &\le C\int_\Omega Z(\delta)^{-1}\left|\Phi(\omega;\delta)-\Phi_{N,h}(\omega;\delta)\right|^2 d\mathbb{P}(\omega)\\
&= C\int_\Omega Z(\delta)^{-1}\left|\rho(\delta-[\mathcal{O}\circ u](\omega))-\rho(\delta-[\mathcal{O}\circ u_{N,h}](\omega))\right|^2 d\mathbb{P}(\omega)\\
&\le C\frac{C_\rho^2\|\mathcal{O}\|^2_{((H_0^{\theta_\mathcal{O}})')^k}}{Z(\delta)}\int_\Omega \|u(\omega)-u_{N,h}(\omega)\|^2_{H^{\theta_\mathcal{O}}} d\mathbb{P}(\omega)\\
&\le \frac{C}{Z(\delta)}\int_\Omega \|u(\omega)-u_N(\omega)\|_V^2+\|u_N(\omega)-u_{N,h}(\omega)\|^2_{H^{\theta_\mathcal{O}}} d\mathbb{P}(\omega)\\
&\le \frac{C}{Z(\delta)}\int_\Omega \|u(\omega)-u_N(\omega)\|_V^2+\|u_N(\omega)-u_{N,h}(\omega)\|_V^{2\theta_\mathcal{O}}\|u_N(\omega)-u_{N,h}(\omega)\|_H^{2(1-\theta_\mathcal{O})} d\mathbb{P}(\omega)\\
&\le \frac{C}{Z(\delta)}\left(\|u-u_N\|^2_{L^2(\Omega,\mathbb{P};V)}+\|u_N-u_{N,h}\|^{2\theta_\mathcal{O}}_{L^{4\theta_\mathcal{O}}(\Omega,\mathbb{P};V)}\|u_N-u_{N,h}\|^{2(1-\theta_\mathcal{O})}_{L^{4(1-\theta_\mathcal{O})}(\Omega,\mathbb{P};H)}\right).
\end{aligned}
$$

然后，我们使用定理 3.10 和 $q:=4\max(\theta_\mathcal{O},(1-\theta_\mathcal{O}))\le 4$ 推导出估计

$$
I\le \frac{C}{Z(\delta)}(2^{-2Nt}+h^{2\theta_\mathcal{O} r}h^{2(1-\theta_\mathcal{O})2r})\le \frac{C}{Z(\delta)}(2^{-Nt}+h^{(2-\theta_\mathcal{O})r})^2.
$$

第二项可由以下式子界定

$$
\begin{aligned}
II &= \left(Z(\delta)^{-1/2}-Z_{N,h}(\delta)^{-1/2}\right)^2 Z_{N,h}(\delta)\\
&\le \frac{Z_{N,h}(\delta)}{\min(Z(\delta),Z_{N,h}(\delta))^3}\left|Z_{N,h}(\delta)-Z(\delta)\right|^2\\
&\le \frac{\|\rho\|_{L^\infty(\mathbb{R}^k)}}{\min(Z(\delta),Z_{N,h}(\delta))^3}\left|Z_{N,h}(\delta)-Z(\delta)\right|^2.
\end{aligned}
$$

与 I类似，利用 ρ 的 Lipschitz 连续性和定理 3.10 可得

$$
\begin{aligned}
|Z_{N,h}(\delta)-Z(\delta)|^2 &= \left|\int_\Omega \rho(\delta-[\mathcal{O}\circ u_{N,h}](\omega))-\rho(\delta-[\mathcal{O}\circ u](\omega))d\mathbb{P}(\omega)\right|^2\\
&\le C(2^{-Nt}+h^{(2-\theta_\mathcal{O})r})^2.
\end{aligned}
$$

因此，由于……，根据引理 4.1 和引理 4.4 以及推论 2.4，即可得出结论

$$
d_{\mathrm{Hell}}(\mathbb{P}_\delta,\mathbb{P}_{\delta,N,h})\le C\frac{\max(1,\|\rho\|_{L^\infty(\mathbb{R}^k)})}{\min(1,Z(\delta),Z_{N,h}(\delta))^{3/2}}(2^{-Nt}+h^{(2-\theta_\mathcal{O})r})\le \frac{C}{c(\lambda)^{3/2}}(2^{-Nt}+h^{(2-\theta_\mathcal{O})r}).
$$

□

令 $\mathbb{E}_{\delta,N,h}(\cdot):=\mathbb{E}_{\mathbb{P}_{\delta,N,h}}(\cdot)$ 表示关于近似后验分布$\mathbb{P}_{\delta,N,h}$的期望。命题 4.3 中的界控制了 $\mathbb{E}_\delta$ 与$\mathbb{E}_{\delta,N,h}$之间的差异。

定理 4.5。 设 $(\mathcal{X},\|\cdot\|_\mathcal{X})$ 为任意巴拿赫空间，且 $\varphi\in L^2(\Omega\ \mathbb{P};\mathcal{X})$。在定理 *3.10* 的假设下，对于任意 $\lambda>0$，存在一个常数 $C(\lambda)>0$，与 φ、 N 和 h 无关，使得

$$
\|\mathbb{E}_\delta(\varphi)-\mathbb{E}_{\delta,N,h}(\varphi)\|_\mathcal{X}\le C(\lambda)\|\varphi\|_{L^2(\Omega,\mathbb{P};\mathcal{X})}(2^{-Nt}+h^{(2-\theta_\mathcal{O})r})\quad \textit{for all } \delta\in B_\lambda(0).
$$

证明。我们固定 $\lambda > 0$，并任取 $\delta \in B_\lambda(0)$。显然，密度 ρ 在 $\mathbb{R}^k$ 上连续且有界，因此 $\|\rho\|_{L^\infty(\mathbb{R}^k)} < \infty$。结合推论 2.4 和 $\varphi \in L^2(\Omega, \mathbb{P}; \mathcal{X})$，可得

$$\begin{aligned}\|\varphi\|^2_{L^2(\Omega,\mathbb{P}_\delta;\mathcal{X})} &= \int_\Omega \|\varphi(\omega)\|^2_{\mathcal{X}} \frac{\exp\left(-\Phi(\omega;\delta)\right)}{Z(\delta)} d\mathbb{P}(\omega) \\ &= \frac{1}{Z(\delta)} \int_\Omega \|\varphi(\omega)\|^2_{\mathcal{X}} \rho(\delta - \mathcal{G}(\omega)) d\mathbb{P}(\omega) \\ &\leq \frac{\|\rho\|_{L^\infty(\mathcal{D})}}{c(\lambda)} \|\varphi\|^2_{L^2(\Omega,\mathbb{P};\mathcal{X})} < \infty.\end{aligned}$$

因此，对于 $\delta \in B_\lambda(0)$，$\varphi \in L^2(\Omega, \mathbb{P}_\delta; \mathcal{X})$ 一致有界。以同样的方式可得，$\varphi \in L^2(\Omega, \mathbb{P}_{\delta,N,h}; \mathcal{X})$ 关于 N、h 以及 $\delta \in B_\lambda(0)$ 一致有界，因为引理 4.4 表明

$$\begin{aligned}\|\varphi\|^2_{L^2(\Omega,\mathbb{P}_{\delta,N,h};\mathcal{X})} &= \frac{1}{Z_{N,h}(\delta)} \int_\Omega \|\varphi(\omega)\|^2_{\mathcal{X}} \rho(\delta - [\mathcal{O} \circ u_{N,h}](\omega)) d\mathbb{P}(\omega) \\ &\leq \frac{\|\rho\|_{L^\infty(\mathcal{D})}}{c(\lambda)} \|\varphi\|^2_{L^2(\Omega,\mathbb{P};\mathcal{X})} < \infty.\end{aligned}$$

根据 [30, 引理 6.37]，我们进而得到

$$\|\mathbb{E}_\delta(\varphi) - \mathbb{E}_{\delta,N,h}(\varphi)\|_{\mathcal{X}} \leq 2 \left(2 \frac{\|\rho\|_{L^\infty(\mathcal{D})}}{c(\lambda)}\right)^{1/2} \|\varphi\|_{L^2(\Omega,\mathbb{P};\mathcal{X})} d_{\mathrm{Hell}}(\mathbb{P}_\delta, \mathbb{P}_{\delta,N,h}),$$

由命题 4.3 可知，结论对 $C(\lambda) := 2\left(2\frac{\|\rho\|_{L^\infty(\mathcal{D})}}{c(\lambda)}\right)^{1/2}$ 成立。

□

5 马尔可夫链蒙特卡洛

我们使用马尔可夫链蒙特卡洛（MCMC）抽样来获得近似后验测度 $\mathbb{P}_{\delta,N,h}$，其中假设对于一个固定的 $\lambda > 0$，$\delta \in B_\lambda(0)$。为简洁起见，我们假设 $h^{(2-\theta_\mathcal{O})r} \simeq 2^{-Nt}$，以平衡截断误差与有限元误差，并在全文中使用缩写 $\mathbb{P}_h := \mathbb{P}_{\delta,N,h}$ 和 $\mathbb{E}_h := \mathbb{E}_{\delta,N,h}$。

5.1 单层马尔可夫链蒙特卡洛

给定当前状态 $\omega^{(k)}$，我们抽取一个候选状态 $v^{(k)} \sim \mathbb{Q}(\omega^{(k)}; \cdot)$，其中 $\mathbb{Q}(\omega^{(k)}; \cdot) : \mathcal{A} \to [0,1]$是给定的、定义在 $(\Omega, \mathcal{A})$ 上的提议概率测度，并依赖于当前状态 $\omega^{(k)}$。进一步，我们在 $\Omega \times \Omega$ 上定义测度 ν 和 $\nu^\top$，使得对任意 $(\omega, v) \in \Omega \times \Omega$，有 $d\nu(\omega, v) := \mathbb{Q}(\omega; dv) d\mathbb{P}_h(\omega)$，以及 $d\nu(\omega, v)^\top := \mathbb{Q}(v; d\omega) d\mathbb{P}_h(v)$，并假设 $\mathbb{Q}$ 的选取满足 $\nu^\top \ll \nu$。以接受概率接受新的提议作为下一个状态，即 $\omega^{(k+1)} = v^{(k)}$，其接受概率为

$$\alpha(\omega^{(k)}, v^{(k)}) := \min\left\{1, \frac{d\nu^\top(\omega^{(k)}; v^{(k)})}{d\nu(\omega^{(k)}; v^{(k)})}\right\}. \tag{40}$$

注意，由于假设 $\nu^\top \ll \nu$，(40) 中的 α 定义良好。如果拒绝 $v^{(k)}$，则保留当前状态 $\omega^{(k+1)} = \omega^{(k)}$。这种方法是 Metropolis-Hastings 算法的一种变体，它生成一个平稳分布为 $\mathbb{P}_h$的马尔可夫链 $(\omega^{(k)}, k \in \mathbb{N})$。显然，生成的样本 $(\omega^{(k)}, k \in \mathbb{N})$ 以非平凡的方式相关。众所周知，选择合适的提议密度 $\mathbb{Q}$ 可以降低相关性并提高算法效率。我们将特别关注独立采样器，其中 $\mathbb{Q}(\omega^{(k)}; dv^{(k)}) = d\mathbb{P}(v^{(k)})$，也就是说，提议 $v^{(k)}$ 从先验测度 $\mathbb{P}$中抽取，且独立于马尔可夫链的当前状态 $\omega^{(k)}$。

现在，设 $\omega^{(1)}, \ldots, \omega^{(M)}, M \in \mathbb{N}$ 是来自 $\mathbb{P}_h$ 的马尔可夫链蒙特卡洛样本序列。我们旨在估计 φ 的后验均值：$\Omega \to \mathbb{R}$, $\omega \mapsto [\Psi \circ u](\omega)$，其中 $\Psi : V \to \mathbb{R}$ 是一个确定性泛函

