arXiv:2601.05850v1 [cs.CC] 9 Jan 2026

低次启发式的严格推论

Jun-Ting Hsieh* Daniel M. Kane† Pravesh K. Kothari‡ Jerry Li§
Sidhanth Mohanty¶ Stefan Tiegel||

2026年1月12日

摘要

过去十年间，低次启发式 一直被用于估计各类平均情形下植入与零假设区分问题的算法阈值。这类结果依赖于这样一个假设：如果植入分布与零假设分布的低阶矩足够接近，那么就不存在能够区分二者的高效（耐噪）算法。由于低次似然比（LDLR）的计算十分简单——它衡量低阶矩之间的相似度——这一假设颇具吸引力。然而，尽管该领域持续受到关注，低次不可区分性是否真的能排除任何有意义的算法类别，仍不清楚。

在这项工作中，我们开启了这一方向的研究，并开发出将低次似然比上界转化为针对具体算法的严格下界的技术工具。因此，我们证明：对于任意置换不变分布 P，

1. 如果 P 定义在 $\{0,1\}^n$ 上，且与 $U = \mathrm{Unif}(\{0,1\}^n)$ 低次数不可区分，那么 P 的含噪版本与U在统计上不可区分。2. 如果 P 定义在 $\mathbb{R}^n$ 上，且与标准高斯分布 $\mathcal{N}(0,1)^n$ 低次数不可区分，那么没有任何基于次数至多为$O(\log n/\log\log n)$ 的对称多项式的统计量，能够区分 P 的含噪版本与 $\mathcal{N}(0,1)^n$。3. 如果 P 定义在 $\mathbb{R}^{n\times n}$ 上，且与 $\mathcal{N}(0,1)^{n\times n}$ 低次数不可区分，那么没有任何常数大小的子图统计量能够区分 P 的含噪版本与 $\mathcal{N}(0,1)^{n\times n}$。

*麻省理工学院。 juntingh@mit.edu。

†加州大学圣迭戈分校。 dakane@ucsd.edu。

‡普林斯顿大学。 kothari@cs.princeton.edu。§华盛顿大学西雅图分校。 jerryzli@cs.washington.edu。

¶西北大学。 sidhanth.mohanty@northwestern.edu。||麻省理工学院。 stefan23@mit.edu。

由瑞士国家科学基金会博士后流动资助（P500-2 235374）提供支撑

目录

1 引言 1
1.1 我们的结果.3

2 技术概述 5
2.1 背景与符号.5
2.2 二项式空间中的下界.6
2.3 高斯空间中的下界.6

3 技术预备知识 9
3.1 统计距离.9
3.2 伯努利分布与二项分布.10
3.3 正交多项式.10

4 二项空间：完全不可区分性 12
4.1 截断植入分布.13
4.2 定理 1.1 的证明 .15

5 高斯空间 I：低次对称多项式 16
5.1 ℓ_∞-密度的接近性.18
5.1.1 傅里叶衰减.20
5.1.2 频率匹配.23
5.2 小引理的证明.24

6 高斯空间 II：连通子图计数 27
6.1 含噪输入的子图计数与含噪子图计数接近.30
6.2 无穷范数界.36

A 二项分布 42
B Krawtchouk 多项式的上界 43
C 特征函数的界 44
D 子图多项式的矩 47

1 引言

过去十年来，人们对理解平均情形下的信号检测与恢复问题兴趣激增。这些问题是多个领域的核心挑战，包括算法设计（例如，在随机图中检测植入团）、统计学（例如，学习高斯混合的参数）和密码学（例如，区分伪随机生成器的输出与均匀随机比特）。这类问题通常可以建模为一对概率分布：具有可调节“强度”信号的植入分布，以及不含任何信号的零假设分布。主要研究问题在于，有效进行信号检测／恢复所需的最小信号强度是否超过统计学上的最低要求。数十年来的研究表明，信息—计算差距不仅普遍存在于上述基础问题中，也存在于学习理论、统计学和密码学中涌现的许多新问题里。现代平均情形复杂性领域致力于建立一套有原则的理论，以理解并预测这类差距。

我们应如何应对这种平均情形复杂性理论？[1] 类比于最坏情形困难性归约，一种策略是从一些典型的平均情形信号恢复问题出发，构建一个归约网络。这项工作已经取得了显著进展 [BR13、BB20、BH25]，但由于难以设计出将一个问题的实例分布转换为另一个问题的目标分布的归约方法，其成功仍局限于一组有限的问题。[2] 相反，平均情形复杂性中有关算法与统计差距的大多数证据都来自第二种方法——证明针对受限算法类的下界（类似于最坏情形复杂性中的电路下界）。

针对受限算法的困难性。一系列长期以来的研究聚焦于针对谱方法 [MRZ16、BMV+18、统计查询算法 [Kea98、FGR+17、DKS17、BBH+21]、马尔可夫链 [Jer92、CMRW23、CSZ24、CMZ25]，近似消息传递 [GJ21、GJW24]，凸规划（例如平方和（SoS）半定规划层级） [BHK+16、HKP+17、MRX20、GJJ+20、KB21、JPR+21、JPRX23、KPX24]的下界。特别是在 SoS 下界的语境中，伪校准方法 [BHK+16] 的发现表明，SoS 下界与植入分布和零假设分布之间的某种低次似然比（LDLR）存在紧密联系；这种低次似然比刻画了两个分布的低阶矩之间的接近程度。

定义 1.1（低次似然比与优势）。设 P（植入）和 N（零假设）是定义在 $\{0,1\}^{\binom{n}{k}}$ 上的两个概率分布。P 和 N 的似然比是函数 $L(x) = d\mathsf{P}/d\mathsf{N}(x)$，我们用 $L^{\leqslant D}$ 表示将 L 投影到次数不超过 $\leqslant D$ 的多项式张成空间所得的投影。

次　数-{style id='1'>D 优势 定义为 P 与 N 之间的 $\sqrt{\mathrm{Var}_{\mathsf{N}}[L^{\leqslant D}]}$，它也等于

$$\max_{f \text{ deg-}D \text{ polynomial}} \frac{\mathbb{E}_{\mathsf{P}} f - \mathbb{E}_{\mathsf{N}} f}{\sqrt{\mathrm{Var}_{\mathsf{N}} f}} . \tag{1}$$

我们用 $\chi^2_D(\mathsf{P} \| \mathsf{N})$ 表示 $\mathrm{Var}_{\mathsf{N}}[L^{\leqslant D}]$，即次数-{style id='16'>D 优势的平方。

换言之，次数-{style id='1'>D 优势衡量的是，在期望意义下，以 N 为基准归一化的次数-{style id='3'>D 多项式能够多好地区分 P 与 N。$\chi^2_D(\mathsf{P} \| \mathsf{N})$ 等于 (1) 的平方，这是一个标准计算；参见 Hopkins 的博士论文 [Hop18] 或 [KWB19]的综述。此外，$\chi^2_D(\mathsf{P} \| \mathsf{N})$ 实际上可解释为 P 与 N 之间 χ^2-散度的次数-{style id='29'>D 截断，因此采用了这一记号。此后，我们用“趋于零的低次似然比”表示 $\chi^2_D(\mathsf{P} \| \mathsf{N}) \leqslant o(1)$。

从 [HS17、HKP+17、Hop18]，等工作的开创以来，人们一直猜测，较小的低次似然比（LDLR）意味着假设检验问题在计算上具有困难性。Hopkins 的论文 [Hop18 中将一个精确的猜想正式表述为猜想 2.2：4]，大致而言，该猜想指出，对于对称分布，若低于 polylog(n) 次的优势很小，就意味着不存在多项式时间的耐噪声区分算法。

[1]经典的理论 [Lev86，BT06，Gol11] 关于平均情形复杂性的理论，无法帮助我们理解上述输入分布下的算法问题。[2]例如，尽管付出了大量努力，我们迄今仍未能将用 $n^{1+\Omega(1)}$ 个子句反驳随机 4-SAT 公式，归约为用 $n^{1+\Omega(1)}$ 个子句反驳随机 3-SAT 公式。

猜想 1.1（低阶假设（非正式表述） [Hop18]）。固定 $d \in \mathbb{N}$。令 N 为定义在 $\{0,1\}^{\binom{n}{d}}$ 上的均匀分布，或为标准高斯分布 $\mathcal{N}(0,1)^{\binom{n}{d}}$（视为对称 d 阶张量）。令 P 为定义在与 N 相同的定义域上的分布，且具有 S_n-对称性[3]，并满足 $\chi^2_D(P \| N) \leqslant o(1)$，其中 $D = (\log n)^{1+\Omega(1)}$。那么，不存在多项式时间可计算检验能够以 $1 - o(1)$ 的概率区分 $T_\varepsilon P$ 与 N。这里，$T_\varepsilon P$ 是 P 的含噪版本（精确定义见第 2 节第 2.1 小节）。

我们首先对猜想 1.1作几点说明。无论是S_n-对称性假设，还是向P 中添加噪声，都是必要的。若没有噪声，就会出现一些简单的反例：高斯消元或格基约化等脆弱算法便是高效区分器 [ZSWB22、DK22]。基于纠错码的反例也表明，对称性同样不可或缺 [HW21]。

在一项近期工作中 [BHJK25]在 [HW21], 的基础上，通过精心构造一个零假设与植入假设检验问题，推翻了（准多项式时间范围内的）Hopkins 猜想。尽管如此，低次似然比方法仍是推导平均情形问题困难性结论的常用方法；例如，参见 [KWB19、BHJK25、Wei25] 及其中列出的众多应用。此外，低次似然比方法的应用不止于算法设计，在统计学（用于推导估计问题的计算困难性）和密码学中也很常见 [ABI+23、BKR23、BJRZ24]（为安全性分析所依据的困难性猜想提供形式化依据）。

低次似然比框架的吸引力源于其简洁性和预测能力。一方面，对于大多数假设检验问题，计算次数为 D 的优势通常十分直接。另一方面，在广泛的平均情形问题中，低次似然比方法对计算困难性的预测，与人们推测的困难区域惊人地一致，这些问题包括植入团、植入着色、植入稠密子图问题、尖峰高斯矩阵与张量模型、随机子空间中的植入向量、稀疏 PCA、非高斯分量分析等（参见 [KWB19]的综述）。在这些问题中，低次似然比方法预测的计算困难性与最佳已知算法的失效点相吻合。

趋于零的低次似然比意味着什么? 尽管趋于零的低次似然比在预测计算阈值方面表现出非凡能力，但其含义仍未得到充分理解。Kunisky、Wein 和 Bandeira [KWB19] 的工作在这一方向上迈出了第一步。他们证明，当N 为独立同分布的 $\mathcal{N}(0,1)$ 或独立同分布的均匀分布$(\{\pm 1\})$[4], 时，若存在一个次数为$p : \mathbb{R}^n \to \mathbb{R}$，次数为 $\leqslant D/2$ 的多项式，满足

- 在 P 下具有较大的期望： $\mathbb{E}_{x\sim P}[p(x)] \geqslant A$,
- 在 N 下集中： $\mathbb{P}_{x\sim N}[p(x) \geqslant B] \leqslant e^{-\Omega(D)}$,

对某些 $A > B > 0$，则 $\chi^2_D(P \| N) \geqslant \Omega(1)$（见 [KWB19，定理 4.3]，具体表述见相关内容）。用逆否命题的形式来说，该结果意味着，趋于零的 $\chi^2_D(P \| N)$ 排除了这样的低次多项式的存在性：它们在 N 下集中，却在 P 下具有较大的期望。

然而，对于低次多项式期望的这类保证，并不能从形式上排除将低次多项式用作区分器。例如，即使在 $\chi^2_D(P \| N) = 0$（即所有次数为 D 的矩都匹配）的情形下，我们也只能知道，任意次数为 D 的多项式在这两个分布下的期望相同。但这种期望的匹配并不能排除这样的可能性：低次多项式的某个其他统计量可以区分 N 和 P。事实上，在最近一次关于该主题的研讨会 [AIM24] 中，有人指出，建立低次似然比方法的严格困难性推论是一个亟待研究的方向。

在这项工作中，我们开启了这一研究方向的初步探索。我们聚焦于或许最自然的第一个问题，并提出：

消失　　低次似然比正式意味着基于输入低次多项式的区分器失效　　*s?*

[3] S_n-对称性意味着分布在S_n的重新标记作用下保持不变。[4]证明只依赖于超收缩性，因此任何合理的分布都足够。

具体而言，我们希望证明，在猜想 1.1 的对称性条件下，并假设$\chi^2_D(\mathsf{P}\,\|\mathsf{N}) \leqslant o(1)$ 对于 $D = \omega(\log n)$，任意低次多项式 p 都满足 $d_{\mathrm{TV}}(p(N), p(T_\varepsilon P)) \leqslant o(1)$，其中常数 $\varepsilon > 0$ 任意。这将排除任何使用 p 的计算结果作为区分器的算法，无论其是否高效。在本文中，我们考虑一个更强的陈述：给定一组低次多项式$\mathcal{P} = (p_1, p_2, \ldots, p_\ell)$，我们希望它们在 N 和 $T_\varepsilon P$ 下的计算结果的联合分布是统计不可区分的，即 $d_{\mathrm{TV}}(\mathcal{P}(N), \mathcal{P}(T_\varepsilon P)) \leqslant o(1)$。

事实上，在布尔向量情形（$d = 1$ 属于 猜想 1.1），我们能够证明最强的可能保证：这两个分布本身是统计不可区分的：$d_{\mathrm{TV}}(\mathsf{N}, T_\varepsilon \mathsf{P}) \leqslant o(1)$（见 定理 1.1）。这一结果利用了 S_n-对称性对向量值分布施加的强限制。相比之下，对于矩阵值或张量值问题（$d \geqslant 2$ 属于 猜想 1.1），我们认为如此强的结论很可能不成立。

有界独立性加噪声。从技术角度来看，我们的结果与伪随机性文献密切相关。条件 $\chi^2_D(\mathsf{P}\,\|\mathsf{N}) \leqslant o(1)$ 可以看作 P 近似 D-重独立的一种类比，而我们研究的是 P的含噪版本。这类似于伪随机生成器（PRG）背景下研究的“有界独立性加噪声”范式[AW85、HLV18、FK18、LV20]。关键区别在于，这些工作着重于欺骗特定的函数类，即证明给定类中的函数在伪随机分布和真正随机分布下具有相近的期望。相比之下，我们旨在证明某些函数在这两种分布下的总变差距离很小，这是一种更强的保证。一个重要区别是，那些工作需要“伪随机噪声”来减少随机比特的数量，而在我们的设定中，根据 $\mathsf{PT}_\varepsilon\mathsf{P}$ 的定义，添加的噪声是独立同分布的。

