arXiv:2603.06027v1 [cs.LG] 6 Mar 2026

通过高斯表面积实现近乎最优时间的无关学习

Lucas Pesenti*
lucas.pesenti@inf.ethz.ch苏黎世联邦理工学院

Lucas Slot†
l.f.h.slot@uva.nl阿姆斯特丹大学

Manuel Wiedmer*
manuel.wiedmer@inf.ethz.ch苏黎世联邦理工学院

2026年3月6日

摘要

在困难的不可知模型中，于高斯边缘分布下学习概念类的复杂度，与其被低次多项式进行 L_1-逼近的能力密切相关。对于任意高斯表面积至多为 Γ的概念类，Klivans等人（2008）证明，次数 $d = O(\Gamma^2/\varepsilon^4)$足以实现 ε-逼近。由此得到了学习多种概念类的复杂度方面目前已知的最佳界。在本注记中，我们改进了他们的分析，证明次数 $d = \tilde{O}(\Gamma^2/\varepsilon^2)$即已足够。结合Diakonikolas等人（2021）给出的下界，这给出了在统计查询模型中不可知学习多项式阈值函数的复杂度的（近）最优界。我们的证明依赖于Feldman等人（2020）构造的直接类比，他们考虑了布尔超立方体上的 L_1-逼近。

1 引言

[KSS92] 的无关学习框架将 [Val84] 的经典PAC模型扩展到含噪输入数据的学习任务，并成为研究这类任务计算复杂性的标准模型之一。它可以描述如下：设 C 为一个概念类（即一组函数 $f: \mathcal{X} \to \{-1, 1\}$），且 $\mathcal{D}$ 为定义在 $\mathcal{X} \times \{-1, 1\}$上的分布。给定从 $\mathcal{D}$中独立抽取的带标签样本$\hat{}$ (x, y)，我们的任务是输出一个假设 f，用于预测从 $\mathcal{D}$中抽取的未来样本的标签。

与PAC模型不同，我们不假设样本 x 与其标签 y之间存在任何关系。因此，通常不可能找到具有非平凡成功率的假设。相反，我们要求$\hat{}$ f 预测标签的效果几乎与 C中最佳概念一样好。更正式地说，如果一个算法以无关方式学习 C，且超额误差至多为 $\varepsilon > 0$，那么我们称它能够学习该概念类；也就是说，它以高概率输出一个假设$\hat{}$ $\hat{}$ f，满足 $\mathbb{P}_{(x,y)\sim\mathcal{D}}(f(x) \neq y) \leq \text{opt} + \varepsilon$，其中 $\text{opt} := \min_{f \in C} \mathbb{P}_{(x,y)\sim\mathcal{D}}(f(x) \neq y)$。参数 opt 衡量问题的“噪声程度”；如果 $\text{opt} = 0$，我们就恢复了PAC模型。

L_1多项式回归算法

在一般分布下进行无关学习的高效算法不太可能存在，即使对于半空间这类简单的概念类也是如此 [Dan16，Tie23]。因此，自然的做法是将范围限制在某些

由瑞士国家科学基金会（SNSF）资助，资助编号：10004947。
+由项目“平方和算法基础：最坏情况及其他”提供支持，该项目属于 Veni ENW 研究计划，项目编号为VI.Veni.242.234，部分经费由荷兰研究理事会（NWO）通过以下资助提供：https://doi.org/10.61686/GIEOW65108。

分布类上。在这里，我们关注这样的设定：$\mathcal{X} = \mathbb{R}^n$ 且样本的边缘分布$\mathcal{D}_{\mathcal{X}} = \mathcal{N}(0,\ I_n)$是标准高斯分布。

在其开创性工作中，[KKMS08] 首次给出了在高斯分布（以及其他一些分布）下不可知学习半空间的高效算法。该算法计算带标签样本的最佳d 次多项式逼近，其度量采用 L_1-范数。这可以在 $n^{O(d)}$时间内完成，例如通过线性规划。假设C 中的每个概念都具有d 次逼近，且 L_1-误差至多为 ε，则该过程（结合阈值化方案）为 C 给出了一个超额误差为 ε的无关学习器。[1] 如果 $d = d(\varepsilon)$ 可以独立于维数 n 进行选择，则可得到一个运行时间为 $n^{O(d)} \cdot \mathrm{poly}(1/\varepsilon)$的多项式时间近似方案。

该 L_1次多项式回归算法仍然是在分布假设下进行高效不可知学习的标准（事实上也是唯一实际采用的）选择。对于各种概念类和分布，它都达到了已知的最佳运行时间 [KKMS08、KOS08、DKN10、DHK+10、BOW10、Kan11a、KKM13、DKR25、CPS26]。事实上，L_1次多项式回归已知在高斯分布下进行不可知学习的统计查询（SQ）模型中（本质上）是最优的 [DKPZ21]。

L1 逼近的次数界与高斯表面积

的结果 [KKMS08] 和 [DKPZ21] 表明，在高斯分布下不可知学习一个概念类的复杂度，取决于它能否由低次数多项式近似：如果 L_1：令 d 为满足以下条件的最小次数：每个 $f \in C$ 都存在一个 ε-逼近，在 L_1 中的次数为 d，那么不可知学习 C 至误差 ε 的 SQ 复杂度约为 $n^{\Theta(d)}$。

对于半空间，[KKMS08] 证明，次数为 $d = O(1/\varepsilon^4)$ 足以实现一个 ε-逼近，在 L_1中。他们的证明首先给出了在 ε-范数下对半空间进行 L_2-逼近所需次数的界；随后可利用柯西-施瓦茨不等式将其转化为 L_1-保证。采用这条间接途径的原因在于，有强大的工具可用于分析 L_2中的逼近，而这些工具在 L_1情形下并不可用，其中最著名的就是 Hermite分析。

为了将这一方法推广到任意概念类，[KOS08] 考察了概念 f的高斯表面积（GSA），它衡量集合 $\{x \in \mathbb{R}^n\colon\ f(x) = 1\}$ 相对于高斯密度函数的表面积。通过 Hermite 分析，他们证明，任何元素的 GSA 至多为 Γ 的概念类，都可以用次数为 $d = O(\Gamma^2/\varepsilon^4)$ 的多项式在 L_2-范数下进行 ε 近似；因此，$d = O(\Gamma^2/\varepsilon^4)$ 在 L_1 中也足够。半空间、欧几里得球和某些锥的 GSA 均为常数 [KOS08]；k 个半空间的交集的 GSA 至多为 $O(\sqrt{\log k})$ [KOS08]；次数为 k 的多项式阈值函数（PTF）的 GSA 至多为$O(k)$ [Kan11a]；凸集的 GSA 至多为 $O(n^{1/4})$ [Bal93]。

L[1]逼近的最优次数界

该方法的一个关键优势 [KOS08] 在于，它确立了 GSA 是无关学习复杂度的通用上界。然而，他们的分析在逼近半空间（从而也在学习半空间）时给出的保证并非最优：通过直接构造 [DKN10] 可知，只需次数为 $d = O(1/\varepsilon^2)$，即可获得半空间的 ε逼近（相对于 L_1范数）（而$d = O(1/\varepsilon^4)$是 [KKMS08和KOS08]得到的）。

.....的界 [DKN10] 在两方面都是最优的：一方面，不可能得到更好的 L_1近似 [Gan08]，；另一方面，相应的 L_1多项式回归算法的运行时间

[1]这里选择 L_1范数至关重要：早期计算最佳 L_2逼近的算法，得到的假设其（总）误差为 $O(\mathrm{opt}) + \varepsilon$。这种保证较弱，只能容忍少量噪声。关于这一情形下的算法结果，可参见 [KLS09、ABL14、Dan15、DKS18、DKTZ20]。

与 [DKPZ21] 在统计查询模型中的下界相匹配（相差$1/\varepsilon$中的多项式因子）。另一方面，他们的构造（显然）无法推广到半空间之外（见第 4.2 节）。

这引出了一个问题：是否能够证明一种既具有普适性（适用于所有高斯表面积有界的概念）又具有最优性（就其对 ε 的依赖而言）的近似保证。在本文中，我们对这一问题给出了（几乎）肯定的回答：我们改进了 [KOS08] 的分析，证明次数为 $d = \tilde{O}(\Gamma^2/\varepsilon^2)$ 即足以实现对 GSA 至多为 Γ 的任意概念类的 ε- 近似（定理 1.1 ）。这恢复了半空间的（最优） [DKN10] 界（相差一个 $\log(1/\varepsilon)$ 因子）。此外，我们将次数为 k 多项式阈值函数（PTFs）的已知最佳界从 $O(k^2/\varepsilon^4)$ 改进至 $\tilde{O}(k^2/\varepsilon^2)$，这与 [DKPZ21] 给出的 $\Omega(k^2/\varepsilon^2)$ 下界（几乎）相匹配。最后，我们将半空间交集和凸集的已知最佳界提高了 $\tilde{\Omega}(1/\varepsilon^2)$ 倍。概览请参见表 1。

我们的技术贡献很小：我们证明， [FKV20] 关于在布尔超立方体上使用噪声算子的构造，可以经过适当调整后推广到高斯情形。然后，我们利用 [KOS08]的结果，以高斯表面积为依据界定由此得到的 L_1-近似误差，并据此得出结论。因此，我们证明中的所有要素在文献中或多或少都已为人所知，只是据我们所知，尚未在这一背景下组合起来。