且 u 是方程 (29) 的解。相应的 $\mathbb{E}_h(\varphi)$ 的马尔可夫链蒙特卡罗估计量记为

$$E_M^h(\varphi) := \frac{1}{M}\sum_{i=1}^{M}\varphi(\omega^{(i)}). \tag{41}$$

马尔可夫链蒙特卡洛估计量的采样误差由下一个结果给出界。

引理 5.1 (独立采样器的几何遍历性)。令 $\mathcal{Q}$ 表示初始样本的分布 $\omega^{(1)}$，并令 $\mathcal{P}_{\mathcal{Q}}$ 表示由马尔可夫链蒙特卡洛独立采样器生成的概率空间上的概率测度。此外，我们用 $\mathcal{E}_{\mathcal{Q}}$ 表示关于 $\mathcal{P}_{\mathcal{Q}}$的期望。存在 $C = C(\lambda) > 0$ （回忆一下，$\delta \in B_\lambda(0)$），使得对所有 $\varphi \in L^2(\Omega\ \mathbb{P})$ 以及 $h \in \mathfrak{H}$，均有

$$\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_h(\varphi) - E_M^h(\varphi)\right)^2\right) \le C\|\varphi\|^2_{L^2(\Omega,\mathbb{P})}M^{-1}. \tag{42}$$

证明。由引理 4.4 可知，对任意 $h \in \mathfrak{H}$ 和 $N \in \mathbb{N}$，一致下界为

$$Z_{N,h}(\delta) \ge c(\lambda) > 0 \quad \text{for all } \delta \in B_\lambda(0),$$

因此

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\ \frac{\exp(-\Phi_{N,h}(\omega;\delta))}{Z_{N,h}(\delta)} \le \frac{\|\rho\|_\infty}{c(\lambda)} < \infty.$$

不失一般性，我们假设 $\frac{\|\rho\|_\infty}{c\ \lambda}() > 1$。令 $\mathbb{Q}^{(n)}(\widetilde{\omega};\cdot) : \Omega \to [0,1]$ 表示从（固定的）$\widetilde{\omega} \in \Omega$ 出发，经过 n 步后马尔可夫链的分布。根据 [29,定理 1 和式 (13)]，对于任意满足以下条件的 $A \in \mathcal{A}$

$$\left|\int_A \mathbb{Q}^{(n)}(\widetilde{\omega}, d\omega) - \mathbb{P}_h(A)\right| \le C\left(1 - \frac{c(\lambda)}{\|\rho\|_\infty}\right)^n,$$

因此，马尔可夫链以几何速度收敛至目标测度（注意，C 取决于初始样本 $\widetilde{\omega}$的分布。）误差界 (42) 随即可完全按照 [12, 引理 B.2（第 41/42 页）], 中对数正态情形的推导得到，因此此处省略。□

注 5.2。我们指出，后续对 (41) 中马尔可夫链蒙特卡洛估计量以及第 5.2 小节中其多层扩展的误差分析，并不 依赖于将独立采样器用作提议核。我们的结果实际上适用于 (40) 中任何满足类似引理 5.1 所述几何遍历性结果的 $\mathbb{Q}$ 选择。

我们只能对近似的感兴趣量（QoI）$\Psi \circ u_{N',\ h'} \approx \varphi$进行采样，其中$h' > 0$ 和 $N' \in \mathbb{N}$ 是定理 3.10 中的离散化参数，不一定需要与 $\mathbb{E}_h$中的 h 和 N 一致。我们对 Ψ 作出以下假设，以界定由此产生的离散化误差。

假设 5.3。令 $\Psi\colon V \to \mathbb{R}$，并令 $u : \Omega \to V$为方程 (29) 的逐路径弱解。

1.) 令 $\theta_\Psi \in [0,1]$，令 $\Psi\colon H^{\theta_\Psi}(\mathcal{D}) \to \mathbb{R}$ 为 Fr´echet 可微泛函，并记

$$\Psi' : H^{\theta_\Psi}(\mathcal{D}) \to \mathcal{L}(H^{\theta_\Psi}(\mathcal{D});\mathbb{R}) = (H^{\theta_\Psi}(\mathcal{D}))'$$

Ψ 的 Fr´echet 导数。存在常数 $C > 0, \rho_1, \rho_2 \ge 0$，使得对所有 $v \in H^{\theta_\Psi}($ $\mathcal{D})$

$$|\Psi(v)| \le C(1 + \|v\|^{\rho_1}_{H^{\theta_\Psi}(\mathcal{D})}), \quad \|\Psi'(v)\|_{\mathcal{L}(H^{\theta_\Psi}(\mathcal{D});\mathbb{R})} \le C(1 + \|v\|^{\rho_2}_{H^{\theta_\Psi}(\mathcal{D})}). \tag{43}$$

2.) 存在 $u \in L^q(\Omega; V)$，并且存在常数 $t > 0, r \in (0,1],$，使得定理 3.10 对 $q := 6\max(\rho_1, \rho_2 + 1)$ 成立。

注 5.4. 假设 5.3 是自然的，尤其包括有界线性泛函Ψ，其中 $\rho_1 = 1$和 $\rho_2 = 0$。此外，第二部分中的条件 $q = 6\max(\rho_1, \rho_2 + 1) \geq 6$对于界定下文定理 5.9 中的均方误差是必要的。不过，这一限制仅适用于 $p = 1$的情况，因为在定理 3.8 和 3.10 中， q 对于 $p > 1$可以任意大。

我们记录以下结果，以界定近似误差。

定理 5.5. 在假设 *5.3* 下，存在一个常数 $C > 0$，使得对任意 $N' \in \mathbb{N}$ 和$h' \in \mathfrak{H}$ ，都有

$$\|\Psi(u) - \Psi(u_{N',h'})\|_{L^6(\Omega)} \leq C\left(2^{-tN'} + h'^{(2-\theta_\Psi)r}\right) \tag{44}$$

证明. 该结论已在 [28, 定理 5.4]证明的第二部分中得到证明。 □

注 5.6. 估计式 (44) 特别蕴含了以下较弱的界

$$\|\Psi(u) - \Psi(u_{N',h'})\|_{L^2(\Omega)} \leq C\left(2^{-tN'} + h'^{(2-\theta_\Psi)r}\right),$$

这将在下文的推论 5.7 中使用。另一方面，为了界定下一小节定理 5.9 中多层 MCMC 估计量的均方误差，我们需要使用 (44) 中关于$L^6(\Omega)$ 的界。

基于 (44)，为简单起见，我们假设 $2^{-tN'} \simeq h'^{(2-\theta_\Psi)r}$ ，并考虑近似的感兴趣量 $\varphi_{h'} := \Psi(u_{N',h'})$。马尔可夫链蒙特卡洛估计量的总体误差取决于泛函 $\mathcal{O}$ 和 Ψ的正则性。为方便记号，我们引入变量

$$\eta_{\mathcal{O}} := 2 - \theta_{\mathcal{O}} \in [1, \frac{3}{2}), \quad \text{and} \quad \eta_\Psi := 2 - \theta_\Psi \in [1, 2]. \tag{45}$$

推论 5.7. 在假设 *5.3* 下，存在一个 $C > 0$，它与 h 和 h' 无关，使得 at

$$\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_\delta(\varphi) - E_M^h(\varphi_{h'})\right)^2\right) \leq C\left(h^{2\eta_{\mathcal{O}} r} + (h')^{2\eta_\Psi r} + M^{-1}\right).$$

证明. 我们考虑误差分解

$$\begin{aligned}
\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_\delta(\varphi) - E_M^h(\varphi_{h'})\right)^2\right) &\leq 3\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_\delta(\varphi) - \mathbb{E}_h(\varphi)\right)^2\right) \\
&\quad + 3\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_h(\varphi) - \mathbb{E}_h(\varphi_{h'})\right)^2\right) \\
&\quad + 3\mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_h(\varphi_{h'}) - E_M^h(\varphi_{h'})\right)^2\right) \\
&:= 3(I + II + III).
\end{aligned}$$

由于 $\varphi \in L^2(\Omega\ \mathbb{P})$ 和 $2^{-Nt} \simeq h^{\eta_{\mathcal{O}} r}$，根据定理 4.5，项 I 可由以下式子界定

$$I = \left(\mathbb{E}_\delta(\varphi) - \mathbb{E}_h(\varphi)\right)^2 = C(\lambda)\|\varphi\|_{L^2(\Omega,\mathbb{P};\mathcal{X})}(2^{-2Nt} + h^{2\eta_{\mathcal{O}} r}) \leq Ch^{2\eta_{\mathcal{O}} r}.$$

我们利用 $2^{-N't} \simeq (h')^{\eta_\Psi r}$、定理 4.5 和 5.5，通过以下式子界定项 II

$$\begin{aligned}
II &= \left(\mathbb{E}(\varphi - \varphi_{h'}) + \mathbb{E}_h(\varphi - \varphi_{h'}) - \mathbb{E}(\varphi - \varphi_{h'})\right)^2 \\
&\leq 2\mathbb{E}(|\varphi - \varphi_{h'}|^2) + 2\left(\mathbb{E}_h(\varphi - \varphi_{h'}) - \mathbb{E}(\varphi - \varphi_{h'})\right)^2 \\
&\leq C\mathbb{E}(|\varphi - \varphi_{h'}|^2)(1 + h^{2\eta_{\mathcal{O}} r}) \\
&\leq C(h')^{2\eta_\Psi r}.
\end{aligned}$$

根据引理 5.1，结论即可得出，因为项 III 有界。

$$III = \mathcal{E}_{\mathcal{Q}}\left(\left(\mathbb{E}_h(\varphi_{h'}) - E_M^h(\varphi_{h'})\right)^2\right) \leq C\|\varphi_{h'}\|^2_{L^2(\Omega,\mathbb{P})} M^{-1} \leq C\|\varphi\|^2_{L^2(\Omega,\mathbb{P})} M^{-1}. \tag{46}$$

□

5.2 多层马尔可夫链蒙特卡洛

现在令 $L \in \mathbb{N}$，并考虑近似参数的细化序列 $h_0 > \cdots > h_L$ 和 $N_0 < \cdots < N_L$。我们进一步将层级 $\ell = 0, \ldots, L$ 上的近似后验分布记为 $\mathbb{E}_\ell := \mathbb{E}_{\delta, N_\ell, h_\ell}$。给定一个固定的 ℓ，我们选取 $L'(\ell) \in \mathbb{N}$，并将层级 $\ell' = 0, \ldots, L'(\ell)$ 上的近似感兴趣量定义为 $\varphi_{\ell'} := \Psi(u_{N_{\ell'}, h_{\ell'}})$。随后，我们通过望远镜求和来近似 $\mathbb{E}_\delta(\varphi)$：

$$
\begin{aligned}
\mathbb{E}_\delta(\varphi) \approx \mathbb{E}_L(\varphi) &= \sum_{\ell=1}^{L} \mathbb{E}_\ell(\varphi) - \mathbb{E}_{\ell-1}(\varphi) + \mathbb{E}_0(\varphi) \\
&\approx \sum_{\ell=1}^{L} \mathbb{E}_\ell(\varphi_{L'(\ell)}) - \mathbb{E}_{\ell-1}(\varphi_{L'(\ell)}) + \mathbb{E}_0(\varphi_{L'(0)}) \\
&= \sum_{\ell=1}^{L} \sum_{\ell'=1}^{L'(\ell)} \mathbb{E}_\ell(\varphi_{\ell'} - \varphi_{\ell'-1}) - \mathbb{E}_{\ell-1}(\varphi_{\ell'} - \varphi_{\ell'-1}) + \sum_{\ell=1}^{L} \mathbb{E}_\ell(\varphi_0) - \mathbb{E}_{\ell-1}(\varphi_0) \\
&\quad + \sum_{\ell'=1}^{L'(0)} \mathbb{E}_0(\varphi_{\ell'} - \varphi_{\ell'-1}) + \mathbb{E}_0(\varphi_0).
\end{aligned}
\tag{47}
$$

