arXiv:2602.04823v1 [math.ST] 4 Feb 2026

球面上 Sobolev 型能量泛函的自适应估计

Claudio Durastanti罗马第一大学工程基础与应用
科学系claudio.durastanti@uniroma1.it

2026年2月5日

摘要

我们研究单位球面上未知密度的二次Sobolev型积分泛函的估计。该泛函通过拉普拉斯–贝尔特拉米算子的分数次幂定义，可提供光滑度和谱能量的全局度量。我们的方法依赖球面 needlet 框架，它在球面上自然的平方可积函数空间中实现局部化的多尺度分解，同时保持紧框架性质。

我们构造了该泛函适当截断版本的无偏估计量，并通过明确的偏差–方差分析推导出精确的预言机风险界。当密度的光滑度未知时，我们提出一种Lepski型数据驱动选择方法来确定分辨率层级。由此得到的自适应估计量在索博列夫类上达到极小极大最优速率，无需采用非线性或基于稀疏性的方法。

关键词：自适应估计；二次泛函；Sobolev 空间；Needletframes；拉普拉斯–贝尔特拉米算子；球面数据

MSC 2020：主分类号 62G05；次分类号 62G20, 42C40

1. 引言

未知密度的二次积分泛函在非参数统计中起着核心作用。它们自然出现在拟合优度检验、能量与光滑度的估计、信号检测，以及反问题与不适定问题的分析中。经典例子包括密度的 L^2-范数、积分平方梯度，以及更一般的Sobolev型半范数。基于独立同分布观测估计此类泛函已有悠久历史（参见，其中包括 [4, 15, 18, 19]）。

本文研究定义在单位球面上的未知密度的二次 Sobolev 型泛函的估计。具体而言，我们考虑如下形式的泛函

$$T_r(f) = \left\| (-\Delta_{\mathbb{S}^d})^{r/2} f \right\|^2_{L^2(\mathbb{S}^d)}, \qquad r \geq 0,$$

其中 $\Delta_{\mathbb{S}^d}$ 表示拉普拉斯–贝尔特拉米算子。此类泛函提供密度光滑度与谱能量的全局度量，并自然地适应球面的几何。它们出现在方向数据与球面数据的分析中，以及

流形上的正则化理论中，以及旋转不变性与几何结构至关重要的应用中（参见，例如，[1, 22]）。

索博列夫型泛函 $T_r(f)$ 在 r 取较小值时具有自然的解释。当 $r = 0$, $T_0(f) = \|f\|^2_{L^2(\mathbb{S}^d)}$ 与密度的经典二次能量相一致，并构成二次泛函最简单的例子。对于 $r = 1$，分部积分表明 $T_1(f) = \|\nabla_{\mathbb{S}^d} f\|^2_{L^2(\mathbb{S}^d)}$，它量化了密度的全局振荡行为或粗糙度。情形 $r = 2$ 对应于拉普拉斯–贝尔特拉米算子作用于 f 后所得结果的 L^2-范数平方，用于度量二阶变差和曲率。更一般地，r 的分数阶值在这些情形之间平滑过渡，并提供了一种谱光滑性概念，这种概念在球面等紧流形上尤为自然。

从统计学角度来看，r 的值越大，密度的高频分量所占权重就越高。因此，对 $T_r(f)$ 的估计会越来越容易受到噪声影响，使截断和多尺度正则化成为该问题的内在组成部分。这一观察促使本文始终采用 Needlet分解和截断估计量。

与欧几里得情形相比，球面情形带来了特有的分析与统计学挑战。经典的偏导数概念和傅里叶分析无法在全局上使用，必须由与拉普拉斯–贝尔特拉米算子相关的谱分解加以替代。从统计学角度来看，球面是方向数据的自然状态空间，而旋转不变性和全局坐标的缺失使得无法直接采用欧几里得方法。此外，Sobolev 泛函 $T_r(f)$ 的估计由于高频分量的放大而具有内在的不适定性，因此截断或正则化成为该问题不可或缺的一部分。在分析方面，球面提供了一个具有离散谱的典型紧黎曼流形，使得 Sobolev 正则性和二次能量泛函能够以自然且内禀的方式加以定义。尽管分析中的若干要素可以推广到配备适当多尺度框架的更一般紧流形，但聚焦于球面情形能够让我们在无需增加额外技术负担的情况下，分离出其中的核心几何与统计特征（例如，参见[4, 15, 18]）。

我们的方法基于球面 needlet 框架，它在保留 $L^2(\mathbb{S}^d)$ 中紧框架性质的同时，为球面上的函数提供局部化的多尺度表示。Needlet 将球谐函数的全局谱结构与精细的空间局部化相结合，并已证明在球面上的多种统计任务中效果显著；相关研究可参见 [2, 10, 22, 25] 等。在当前背景下，Needlet分解使我们能够将密度的 Sobolev 能量按二进频带重新组织，并构造目标泛函自然的截断版本。

我们首先引入适当截断的索博列夫泛函的无偏二次估计量，并通过明确的偏差–方差分析推导出精确的预言机风险界。当底层密度的光滑度已知时，这些估计量在索博列夫类上达到极小极大最优速率。所得预言机界阐明了分辨率层级在平衡截断偏差与随机变异方面的作用，并为自适应程序提供了基准（例如，参见 [15, 19, 26]）。

本文的核心贡献之一是构造了一种不依赖于光滑度参数先验知识的自适应估计量。

与现有主要关注密度估计或局部

特征恢复的基于小波包的方法不同，本文聚焦于全局 Sobolev型能量泛函及其自适应估计，因此需要采用根本不同的偏差–方差和截断分析。

利用多尺度 Needlet 分解所产生的偏差项和方差项的单调性，我们通过 Lepski 型程序以数据驱动的方式选择分辨率层级 [16, 20]。我们证明，所得自适应估计量在一系列 Sobolev 正则性范围内均能达到与预言估计量相同的收敛速率。值得注意的是，这种自适应性通过选择单一的全局分辨率参数实现，无需采用非线性技术或基于稀疏性的技术，体现了 Sobolev 泛函内在的二次性和希尔伯特空间性质。

本文组织如下。第 2 节介绍解析框架，回顾球面上的 Sobolev 几何及其通过球面小波实现的多尺度刻画。第3 节阐述统计方法，包括 Sobolev 泛函的截断、基于小波的无偏估计量的构造，以及预言机偏差–方差界。第 4 节通过 Lepski 型程序选择分辨率层级来研究自适应估计，并建立 Sobolev 类上的极小极大最优自适应速率。第 5 节给出理论偏差–方差权衡和分辨率选择的数值示例。所有证明均汇集于第6节。

2. 球面上的索伯列夫几何与多尺度表示

本节介绍本文所用估计程序背后的解析框架。我们回顾球面调和分析的基本内容，并说明如何通过与拉普拉斯–贝尔特拉米算子相关的谱分解来刻画 Sobolev 正则性。随后，我们引入球面 needlet 框架，它在 $L^2(\mathbb{S}^d)$ 中保持紧框架性质，同时提供多尺度且具有空间局部性的表示。谱 Sobolev 几何与 needlet 分解相结合，使我们能够将 Sobolev 型能量重新组织到各个二进频带中；这一特性对于截断、偏差控制和自适应估计至关重要。本节介绍的所有材料均属标准内容，可见于球面调和分析和多尺度方法的相关文献（例如，参见 [1, 22, 25]）。

全文中，对于两个非负量 a 和 b，若存在一个与样本量 n 及分辨率层级无关的常数 $C > 0$，使得 $a \le C\, b$，则记作 $a \lesssim b$。类似地，$a \gtrsim b$ 表示 $b \lesssim a$，而 $a \asymp b$ 表示 $a \lesssim b$ 和$a \gtrsim b$ 均成立。除非另有说明，所有隐含常数都可能依赖于固定的结构参数，例如维数 d、索博列夫阶数 r以及小波包窗函数的选择，但绝不依赖于 n或分辨率层级。

2.1. 球面上的谱分析与 Sobolev 空间

设 $\Delta_{\mathbb{S}^d}$ 表示 $\mathbb{S}^d$上的拉普拉斯–贝尔特拉米算子。其特征函数为球谐函数

$$\{Y_{\ell,m} : \ell \ge 0,\ m = 1, \dots, M_{\ell;d}\},$$

它们构成 $L^2(\mathbb{S}^d)$的完备正交归一基。相应的特征值由下式给出

$$e_{\ell,d} = \ell(\ell + d - 1), \qquad \ell \geq 0.$$

注意，$e_{\ell,\ d} \asymp \ell^2$ 当 $\ell \to \infty$ 时。对于每个固定的 ℓ，重数 $M_{\ell;d}$ 表示与 $e_{\ell,\ d}$ 相关联的特征空间维数，其显式表达式为

$$M_{\ell;d} = \frac{(2\ell + d - 1)(\ell + d - 2)!}{\ell!\,(d-1)!}, \qquad \ell \geq 0;$$

例如，参见 [1]。特别地，$M_{\ell;d} \asymp \ell^{d-1}$ 当 $\ell \to \infty$。

任意函数 $f \in L^2(\mathbb{S}^d)$ 都可以展开为球谐级数

$$f(x) = \sum_{\ell=0}^{\infty} \sum_{m=1}^{M_{\ell;d}} a_{\ell,m}\, Y_{\ell,m}(x), \qquad a_{\ell,m} = \langle f, Y_{\ell,m} \rangle_{L^2(\mathbb{S}^d)}.$$

球谐表示直接建立了观测样本与目标泛函之间的联系。事实上，对于任意 $(\ell,\ m)$,

$$a_{\ell,m} = \mathbb{E}_f[Y_{\ell,m}(X_1)]\,,$$

因此，可以通过数据的经验对应量来估计 f 的球谐系数

$$\hat{a}_{\ell,m} = \frac{1}{n} \sum_{i=1}^{n} Y_{\ell,m}(X_i).$$

因此，以下介绍的这类二次泛函可以通过其谱表示，利用观测值进行自然估计。

注 2.1 (球谐表示与多尺度表示). 尽管 *Sobolev* 泛函可以通过球谐系数 $a_{\ell,m}$ 直接表示为谱表示，但基于对球谐展开进行硬截断的估计量会导致僵化的偏差—方差结构，因此并不适合进行自适应分辨率选择。具体而言，在球谐域中截断并不能在不同尺度之间形成自然的估计量单调排序，这使得数据驱动选择规则的实施与分析颇为棘手。