概念类	**LB**	**上界** （前一项）	**上界** （定理 1.1）
半空间	$\Omega(1/\varepsilon^2)$ [Gan08]	$O(1/\varepsilon^2)$ [DKN10]	$\tilde{O}(1/\varepsilon^2)$
k 次多项式阈值函数	$\Omega(k^2/\varepsilon^2)$ [DKPZ21]	$O(k^2/\varepsilon^4)$ [Kan11a]	$\tilde{O}(k^2/\varepsilon^2)$
半空间交集 k	$\tilde{\Omega}(\sqrt{\text{对数}\, k/\varepsilon})$ [DKPZ21]	O（对数 k/ε^4） [KOS08]	$\tilde{O}(\log\, k/\varepsilon^2)$
凸集	—	$O(\sqrt{n}/\varepsilon^4)$ [KOS08]	$\tilde{O}(\sqrt{n}/\varepsilon^2)$
GSA 至多 Γ	—	$O(\Gamma^2/\varepsilon^4)$ [KOS08]	$\tilde{O}$（Γ^2/ε^2）

表 1：对概念类 C 在 $L_1(\mathcal{N}_n)$ 中实现 ε-近似所需最小次数的上界（上界）和下界（下界）。这些分别对应于 $n^{O(d)}$ 和 $n^{\Omega(d)}$ 中关于不可知学习 C 直至误差 ε 的（SQ-）复杂度的上界和下界。

1.1 主要结果

为陈述我们的主要结果，我们需要介绍两个与布尔函数相关的（标准）概念。首先，函数 $f\colon \mathbb{R}^n \to \{\pm 1\}$ 在 $\rho \in [0,1]$ 级别的高斯噪声敏感度 $\mathrm{GNS}_\rho(f)$，是指两个 $(1-\rho)$-相关高斯随机变量 $X, Y \sim \mathcal{N}_n$ 满足 $f(X) \neq f(Y)$ 的概率（见定义 2.2）。其次，记 $K(f) = \{x \in \mathbb{R}^n\colon f(x) = 1\}$，则 f 的高斯表面积为 [2]

$$\mathrm{GSA}(f) := \lim_{\delta \to 0} \frac{\mathrm{vol}_{\mathcal{N}_n}(K(f)_\delta \setminus K(f))}{\delta},$$

其中 $A_\delta := \{x \in \mathbb{R}^n : \mathrm{dist}(x,\ A) \le \delta\}$ 是 δ-邻域 $A \subseteq \mathbb{R}^n$，且 $\mathrm{vol}_{\mathcal{N}_n}(\cdot)$ 表示高斯概率质量（见定义 2.4）。

我们的主要结果是以下关于 f 的 L_1-近似保证，其界由其 GSA 给出 .

定理 1.1。 设 $f\colon \mathbb{R}^n \to \{\pm 1\}$ 是一个（可测）函数。对于每个 $\varepsilon > 0$，都存在一个次数为 p $(\log(1/\varepsilon)\cdot \mathrm{GSA}(f)^2/\varepsilon^2)$ 的多项式 $\mathbb{E}_{x\sim\mathcal{N}_n}\left[|f(x) - p(x)|\right] \le \varepsilon$。

[2]在 [KOS08], 的定义中，高斯表面积是 $\mathbb{R}^n$的一个子集的性质。我们将布尔函数 f视为子集 $K(f) = \{x\colon f(x) = 1\} \subseteq \mathbb{R}^n$的指示函数，这使我们能够为 f其本身赋予一个表面积。

作为中间结果，我们得到了以下关于 f 的噪声敏感度界。

命题 1.2。令 $f\colon \mathbb{R}^n \to \{\pm 1\}$ 为一个（可测）函数。对于每个 $\rho \in [0,1]$ 和 $d \in \mathbb{N}$，存在一个次数为 d 的多项式 p，使得

$$\mathbb{E}_{x\sim\mathcal{N}_n}\left[|f(x)-p(x)|\right] \le 2\mathrm{GNS}_{1-\rho}(f) + \rho^{d+1}.$$

结合 [KKMS08、定理 5] 以及 [Kan11a给出的高斯表面积界，KOS08], 我们立即得到定理 1.1的以下推论。

推论 1.3。令 C 为布尔函数 $f\colon \mathbb{R}^n \to \{\pm 1\}$构成的概念类。令 Γ 为每个 $f \in C$的 *GSA* 的上界。令 $\varepsilon > 0$。那么，L_1次多项式回归算法能够在时间和样本复杂度为 $n^{\tilde{O}(\Gamma^2/\varepsilon^2)}\cdot\mathrm{poly}(1/\varepsilon)$的情况下，以不超过 ε 的误差无偏学习 C。特别地，L_1 次多项式回归算法能够无偏学习

a) 次数为k 的多项式阈值函数，所需时间为 $n^{\tilde{O}(k^2/\varepsilon^2)}$；*b)* k 个半空间的交集，所需时间为 $n^{\tilde{O}(\log(k)/\varepsilon^2)}$；
c) 凸集，所需时间为 $n^{\tilde{O}(\sqrt{n}/\varepsilon^2)}$。

概念的 L_1-近似次数下界

如前所述，[DKPZ21] 证明，学习一个概念类C 且误差不超过 ε 时，其 SQ 复杂度约为 $n^{\Omega(d)}$，其中 $d = d(\varepsilon)$是在 ε中对任意元素进行 C -近似所需的最小多项式次数 L_1。作为这一结果的补充，他们还给出了该次数的下界：对于次数为 k的多项式阈值函数，他们证明 $d = \Omega(k^2/\varepsilon^2)$。情形 $k = 1$ 对应于半空间，因此需要 $d = \Omega(1/\varepsilon^2)$。

他们还根据噪声敏感度证明了以下一般界：

定理 1.4([DKPZ21])。设 $f\colon \mathbb{R}^n \to \{\pm 1\}$ 是一个（可测）函数。对于每个次数为 p 的多项式 d，我们有

$$\mathbb{E}_{x\sim\mathcal{N}_n}\left[|f(x)-p(x)|\right] \ge \Omega(1/\log d)\cdot \mathrm{GNS}_{(\log d/d)^2}(f)$$

作为一项应用，定理 1.4 表明，要近似 k 个半空间的交集，需要 $d = \tilde{\Omega}(\sqrt{\log k}/\varepsilon)$。请注意，定理 1.4 通常并非紧界，也无法达到命题 1.2 的上界。例如，当 f 是一个半空间时，我们有 $\mathrm{GNS}_\delta(f) \approx \sqrt{\delta}$。于是，定理 1.4 给出的次数界为 $d = \tilde{\Omega}(1/\varepsilon)$，而命题 1.2 表明界为 $d = \tilde{O}(1/\varepsilon^2)$。正确的次数为 $d = \Theta(1/\varepsilon^2)$ [Gan08, DKN10]。

概述

本文其余部分安排如下：第2节第2节中，我们介绍 Hermite分析和 Ornstein-Uhlenbeck（噪声）算子的预备知识；第 3 节第 3 节中，我们给出主要结果 定理 1.1的证明；第4节第4节 中，我们将本文的构造与 [KKMS08、DKN10、FKV20]的构造进行比较。

2 Hermite 分析与 Ornstein–Uhlenbeck 算子

在本节中，我们回顾关于Hermite分析和Ornstein-Uhlenbeck 算子的一些基本事实。参见 [LT91，Bak94，Jan97] 以获取一般参考文献，以及 [O'D14，第 11] 章，以获得一个很好的概述。

2.1 Hermite分析

我们用 $\mathcal{N}_n = \mathcal{N}(0, I_n)$ 表示 $\mathbb{R}^n$ 上的标准高斯分布。在全文中，我们假设所有函数 $f\colon \mathbb{R}^n \to \mathbb{R}$ 均为（博雷尔）可测的。我们用 $L_2(\mathcal{N}_n)$ 表示满足 $\|f\|^2_{L_2(\mathcal{N}_n)} := \mathbb{E}_{X\sim\mathcal{N}_n}[f(X)^2] < \infty$ 的函数 $f\colon \mathbb{R}^n \to \mathbb{R}$ 所构成的空间。它具有由 $f, g_{\mathcal{N}_n} := \mathbb{E}_{X\sim\mathcal{N}_n}[f(X)g(X)]$ 给出的内积。类似地，$L_1(\mathcal{N}_n)$ 表示满足 $\|f\|_{L_1(\mathcal{N}_n)} := \mathbb{E}_{X\sim\mathcal{N}_n}[|f(X)|] < \infty$ 的函数 $f\colon \mathbb{R}^n \to \mathbb{R}$ 所构成的空间。我们可以使用柯西-施瓦茨不等式来界定

$$\|f\|_{L_1(\mathcal{N}_n)} \le \|f\|_{L_2(\mathcal{N}_n)}. \tag{1}$$

对于 $n = 1$，空间 $L_2(\mathcal{N}_1)$ 有一组标准正交基，由厄米多项式 $(H_k)_{k\ge 0} \subseteq \mathbb{R}[x]$构成；它们是首项系数为正、满足 H_k 次数为 k 且.....的唯一多项式

$$\langle H_i, H_j\rangle_{\mathcal{N}_1} = \mathbb{E}_{X\sim\mathcal{N}_1}[H_i(X)H_j(X)] = \delta_{ij} \quad (\forall i, j \in \mathbb{N}).$$

这个定义对应于（首一）概率论家的厄米多项式，并缩放了因子 $1/\sqrt{k!}$。对于 $n > 1$，我们有一组由多元厄米多项式构成的标准正交基，定义如下：

$$H_\alpha(x) := H_{\alpha_1}(x_1)\cdot H_{\alpha_2}(x_2)\cdot \ldots \cdot H_{\alpha_n}(x_n) \quad (\alpha \in \mathbb{N}^n). \tag{2}$$