我们进一步引入截断函数

$$
\mathcal{I}_\ell(\omega) := \mathbb{1}_{\{\Phi_\ell(\omega,\delta) \le \Phi_{\ell-1}(\omega,\delta)\}} \in \{0,1\}, \quad \ell = 1, \ldots, L, \tag{48}
$$

其中 $\Phi_\ell(\omega, \delta)$ 是贝叶斯势的 ℓ 级近似，即

$$
\Phi_\ell(\omega, \delta) := -\log\left(\rho(\delta - [\mathcal{O} \circ u_{N_\ell, h_\ell}](\omega))\right). \tag{49}
$$

Foll根据 [12, 第4节和命题 A.1], , 我们可以据此表示式 (47) 中的展开 v ia

$$
\begin{aligned}
&\sum_{\ell=1}^{L} \mathbb{E}_\ell(\varphi_{L'(\ell)}) - \mathbb{E}_{\ell-1}(\varphi_{L'(\ell)}) + \mathbb{E}_0(\varphi_{L'(0)}) \\
=&\sum_{\ell=1}^{L} \sum_{\ell'=0}^{L'(\ell)} \mathbb{E}_\ell(A^{(1)}_{\ell,\ell'}) + \mathbb{E}_{\ell-1}(A^{(2)}_{\ell,\ell'}) + \mathbb{E}_\ell(A^{(3)}_\ell)\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) \\
&+ \sum_{\ell=1}^{L} \sum_{\ell'=0}^{L'(\ell)} \mathbb{E}_{\ell-1}(A^{(5)}_\ell)\mathbb{E}_\ell(A^{(6)}_{\ell,\ell'} + A^{(7)}_{\ell,\ell'}) + \sum_{\ell'=1}^{L'(\ell)} \mathbb{E}_0(\varphi_{\ell'} - \varphi_{\ell'-1}) + \mathbb{E}_0(\varphi_0),
\end{aligned}
\tag{50}
$$

其中，项 $A^{(1)} - A^{(8)}$ 对于 $\varphi_{-1} := 0$ 由下式给出

$$
\begin{aligned}
A^{(1)}_{\ell,\ell'} &:= (1 - \exp(\Phi_\ell(\omega,\delta) - \Phi_{\ell-1}(\omega,\delta)))(\varphi_{\ell'} - \varphi_{\ell'-1})\mathcal{I}_\ell, \\
A^{(2)}_{\ell,\ell'} &:= (\exp(\Phi_{\ell-1}(\omega,\delta) - \Phi_\ell(\omega,\delta)) - 1)(\varphi_{\ell'} - \varphi_{\ell'-1})(1 - \mathcal{I}_\ell), \\
A^{(3)}_{\ell} &:= (\exp(\Phi_\ell(\omega,\delta) - \Phi_{\ell-1}(\omega,\delta)) - 1)\mathcal{I}_\ell, \\
A^{(4)}_{\ell,\ell'} &:= (\varphi_{\ell'} - \varphi_{\ell'-1})\mathcal{I}_\ell, \\
A^{(5)}_{\ell} &:= (1 - \exp(\Phi_{\ell-1}(\omega,\delta) - \Phi_\ell(\omega,\delta)))(1 - \mathcal{I}_\ell), \\
A^{(6)}_{\ell,\ell'} &:= \exp(\Phi_\ell(\omega,\delta) - \Phi_{\ell-1}(\omega,\delta))(\varphi_{\ell'} - \varphi_{\ell'-1})\mathcal{I}_\ell, \\
A^{(7)}_{\ell,\ell'} &:= (\varphi_{\ell'} - \varphi_{\ell'-1})(1 - \mathcal{I}_\ell), \\
A^{(8)}_{\ell,\ell'} &:= \exp(\Phi_{\ell-1}(\omega,\delta) - \Phi_\ell(\omega,\delta))(\varphi_{\ell'} - \varphi_{\ell'-1})(1 - \mathcal{I}_\ell).
\end{aligned}
\tag{51}
$$

现在，我们将 (50) 中的期望 $\mathbb{E}_\ell$ 替换为马尔可夫链蒙特卡洛估计量 $E^\ell_{M_{\ell,\,\ell'}}(\cdot)$，其中样本数 $M_{\ell,\,\ell'}$ 取决于离散化层级 ℓ 和 ℓ'。由此得到多层马尔可夫链蒙特卡罗 (ML-MCMC) 估计量

$$E_L(\varphi) := \sum_{\ell=1}^{L}\sum_{\ell'=0}^{L'(\ell)} E^{\ell}_{M_{\ell,\ell'}}(A^{(1)}_{\ell,\ell'}) + E^{\ell-1}_{M_{\ell,\ell'}}(A^{(2)}_{\ell,\ell'}) \\ + E^{\ell}_{M_{\ell,\ell'}}(A^{(3)}_{\ell})E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) + E^{\ell-1}_{M_{\ell,\ell'}}(A^{(5)}_{\ell})E^{\ell}_{M_{\ell,\ell'}}(A^{(6)}_{\ell,\ell'} + A^{(7)}_{\ell,\ell'}) \\ + \sum_{\ell'=1}^{L'(\ell)} E^{0}_{M_{0,\ell'}}(\varphi_{\ell'} - \varphi_{\ell'-1}) + E^{0}_{M_{0,0}}(\varphi_0). \tag{52}$$

出于技术上的原因，我们要求 ML-MCMC 估计量满足以下假设 r.

假设 5.8。 给定 $L \in \mathbb{N}$，并固定离散化层级 $\ell \in \{0,\dots,L\}$。式 (52) 中的估计量 $(E^{\ell}_{M_{\ell,\ell'}}(A^{(3)}_{\ell})\, E^{\ell}_{M_{\ell,\ell'}}(A^{(6)}_{\ell,\ell'} + A^{(7)}_{\ell,\ell'}))$ 与 $(E^{\ell-1}_{M_{\ell,\ell'}}(A^{(5)}_{\ell})\, E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}))$ 相互独立。此外，式 (52) 中的所有估计量关于 $\ell \in \{0,\dots,L\}$ 均相互独立，这意味着针对每个后验细化层级 $\ell \in \{0,\dots,L\}$ 生成的马尔可夫链彼此独立。

推导定理 5.9 中的均方误差时，假设 5.8 是必要的；否则就必须将引理 5.1 加强到四阶矩。对于固定的 $\ell = 0,\dots,L$，我们用 $\mathcal{P}_\ell$ 表示由后验细化层级为 ℓ 的所有马尔可夫链所生成的概率空间上的概率测度。所有 $L+1$ 测度 $\mathcal{P}_\ell$ 的组合，结合假设 5.8 的第二部分，得到乘积概率测度 $\mathcal{P}^{\mathrm{ML}}_L := \mathcal{P}_0 \otimes \cdots \otimes \mathcal{P}_L$；我们用 $\mathcal{E}^{\mathrm{ML}}_L$ 表示相应的期望。由此，我们可以量化多层马尔可夫链蒙特卡洛算法的均方误差：

定理 5.9。 设假设*5.3*和*5.8*成立，并设 $h_0 \in \mathfrak{H}$ 表示最粗层有限元细化参数。对于任意固定的 $0 < \varepsilon < 1$ 集合

$$L := \left\lceil \frac{-\log_2(\varepsilon)}{\eta_{\mathcal{O}} r} + \log_2(h_0) \right\rceil, \quad L'(\ell) = L' := \left\lceil L\frac{\eta_{\mathcal{O}}}{\eta_\Psi} \right\rceil \quad \textit{for } \ell = 0,\dots,L, \quad \textit{and} \\ h_\ell = h_0 2^{-\ell}, \quad h_{\ell'} = h_0 2^{-\ell'}, \quad N_\ell = \left\lceil -\log_2(h_\ell)\frac{\eta_{\mathcal{O}} r}{t} \right\rceil, \quad N_{\ell'} = \left\lceil -\log_2(h_{\ell'})\frac{\eta_\Psi r}{t} \right\rceil, \\ \textit{for } \ell' = 0,\dots,L'(\ell),\ \ell = 0,\dots,L.$$

此外，将每个层级上的样本数 $M_{\ell,\ell'}$ 设为

$$M_{\ell,\ell'} := \begin{cases} \lceil h_L^{-2r\eta_{\mathcal{O}}} w_{0,0} \rceil, & \textit{for } \ell = \ell' = 0, \\ \lceil h_L^{-2r\eta_{\mathcal{O}}} h_\ell^{2r\eta_{\mathcal{O}}} w_{\ell,0} \rceil, & \textit{for } \ell = 1,\dots,L \textit{ and } \ell' = 0, \\ \lceil h_L^{-2r\eta_{\mathcal{O}}} h_{\ell'}^{2r\eta_\Psi} w_{0,\ell'} \rceil, & \textit{for } \ell = 0 \textit{ and } \ell' = 1,\dots,L'(0), \\ \lceil h_L^{-2r\eta_{\mathcal{O}}} h_\ell^{2r\eta_{\mathcal{O}}} h_{\ell'}^{2r\eta_\Psi} w_{\ell',\ell} \rceil, & \textit{for } \ell = 1,\dots,L \textit{ and } \ell' = 1,\dots,L'(\ell), \end{cases}$$

其中，权重 $w_{\ell,\ell'} > 0$ 的选取满足存在一个 $C_w > 0$，且其与 L 无关，使得

$$w_{0,0}^{-1/2} + \sum_{\ell=1}^{L'(0)} w_{\ell,0}^{-1/2} + \sum_{\ell'=1}^{L'(0)} w_{0,\ell'}^{-1/2} + \sum_{\ell=1}^{L}\sum_{\ell'=1}^{L'(\ell)} w_{\ell,\ell'}^{-1/2} \le C_w < \infty. \tag{53}$$

那么，存在一个与 ε 无关的 $C > 0$，使得

$$\mathcal{E}^{\mathrm{ML}}_L\left(\left(\mathbb{E}_\delta(\varphi) - E_L(\varphi) \right)^2 \right)^{1/2} \le C\varepsilon.$$

我们指出，总能选取满足（53）中一致有界条件的可容许权重 $w_{\ell,\ell'}$。要实现（准）最优计算复杂度，$w_{\ell,\ell'}$ 的适当选取取决于参数 r，$\eta_{\mathcal{O}}$，η_Ψ 和 d，具体见下文定理5.11。

定理 *5.9* 的证明。我们采用误差分解

$$
\begin{aligned}
&\mathcal{E}_L^{\mathrm{ML}}\left(\left(\mathbb{E}_\delta(\varphi)-E_L(\varphi)\right)^2\right)^{1/2}\\
&\quad\le \left|\mathbb{E}_\delta(\varphi)-\mathbb{E}_L(\varphi)\right|\\
&\qquad+\left|\mathbb{E}_L(\varphi)-\left(\sum_{\ell=1}^{L}\mathbb{E}_\ell(\varphi_{L'(\ell)})-\mathbb{E}_{\ell-1}(\varphi_{L'(\ell)})+\mathbb{E}_0(\varphi_{L'(0)})\right)\right|\\
&\qquad+\mathcal{E}_L^{\mathrm{ML}}\left(\left(\sum_{\ell=1}^{L}\mathbb{E}_\ell(\varphi_{L'(\ell)})-\mathbb{E}_{\ell-1}(\varphi_{L'(\ell)})+\mathbb{E}_0(\varphi_{L'(0)})-E_L(\varphi)\right)^2\right)^{1/2}\\
&\quad= I+II+III.
\end{aligned}
$$