为引入作为目标泛函基础的光滑度概念，我们现在回顾球面上 Sobolev 空间的定义；在全局坐标中，经典的偏导数概念并不适用。因此，Sobolev 空间自然地通过拉普拉斯–贝尔特拉米算子的谱性质来定义；该算子是与 $\mathbb{S}^d$ 的黎曼结构相联系的典范椭圆算子。在这一框架下，$(-\Delta_{\mathbb{S}^d})$ 的分数次幂在谱意义上定义了导数，这与欧几里得空间中的傅里叶乘子完全类似。众所周知，这一定义等价于球面上 Sobolev 空间的其他解析刻画，包括基于再生核、平方函数和积分表示的刻画；例如参见 [1, 3, 7, 17, 22]。根据这些等价定义，Sobolev 空间 $H^r(\mathbb{S}^d)$ 由球谐系数满足适当加权可求和性条件的函数构成。

对于 $r \geq 0$，此处及全文中，$f^{(r)}$ 表示由拉普拉斯–贝尔特拉米算子诱导的谱意义下的 f 的 r 阶导数，即

$$f^{(r)}(x) = (-\Delta_{\mathbb{S}^d})^{r/2} f(x) \qquad x \in \mathbb{S}^d. \tag{1}$$

因此， $f^{(r)}$ 可作如下球谐展开：

$$f^{(r)}(x) = \sum_{\ell=0}^{\infty} \sum_{m=1}^{M_{\ell;d}} a_{\ell,m}^{(r)} Y_{\ell,m}(x), \qquad x \in \mathbb{S}^d \tag{2}$$

其中，对任意 $\ell \geq 0$, $m = 1,\ldots, M_{\ell;d}$,

$$a_{\ell,m}^{(r)} = e_{\ell,d}^{r/2} a_{\ell,m} \tag{3}$$

表示相应的球谐系数。

2.2. Sobole夫型二次泛函

本文关注的对象是索博列夫型二次泛函

$$T_r(f) = \left\| (-\Delta_{\mathbb{S}^d})^{r/2} f \right\|_{L^2(\mathbb{S}^d)}^2 = \int_{\mathbb{S}^d} \left| f^{(r)}(x) \right|^2 \mathrm{d}x. \tag{4}$$

我们强调， $T_r(f)$ 对应于阶数为 $r > 0$ 的 Sobolev 半范数，因为常值函数属于 $(-\Delta_{\mathbb{S}^d})^{r/2}$ 的核空间。当且仅当 f 属于 Sobolev 空间 $H^r(\mathbb{S}^d)$ 时，该半范数才是有限的。利用公式（2）、（3）和（4），$T_r(f)$ 可表示为精确形式

$$\begin{aligned} T_r(f) &= \sum_{\ell=0}^{\infty} \sum_{m=1}^{M_{\ell;d}} \left| a_{\ell,m}^{(r)} \right|^2 \\ &= \sum_{\ell=0}^{\infty} \sum_{m=1}^{M_{\ell;d}} e_{\ell,d}^{r} \left| a_{\ell,m} \right|^2 . \end{aligned} \tag{5}$$

二次泛函估计在非参数统计中有着悠久的历史；参见例如 [4, 15, 18, 19]。泛函 $T_r(f)$ 对球面上密度 f 的平滑度和谱能量提供了定量度量。 r 的值较小时对应于全局能量度量，而 r 的值越大，则越着重于 f 的高频振荡和精细尺度不规则性。

从统计学的角度来看，阶数 r 决定了估计问题的不适定程度。较大的 r 意味着通过 Sobolev 权重 $e_{\ell,\,d}^{r}$对高频分量的放大作用更强，从而使经验估计量的方差增长更快。因此，随着 r 增大，对谱展开或多尺度展开进行截断变得愈发关键。第 3 节所构建的截断估计量的偏差–方差权衡将明确体现这一特征。

全文假设密度 f 属于 $H^s(\mathbb{S}^d)$，其中 $s > r$，且本质有界。特别地，条件 $s > d/2$ 根据Sobolev嵌入可保证 $f \in L^\infty(\mathbb{S}^d)$，这足以满足下文推导方差界的需要。当$s \leq d/2$时，也可以直接将有界性作为一项附加假设。

2.3. Needletframe 与函数分解

Needlet为球面上的函数提供了局部化的多尺度表示，将球谐函数的谱性质与空间局部化相结合。与全局

调和基不同，Needlet同时在频率和空间上局部化，这一特性在非参数统计应用中尤为有利。基于小波包的方法已被广泛用于密度估计、回归、阈值化、检验以及球面随机场的分析；例如，参见 [2, 10, 12, 14, 22, 23, 24, 25]。

Fix $B > 1$ 并令 $b \in C^\infty(\mathbb{R})$ 为非负函数，且支撑于 $[B^{-1}, B]$，使得 t

$$\sum_{j\geq 0} b^2\left(\frac{\ell}{B^j}\right) = 1, \qquad \text{for all } \ell \geq 1.$$

指标 $j \geq 0$ 表示分辨率层级，对应于宽度为 B^j 的二进频带。对于每个固定的 j，令 $\{(\xi_{j,k}, \lambda_{j,k}) : k = 1, \dots, K_j\}$ 为球面 $\mathbb{S}^d$ 上的求积公式，该公式对次数不超过 B^{j+1} 阶的球面多项式精确，其中指标 k 标记球面上的空间位置，且 $K_j \asymp B^{dj}$。

needlet 系统定义为

$$\psi_{j,k}(x) = \sqrt{\lambda_{j,k}} \sum_{\ell \in \Lambda_j} b\left(\frac{\ell}{B^j}\right) \sum_{m=1}^{M_{\ell;d}} Y_{\ell,m}(\xi_{j,k})\, Y_{\ell,m}(x),$$

其中 $\Lambda_j = \{\ell : B^{j-1} \leq \ell \leq B^{j+1}\}$。集合 $\{\psi_{j,k} : j \geq 0;\ k = 1, \dots, K_j\}$ 构成了 $L^2(\mathbb{S}^d)$ 的紧框架，即

$$\|g\|^2_{L^2(\mathbb{S}^d)} = \sum_{j\geq 0} \sum_{k=1}^{K_j} |\langle g, \psi_{j,k}\rangle_{L^2(\mathbb{S}^d)}|^2, \qquad g \in L^2(\mathbb{S}^d).$$

这种多尺度分解使我们能够将 f 的谱能量重新组织为局部化频带，这一特性对于截断、偏差控制和自适应估计至关重要。

此外，Needlet 在频率和空间上都具有良好的局部化性质。具体而言，它们的谱支撑位于二进频带内，并且在求积点 $\xi_{j,k}$ 附近具有空间局部化性质。具体来说，对于每个整数 $M \geq 1$，都存在一个常数 $C_M > 0$，使得

$$|\psi_{j,k}(x)| \leq C_M\, B^{jd/2} \left(1 + B^j d_{\mathbb{S}^d}(x, \xi_{j,k})\right)^{-M}, \qquad x \in \mathbb{S}^d,$$

其中，$d_{\mathbb{S}^2}(\cdot,\cdot)$ 表示 $\mathbb{S}^d$ 上的测地距离。

因此，对于每个 $1 \leq p \leq \infty$，都存在一个常数 $C_p > 0$，使得

$$\|\psi_{j,k}\|_{L^p(\mathbb{S}^d)} \asymp B^{jd\left(\frac{1}{2} - \frac{1}{p}\right)}, \qquad j \geq 0,\ k = 1, \dots, N_j. \tag{6}$$

公式 (6) 将在后续统计分析中反复用于控制经验小波系数的矩和方差。特别地，$\|\psi_{j,k}\|_{L^2(\mathbb{S}^d)} \asymp 1$ 和$\|\psi_{j,k}\|_{L^\infty(\mathbb{S}^d)} \asymp B^{jd/2}$。

现在我们将 Needlet分解 与 Sobolev 型正则性联系起来。令 $r \geq 0$，并令$f \in H^r$（$\mathbb{S}^d$）。回顾一下，f 在谱意义下的 r 阶导数定义为

$$f^{(r)} = (-\Delta_{\mathbb{S}^d})^{r/2} f.$$

由于 needlet 系统构成了 $L^2(\mathbb{S}^d)$ 的紧框架，将框架恒等式应用于函数 $f^{(r)}$，可得

$$T_r(f) = \|f^{(r)}\|^2_{L^2(\mathbb{S}^d)} = \sum_{j,k} \left|\langle f^{(r)}, \psi_{j,k}\rangle\right|^2 .$$

我们定义 f 及其 Sobolev 导数的 needlet 系数为

$$\beta_{j,k} = \langle f, \psi_{j,k}\rangle_{L^2(\mathbb{S}^d)}, \qquad \beta^{(r)}_{j,k} = \langle f^{(r)}, \psi_{j,k}\rangle_{L^2(\mathbb{S}^d)}.$$