注意 H_α 的次数为 $|\alpha| := \sum_{i=1}^n \alpha_i$。任意 $f \in L_2(\mathcal{N}_n)$ 都有一个（唯一的）厄米展开

$$f = \sum_{\alpha\in\mathbb{N}^n} \hat{f}(\alpha)\cdot H_\alpha, \quad \text{with} \quad \hat{f}(\alpha) \in \mathbb{R}.$$

^我们将 $f(\alpha)$ 称为 f 的厄米系数。由厄米多项式的标准正交性可得 $\|f\|^2_{L_2(\mathcal{N}_n)} = \sum_\alpha \hat{f}(\alpha)^2$（帕塞瓦尔恒等式）。对于 $d \in \mathbb{N}$，我们记 $\Pi_d\colon L_2(\mathcal{N}_n) \to L_2(\mathcal{N}_n)$，表示投影到 $\operatorname{span}\{H_\alpha : |\alpha| \le d\}$ 上的正交投影。也就是说，

$$\Pi_d f = \sum_{|\alpha|\le d} \hat{f}(\alpha)\cdot H_\alpha.$$

根据定义，$\Pi_d f$ 是 f 的最佳 L_2近似，由一个次数为d 的多项式给出。它也称为 f的d 次厄米展开。其近似误差可用 f 的厄米系数表示如下：

$$\|f - \Pi_d f\|^2_{L_2(\mathcal{N}_n)} = \Big\| \sum_{|\alpha|>d} \hat{f}(\alpha)\cdot H_\alpha \Big\|^2_{L_2(\mathcal{N}_n)} = \sum_{|\alpha|>d} \hat{f}(\alpha)^2. \tag{3}$$

2.2 Ornstein-Uhlenbeck 算子

对于 $\rho \in [0,1]$，*Ornstein-Uhlenbeck* 算子 $T_\rho\colon L_2(\mathcal{N}_n) \to L_2(\mathcal{N}_n)$ 定义如下：

$$T_\rho f(x) = \mathbb{E}_{Y\sim\mathcal{N}_n}\left[f\left(\rho x + \sqrt{1-\rho^2}Y\right)\right].$$

根据构造，T_ρ 是一个线性算子。它可由厄米多项式对角化；即

$$T_\rho H_\alpha = \rho^{|\alpha|} H_\alpha \quad (\forall \alpha \in \mathbb{N}^n). \tag{4}$$

为完整起见，我们在引理 A.2中给出这一事实的证明。由此直接可知，$T_\rho f$的厄米系数对于任意 $f \in L_2(\mathcal{N}_n)$都会 $T_\rho f$。根据（3），这意味着 $T_\rho f$ 可以在 $L_2(\mathcal{N}_n)$ 中由低次多项式很好地逼近：

引理 2.1。 设 $f \in L_2(\mathcal{N}_n)$。对于任意 $\rho \in [0,1]$ 和 $d \in \mathbb{N}$，有

$$\|T_\rho f - \Pi_{\le d}(T_\rho f)\|^2_{L_2(\mathcal{N}_n)} \le \rho^{2d+2} \cdot \|f\|^2_{L_2(\mathcal{N}_n)}.$$

^证明。 记 $f(x) = \sum_\alpha f(\alpha) H_\alpha(x)$，利用 (4，可得

$$T_\rho f(x) = \sum_{\alpha \in \mathbb{N}^n} \hat{f}(\alpha) \cdot T_\rho H_\alpha(x) = \sum_{\alpha \in \mathbb{N}^n} \rho^{|\alpha|} \hat{f}(\alpha) \cdot H_\alpha(x).$$

因此，由 (3) 可得

$$\|T_\rho f - \Pi_{\le d}(T_\rho f)\|^2_{L_2(\mathcal{N}_n)} = \sum_{|\alpha|>d} \hat{f}(\alpha)^2 \rho^{2|\alpha|} \le \rho^{2(d+1)} \cdot \sum_{|\alpha|>d} \hat{f}(\alpha)^2 \le \rho^{2(d+1)} \cdot \|f\|^2_{L_2(\mathcal{N}_n)}. \qquad \square$$

2.3 高斯噪声敏感度与高斯表面积

对于 $f: \mathbb{R}^n \to \{\pm 1\}$，$f$ 与 $T_\rho f$ 之间的期望差值等于 高斯噪声敏感度 的两倍，其中 f。（事实上，有时会用这个性质来定义噪声敏感度。）为完整起见，我们在 A.2 节中给出这一事实的证明。

定义 2.2 (参见 [KOS08，定义 12])。函数 的高斯噪声敏感度 $f: \mathbb{R}^n \to \{\pm 1\}$ 在 $\delta \in [0,1]$ 处定义为

$$\mathrm{GNS}_\delta(f) := \mathbb{P}_{(X,Y)}[f(X) \ne f(Y)],$$

其中 $(X,$ Y 是两个相关系数为 $(1-\delta)$ 的标准高斯变量，即

$$(X,Y) \sim \mathcal{N}\left(\begin{pmatrix}0_n\\0_n\end{pmatrix}, \begin{pmatrix} I_n & (1-\delta)I_n \\ (1-\delta)I_n & I_n \end{pmatrix}\right).$$

引理 2.3 （见 A.2 节）。 设 $f: \mathbb{R}^n \to \{\pm 1\}$。对任意 $\rho \in [0,1]$，我们有

$$\|f - T_\rho f\|_{L_1(\mathcal{N}_n)} = 2\mathrm{GNS}_{1-\rho}(f).$$

进而，函数的噪声敏感度可以用其高斯表面积来界定。这归功于 [KOS08], , 他们利用 [Led 94]的一个结果证明了这一点。

定义 2.4 （参见 [KOS08，定义 1]）。 设 $K \subseteq \mathbb{R}^n$ 为博雷尔集。其高斯表面积为

$$\mathrm{GSA}(K) := \lim_{\delta \to 0} \frac{\mathrm{vol}_{\mathcal{N}_n}(K_\delta \setminus K)}{\delta},$$

其中 $K_\delta := \{x \in \mathbb{R}^n : \mathrm{dist}(x, K) \le \delta\}$ 是 K 的 δ-邻域，且 $\mathrm{vol}_{\mathcal{N}_n}(\cdot)$ 表示高斯概率质量。

对 $f: \mathbb{R}^n \to \{\pm 1\}$，我们记 $K(f) = \{x \in \mathbb{R}^n : f(x) = 1\}$，并令 $\mathrm{GSA}(f) := \mathrm{GSA}(K(f))$。我们始终如 [KOS08] 中那样假设 $\mathrm{vol}_{\mathcal{N}_n}(\delta K(f)) = 0$，其中 $\delta K(f)$ 表示 $K(f)$ 的边界。

引理 2.5 ([KOS08]，推论 14])。 设 $f: \mathbb{R}^n \to \{\pm 1\}$。对于每个 $\rho \in [0,1]$，我们有

$$\mathrm{GNS}_{1-\rho}(f) \le \sqrt{\pi}\sqrt{1-\rho} \cdot \mathrm{GSA}(f).$$

3 定理 1.1 的证明

我们对 定理 1.1 的证明分两步进行。首先，我们用 f来近似 $T_\rho f$，其中 $\rho \in [0,1]$ 将在稍后选定。然后，我们用其低次数厄米展开 $\Pi_d(T_\rho f)$来近似 $T_\rho f$。根据三角不等式，总近似误差至多为

$$\|f - \Pi_d(T_\rho f)\|_{L_1(\mathcal{N}_n)} \le \|f - T_\rho f\|_{L_1(\mathcal{N}_n)} + \|T_\rho f - \Pi_d(T_\rho f)\|_{L_1(\mathcal{N}_n)}.$$

右侧的第一项可以直接用 f（引理 2.3）中的噪声敏感度来界定。对于第二项，注意，由（1）和 引理 2.1，

$$\|T_\rho f - \Pi_d(T_\rho f)\|_{L_1(\mathcal{N}_n)} \le \|T_\rho f - \Pi_d(T_\rho f)\|_{L_2(\mathcal{N}_n)} \le \rho^{d+1}.$$

综上可得：

命题 1.2。 设 $f\colon \mathbb{R}^n \to \{\pm 1\}$ 是一个（可测）函数。对于每个 $\rho \in [0,1]$ 和 $d \in \mathbb{N}$，都存在一个次数为 d 的多项式 p，使得

$$\mathbb{E}_{x\sim\mathcal{N}_n}\left[|f(x) - p(x)|\right] \le 2\mathrm{GNS}_{1-\rho}(f) + \rho^{d+1}.$$

证明定理 *1.1*。 令 p 为 d 次数的多项式，它来自命题 1.2。根据引理 2.5，我们可以进一步界定

$$\|f - p\|_{L_1(\mathcal{N}_n)} \le 2\sqrt{\pi}\sqrt{1-\rho}\cdot \mathrm{GSA}(f) + \rho^{d+1}. \tag{5}$$

给定 $\varepsilon > 0$，我们希望确保（5）右侧的两项都不超过 $\varepsilon/2$。对于第一项，只需令

$$\rho = \max\left(0, 1 - \frac{\varepsilon^2}{16\pi \mathrm{GSA}(f)^2}\right).$$

现在，给定 ρ 的这一选择，我们选取 $d \in \mathbb{N}$，使得 $\rho^{d+1} \le \varepsilon/2$。不失一般性，假设 $\rho > 0$（否则，$d = 0$ 就足够了）。我们有

$$\begin{aligned}
\rho^{d+1} \le \frac{\varepsilon}{2} &\iff (d+1)\log(\rho) \le \log(\varepsilon/2) \\
&\impliedby (d+1)(\rho - 1) \le \log(\varepsilon/2) && (\text{as } d \ge 0 \text{ and } \log(t) \le t-1 \text{ for all } t > 0) \\
&\iff (d+1) \ge \frac{\log(\varepsilon/2)}{\rho - 1} && (\text{as } \rho - 1 < 0) \\
&\iff d \ge \frac{\log(2/\varepsilon)}{1-\rho} - 1.
\end{aligned}$$