尽管与上述内容存在差异，整体策略相似：给定一个函数的傅里叶展开，利用 D-路独立性来界定低次项，并利用噪声来界定高次项。在我们的情形中，为了证明全变差距离接近，我们需要对概率密度函数之间的差异给出逐点界，这些界是通过傅里叶逆变换以及对特征函数的界得到的 $\widehat{p}(\xi) = \mathbb{E}[e^{i\xi p}]$。我们遵循一般的框架：通过矩匹配界定低频项，通过噪声界定高频项；不过，分析特征函数的技术细节与此前关于伪随机生成器的工作有很大不同。详细概述见第 2 节。

我们还提及一系列仅依赖有界独立性（不加噪声）的研究，这些研究证明了有界独立性能够愚弄若干类函数，包括 AC^0 电路、半空间以及低次多项式阈值函数（在布尔和高斯情形下均有研究） [Bra08、DGJ+10、DKN10、Kan11、BIVW16、BDF+22]。不过我们强调，与低次似然比假设中的情形一样，在我们所有关于全变差距离接近性的结果中，加入噪声都是必要的——例如，存在一个离散分布 P，它与 $\mathsf{N} = \mathcal{N}(0, \mathbb{I}_n)$的所有低阶矩相匹配，但它的任何函数都必须具有较大的全变差距离。

1.1 我们的结果

在本节中，我们将使用 $T_\varepsilon P$ （对于布尔设置）和 $U_\varepsilon P$ （对于高斯情形）表示ε-含噪版本的植入分布P；参见第 2.1 节以了解精确定义。首先，我们将证明，具有有界 LDLR（含噪）的 S_n-对称分布在 $\{\pm 1\}^n$ 上实际上与零假设分布的全变差距离很小。

定理 1.1。令 $\gamma, \varepsilon \in (0,1)$ 为绝对常数，并令 $\delta \geqslant 0$。令 $D \in \mathbb{N}$ 满足 $D \geqslant C_\varepsilon^\gamma \log n$，其中$C\gamma$ 是仅依赖于 γ 的常数。令 $\mathsf{N} =$ 伯努利分布$(\gamma)^n$，并令 P 为关于 $\{\pm 1\}^n$ 的 S_n-对称分布，使得$\chi^2_D(\mathsf{P}\,\|\mathsf{N}) \leqslant \delta$。则，

$$d_{\mathrm{TV}}(\mathsf{N}, T_\varepsilon \mathsf{P}) \leqslant O_\gamma\left(\frac{\delta}{\varepsilon^2} + \frac{1}{\sqrt{n}\cdot\varepsilon^2}\right).$$

我们注意到 t 他的结果意味着，没有任何算法能够区分这两个分布。注意 t

我们证明了 N 与含噪版本 $\mathrm{T}_\varepsilon \mathrm{P}$ 之间的接近程度保证，而不是 P 本身。这是获得全变差保证所必需的，因为在没有噪声的情况下，存在简单的反例。

关键在于，这一结果利用了 S_n-对称性将问题化简为一维命题（更多细节见 第 2.2 节 ），因为任何 S_n-不变区分器都只能依赖输入的汉明重量。这使得分析可以借助研究低次下界时所用的相关工具。具体而言，我们可以通过给出 χ^2-距离（条件化后）的上界，来给出全变差距离的上界。

高斯空间。 与布尔对应情形相比，高斯情形要困难得多，因为S_n-对称性不足以将问题化简为一维问题。因此，高斯情形的分析难度显著增加。尽管如此，我们仍能得出以下两个结果。第一个涉及向量值情形，以及基于低次对称多项式的任意检验统计量。第二个涉及矩阵值情形，以及一类自然的检验统计量。

我们先讨论向量值情形。对于 S_n-对称的植入分布，自然会考虑基于低次对称多项式的区分器。然而，我们证明，基于此类多项式的统计量均无法成功。具体而言，如果低次似然比至多为 $o(k^{-O(k)}\log^{-1} n)$，那么任何此类检验都无法以非零概率区分 N 和 $\mathrm{U}_\varepsilon \mathrm{P}$。

定理 1.2 （见 定理 5.1 完整版本）。 设 $\varepsilon \in (0,1)$ 为一个常数。设 k，$D \in \mathbb{N}$ 满足 $D \geqslant C\,\varepsilon^2 k^2 \log n$ 且 $k \leqslant n^{\varepsilon}/\log n\, C\, \mathrm{loglog}\, n$，~~其中 $C \longrightarrow 0$~~ 为一个较大的普适常数。设 $\mathsf{N} = \mathcal{N}(0, \mathbb{I}_n)$，且 P 是 $\mathbb{R}^n$ 上的一个分布，满足 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) = \delta$。令 $\mathcal{P}(\cdot)$ 为次数至多为 k 的对称多项式向量。则有

$$d_{\mathrm{TV}}\left(\mathcal{P}(\mathsf{N}), \mathcal{P}(\mathrm{U}_\varepsilon \mathsf{P})\right) \leqslant \delta^{\Omega(1)} \cdot k^{k/2} + n^{-1/2}.$$

注意到，由于在定理 1.2中我们只考虑对称多项式，因此实际上不需要假设P 是S_n-对称的。

注 1.1。尽管证明这一点并非我们工作的重点，但我们认为，经过一些进一步的工作（例如，更仔细地分析定理 5.3），我们的证明有望得到局部中心极限定理（参见[Pet12]）。具体而言，经过对 $\mathcal{P}$作适当归一化后，我们期望得到 $\mathrm{d_{TV}}(\mathcal{P}(\mathsf{N}), \mathcal{N}(0, \mathbb{I}_k)) \leqslant o(1)$ 和 $\mathrm{d_{TV}}(\mathcal{P}(\mathrm{U}_\varepsilon \mathsf{P}), \mathcal{N}(0, \mathbb{I}_k)) \leqslant o(1)$，这将推出定理 1.2。

另一方面，这可以视为将定理 1.2 扩展到$k \geqslant \Omega(\log n)$所面临的固有障碍。我们预计，次数为 $\Omega(\log n)$的对称多项式（例如幂和多项式 $\sum_{i=1}^n x_i^k$）即使在 N 下也不满足局部中心极限定理。

矩阵模型中的子图计数。 在矩阵值情形中，我们考虑这样的情形：N 是由$n \times n$ 个对称矩阵组成的分布，这些矩阵的高斯条目相互独立且具有单位方差（对称位置上的条目除外）；而 P 是某个分布，满足 $\chi^2_D(\mathsf{P}\|\mathsf{N}) \leqslant \delta$ 其中$D \geqslant \log n \cdot \mathrm{loglog}\, n$。

我们研究带符号子图计数的区分能力。具体而言，对于一个有v个顶点和e 条边的连通图 ϑ，以及一个矩阵 M，我们将其带符号 ϑ计数 定义为

$$\chi_\vartheta(M) \coloneqq \sum_{\substack{H \text{ copy of } \vartheta \\ \text{in } K_n}} \prod_{ab \text{ edge in } H} M_{ab}.$$

关于带符号子图计数（缺乏）区分能力的主要结果是：任何基于常数大小子图的带符号子图计数的统计量，都无法区分 N 与 P 的含噪版本。形式化地说：

定理 1.3 (参见 定理 6.1 完整版本)。 令 ν 表示当 $\boldsymbol{M} \sim \mathsf{N}$ 时 $\chi_\vartheta(\boldsymbol{M})$ 的分布，并令 π 表示当 $\boldsymbol{M} \sim \mathrm{U}_\varepsilon \mathsf{P}$ 时其分布，其中 $\varepsilon > 0$ 是一个绝对常数。假设 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) \leqslant \delta$，其中 $D \geqslant \log n \cdot \mathrm{loglog}\, n$。存在一个足够小的常数 $\alpha > 0$，使得：

$$d_{\mathrm{TV}}(\nu, \pi) \leqslant O\left(\delta^{\alpha\varepsilon^2} + n^{-\alpha\varepsilon^2}\right),$$

其中 $O(\cdot)$ 隐藏了依赖于 ε 和 ϑ的常数因子。

注 1.2。 在 P $=$ N这一特殊情形下，我们的结果建立了关于高斯矩阵中子图计数多项式的 局部中心极限定理 。为埃尔德什–雷尼随机图建立此类结果，人们已付出大量努力 ˝os–R´enyi 随机图 [GK16、Ber18、SS 22] ，通过对这些分布的傅里叶谱进行“赤手空拳”的分析来实现。我们对高斯情形下傅里叶谱的分析，将其与 Chatterjee 的二阶 Poincar´e 不等式 [Cha09], 以及在平方和下界背景下研究的图矩阵算子范数界联系起来 [AMP16、 BHK+16]。

注 1.3。 为了排除最先进的算法（如谱方法）或基于矩阵最大元素的区分器，我们需要一个适用于 定理 1.3 且对 $O(\log n)$ 规模子图成立的版本。尽管要实现这一点仍有技术障碍需要克服（参见 注 1.1），我们的结果是朝这一方向迈出的第一步。

例如，考虑 Wigner 稀疏主成分分析问题，其植入分布由$\lambda uu^\top + W$ 描述，其中 $\lambda > 0$， W $\sim$ N，且 u 是一个 ℓ稀疏单位向量，其非零项例如为 $\pm 1/\sqrt{\ell}$，并且 $\ell \ll \sqrt{n}$。注意，在这种情况下，选择 $\varepsilon = \Theta(1)$ 时，噪声算子只会使信号强度λ 改变一个很小的常数。我们关注的是， λ 需要多大，才能区分 N 和 P。从统计角度看，当且仅当 $\lambda = \widetilde{\Omega}(\sqrt{\ell})$ 时，这是可能的。另一方面，所有已知的高效算法都要求 $\lambda = \Omega(\ell)$ [JL09]。

具体而言，考虑 $\ell = n^{0.4}$ 且 $\lambda = n^{0.3}$ 的情形。此时，从统计角度看，N 和 P 是可区分的，因为 $\lambda \gg n^{0.2}$。然而，已知算法无法成功，因为 $\lambda \ll \widetilde{\Omega}(n^{0.4})$。事实上， [DKWB24 中定理 2.24 (a) 的证明]已验证，在这种情形下，次数至多为 $o(n)$ 的低次似然比被 $o(1)$ 所界定。因此，我们的定理意味着，任何基于单个常数大小子图统计量的检验，其成功概率必然趋于零。

2 技术概述

2.1 背景与符号

我们使用小写字母表示标量和向量，使用大写字母表示矩阵。在本文中，我们将用g （表示标量和向量）或 G （表示矩阵）表示标准正态随机变量。
我们首先定义所使用的噪声模型。

定义 2.1（噪声模型）。

1. 向量值二项分布情形：设 P 是定义在 $\{\pm 1\}^n$ 和 y $\sim$ P 上的分布。设 z $\sim$ N 与 y 相互独立。我们用 TεP 表示 ˜y 的分布，其中对每个 i，我们独立地以 $1-\varepsilon$ 的概率设置˜$y_i = y_i$，并以 ε 的概率设置 ˜ $y_i = z_i$。2. 向量值高斯情形：设 P 是定义在 $\mathbb{R}^n$ 和 y $\sim$ P 上的分布。设 g $\sim \mathcal{N}(0, I_n)$ 与 y 相互独立。我们用 Uε P 表示 $\sqrt{1-\varepsilon}$y $+ \sqrt{\varepsilon}$g 的分布。3. 矩阵值高斯情形：设 P 是定义在 $\mathbb{R}^{n\times n}$ 中的对称矩阵上，且 M $\sim$P。设 G $\in \mathbb{R}^{n\times n}$ 为对称矩阵，与 M 相互独立，并具有独立同分布的 $\mathcal{N}(0,1)$ 条目（根据对称性取值）。我们用 Uε P 表示 $\sqrt{1-\varepsilon}$M $+ \sqrt{\varepsilon}$G 的分布。

傅里叶变换。 对于任意足够良好的函数 $f\colon \mathbb{R}^d \to \mathbb{R}$，我们记 $\widehat{f(\xi)}$ 为f的标准傅里叶变换:

$$\widehat{f}(\xi) = \tfrac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} e^{-i\langle \xi, x\rangle} f(x)\, dx\,.$$

对于任意足够光滑的、取值于 $\mathbb{R}^d$ 的随机变量 x，令 π_x 表示 x 的概率密度函数，令 $\widehat{\pi_x}$ 表示 x 的傅里叶变换:

$$\widehat{\pi_x}(\xi) = \mathop{\mathbb{E}}_{x} [\exp(i\langle \xi, x\rangle)]\,.$$

等价地， $\widehat{\pi_x}(\xi)$ 是 π_x 在 ξ处的傅里叶变换。[5]

2.2 BinomialSpace 中的下界

对于取值于 $\{\pm 1\}^n$ 的 S_n-对称分布，定理 1.1 的分析可归结为研究一个统计量——元素之和。在此概述中，我们稍微滥用一下记号，用 N 和 P 表示元素之和的分布。零假设分布恰为二项分布 $\mathrm{Bin}(n,\gamma)$，相应的正交归一多项式是（平移后的）克拉夫丘克多项式 $(\mathrm{Kr}_i)_{i\leqslant n}$，它们取决于 γ 和 n （见第 3.3 节）。注意，根据中心极限定理，$\mathrm{Bin}(n,\gamma)$ 收敛于平移高斯分布，因此我们预期 Kr_i 的行为大致类似于埃尔米特多项式。

自然的策略是对 $\chi^2(\mathrm{T}_\varepsilon \mathrm{P} \,\|\mathrm{N})$给出上界，这将推出 TV distance 的上界。给定正交归一基，我们可以写出

$$\chi^2(\mathrm{T}_\varepsilon \mathrm{P} \,\|\, \mathrm{N}) = \sum_{i\geqslant 1}\left(\mathop{\mathbb{E}}_{\mathrm{T}_\varepsilon \mathrm{P}}[\mathrm{Kr}_i]\right)^2 = \sum_{i\geqslant 1}(1-\varepsilon)^i\left(\mathop{\mathbb{E}}_{\mathrm{P}}[\mathrm{Kr}_i]\right)^2.$$

考虑到 $(1-\varepsilon)^{\,i}$ 衰减，将求和拆分为低次数部分和高次数部分是很自然的。对于低次数部分，有界 LDLR，即 $\chi^2_D(\mathrm{P}\,\|\mathrm{N}) \leqslant \delta$，意味着 $\sum_{i=1}^{D}(1-\varepsilon)^i\,\mathbb{E}\mathrm{P}[\mathrm{Kr}_i]^2 \leqslant \delta$。对于高次数部分，我们有 $\sum_{i>D}(1-\varepsilon)^i(\mathbb{E}\mathrm{P}[\mathrm{Kr}_i])^2 \leqslant \mathrm{e}^{-\varepsilon}\mathrm{D}\sum_{i\geqslant 1}(\mathbb{E}\mathrm{P}[\mathrm{Kr}_i])^2$,，并注意到 $\sum_{i\geqslant 1}(\mathbb{E}\mathrm{P}[\mathrm{Kr}_i])^2$ 恰好是 $\chi^2(\mathrm{P}\,\|\mathrm{N})$。