利用定理 4.5 以及 L 和 N_L 的取值，第一个项有如下界：

$$
I\le C(2^{-N_L t}+h_L^{\eta\wp r})\le C(h_0 2^{-L})^{\eta\wp r}\le C\varepsilon.
$$

我们对 II 中的第一项进行展开，并利用定理 4.5 和 5.5 得到如下界：

$$
\begin{aligned}
II&=\left|\sum_{\ell=1}^{L}\mathbb{E}_\ell(\varphi)-\mathbb{E}_{\ell-1}(\varphi)-\left(\mathbb{E}_\ell(\varphi_{L'(\ell)})-\mathbb{E}_{\ell-1}(\varphi_{L'(\ell)})\right)+\mathbb{E}_0(\varphi)-\mathbb{E}_0(\varphi_{L'(0)})\right|\\
&\le\sum_{\ell=1}^{L}\left|\mathbb{E}_\ell(\varphi-\varphi_{L'(\ell)})-\mathbb{E}_{\ell-1}(\varphi-\varphi_{L'(\ell)})\right|+\left|\mathbb{E}_0(\varphi)-\mathbb{E}_0(\varphi_{L'(0)})\right|\\
&\le C\sum_{\ell=0}^{L}\|\varphi-\varphi_{L'(\ell)}\|_{L^2(\Omega)}h_\ell^{\eta\wp r}\\
&\le C\sum_{\ell=0}^{L}h_{L'(\ell)}^{\eta_\Psi r}h_\ell^{\eta\wp r}\\
&\le C(h_0 2^{-L})^{\eta\wp r}\sum_{\ell=0}^{L}(h_0 2^{-\ell})^{\eta\wp r}\\
&\le C\varepsilon.
\end{aligned}
$$

为了给 III 取界，我们将使用表示式 (50)，并分别对所有项 $A^{(1)}_{\ell\,\ell'}$、……、$A^{(8)}_{\ell\,\ell'}$ 的估计误差取界。令 $\mathcal{E}_\ell$ 表示关于 $\mathcal{P}_\ell$ 的期望，其中 $\mathcal{P}_\ell$ 是由第 ℓ 层上的马尔可夫链生成的空间上的概率测度。根据泰勒展开、(48)、(9) 和 (49)，我们有

$$
\begin{aligned}
&\left|\left(1-\exp(\Phi_\ell(\omega,\delta)-\Phi_{\ell-1}(\omega,\delta)\right)\mathcal{I}_\ell\right|\\
&\le\left|\Phi_\ell(\omega,\delta)-\Phi_{\ell-1}(\omega,\delta)\right|\\
&=\frac{1}{2}[\mathcal{O}\circ(u_{N_\ell,h_\ell}-u_{N_{\ell-1},h_{\ell-1}})](\omega)^\top\Sigma^{-1}(2\delta+[\mathcal{O}\circ(u_{N_\ell,h_\ell}+u_{N_{\ell-1},h_{\ell-1}})](\omega))\\
&\le C\|u_{N_\ell,h_\ell}(\omega)-u_{N_{\ell-1},h_{\ell-1}}(\omega)\|_{H^{\theta\wp}}\left(1+\|u_{N_\ell,h_\ell}(\omega)+u_{N_{\ell-1},h_{\ell-1}}(\omega)\|_{H^{\theta\wp}}\right)\\
&\le C\left(\|u(\omega)-u_{N_{\ell-1},h_{\ell-1}}(\omega)\|_{H^{\theta\wp}}+\|u(\omega)-u_{N_\ell,h_\ell}(\omega)\|_{H^{\theta\wp}}\right)\\
&\qquad\cdot\left(1+\|u_{N_\ell,h_\ell}(\omega)\|_{H^{\theta\wp}}+\|u_{N_{\ell-1},h_{\ell-1}}(\omega)\|_{H^{\theta\wp}}\right),
\end{aligned}
\tag{54}
$$

wh其中 $C=C(\Sigma,\ \lambda,\ \mathcal{O})$ 与 ℓ 无关。现在先考虑 $\ell'\ge 1$ 的情形。我们得到 by

引理 5.1、式 (54) 中的估计以及 Hölder 不等式

$$
\begin{aligned}
&\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A_{\ell,\ell'}^{(1)})-E_{M_{\ell,\ell'}}^{\ell}(A_{\ell,\ell'}^{(1)})\right)^2\right)^{1/2}\\
&\quad\le CM_{\ell,\ell'}^{-1/2}\|(1-\exp(\Phi_\ell(\cdot,\delta)-\Phi_{\ell-1}(\cdot,\delta)))\mathcal{I}_\ell(\varphi_{\ell'}-\varphi_{\ell'-1})\|_{L^2(\Omega)}\\
&\quad\le CM_{\ell,\ell'}^{-1/2}\||\Phi_\ell(\cdot,\delta)-\Phi_{\ell-1}(\cdot,\delta)||\varphi_{\ell'}-\varphi_{\ell'-1}|\|_{L^2(\Omega)}\\
&\quad\le CM_{\ell,\ell'}^{-1/2}\left(\|u-u_{N_{\ell-1},h_{\ell-1}}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}+\|u-u_{N_\ell,h_\ell}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}\right)\\
&\qquad\cdot\left(1+\|u_{N_\ell,h_\ell}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}+\|u_{N_{\ell-1},h_{\ell-1}}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}\right)\|\varphi-\varphi_{\ell'}\|_{L^6(\Omega)}.
\end{aligned}
$$

与命题 4.3 的证明中一样，接下来我们利用定理 3.10 证明

$$
\begin{aligned}
\|u-u_{N_\ell,h_\ell}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}&\le\|u-u_{N_\ell}\|_{L^6(\Omega;V)}+\|u_{N_\ell}-u_{N_\ell,h_\ell}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}\\
&\le C2^{-N_\ell t}+\|u_{N_\ell}-u_{N_\ell,h_\ell}\|_{L^6(\Omega;V)}^{\theta_{\mathcal{O}}}\|u_{N_\ell}-u_{N_\ell,h_\ell}\|_{L^6(\Omega;H)}^{1-\theta_{\mathcal{O}}}\\
&\le C(2^{-N_\ell t}+h_\ell^{(2-\theta_{\mathcal{O}})r}).
\end{aligned}
\tag{55}
$$

这尤其表明

$$
\|u_{N_\ell,h_\ell}\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}\le\|u_{N_\ell,h_\ell}-u\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}+\|u\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}\le C(1+\|u\|_{L^6(\Omega;H^{\theta_{\mathcal{O}}})}),
\tag{56}
$$

其中 $C=C(N_0,\ h_0)$，且最后一个估计与 N_l 和 h_ℓ 无关。由于根据 N_ℓ 的选取，$h_{\ell-1}=2h_\ell$ 和 $2^{-N_\ell t}=h^{(2-\theta_{\mathcal{O}})_\ell}\qquad^r=h_\ell^{\eta_{\mathcal{O}}r}$，由此可知，估计式 (55)、(56) 以及定理 5.5 表明

$$
\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A_{\ell,\ell'}^{(1)})-E_{M_{\ell,\ell'}}^{\ell}(A_{\ell,\ell'}^{(1)})\right)^2\right)^{1/2}\le CM_{\ell,\ell'}^{-1/2}h_\ell^{\eta_{\mathcal{O}}r}h_{\ell'}^{\eta_\Psi r}.
$$

类似地，对于 $\ell'\ge1$，并且由于 $h_{\ell-1}=2h_\ell$，我们得到

$$
\mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(2)})-E_{M_{\ell,\ell'}}^{\ell-1}(A_{\ell,\ell'}^{(2)})\right)^2\right)^{1/2}\le CM_{\ell,\ell'}^{-1/2}h_\ell^{\eta_{\mathcal{O}}r}h_{\ell'}^{\eta_\Psi r}.
$$

为处 理 (50) 中第三项对应的误差，我们使用三角不等式得到

$$
\begin{aligned}
&\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A_\ell^{(3)})\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})-E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})E_{M_{\ell,\ell'}}^{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})\right)^2\right)^{1/2}\\
&\le\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A_\ell^{(3)})-E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})\right)^2\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})^2\right)^{1/2}\\
&\quad+\mathcal{E}_\ell\left(E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})^2\left(\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})-E_{M_{\ell,\ell'}}^{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})\right)^2\right)^{1/2}\\
&\le|\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})|\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A_\ell^{(3)})-E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})\right)^2\right)^{1/2}\\
&\quad+\sqrt{2}|\mathbb{E}_\ell(A_\ell^{(3)})|\mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})-E_{M_{\ell,\ell'}}^{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})\right)^2\right)^{1/2}\\
&\quad+\sqrt{2}\mathcal{E}_\ell\left(\left(E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})-\mathbb{E}_\ell(A_\ell^{(3)})\right)^2\right)^{1/2}\\
&\qquad\cdot\mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})-E_{M_{\ell,\ell'}}^{\ell-1}(A_{\ell,\ell'}^{(4)}+A_{\ell,\ell'}^{(8)})\right)^2\right)^{1/2}.
\end{aligned}
\tag{57}
$$

最后一行由关于 $c_1,c_2\in\mathbb{R}$ 的基本不等式 $(c_1+c_2)^2\le2(c_1^2+c_2^2)$ 得出，同时利用假设 5.8 中关于 $E_{M_{\ell,\ell'}}^{\ell}(A_\ell^{(3)})$ 和 $E_{M_\ell}^{\ell-1}{}_{,\ell'}(A_\ell^{(4)}{}_{,\ell'}+A_{\ell\,\ell'}^{(8)})$ 相对于

测度 $\mathcal{P}_\ell$ 的独立性。定理 4.5 和 5.5 对于满足 $|\mathcal{I}_\ell| \le 1$ 的 $\ell' \ge 1$ 给出

$$|\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'})| \le \|A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}\|_{L^2(\Omega)} \le C h_{\ell'}^{\eta_\Psi r}. \tag{58}$$

关于界 (55)，由 (54)、H¨older 不等式和定理 3.10 可得

$$\begin{aligned}
|\mathbb{E}_\ell(A^{(3)}_\ell)| \le \|A^{(3)}_\ell\|_{L^2(\Omega)} &\le \mathbb{E}_\ell\left(|(\exp(\Phi_\ell(\cdot,\delta) - \Phi_{\ell-1}(\cdot,\delta)) - 1)\mathcal{I}_\ell|^2\right)^{1/2} \\
&\le C\left(\|u - u_{N_{\ell-1},h_{\ell-1}}\|_{L^4(\Omega,\mathbb{P};H^{\theta\mathcal{O}})} + \|u - u_{N_\ell,h_\ell}\|_{L^4(\Omega,\mathbb{P};H^{\theta\mathcal{O}})}\right) \\
&\quad \cdot (1 + \|u\|_{L^4(\Omega,\mathbb{P};H^{\theta\mathcal{O}})}) \\
&\le C h_\ell^{\eta_\mathcal{O} r}.
\end{aligned} \tag{59}$$