代入 $1-\rho$ 的值后，只需令

$$d = 16\pi \cdot \mathrm{GSA}(f)^2 \cdot \frac{\log(2/\varepsilon)}{\varepsilon^2} - 1. \qquad \square$$

4 与以往工作的比较

在本节中，我们将我们对定理 1.1 的证明与三项先前关于无关学习的工作进行比较，这些工作在 L_1 中建立了多项式逼近结果。它们的区别在于其对 f的低次数逼近的选择：

1. f [的截断埃尔米特展开KOS08];

2. 对 f进行平滑处理后，取其厄米截断展开；平滑处理通过与精心选取的函数 [DKN10]卷积得到；[3]

3. 对 f进行平滑处理后，取其截断傅里叶展开；平滑处理通过应用布尔噪声算子 [FKV20]。

4.1 与 [KOS08]比较

在 [KKMS08] 关于半空间的论证基础上， [KOS08] 更一般地证明， f 的次数为 d 的厄米截断（记为 $\Pi_d f$）满足 $\|f-\Pi_d f\|_{L_1(\mathcal{N}_n)} \le \varepsilon$，其中 $d = O(\mathrm{GSA}(f)^2/\varepsilon^4)$。关键在于，他们的证明将所需的 L_1 近似转化为一个 L_2 界：

$$\begin{aligned}\|f - \Pi_d f\|_{L_1(\mathcal{N}_n)} &\le \|f - \Pi_d f\|_{L_2(\mathcal{N}_n)} \\ &\le O(1)\cdot \mathrm{GNS}_{1/d}(f)^{1/2} && \text{(by [KOS02, Prop. 16])} \\ &\le O(1)\cdot d^{-1/4}\cdot \mathrm{GSA}(f)^{1/2}\,. && \text{(by Lemma 2.5)}\end{aligned}$$

此论证所得界与定理 1.1 所给保证之间的差距，完全可以归因于第一步：即使对于原点中心半空间，次数为 L_2的厄米截断的误差为d $\Omega(d^{-1/4})$ [KKMS08]。由于 $\Pi_d f$ 是对d 的最佳次数为 f 的近似 L_2，这表明将整个问题化归为一个 L_2 问题的策略本质上会造成损失。

然而，我们尚未发现任何例子能够表明 $\Pi_d f$ 无法达到 定理 1.1 在 L_1中的保证。另一方面，我们证明 $\Pi_d f$ 确实能对原点中心半空间达到这些保证。将问题归约到 $n = 1$之后，只需证明以下结论：

命题 4.1。 令 $\mathrm{sign} : \mathbb{R} \to \{\pm 1\}$ 表示符号函数。我们有

$$\left\|\mathrm{sign} - \Pi_d \mathrm{sign}\right\|_{L_1(\mathcal{N}_1)} \le O\left(\frac{\log d}{\sqrt{d}}\right).$$

证明。 参见A.4 节。 □

总之，至少对于原点中心半空间而言， [KKMS08、KOS08] 相对于定理 1.1 的次优次数界，并非源于逼近多项式的选择（$\Pi_d f$ 与 $\Pi_d T_\rho f$相比），而是源于对柯西-施瓦茨不等式的宽松应用。

4.2 与 [DKN10]<style id="">的比较

[DKN10] 提出了一种构造低次逼近的不同技术，特别证明了次数为 $d = O(1/\varepsilon^2)$ 即可在 L_1中逼近半空间。在这一特殊情形下，其保证比定理 1.1 提高了一个对数因子。

鉴于我们的证明，他们的方法可以如下说明其动机。首先，对于半空间$f(x) = \mathrm{sign}(c - w, x)$，通过投影到$\mathrm{span}(w)$上，我们可以不失一般性地假设 $n = 1$。我们现在确定希望 f 的平滑 g 具备哪些性质，以便

[3] [DKN10] 实际上考虑的是如下对象这一平滑版本的泰勒近似 f。他们关注的是比这里所考虑的更强的近似概念（与证明普适性相关），因此这一点很重要。在我们的语境中，我们可以等价地考虑截断埃尔米特展开。

实施来自第 3 节的策略。通过高斯分部积分，在引理 2.1 中被界定的厄米系数可以解释为（参见引理 A.4）

$$\langle g, h_k \rangle = \frac{\mathbb{E}_{x\sim\mathcal{N}(0,1)}\, g^{(k)}(x)}{\sqrt{k!}}\,. \tag{6}$$

因此，我们希望在 L_1 中用 g 逼近 f（以替代引理 2.3），同时保持 g 的导数有界。在一维情形中， [DKN10] 的作者证明了这在强意义下是可行的：对于任意固定的 $\varepsilon > 0$，都可以选取 g，使得 (a) $\|g - f\|_{L_1(\mathcal{N})} \le \varepsilon$，且 (b) 对所有 $k \ge 1$， $\|g^{(k)}\|_\infty \le O(1/\varepsilon)^k$。在证明中， g 是通过将 f 与一个傅里叶变换具有有界支撑的合适函数进行卷积而定义的。结合 (6) 并沿用第 3 节中的论证，(a) 和 (b) 可直接推出，对于 $d = O(1/\varepsilon^2)$， $\|\Pi_d g - f\|_1 \le \varepsilon$，正如所需。

这一构造的主要难点在于将其推广到更高维度。对于具有有界高斯表面积的一般概念类上的无关学习，重复前面的论证似乎会引入一个依赖于维度的因子。对于多项式阈值函数这一特殊情形， [DKN10、Kan11b]证明了仍可将问题化归到低维情形，但代价是对 ε 的依赖（大幅）变差。

4.3 与 [FKV20]的比较

中的构造 [FKV20] 是我们构造的布尔类比。事实上， [FKV20，引理 3.1]表明，对于任意布尔函数 $f\colon \{0,1\}^n \to \mathbb{R}$ 以及 $\rho \in (0,1]$,，存在一个次数为 p 的多项式$d = O(\log(1/\varepsilon)/(1-\rho))$ 使得

$$\|f - p\|_{L_1(\mathcal{U}(\{0,1\}^n))} \le 2\mathrm{BNS}_{1-\rho}(f) + \rho^{d+1} \cdot \|f\|_{L_2(\mathcal{U}(\{0,1\}^n)}\,, \tag{7}$$

其中 $\mathcal{U}(\{0,1\}^n)$ 是 均匀分布在 $\{0,1\}^n$ 上，而 BNS 是布尔噪声敏感度。

公式（7）是命题 1.2的布尔类比。事实上，7式的证明 [FKV20]与我们对命题 1.2的证明十分相似。首先将 f 近似为 $T_\rho f$，其中 T_ρ 是超立方体上的噪声算子；所得 L_1 误差可直接由布尔噪声敏感度控制。然后将 $T_\rho f$ 近似为其次数为d 的傅里叶截断，并利用与高斯情形中基本相同的论证，界定该截断的 L_1 误差。

致谢

我们感谢 Pjotr Buys 在本研究早期阶段进行了一次有益的讨论。

参考文献

[ABL14] Pranjal Awasthi、Maria Florina Balcan 和 Philip M. Long。利用局部化的力量，高效学习带噪声的线性分隔器。载于*ACM* 计算理论研讨会论文集（*STOC 2014*），第 449–458 页。ACM，纽约，2014 年。2[Bak94] Dominique Bakry。超收缩性及其在半群理论中的应用，数学讲义第 1581 卷，数学讲义，第 1–114 页。Springer，柏林，1994 年。4，12[Bal93] Keith Ball。高斯测度的反向等周问题。离散与计算几何，10(4):411–420，1993 年。2

[BOW10] Eric Blais、Ryan O’Donnell 和 Karl Wimmer。在任意乘积分布下的多项式回归。机器学习，80(2-3):273–294，2010 年。2[CPS26] Xi Chen、Shyamal Patel 和 Rocco A. Servedio。容错合取测试与不可知学习合取式之间一种神秘的联系。载于第 58 届 ACM 计算理论研讨会年会论文集（STOC 2026），2026 年。即将发表。2[Dan15] Amit Daniely。不可知学习半空间的多项式时间近似方案。载于第 28 届学习理论会议（COLT 2015）论文集，第 40 卷，第 484–502 页，法国巴黎，2015 年。PMLR。2[Dan16] Amit Daniely。学习半空间的复杂性理论限制。载于第 48 届 ACM SIGACT 计算理论研讨会年会论文集（STOC 2016），第 105–117 页。ACM，纽约，2016 年。1[DHK+10] Ilias Diakonikolas、Prahladh Harsha、Adam Klivans、Raghu Meka、Prasad Raghavendra、Rocco A. Servedio 和 Li-Yang Tan。多项式阈值函数的平均敏感度与噪声敏感度界。载于 2010 年 ACM 计算理论国际研讨会论文集（STOC 2010），第 533–542 页。ACM，纽约，2010 年。2[DKN10] Ilias Diakonikolas、Daniel M. Kane 和 J elani Nelson。有界独立性能够欺骗二次阈值函数。载于 2010 年 IEEE 第 51 届计算机科学基础年会（FOCS 2010），第 11–20 页。IEEE 计算机学会，洛斯阿拉米托斯，加利福尼亚州，2010 年。2, 3, 4, 8, 9[DKPZ21] Ilias Diakonikolas、Daniel M. Kane、Thanasis Pittas 和 Nikos Zarifis。统计查询模型中高斯边缘分布下不可知学习的多项式回归最优性。载于第三十四届学习理论会议（COLT 2021）论文集，第 134 卷，第 1552–1584 页。PMLR，2021 年 8 月 15–19 日。2, 3, 4[DKR25] Ilias Diakonikolas、Daniel M. Kane 和 Lisheng Ren。更快地学习析取式的不可知算法及其影响。载于第三十八届学习理论会议（COLT 2025）论文集，第 291 卷，第 1531–1558 页。PMLR，2025 年。2[DKS18] Ilias Diakonikolas、Daniel M. Kane 和 Alistair Stewart。在恶劣噪声下学习几何概念。载于第 50 届 ACM SIGACT 计算理论研讨会年会论文集（STOC 2018），第 1061–1073 页。ACM，纽约，2018 年。2[DKTZ20] Ilias Diakonikolas、Vasilis Kontonis、Christos Tzamos 和 Nikos Zarifis。非凸随机梯度下降能够在对抗性标签噪声下学习半空间。载于神经信息处理系统进展，第 33 卷，第 18540–18549 页。Curran Associates, Inc.，2020 年。2[DLMF] NIST数学函数数字图书馆。https://dlmf.nist.gov/，2025-12-15 发布的 1.2.5 版。编者：F. W. J. Olver、A. B. Olde Daalhuis、D. W. Lozier、B. I. Schneider、R. F. Boisvert、C. W. Clark、B. R. Miller、B. V. Saunders、H. S. Cohl 和 M. A. McClain。19[FKV20] Vitaly Feldman、Pravesh Kothari 和 Jan Vondrák。自界函数的 L^1近似与学习的紧确界。理论计算机科学，808:86–98，2020 年。3, 4, 8, 9