遗憾的是，先验地，我们无法对 $\chi^2(\mathrm{P}\,\|\mathrm{N})$ 给出任何上界，因为 P 可能存在一些概率很小的事件，使 χ^2-散度急剧增大。为绕过这一障碍，我们注意到，只需匹配低阶矩，就足以为 P 获得良好的尾界。因此，我们考虑截断分布 P，即将 P 限制在 $[-\tau,\tau]$ for $\tau=\Theta(\sqrt{\log n})$ 范围内得到的条件分布——这是标准高斯分布的典型区域。即使在 P 下，该事件的概率也为 $1-\frac{1}{\mathrm{poly}(n)}$，因此 $d_{\mathrm{TV}}(\mathrm{P},\mathrm{P}) \leqslant \frac{1}{\mathrm{poly}(n)}$。

经过截断后，容易看出 $\chi^2(\mathrm{P}\,\|\,\mathrm{N}) \leqslant \mathrm{poly(n)}$。因此，令 $\mathrm{D} = \mathrm{O}(\frac{1}{\varepsilon}\log \mathrm{n})$ 就足以使 $(1-\varepsilon)^{\mathrm{D}}\,\mathrm{poly(n)} \leqslant \mathrm{o(1)}$。需要仔细证明，截断后 $\mathbb{E}\mathrm{P}[\mathrm{Kr}_i]$ 对 $\mathrm{i} \leqslant \mathrm{D}$ 仍然很小。我们使用低次似然比和克拉夫丘克多项式的性质，在定理 4.6 中证明这一点。

2.3 高斯空间中的下界

接下来，我们讨论关于高斯空间中低次不可区分性的结果。如前所述，这个空间中的情况比二项式空间中的情况复杂得多。直观地说，这是因为 S_n-对称性在实数上远不如在离散分布中那样严格。为了证明这个空间中的界，我们大幅偏离了低次似然比下界文献中的“标准”技术。我们不再像上一节那样，将全变差距离与类似 χ^2-散度的量联系起来，而是直接界定高斯变量的多项式与含噪植入分布的多项式的概率密度函数之间的 L_1-距离。值得注意的是，我们的策略与基于傅里叶反演的局部极限定理的基本思路相似[SM62、Pet12、GK16、Ber18、SS22]。

作为贯穿全文的示例，本节将概述我们如何针对单个假设的 k 次多项式 $\mathrm{p}:\ \mathbb{R}^n \to \mathbb{R}$完成这一过程。需要注意的是，对于我们的实际结果，我们实际上必须考虑任意的向量值对称多项式，以及定义在矩阵输入上的结构化多项式，即 $\mathrm{p}:\ \mathbb{R}^{n\times n} \to \mathbb{R}$。随着差异出现，我们将解释实际技术与这一简化论述之间的不同之处。令 $\mathrm{y} \sim \mathrm{P}$ 为一个随机变量，使得 $\chi^2_D(\mathrm{P}\,\|\mathrm{N}) \leqslant \delta$ 对于某个充分大的 D 成立，并令 ε为一个较小的常数。令 $\mathrm{z} = \mathrm{U}_\varepsilon(\mathrm{y})$ 为 y 的含噪版本。我们的目标是证明 p(g) 与 p(z) 之间的全变差距离极小。为使论述稍微简化一些，假设 y 不仅与高斯分布的 D 阶矩近似匹配，事实上还精确匹配所有次数至多为 D 的矩。此外，我们将 p 归一化，使得 $\mathrm{Var}[\mathrm{p(g)}] = 1$，并令 $\mathrm{T} = \mathrm{D/k}$。

[5]注意，这与x的特征函数一致。

第 1 步：从全变差距离到傅里叶界。 我们的第一步相当标准。令 R 为稍后确定的量。全变差距离在相差一个 1/2 的因子下，可以写成 $p(g)$ 与 $p(z)$ 之间的

$$\int_{\mathbb{R}} \left|\pi_{p(g)}(x) - \pi_{p(z)}(x)\right| dx = \int_{|x|\leqslant R} \left|\pi_{p(g)}(x) - \pi_{p(z)}(x)\right| dx + \int_{|x|>R} \left|\pi_{p(g)}(x) - \pi_{p(z)}(x)\right| dx\,. \quad (2)$$

根据标准集中界，我们知道，由于p 具有单位方差且次数较低，若选择$R = \omega(1)$，则第二项趋于零，因为 $p(g)$ 和$p(z)$ 落在某个常数长度区间之外的概率质量都趋于零。

为了界定第一项，我们的策略是证明点态界 $|\pi_{p(g)}(x) - \pi_{p(z)}(x)|$。为证明这样的界，我们主要通过分析 $p(g)$ 和$p(z)$的傅里叶变换。具体而言，对于任意固定的y，根据傅里叶逆变换，我们有

$$\begin{aligned} 2\pi \cdot |\pi_{p(g)}(y) - \pi_{p(z)}(y)| &= \left| \int_{\mathbb{R}} e^{i\xi y} \left(\widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right) d\xi \right| \\ &\leqslant \int_{\mathbb{R}} \left| \widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right| d\xi \\ &= \underbrace{\int_{|\xi|\leqslant R'} \left| \widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right| d\xi}_{\text{(A)}} + \underbrace{\int_{|\xi|>R'} \left| \widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right| d\xi}_{\text{(B)}}, \end{aligned} \quad (3)$$

对于参数 R'作适当选择即可。现在只需界定式（3）中的两项。

步骤 2：频率匹配。 我们先说明如何界定式（3）中的 **(A)**。这是使用矩匹配的情形。回顾一下，$T = D/k$。根据泰勒级数展开，我们得到

$$\begin{aligned} \widehat{p(g)}(\xi) &= \mathbb{E}[\exp(i\xi p(g))] \\ &= \sum_{\ell=0}^{T-1} \mathbb{E}\left[\frac{(i\xi p(g))^\ell}{\ell!}\right] + O\left(\mathbb{E}\left[\frac{(\xi p(g))^T}{T!}\right]\right), \end{aligned}$$

同理

$$\widehat{p(z)}(\xi) = \sum_{\ell=0}^{T-1} \mathbb{E}\left[\frac{(i\xi p(z))^\ell}{\ell!}\right] + O\left(\mathbb{E}\left[\frac{(\xi p(z))^T}{T!}\right]\right).$$

由于G 和Z 匹配D 阶矩，因此有 $\mathbb{E}\left[p(z)^\ell\right] = \mathbb{E}\left[p(g)^\ell\right]$ 对所有 $\ell \leqslant D/k$均成立。因此，

$$\left| \widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right| \leqslant O\left(\mathbb{E}\left[\frac{(\xi p(g))^T}{T!}\right]\right).$$

我们对 p 的关键要求是：$p(g)$ 的T 次矩必须严格慢于$O(T^T)$增长；也就是说，$p(g)$ 的T 次矩必须小于次指数随机变量的T 次矩。例如，如果$p(g)$ 的T阶矩的增长方式与次高斯随机变量相同，即 $\mathbb{E}[p(g)^T] = O(T^{T/2})$，那么我们就有

$$\left| \widehat{p(g)}(\xi) - \widehat{p(z)}(\xi) \right| \leqslant O\left(\frac{|\xi|^2}{T}\right)^{T/2}, \quad (4)$$

因此我们将得到

$$\text{(A)} \leqslant O(T) \cdot \exp(-\Omega(T)) \ll 1,$$

只要 $|\text{实数域}'|^2 \leqslant O(T)$。

对于我们考虑的具有矩阵值输入的结构化多项式（称为子图计数多项式），文献中已知此类矩界适用于 T 为较大的常数时（例如，参见 [SS22]）。遗憾的是，我们需要 n 缓慢增长的 T（即 $O(\log n/\log\log n)$），因此这些界并不足够。我们将通过复杂的组合论论证，手动推导这些界。

对于具有向量值输入的低次对称多项式，情况要复杂一些。我们不仅需要同时控制所有对称多项式的整个分布，而且，即便是单个高斯随机变量的低次对称多项式，也显然并非总具有次高斯矩。例如，考虑以下二次对称多项式

$$p(x) = \left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} x_i\right)^2 .$$

Then, $p(g)$ 分布 为卡方随机变量，显然其不具有次高斯尾部 .

不过，我们发现，尽管并非所有对称多项式都具有次高斯尾部，但确实存在一类性质良好的对称多项式，它们具有这种尾部行为，并且构成所有低次对称多项式的充分统计量。具体而言，我们将利用这样一个事实：任意对称的k 次多项式都可以表示为以下多项式的确定性函数

$$\phi_\ell = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} h_\ell(x_j) , \tag{5}$$

对 $\ell = 1, \ldots, k$，其中 h_ℓ 是 ℓ 次（概率论中的）埃尔米特多项式。请注意，这是一类不太标准的对称多项式，不过事实证明，它非常适合我们的目的。关键在于，$\phi_\ell(g)$ 是 n 个独立随机变量之和，因此很容易验证，只要$k \leqslant O(\log n/\log\log n)$，那么对于我们所需的所有$T$ 值，这些多项式都具有次高斯矩界。

我们的策略是直接证明向量值多项式 $F_k\colon \mathbb{R}^n \to \mathbb{R}^k$的分布在 g 和 z 下在全变差距离意义下彼此接近，具体方法是证明它们的多元傅里叶变换逐点接近。其中，$F_k(g)$ 和 $F_k(z)$ 的分布彼此接近。根据数据处理不等式，这意味着 g 和 z 下所有对称低次多项式的联合分布彼此接近。这又立即意味着，不存在任何能够利用任意数量的对称低次信息来区分 g 和 z的区分器。我们强调，事实证明，这种充分统计量的选择对于我们的证明非常重要：它不仅给出了这里所需的矩界，而且这些多项式中的独立性结构对于获得正确的傅里叶衰减界也很重要，下面将对此进行解释。

步骤 3：傅里叶衰减。最后，我们转向对（3）中的项 (B) 进行界定。我们首先写出

$$(\mathrm{B}) \leqslant \int_{|\xi|>T} \left|\widehat{p(g)}(\xi)\right| d\xi + \int_{|\xi|>T} \left|\widehat{p(z)}(\xi)\right| d\xi ,$$

我们将分别对每一项进行界定，尽管第二项的控制几乎严格地比第一项更困难，因此这里我们将主要讨论如何控制第二项。

也就是说，最后剩下的概念性挑战是证明$p(z)$具有傅里叶衰减性。在这里，我们必须设法利用$z = \mathrm{U}_\varepsilon[y]$, 这一事实，即$z$ 中内置了高斯噪声。但仅凭这一点还不够，因为我们还需要控制y的行为。事实证明，在我们考虑的两种情形中，最方便的做法都不是直接处理$p(z)$，而是处理一个在全变差距离下接近$p(z)$的随机变量，并且该随机变量满足某些正则性条件。因此，在实际证明中，我们并不是按照上文概述的方式对$p(z)$执行步骤1—3，而是对$p(z)$的正则化版本执行这些步骤，最后再利用三角不等式，将所得的界与$p(z)$ 联系起来。

我们所要施加的正则性确切概念将取决于具体情形。在对称情形下，回顾我们使用的是多项式基 ϕ_ℓ，其定义见 方程 (5)，并且该基

具有一个很好的性质：当输入为高斯分布时，这些多项式是 n 个独立随机变量之和。由于这种独立结构，事实证明，傅里叶系数的大小主要取决于多项式 $q_j(x) = h_\ell(\sqrt{1-\varepsilon^2}\boldsymbol{y}_j + \varepsilon x)$ 在高斯输入 x 下对"典型"$\boldsymbol{y}$ 的方差。尽管对于所有坐标 $j = 1, \ldots, d$，该方差并不一定会对典型的 $\boldsymbol{y}$ 得到良好控制，但我们可以证明，以对 $\boldsymbol{y}$ 的选择而言的高概率下，存在许多坐标使得该方差得到良好控制。因此，如果我们对该事件发生的条件化进行处理，就能够得到原多项式傅里叶系数所需的界。总结—

因此，我们证明存在一个关于 $\boldsymbol{y}$ 的高概率事件，使得在以该事件为条件时，$p(\mathrm{U}_\varepsilon[\boldsymbol{y}])(\zeta)$ 有界。因此，在这里，$p(\mathrm{U}_\varepsilon[\boldsymbol{y}])$ 的正则化版本就是 $p(\mathrm{U}_\varepsilon[\tilde{\boldsymbol{y}}])$，其中 $\tilde{\boldsymbol{y}}$ 服从与 $\boldsymbol{y}$ 相同的分布，并以该高概率事件为条件。

子图计数多项式。对于子图计数多项式，不幸的是，我们失去了这种独立性结构，因此无法使用这类论证来界定傅里叶系数。然而，对于任意一个足够良好的函数 f，Chatterjee 给出了强有力的局部极限定理（见定理 6.2），指出 $\mathrm{f}(\mathrm{U}_\varepsilon[\mathrm{Y}])$ 在全变差距离下接近 $\mathrm{f}(\mathrm{Y}) + \sigma\mathrm{G}'$，其中 G' 是高斯变量，σ 是某个适当的尺度因子。由于计算高斯分布的傅里叶变换很容易，在这种情形下，我们可以将其作为 $\mathrm{f}(\mathrm{U}_\varepsilon[\mathrm{Y}])$ 的正则化版本。这里的主要技术挑战是证明 Chatterjee 的界足够紧，从而能为这些子图计数多项式得到一个非平凡的局部极限定理（LLT）；这需要界定梯度的 ℓ_2-范数、海森矩阵的谱范数，并给出这些多项式方差的下界。为此，需要处理相当复杂的图矩阵矩；我们通过仔细的组合界以及最初由 [AMP16] 为证明平方和下界而发展出的技术来控制这些矩。