为便于后续使用，我们引入 needlet 原子的 Sobolev 导数，

$$\psi^{(r)}_{j,k} := (-\Delta_{\mathbb{S}^d})^{r/2}\psi_{j,k}. \tag{7}$$

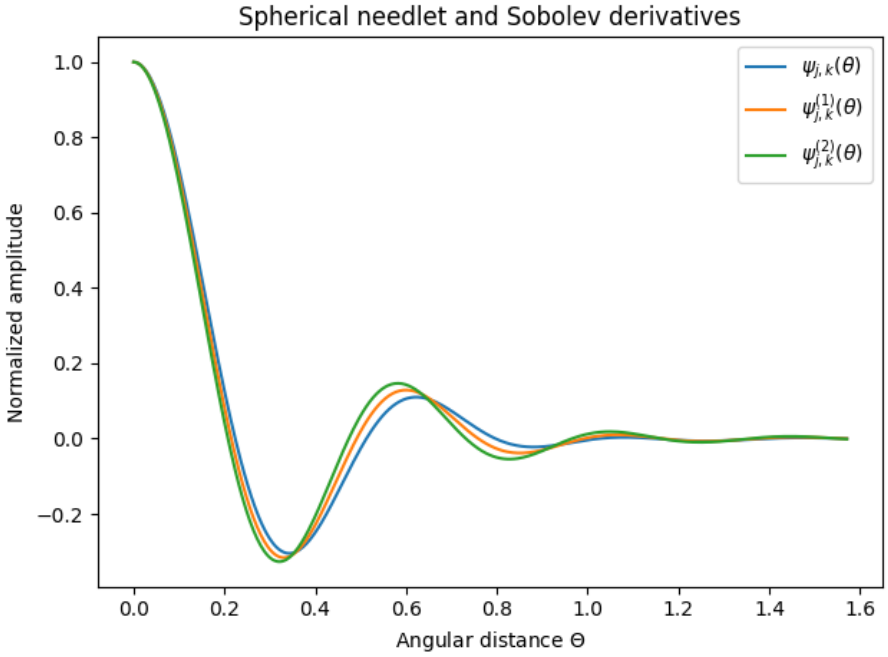


图 1： 分辨率层级为 $j = 3$ 的球面 needlet $\psi_{j,k}$ 及其一阶和二阶 Sobolev 导数 $\psi^{(1)}_{j,k} = (-\Delta_{\mathbb{S}^2})^{1/2}\psi_{j,k}$ 和 $\psi^{(2)}_{j,k} = (-\Delta_{\mathbb{S}^2})\psi_{j,k}$，绘制为关于中心 $\xi_{j,k}$的角距离 Θ 的函数。求导会增强振荡行为，同时保持空间局部化，展示了 Sobolev 导数所引起的依赖尺度的放大效应。

引理 2.2（导数与 needlet 系数）。设 $r \geq 0$ 且设 $f \in H^r(\mathbb{S}^d)$。则 *Sobolev* 导数 $f^{(r)}$ 具有如下表示

$$f^{(r)} = \sum_{j\geq 0}\sum_{k=1}^{K_j} \beta^{(r)}_{j,k}\,\psi_{j,k} = \sum_{j,k} \beta_{j,k}\,\psi^{(r)}_{j,k},$$

并在 $L^2(\mathbb{S}^d)$ 中收敛。此外，对于 $j \geq 0$，以及 $k = 1$、…、K_j，存在常数$0 < c_r \leq C_r < \infty$，使得

$$c_r\, B^{jr}|\beta_{j,k}| \leq |\beta^{(r)}_{j,k}| \leq C_r\, B^{jr}|\beta_{j,k}|.$$

图 1 通过数值示例展示了这种现象在 $d = 2$情形下的表现，表明索伯列夫微分会放大高频振荡，同时保留空间局部化特性。

鉴于前述说明，Sobolev 泛函 $T_r(f)$ 有两种等价的、基于 needlet 的表示。一方面，由 needlet 框架的紧性可得，

$$T_r(f) = \sum_{j\geq 0}\sum_{k=1}^{K_j} \left(\beta_{j,k}^{(r)}\right)^2, \tag{8}$$

这是 $L^2(\mathbb{S}^d)$ 中的精确恒等式。另一方面，根据球面上的 Littlewood–Paley理论，

$$T_r(f) \asymp \sum_{j\geq 0}\sum_{k=1}^{K_j} B^{2jr}\,\beta_{j,k}^2, \tag{9}$$

这种表示表明，Sobolev 泛函取决于 needlet 系数的全局二次能量，而非其空间分布；这与 Besov范数不同，后者刻画局部或稀疏正则性模式。等价常数仅取决于 r、 d 以及所选的小波包窗函数。

注 2.3 （多尺度结构与自适应分辨率选择）。 *needlet* 分解将球谐系数中蕴含的谱信息重新组织为按分辨率层级 j 编号的二进频带。这种多尺度结构使不同分辨率层级间呈现单调的偏差–方差行为：随着 j 增大，截断偏差减小而方差增大。这种单调排序对于 *Lepski* 型自适应程序的实施与理论分析至关重要，也是本框架采用 *needlet*表示的主要动机之一。

注记 2.4 （索伯列夫几何与框架紧性）。 在全文中，我们将自由地在精确表示 (8) 与等价的加权 *needlet* 表示 (9)之间切换，具体取决于所需的是代数恒等式还是渐近论证。

尽管 *Sobolev* 小波原子 $\{\psi_{j\,k}^{(r)}\,,\,\}$ 在 $L^2(\mathbb{S}^d)$ 中并不构成紧框架，但这并不影响 *Sobolev* 泛函 $T_r(f)$ 的表示。

更精确地说，*Sobolev* 泛函满足如下等价关系

$$\sum_{j\geq 0}\sum_{k=1}^{K_j}\left|\langle f, \psi_{j,k}^{(r)}\rangle_{L^2(\mathbb{S}^d)}\right|^2 \asymp \sum_{j\geq 0}\sum_{k=1}^{K_j}\left|\langle f^{(r)}, \psi_{j,k}\rangle_{L^2(\mathbb{S}^d)}\right|^2,$$

其等价常数与分辨率层级无关。

Sobolev 小波原子层面框架紧性的损失可以显式量化。若 $\psi_{j\,k}^{(r)}\,,\,:=(-\Delta_{\mathbb{S}^d})^{r/2}\psi_{j,\,k}$，则小波包窗函数的谱局部化蕴含依赖于尺度的范数界

$$\|\psi_{j,k}^{(r)}\|_{L^p(\mathbb{S}^d)} \asymp B^{j\left(r+d\left(\frac{1}{2}-\frac{1}{p}\right)\right)},\ \textit{for } p\geq 1; \qquad \|\psi_{j,k}^{(r)}\|_{L^\infty(\mathbb{S}^d)} \asymp B^{jr+\frac{d}{2}j},$$

对 $j\geq 0$和 $k=1,\ldots,N_j$均一致成立。这些界限反映了这样一个事实：*Sobolev* 微分会对 *needlet* 原子产生频率依赖放大。具体而言，尽管原始系统 $\{\psi_{j,\,k}\}$具有一致有界的 L^2-范数并构成紧框架，

$\|\psi_j^{(r)}\,,\,_k\|_{L^2}$ 随分辨率层级 j 的增长，使 *Sobolev* 小波原子系统无法在 $L^2(\mathbb{S}^d)$ 中构成紧框架。

关键在于，紧性保留在经过求导的函数层面，而非 *needlet* 原子的 *Sobolev* 导数层面。这一观点使我们能够使用固定且与 r 无关的 *needlet* 框架，并通过依赖尺度的权重来编码 *Sobolev* 正则性；这种选择在分析上自然，也非常适合自适应统计估计。

注 2.5（通过 needlet 表示 Sobolev 空间和 Besov 空间，以及关于 r依赖型构造）。更一般地，同一组 *needlet* 系数给出了 $\mathbb{S}^d$ 上 *Besov* 空间的标准刻画，其中 *Sobolev* 空间对应于特殊情形 $p = q = 2$。

近年来，基于收缩带宽的相关构造已在 *[11,13]* 中得到研究。在这一框架下，多尺度几何由预先设定的膨胀序列决定，该序列控制相对带宽以及相邻频率窗口的重叠程度，从而形成不同的局部化和冗余模式。当目标索博列夫阶数 r 事先已知时，可以对这类构造进行调节，以获得非常精确的表示。然而，在当前设定中，r 未知，必须以自适应方式处理，因此将框架几何与特定的光滑度水平绑定，不符合一致自适应性的要求。

因此，我们采用固定的 *needlet* 框架，并通过依赖尺度的权重来刻画 *Sobolev* 正则性。这一选择既保留了普适性，也适用于自适应估计；在自适应估计中，截断展开是方差控制的内在环节。

3. 截断 Sobolev 泛函的估计

未知密度的二次泛函估计已在非参数统计中得到广泛研究；例如，参见 [4, 15, 18, 19]。这类方法的一个共同特点是，由于高频分量不断累积，直接估计完整泛函是不适定的，因此截断或正则化是问题本身不可或缺的一部分。

在当前的球面情形中，Sobolev 泛函 $T_r(f)$ 可自然地表示为关于 needlet 系数的多尺度表示。借助这一表示，我们可以构造适当截断版本 $T_r(f)$ 的无偏二次估计量，并通过偏差–方差分解分析其风险；这一分析与经典的欧几里得情形十分相似。下面我们将介绍截断方案及相应的估计策略。

3.1. 球面上的统计模型

令 $(\mathbb{S}^d, \mathrm{d}x)$ 表示配备均匀表面测度的 $\mathbb{R}^{d+1}$ 单位球面。令 $X_1, \dots, X_n$ 为具有共同分布 P_f 的独立同分布随机变量，并且相对于 $\mathrm{d}x$ 具有密度 f。我们考虑非参数统计模型

$$\mathcal{F} = \left\{ f \in L^2(\mathbb{S}^d) : f \geq 0,\ \int_{\mathbb{S}^d} f(x)\,\mathrm{d}x = 1 \right\},$$

以及相应的乘积测度族 $\{P_f^{\otimes n} :\ f \in \mathcal{F}\}$。

推断仅基于观测样本 $X_1, \dots, X_n$，并将密度 f 视为 $\mathcal{F}$ 中的未知元素。这一框架对应于经典的方向

数据分析，其中观测值取自单位球面，并在旋转不变性的假设下进行统计推断（例如，参见 [21]）。

我们关注的对象是二次 Sobolev 泛函

$$T_r(f) = \|(-\Delta_{\mathbb{S}^d})^{r/2} f\|^2_{L^2(\mathbb{S}^d)}, \qquad r \geq 0,$$

假设它在 $f \in H^r(\mathbb{S}^d)$ 下有限。统计目标是根据数据估计 $T_r(f)$。在全文中，我们默认密度 f 有界；这在 Sobolev 正则性 $s > d/2$ 下自动成立，并足以保证下文推导的所有方差界。

鉴于第 2.3 节引入了 Sobolev 导数的基于 needlet 的表示，我们可以直接根据与求导后的 needlet 原子相关联的经验系数来构建推断。具体而言，对于每个分辨率层级 j 和空间索引 k，我们定义经验 Sobolev–needlet 系数

$$\widehat{\beta}^{(r)}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} \psi^{(r)}_{j,k}(X_i),$$

其中 $\psi^{(r)}_{j\,k}$，由（7）给出。根据期望的线性性以及 Laplace–Beltrami 算子的自伴性，

$$\mathbb{E}_f\left[\widehat{\beta}^{(r)}_{j,k}\right] = \langle f^{(r)}, \psi_{j,k} \rangle_{L^2(\mathbb{S}^d)} = \beta^{(r)}_{j,k},$$