[Gan08] Michael I. Ganzburg。带指数权的多项式逼近的极限定理。《美国数学学会回忆录》，192(897):viii+161，2008。2, 3, 4[Jan97] Svante Janson。《高斯希尔伯特空间》。剑桥大学出版社，1997。4[Kan11a] Daniel M. Kane。次数-D 多项式阈值函数的高斯表面积与噪声敏感度。《计算复杂性》，20(2):389–412，2011。2, 3, 4[Kan11b] Daniel M. Kane。 k-独立高斯变量可以欺骗多项式阈值函数。载于第26届年度IEEE计算复杂性会议（CCC 2011），第252–261页。IEEE计算机学会，洛斯阿拉米托斯（加利福尼亚州），2011。9[KKM13] Daniel M. Kane、Adam Klivans和Raghu Meka。对数凹密度下的半空间学习：多项式逼近与矩匹配。载于第26届年度学习理论会议论文集（COLT 2013），第30卷，第522–545页，美国新泽西州普林斯顿，2013。PMLR。2[KKMS08] Adam Tauman Kalai、Adam R. Klivans、Yishay Mansour和Rocco A. Servedio。不可知学习半空间。《SIAM计算杂志》，37(6):1777–1805，2008。2, 4, 8[KLS09] Adam R. Klivans、Philip M. Long和Rocco A. Servedio。带恶意噪声的半空间学习。《机器学习研究期刊》（JMLR），10:2715–2740，2009。2[KOS02] Adam Klivans、Ryan O’Donnell和Rocco A. Servedio。半空间的交集与阈值的学习。载于第43届计算机科学基础研讨会论文集（FOCS 2002），第177–186页，美国，2002。IEEE计算机学会。8[KOS08] Adam R. Klivans、Ryan O’Donnell和Rocco A. Servedio。通过高斯表面积学习几何概念。载于第49届年度IEEE计算机科学基础研讨会（FOCS 2008），第541–550页，2008。2, 3, 4, 6, 8[KSS92] Michael J. Kearns、Robert E. Schapire和Linda M. Sellie。迈向高效的无关学习。载于第五届年度计算学习理论研讨会论文集，第341–352页。美国计算机协会，1992。1[Led94] Michel Ledoux。欧氏空间与高斯空间中等周不等式的半群证明。《数学科学通报》，118(6):485–510，1994。6[LT91] Michel Ledoux和Michel Talagrand。《巴拿赫空间中的概率》，第23卷。Springer-Verlag，柏林，1991。4[O’D14] Ryan O’Donnell。《布尔函数分析》。剑桥大学出版社，纽约，2014。4[Sze75] Gábor Szegő。《正交多项式》，美国数学学会专题论文集第XXIII卷。美国数学学会，普罗维登斯（罗得岛州），第四版，1975。15, 16[Tie23] Stefan Tiegel。最坏情形格问题带来的不可知学习半空间的困难性。载于第三十六届学习理论会议论文集（COLT 2023），第195卷，第3029–3064页。PMLR，2023。1[Val84] Leslie G. Valiant。可学习性理论。《ACM通讯》，27:1134–1142，1984。1

A 技术性证明

在本附录中，我们给出正文中省略的一些技术性证明。

A.1 Ornstein-Uhlenbeck 算子的特征值

我们证明，厄米多项式将 Ornstein-Uhlenbeck 算子对角化（方程（4））。这是一个标准结果，为求完整起见，我们在此给出，参见例如 [Bak94]。我们需要关于厄米多项式生成函数的以下事实。

事实 A.1. 对于任意 x、$t \in \mathbb{R}$，有

$$\exp\left(xt - \frac{1}{2}t^2\right) = \sum_{n=0}^{\infty} H_k(x)\frac{t^k}{\sqrt{k!}}.$$

By T_ρ 的定义，对于任意形如 $f(x) = f_1(x_1) \cdot f_2(x_2) \cdot \ldots \cdot f_n(x_n)$ 的 f，有

$$T_\rho f(x) = T_\rho f_1(x_1) \cdot T_\rho f_2(x_2) \cdot \ldots \cdot T_\rho f_n(x_n).$$

连同 h (2)，因此只需证明以下一维情形即可得出结论 (4).

引理 A.2. 对于任意 $n \in \mathbb{N}$，我们有

$$T_\rho H_n = \rho^n H_n.$$

下面给出的证明遵循 [Bak94，第 0.I 节]。

证明。 我们要用到 事实 A.1。定义 $g(x) = \exp(xt - \frac{1}{2}t^2)$。于是，根据 T_ρ 和 g的定义，我们得到

$$\begin{aligned} T_\rho g(x) &= \mathbb{E}_{Y\sim\mathcal{N}_n}\left[g\left(\rho x + \sqrt{1-\rho^2}Y\right)\right] \\ &= \int_{y\in\mathbb{R}} \exp\left(\rho xt + \sqrt{1-\rho^2}yt - \frac{1}{2}t^2\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right)\mathrm{d}y. \end{aligned}$$

重新整理各项并配方后，我们得到

$$\begin{aligned} T_\rho g(x) &= \exp\left(\rho xt - \frac{1}{2}t^2\right)\int_{y\in\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y - \sqrt{1-\rho^2}t\right)^2\right)\exp\left(\frac{1}{2}\left(1-\rho^2\right)t^2\right)\mathrm{d}y \\ &= \exp\left(\rho xt - \frac{1}{2}\rho^2t^2\right)\int_{y\in\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y - \sqrt{1-\rho^2}t\right)^2\right)\mathrm{d}y. \end{aligned}$$

现在代入 $y' = y - \sqrt{1-\rho^2}t$，得到

$$\int_{y\in\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y - \sqrt{1-\rho^2}t\right)^2\right)\mathrm{d}y = \int_{y\in\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y'\right)^2\right)\mathrm{d}y' = 1$$

因此我们得到

$$T_\rho g(x) = \exp\left(\rho xt - \frac{1}{2}\rho^2t^2\right).$$

现在我们使用事实 A.1两次。首先，我们得到（使用 事实 A.1 对于 $t' = \rho t$）

$$T_\rho g(x) = \exp\left(\rho x t - \frac{1}{2}\rho^2 t^2\right) = \sum_{n=0}^{\infty} H_n(x)\rho^n \frac{t^n}{\sqrt{n!}}.$$

另一方面，首先使用 事实 A.1 然后利用 T_ρ的线性性，可得

$$T_\rho g(x) = T_\rho\left(\sum_{n=0}^{\infty} H_n(x)\frac{t^n}{\sqrt{n!}}\right) = \sum_{n=0}^{\infty} T_\rho H_n(x)\frac{t^n}{\sqrt{n!}}.$$

因此，我们可以得出

$$\sum_{n=0}^{\infty} H_n(x)\rho^n \frac{t^n}{\sqrt{n!}} = \sum_{n=0}^{\infty} T_\rho H_n(x)\frac{t^n}{\sqrt{n!}}$$

由于这对所有 $t \in \mathbb{R}$ 都成立，因此必有

$$T_\rho H_n = \rho^n H_n.$$ □

A.2 Ornstein-Uhlenbeck 算子与噪声敏感性

这里，我们给出引理 2.3的证明，表明 $\|f - T_\rho f\|_{L_1(\mathcal{N}_n)} =$ 2高斯噪声敏感性$_{1-\rho}(f)$ 对于 $f\colon \mathbb{R}^n \to \{\pm 1\}$。

引理 *2.3* 的证明引理 *2.3*。 我们有

$$\begin{aligned}
\|f - T_\rho f\|_{L_1(\mathcal{N}_n)} &= \mathbb{E}_{X\sim\mathcal{N}_n}\left[|f(X) - T_\rho f(X)|\right] \\
&= \mathbb{E}_{X\sim\mathcal{N}_n}\left[\left|f(X) - \mathbb{E}_{Y\sim\mathcal{N}_n}\left[f\left(\rho X + \sqrt{1-\rho^2}Y\right)\right]\right|\right] \\
&= \mathbb{E}_{X\sim\mathcal{N}_n}\left[\left|\mathbb{E}_{Y\sim\mathcal{N}_n}\left[f(X) - f\left(\rho X + \sqrt{1-\rho^2}Y\right)\right]\right|\right] \\
&= \mathbb{E}_{X,Y\sim\mathcal{N}_n}\left[\left|f(X) - f\left(\rho X + \sqrt{1-\rho^2}Y\right)\right|\right].
\end{aligned}$$