备注：与局部极限定理的关系。简要补充一点，这类傅里叶衰减界与局部极限定理（LLT）密切相关，因此局部极限定理出现在我们的证明中也就不足为奇了。局部极限定理是概率论中经典中心极限定理（CLT）的强化版本。中心极限定理，或者更一般地说，不变性原理指出，对于具有"良好"独立同分布坐标的随机变量 $\boldsymbol{x} \in \mathbb{R}^n$，以及"良好"函数 $f\colon \mathbb{R}^n \to \mathbb{R}$，$f(\boldsymbol{x})$ 的分布与经过适当缩放的高斯分布接近，也就是说，它的累积分布函数在所有位置都接近高斯分布的累积分布函数。相比之下，局部极限定理指出，$f(\boldsymbol{x})$ 的概率密度函数 在所有位置都接近高斯分布的概率密度函数。特别是，这意味着$f(\boldsymbol{x})$ 的所有局部行为也与高斯分布相似，因此得名。由于高斯分布的高阶傅里叶系数呈指数级衰减，要证明局部极限定理，通常需要先证明$f(\boldsymbol{x})$ 的高阶傅里叶系数也呈指数级小。虽然已有此类局部极限定理适用于某些有限类别的多项式，但遗憾的是，这些结果不足以满足我们的需要，因此我们针对证明中出现的多项式，手动证明这些傅里叶衰减界。事实上，我们的界可以解释为一种新的局部极限定理，适用于形如$p(\mathrm{U}_\varepsilon[\boldsymbol{y}])$的随机变量。

3 技术预备知识

3.1 统计距离

定义 3.1. χ^2-散度，记作 $\chi^2(P \,\|\, Q)$，定义为

$$\chi^2(P \,\|\, Q) = \mathop{\mathbb{E}}_{x\sim Q}\left[\left(\frac{P(\boldsymbol{x})}{Q(\boldsymbol{x})} - 1\right)^2\right] = \mathop{\mathbb{E}}_{x\sim Q}\left[\left(\frac{P(\boldsymbol{x})}{Q(\boldsymbol{x})}\right)^2\right] - 1\,.$$

换言之，$\chi^2(P \,\|\, Q)$ 是似然比 P/Q 在 Q 下的方差。

请注意，χ^2-散度并非距离，因为它不对称。不过，根据詹森不等式，我们可以得到以下简单事实，将 χ^2-散度与全变差距离联系起来。

事实 3.1。$d_{\mathrm{TV}}(P、Q) \leqslant \frac{1}{2}\sqrt{\chi^2(P \,\|\, Q)}$。

当Q是Ω^n上的乘积分布时，它具有一个相关的标准正交基$\{\psi_\alpha\}_{\alpha\in\mathbb{N}^n}$，其对应函数$f$：$\Omega^n \to \mathbb{R}$。于是，我们得到$\chi^2$-散度的如下表达式。

引理 3.2。设Q是Ω^n上的一个乘积分布，$\{\psi_\alpha\}_{\alpha\in\mathbb{N}^n}$是关于$Q$的正交归一基。那么，$\chi^2(P\,\|\,Q)=\sum_{\alpha\neq 0}(\mathbb{E}_{x\sim P}[\psi_\alpha(x)])^2$。

证明。令$\mathrm{L(x)} = \mathrm{P(x)/Q(x)}$。我们可以用该基表示L：$\mathrm{L(x)}=\sum_\alpha \widehat{\mathrm{L}}(\alpha)\psi_\alpha(\mathrm{x})$，其中$\widehat{\mathrm{L}}(\alpha) = \mathbb{E}_{x\sim Q}[\mathrm{L(x)}\psi_\alpha(\mathrm{x})] = \sum_{x\in\Omega^n}\mathrm{Q(x)}\cdot\mathrm{P(x)}\,\mathrm{Q(x)}^{\psi_\alpha(\mathrm{x})} = \mathbb{E}_{x\sim P}[\psi_\alpha(\mathrm{x})]$。我们还有$\widehat{\mathrm{L}}(0)=\mathbb{E}_Q[\mathrm{L}]=1$。

根据 定义 3.1，我们可以将$\chi^2(P\,\|\,Q)$写作$\mathbb{E}_Q[L^2]-1$。根据帕塞瓦尔定理以及上述对$\widehat{L}(\alpha)$的表达式，可得$\chi^2(P\,\|\,Q)=\sum_\alpha \widehat{L}(\alpha)^2-1=\sum_{\alpha\neq 0}(\mathbb{E}_P[\psi_\alpha])^2$。 □

以下是一个简单但至关重要的结论，我们将需要利用它来界定两个分布之间的χ^2-散度。

引理 3.3。假设$\mathrm{supp}(P)\subseteq\mathrm{supp}(Q)$，则$\chi^2(P\,\|\,Q)\leqslant \frac{1}{\min_{x\in\mathrm{supp}(P)}Q(x)}$。

Proof. From Definition 3.1, $\chi^2(P\,\|\,Q)=\mathbb{E}_Q[(P/Q)^2]-1=\sum_{x\in\mathrm{supp}(P)}\frac{P(x)^2}{Q(x)}-1\leqslant\frac{1}{\min_{x\in\mathrm{supp}(P)}Q(x)}$. □

我们还需要以下著名的数据处理不等式：

事实 3.4 (数据处理不等式)。如果κ是一个任意信道，分别将P和Q转换为P_κ和Q_κ，那么$d_{\mathrm{TV}}(P_\kappa,Q_\kappa)\leqslant d_{\mathrm{TV}}(P,Q)$。

3.2 伯努利分布和二项分布

以下是一个由泰勒展开推导出的标准事实。

事实 3.5。记伯努利分布(γ)与伯努利分布$(\tilde{\gamma})$之间的KL散度为$D(\gamma\|\tilde{\gamma}):=\gamma\log\frac{\gamma}{\tilde{\gamma}}+(1-\gamma)\log\frac{1-\gamma}{1-\tilde{\gamma}}$。则对于$\gamma\leqslant 1/2$和$\varepsilon\leqslant\gamma/2$，有$D(\gamma+\varepsilon\|\gamma)\leqslant\frac{\varepsilon^2}{2\gamma(1-\gamma)}+O(|\varepsilon|^3/\gamma^2)$。

以下是一个民间结果。为完整起见，我们在附录 A 中给出证明。

引理 3.6。设$\gamma\leqslant 1/2$。对于任意$y\in\mathbb{R}$，若满足$|y|\leqslant o(\sqrt{\gamma n})$且$n\gamma+y\sqrt{n\gamma(1-\gamma)}\in\{0,1,\ldots,n\}$，

$$\mathbb{P}\left[\mathrm{Bin}(n,\gamma)=n\gamma+y\sqrt{n\gamma(1-\gamma)}\right]\geqslant\frac{1}{\sqrt{2n}}\exp\left(-\frac{y^2}{2}-O(|y|^3/\sqrt{\gamma n})\right).$$

我们给出了一个定量版本的著名事实：中心化二项随机变量的矩与标准高斯分布的矩相匹配，并在附录 A中给出证明。

引理 3.7 (二项分布矩)。令$\gamma\leqslant 1/2$，并令$\boldsymbol{x}\sim$伯努利分布$(\gamma)^n$且$\boldsymbol{y}=\frac{1}{2\sqrt{n\gamma(1-\gamma)}}\sum_{i=1}^n(\boldsymbol{x}_i-(2\gamma-1))$。那么，对任意$k\in\mathbb{N}$, $\mathbb{E}[\boldsymbol{y}^{2k}]\leqslant(2k-1)!!\cdot(1+O(k^3\gamma n))$。

以下事实是标准结果；例如参见 [O'D14，定理 10。21]。

引理 3.8 (超收缩性)。固定一个常数$\gamma\in(0,1)$。存在常数$C_\gamma>0$，使得对于任意次数为d的n元多项式p，$\mathbb{E}_{x\sim\mathrm{Ber}(\gamma)^n}[p(x)^4]\leqslant C_\gamma^d\cdot\mathbb{E}_{x\sim\mathrm{Ber}(\gamma)^n}[p(x)^2]^2$。

3.3 正交多项式

本文将根据具体应用，关键性地使用几类正交多项式，并在此回顾这些多项式。

概率论者的埃尔米特多项式 对于非负整数 k，令 $h_k:\ \mathbb{R}\to\mathbb{R}$ 表示 k 阶标准化（概率论者的）埃尔米特多项式 [Sze75]:

$$h_k(x) := \frac{1}{\sqrt{k!}}\mathrm{He}_k(x) = \frac{1}{\sqrt{k!}}(-1)^k e^{x^2/2}\frac{d^k}{dx^k}e^{-x^2/2}\,.$$

这里，He 是（未归一化）埃尔米特多项式的通常记号。埃尔米特多项式的一个标准性质是，对于所有 $k \geqslant 0$，多项式 $\{h_\ell\}_{\ell=0}^k$ 构成关于高斯内积 $\langle\cdot,\cdot\rangle$的所有次数为 k 的多项式的完备正交归一基，该内积定义为 $\langle p,q\rangle = \mathbb{E}_{g\sim\mathcal{N}(0,1)}[p(g)q(g)]$。特别地，

$$\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[h_k(g)h_\ell(g)] = \mathbf{1}(k=\ell)\,.$$

此外，对于所有 x，我们有递推关系 $\mathrm{He}'_\ell(x) = \ell\cdot\mathrm{He}_{\ell-1}(x)$，因此对于所有$\ell\geqslant 1$，有 $h'_\ell(x) = \sqrt{\ell}\cdot h_{\ell-1}(x)$。

我们还将使用以下标准估计，这些估计将很有用：

引理 3.9。设 $g\sim\mathcal{N}(0,1)$。对于任意次数为 k 的多项式 p，有 $\mathrm{Var}[p(g)]\leqslant\mathbb{E}[p'(g)^2]\leqslant k\cdot\mathrm{Var}[p(g)]$。证明。将 $p(g)=\sum_{\ell=0}^k a_\ell h_\ell(g)$ 写成上述形式，则 $\mathrm{Var}[p(g)]=\sum_{\ell=1}^k a_\ell^2$。于是，由于对于所有 $\ell>0$，$h'_\ell=\sqrt{\ell}\cdot h_{\ell-1}$，且 $h'_0=0$，我们有 $\mathbb{E}[p'(g)^2]=\sum_{\ell=1}^k \ell a_\ell^2$，从而结论显然成立。 □

我们回顾高斯多项式的超收缩性，例如参见 [O'D14，定理 9.21 以及第 9] 和 [Nel73] 的注记。

事实 3.10（高斯超收缩性）。设 $k\in\mathbb{N}$、$q\geqslant 2$，并令 p 为任意次数至多为 k 的多项式。则 $\mathbb{E}[|p(g)|^q]\leqslant(q-1)^{qk/2}\cdot\mathbb{E}[p(g)^2]^{q/2}$。

我们还需要以下关于高斯空间中多项式的标准反集中界：

事实 3.11（Carbery-Wright 不等式 [CW01]）。令 $g\sim\mathcal{N}(0,\mathbb{I}_n)$，并令 $p:\mathbb{R}^n\to\mathbb{R}$ 为一个 k 次多项式。则存在某个普适常数 $B>0$使得

$$\mathbb{P}\left[|p(g)|\leqslant\varepsilon\mathrm{Var}[p(g)]^{1/2}\right]\leqslant Bk\varepsilon^{1/k}\,.$$

对称埃尔米特多项式 我们还将关键性地使用以下厄米特多项式的对称多元扩展：

定义 3.2. 设k、n 为正整数。定义向量值多项式$F_k:\ \mathbb{R}^n\to\mathbb{R}^k$ 如下：

$$(F_k)_i := \frac{1}{\sqrt{n}}\sum_{j=1}^n h_i(x_j)\,.$$

根据对称多项式基本定理，任意次数为k 的对称多项式都可以表示为次数不超过k的初等对称多项式的函数。不难看出，次数不超过k 的初等对称多项式可以表示为上述向量F_k 的函数。因此，任意次数为k 的对称多项式都可以表示为F_k的某个函数。

请注意，这并非高斯空间中对称多项式的标准基底；通常会采用初等对称多项式的厄米变体。不过，这个基底对我们来说要方便得多。

克拉夫丘克多项式 变量 $x_1,x_2,\ldots,x_n$ 上的任意对称多项式，都可以唯一地表示为初等对称多项式的多项式。特别地，设 $\mathcal{D}$ 为如下分布：对于每个 $i\in[n]$，$x_i=1$ 以概率 $\gamma\in(0,1)$ 取值，且 -1 以概率 $1-\gamma$ 取值，并且各变量相互独立，则以下多项式构成相应的正交基：

$$s_k(x) = \sum_{S\subseteq[n]:|S|=k}\chi_S(x)\,,$$

其中 $\chi_S(x) = \prod_{i\in S} \frac{x_i-(2\gamma-1)}{2\sqrt{\gamma(1-\gamma)}}$。（为简化记号，我们将省略对 γ 的依赖关系。）事实上，由于 $\mathbb{E}[x_i] = 2\gamma-1$ 且 $\mathrm{Var}[x_i] = 4\gamma(1-\gamma)$，我们有 $\mathbb{E}_{x\sim\mathcal{D}}[s_k(x)] = 0$，其中 $k \geqslant 1$，并且对于所有 $k,\ell \leqslant n$，都有 $\mathbb{E}_{x\sim\mathcal{D}}[s_k(x)s_\ell(x)] = \binom{n}{k}\delta_{k\ell}$。

此外，$s_k(x)$ 可以表示为关于 k 次的单变量多项式 $\sum_{i=1}^n x_i$。这类多项式称为（平移且归一化的）克拉夫丘克多项式 [Kra29]。

定义 3.3（平移克拉夫丘克多项式）。固定 $\gamma \in (0,1)$，并令 $n \in \mathbb{N}$。定义 Kr_k 为唯一的、依赖于 n 和 γ 的 k 次单变量多项式，使得

$$\mathrm{Kr}_k(y;\, n,\gamma) = \binom{n}{k}^{-1/2} s_k(x)$$

对所有 $x \in \{\pm 1\}^n$ 和 $y = \frac{1}{\sqrt{n}} s_1(x) = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^n x_i-(2\gamma-1)}{2\sqrt{\gamma(1-\gamma)}}$。当上下文明确时，我们将省略对 n 的依赖，γ。

容易看出，$\{\mathrm{Kr}_k\}_{k\leqslant n}$ 关于 y 的分布构成一组标准正交多项式。直观地说，由于 $y \xrightarrow{d} \mathcal{N}(0,1)$ 当 $n \to \infty$ 时，根据中心极限定理，$\mathrm{Kr}_k(y)$ 收敛于归一化的（概率论学派）埃尔米特多项式 h_k，它们关于标准高斯分布构成一组标准正交多项式 [Sze75]。以下是我们将用到的一个关键事实：

引理 3.12。固定一个常数 $\gamma \in (0,1)$。对于任意 $k \leqslant n/2$ 以及 $|y| \leqslant o(n^{1/6})$，我们有 $|\mathrm{Kr}_k(y)| \leqslant O(k^{1/4}) \cdot e^{y^2/4}$。

定理 3.12 类似于厄米特多项式 [BC90] 的界 $|h_k(y)| \leqslant O(e^{y^2/4})$。对于 $\gamma = 1/2$ 的情形，也称为二元克拉夫丘克多项式，精确的界（没有额外的 $k^{1/4}$ 因子）已在 [Kra01] 中给出，最近 [Pol19, DILV24] 也给出了相关结果。然而，我们未能找到关于任意 $\gamma \in (0,1)$ 情形下定理 3.12 的参考文献。因此，我们在 B 节给出证明，该证明紧密沿用了 [Pol19, DILV24] 中的证明。