将 (58) 和 (59) 代入 (57)，即可得到

$$\begin{aligned}
&\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A^{(3)}_\ell)\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) - E^\ell_{M_{\ell,\ell'}}(A^{(3)}_\ell)E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'})\right)^2\right)^{1/2} \\
&\le C h_{\ell'}^{\eta_\Psi r}\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A^{(3)}_\ell) - E^\ell_{M_{\ell,\ell'}}(A^{(3)}_\ell)\right)^2\right)^{1/2} \\
&\quad + C h_\ell^{\eta_\mathcal{O} r}\mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) - E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'})\right)^2\right)^{1/2} \\
&\quad + \sqrt{2}\mathcal{E}_\ell\left(\left(E^\ell_{M_{\ell,\ell'}}(A^{(3)}_\ell) - \mathbb{E}_\ell(A^{(3)}_\ell)\right)^2\right)^{1/2} \\
&\qquad \cdot \mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) - E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'})\right)^2\right)^{1/2}.
\end{aligned}$$

接着，我们再次使用 (58)、(59) 以及与 $A^{(1)}_{\ell\,\ell'}$,的界相同的论证，得到

$$\begin{aligned}
&\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A^{(3)}_\ell)\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}) - E^\ell_{M_{\ell,\ell'}}(A^{(3)}_\ell)E^{\ell-1}_{M_{\ell,\ell'}}(A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'})\right)^2\right)^{1/2} \\
&\le C\left(h_{\ell'}^{\eta_\Psi r}\frac{\|A^{(3)}_\ell\|_{L^2(\Omega)}}{M^{1/2}_{\ell,\ell'}} + h_\ell^{\eta_\mathcal{O} r}\frac{\|A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}\|_{L^2(\Omega)}}{M^{1/2}_{\ell,\ell'}} + \frac{\|A^{(3)}_\ell\|_{L^2(\Omega)}\|A^{(4)}_{\ell,\ell'} + A^{(8)}_{\ell,\ell'}\|_{L^2(\Omega)}}{M_{\ell,\ell'}}\right) \\
&\le C\left(\frac{h_\ell^{\eta_\mathcal{O} r}h_{\ell'}^{\eta_\Psi r}}{M^{1/2}_{\ell,\ell'}} + \frac{h_\ell^{\eta_\mathcal{O} r}h_{\ell'}^{\eta_\Psi r}}{M_{\ell,\ell'}}\right) \\
&\le C M^{-1/2}_{\ell,\ell'}h_\ell^{\eta_\mathcal{O} r}h_{\ell'}^{\eta_\Psi r},
\end{aligned}$$

对 $\ell' \ge 1$ 成立，其中最后一行是因为 $M_{\ell,\ell'} \ge 1$。类似地，我们可推出

$$\mathcal{E}_\ell\left(\left(\mathbb{E}_{\ell-1}(A^{(5)}_\ell)\mathbb{E}_\ell(A^{(6)}_{\ell,\ell'} + A^{(7)}_{\ell,\ell'}) - E^{\ell-1}_{M_{\ell,\ell'}}(A^{(5)}_\ell)E^\ell_{M_{\ell,\ell'}}(A^{(6)}_{\ell,\ell'} + A^{(7)}_{\ell,\ell'})\right)^2\right)^{1/2} \le C M^{-1/2}_{\ell,\ell'}h_\ell^{\eta_\mathcal{O} r}h_{\ell'}^{\eta_\Psi r}.$$

对于 $\ell' = 0$ 这种情况，由于 $\varphi_{\ell'-1} = 0$ 按定义成立，我们在关于 III 的所有估计中都会失去因子 $h_{\ell'}^{\eta_\Psi r}$。重复前面的论证即可得到 $\ell = 1, \dots, L$

$$\begin{aligned}
\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A^{(1)}_{\ell,0}) - E^\ell_{M_{\ell,0}}(A^{(1)}_{\ell,0})\right)^2\right)^{1/2} &\le C M^{-1/2}_{\ell,0}h_\ell^{\eta_\mathcal{O} r}, \\
\mathcal{E}_\ell\left(\left(\mathbb{E}_\ell(A^{(2)}_{\ell,0}) - E^\ell_{M_{\ell,0}}(A^{(2)}_{\ell,0})\right)^2\right)^{1/2} &\le C M^{-1/2}_{\ell,0}h_\ell^{\eta_\mathcal{O} r}, \\
\mathcal{E}_\ell\left(\mathbb{E}_\ell(A^{(3)}_\ell)\mathbb{E}_{\ell-1}(A^{(4)}_{\ell,0} + A^{(8)}_{\ell,0}) - E^\ell_{M_{\ell,0}}(A^{(3)}_\ell)E^{\ell-1}_{M_{\ell,0}}(A^{(4)}_{\ell,0} + A^{(8)}_{\ell,0})\right) &\le C M^{-1/2}_{\ell,0}h_\ell^{\eta_\mathcal{O} r}, \\
\mathcal{E}_\ell\left(\mathbb{E}_{\ell-1}(A^{(5)}_\ell)\mathbb{E}_\ell(A^{(4)}_{\ell,0} + A^{(8)}_{\ell,0}) - E^{\ell-1}_{M_{\ell,0}}(A^{(5)}_\ell)E^\ell_{M_{\ell,0}}(A^{(4)}_{\ell,0} + A^{(8)}_{\ell,0})\right) &\le C M^{-1/2}_{\ell,0}h_\ell^{\eta_\mathcal{O} r}.
\end{aligned}$$

最后，我们还有

$$\mathcal{E}_\ell\left(\left(\mathbb{E}_0(\varphi_{\ell'}-\varphi_{\ell'-1})-E^0_{M_{0,\ell'}}(\varphi_{\ell'}-\varphi_{\ell'-1})\right)^2\right)^{1/2}\le CM_{0,\ell'}^{-1/2}h_{\ell'}^{\eta_\Psi r},$$

以及

$$\mathcal{E}_\ell\left(\left(\mathbb{E}_0(\varphi_0)-E^0_{M_{0,0}}(\varphi_0)\right)^2\right)^{1/2}\le CM_{0,0}^{-1/2}.$$

现在我们汇总所有估计，并利用 (50) 和 (53) 将 III 界定为

$$\begin{aligned}III&\le C\left(M_{0,0}^{-1/2}+\sum_{\ell'=1}^{L'(0)}M_{0,\ell'}^{-1/2}h_{\ell'}^{\eta_\Psi r}+\sum_{\ell=1}^{L}M_{\ell,0}^{-1/2}h_\ell^{\eta_\mathcal{O} r}+\sum_{\ell=0}^{L}\sum_{\ell'=1}^{L'(\ell)}M_{\ell,\ell'}^{-1/2}h_\ell^{\eta_\mathcal{O} r}h_{\ell'}^{\eta_\Psi r}\right)\\&\le Ch_L^{\eta_\mathcal{O} r}\left(w_{0,0}^{-1/2}+\sum_{\ell'=1}^{L'(0)}w_{0,\ell'}^{-1/2}+\sum_{\ell=1}^{L}w_{\ell,0}^{-1/2}+\sum_{\ell=1}^{L}\sum_{\ell'=1}^{L'(\ell)}w_{\ell,\ell'}^{-1/2}\right)\\&\le C\varepsilon.\end{aligned}$$

□

为了推导多层马尔可夫链蒙特卡罗估计量的复杂度估计，我们还需要对采样成本作出另一个假设。

假设 5.10. 生成一个 $\Phi_\ell(\cdot;\,\delta)\;=\;-\log(\rho(\delta-u_{N_\ell,h_\ell}))$ 和 $\varphi_\ell\;=\;\Psi(u_{N_\ell,h_\ell})$ 样本，其中 $u_{N_\ell,h_\ell}\in V_{h_\ell}$ 和 $n_\ell:=\dim(V_{h_\ell})\;=\mathcal{O}(h_\ell^{-d})$，需要 $\mathcal{O}(n_\ell)$ 的计算工作量和内存。

定理5.11. 设假设 *5.3*、*5.8* 和 *5.10* 成立。对于任意给定的 $\varepsilon>0$，存在多层马尔可夫链蒙特卡洛参数 $L, L', h_\ell, N_\ell, M_{\ell,\ell'}$，使得多层马尔可夫链蒙特卡洛估计量满足

$$\mathcal{E}_L^{\mathrm{ML}}\left(\left(\mathbb{E}_\delta(\varphi)-E_L(\varphi)\right)^2\right)^{1/2}\le C\varepsilon,\tag{60}$$

计算成本为 $\mathcal{C}_{MLMC}$，其 $\varepsilon\to0$ 的阶为

$$\mathcal{C}_{MLMC}=\begin{cases}\mathcal{O}(\varepsilon^{-2}) & \text{if } 2r\min(\eta_\mathcal{O},\eta_\Psi)>d,\\ \mathcal{O}(\varepsilon^{-2}|\log_2(\varepsilon)|^3) & \text{if } 2r\max(\eta_\mathcal{O},\eta_\Psi)>d \text{ and } 2r\min(\eta_\mathcal{O},\eta_\Psi)=d,\\ \mathcal{O}(\varepsilon^{-2}|\log_2(\varepsilon)|^5) & \text{if } 2r\eta_\mathcal{O}=2r\eta_\Psi=d,\\ \mathcal{O}(\varepsilon^{-d/(r\min(\eta_\mathcal{O},\eta_\Psi))}) & \text{if } 2r\max(\eta_\mathcal{O},\eta_\Psi)=d \text{ and } 2r\min(\eta_\mathcal{O},\eta_\Psi)<d,\\ \mathcal{O}(\varepsilon^{-d/(r\min(\eta_\mathcal{O},\eta_\Psi))-\epsilon}) & \text{if } 2r\max(\eta_\mathcal{O},\eta_\Psi)<d.\end{cases}\tag{61}$$

关于 $2r\max(\eta_\mathcal{O},\eta_\Psi)<d$ 的最后一个复杂度估计对任意 $\epsilon>0$均成立。

注 5.12。 </style} C_{MLMC} 的前三个估计要求 $d\in\{1,2\}$，因为 $r\in(0,1]$,$\eta_\mathcal{O}\in[1,\frac{3}{2})$ 且 $\eta_\Psi\in[1,2]$。此外，对于常用参数集 $r=\eta_\mathcal{O}=\eta_\Psi=1$，若 $d\ge2$，我们可以（基本上）恢复阶为 $\mathcal{O}(\varepsilon^{-d})$ 的渐近复杂度，这对应于空间分辨率为 ε 的单个样本的计算成本。

*Pro*定理*5.11*的对于给定的 $\varepsilon>0$，我们将多层马尔可夫链蒙特卡罗参数设为 $L,\ L',\ h_\ell,\ N_\ell,\ M_{\ell,\ \ell'}$ as

在定理 5.9 中。权重 $w_{\ell,\ \ell'}$ 在 $M_{\ell,\ \ell'}$ 中由 $w_{\ell,\ \ell'} = w_\ell w_{\ell'}$ 给出，其中我们选择

$$w_\ell := \begin{cases} (\ell+1)^{\alpha_1} & \text{if } 2\eta_{\mathcal{O}} r > d, \\ 1 + L^2 \mathbb{1}_{\{\ell>0\}} & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r \geq d, \\ 2^{\alpha_2 \ell} & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r < d, \\ 2^{(d-2\eta_{\mathcal{O}} r)(L-\ell)\alpha_3} & \text{if } 2\eta_{\mathcal{O}} r < d, \end{cases} \quad \text{and}$$

$$w_{\ell'} := \begin{cases} (\ell'+1)^{\alpha_1} & \text{if } 2\eta_\Psi r > d, \\ 1 + (L')^2 \mathbb{1}_{\{\ell>0\}} & \text{if } 2\eta_\Psi r = d \text{ and } 2\eta_{\mathcal{O}} r \geq d, \\ 2^{\alpha_2 \ell'} & \text{if } 2\eta_\Psi r = d \text{ and } 2\eta_{\mathcal{O}} r < d, \\ 2^{(d-2\eta_\Psi r)(L-\ell')\alpha_3} & \text{if } 2\eta_\Psi r < d, \end{cases}$$