在最后一步中，我们使用了三角不等式。对于函数 $f\colon \mathbb{R}^n \to \{\pm 1\}$，实际上等号成立，因为对于固定的 x，项 $f(x) - f(\rho x + \sqrt{1-\rho^2}y)$ 要么始终非负（如果$f(x) = 1$），要么始终非正（如果 $f(x) = -1$）。

现在，如果 $X, Y \sim \mathcal{N}(0, I_n)$，那么 X 和 $\rho X + \sqrt{1-\rho^2}Y$ 是 ρ-相关高斯随机变量。因此，

$$\left|f(x) - f\left(\rho x + \sqrt{1-\rho^2}y\right)\right| = 2 \cdot \mathbb{1}\left\{f(x) \neq f\left(\rho x + \sqrt{1-\rho^2}y\right)\right\}$$

上述期望等于 2GNS$_{1-\rho}(f)$。 □

A.3 Hermite 系数的光滑函数

我们展示如何根据光滑函数的导数计算其厄米系数。我们需要以下关于厄米多项式的事实。令 φ 表示 $\mathbb{R}$ 上标准高斯分布的密度函数。

事实 A.3. 对于 $x \in \mathbb{R}$ 和 $k \in \mathbb{N}_0$，我们有

$$H_k(x)\varphi(x) = \frac{(-1)^k}{\sqrt{k!}} \cdot \frac{d^k \varphi}{dx^k}(x).$$

有了这一事实，我们现在可以证明下面的引理。

引理 A.4. 设 $g\colon \mathbb{R}^n \to \mathbb{R}$ 为一个光滑函数，且对所有多重指标 α 满足 $\left\|\frac{\partial g}{\partial \alpha}\right\|_\infty < \infty$。设 β 为一个多重指标。则我们有

$$\hat{g}(\beta) \coloneqq \langle g, H_\beta \rangle_{\mathcal{N}_n} = \frac{\mathbb{E}_{X\sim\mathcal{N}_n}\left[\frac{\partial g}{\partial \beta}(X)\right]}{\sqrt{\beta!}}.$$

证明。 为了证明该引理，我们希望对每个坐标使用高斯分部积分。为此，首先考虑情形 $n = 1$。在此情形下，由事实 A.3，

$$\langle g, H_k \rangle_{\mathcal{N}_1} = \frac{(-1)^k}{\sqrt{k!}} \int_{x\in\mathbb{R}} g(x) \cdot \frac{d^k \varphi}{dx^k}(x) \mathrm{d}x.$$

使用分部积分 k 次，我们得到

$$\langle g, H_k \rangle_{\mathcal{N}_1} = \frac{1}{\sqrt{k!}} \int_{x\in\mathbb{R}} \frac{d^k g}{dx^k}(x) \cdot \varphi(x) \mathrm{d}x = \frac{1}{\sqrt{k!}} \mathbb{E}_{X\sim\mathcal{N}_n}\left[\frac{d^k g}{dx^k}(X)\right].$$

注意，边界条件始终为 0，因为我们假设 $\left\|\frac{d^\ell g}{dx^\ell}\right\|_\infty < \infty$。这表明了一维情形。对于多元情形，注意我们可以依次迭代考虑每个坐标，从而得到

$$\begin{aligned}
\hat{g}(\beta) &\coloneqq \langle g, H_\beta \rangle_{\mathcal{N}_n} \\
&= \mathbb{E}_{X\sim\mathcal{N}_n}[g(X) \cdot H_{\beta_1}(X_1) \cdot \ldots \cdot H_{\beta_n}(X_n)] \\
&= \mathbb{E}_{X_2,\ldots,X_n\sim\mathcal{N}_1}[\mathbb{E}_{X_1\sim\mathcal{N}_1}[g(X) \cdot H_{\beta_1}(X_1) \cdot \ldots \cdot H_{\beta_n}(X_n) \mid X_2, \ldots, X_n]] \\
&= \mathbb{E}_{X_2,\ldots,X_n\sim\mathcal{N}_1}[\mathbb{E}_{X_1\sim\mathcal{N}_1}[g(X) \cdot H_{\beta_1}(X_1) \mid X_2, \ldots, X_n] \cdot H_{\beta_2}(X_2) \ldots \cdot H_{\beta_n}(X_n)].
\end{aligned}$$

现在，我们可以利用一维情形得到

$$\mathbb{E}_{X_1\sim\mathcal{N}_1}[g(X)H_{\beta_1}(X_1) \mid X_2, \ldots, X_n] = \frac{1}{\sqrt{\beta_1!}} \mathbb{E}_{X_1\sim\mathcal{N}_1}\left[\frac{\partial^{\beta_1} g}{\partial x_1^{\beta_1}}(X) \mid X_2, \ldots, X_n\right]$$

因此

$$\hat{g}(\beta) = \frac{1}{\sqrt{\beta_1!}} \mathbb{E}_{X\sim\mathcal{N}_n}\left[\frac{\partial^{\beta_1} g}{\partial x_1^{\beta_1}}(X) \cdot H_{\beta_2}(X_2) \ldots \cdot H_{\beta_n}(X_n)\right].$$

我们可以对其余坐标迭代使用同一论证，从而得到该结果。 □

A.4 半空间的厄米截断

本小节的目标是证明命题 4.1。所有 O 记号均相对于 $d \to \infty$。我们用 φ 表示标准高斯分布在 $\mathbb{R}$ 上的密度。

注意，$\mathrm{sign} \in L^2(\mathcal{N}_1)$，因此 sign 的厄米展开是良定义的。此外，由于 sign 是奇函数，sign 的所有偶次厄米系数均为 0，因此只需考虑 d 为奇数的情况。利用同样的对称性，我们可以简化为

$$\|\mathrm{sign} - \Pi_d \mathrm{sign}\|_{L_1(\mathcal{N}_1)} = \int_{-\infty}^{\infty} |\mathrm{sign}(x) - \Pi_d \mathrm{sign}(x)| \, \mathrm{d}\varphi(x) = 2 \int_0^{\infty} |1 - \Pi_d \mathrm{sign}(x)| \, \mathrm{d}\varphi(x) \,. \tag{8}$$

为了给右侧的积分设定上界，我们首先用厄米多项式表示 sign 的低次厄米投影：

引理 A.5。若 d 为奇数，则对于任意 $x \geq 0$,

$$\Pi_d \mathrm{sign}(x) = \sqrt{\frac{2d}{\pi}} \, H_{d-1}(0) \int_0^x \frac{H_d(t)}{t} \, \mathrm{d}t \,.$$

对于任意固定的奇数 d，右侧的积分都是良定义的，因为如 $t \to 0$ 所示，我们有 $H_d(t) \sim C(d)t$，其中 $C(d) > 0$ 为某个常数。

证明。当 k 为奇数时，sign 的第 k 个厄米系数为

$$\langle \mathrm{sign}, H_k \rangle_{\mathcal{N}_1} = \int_{-\infty}^{\infty} \mathrm{sign}(y) H_k(y) \, \mathrm{d}\varphi(y) = 2 \int_0^{\infty} H_k(y) \, \mathrm{d}\varphi(y) \,.$$

接下来，我们利用恒等式 $H_k \varphi = -\frac{1}{\sqrt{k}} (H_{k-1}\varphi)'$（它由正交归一化厄米多项式满足的递推关系推出），得到

$$\langle \mathrm{sign}, H_k \rangle_{\mathcal{N}_1} = \frac{2}{\sqrt{k}} \left[-H_{k-1}(y)\varphi(y) \right]_0^{\infty} = \sqrt{\frac{2}{k\pi}} H_{k-1}(0) \,.$$

因此，对于任意 $x \in \mathbb{R}$,

$$\Pi_d \mathrm{sign}(x) = \sum_{k=1}^{d} \langle \mathrm{sign}, H_k \rangle_{\mathcal{N}_1} H_k(x) = \sqrt{\frac{2}{\pi}} \sum_{k=1}^{d} \frac{H_k(x) H_{k-1}(0)}{\sqrt{k}} \,.$$

对两边关于 x 求导，并应用恒等式 $H_k' = \sqrt{k} H_{k-1}$,

$$(\Pi_d \mathrm{sign})'(x) = \sqrt{\frac{2}{\pi}} \sum_{k=1}^{d} H_{k-1}(x) H_{k-1}(0) \,.$$

现在，我们应用克里斯托费尔–达布公式 [Sze75，定理 3.2。2]:

$$\sum_{k=0}^{d-1} H_k(x) H_k(0) = \sqrt{d} \, \frac{H_d(x) H_{d-1}(0) - H_{d-1}(x) H_d(0)}{x} \,.$$

由于 d 是奇函数，$H_d(0) = 0$，因此我们只需考虑：

$$(\Pi_d \mathrm{sign})'(x) = \sqrt{\frac{2d}{\pi}} H_{d-1}(0) \frac{H_d(x)}{x} \,,$$

对这个恒等式进行积分即可得到所述结论。 □

A.4.1 Hermite多项式的渐近式

我们将 使用厄米多项式 $H_d(x)$ 在 $d \to \infty$ 时的逐点渐近展开 .