4 BinomialSpace：完全不可区分性

在本节中，我们证明定理 1.1（重述如下）。

定理（定理 1.1 的重述）。设 $\gamma, \varepsilon \in (0,1)$ 为绝对常数，且 $\delta \geqslant 0$。设 $D \in \mathbb{N}$ 满足 $D \geqslant C_\varepsilon^\gamma \log n$，其中 C^γ 是仅依赖于 γ 的常数。设 $\mathsf{N} =$ 伯努利分布$(\gamma)^n$，并设 P 为定义在 $\{\pm 1\}^n$ 上的 S_n-对称分布，使得 $\chi^2_D(\mathsf{P}\,\|\mathsf{N}) \leqslant \delta$。那么，

$$d_{\mathrm{TV}}(\mathsf{N}, \mathsf{T}_\varepsilon \mathsf{P}) \leqslant O_\gamma\left(\frac{\delta}{\varepsilon^2} + \frac{1}{\sqrt{n}\cdot\varepsilon^2}\right).$$

在本节中，我们假设 γ 为固定常数，并为使记号简洁而省略对 γ 的依赖。

给定从 N（零假设）或 P（植入）中抽取的 $\boldsymbol{x} \in \{\pm 1\}^n$，我们将坐标的中心化和归一化和记作

$$\boldsymbol{y} = \frac{1}{2\sqrt{n\gamma(1-\gamma)}} \sum_{i=1}^n (\boldsymbol{x}_i - (2\gamma-1)). \tag{6}$$

设 ν, π 分别为 $\boldsymbol{y}$ 在 $\boldsymbol{x}$ 分别从 N 和 P 中抽样时的分布。注意，$\mathbb{E}_\nu[\boldsymbol{y}] = 0$ 和 $\mathbb{E}_\nu[\boldsymbol{y}^2] = 1$，因为在零假设下，$\boldsymbol{x}_i \sim$ 伯努利分布(γ) 的均值为 $2\gamma-1$，方差为 $4\gamma(1-\gamma)$。

稍微滥用一下记号，令 $\mathsf{T}_\varepsilon \pi$ 表示当 $\boldsymbol{x} \sim \mathsf{T}_\varepsilon \mathsf{P}$ 时 $\boldsymbol{y}$ 的分布。由于 N 和 P（以及 $\mathsf{T}_\varepsilon \mathsf{P}$）都是对称分布，我们有以下引理，它指出我们可以专注于 $\boldsymbol{y}$ 的分布：

引理 4.1。对于对称分布 N 和 P，$d_{\mathrm{TV}}(\mathsf{N}, \mathsf{T}_\varepsilon \mathsf{P}) = d_{\mathrm{TV}}(\nu, \mathsf{T}_\varepsilon \pi)$。

证明。 注意，任意对称函数$f\colon \{\pm 1\}^n \to \mathbb{R}$都可以唯一地表示为坐标和的函数，即对于所有$x \in \{\pm 1\}^n$，以及某个函数$g\colon \mathbb{R} \to \mathbb{R}$，均有$f(x) = g(\sum_{i=1}^n x_i)$。对于对称分布，其概率质量函数（因而似然比）都是对称函数，这意味着它们仅取决于y。因此结论立即成立，因为 $\mathrm{T}_\varepsilon \mathrm{P}(x)/\mathrm{N}(x) = \mathrm{T}_\varepsilon \pi(y)/\nu(y)$。

□

因此，要证明 定理 1.1，等价于给出 $d_{\mathrm{TV}}(\nu, \mathrm{T}_\varepsilon \pi)$ 的上界。

我们还将用到以下事实：由 $\mathrm{T}_\varepsilon \pi$ 诱导的噪声算子相对于 Krawtchouk 基的作用符合预期。

事实 4.2。 设 π 由 $\{\pm 1\}^n$上的任意分布诱导。对于任意 $\ell \geqslant 1$，都有

$$\mathop{\mathbb{E}}_{\widetilde{\boldsymbol{y}} \sim \mathrm{T}_\varepsilon \pi} \mathrm{Kr}_\ell(\widetilde{\boldsymbol{y}}) = (1-\varepsilon)^\ell \mathop{\mathbb{E}}_{\boldsymbol{y} \sim \pi} \mathrm{Kr}_\ell(\boldsymbol{y}) .$$

证明。 设 $\boldsymbol{x} \sim \mathrm{P}$，且 $\boldsymbol{z} \sim \mathrm{N}$，二者相互独立。令 $\widetilde{\boldsymbol{x}} = \mathrm{T}_\varepsilon[\boldsymbol{x}],$，即对每个 i，独立地有 $\widetilde{\boldsymbol{x}}_i = \boldsymbol{x}_i$，其概率为 $1-\varepsilon$；并且有 $\widetilde{\boldsymbol{x}}_i = \boldsymbol{z}_i$，其概率为 ε。那么

$$\mathrm{Kr}_\ell(\widetilde{\boldsymbol{y}}) = \binom{n}{\ell}^{-1/2} \sum_{S \subseteq [n]\colon |S| = \ell} \chi_S(\widetilde{\boldsymbol{x}}) \qquad \text{and} \qquad \mathrm{Kr}_\ell(\boldsymbol{y}) = \binom{n}{\ell}^{-1/2} \sum_{S \subseteq [n]\colon |S| = \ell} \chi_S(\boldsymbol{x}) .$$

因此，由线性性可知，只需证明对于每个大小为 ℓ 的 $S \subseteq [n]$，都有 $\mathbb{E}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} \chi_S(\widetilde{\boldsymbol{x}}) = (1-\varepsilon)^\ell \mathbb{E}_{\boldsymbol{x} \sim \mathrm{P}} \chi_S(\boldsymbol{x})$。由于在给定 $\boldsymbol{x}$ 的条件下，$\widetilde{\boldsymbol{x}}$ 的各项相互独立，因此

$$\mathop{\mathbb{E}}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} \chi_S(\widetilde{\boldsymbol{x}}) = \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathrm{P}} \mathop{\mathbb{E}}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} [\chi_S(\widetilde{\boldsymbol{x}}) \mid \boldsymbol{x}] = \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathrm{P}} \mathop{\mathbb{E}}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} \left[\prod_{j \in S} \frac{\widetilde{\boldsymbol{x}}_i - (2\gamma - 1)}{2\sqrt{\gamma(1-\gamma)}} \mid \boldsymbol{x} \right] = \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathrm{P}} \prod_{j \in S} \mathop{\mathbb{E}}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} \left[\frac{\widetilde{\boldsymbol{x}}_i - (2\gamma - 1)}{2\sqrt{\gamma(1-\gamma)}} \mid \boldsymbol{x} \right]$$

现在，在$\boldsymbol{x}$的条件下，要么 $\widetilde{\boldsymbol{x}}^j = \boldsymbol{x}^j$，要么 $\widetilde{\boldsymbol{x}}^j = \boldsymbol{z}^j$。在第二种情况下，根据构造，内层期望为 0。因此，

$$\mathop{\mathbb{E}}_{\widetilde{\boldsymbol{x}} \sim \mathrm{T}_\varepsilon \mathrm{P}} \chi_S(\widetilde{\boldsymbol{x}}) = \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathrm{P}} \prod_{j \in S} (1-\varepsilon) \cdot \left(\frac{\boldsymbol{x}_i - (2\gamma - 1)}{2\sqrt{\gamma(1-\gamma)}} \right) = (1-\varepsilon)^\ell \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathrm{P}} \chi_S(\boldsymbol{x}) .$$

□

4.1 截断植入分布

本节中，我们分析截断植入分布 π的性质，即在某个阈值 τ下，π 以$|\boldsymbol{y}| \leqslant \tau$为条件的分布。我们先给出一个简单的界，它可直接由二项分布概率质量函数的下界推出。

引理 4.3。 设$\overline{\pi}$为 π 在$|\boldsymbol{y}| \leqslant \tau = 2\sqrt{\log n}$条件下的分布。那么，$\chi^2(\overline{\pi} \,\|\, \nu) \leqslant O(n^{3 \quad -} \qquad)$.

证明。 注意，$\mathrm{supp}(\pi)^{-} \subseteq \mathrm{supp}(\nu)$。因此，根据定理 3.3，只需给出 $\nu(y)$ 在 $|y| \leqslant \tau$上的下界。由于 $\tau \leqslant o(\sqrt{\gamma n})$，根据定理 3.6 可得

$$\nu(y) = \mathop{\mathbb{P}}_{\boldsymbol{x} \sim \mathrm{Ber}(\gamma)^n} \left[\sum_{i=1}^n \boldsymbol{x}_i = (2\gamma - 1)n + 2\sqrt{n\gamma(1-\gamma)} \cdot y \right] \geqslant \frac{1}{\sqrt{2n}} \exp\left(-\frac{y^2}{2} - O\left(|y|^3 / \sqrt{\gamma n} \right) \right) .$$

对于 $\tau = 2\sqrt{\log n}$，上述值至少为 $\Omega(n^{-5/2})$。于是，根据 定理 3.3，$\chi^2(\pi \,\|\, \nu) \leqslant$ ^{1} $^{\text{最小值}_{|y| \leqslant \tau} \nu(y)} \leqslant O(n^3)$.

□

由于 ν 是（经过中心化和标准化的）二项分布，我们知道其矩和尾界接近标准高斯分布的相应结果（定理 3.7）。我们证明以下一般性结果：假设具有低次不可区分性，则在N下具有类高斯矩的任何统计量，在P下也具有类高斯尾界。

引理 4.4（类高斯尾界）。设 N 和 P 满足 $\chi^2_D(P\ \|N) \leqslant \delta$，并设 p 是次数为 d 的多元多项式，使得对于所有 $k \leqslant D2d$，都有 $\mathbb{E}_{x\sim \mathsf{N}}[p(x)^{2k}] \leqslant B(2k-1)!!$，其中某个 $B \geqslant 1$。那么，对于任意 $t \leqslant \sqrt{D/d}$，我们有 $\mathbb{P}_{x\sim \mathsf{P}}[|p(x)| \geqslant t] \leqslant O(B)\cdot(\delta + 2^{-t^2/4})e^{-t^2/4}$。

证明。设 $k = \lfloor t^2/4 \rfloor \leqslant D/4d$。若 $t \leqslant 2$，结论显然成立，因此不妨设 $k \geqslant 1$。令 $\eta := \mathbb{P}_{x\sim \mathsf{P}}[|p(x)| \geqslant t]$。考虑这个次数为 $2dk$ 的多项式 p^{2k}。首先，我们有

$$\mathop{\mathbb{E}}_{x\sim \mathsf{P}}[p(x)^{2k}] \geqslant \eta \cdot \mathop{\mathbb{E}}_{x\sim \mathsf{P}}\left[p(x)^{2k} \mid |p(x)| \geqslant t\right] \geqslant \eta t^{2k}.$$

另一方面，$\chi^2_D(\mathsf{P}\ \|\mathsf{N}) \leqslant \delta$ 以及 $4dk \leqslant D$ 意味着

$$\mathop{\mathbb{E}}_{x\sim \mathsf{P}}[p(x)^{2k}] \leqslant \mathop{\mathbb{E}}_{x\sim \mathsf{N}}[p(x)^{2k}] + \delta\sqrt{\mathop{\mathrm{Var}}_{x\sim \mathsf{N}}[p(x)^{2k}]} \leqslant \mathop{\mathbb{E}}_{x\sim \mathsf{N}}[p(x)^{2k}] + \delta\sqrt{\mathop{\mathbb{E}}_{x\sim \mathsf{N}}[p(x)^{4k}]}.$$

根据假设，我们有 $\mathbb{E}_{\mathsf{N}}[p^{2k}] \leqslant B(2k-1)!!$，由斯特林近似可知，它至多为 $B\cdot\sqrt{2}(\frac{2k}{e})^k$。类似地，我们有 $\mathbb{E}_{\mathsf{N}}[p^{4k}] \leqslant B\cdot\sqrt{2}(\frac{4k}{e})^{2k}$，因为 $2k \leqslant D/2d$。因此，

$$\eta \leqslant t^{-2k}\mathop{\mathbb{E}}_{x\sim \mathsf{P}}[p(x)^{2k}] \leqslant t^{-2k}\cdot 2B\left(\left(\frac{2k}{e}\right)^k + \delta\left(\frac{4k}{e}\right)^k\right) \leqslant 2B(\delta + 2^{-k})\left(\frac{4k}{et^2}\right)^k.$$

由于 $k = \lfloor t^2/4 \rfloor \geqslant t^2/4 - 1$，上式为 $O(B)\cdot(\delta + 2^{-t^2/4})e^{-t^2/4}$。 □

接下来，我们研究截断下函数期望的行为。假设分布具有指数尾界且函数有界，我们证明以下一般界。

引理 4.5（截断下的有界函数）。设 $\tau > 0$ 和 $\alpha > 0$。设 π 为一个分布，使得 $\mathbb{P}_{y\sim\pi}[|y| \geqslant \tau] \leqslant \alpha e^{-\tau^2/4}$。那么，对于任意 $f\colon \mathbb{R} \to \mathbb{R}$，

$$\left|\mathop{\mathbb{E}}_{y\sim\pi}[f(y)\cdot\mathbf{1}(|y| \leqslant \tau)]\right| \leqslant \left|\mathop{\mathbb{E}}_{y\sim\pi}[f(y)]\right| + \sqrt{\alpha e^{-\tau^2/4}\mathop{\mathbb{E}}_{y\sim\pi}[f(y)^2]}. \tag{7}$$

此外，设 $T \geqslant \tau \geqslant 2$ 和 $\beta > 0$ 满足 $\mathbb{P}_{y\sim\pi}[|y| \geqslant t] \leqslant \alpha e^{-t^2/4}$，对所有 $t \in [\tau, T]$ 成立，且 $|f(y)| \leqslant \beta e^{y^2/4}$ 对所有 $|y| \in [\tau, T]$ 成立，则

$$\left|\mathop{\mathbb{E}}_{y\sim\pi}[f(y)\cdot\mathbf{1}(|y| \leqslant \tau)]\right| \leqslant \left|\mathop{\mathbb{E}}_{y\sim\pi}[f(y)]\right| + \frac{1}{4}\alpha\beta T^2 + \sqrt{\alpha e^{-T^2/4}\mathop{\mathbb{E}}_{y\sim\pi}[f(y)^2]}. \tag{8}$$