参数 $\alpha_1 > 2$, $\alpha_2 \in (0,\ d)$ 和 $\alpha_3 \in (0,\ 1)$ 将在下文进一步指定。在每种情形下，$w_{\ell,\ \ell'}$ 的选取均满足 (53)，且常数 $C_w > 0$ 对 L 一致，因此我们有

$$\mathcal{E}_L^{\mathrm{ML}} \left((\mathbb{E}_\delta(\varphi) - E_L(\varphi))^2 \right)^{1/2} \leq C\varepsilon,$$

根据定理 5.9，接下来只需界定计算复杂度。

在假设 5.10 下，且由于采样 $E_L(\varphi)$ 的代价有如下界

$$\begin{aligned} \mathcal{C}_{MLMC} &\leq C \left(M_{0,0} h_0^{-d} + \sum_{\ell=1}^{L} M_{\ell,0} h_\ell^{-d} + \sum_{\ell'=1}^{L'(0)} M_{0,\ell'} h_{\ell'}^{-d} + \sum_{\ell=1}^{L} \sum_{\ell'=1}^{L'(\ell)} M_{\ell,\ell'} (h_\ell^{-d} + h_{\ell'}^{-d}) \right) \\ &\leq C\varepsilon^{-2} \left(h_0^{-d} w_{0,0} + h_0^{2\eta_{\mathcal{O}} r - d} \sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} w_\ell + h_0^{2\eta_\Psi r - d} \sum_{\ell'=1}^{L'(0)} 2^{(d-2\eta_\Psi r)\ell'} w_{\ell'} \right) \\ &\quad + C\varepsilon^{-2} h_0^{2(\eta_{\mathcal{O}}+\eta_\Psi)r-d} \sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} w_\ell \sum_{\ell'=1}^{L'(\ell)} 2^{-2\eta_\Psi r \ell'} w_{\ell'} \\ &\quad + C\varepsilon^{-2} h_0^{2(\eta_{\mathcal{O}}+\eta_\Psi)r-d} \sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r \ell} w_\ell \sum_{\ell'=1}^{L'(\ell)} 2^{(d-2\eta_\Psi r)\ell'} w_{\ell'} \\ &\leq C\varepsilon^{-2} \Bigg(1 + \sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} w_\ell \left(1 + \sum_{\ell'=1}^{L} 2^{-2\eta_\Psi r \ell'} w_{\ell'} \right) \\ &\qquad + \sum_{\ell'=1}^{L'} 2^{(d-2\eta_\Psi r)\ell'} w_{\ell'} \left(1 + \sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r \ell} w_\ell \right) \Bigg). \end{aligned} \tag{62}$$

这里，最后一行成立是因为 $L'(\ell) = L'$ 与 ℓ 无关。因此，关于 ℓ 的第一个求和有如下界

$$\sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} w_\ell = \begin{cases} \sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} (\ell+1)^{\alpha_1} & \leq C_1 & \text{if } 2\eta_{\mathcal{O}} r > d, \\ \sum_{\ell=1}^{L} 1 + L^2 & \leq C_2 L^3 & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r \geq d, \\ \sum_{\ell=1}^{L} 2^{\alpha_2 \ell} & \leq C_3 2^{\alpha_2 L} & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r < d, \\ \sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)(\alpha_3 L + (1-\alpha_3)\ell)} & \leq C_4 2^{(d-2\eta_{\mathcal{O}} r)L} & \text{if } 2\eta_{\mathcal{O}} r < d. \end{cases}$$

常数 C_1, C_2, C_3, $C_4 \in (0,\ \infty)$ 与 L 无关，因此也与 ε 无关。由于 $L \leq C|\text{对数}_2(\varepsilon)|$ 且 $\varepsilon \leq C 2^{-L\eta_{\mathcal{O}} r}$，存在一个与 ε 无关的 $C > 0$，使得

$$\sum_{\ell=1}^{L} 2^{(d-2\eta_{\mathcal{O}} r)\ell} w_\ell \leq \begin{cases} C & \text{if } 2\eta_{\mathcal{O}} r > d, \\ C|\log_2(\varepsilon)|^3 & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r \geq d, \\ C\varepsilon^{-\alpha_2/(\eta_{\mathcal{O}} r)} & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r < d, \\ C\varepsilon^{2-d/(\eta_{\mathcal{O}} r)} & \text{if } 2\eta_{\mathcal{O}} r < d. \end{cases} \tag{63}$$

类似地，我们由 $\varepsilon \simeq 2^{-L\eta_{\mathcal{O}} r} \simeq 2^{-L'\eta_\Psi r}$ 得出，存在一个与 ε 无关的 $C>0$，使得

$$\sum_{\ell'=1}^{L'(\ell)} 2^{(d-2\eta_\Psi r)\ell'} w_{\ell'} \leq \begin{cases} C & \text{if } 2\eta_\Psi r > d, \\ C|\log_2(\varepsilon)|^3 & \text{if } 2\eta_\Psi r = d \text{ and } 2\eta_{\mathcal{O}} r \geq d, \\ C\varepsilon^{-\alpha_2/(\eta_\Psi r)} & \text{if } 2\eta_\Psi r = d \text{ and } 2\eta_{\mathcal{O}} r < d, \\ C\varepsilon^{2-d/(\eta_\Psi r)} & \text{if } 2\eta_\Psi r < d. \end{cases} \tag{64}$$

每当 $2\eta_{\mathcal{O}} r = d$ 和 $2\eta_\Psi r \geq d$时，根据 $L \leq C|\log_2(\varepsilon)|$，有

$$\sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r\ell} w_\ell \leq L^2 \sum_{\ell=1}^{L} 2^{-d\ell} < C|\log_2(\varepsilon)|^2.$$

接下来，若 $2\eta_{\mathcal{O}} r = d$ 和 $2\eta_\Psi r < d$ 成立，则根据 $\alpha_2 < d = 2\eta_{\mathcal{O}} r$ 有

$$\sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r\ell} w_\ell = \sum_{\ell=1}^{L} 2^{(-d+\alpha_2)\ell} < \infty.$$

此外，若 $2\eta_{\mathcal{O}} r < d$ 成立，则有 $w_\ell \leq 2^{(d-2\eta_{\mathcal{O}} r)\alpha_3 L}$，因此

$$\sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r\ell} w_\ell \leq 2^{(d-2\eta_{\mathcal{O}} r)\alpha_3 L} \sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r\ell} \leq C\varepsilon^{\alpha_3(2-d/(\eta_{\mathcal{O}} r))}.$$

综上可知，存在 $C>0$，使得对所有 $L \geq 1$

$$\sum_{\ell=1}^{L} 2^{-2\eta_{\mathcal{O}} r\ell} w_\ell \leq \begin{cases} C|\log_2(\varepsilon)|^2 & \text{if } 2\eta_{\mathcal{O}} r = d \text{ and } 2\eta_\Psi r \geq d, \\ C\varepsilon^{\alpha_3(2-d/(\eta_{\mathcal{O}} r))} & \text{if } 2\eta_{\mathcal{O}} r < d, \\ C & \text{otherwise.} \end{cases} \tag{65}$$

由于 $\alpha_2 < d$ 对 $2\eta_{\mathcal{O}} r < d$ 和 $2\eta_\Psi r = d$也成立，因此可以类似地得出

$$\sum_{\ell'=1}^{L'(\ell)} 2^{-2\eta_\Psi r\ell'} w_{\ell'} \leq \begin{cases} C|\log_2(\varepsilon)|^2 & \text{if } 2\eta_\Psi r = d \text{ and } 2\eta_{\mathcal{O}} r \geq d, \\ C\varepsilon^{\alpha_3(2-d/(\eta_\Psi r))} & \text{if } 2\eta_\Psi r < d, \\ C & \text{otherwise.} \end{cases} \tag{66}$$

将 (63)– (66) 中的估计与 (62) 结合，便可立即得到 (61) 中关于 C_{MLMC} 的前三个界。

现在我们考虑 $2\eta_{\mathcal{O}} r = d$ 和 $d > 2\eta_\Psi r$的情形，其中我们选择

$$\alpha_2 \in \left(0, \frac{d}{2}\left(\frac{d}{2\eta_\Psi r} - 1\right)\right] \cap (0, d)$$

并且令 $\alpha_3 = \frac{1}{2}$，结合 $\frac{d}{\eta_\Psi r} > 2$ 和 $d = 2\eta_{\mathcal{O}} r$ 可得

$$\begin{aligned} C_{MLMC} &\leq C\varepsilon^{-2}\left(1 + \varepsilon^{-\alpha_2/(\eta_{\mathcal{O}} r)}(1 + \varepsilon^{\alpha_3(2-d/(\eta_\Psi r))}) + \varepsilon^{2-d/(\eta_\Psi r)}(1+C)\right) \\ &\leq C\left(\varepsilon^{-2-\alpha_2/(\eta_{\mathcal{O}} r)+\alpha_3(2-d/(\eta_\Psi r))} + \varepsilon^{-d/(\eta_\Psi r)}\right) \\ &\leq C\left(\varepsilon^{-\alpha_2 2/d-1-d/(2\eta_\Psi r)} + \varepsilon^{-d/(\eta_\Psi r)}\right) \\ &\leq C\varepsilon^{-d/(\eta_\Psi r)}. \end{aligned}$$

类似地，如果 $2\eta_{\mathcal{O}} r < d$ 和 $2\eta_\Psi r = d$，我们令

$$\alpha_2 \in \left(0, \frac{d}{2}\left(\frac{d}{2\eta_{\mathcal{O}} r} - 1\right)\right] \cap (0, d),$$

并令 $\alpha_3 = -\frac{1}{2}$，得到 $C_{MLMC} \le C\varepsilon^{-d/(\eta_{\mathcal{O}} r)}$。

在满足 $\max(2\eta_{\mathcal{O}} r, 2\eta_{\mathcal{O}} r) < d$ 的最后一种情况下，我们有

$$\begin{aligned} C_{MLMC} &\le C\varepsilon^{-2}\left(1+\varepsilon^{2-d/(\eta_{\mathcal{O}} r)}(1+\varepsilon^{\alpha_3(2-d/(\eta_\Psi r))}) + \varepsilon^{2-d/(\eta_\Psi r)}(1+\varepsilon^{\alpha_3(2-d/(\eta_{\mathcal{O}} r))})\right) \\ &\le C\left(\varepsilon^{\alpha_3(2-d/(r\min(\eta_{\mathcal{O}},\eta_\Psi))}\right)\varepsilon^{-d/(r\min(\eta_{\mathcal{O}},\eta_\Psi))} \\ &\le C\varepsilon^{-d/(\min(\eta_{\mathcal{O}},\eta_\Psi)r)-\epsilon}, \end{aligned}$$

其中，$\epsilon := -\alpha_3(2 - d/(r\min(\eta_{\mathcal{O}}, \eta_\Psi)) > 0$ 可通过选择足够小的 $\alpha_3 \in (0,1)$ 变得任意小（不过请注意，式（60）中的 $C = C(\alpha_3)$ 仅在 $\alpha_3 > 0$ 时关于 L 和 ε 一致，但我们得到，对于 $\alpha_3 = 0$，$C(\alpha_3) = \mathcal{O}(L^2) = \mathcal{O}(|\log_2(\varepsilon)|^2)$。 □

6 数值实验

6.1 一维贝叶斯逆问题

令 $\mathbb{T}^1 = [0,1]$ 为一维环面，令 $\mathcal{D} := \mathbb{T}$，并考虑如下椭圆型（正向）问题：对于给定的 $\omega \in \Omega$，求 $u(\omega) : \mathcal{D} \to \mathbb{R}$，使得

$$-\nabla \cdot (\exp(b_T(\omega))\nabla u(\omega)) = 10 \quad \text{in } (0,1), \quad u(\omega) = 0 \quad \text{on } \{0,1\}. \tag{67}$$