该展开在区域 $|x| \ll \sqrt{d}$ 内一致成立，称为*Plancherel-Rotach*渐近式 [Sze75]。其形式为

$$H_d(x) = \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \left(\sin\left[\frac{1-d}{2}\pi + \sqrt{d}x\right] + r_d(x)\right), \tag{9}$$

这定义了余项 $r_d(x)$，我们接下来将对其进行界定。注意，正弦函数的自变量可根据 d 的奇偶性进行化简：

$$\sin\left[\frac{1-d}{2}\pi + \sqrt{d}x\right] = \begin{cases} (-1)^{\frac{d-1}{2}} \sin(\sqrt{d}x) & \text{if } d \text{ is odd} \\ (-1)^{\frac{d}{2}} \cos(\sqrt{d}x) & \text{if } d \text{ is even} \end{cases}$$

引理 A.6。 对于所有 $|x| \le \sqrt{d}$, 我们有

$$|r_d(x)| \le O\left(\max\left\{\frac{|x|^3}{\sqrt{d}}, \frac{|x|}{\sqrt{d}}, \frac{x^2}{d}, \frac{1}{d}\right\}\right),$$

其中， O 对所有$|x| \le \sqrt{d}$均一致。

证明。 我们使用以下公式 [Sze75，定理 8.22。9]:

$$H_d(x) = \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \sin\left[\arccos\left(\frac{x}{\sqrt{4d+2}}\right)\right]^{-\frac{1}{2}} \cdot \left(\sin\left[\left(\frac{d}{2}+\frac{1}{4}\right)\left(2\frac{x}{\sqrt{4d+2}}\sqrt{1-\frac{x^2}{4d+2}} - 2\arccos\left(\frac{x}{\sqrt{4d+2}}\right)\right) + \frac{3\pi}{4}\right] + O\left(\frac{1}{d}\right)\right).$$

只要满足 $|x| \le \frac{1}{2}\sqrt{4d+2}$，且 O 在所有此类 x 上是一致的，这一结论就成立。我们利用泰勒展开得到 $\sin[\arccos(x/\sqrt{4d+2})] = 1 + O(x^2/d)$，其中 O 在所有 $|x| \le \frac{1}{2}\sqrt{4d+2}$ 上也同样一致。定义 $g(y) = 2y\sqrt{1-y^2} - 2\arccos(y)$，于是可以写成

$$H_d(x) = \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \left(\sin\left[\left(\frac{d}{2}+\frac{1}{4}\right) \cdot g\left(\frac{x}{\sqrt{4d+2}}\right) + \frac{3\pi}{4}\right] + O\left(\max\left\{\frac{x^2}{d}, \frac{1}{d}\right\}\right)\right).$$

对 g 在 0 处进行泰勒展开，得到 $g(y) = -\pi + 4y + \frac{1}{2}g''(\xi)y^2$，其中某个 ξ 介于 0 和 y 之间。我们可以用 $O(|y|)$ 来界定导数 $g''(\xi)$，其中对于所有 $|y| \le \frac{1}{2}$，该界都由某个统一常数给出。综合以上结果，并利用正弦函数是 1-利普希茨的这一事实，可得

$$\begin{aligned} H_d(x) &= \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \left(\sin\left[\left(\frac{d}{2}+\frac{1}{4}\right) \cdot \left(-\pi + 4\frac{x}{\sqrt{4d+2}}\right) + \frac{3\pi}{4}\right] + O\left(\max\left\{\frac{|x|^3}{\sqrt{d}}, \frac{x^2}{d}, \frac{1}{d}\right\}\right)\right) \\ &= \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \left(\sin\left[\frac{1-d}{2}\pi + \sqrt{d+\frac{1}{2}}x\right] + O\left(\max\left\{\frac{|x|^3}{\sqrt{d}}, \frac{x^2}{d}, \frac{1}{d}\right\}\right)\right). \end{aligned}$$

接下来，注意到 $\sqrt{d \pm \frac{1}{2}} \equiv \sqrt{d} \pm O(1/\sqrt{d})$，因此再次利用正弦函数的利普希茨连续性，可得

$$H_d(x) = \exp\left(\frac{x^2}{4}\right) \cdot \left(\frac{2}{\pi d}\right)^{\frac{1}{4}} \cdot \left(\sin\left[\frac{1-d}{2}\pi + \sqrt{d}x\right] + O\left(\max\left\{\frac{|x|^3}{\sqrt{d}}, \frac{|x|}{\sqrt{d}}, \frac{x^2}{d}, \frac{1}{d}\right\}\right)\right),$$

如所声称。 □

推论 A.7. 若 d 为偶数，则

$$H_d(0) = (-1)^{\frac{d}{2}} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} d^{-\frac{1}{4}} + O(d^{-\frac{5}{4}}).$$

推论 A.8. 我们有

$$\sup_{t \in [0, d^{1/6}]} |H_d(t)| e^{-t^2/4} \le O(d^{-\frac{1}{4}}).$$

推论 A.9. 若 d 为奇数，则

$$\sup_{t \in (0, d^{-1/2}]} \frac{|r_d(t)|}{t} \le O\left(\frac{1}{\sqrt{d}}\right).$$

证明推论 A.9 由引理 A.6。我们应用泰勒不等式（当 d 为奇数时，我们有 $r_d(0) = 0$）：

$$\sup_{t \in (0, d^{-1/2}]} \frac{|r_d(t)|}{t} \le \sup_{t \in (0, d^{-1/2}]} |r_d'(t)|. \tag{10}$$

接下来，我们计算 $r_d'(t)$。逐项对（9）求导，得到

$$H_d'(t) = \left(\frac{2}{\pi d}\right)^{1/4} e^{t^2/4} \left(\frac{t}{2}((-1)^{\frac{d-1}{2}} \sin(t\sqrt{d}) + r_d(t)) + \sqrt{d}(-1)^{\frac{d-1}{2}} \cos(t\sqrt{d}) + r_d'(t)\right).$$

利用恒等式 $H_d'(t) = \sqrt{d} H_{d-1}(t)$，整理后可得：

$$r_d'(t) = \sqrt{d}\left[\left(\frac{\pi d}{2}\right)^{1/4} e^{-t^2/4} H_{d-1}(t) - (-1)^{\frac{d-1}{2}} \cos(t\sqrt{d})\right] - \frac{t}{2}\left[(-1)^{\frac{d-1}{2}} \sin(t\sqrt{d}) + r_d(t)\right].$$

我们分别证明第一项和第二项的绝对值都很小。第二项在 $O(d^{-1/2})$ 时为 $t \le d^{-1/2}$。对于第一项，我们对 引理 A.6 应用于 $H_{d-1}(t)$（在$t \le d^{-1/2}$的情形下，误差项$1/d$ 占主导）：

$$\sup_{t \in [0, d^{-1/2}]} \left|\left(\frac{\pi d}{2}\right)^{1/4} e^{-t^2/4} H_{d-1}(t) - (-1)^{\frac{d-1}{2}} \cos(t\sqrt{d})\right| \le O\left(\frac{1}{d}\right).$$

证明可通过乘以$\sqrt{d}$ 并代入（10）中的上界得到。 □

A.4.2 命题 4.1 的证明

将 $H_{d-1}(0)$ 的 Hermite 渐近展开应用于我们的目标（8），

$$\begin{aligned}\Pi_d\mathrm{sign}(x) &= \sqrt{\frac{2d}{\pi}}H_{d-1}(0)\int_0^x \frac{H_d(t)}{t}\,\mathrm{d}t && \text{(by Lemma A.5)}\\ &= \left((-1)^{\frac{d-1}{2}}\left(\frac{2}{\pi}\right)^{\frac{3}{4}}d^{\frac{1}{4}} + O(d^{-\frac{3}{4}})\right)\int_0^x \frac{H_d(t)}{t}\,\mathrm{d}t\,. && \text{(by Corollary A.7)}\end{aligned}$$

论证中主要的技术难点在于估计 $t \mapsto H_d(t)/t$ 的积分。

引理 A.10（小 t）。 如果 d 是奇数，那么对于所有 $\tau = \tau(d) \le d^{1/6}$，

$$\int_0^\tau \frac{|H_d(t)|}{t}\,\mathrm{d}t \le O(d^{\frac{1}{4}})\cdot\tau e^{\tau^2/4}\,.$$

证明。 由于 $H_d(0) = 0$（当 d 是奇数时），由泰勒不等式可得：

$$\left|\int_0^\tau \frac{H_d(t)}{t}\,\mathrm{d}t\right| \le \tau\cdot\sup_{t\in[0,\tau]}|H_d'(t)|\,.$$

结论由恒等式 $H_d'(t) = \sqrt{d}\,H_{d-1}(t)$ 以及 推论 A.8得出。 □

引理 A.11（大 t）。对于所有 $\tau \ge 1$，

$$\int_0^\infty\int_\tau^x \frac{|H_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x) \le e^{-\tau^2/4}\,.$$

证明。使用记号 $\bar\Phi(t) = \Pr^{x\sim\mathcal{N}(0,1)}(x > t)$，我们有

$$\begin{aligned}\int_0^\infty\int_\tau^x \frac{|H_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x) &= \int_\tau^\infty \frac{|H_d(t)|}{t}\bar\Phi(t)\,\mathrm{d}t && \text{(by Fubini's theorem)}\\ &\le \left(\int_\tau^\infty H_d(t)^2\varphi(t)\,\mathrm{d}t\right)^{\frac{1}{2}}\left(\int_\tau^\infty \frac{\bar\Phi(t)^2}{t^2\varphi(t)}\,\mathrm{d}t\right)^{\frac{1}{2}} && \text{(by Cauchy-Schwarz)}\\ &\le \left(\int_\tau^\infty \varphi(t)\,\mathrm{d}t\right)^{\frac{1}{2}} && \text{(orthonormality, } \bar\Phi(t)\le\varphi(t)\text{, and } \tau\ge 1)\\ &\le e^{-\tau^2/4}\,. && (\bar\Phi(\tau) \le e^{-\tau^2/2})\end{aligned}$$