证明。设 $\mathcal{A}$ 为事件 $|y| > T$，并设 $\mathcal{B}$ 为事件 $\tau < y \leqslant T$。那么，

$$\mathop{\mathbb{E}}_{\pi}[f] = \mathop{\mathbb{E}}_{\pi}[f\cdot\mathbf{1}(|y| \leqslant \tau)] + \mathop{\mathbb{E}}_{\pi}[f\cdot\mathbf{1}(\mathcal{B})] + \mathop{\mathbb{E}}_{\pi}[f\cdot\mathbf{1}(\mathcal{A})]. \tag{9}$$

为证明式（7），我们令 $T = \tau$，因此第二项为 0。对于第三项，根据柯西-施瓦茨不等式，我们有 $\mathbb{E}_\pi[f\cdot\mathbf{1}(\mathcal{A})] \leqslant \sqrt{\mathbb{E}_\pi[f^2]\cdot\mathbb{P}_\pi[\mathcal{A}]} \leqslant \sqrt{\alpha e^{-T^2/4}\mathbb{E}_\pi[f^2]}$。

为证明式（8），我们考虑 $T > \tau$，并利用假设对式（9）的第二项进行界定：$|f(y)| \leqslant \beta e^{y^2/4}$，其中 $|y| \in [\tau, T]$：

$$\left|\mathop{\mathbb{E}}_{\pi}[f\cdot\mathbf{1}(\mathcal{B})]\right| \leqslant \mathop{\mathbb{E}}_{\pi}[|f|\cdot\mathbf{1}(\mathcal{B})] \leqslant \beta\cdot\mathop{\mathbb{E}}_{\pi}\left[e^{y^2/4}\cdot\mathbf{1}(\mathcal{B})\right].$$

记 $z := e^{y^2/4}\cdot\mathbf{1}(\mathcal{B})$，它是一个非负随机变量，并且 $z = 0$，除非 $\tau < y \leqslant T$。因此，对于任意 $\theta > 0$，

$$\mathop{\mathbb{P}}_{\pi}[z \geqslant \theta] \leqslant \begin{cases} \mathbb{P}_\pi[|y| \geqslant \tau], & 0 < \theta < e^{\tau^2/4}. \\ \mathbb{P}_\pi[|y| \geqslant 2\sqrt{\log\theta}], & e^{\tau^2/4} \leqslant \theta \leqslant e^{T^2/4}. \\ 0, & \theta > e^{T^2/4}. \end{cases}$$

对于前两种情况，根据假设，我们有 $\mathbb{P}[|\boldsymbol{y}| \geqslant 2\sqrt{\text{对数 }\theta}] \leqslant \alpha e^{-\text{对数}\theta} = \alpha\theta^{-1}$（因为 $2\sqrt{\text{对数 }\theta} \leqslant T$）。因此，

$$\mathop{\mathbb{E}}_{\pi}[\boldsymbol{z}] = \int_0^{\infty} \mathop{\mathbb{P}}_{\pi}[\boldsymbol{z} \geqslant \theta]\, d\theta \leqslant \alpha e^{-\tau^2/4} \cdot e^{\tau^2/4} + \int_{e^{\tau^2/4}}^{e^{T^2/4}} \alpha\theta^{-1}\, d\theta \leqslant \alpha + \frac{\alpha}{4}(T^2 - \tau^2) \leqslant \frac{1}{4}\alpha T^2 .$$

然后，$|\mathbb{E}_{\pi}[f \cdot \mathbf{1}(\mathcal{B})]| \leqslant \frac{1}{4}\alpha\beta T^2$。代入式（9）即可完成证明。 □

作为推论，我们得到如下结果。

引理 4.6。设 $\delta \in [n^{-1/2}, 1]$、$\tau = 2\sqrt{\log n}$，且 $2\tau^2 \leqslant D \leqslant o(\gamma n)^{2/3}$。假设 ν 和 π 满足 $\chi^2_D(P \,\|\, N) \leqslant \delta$。那么，存在一个常数 $C_\gamma > 1$（仅取决于 γ），使得对于所有 $1 \leqslant \ell \leqslant D/C\gamma$，均有 $|\mathbb{E}_{\pi}[Kr_\ell(y) \cdot \mathbf{1}(|y| \leqslant \tau)]| \leqslant O(\delta\ell^{5/4})$。

证明。我们首先利用定理 4.4 建立一个尾部区域界。根据定理 3.7，对于所有 $k \leqslant o(\gamma n)^{1/3}$，有 $\mathbb{E}_{\nu}[y^{2k}] = (1 + o(1))(2k-1)!!$。因此，应用定理 4.4，并令 $f(x)$ 如式（6）所示（一个次数为 1 的多项式），则对于所有 $\tau \leqslant t \leqslant \sqrt{D}$，

$$\mathop{\mathbb{P}}_{\pi}[|\boldsymbol{y}| > t] \leqslant O(\delta + 2^{-t^2/4}) \cdot e^{-t^2/4} \leqslant O(\delta) \cdot e^{-t^2/4},$$

因为我们假设 $\delta \geqslant \Omega(n^{-1/2}) \geqslant 2^{-\tau^2/4}$。特别地，$\mathbb{P}_{\pi}[|\boldsymbol{y}| > \tau] \leqslant O(\delta) \cdot n^{-1}$。

接下来，我们给出 $\mathbb{E}_{\pi}[Kr_\ell^2]$ 的界。$\chi^2_D(P \,\|\, N) \leqslant \delta$ 意味着

$$\mathop{\mathbb{E}}_{\pi}[Kr_\ell^2] \leqslant \mathop{\mathbb{E}}_{\nu}[Kr_\ell^2] + \delta\sqrt{\mathop{\mathbb{E}}_{\nu}[Kr_\ell^4]} \leqslant 1 + \delta e^{C'_\gamma \ell},$$

由超收缩性（定理 3.8）可知，存在某个常数 C'_γ。

对于 $\ell \leqslant \tau^2/8C'_\gamma$，我们可以直接使用定理 4.5 中的第一个界（其中 $\alpha = O(\delta)$）：

$$\begin{aligned}|\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y}) \cdot \mathbf{1}(|\boldsymbol{y}| \leqslant \tau)]| &\leqslant |\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y})]| + \sqrt{O(\delta)e^{-\tau^2/4}(1 + \delta e^{\tau^2/8})} \\ &\leqslant |\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y})]| + O(\sqrt{\delta}n^{-1/2}) + O(\delta n^{-1/4}).\end{aligned}$$

对于 $\tau^2/8C'_\gamma \leqslant \ell \leqslant D/8C'_\gamma$，令 $T = \sqrt{8C'_\gamma \ell} \leqslant \sqrt{D}$，并应用定理 4.5，其中 $\alpha = O(\delta)$ 且 $\beta = O(\ell^{1/4})$（因为根据定理 3.12，$|Kr_\ell(y)| \leqslant O(\ell^{1/4}) \cdot e^{y^2/4}$）：

$$\begin{aligned}|\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y}) \cdot \mathbf{1}(|\boldsymbol{y}| \leqslant \tau)]| &\leqslant |\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y})]| + O(\delta\ell^{5/4}) + \sqrt{O(\delta)e^{-2C'_\gamma \ell}(1 + \delta e^{C'_\gamma \ell})} \\ &\leqslant |\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y})]| + O(\delta\ell^{5/4}) + O(\delta n^{-1/2}).\end{aligned}$$

最后，$\chi^2_D(P \,\|N) \leqslant \delta$ 意味着 $|\mathbb{E}_{\boldsymbol{y}\sim\pi}[Kr_\ell(\boldsymbol{y})]| \leqslant \delta$，从而完成证明。 □

4.2 定理 1.1 的证明

定理 1.1 的证明。对于我们的假设 $\chi^2_D(P \,\|N) \leqslant \delta$，由于我们的目标界为 $O(\delta + 1/\sqrt{n})/\varepsilon^2$，我们可以假设 $\delta \geqslant 1/\sqrt{n}$。设 ν, π 分别为当 $\boldsymbol{x}$ 从 N（零假设）和 P（植入）中采样时 $\boldsymbol{y} = \frac{1}{2\sqrt{n\gamma(1-\gamma)}} \sum_{i=1}^{n} (x_i - (2\gamma - 1))$ 的分布。根据定理 4.1，我们可以等价地对 $d_{TV}(\nu, T_\varepsilon \pi)$ 给出上界。

令 $\tau = 2\sqrt{\log n}$。我们首先考虑 $\overline{\pi}$，即在给定 $|y| \leqslant \tau$ 的条件下 π 的分布。由于 $D \gg \tau^2$，且 ν 的矩与标准高斯分布的矩相匹配（定理 3.7），根据定理 4.4，我们有 $\mathbb{P}_{y\sim\pi}[|y| > \tau] \leqslant O(\delta + 2^{-\tau^2/4}) \cdot e^{-\tau^2/4} \leqslant O(\delta n^{-1})$。因此，根据数据处理不等式（定理 3.4），

$$d_{TV}(T_\varepsilon \pi, T_\varepsilon \overline{\pi}) \leqslant d_{TV}(\pi, \overline{\pi}) \leqslant O(\delta n^{-1}). \tag{10}$$

因此，只需对 $d_{\mathrm{TV}}(\nu,\mathrm{T}_\varepsilon\pi)$ 作界 $\chi^2(\mathrm{T}_\varepsilon\overline{\pi}\|\nu)$，并且我们将通过对 $\chi^2(\mathrm{T}_\varepsilon\pi\|\nu)$ 作界定理 3.1并应用定理 3.2（取 $n=1$）和定理 4.2，

$$\chi^2(\mathrm{T}_\varepsilon\overline{\pi}\|\nu)=\sum_{\ell\geqslant1}\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\mathrm{T}_\varepsilon\overline{\pi}}[\mathrm{Kr}_\ell(\boldsymbol{y})]^2=\sum_{\ell\geqslant1}(1-\varepsilon)^\ell\mathop{\mathbb{E}}_{\boldsymbol{y}\sim\overline{\pi}}[\mathrm{Kr}_\ell(\boldsymbol{y})]^2\,. \tag{11}$$

接下来，我们将求和拆分为 $\ell\leqslant T$ 和 $\ell>T$，其中 $T=\frac{4}{\varepsilon}\log n$。

对于 $\ell\leqslant T$，我们使用假设 $\chi^2_D(\mathsf{P}\,\|\,\mathsf{N})\leqslant\delta$。根据 定理 4.6，可得 $|\mathbb{E}_\pi[\mathrm{Kr}_\ell]|\leqslant O(\delta\ell^{5/4})$ 对所有 $\ell\leqslant T$均成立。这里，我们需要 $D\geqslant C_\gamma T$，其中 C_γ 是 定理 4.6中的常数。

对于 $\ell>T$，我们使用 $\chi^2(\pi\|\nu)\leqslant O(n^3)$ 这一界，它来自 定理 4.3 （因为 $\tau=2\log n$），以及事实$\chi^2(\pi\|\overline{\nu})=\sum_{\ell\geqslant1}\mathbb{E}_\pi[\mathrm{Kr}_\ell]^2$，根据 定理 3.2可知。因此，我们可以将 公式（11） 上界估计如下：

$$\begin{aligned}\chi^2(\mathrm{T}_\varepsilon\overline{\pi}\|\nu)&\leqslant\sum_{1\leqslant\ell\leqslant T}(1-\varepsilon)^\ell\cdot O_\gamma(\delta\ell^{5/4})^2+(1-\varepsilon)^T\sum_{\ell>T}\mathop{\mathbb{E}}_{\overline{\pi}}[\mathrm{Kr}_\ell]^2\\&\leqslant O_\gamma(\delta^2)\sum_{1\leqslant\ell\leqslant T}\ell^3e^{-\ell\varepsilon}+e^{-\varepsilon T}\cdot\chi^2(\overline{\pi}\|\nu)\\&\leqslant O_\gamma(\delta^2)\sum_{1\leqslant\ell\leqslant T}\ell^3e^{-\ell\varepsilon}+O(1/n)\,.\end{aligned}$$

函数 $z^3e^{-\varepsilon z}$ 在 $3/\varepsilon$ 处达到最大值。因此，$\sum_{\ell=1}^T\ell^3e^{-\ell\varepsilon}\leqslant O(1/\varepsilon^4)+\int_T^1 z^3e^{-\varepsilon z}dz\leqslant O(1/\varepsilon^4)$。接着，由于 $d_{\mathrm{TV}}(\nu,\mathrm{T}_\varepsilon\pi)\leqslant\frac{1}{2}\sqrt{\chi^2(\mathrm{T}_\varepsilon\pi\|\nu)}$（定理 3.1），我们得到

$$d_{\mathrm{TV}}(\nu,\mathrm{T}_\varepsilon\overline{\pi})\leqslant O_\gamma(\delta/\varepsilon^2+1/\sqrt{n})\,.$$

最后，根据 公式（10） 和三角不等式，我们有 $d_{\mathrm{TV}}(\nu,\mathrm{T}_\varepsilon\pi)\leqslant d_{\mathrm{TV}}(\nu,\mathrm{T}_\varepsilon\pi)+d_{\mathrm{TV}}(\mathrm{T}_\varepsilon\pi,\mathrm{T}_\varepsilon\pi)\leqslant O\gamma(\delta+1/\sqrt{n})/\varepsilon^2$，证明完毕。 □

5 高斯空间 I：低次数对称多项式

记号。在本节中，令 $\mathsf{N}=N(0,\mathrm{I}_n)$，且 P 是 $\mathbb{R}^n$ 上的一个分布。我们用 δ 表示 $\chi^2_D(\mathsf{P}\,\|\mathsf{N})$。

本节的主要目标是证明，当 δ 足够小时，向量 $\nu_k:=\mathrm{F}_k(\mathsf{N})$ 和 $\pi_{k,\varepsilon}:=\mathrm{F}_k(\mathrm{U}_\varepsilon\mathsf{P})$ 上的分布在全变差距离意义下足够接近。回忆一下，$\mathrm{F}_k(\cdot)=(\widetilde{\mathrm{h}}_i(\cdot))_{i=1}^k$ 是向量值多项式，其中 $\widetilde{\mathrm{h}}_i(\mathrm{x})=\frac{1}{\sqrt{n}}\sum_{j=1}^n\mathrm{h}_i(\mathrm{x}_j)$ 是第 i 个归一化埃尔米特多项式 $\mathrm{h}_i(\cdot)$。具体而言，我们将证明以下定理。

定理 5.1（定理 1.2的形式化版本）。设 $0<\varepsilon\leqslant1$，$\delta>0$，且 n、k、$D\in\mathbb{N}$。设 P，N 满足$\chi^2_D(\mathsf{P}\,\|\,\mathsf{N})\leqslant\delta$。此外，假设对于足够大的常数 C：

$$k\leqslant\frac{\varepsilon\log(n)}{C\log\log(n/\varepsilon)},\qquad D\geqslant C\cdot\frac{k^2\log(k)\cdot\log(1/\varepsilon)}{\varepsilon}\,.$$