因此，$\widehat{\beta}^{(r)}_{j\,k}$，是相应 Sobolev–needlet 系数的无偏估计量。这些经验量构成了后续小节中用于估计截断 Sobolev 泛函的基本构件。

注 3.1 (截断方案的模型无关性)。 虽然本文为具体说明而聚焦于密度估计，但本节提出的方法更广泛地适用于球面上的一类非参数统计模型，其中研究对象可表示为基于 *needlet* 的表示，且其系数（可能经过微分）可通过经验数据估计。具体而言，这包括球面密度的全局阈值化和自适应估计问题 *[2, 10]*，，以及基于球面设计的回归型模型 *[12, 23]*。在这些情形下，经验小波系数具有类似的偏差和方差性质，而二次 *Sobolev* 型泛函的估计可在同一基于截断的框架中处理。

重要的是，截断泛函及其相关偏差项的定义纯粹是解析性的，并不依赖于具体的观测模型。统计模型只通过经验小波系数的随机行为发挥作用，影响估计量的方差，但不影响其结构形式。解析逼近与概率波动之间的这种分离，构成了所提出的估计策略能够稳健适用于不同球面模型的基础，也促使本文采用统一的处理方式。

3.2. 目标泛函的截断

尽管 Sobolev 泛函 $T_r(f)$ 可以用 needlet 系数精确表示，但由于高频率分量不断累积，直接从数据中对其进行估计是不适定的。随着 r 增大，Sobolev 权重会放大精细尺度振荡，

导致朴素经验估计量的方差发散。因此，多尺度展开的截断是统计问题本身不可或缺的一部分，而非技术性手段；例如参见 [4, 15, 19]。

鉴于恒等式

$$T_r(f) = \sum_{j\geq 0}\sum_{k=1}^{K_j}\left(\beta_{j,k}^{(r)}\right)^2, \qquad \beta_{j,k}^{(r)} = \langle f^{(r)}, \psi_{j,k}\rangle_{L^2(\mathbb{S}^d)},$$

该恒等式由将紧框架性质应用于 Sobolev 导数 $f^{(r)}$ 上的 needlet 系统而得，因此，自然可以通过将求和限制在低分辨率层级，来定义该泛函的截断版本。对于称为截断分辨率层级的整数 $J \geq 0$，我们定义截断索伯列夫泛函

$$T_r^{(J)}(f) := \sum_{j\leq J}\sum_{k=1}^{N_j}\left(\beta_{j,k}^{(r)}\right)^2. \tag{10}$$

该量表示 f 在截至 B^J 阶的频带中所包含的索伯列夫能量，也可解释为将 $f^{(r)}$ 投影到由分辨率不高于 J 的 needlet 张成的子空间上。

定义（10）是球面索伯列夫几何所固有的，不依赖任何近似框架表示。特别是，截断偏差完全源于舍弃高频 Sobolev 分量，不受 Sobolev 小波原子不具紧框架性质的影响。这一选择明确地区分了解析逼近误差与后续的统计考量。

在渐近分析中，通常可将 $T_{(J)r}(f)$ 表示为原始 needlet 系数的形式。根据球面上的 Littlewood–Paley理论以及小波包窗函数的谱局部化性质，Sobolev 系数与 needlet 系数满足

$$\left|\beta_{j,k}^{(r)}\right| \asymp B^{jr}\,|\beta_{j,k}|,$$

在 $j \geq 0$和 $k = 1, \ldots, K_j$上一致。因此，截断泛函可以表示为等价形式

$$T_r^{(J)}(f) \asymp \sum_{j\leq J}\sum_{k=1}^{N_j} B^{2jr}\,\beta_{j,k}^2, \tag{11}$$

其中常数仅取决于 r, d以及小波包窗函数的选择，而与 J无关。全文中，我们将根据需要精确恒等式还是渐近界，在（10）与（11）之间灵活切换。

现在可以清晰地刻画由（10）引入的截断偏差。若$f \in H^s(\mathbb{S}^d)$ 对某个 $s > r$成立，则

$$T_r(f) - T_r^{(J)}(f) = \sum_{j>J}\sum_{k}\left(\beta_{j,k}^{(r)}\right)^2 \lesssim B^{-2J(s-r)},$$

其中，该界由等价关系（11）以及加权小波包系数的可求和性推出。条件 $s > r$ 是确保截断偏差随 $J \to \infty$而消失所必需的；当 $s = r$时，该泛函处于可估性的边界，若无进一步假设，则无法进行相合估计。本文始终

因此，我们将讨论范围限制在 $s > r$这一情形，这对于对 $T_r(f)$ 进行相合估计是必要的。

这种偏差行为凸显了分辨率层级 J 作为正则化参数的作用，它控制着逼近误差与随机变异之间的权衡。在下一小节中，我们将利用这一结构构造 $T_{(J)r}(f)$ 的无偏估计量，并通过经典的偏差–方差分解分析其风险。因此，截断所引入的偏差仅由底层密度的超额光滑度 $s-r$和分辨率层级 J决定。

3.3. 估计量的构造

截断索伯列夫泛函 $T_{(J)r}(f)$ 在 (10) 中定义，是未知 Sobolev–needlet 系数 $\beta_{j\,k}^{(r)}$,的二次泛函。基于单个系数经验估计值的朴素代入估计量通常会产生不可忽略的对角偏差。这一现象源于估计误差与二次聚合之间的相互作用，即使在有限维情形下也依然存在。

消除这一偏差的一种标准且有效的方法是采用样本拆分策略，从而构造方差结构清晰的无偏二次估计量。这类构造可追溯至欧几里得空间中有关二次泛函估计的开创性研究，并且特别适用于希尔伯特空间情形；参见 [4, 15, 19] 等文献。

因此，我们将样本$\{X_1,\dots,X_n\}$ 分成两个大小相等的独立子样本，分别记为$\mathcal{D}_1$ 和$\mathcal{D}_2$。对于 $a=1$,2，我们定义经验 Sobolev–needlet 系数

$$\widehat{\beta}^{(r)}_{j,k;(a)} = \frac{2}{n}\sum_{X_i\in\mathcal{D}_a} \psi^{(r)}_{j,k}(X_i), \qquad \psi^{(r)}_{j,k} = (-\Delta_{\mathbb{S}^d})^{r/2}\psi_{j,k}.$$

根据期望的线性性和拉普拉斯–贝尔特拉米算子的自伴性，

$$\mathbb{E}_f\left[\widehat{\beta}^{(r)}_{j,k;(a)}\right] = \beta^{(r)}_{j,k}, \qquad a=1,2,$$

并且这两组$\{\widehat{\beta}^{(r)}_{j\,k\,;;(1)}\}$和$\{\widehat{\beta}^{(r)}_{j\,k\,;;(2)}\}$相互独立。

我们将截断 Sobolev 泛函的估计量定义为

$$\widehat{T}^{(J)}_r = \sum_{j\le J}\sum_{k=1}^{N_j} \widehat{\beta}^{(r)}_{j,k;(1)}\,\widehat{\beta}^{(r)}_{j,k;(2)}. \tag{12}$$

B由于两个子样本相互独立，且经验系数具有无偏性 s,

$$\mathbb{E}_f\left[\widehat{T}^{(J)}_r\right] = \sum_{j\le J}\sum_{k=1}^{N_j}\left(\beta^{(r)}_{j,k}\right)^2 = T^{(J)}_r(f),$$

因此， $\widehat{T}^{(J)}_r$是截断目标泛函的无偏估计量。

这一构造与欧几里得空间中二次泛函的经典估计量相呼应，并通过使用 Sobolev–needlet 系数自然地推广到球面情形。该估计量继承了 needlet 框架所诱导的多尺度结构，其中

分辨率层级 J 充当全局正则化参数。随着 J 增大，估计量会纳入更精细的频率分量，从而降低截断偏差，但代价是方差增大。下一小节将利用分辨率层级之间这种单调的偏差–方差行为，推导 oracle 风险界并构造自适应估计量。

类似的样本拆分策略已成功应用于基于 needlet 的球面估计问题，包括密度估计、阈值化和回归型模型；例如，参见 [2, 10, 14]。

注 3.2（样本拆分与 U-统计量的构造）。截断泛函 $T_{(J)_r}(f)$ 的二次结构允许采用不同的无偏估计策略。特别是，也可以通过对在不同观测值处计算的 *Sobolev–needlet* 原子的乘积进行对称化，构造基于单样本二阶退化U-统计量的估计量。这类估计量是无偏的，并且达到与本文所考虑的样本拆分估计量相同的极小极大速率。

然而，在当前的 L^2-*Sobolev* 框架下，样本拆分构造具有若干优势。首先，它带来更简单的方差结构，避免了单样本 U-统计量的霍夫丁分解中出现的线性投影项。其次，它能够得到更精确的常数，并且对微分 *needlet* 原子的矩条件要求更弱。最后，也是对本文目的而言最重要的一点，样本拆分保留了各分辨率层级上清晰的偏差–方差分解，这对自适应选择规则的实施和分析至关重要。

U-统计量构造在估计非二次或非希尔伯特型泛函时发挥着更核心的作用，例如 L^p 型能量，或带有 $p \neq 2$的 *Besov*型范数。在这类情形下，对称化是内在要求，而样本拆分无法带来同等程度的简化。关于在紧流形上采用相关的基于 U-统计量的方法处理 *Sobolev* 和 *Besov* 泛函，请参见例如 *[6]*。

3.4. 偏差–方差权衡

估计量 $\widehat{T}_r^{(J)}$的均方误差由经典的偏差–方差分解所决定。偏差源于索博列夫展开的截断，且仅取决于底层密度的光滑度；方差则反映经验 Sobolev–needlet 系数的随机波动。两项均随分辨率层级 J呈单调变化，这是自适应分析所利用的一项关键结构特征。

命题 3.3 (截断索博列夫估计量的偏差–方差分解**)**。设$r \geq 0$ 且设 $f \in H^s(\mathbb{S}^d)$ 对某个 $s > r$成立。考虑估计量 $\widehat{T}_r^{(J)}$，其定义见 (12)。则存在常数 $C > 0$，它仅取决于 r、s、d以及小波包窗函数的选取，使得