□

命题 *4.1* 的证明。 我们首先利用 $H_{d-1}(0)$ 展开推论 A.7：

$$\begin{aligned}&\int_0^\infty |1-\Pi_d\mathrm{sign}(x)|\,\mathrm{d}\varphi(x)\\ &= \int_0^\infty \left|1-(-1)^{\frac{d-1}{2}}\left(\frac{2}{\pi}\right)^{\frac{3}{4}}d^{\frac{1}{4}}\int_0^x\frac{H_d(t)}{t}\,\mathrm{d}t\right|\mathrm{d}\varphi(x) + O(d^{-\frac{3}{4}})\cdot\int_0^\infty\int_0^x\frac{|H_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x)\,.\end{aligned}$$

我们利用引理 A.10 在 $t \le d^{1/6}$ 时，以及引理 A.11 在 $t > d^{1/6}$时，对第二项进行估计：

$$d^{-\frac{3}{4}} \int_0^\infty \int_0^x \frac{|H_d(t)|}{t} \, \mathrm{d}t \, \mathrm{d}\varphi(x) \le d^{-\frac{3}{4}} \left(\int_0^\infty O(d^{1/4}) \cdot x e^{x^2/4} \, \mathrm{d}\varphi(x) + e^{-d^{1/3}/4} \right) \le O(d^{-\frac{1}{2}}) .$$

至此，第二项的分析结束。接下来，我们将第一项的内层积分拆分为 $t \in [0, \tau]$ 和 $t \in [\tau, \infty)$（当 $\tau = \sqrt{100 \log d}$时）。我们先考虑后者，根据 引理 A.11，它可以忽略不计：

$$d^{\frac{1}{4}} \int_0^\infty \int_\tau^x \frac{|H_d(t)|}{t} \, \mathrm{d}t \, \mathrm{d}\varphi(x) \le d^{\frac{1}{4}} e^{-\tau^2/4} = O(d^{-\frac{1}{2}}) .$$

我们转而考虑 $t \in [0, \tau]$，它是主要贡献项。将其分解如下：

$$\begin{aligned} \int_0^\infty \left| 1 - (-1)^{\frac{d-1}{2}} \left(\frac{2}{\pi}\right)^{\frac{3}{4}} d^{\frac{1}{4}} \int_0^{\min(x,\tau)} \frac{H_d(t)}{t} \, \mathrm{d}t \right| \mathrm{d}\varphi(x) &\le \int_0^\infty \left| 1 - \frac{2}{\pi} \int_0^{\min(x,\tau)} \frac{\sin(t\sqrt{d})}{t} \, \mathrm{d}t \right| \mathrm{d}\varphi(x) \\ &\quad + \frac{2}{\pi} \int_0^\infty \left| \int_0^{\min(x,\tau)} \sin(t\sqrt{d}) \cdot \frac{e^{t^2/4} - 1}{t} \right| \mathrm{d}\varphi(x) \\ &\quad + \frac{2}{\pi} \int_0^\infty \left| \int_0^{\min(x,\tau)} \frac{e^{t^2/4} r_d(t)}{t} \right| \mathrm{d}\varphi(x) . \end{aligned}$$

这里回顾一下，$r_d(t)$ 在（9）中定义。我们将分别估计这三项：

$$\int_0^\infty \left| 1 - \frac{2}{\pi} \int_0^{\min(\tau,x)} \frac{\sin(t\sqrt{d})}{t} \, \mathrm{d}t \right| \mathrm{d}\varphi(x) \le O\left(\frac{\log d}{\sqrt{d}} \right) \tag{11}$$

$$\int_0^\infty \left| \int_0^{\min(x,\tau)} \frac{e^{t^2/4} - 1}{t} \sin(t\sqrt{d}) \, \mathrm{d}t \right| \mathrm{d}\varphi(x) \le O\left(\frac{1}{\sqrt{d}} \right) \tag{12}$$

$$\int_0^\infty \int_0^{\min(x,\tau)} e^{t^2/4} \frac{|r_d(t)|}{t} \, \mathrm{d}t \, \mathrm{d}\varphi(x) \le O\left(\frac{1}{\sqrt{d}} \right) . \tag{13}$$

假设（11）、（12）和（13）成立，则由三角不等式可得命题 4.1 的证明。 □

（11）的证明)。 首先，通过变量替换可得：

$$\int_0^{\min(\tau,x)} \frac{\sin(t\sqrt{d})}{t} \, \mathrm{d}t = \int_0^{\min(\tau,x)\cdot\sqrt{d}} \frac{\sin u}{u} \, \mathrm{d}u .$$

断言 A.12 （例如，参见 [DLMF，§6。12]）。 存在一个普适常数 $C > 0$，使得对任意 $z > 0$,

$$\left| 1 - \frac{2}{\pi} \int_0^z \frac{\sin t}{t} \, \mathrm{d}t \right| \le C \min\left(1, \frac{1}{z}\right) .$$

应用来自断言 A.12的界，我们得到

$$\int_0^\infty \left| 1 - \frac{2}{\pi} \int_0^{\min(\tau,x)} \frac{\sin(t\sqrt{d})}{t} \, \mathrm{d}t \right| \mathrm{d}\varphi(x) \le O(1) \cdot \int_0^\infty \min\left(1, \frac{1}{\sqrt{d}\min(\tau, x)}\right) \mathrm{d}\varphi(x) .$$

我们进一步根据最小值中哪个参数占主导，将这个积分分成三部分：

1. 在 0 和 $\frac{1}{\sqrt{d}}$ 之间，根据第一个界，积分的上界为 $O\left(\frac{1}{\sqrt{d}}\right)$。

2. 在 $\frac{1}{\sqrt{d}}$ 和 τ 之间，$\min(\tau,\ x) = x$ 因此，根据第二个界，积分为 $O\left(\frac{\log d}{\sqrt{d}}\right)$。

3. 在 τ 和 $+\infty$ 之间，根据第二个界，积分为 $O\left(\frac{1}{\sqrt{d}}\right)$。

综合这些观察，我们完成了证明。 □

（12）的证明。设 $u(t) = \frac{e^{t^2/4}-1}{t}$ 通过连续性延拓，并令 $u(0) = 0$。计算可得

$$\max(|u(t)|, |u'(t)|) \le \frac{e^{t^2/4}}{2} \quad \text{for all } t \ge 0. \tag{14}$$

我们要研究的内层积分是 u 与频率为$\sqrt{d}$的振荡因子的乘积，因此，当 $d \to \infty$时，预计会发生显著的相消。我们通过分部积分将这一点明确写出：

$$\int_0^{\min(x,\tau)} \sin(t\sqrt{d})u(t)\,\mathrm{d}t = \left[-\frac{\cos(t\sqrt{d})}{\sqrt{d}}u(t)\right]_0^{\min(x,\tau)} + \frac{1}{\sqrt{d}}\int_0^{\min(x,\tau)} \cos(t\sqrt{d})u'(t)\,\mathrm{d}t\,.$$

因此，根据三角不等式和（14），

$$\left|\int_0^{\min(x,\tau)} \sin(t\sqrt{d})u(t)\,\mathrm{d}t\right| \le \frac{(1+x)e^{x^2/4}}{2\sqrt{d}}\,.$$

对右侧进行积分，得到 $\varphi\ O(d^{-\frac{1}{2}})$，正如所需。 □

证明 (13)。我们根据 $t \le d^{-1/2}$ 还是 $t > d^{-1/2}$，将内部积分拆分，并对第一部分应用 推论 A.9：

$$\int_0^\infty \int_0^{\min(x,\tau)} e^{t^2/4}\frac{|r_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x) \le O\left(\frac{1}{\sqrt{d}}\right)\cdot\int_0^\infty xe^{x^2/4}\,\mathrm{d}\varphi(x) + \int_{d^{-1/2}}^\infty \int_{d^{-1/2}}^{\min(x,\tau)} e^{t^2/4}\frac{|r_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x)\,.$$

第一项为 $O(d^{-1/2})$，对于第二项，我们使用富比尼定理：

$$\int_{d^{-1/2}}^\infty \int_{d^{-1/2}}^{\min(x,\tau)} e^{t^2/4}\frac{|r_d(t)|}{t}\,\mathrm{d}t\,\mathrm{d}\varphi(x) = \int_{d^{-1/2}}^\tau \frac{|r_d(t)|}{t}e^{t^2/4}\bar{\Phi}(t)\,\mathrm{d}t \le \int_{d^{-1/2}}^\tau \frac{|r_d(t)|}{t}e^{-t^2/4}\,\mathrm{d}t\,,$$

其中，我们采用了引理 A.11 的证明中所用的记号 $\Phi(t)$ 表示高斯尾概率，并应用标准上界 $\bar{\Phi}(t) \le e^{-t^2}/2$。

当 $t \le \tau$ 且 $t \ge \frac{1}{\sqrt{d}}$（即 $\frac{t}{\sqrt{d}} \ge \frac{1}{d}$）时，引理 A.6 给出的上界可化简为：

$$|r_d(t)| \le O\left(\frac{t^3}{\sqrt{d}} + \frac{t}{\sqrt{d}} + \frac{t^2}{d}\right).$$

将其代入上面的积分，我们得到：

$$\int_{d^{-1/2}}^\tau \frac{|r_d(t)|}{t}e^{-t^2/4}\,\mathrm{d}t \le O\left(\frac{1}{\sqrt{d}}\right),$$

如所愿。 □