那么，

$$d_{\mathrm{TV}}(\nu_k,\pi_{k,\varepsilon})\leqslant\exp\left(-\Omega\left(\frac{\varepsilon D}{k}\right)\right)+O(\delta)^{\Omega(1)}\cdot k^{k/2}+n^{-1/2}\,.$$

本节概述。为了证明定理 5.1，我们通过直接计算两个密度函数之差的积分，来界定 ℓ_1 距离。界定该积分的策略是将其拆成两项：一项是在“主体区域”上的积分，即适当选取参数 τ 后半径为 τ 的球内的积分；另一项是在“尾部区域”上的积分，即半径为 τ 的球外区域上的积分。我们证明，对于 ν_k 和 $\pi_{k,\varepsilon}$，尾部衰减都足够快，因此尾部积分可以忽略。为控制主体区域

上的积分，我们采用傅里叶分析论证，证明密度在 ℓ_∞-范数意义下接近（定理 5.4）；这构成了我们的主要技术工作，详见第 5.1 节。

定义与辅助引理。 我们还将使用以下定义。第一个是“正则性”条件，它确保 $\pi_{k,\varepsilon}$ 的傅里叶变换具有足够快的衰减（见 第 5.1.1 节）。

定义 5.1（正则性）。 固定向量 $y \in \mathbb{R}^n$ 对于正整数 ℓ 是 ℓ-正则的，如果对于每个单变量、次数为 $\leqslant 2\ell$ 的平方和多项式 p，都有

$$\frac{1}{2}\mathop{\mathbb{E}}_{g\sim \mathsf{N}}[p(g)] \leqslant \frac{1}{n}\sum_{i=1}^{n} p(y_i) \leqslant \frac{3}{2}\mathop{\mathbb{E}}_{g\sim \mathsf{N}}[p(g)]\,.$$

我们用 $\mathcal{R}_\ell \subseteq \mathbb{R}^n$ 表示所有 ℓ-正则点的集合。

定义 5.2（条件种植分布）。 我们将 条件种植分布 P'_ℓ 定义为 $\mathrm{P}|\mathcal{R}_\ell$。在本节中，我们令 $\ell = 4k$，并用 P' 表示 P'_{4k}。我们用 $\pi'_{k,\varepsilon}$ 表示分布 $F_k(\mathrm{U}\quad \varepsilon\mathrm{P}')$。

我们可以证明，从一个矩匹配分布中抽取的样本以高概率是(ℓ,m)-正则的。

引理 5.2。 令 D，$\ell \in \mathbb{N}$ 且 $c > 0$ 为充分小的常数，使得 $D = 4\ell$ 且 $\ell < c\log n$。设P 为一个分布，满足 $\chi^2_D(\mathrm{P}\,\|\,\mathsf{N}) \leqslant \delta$。那么$\boldsymbol{y} \sim \mathrm{P}$ 以至少 $1-(1+\delta)^{2^{O(\ell)}}n$ 的概率是 ℓ-正则的。 ——

下一个引理说明，分布 ν_k 和 $\pi_{k,\varepsilon}$ 在每个方向上的次高斯矩都达到 $\exp(\log n/k)$ 阶。

引理 5.3 (矩界)。 令 $0 < \varepsilon \leqslant 1$，$0 < \delta$。令 T，k，$D \in \mathbb{N}$ 满足 $kT \leqslant D$ 且 $T \leqslant \exp(\log(nC)/k)$，其中 C 为充分大的常数。对于满足 $\chi^2_D(\mathrm{P}\,\|\,\mathsf{N}) \leqslant \delta$ 的P 以及单位 $\xi \in \mathbb{R}^k$，我们有：

1. $\mathbb{E}_{z\sim\nu_k}\left[\langle z,\xi\rangle^T\right] \leqslant O(T)^{T/2}$,
2. $\mathbb{E}_{z\sim\pi_{k,\varepsilon}}\left[\langle z,\xi\rangle^T\right] \leqslant (1+\delta)\cdot O(T)^{T/2}$.

ℓ_∞-接近性以及整合各部分。 以下是关于正则点处密度的 ℓ_∞-接近性的主定理。

定理 5.4。 设 $\delta \geqslant 0$，$D \in \mathbb{N}$ 和 P，N 满足 $\chi^2_D(\mathrm{P}\,\|\,\mathsf{N}) \leqslant \delta$。设 C，c 分别为充分大的（相应地，充分小的）绝对常数，并设 $0 < \varepsilon \leqslant 1$，$R > 0$ 和 $k \in \mathbb{N}$ 满足：

1. $k \leqslant \frac{\varepsilon\log(n)}{C\log\log(n/\varepsilon)}$,
2. $D \geqslant C\cdot\frac{k^2\log(1/\varepsilon)}{\varepsilon}$.

那么：

$$\sup_{y\in\mathcal{R}_k}\left|\nu_k(y)-\pi'_{k,\varepsilon}(y)\right| \leqslant e^{-\Omega\left(\varepsilon\cdot\frac{D}{k}\right)} + O(\delta) + \frac{2^{O(k)}}{n}\,.$$

特别地，上述量总是不超过

$$e^{-\Omega\left(\varepsilon\cdot\frac{D}{k}\right)} + O(\delta) + n^{-1+o(1)}\,.$$

利用这些工具，我们可以证明定理 5.1。

定理 *5.1*的证明。 对于稍后选定的 $\tau \geqslant C\sqrt{k}$，定义 B_τ 为半径为 τ 的 ℓ_2-球。回顾一下， $\mathcal{R}_\ell$ 是所有 ℓ-正则点的集合，而 $\pi'_{k,\varepsilon}$ 是$F_k(\mathrm{U}_\varepsilon[\mathrm{P}|\mathcal{R}_{4k}])$的分布。我们可以写成：

$$\begin{aligned} d_{\mathrm{TV}}(\nu_k, \pi_{k,\varepsilon}) &\leqslant \int_{\mathbb{R}^k} \left|\nu_k(x) - \pi'_{k,\varepsilon}(x)\right| \cdot \mathbf{1}[\nu_k(x) \geqslant \pi'_{k,\varepsilon}(x)]\,\mathrm{d}x + d_{\mathrm{TV}}(\pi_{k,\varepsilon}, \pi'_{k,\varepsilon}) \\ &\leqslant \nu_k(\overline{B_\tau}) + \mathrm{Vol}(B_\tau) \cdot \sup_{x \in B_\tau \cap \mathcal{R}_{4k}} \left|\nu_k(x) - \pi_{k,\varepsilon}(x)\right| + d_{\mathrm{TV}}(\pi_{k,\varepsilon}, \pi'_{k,\varepsilon}). \end{aligned}$$

现在我们来界定上述总和中的每一项。

关于 $\nu_k(B_\tau)$的界。 根据 定理 5.3，取 $T = k$，并利用马尔可夫不等式，我们知道，对于任意固定的单位向量 $\xi \in \mathbb{R}^k$，

$$\mathop{\mathbb{P}}_{z \sim \nu_k}[\langle \xi, z\rangle \geqslant \tau] \leqslant \left(\frac{O(k)}{\tau^2}\right)^{k/2}.$$

注意，这至多为 $2^{-10}k$，但具体值可能更小，取决于 τ 的具体选择。因此，通过对网中 ξ 的各种取值应用并集界，$z \sim \nu_k$ 属于B_τ，除非发生概率至多为 $\left(\frac{O(k)}{\tau^2}\right)^{k/2}$ 的情况。 ——

关于 $\mathrm{Vol}(B_\tau) \cdot$上确界$_{x \in B_\tau \cap \mathcal{R}_{k,\ m}} \left|\nu_k(x) - \pi_{k,\ \varepsilon}(x)\right|$的界。 根据 定理 5.4，我们有：

$$\mathrm{Vol}(B_\tau) \cdot \sup_{x \in B_\tau \cap \mathcal{R}_{k,m}} \left|\nu_k(x) - \pi_{k,\varepsilon}(x)\right| \leqslant O(\tau^k) \cdot \left(e^{-\Omega\left(\frac{\varepsilon D}{k}\right)} + O(\delta) + n^{-1+o(1)}\right).$$

关于 $d_{\mathrm{TV}}(\pi_{k,\varepsilon}, \pi'_{k,\varepsilon})$的界。 根据定理 5.2，其界为$(1+\delta)^{2^{O(k)}} n$。 ——

综合以上结果 由于$t^k \geqslant 1$ 和 $2^{O(k)} = n^{o(1)}$，总体而言，这表明

$$d_{\mathrm{TV}}(\nu_k, \pi_{k,\varepsilon}) \leqslant O\left(\frac{O(k)}{\tau^2}\right)^{k/2} + O(\tau^k) \cdot \left(e^{-\Omega\left(\varepsilon \cdot \frac{D}{k}\right)} + O(\delta) + n^{-1+o(1)}\right).$$

Choosing $\tau = C\sqrt{k} + \min\left\{e^{\varepsilon \cdot \frac{D}{Ck^2}}, O\left(\frac{1}{\delta}\right)^{1/Ck}, n^{1/Ck}\right\}$ completes the proof. □

5.1 ℓ_∞-密度的接近性

本节将证明 定理 5.4，其重述如下。

定理 （定理 5.4的重述）。 设 $\delta \geqslant 0$， $D \in \mathbb{N}$ 和 P，N 满足 $\chi^2_D(\mathrm{P} \,\|\, \mathrm{N}) \leqslant \delta$。设 C， c 为充分大的（分别为充分小的）绝对常数，并设 $0 < \varepsilon \leqslant 1$， $R > 0$ 和 $k \in \mathbb{N}$ 满足：

1. $k \leqslant \frac{\varepsilon \log(n)}{C \log\log(n/\varepsilon)}$，
2. $D \geqslant C \cdot \frac{k^2 \log(1/\varepsilon)}{\varepsilon}$.

那么：

$$\sup_{y \in \mathcal{R}_k} \left|\nu_k(y) - \pi'_{k,\varepsilon}(y)\right| \leqslant e^{-\Omega\left(\varepsilon \cdot \frac{D}{k}\right)} + O(\delta) + \frac{2^{O(k)}}{n}.$$

特别地，上述值总是不超过

$$e^{-\Omega\left(\varepsilon \cdot \frac{D}{k}\right)} + O(\delta) + n^{-1+o(1)}.$$

我们将通过考察两个分布的傅里叶变换来证明这一点。具体而言，我们将证明这两个分布都具有足够快的傅里叶衰减，因此可以放心地忽略范数较大的频率向量所作的贡献；对于其余频率向量，我们则利用矩匹配性质证明这两个分布必定相近。我们将正式陈述所需条件，然后给出一个证明

定理 5.4。

更具体地说，我们将证明以下傅里叶衰减结果，其证明见第 5.1.1 节。

引理 5.5 (傅里叶衰减)。令 $\varepsilon \in (0,1]$，令 $k \leqslant \log(\varepsilon n) C \log\log n$，且 $R > C\sqrt{k}/\varepsilon$，其中 $C > 0$ 为充分大的常数。则，

$$\int_{\|\xi\|>R} |\widehat{\nu}_k(\xi)|\, d\xi \quad \text{and} \quad \int_{\|\xi\|>R} \left|\widehat{\pi}'_{k,\varepsilon}(\xi)\right| d\xi \leqslant \Omega(\varepsilon)^{-k/2} \cdot e^{-\Omega(\varepsilon R^2)} + e^{-\sqrt{n}}.$$

对于范数较小的频率向量，我们将证明以下结果（证明见第 5.1.2 节）。

引理 5.6 (频率匹配)。令 δ，$R \geqslant 0$ 且 k，$D \in \mathbb{N}$ 满足 $R^2 \leqslant \min\left\{c \cdot Dk, \exp(\log(cn)/k)\right\}$，其中 c 充分小。则

$$\sup_{\|\xi\|\leqslant R} \left|\widehat{\pi}'_{k,\varepsilon}(\xi) - \widehat{\nu}_k(\xi)\right| \leqslant 2^{-\Omega(R^2)} + O(\delta) + (1+\delta) \cdot \frac{2^{O(k)}}{n}.$$

推导 ℓ_∞ 界。接下来我们将推导定理 5.4。

定理 *5.4* 的证明。令实数域 R 取如下值。则对于所有 $y \in \mathcal{R}_{k,m}$，

$$\begin{aligned}
\left|\nu_k(y) - \pi'_{k,\varepsilon}(y)\right| &= \left|\int_{\mathbb{R}^k} e^{-i\langle \xi, y\rangle} \left(\widehat{\nu}_k(\xi) - \widehat{\pi}'_{k,\varepsilon}(\xi)\right) \mathrm{d}\xi\right| \\
&\leqslant \int_{\mathbb{R}^k} \left|\widehat{\nu}_k(\xi) - \widehat{\pi}'_{k,\varepsilon}(\xi)\right| \mathrm{d}\xi \\
&\leqslant \int_{\|\xi\|\leqslant R} \left|\widehat{\nu}_k(\xi) - \widehat{\pi}'_{k,\varepsilon}(\xi)\right| \mathrm{d}\xi + \int_{\|\xi\|>R} |\widehat{\nu}_k(\xi)|\, \mathrm{d}\xi + \int_{\|\xi\|>R} \left|\widehat{\pi}'_{k,\varepsilon}(\xi)\right| \mathrm{d}\xi.
\end{aligned}$$

参数选择。接下来我们选择实数域 R，使得定理 5.5 和5.6的前置条件均得到满足。令 $R^2 = \min\left\{c \cdot Dk, \exp(\log(cn)/k)\right\}$，其中 c 为充分小的常数。我们首先验证，对于某个足够大的常数 C'，有 $R^2 \geqslant C' k\varepsilon$。注意，根据我们对 D 的条件，有 $Dk \geqslant C\frac{k}{\varepsilon}$。还需检查

$$C \cdot \frac{k}{\varepsilon} \leqslant \exp(\log(cn)/k).$$

对两边取对数可由下式推出

$$Ck \log(k/\varepsilon) \leqslant \log n,$$

当 $k \leqslant \varepsilon \log(n)/C' \log\log(n/\varepsilon)$ 时成立。最后，根据定理 5.5的要求，还可得到 $k \leqslant \log(n/\varepsilon)/C \log\log n$。

最终界。根据 定理 5.5，最后两项至多为

$$\begin{aligned}
\Omega(\varepsilon)^{-k/2} \cdot e^{-\Omega(\varepsilon R^2)} + e^{-\sqrt{n}} &= \Omega(\varepsilon)^{-k/2} \cdot \left(e^{-\Omega\left(\varepsilon \cdot \frac{D}{k}\right)} + e^{-\varepsilon \cdot \Omega(n)^{1/k}}\right) + e^{-\sqrt{n}} \\
&= \Omega(\varepsilon)^{-k/2} \cdot \left(e^{-\Omega\left(\varepsilon \cdot \frac{D}{k}\right)} + e^{-\varepsilon \cdot \Omega(n)^{1/k}}\right).
\end{aligned}$$