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(J)} - T_r(f)\right)^2\right] \leq C\left(B^{-4J(s-r)} + \frac{B^{J(d+4r)}}{n}\right), \qquad \forall J \geq 0.$$

特别地：

- 截断偏差满足

$$T_r(f) - T_r^{(J)}(f) \lesssim B^{-2J(s-r)};$$

- 估计量的方差满足

$$\mathrm{Var}\left(\widehat{T}_r^{(J)}\right) \asymp \frac{B^{J(d+4r)}}{n}.$$

此外，分辨率层级的选择

$$B^{J^\star} \asymp n^{1/(2s+d+4r)}$$

平衡了偏差与方差项，并得到

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(J^\star)} - T_r(f)\right)^2\right] \asymp n^{-4(s-r)/(2s+d+4r)}.$$

命题 3.3中的界凸显了分辨率层级 J 作为全局正则化参数的作用。随着 J 增大，偏差单调减小，而方差单调增大，这反映了高频 Sobolev 分量的累积。这种跨尺度的单调结构为J的自适应选择奠定了基础，该选择将在下一节中加以发展。

4. 分辨率层级的自适应选择

在第3 节中所发展的估计方法关键依赖于分辨率层级 J的选择，它控制着截断偏差与随机变异性之间的平衡。当光滑度参数 s （底层密度的）已知时，这一权衡会导向对 J 的预言选择，并得到极小极大最优收敛速率，正如命题 3.3所确立的那样。然而在实践中， f 的正则性通常未知，分辨率层级必须以完全数据驱动的方式加以选择。

二次泛函的自适应估计是非参数统计中的核心问题。奠基性贡献包括 [20]，的开创性工作，以及随后针对二次泛函和基于 U 统计量的泛函所取得的进展 [15, 19]。关于带宽和分辨率选择的较新统一视角由 [16] 的框架提供。在当前情形下，通过选择单个全局分辨率层级实现自适应性，利用截断 needlet 展开所导致的偏差项和方差项的单调变化特性。

4.1. 通过 Lepski 型程序进行自适应分辨率选择

自适应过程的目标是在未知光滑度参数的情况下，模仿分辨率层级的预言选择 s。正如命题 3.3 所示，最优分辨率层级在递减的截断偏差与递增的方差项之间取得平衡。由于这种平衡明确取决于 s，而该参数在实际中未知，因此必须完全根据数据选择分辨率层级。

当前问题的一个关键结构特征是，估计量族 $\{\widehat{T}_r^{(J)}\colon J \geq 0\}$ 自然地按尺度排序。对于 $J' > J$，估计量 $\widehat{T}_{(J')_r}$包含 $\widehat{T}_r^{(J)}$ 所使用的所有频率分量，以及额外的高频贡献。因此，差值

$$\widehat{T}_r^{(J)} - \widehat{T}_r^{(J')}$$

将中间尺度的贡献单独分离出来，并直接揭示这些额外分量究竟能否降低截断偏差，还是会被随机波动所主导。

这种嵌套结构使该问题处于 Lepski 型自适应 [5, 16, 20]程序的标准框架中。其基本原则是比较不同分辨率层级下的估计量，并选择一个最大的分辨率，使得较粗与较细估计量之间的差异仍与随机变异相容。超过这一点后，提高分辨率主要会放大噪声，而非提升近似精度。

为形式化这一思路，我们首先控制不同分辨率层级下估计量之差的随机波动。

引理 4.1 (不同分辨率下的随机波动界)。设 $r \geq 0$，并令 $f \in H^s(\mathbb{S}^d)$，其中某个 $s > r$。存在一个常数 $C_0 > 0$，它仅依赖于 r、d 和小波包窗函数，使得对所有 $J\prime > J$，

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(J)} - \widehat{T}_r^{(J')}\right)^2\right] \leq C_0^2 \, \frac{B^{J'(d+4r)}}{n}.$$

In 特别地，$\widehat{T}_r^{(J} - \widehat{T}_r^{(J\prime)}$)的随机波动量级为 $n^{-1/2}B^{J'(d/2+2r)}$.

引理 4.1 表明，分辨率为 J和 J'的估计量之差的波动性由较精细尺度 $J\prime$上的方差决定。这促使我们定义尺度相关波动界

$$\omega(J) = C_0 \, n^{-1/2} B^{J(d/2+2r)},$$

该界用作阈值，用于区分高频分量增量贡献中的信号与噪声。

我们考虑一组有限的可行分辨率层级

$$\mathcal{J}_n = \{J_{\min}, \ldots, J_{\max}\}, \qquad J_{\max} \asymp \log_B n,$$

其中，$J_{\min} \geq 0$为固定值，独立于 n，且取值足够小，使得对于所有 $f \in H^s(\mathbb{S}^d)$，$T_r^{J_{\min}}(f) < \infty$。然后，我们计算所有 $J \in \mathcal{J}_n$对应的估计量 $\widehat{T}_r^{(J)}$。遵循经典的 Lepski 原则，我们定义数据驱动的分辨率层级

$$\widehat{J} = \min\left\{J \in \mathcal{J}_n : \forall J' > J, \ \left|\widehat{T}_r^{(J)} - \widehat{T}_r^{(J')}\right| \leq \omega(J')\right\}. \tag{13}$$

Sobolev 泛函的自适应估计量定义为

$$\widetilde{T}_r = \widehat{T}_r^{(\widehat{J})}.$$

选择规则（13）选取这样一个最大的分辨率层级，使得更精细尺度的贡献仍处于随机波动的范围内。过小的分辨率层级会被舍弃，因为相应的估计量与较高分辨率层级下的估计量差异显著；过大的分辨率层级也会被排除，因为额外增量主要受噪声支配。这种机制是自适应估计中的经典方法，也是非参数统计中许多最优选择程序的基础[5, 16, 20]。

4.2. 自适应风险界

以下结果表明，自适应估计量在一系列索博列夫类上均达到与命题 3.3 中预言估计量相同的收敛速率。这一最优性源于命题 3.3 所确立的偏差主导区间与方差主导区间之间的清晰分离，也得益于 Lepski 型选择规则能够以相差一个乘法常数的精度识别这一转变。

从概率的角度来看，自适应过程通过确保所选分辨率层级以高概率保持在预言选择附近来控制风险。过小的分辨率层级会被舍弃，因为相应的估计量与更高分辨率下的估计量差异显著；过大的分辨率层级则会被排除，因为额外增量会被随机波动所淹没。这一机制是自适应估计中的经典方法，也是非参数统计中许多最优选择规则的基础 [5, 16, 20]。

在整个自适应分析中，我们假设未知的光滑度参数s 属于一个紧区间 $[r+\varepsilon, s_{\max}]$，其中某个固定的 $\varepsilon > 0$ 和 $s_{\max} < \infty$。下界 $s \geq r+\varepsilon$ 确保与 Sobolev 泛函可估性边界之间保持一致的间隔，而上界 $s \leq s_{\max}$ 则保证在整个自适应分析中使用的方差界和集中界中的常数均受到一致控制。这两个条件都是二次泛函自适应估计中的标准条件。

引理 4.2（控制过度平滑事件）。 在定理 *4.3*的假设下，存在常数 c_1， $c_2 > 0$ 使得

$$\sup_{f \in H^s(\mathbb{S}^d)} \mathbb{P}_f\Big(\widehat{J} > J^\star\Big) \leq c_1 \exp(-c_2 \log n),$$

一致地对 $s \in [r+\varepsilon, s_{\max}]$ 成立。

引理 4.2 保证 Lepski 选出的分辨率层级以高概率不会超过预言选择，因此，除去可忽略事件，自适应估计量与预言估计量处于相同的偏差—方差区域。

定理 4.3（索博列夫类上的自适应估计）。 设 $r \geq 0$，且 s 属于某个紧区间 $[r+\varepsilon, s_{\max}]$，其中 $\varepsilon > 0$。则存在常数 $C > 0$，使得

$$\sup_{f \in H^s(\mathbb{S}^d)} \mathbb{E}_f\left[\Big(\widetilde{T}_r - T_r(f)\Big)^2\right] \leq C\, n^{-4(s-r)/(2s+d+4r)},$$

一致地对 $s \in [r+\varepsilon, s_{\max}]$ 成立。

证明依赖于 Lepski 型程序的标准论证，结合了 $\widehat{T}_{(J)_r}$ 的无偏性、关于二次 U-统计量的集中不等式，以及偏差项和方差项关于分辨率层级的单调性；例如参见 [15, 16, 20]。

注 4.4（全局自适应性与索伯列夫几何）。 上述自适应过程无需事先掌握光滑度参数s 的先验知识，也无需借助非线性或逐系数阈值化技术，即可达到最优速率。这反映了 *Sobolev* 泛函 $T_r(f)$ 内在的二次性与希尔伯特空间性质，它依赖于全局能量聚合，而非密度的局部或稀疏特征。

需要强调的是，此处的自适应性仅通过选择一个全局分辨率层级来实现。这与 *Besov* 自适应程序形成对比，这些程序

依赖于局部阈值化和稀疏性假设。对于本文所考虑的 *Sobolev* 泛函，此类非线性技术既非必要，也不自然，因为目标泛函并不编码底层密度的局部特征或稀疏特征。

更广泛地说，在基本模型的温和扩展下，自适应构造依然稳定。例如，存在加性观测噪声或与精确 *Sobolev* 正则性略有偏离，主要会影响方差界中的常数，而不会改变自适应规则的结构或由此得到的收敛速率。这进一步表明，在当前设定中，自适应性取决于整体几何考量，而非数据的精细尺度特征。

注 4.5（关于超越索博列夫类的自适应性）。 定理 *4.3*的自适应结果是在索博列夫类 $H^s(\mathbb{S}^d)$ 上表述的，这些类与 *Besov* 空间 $B^s_{2,\,2}(\mathbb{S}^d)$ 重合。尽管球面 *needlet* 系数能够完整刻画更一般的 *Besov* 空间 $B^s_{p,\,q}(\mathbb{S}^d)$，其中 $p \neq 2$ *[2, 14, 24, 25]*，但在这类空间上自适应估计二次 *Sobolev* 泛函 $T_r(f)$ 并不自然，也无法通过当前方法实现。

这一局限性内在于目标泛函的几何结构，而非估计技术。事实上，$T_r(f)$ 编码了 *Sobolev* 导数 $f^{(r)}$的全局 L^2-能量，因而本质上是希尔伯特型的。因此，其极小极大最优估计仅依赖于 *needlet* 系数的总能量，而不依赖于它们在各空间位置上的分布，这与二次泛函估计的经典结果 *[15, 19]*完全类似。

相比之下，具有 $p \neq 2$ 的 *Besov* 空间描述了非齐次或稀疏正则性模式；在这些模式中，自适应程序必须依靠逐系数或分块阈值化来利用稀疏性 *[9, 26]*。当目标泛函本身对局部行为敏感时，此类非线性技术至关重要 *[8]*。然而，在当前设定下，阈值化会改变二次泛函的结构并引入额外的偏差项，却无法改进 $T_r(f)$的极小极大速率。

基于这些原因，通过选择单一全局分辨率层级实现自适应性，对于估计 $T_r(f)$而言既充分又最优，而扩展到带有$p \neq 2$ 的 *Besov* 类则超出了当前框架的范围。

与其扩展 *Sobolev* 泛函 $T_r(f)$ 本身，不如扩展到 *Besov* 型泛函，这将是未来研究的一个自然方向，并且需要根本不同的估计技术。

5. 数值示例研究

本节通过数值示例展示前几节得出的理论结果，特别关注截断所引起的偏差–方差权衡以及预言机分辨率层级的行为。我们并不旨在对基于 needlet 的估计量进行全面的蒙特卡洛研究，因为这会掩盖其内在机制；相反，我们采用了一个刻意简化的数值设定，以反映命题 3.3.