式（67）中的对数扩散系数 b_T 是具有参数 $s = \frac{8}{5}$、$p = \frac{5}{3}$ 和小波密度 $\beta = \frac{4}{5}$ 的 Besov 随机树先验。对于贝叶斯逆问题，我们针对给定的 ω 对 $u(\omega)$ 进行一次实现抽样（也称为“真实值”），并考虑参数到观测映射

$$\mathcal{G} : \Omega \to \mathbb{R}^k, \quad \omega \mapsto \big(u(\omega, x_i),\, i = 1, \dots, k\big)^\top, \quad \text{for } 0 < x_1 < \cdots < x_k < 1,$$

其中 $k = 9$ 和 $x_i := 0.1 \cdot i$，对于 $i = 1, \dots, 9$。因此，观测泛函 $\mathcal{O}$ 是一个线性泛函 $\mathcal{O} \in (V')^k$。为生成合成数据，我们在 $[0,1]$上的有限元网格上近似 u，该网格具有 2^{11} 个等距节点；并在 $N = 11$ 个尺度后截断 Besov 随机树先验 b_T，以获得可行的对数扩散 $b_{T,\,N} \approx b_T$（因为根据定理 3.10，所得逐路径误差的阶为 $\mathcal{O}(2^{-11})$）。真实值、相应的u 精细近似以及观测值的图示见图1左图。含噪观测为

$$\delta = \mathcal{G}(\omega) + \vartheta,$$

其中 $\vartheta \sim \mathcal{N}(0,\ \sigma^2 I_9)$，且 $\sigma = 0.1$。

我们的目标是近似 $\mathbb{E}_\delta(\Psi_i(u))$ 的后验期望，其中 $i = 1,\ 2$，且

$$\Psi_1(u) := \left(\int_{\mathcal{D}} \nabla u \cdot \nabla u \, dx\right)^{\frac{1}{2}} \quad \text{and} \quad \Psi_2(u) := \int_{\mathcal{D}} u \, dx$$

分别是 u 的能量范数和空间均值。对于这些感兴趣量，假设 5.3 在 $\theta_{\Psi_1} = 1$、$\theta_{\Psi_2} = 0$ 以及 $\rho_1 = 1$、$\rho_2 = 0$ 两种情况下均成立。我们采用第 5.2 节中的多层马尔可夫链蒙特卡罗估计量，初始有限元网格宽度为 $h_0 := 2^{-3}$，并测试情况 $L \in \{2, \dots, 6\}$。多层马尔可夫链蒙特卡罗参数按照定理 5.9（针对 $t = r = s - \frac{d}{p} = 1$）选取；为简单起见，能量范数和空间均值均采用 $\eta_\Psi = 1$。空间均值采用 $\eta_\Psi = 2$ 时需要较大的参数 α，否则对于 $\ell,\ \ell' \ge 1$，我们得到的基本上是 $M_{\ell,\,\ell'} = 1$。但若 α 较大，计算时间并不会显著减少。由于对所有 $\eta_\Psi \in [1,2]$，$2r \min(\eta_{\mathcal{O}},\ \eta_\Psi) > d$，因此这种简化不会影响阶为 $\mathcal{O}(\varepsilon^{-2})$ 的渐近复杂度。我们选取多层马尔可夫链蒙特卡罗权重 $w_\ell = (\ell+1)^\alpha$，$w_{\ell'} = (\ell'+1)^\alpha$，其中 $\alpha = 3$。根据定理5.11，对于任意 L，这将得到阶为 $\mathcal{O}(2^{-(L+3)})$ 的均方根误差，计算工作量为 $\mathcal{O}(2^{2L})$。

我们采用第4节中的单层蒙特卡罗比率估计量，使用有限元网格宽度 $h_{ref} =$、尺度截断 $N_{ref} = 11$ 和 $M_{ref} = 2^{22}$ 个样本，在我们的

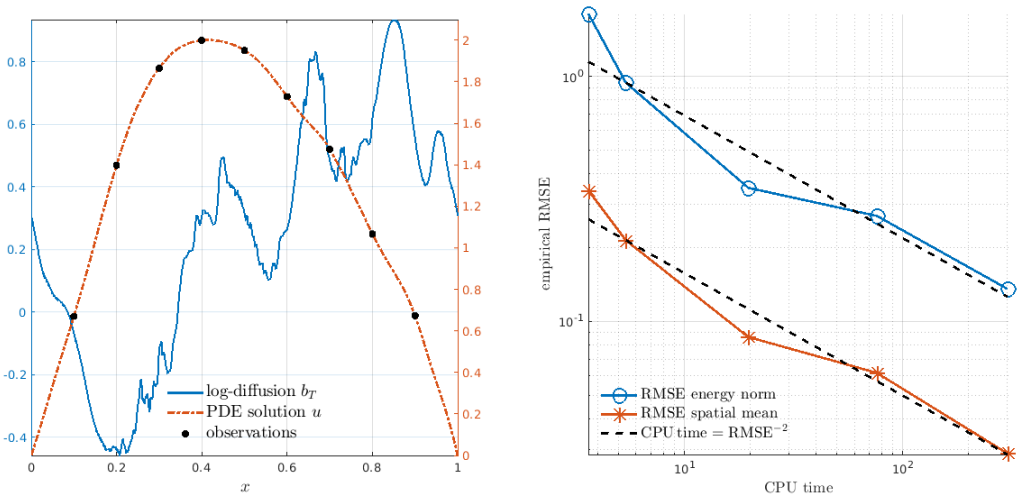

图1：左：合成数据图，即采样得到的 Besov 随机树先验、对应的偏微分方程解 u 以及点观测（黑点）。右：多层马尔可夫链蒙特卡罗估计量的能量范数和空间均值时间-误差图，最大细化层级为 $h_L = 2^{-3-L}$，其中 $L = 2$、……、6。正如预测的那样，误差达到 $\mathcal{O}(\varepsilon)$ 阶，计算复杂度达到 $\mathcal{O}(\varepsilon^{-2})$ 阶。

测试算例中得到参考解。该参考解的误差量级为 $\mathcal{O}(2^{-11})$，因此与具有 $L \leq 6$ 的多层马尔可夫链蒙特卡罗估计量相比可以忽略不计。对于给定的 L，我们对每个多层马尔可夫链蒙特卡罗估计量独立采样 $M_{ML} = 64$ 次实现，并基于参考解计算经验均方根误差。结果显示在图1右侧。显然，在 $\mathcal{O}(\varepsilon^{-2})$ 的计算时间内，经验误差达到了 $\mathcal{O}(\varepsilon)$ 的量级，这证实了我们在第 5.2 节中的理论分析。

6.2 二维贝叶斯逆问题

设 $\mathbb{T}^2 = [0,1]^2$ 为二维环面，令 $\mathcal{D} := \mathbb{T}^2$，并考虑椭圆型（正向）问题：对于给定的 $\omega \in \Omega$，求 $u(\omega) : \mathcal{D} \to \mathbb{R}$，使得

$$-\nabla \cdot (\exp(b_T(\omega))\nabla u(\omega)) = 10 \quad \text{in } (0,1)^2, \quad u(\omega) = 0 \quad \text{on } \partial\mathcal{D}. \tag{68}$$

式（68）中的对数扩散系数 b_T 是一个 Besov 随机树先验，其参数为 $s = \frac{12}{5}$、$p = \frac{5}{3}$，小波密度为 $\beta = \frac{1}{2}$。对于给定的 ω，我们再次对 $u(\omega)$ 进行采样，得到一个实现并将其视为“真实值”，现在考虑参数到观测映射

$$\mathcal{G} : \Omega \to \mathbb{R}^k, \quad \omega \mapsto \left(u(\omega, (x_i, y_j)), i, j = 1, \ldots, \sqrt{k}\right)^\top,$$

其中 $k = 36$，并且观测点为 $x_i,\ y_j \in \{0.1, 0.26, 0.42, 0.58, 0.74, 0.9\}$。这得到一个观测泛函 $\mathcal{O} \notin (V')^k$，而不是 $\mathcal{O} \in ((H_0^{1+\varepsilon}(\mathcal{D}))')^k$，对任意 $\varepsilon > 0$ 都成立。然而，由于 ε 可能任意小，我们在实验中将 $\mathcal{O}$ 视为满足 $\eta_{\mathcal{O}} = 1$。合成数据通过在每个坐标方向上含有 2^{10} 个节点的等距网格上使用双线性有限元生成，并将 Besov 随机树先验 b_T 截断至 $N = 10$ 个尺度。根据定理 3.10，所得逐路径误差的阶为 $\mathcal{O}(2^{-10})$。图2 展示了真实值、u 的相应精细近似以及观测结果。

我们再次近似后验期望 $\mathbb{E}_\delta(\Psi_i(u))$，其中 $i = 1,2$，Ψ_1 和 Ψ_2分别是对应于能量范数和空间均值的泛函。我们采用双线性有限元和初始网格宽度 $h_0 := 2^{-3}$及 $L \in \{2, \ldots, 5\}$，测试第 5.2 节中的多层马尔可夫链蒙特卡罗估计量。多层蒙特卡罗参数按照定理5.11（针对 $t = r = s - \frac{d}{p} = 1$）选取：

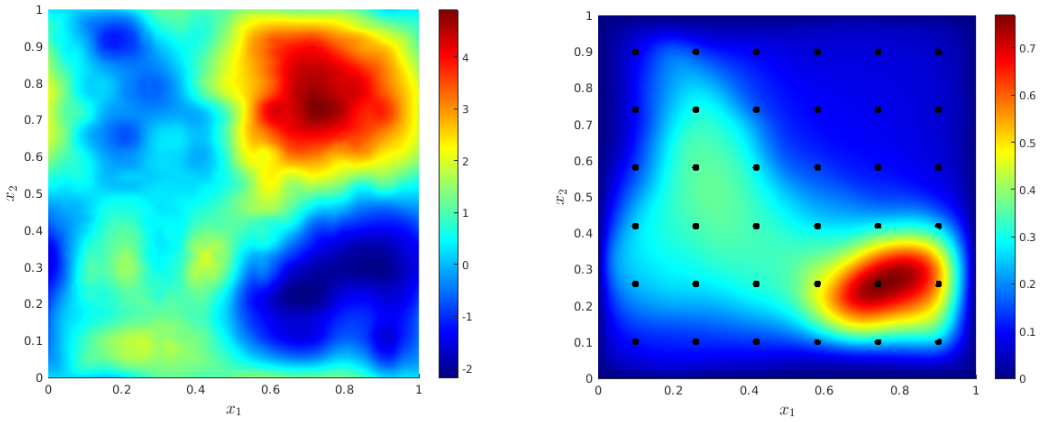

图2：左：合成数据的绘图，即在 $\mathbb{T}^2$ 上采样得到的、具有参数 $s = \frac{12}{5}$、$p = \frac{5}{3}$ 和 $\beta = \frac{1}{2}$ 的 Besov 随机树先验。右：对应的偏微分方程解 u 和点观测（黑点）。