特别是，所有数值实验均基于理论推导出的渐近偏差–方差表达式，旨在展示定性尺度变化规律，而非完整实现的 needlet 估计量的有限样本性能。

这种选择使我们能够分离出分辨率参数的作用，并以直观的方式展示截断偏差与随机变异性在不同尺度上的相互作用。

因此，数值实验旨在验证理论的定性预测，而非评估有限样本最优性。具体而言，实验展示了预言机分辨率层级随样本量呈对数增长，并突显不同光滑度区间会导致截然不同的截断行为。这一视角在当前设定下尤其重要，因为分辨率层级充当全局正则化参数，在预言机估计和自适应估计中都发挥着核心作用。

5.1. 预言机分辨率与偏差–方差权衡

我们对截断 Sobolev 泛函估计中的偏差–方差权衡，以及命题 3.3 所预测的预言机分辨率层级，给出数值示例。我们不模拟完整的 needlet 展开，而是聚焦于理论分析中推导出的渐近偏差–方差模型；该模型刻画了均方误差随分辨率层级变化的主导尺度行为，足以展示预言选择的定性特征。

更确切地说，对于固定参数（d， r， s）和样本量 n，我们考虑如下分解

$$\mathrm{MSE}(J) = \mathrm{Bias}^2(J) + \mathrm{Var}(J), \qquad \mathrm{Bias}^2(J) \asymp B^{-4J(s-r)}, \qquad \mathrm{Var}(J) \asymp \frac{B^{J(d+4r)}}{n},$$

并研究 MSE(J) 随分辨率层级 J 变化的规律。预言机分辨率层级

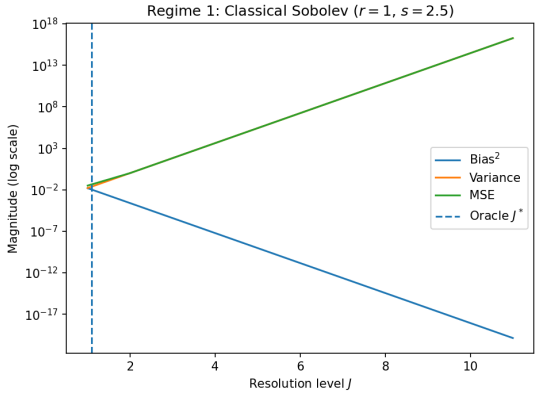
$$J^\star \asymp \frac{\log n}{(2s+d+4r)\log B}$$

对应于截断偏差与随机方差贡献之间的平衡点。

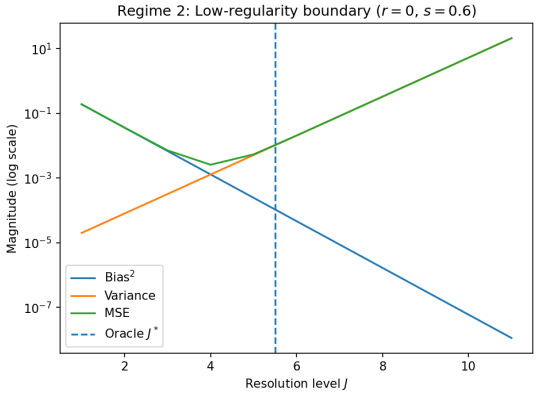
总体而言，这些数值示例证实了理论分析所预测的定性规律。特别是，它们表明，预言机分辨率层级仅随样本量对数增长，在经典 Sobolev 情形下可能保持较小，而在正则性较低的情形下则会移至中间尺度。这一观察进一步说明了自适应选择程序的必要性。下一节将介绍这类程序，它们能够根据底层密度未知的光滑度自动进行调整。

5.2. 预言机风险与自适应风险：数值证据

我们给出定理4.3的数值示例，重点比较预言估计量及其自适应对应估计量。数值实验旨在用于诊断，而非全面的实证研究，以展示理论所预测的渐近偏差—方差机制。因此，我们不模拟完整的 needlet 展开，而是采用一种直接基于命题3.3中推导出的理论偏差—方差分解的简化数值框架。这种方法能够分离分辨率选择的影响，并清晰地展示预言与自适应行为。



经典 Sobolev 制度（$d = 2$, $r = 1$, $s = 2.5$, $n = 4000$）。



低正则性边界制度（$d = 2$, $r = 0$,$s = 0.6$, $n = 2 \times 10^5$）。

图 2：两种索博列夫情形下的偏差–方差权衡与预言机分辨率层级。左：中等光滑度的经典制度，其中预言机分辨率仍接近最粗尺度。右：接近可估性边界的低正则性制度，其中预言机向中间尺度偏移。在两种情形下，预言机行为均反映了截断偏差与随机方差之间的平衡，正如命题3.3所预测的那样。

对于（d, r）的固定取值以及若干索博列夫光滑度层级 s，我们考虑渐近均方误差模型

$$\mathrm{MSE}(J) = B^{-4J(s-r)} + \frac{B^{J(d+4r)}}{n},$$

并同时计算预言机分辨率层级 $J^\star$ 以及离散 Lepski 型自适应分辨率 $\widehat{J}$。预言机分辨率平衡偏差项与方差项，而自适应分辨率则通过比较相邻分辨率层级得到，以模拟第4节引入的选择规则。

图3展示了相关结果。左图展示了偏差–方差权衡随分辨率层级的变化，以及预言机分辨率 $J^\star$。右图比较了预言机风险和自适应风险随样本量 n的变化。在所考虑的所有光滑度区间内，自适应估计量都紧密跟踪预言机风险，并呈现相同的衰减速率，这与定理4.3一致。

表1报告了预言机和自适应分辨率及风险的数值结果。这些结果证实，自适应过程无需预先了解光滑度参数，即可在相差一个乘法常数的范围内重现预言机的表现。

6. 证明

本节汇集了本文主要结果的证明。我们首先介绍联系 Sobolev 导数与 needlet 系数的预备恒等式，然后证明预言机偏差–方差界和自适应风险结果。下文中，常数可能随行变化，且与 n 及分辨率层级无关。

引理 *2.2* 的证明。 needlet 系统的紧框架性质意味着任意 $g \in$

表 1： 不同光滑度参数和样本量下的预言机与自适应分辨率层级及风险

$\widehat{s}$ n自适应分辨率 J 紧密跟随预言机分辨率 $J^\star$，相应的风险处于同一数量级，这与定理 4.3一致。

s	n	$J^\star$	$\widehat{J}$	预言机风险	自适应风险
2.2	1000	1	1	1.94e–02	2.11e–02
2.2	3000	1	1	7.42e–03	7.85e–03
2.2	8000	2	2	2.99e–03	3.02e–03
2.2	20000	2	2	1.18e–03	1.21e–03
2.6	1000	1	1	8.91e–03	9.32e–03
2.6	3000	1	1	3.21e–03	3.38e–03
2.6	8000	2	2	1.15e–03	1.17e–03
2.6	20000	2	2	4.36e–04	4.51e–04
3.0	1000	1	1	4.30e–03	4.59e–03
3.0	3000	1	1	1.45e–03	1.53e–03
3.0	8000	2	2	4.89e–04	5.02e–04
3.0	20000	2	2	1.79e–04	1.87e–04

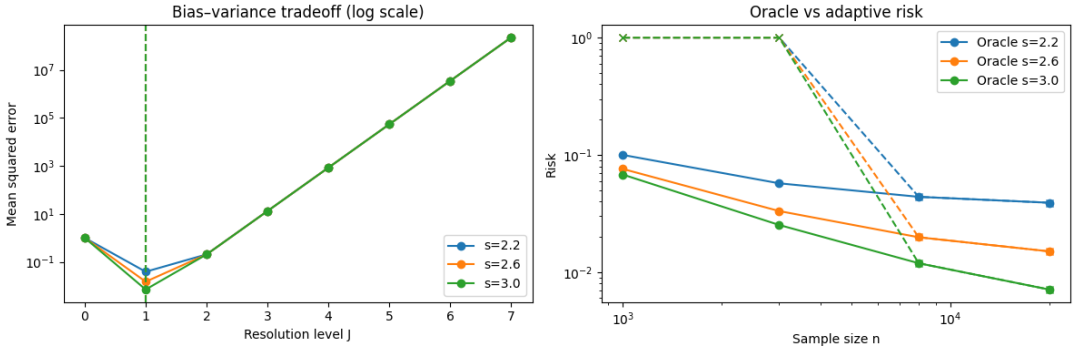


图 3： 定理 4.3 的数值示例。左：不同索博列夫情形下的偏差–方差权衡和预言机分辨率层级（见图 2）。右：不同索博列夫光滑度层级下，预言机风险与自适应风险随样本量变化的关系。

$L^2(\mathbb{S}^d)$ 可展开为

$$g = \sum_{j,k} \langle g, \psi_{j,k} \rangle_{L^2(\mathbb{S}^d)} \, \psi_{j,k},$$

并在 $L^2(\mathbb{S}^d)$ 中收敛。将此恒等式应用于 $g = f^{(r)}$，得到

$$f^{(r)} = \sum_{j,k} \beta_{j,k}^{(r)} \, \psi_{j,k}, \qquad \beta_{j,k}^{(r)} = \langle f^{(r)}, \psi_{j,k} \rangle_{L^2(\mathbb{S}^d)}.$$

由于拉普拉斯–贝尔特拉米算子在 $L^2(\mathbb{S}^d)$ 上是自伴的，

$$\beta_{j,k}^{(r)} = \langle (-\Delta_{\mathbb{S}^d})^{r/2} f, \psi_{j,k} \rangle_{L^2(\mathbb{S}^d)} = \langle f, (-\Delta_{\mathbb{S}^d})^{r/2} \psi_{j,k} \rangle_{L^2(\mathbb{S}^d)} = \langle f, \psi_{j,k}^{(r)} \rangle_{L^2(\mathbb{S}^d)},$$

由此得到第二种表示形式。

利用窗函数 b 的谱局部化性质以及 $b(\ell/B^j)$ 的支撑集为 $\ell \in [B^{j-1}, B^{j+1}]$ 这一事实，特征值在支撑集上满足 $e_\ell^{r/2}$，$d \asymp B^{jr}$。将这一项因式分解出来，即可得到所述的系数界。 □

命题 3.3 的证明。我们将均方误差分解为

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(J)} - T_r(f)\right)^2\right] = \left(T_r(f) - T_r^{(J)}(f)\right)^2 + \mathrm{Var}\left(\widehat{T}_r^{(J)}\right),$$

其中偏差项是纯解析的，而方差项取决于经验 Sobolev–needlet 系数的随机行为。

偏差。 根据截断泛函的定义，

$$T_r(f) - T_r^{(J)}(f) = \sum_{j>J} \sum_{k=1}^{N_j} \left(\beta_{j,k}^{(r)}\right)^2.$$

假设 $f \in H^s(\mathbb{S}^d)$ 对某个 $s > r$。由引理 2.2以及 Sobolev 范数与加权 needlet 能量之间的等价性，我们有

$$\sum_{j,k} B^{2js}\, \beta_{j,k}^2 < \infty, \qquad |\beta_{j,k}^{(r)}| \asymp B^{jr} |\beta_{j,k}|.$$

因此，

$$\begin{aligned} T_r(f) - T_r^{(J)}(f) &\asymp \sum_{j>J} \sum_{k} B^{2jr}\, \beta_{j,k}^2 \\ &= \sum_{j>J} \sum_{k} B^{-2j(s-r)} \left(B^{2js} \beta_{j,k}^2\right) \\ &\lesssim \sum_{j>J} B^{-2j(s-r)} \lesssim B^{-2J(s-r)}, \end{aligned}$$

其中我们用到了 $\sum_k B^{2js} \beta_{j,\,k}^2 \le C_s$ 关于 $j \ge 0$ 对 $f \in H^s(\mathbb{S}^d)$ 一致成立。由此得到所述的截断偏差界。

方差。 回顾该估计量定义为

$$\widehat{T}_r^{(J)} = \sum_{j \le J} \sum_{k=1}^{N_j} \widehat{\beta}_{j,k;(1)}^{(r)}\, \widehat{\beta}_{j,k;(2)}^{(r)},$$

其中两组经验系数分别由相互独立的子样本计算得到。根据子样本之间的独立性，

$$\mathrm{Var}\left(\widehat{T}_r^{(J)}\right) = \sum_{j \le J} \sum_{k=1}^{N_j} \mathrm{Var}\left(\widehat{\beta}_{j,k;(1)}^{(r)} \widehat{\beta}_{j,k;(2)}^{(r)}\right).$$

这源于小波包窗函数的谱局部化以及 Laplace–Beltrami 特征值 $e_{\ell,d} \asymp \ell^2$ 的渐近行为；它们意味着，r 阶微分会将二进频带 $\ell \asymp B^j$ 中的频率放大 B^{jr} 倍。根据小波包窗函数的谱局部化，

$$\|\psi_{j,k}^{(r)}\|_{L^2(\mathbb{S}^d)}^2 \asymp B^{2jr}.$$

对于固定的 (j,k)，由独立性可得

$$\mathrm{Var}\left(\widehat{\beta}_{j,k;(1)}^{(r)} \widehat{\beta}_{j,k;(2)}^{(r)}\right) = \mathrm{Var}\left(\widehat{\beta}_{j,k;(1)}^{(r)}\right) \mathrm{Var}\left(\widehat{\beta}_{j,k;(2)}^{(r)}\right) + \beta_{j,k}^{(r)\,2} \left[\mathrm{Var}\left(\widehat{\beta}_{j,k;(1)}^{(r)}\right) + \mathrm{Var}\left(\widehat{\beta}_{j,k;(2)}^{(r)}\right)\right].$$

由于每个子样本的大小为 $n/2$且

$$\widehat{\beta}^{(r)}_{j,k;(a)} = \frac{2}{n} \sum_{X_i \in \mathcal{D}_a} \psi^{(r)}_{j,k}(X_i),$$

我们得到

$$\mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(a)}\right) = \frac{2}{n} \int_{\mathbb{S}^d} \left|\psi^{(r)}_{j,k}(x)\right|^2 f(x)\,\mathrm{d}x \lesssim \frac{1}{n} \left\|\psi^{(r)}_{j,k}\right\|^2_{L^2(\mathbb{S}^d)}.$$

结合上述界并对 (j,k) 求和，得到

$$\mathrm{Var}\left(\widehat{T}^{(J)}_r\right) \lesssim \frac{1}{n} \sum_{j \le J} \sum_{k=1}^{N_j} B^{4jr} \asymp \frac{1}{n} \sum_{j \le J} B^{j(d+4r)}.$$

由于 $N_j \asymp B^{dj}$，该和由其最大项主导，因此得到

$$\mathrm{Var}\left(\widehat{T}^{(J)}_r\right) \asymp \frac{B^{J(d+4r)}}{n}.$$

结合偏差界和方差界，我们得到

$$\mathbb{E}_f\left[\left(\widehat{T}^{(J)}_r - T_r(f)\right)^2\right] \le C\left(B^{-4J(s-r)} + \frac{B^{J(d+4r)}}{n}\right),$$

其中常数 $C > 0$ 仅依赖于 r、 s、 d 和小波包窗函数。
平衡这两项可得预言机分辨率层级 $B^{J^\star} \asymp n^{1/(2s+d+4r)}$ 及其对应的速率 $n^{-4(s-r)/(2s+d+4r)}$，证明完毕。 □

引理 *4.1* 的证明。 固定 $J\prime > J$。根据估计量的定义及线性性，

$$\widehat{T}^{(J')}_r - \widehat{T}^{(J)}_r = \sum_{j=J+1}^{J'} \sum_{k=1}^{N_j} \widehat{\beta}^{(r)}_{j,k;(1)}\, \widehat{\beta}^{(r)}_{j,k;(2)}.$$

由于估计量 $\widehat{T}_{(J)r}$是 $T_{(J)r}(f)$ 的无偏估计量，因此差值 $\widehat{T}_{(J')_r} - \widehat{T}^{(J)r}$是 $T_{(J')r}(f) - T_{(J)r}(f)$ 的无偏估计量。因此，

$$\mathbb{E}_f\left[\left(\widehat{T}^{(J)}_r - \widehat{T}^{(J')}_r\right)^2\right] = \mathrm{Var}\left(\widehat{T}^{(J')}_r - \widehat{T}^{(J)}_r\right).$$

由于两个子样本 $\mathcal{D}_1$ 和 $\mathcal{D}_2$, 相互独立，且由 needlet 框架的谱局部化所导致的不同分辨率层级之间具有准正交性，方差可分解为

$$\mathrm{Var}\left(\widehat{T}^{(J')}_r - \widehat{T}^{(J)}_r\right) = \sum_{j=J+1}^{J'} \sum_{k=1}^{N_j} \mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(1)}\, \widehat{\beta}^{(r)}_{j,k;(2)}\right).$$

对于每个固定的 $(j,\ k)$，子样本之间的独立性意味着

$$\begin{aligned}\mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(1)}\, \widehat{\beta}^{(r)}_{j,k;(2)}\right) &= \mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(1)}\right) \mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(2)}\right) \\ &\quad + {\beta^{(r)}_{j,k}}^2 \left[\mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(1)}\right) + \mathrm{Var}\left(\widehat{\beta}^{(r)}_{j,k;(2)}\right)\right].\end{aligned}$$

每个经验系数均由大小为 $n/2$的子样本计算得出，因此

$$\widehat{\beta}^{(r)}_{j,k;(a)} = \frac{2}{n} \sum_{X_i \in \mathcal{D}_a} \psi^{(r)}_{j,k}(X_i), \qquad a = 1, 2.$$

因此，

$$\operatorname{Var}\Big(\widehat{\beta}^{(r)}_{j,k;(a)}\Big) = \frac{2}{n} \int_{\mathbb{S}^d} \Big|\psi^{(r)}_{j,k}(x)\Big|^2 f(x)\,\mathrm{d}x \lesssim \frac{1}{n} \|\psi^{(r)}_{j,k}\|^2_{L^2(\mathbb{S}^d)}.$$

由小波包窗函数的谱局部化以及 $\psi^{(r)}_{j,k} = (-\Delta_{\mathbb{S}^d})^{r/2}\psi_{j,k}$ 的定义，

$$\|\psi^{(r)}_{j,k}\|^2_{L^2(\mathbb{S}^d)} \asymp B^{2jr},$$

对 j 和 k 一致成立。又利用 $|\beta^{(r)}_{j,k}|^2 \le T_r(f)$，并将常数吸收后，得到

$$\operatorname{Var}\Big(\widehat{\beta}^{(r)}_{j,k;(1)}\,\widehat{\beta}^{(r)}_{j,k;(2)}\Big) \lesssim \frac{B^{4jr}}{n^2} + \frac{B^{2jr}}{n}.$$

对 $k = 1, \dots, N_j$ 求和，并回忆 $N_j \asymp B^{dj}$，

$$\sum_{k=1}^{N_j} \operatorname{Var}\Big(\widehat{\beta}^{(r)}_{j,k;(1)}\,\widehat{\beta}^{(r)}_{j,k;(2)}\Big) \lesssim \frac{B^{j(d+4r)}}{n^2} + \frac{B^{j(d+2r)}}{n}.$$