对于能量范数（Ψ_1），有 $\eta_\Psi = 1$，因此 $2r\eta_\mathcal{O} = 2r\eta_\Psi = d$，我们现在将多层马尔可夫链蒙特卡罗权重设为 $w_\ell = w_{\ell'} = 1 + 3 \cdot L^2 \mathbb{1}_{\{\ell>0\}}$。我们发现，对 $w_\ell, w_{\ell'}$ 乘以因子 3（当 $\ell, \ell' > 0$ 时）可以使收敛更加稳定，而这显然不会影响估计量的渐近成本。对于空间均值（Ψ_2），有 $\eta_\Psi = 2$，因此 $2r \min(\eta_\mathcal{O}, \eta_\Psi) = d$且 $2r \max(\eta_\mathcal{O}, \eta_\Psi) > d$，我们利用 Ψ_2 更高的光滑性来降低计算成本。因此，我们将多层马尔可夫链蒙特卡罗权重设为 $w_\ell = 1 + 3 \cdot L^2 \mathbb{1}_{\{\ell>0\}}$ 和 $w_{\ell'} = (\ell' + 1)^6$。

对于空间维数为 $d = 2$ 的示例，在每个离散化层级对第一个马尔可夫链设置预热期后，所有多层马尔可夫链蒙特卡洛估计量都显著减少了 CPU 时间，具体如下：对于固定的离散化层级 ℓ（后验近似的离散化层级）和 $\ell' = 0$，我们丢弃了与该层级对应的最长马尔可夫链前 20% 的样本（ℓ，0）。对于 $\ell' \geq 1$，随后我们使用前一条链在（ℓ，$\ell' - 1$）上的最后一个接受样本，根据层级（ℓ，ℓ'）初始化新链，而不再进行预热阶段。我们对所有 $\ell = 0, \ldots, L$ 重复这一过程，其中第一条链的初始值相对于 ℓ 独立选取。该修改后的估计量尤其满足假设 5.8。为论证我们的预热方法，我们报告了在 $\ell' = 0$ 下不设置预热阶段的相应多层马尔可夫链蒙特卡洛估计量的结果；这些估计量对每一对（ℓ，ℓ'）独立地初始化马尔可夫链。

我们再次使用 [26, 第4节1] 中的单层蒙特卡罗比率估计量，采用有限元网格宽度$h_{ref} = 2^{-9}$、尺度截断 $N_{ref} = 9$和 $M_{ref} = 2^{18}$ 个样本，为测试算例获得参考解。现在参考解的误差阶为 $\mathcal{O}$（2^{-9}），这对于我们的实验似乎仍然足够。对于给定的 L，我们对每个多层马尔可夫链蒙特卡洛估计量独立抽样 $M_{ML} = 64$ 次，并基于参考解计算经验均方根误差。结果见图3和表1。可以清楚地看到，对于能量范数和空间均值，采用预热估计量时，均能在 $\mathcal{O}$（ε^{-2}|对数（ε）|^2）的计算时间内达到阶为$\mathcal{O}$（ε）的经验误差。乍看之下，这有些令人意外，因为根据我们的复杂度分析，依据定理5.11和多层马尔可夫链蒙特卡罗权重的选取，我们预期能量范数的复杂度阶为 $\mathcal{O}$（ε^{-2}|对数（ε）|^5）。

我们还发现，相对于ℓ'，初始预热阶段和顺序初始化显著降低了经验均方根误差，而两种估计量的计算时间相当，见表1。这一效应在空间均值的估计量上尤其明显；若不进行预热，它似乎完全无法收敛。这是因为在这种情况下生成的样本极少（基本上为 $M_{\ell,\ \ell'} = \mathcal{O}(1)$），用于 $\ell' \geq 1$。因此，预热阶段和前一层的初始化对这些短链大有裨益；而若在最粗层不进行预热并采用适当初始化，它们便无法进入渐近区间。

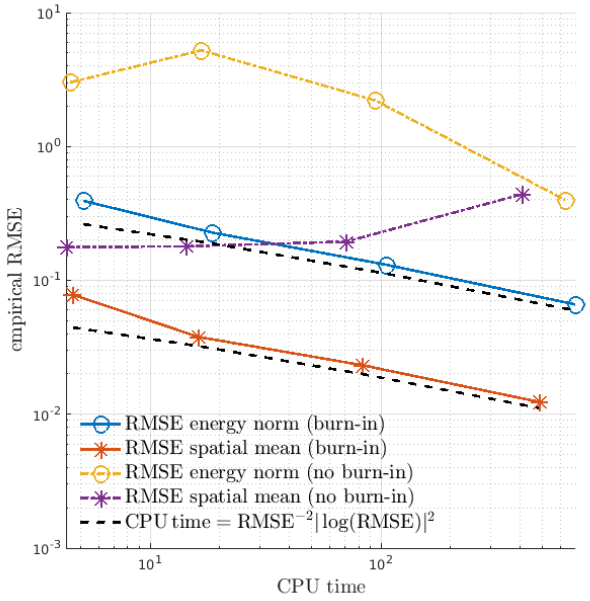

图3：能量范数（蓝色圆圈）和空间均值（橙色星号）的多层马尔可夫链蒙特卡罗估计量的时间-误差图。在最粗层进行预热 $\ell' = 0$ 并采用顺序初始化的估计量能够达到误差 $\mathcal{O}(\varepsilon)$，其计算复杂度为 $\mathcal{O}(\varepsilon^{-2}|\log(\varepsilon)|^2)$ 阶，并且经验误差显著低于未进行预热阶段的对应估计量。

层级 L （最精细分辨率）	2	3	4	5
$\frac{\text{未预热时的均方根误差}}{\text{预热时的均方根误差}}$	7.6692 2.2673	22.9367 4.7504	16.8583 8.3776	6.0254 35.5614
$\frac{\text{预热时的 CPU 时间}}{\text{未预热时的 CPU 时间}}$	1.1372 1.0692	1.1109 1.1341	1.1142 1.0811	1.1045 1.1812

表1：能量范数（第一行）和
空间均值（第二行）的多层马尔可夫链蒙特卡洛估计量的经验均方根误差与 CPU 时间之比。预热和顺序初始化显著降低了均方根误差，额外成本不到 20%。

致谢

AS 的部分经费由 ETH 数据科学基金会倡议（ETH-FDS）资助，谨此致谢。作者感谢 Christoph Schwab 教授在讨论中提出真知灼见，这些讨论促使稿件得到显著改进。

参考解

[1] R. Abraham 和 J.-F. Delmas。《Galton-Watson 树及其局部极限导论》。*arXiv* 预印本 *arXiv:1506.05571*，2015 年。[2]C. D. Aliprantis 和 K. Border。《无限维分析：搭便车指南》。施普林格，2006 年。[3] D. Calvetti 和 E. Somersalo。《贝叶斯科学计算导论：主观计算十讲》，第 2 卷。施普林格科学与商业媒体，2007 年。[4] S. L. Cotter、M. Dashti、J. C. Robinson 和 A. M. Stuart。函数及其在流体力学中的应用的贝叶斯反问题。《反问题》，25(11):115008，2009 年。

[5] S . L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC methods for functi
等：修改旧算法以提高其速度。《统计科学》, 28(3):424–446, 2013.[6] M. Dashti, S. Harris, and A. Stuart. 用于贝叶斯反问题的 Besov 先验。《反问题与成像》, 6(2):183–200, 2012.[7] M. Dashti and A. M. Stuart. 反问题的贝叶斯方法。载于《不确定性量化手册》, pages 311–428. 施普林格, 2017.[8] I. Daubechies. 《小波十讲》. SIAM, 1992.[9] T. J. Dodwell, C. Ketelsen, R. Scheichl, and A. L. Teckentrup. 多层马尔可夫链蒙特卡罗。《*SIAM*评论》, 61(3):509–545, 2019.[10] E. Galetti and A. Curtis. 跨维度电阻率层析成像。《地球物理研究杂志：固体地球》, 123(8):6347–6377, 2018.[11] V. H. Hoang. 测度空间中的贝叶斯反问题及其在白噪声驱动的 Burgers 方程和 Hamilton–Jacobi方程中的应用。《反问题》, 28(2):025009, 2012.[12] V. H. Hoang, J. H. Quek, and C. Schwab. 用于对数正态扩散贝叶斯反演的多层马尔可夫链蒙特卡罗有限元方法分析。《反问题》, 36(3):035021, 2020.[13] V. H. Hoang, J. H. Quek, and C. Schwab. 高斯先验下抛物型偏微分方程贝叶斯反演的多层马尔可夫链蒙特卡罗。*SIAM/ASA* 不确定性量化杂志, 9(2):384–419, 2021.[14] V. H. Hoang, C. Schwab, and A. M. Stuart. 贝叶斯反演加速马尔可夫链蒙特卡罗方法的复杂度分析。《反问题》, 29(8):085010, 2013.[15] B. Hosseini and N. Nigam. 适定的贝叶斯反问题：指数尾先验。*SIAM/ASA* 不确定性量化杂志, 5(1):436–465, 2017.[16] J. Kaipio and E. Somersalo. 《统计与计算反问题》, 第160卷. 施普林格科学与商业媒体, 2006.[17] H. Kekkonen, M. Lassas, E. Saksman, and S. Siltanen. 随机树Besov先验——迈向分形成像。*arXiv*预印本 *arXiv:2103.00574*, 2021.[18] J. Latz. 关于贝叶斯反问题的适定性。*SIAM/ASA* 不确定性量化杂志, 8(1):451–482, 2020.[19] J. Latz, J. P. Madrigal-Cianci, F. Nobile, and R. Tempone. 贝叶斯反问题上的广义并行退火。《统计与计算》, 31(5):1–26, 2021.[20] J. Latz, I. Papaioannou, and E. Ullmann. 用于贝叶斯反问题的多层序贯^2 蒙特卡罗。《计算物理学杂志》, 368:154–178, 2018.[21] J. P. Madrigal-Cianci, F. Nobile, and R. Tempone. 基于独立Metropolis-Hastings算法的一类多层马尔可夫链蒙特卡罗算法分析。*arXiv*预印本 *arXiv:2105.02035*, 2021.[22] F. Monard, R. Nickl, and G. P. Paternain. 高斯过程先验下非线性反问题贝叶斯不确定性量化的统计保证。《统计年刊》, 49(6):3255–3298, 2021.[23] A. R¨orich, T. A. Werthmann, D. G¨oddeke, and L. Grasedyck. 使用低秩张量格式进行肌电图贝叶斯反演。《反问题》, 37(5):055003, 2021.

[24]D. Rudolf and B. Sprungk. 关于预条件 Crank–Nicolson Metropolis 算法的一种推广。计算数学基础, 18(2):309–343, 2018.[25] E. Saksman, M. Lassas, and S. Siltanen. 离散化不变的贝叶斯反演与 Besov 空间先验。《反问题与成像》, 3(1):87–122, 2009.[26] R. Scheichl, A. M. Stuart, and A. L. Teckentrup. 用于计算椭圆型反问题后验期望的准蒙特卡罗与多层蒙特卡罗方法。*SIAM/ASA* 不确定性量化杂志, 5(1):493–518, 2017.[27] C. Schillings and C. Schwab. 用于贝叶斯反问题的稀疏自适应 Smolyak 求积。《反问题》, 29(6):065011, 2013.[28] C. Schwab and A. Stein. 采用 Besov 随机树先验求解椭圆型偏微分方程的多层蒙特卡罗有限元方法。*SAM* 报告, 2022.[29] R. L. Smith and L. Tierney. 独立 Metropolis 采样器的精确转移概率。预印本, 1996.[30] A. M. Stuart. 反问题：贝叶斯视角。《数值学报》, 19:451–559, 2010.[31] H. Triebel. 函数空间理论 *II*. Modern Birkh¨auser Classics. Birkh¨auser, 第2版, 2000.[32] H. Triebel. 区域上的函数空间与小波, 数学丛书第7卷。*EMS*数学论丛. 欧洲数学学会（EMS）, Z¨urich, 2008.[33]H. Triebel. 函数空间理论 *IV*, 数学专著丛书第107卷。Birkh¨auser, 2020.