对所有 $n \ge 1$ 和 $j \ge 0$，第二项占主导。因此，

$$\operatorname{Var}\Big(\widehat{T}^{(J')}_r - \widehat{T}^{(J)}_r\Big) \lesssim \frac{1}{n} \sum_{j=J+1}^{J'} B^{j(d+2r)}.$$

由于该和由其中最大的一项主导，

$$\sum_{j=J+1}^{J'} B^{j(d+2r)} \lesssim B^{J'(d+2r)},$$

我们得出结论，

$$\mathbb{E}_f\Big[\Big(\widehat{T}^{(J)}_r - \widehat{T}^{(J')}_r\Big)^2\Big] \le C_0^2\,\frac{B^{J'(d+4r)}}{n},$$

对于仅依赖于 r、d 以及小波包窗函数的适当常数 $C_0 > 0$。取平方根即得到所述的波动率 $n^{-1/2}B^{J'(d/2+2r)}$。 □

^引理 *4.2* 的证明。 根据 Lepski 选择器的定义（13），事件 $\{J > J^\star\}$ 意味着存在 $J' > J^\star$ 使得

$$\Big|\widehat{T}^{(J^\star)}_r - \widehat{T}^{(J')}_r\Big| > \omega(J').$$

根据引理 4.1，随机变量 $\widehat{T}^{(J^\star)}_r - \widehat{T}^{(J')}_r$ 的方差阶为 $n^{-1}B^{J'(d+4r)}$，并且可以表示为经验 Sobolev–needlet 系数的独立乘积的有限和。应用针对具有有界核的退化二次 U 统计量的 Bernstein 型浓缩不等式（例如参见 [15]），对于每个固定的 $J' > J^\star$，

$$\mathbb{P}_f\Big(\Big|\widehat{T}^{(J^\star)}_r - \widehat{T}^{(J')}_r\Big| > \omega(J')\Big) \le 2\exp(-c\,\log n)\,,$$

对某个常数 $c > 0$。

^根据事件 $\{J > J^\star\}$ 的定义，

$$\{\widehat{J} > J^\star\} \subset \bigcup_{J' \in \mathcal{J}_n : J' > J^\star} \left\{ \left| \widehat{T}_r^{(J^\star)} - \widehat{T}_r^{(J')} \right| > \omega(J') \right\}.$$

由于 $\mathcal{J}_n$ 的基数为 $\log n$ 量级，由并集界可得

$$\mathbb{P}_f\left(\widehat{J} > J^\star\right) \lesssim (\log n) \exp(-c \log n) \le c_1 \exp(-c_2 \log n),$$

对于适当的常数 $c_1, c_2 > 0$。 □

定理 *4.3* 的证明。 固定 $r \ge 0$，并令 $s \in [r+\varepsilon, s_{\max}]$。 令 $J^\star$ 表示如下定义的预言机分辨率层级

$$B^{J^\star} \asymp n^{1/(2s+d+4r)},$$

它平衡了命题 3.3 中的偏差项和方差项。

我们将自适应风险分解为

$$\mathbb{E}_f\left[\left(\widetilde{T}_r - T_r(f)\right)^2\right] = \mathbb{E}_f\left[\left(\widehat{T}_r^{(\widehat{J})} - T_r^{(J^\star)}(f)\right)^2\right] + \left(T_r^{(J^\star)}(f) - T_r(f)\right)^2.$$

偏差项。 根据命题 3.3，由于 $f \in H^s(\mathbb{S}^d)$ 且 $s > r$，

$$\left(T_r^{(J^\star)}(f) - T_r(f)\right)^2 \lesssim B^{-4J^\star(s-r)} \asymp n^{-4(s-r)/(2s+d+4r)}.$$

随机项。 记作

$$\widehat{T}_r^{(\widehat{J})} - T_r^{(J^\star)}(f) = \left(\widehat{T}_r^{(\widehat{J})} - \widehat{T}_r^{(J^\star)}\right) + \left(\widehat{T}_r^{(J^\star)} - T_r^{(J^\star)}(f)\right).$$

由初等不等式 $(a+b)^2 \le 2a^2 + 2b^2$，只需分别对右侧的两项进行界定。

与预言机尺度的偏离。 根据 Lepski 选择器的定义（13），在事件 ^$\{J \le J^\star\}$ 上，我们有

$$\left| \widehat{T}_r^{(\widehat{J})} - \widehat{T}_r^{(J^\star)} \right| \le \omega(J^\star) = C_0\, n^{-1/2} B^{J^\star(d/2+2r)}.$$

^在互补事件 $\{J > J^\star\}$上，Lepski 型程序的标准论证表明，该事件发生的概率呈指数级小（例如参见 [16, 20]）。利用 ^Lepski 选择器的定义，并结合引理 $\{J > J^\star\}$给出的事件 4.2的指数级控制，可得

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(\widehat{J})} - \widehat{T}_r^{(J^\star)}\right)^2\right] \lesssim \omega(J^\star)^2 \asymp \frac{B^{J^\star(d+4r)}}{n}.$$

预言机尺度下的方差。根据命题3.3，可得

$$\mathbb{E}_f\left[\left(\widehat{T}_r^{(J^\star)} - T_r^{(J^\star)}(f)\right)^2\right] = \mathrm{Var}\left(\widehat{T}_r^{(J^\star)}\right) \asymp \frac{B^{J^\star(d+4r)}}{n}.$$

结合上述界可得

$$\mathbb{E}_f\Big[\big(\widetilde{T}_r - T_r(f)\big)^2\Big] \lesssim B^{-4J^\star(s-r)} + \frac{B^{J^\star(d+4r)}}{n}.$$

通过定义 $J^\star$ 平衡这两项可得

$$\mathbb{E}_f\Big[\big(\widetilde{T}_r - T_r(f)\big)^2\Big] \lesssim n^{-4(s-r)/(2s+d+4r)},$$

在 $s \in [r+\varepsilon, s_{\max}]$ 上一致成立。至此，证明完成。 □

资助

作者部分获得了 PRIN2022-GRAFIA-202284Z9E4 和萨皮恩扎大学项目 RM123188F69A66C1（2023）、RG1241907D2FF327（2024）的资助。

参考文献

[1] K. Atkinson 和 W. Han。《单位球面上的球谐函数与逼近》。施普林格，2012 年。[2] P. Baldi、G. Kerkyacharian、D. Marinucci 和 D. Picard。使用 Needlet 对方向数据进行自适应密度估计。《统计年刊》，37(6B):3362–3395，2009 年。[3] J. A. Barceló、T. Luque 和 S. Pérez-Esteva。球面上 Sobolev 空间的刻画。《数学分析及其应用杂志》，491(1):124240，2020 年。[4] P. J. Bickel 和 Y. Ritov。密度导数平方积分的估计：最优收敛速率的精确估计。*Sankhy¯a Ser. A*，50(3):381–393，1988 年。[5] L. Birgé 和 P. Massart。从模型选择到自适应估计。载于纪念 *Lucien Le Cam* 文集，第 55–87 页。施普林格，1997 年。[6] S. Bourguin 和 C. Durastanti。紧流形上 Besov 空间中 U-统计量的高频极限。《伊利诺伊数学杂志》，61(1–2):97–125，2017 年。[7] J. S. Brauchart 和 J. Dick。球面上 Sobolev 空间的刻画，以及 Stolarsky 不变性原理向任意光滑度的推广。《构造逼近》，38:397–445，2013 年。[8] T. T. Cai 和 M. G. Low。二次泛函的非二次估计量。《统计年刊》，33(6):2930–2956，2005 年。[9] D. L. Donoho、I. M. Johnstone、G. Kerkyacharian 和 D. Picard。通过小波阈值化进行密度估计。《统计年刊》，24(2):508–539，1996 年。[10] C. Durastanti。球面上密度估计的全局阈值化。《多元分析杂志》，147:80–95，2016 年。[11] C. Durastanti。灵活带宽小波构造中的尺度膨胀动力学。arXiv:2507.05075，2025 年。

[12] C. Durastanti、D. Geller 和 D. Marinucci。自旋纤维丛上的自适应非参数回归。《多元分析杂志》，104(1):16–38，2012 年。[13] C. Durastanti、D. Marinucci 和 A. P. Todino。球面上的灵活带宽小波。伯努利，30(1):22–45，2024 年。[14]C. Durastanti 和 R. Shevchenko。球面上基于小波的非参数异方差方差估计的自适应阈值处理。arXiv 预印本 arXiv:2601.03920 [math.ST], 2026。[15] E. Giné 和 R. Nickl。密度函数积分平方的一个简单自适应估计量。伯努利，14(1):47–61，2008 年。[16] A. Goldenshluger 和 O. Lepski。核密度估计中的带宽选择：预言机不等式与自适应极小极大最优性。《统计年刊》，39(3):1608–1632，2011 年。[17] E. Hebey。《黎曼流形上的 Sobolev 空间》。施普林格，1996 年。[18] I. A. Ibragimov 和 R. Z. Khas'minskii。高斯白噪声中线性泛函值的非参数估计。《概率论及其应用》，29(1):18–32，1985 年。[19] B. Laurent。密度积分泛函的有效估计。《统计年刊》，24(2):659–681，1996 年。[20] O. V. Lepskiĭ。渐近极小极大自适应估计。I. 上界。最优自适应估计量。《概率论及其应用》，36(4):682–697，1991 年。[21] C. Ley 和 T. Verdebout。《现代方向统计学》。Chapman & Hall/CRC，2017 年。[22] D. Marinucci 和 G. Peccati。《球面上的随机场》。剑桥大学出版社，2011 年。[23] J.-B. Monnier。均匀设计下超球面上的非参数回归。《*TEST*》，20:412–446，2011 年。[24] F. Narcowich、P. Petrushev 和 J. Ward。球面上 Besov 和 Triebel–Lizorkin 空间的分解。《泛函分析杂志》，238(2):530–564，2006 年。[25] F. J. Narcowich、P. Petrushev 和 J. D. Ward。球面上的局部化紧框架。《*SIAM*数学分析杂志》，38(2):574–594，2006 年。[26] Alexandre B. Tsybakov。《非参数估计导论》。施普林格，纽约，2009 年。