我们声称，在我们的参数选择下，这可以简化为

$$e^{-\Omega\left(\varepsilon\cdot\frac{D}{k}\right)}+e^{-\varepsilon\cdot\Omega(n)^{1/k}}.$$

为此，只需证明 $\Omega(n)^{1/k}\geqslant C_\varepsilon' k\log(1/\varepsilon)$ 以及 $\varepsilon\cdot\frac{D}{k}\geqslant C'k\log(1/\varepsilon)$。第二个不等式由我们对 D 的下界立即得到。对于第一个，注意到 $k\leqslant \varepsilon\log n/C\log\log(n/\varepsilon)$ so $C_\varepsilon' k\log(1/\varepsilon)\leqslant\log n\log(1/\varepsilon)$。取对数可得

$$\frac{\log(cn)}{k}\geqslant\frac{\log n}{2k}\geqslant\frac{C\cdot\log\log(n/\varepsilon)}{2\varepsilon}\geqslant\frac{C\cdot(\log\log n+\log\log(1/\varepsilon))}{2}\geqslant C'\log k+C'\log\log(1/\varepsilon).$$

此外，根据定理 5.6，第一项至多为

$$2^{-\Omega(R^2)}+O(\delta)+(1+\delta)\cdot\frac{2^{O(k)}}{n}.$$

注意，$2^{-\Omega(R^2)}\leqslant e^{-\Omega(\varepsilon R^2)}$，且 $\delta\cdot 2^k n\leqslant O(\delta)$（由我们对 k 的选择可知）。此外，由于 $\log(cn)/k\geqslant C\cdot\log\log n\varepsilon$，因此有 $\Omega(n)^{1/k}\geqslant\log(C/\varepsilon)(n)\geqslant\log^C(n)/\varepsilon$。所以 $e^{-\varepsilon\cdot\Omega(n)^{1/k}}\leqslant e^{-\log^c(n)}\ll 2^{O(k)}/n$。因此，我们得到总体界

$$\left|\nu_k(y)-\pi'_{k,\varepsilon}(y)\right|\leqslant e^{-\Omega\left(\varepsilon\cdot\frac{D}{k}\right)}+O(\delta)+\frac{2^{O(k)}}{n}.$$ □

对于简化后的形式，我们注意到 $2^{O(k)}=n^{o(1)}$。

5.1.1 傅里叶衰减

本节将证明定理 5.5，重述如下。

引理 (重述定理 5.5)。 设 $\varepsilon\in(0,1]$, 设 $k\leqslant\log(\varepsilon n)C\log\log n$，且 $R\geqslant C\sqrt{k/\varepsilon}$，其中 $C>0$为充分大的常数。那么，

$$\int_{\|\xi\|>R}|\widehat{\nu_k}(\xi)|\,d\xi\quad and\quad\int_{\|\xi\|>R}\left|\widehat{\pi'_{k,\varepsilon}}(\xi)\right|\,d\xi\leqslant\Omega(\varepsilon)^{-k/2}\cdot e^{-\Omega(\varepsilon R^2)}+e^{-\sqrt{n}}.$$

证明该引理所需的关键估计，是关于低次多项式的傅里叶特征期望的一组界。证明见附录 C。

引理 5.7。 设 p 是一个次数为 $\leqslant k$ 的单变量多项式。那么：

- 若 $\operatorname{方差}_{g\sim\mathcal{N}(0,1)}[p(g)]\leqslant 9^{-k}$，则有

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right|\leqslant 1-\frac{1}{4}\operatorname{Var}_{g\sim\mathcal{N}(0,1)}[p(g)].$$

- 若 $\operatorname{方差}_{g\sim\mathcal{N}(0,1)}[p(g)]=kck$ 对任意 $c\in\mathbb{R}$均成立，则有

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right|\leqslant 1-\frac{1}{16k}k^{-|c|k}.$$

- 最后，若 $\operatorname{方差}_{g\sim\mathcal{N}(0,1)}[p(g)]\geqslant kCk$ 对足够大的 $C>0$成立，则有

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right|\leqslant O\left(\operatorname{Var}_{g\sim\mathcal{N}(0,1)}[p(g)]^{-\frac{1}{4k}}\right).$$

在 定理 5.5 的证明中，我们还需要用到几个众所周知的事实。

事实 5.8。 设 x 是 k 维的。对于任意的 $r > 0$、$c > 0$ 和实数域 $> 4\sqrt{k/c}$,

$$\mathrm{Vol}(B_r) = \int_{\|x\| \leqslant r} dx \leqslant O(r^k), \quad \int_{\|x\| \geqslant R} e^{-c\|x\|^2}\, dx \leqslant (2\pi/c)^{k/2} \cdot e^{-cR^2/2}.$$

利用这些估计，我们现在可以完成定理 5.5的证明：

定理 5.5 的证明。对于任意 $\xi = (\xi_1, \ldots, \xi_k) \in \mathbb{R}^k$，令 $p^\xi(x) : \mathbb{R} \to \mathbb{R}$ 为单变量多项式 $p^\xi(x) = \sum_{\ell=1}^k \xi_\ell h_\ell(x)$。注意，根据厄米特多项式的正交性，$\mathrm{Var}\, p^\xi(g) = \|\xi\|^2$。回忆向量 F_k（在输入 $x \in \mathbb{R}^n$ 时）的坐标为 $\frac{1}{\sqrt{n}} \sum_{j=1}^n h_\ell(x_j)$，其中 $\ell = 1, \ldots, k$。因此，

$$\langle F_k(x), \xi \rangle = \sum_{\ell=1}^k (F_k(x))_\ell \cdot \xi_\ell = \frac{1}{\sqrt{n}} \sum_{\ell=1}^k \left(\sum_{j=1}^n h_\ell(x_j) \right) \xi_\ell = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\sum_{\ell=1}^k h_\ell(x_j) \xi_\ell \right) = \frac{1}{\sqrt{n}} \sum_{j=1}^n p_\xi(x_j).$$

傅里叶系数 $\widehat{\nu}_k(\xi)$ 和 $\widehat{\pi' k,\varepsilon}(\xi) \in \mathbb{C}$ 分别是在 N 和 $U_\varepsilon P'$ 下 $\langle F_k, \xi \rangle$ 的特征函数（其中 P' 是在条件为 4k-正则时的植入分布；参见定义 5.1）：$\widehat{\nu}_k(\xi) = \mathbb{E}_{x \sim N}[e^{i\langle F_k(x), \xi \rangle}]$ 以及 $\widehat{\pi' k,\varepsilon}(\xi) = \mathbb{E}_{x \sim U_\varepsilon P'}[e^{i\langle F_k(x), \xi \rangle}]$。我们将重点给出 $\widehat{\pi' k,\varepsilon}(\xi)$ 的界，因为只要设定 $\varepsilon = 1$，$\widehat{\nu}_k(\xi)$ 也有相同的界。

注意到，可以先采样 $y \sim P'$，再令 $x \sim U_\varepsilon P'$ $x = \sqrt{1-\varepsilon} y + \sqrt{\varepsilon} g$，其中 $g \sim \mathcal{N}(0, \mathbb{I}_n)$，从而得到 $x \sim U_\varepsilon P'$。我们的策略是固定 $y \in \mathbb{R}^n$（假设它是正则的），并利用 g 的随机性来给特征函数（对于较大的频率 $\|\xi\| > R$）建立上界。

对于固定的 $y \in \mathbb{R}^n$，令 $z = \sqrt{1-\varepsilon} y + \sqrt{\varepsilon} g$（隐式地依赖于 y），并且

$$\begin{aligned} \mathop{\mathbb{E}}_{g \sim \mathcal{N}(0, \mathbb{I}_n)} [e^{i\langle F_k(z), \xi \rangle}] &= \mathop{\mathbb{E}}_{g \sim \mathcal{N}(0, \mathbb{I}_n)} \left[\exp\left(\frac{i}{\sqrt{n}} \sum_{j=1}^n p_\xi(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon} g_j) \right) \right] \\ &= \prod_{j=1}^n \mathop{\mathbb{E}}_{g_j \sim \mathcal{N}(0,1)} \left[\exp\left(\frac{i}{\sqrt{n}} \cdot p_\xi(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon} g_j) \right) \right]. \end{aligned} \tag{12}$$

接下来，令 $\mu(y_j) := \mathbb{E}_{g \sim \mathcal{N}(0,1)}[p^\xi(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon} g)]$，且 $v(y_j) :=$ 方差$_{g \sim \mathcal{N}(0,1)}[p^\xi(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon} g)] = \mathbb{E}_g[p^\xi(z_j)^2] - \mu(y_j)^2$（为方便起见，此处省略对 ξ, ε 的依赖）。注意，根据噪声算子的定义，$\mu(y_j) = \sum_{\ell=1}^k (1-\varepsilon)^{\ell/2} \xi_\ell h_k(y_j)$。此外，注意到 $v(y_j)$ 是次数至多为 2k 的单变量平方和多项式，因为方差可以写成平方的期望。因此，根据 y 的 4k-正则性（定义 5.1），我们有

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n v(y_j) &\geqslant \frac{1}{2} \mathop{\mathbb{E}}_{h \sim \mathcal{N}(0,1)} [v(h)] = \frac{1}{2} \mathop{\mathbb{E}}_h \left[\mathop{\mathbb{E}}_g \left[p_\xi(\sqrt{1-\varepsilon} h + \sqrt{\varepsilon} g)^2 \right] - \mu(h)^2 \right] \\ &= \frac{1}{2} \left(\mathop{\mathbb{E}}_g [p_\xi(g)^2] - \mathop{\mathbb{E}}_g [\mu(g)^2] \right) = \frac{1}{2} \left(\|\xi\|^2 - \sum_{\ell=1}^k (1-\varepsilon)^\ell \xi_\ell^2 \right) \\ &\geqslant \frac{\varepsilon}{2} \|\xi\|^2. \end{aligned} \tag{13}$$

另一方面，根据 y 的 $4k$ 正则性，我们再次得到

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n v(y_j)^2 &\leqslant \frac{3}{2} \mathop{\mathbb{E}}_{h \sim \mathcal{N}(0,1)} [v(h)^2] \leqslant \frac{3}{2} \mathop{\mathbb{E}}_h \left[\mathop{\mathbb{E}}_g \left[p_\xi(\sqrt{1-\varepsilon} h + \sqrt{\varepsilon} g)^2 \right]^2 \right] \\ &\leqslant \frac{3}{2} \mathop{\mathbb{E}}_g \left[p_\xi(g)^4 \right] \leqslant \frac{3}{2} \cdot 9^k \mathop{\mathbb{E}}_g [p_\xi(g)^2]^2 = \frac{3}{2} \cdot 9^k \|\xi\|^4, \end{aligned} \tag{14}$$

其中我们使用了詹森不等式以及高斯空间中多项式的 2 到 4 超收缩性（定理 3.10）。上述上界可以解释为 $\{v(y_j)\}_{j \in [n]}$ 的一种“分散性”条件，这一点将在后文中发挥重要作用。

为了给方程 (12)确定界，我们将使用定理 5.7 ，该定理根据方差给出了上界。在我们的情形中，对每个$j \in [n],$，我们有方差$\left[\frac{1}{\sqrt{n}} p_{\xi}(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon g_j}\right] = v(y^j)/n$。因此，

$$\prod_{j=1}^{n} \left| \mathop{\mathbb{E}}_{g_j \sim \mathcal{N}(0,1)} \left[\exp\left(\frac{i}{\sqrt{n}} \cdot p_{\xi}(\sqrt{1-\varepsilon} y_j + \sqrt{\varepsilon} g_j) \right) \right] \right| \leqslant \exp\left(- \sum_{j=1}^{n} \rho(v(y_j)/n) \right),$$

其中 ρ 定义为

$$\rho(a) := \Omega(1) \cdot \begin{cases} a & 0 \leqslant a \leqslant 2^{-Ck}, \\ k^{-Ck} & 2^{-Ck} \leqslant a \leqslant k^{Ck}, \\ \frac{1}{k} \log a & a \geqslant k^{Ck}, \end{cases}$$

对于足够大的普适常数 $C > 1$。注意，在 $a \geqslant 2^{-Ck}$ 的情形下， $\rho(a) \geqslant \Omega(k^{-Ck})$。

对于固定的 y，令 $B_1 := \{j \in [n] : v(y^j)/n \leqslant 2^{-Ck}\}$ 和 $B_2 := \{j \in [n] : v(y^j)/n > 2^{-Ck}\}$。根据公式（13），必有以下两种情况之一：

(1) $\sum_{j \in B_1} v(y^j)/n \geqslant \frac{\varepsilon}{4}\|\xi\|^2$,
或 (2) $\sum_{j \in B_2} v(y^j)/n \geqslant \frac{\varepsilon}{4}\|\xi\|^2$。

在第一种情况下，

$$\exp\left(- \sum_{j=1}^{n} \rho(v(y_j)/n) \right) \leqslant \exp\left(- \sum_{j \in B_1} \rho(v(y_j)/n) \right) \leqslant \exp\left(-\Omega(1) \sum_{j \in B_1} v(y_j)/n \right) \leqslant \exp\left(-\Omega(\varepsilon)\|\xi\|^2 \right).$$

在第二种情况下，我们使用 公式（14）对 $|B_2|$ 求下界：

$$\frac{\varepsilon}{4}\|\xi\|^2 \leqslant \sum_{j \in B_2} \frac{v(y_j)}{n} \leqslant \sqrt{|B_2| \cdot \sum_{j \in B_2} \frac{v(y_j)^2}{n^2}} \leqslant \sqrt{|B_2| \cdot \frac{1}{n} \cdot \frac{3}{2} 9^k \|\xi\|^4}.$$

这意味着 $|B_2| \geqslant \Omega(\varepsilon^2 9^{-k} n)$。根据 B_2 和 ρ的定义，对于每个$j \in B_2$，有 $\rho(v(y^j)/n) \geqslant \Omega(k^{-Ck})$ 。因此，

$$\exp\left(- \sum_{j=1}^{n} \rho(v(y_j)/n) \right) \leqslant \exp\left(-|B_2| \cdot \Omega(k^{-Ck}) \right) \leqslant \exp\left(-\Omega(\varepsilon^2) k^{-Ck} n \right).$$

T因此，结合这两种情况可知，任何满足 $4k$-正则条件的$y \in \mathbb{R}^n$ 都满足

$$\exp\left(- \sum_{j=1}^{n} \rho(v(y_j)/n) \right) \leqslant e^{-\Omega(\varepsilon\|\xi\|^2)} + e^{-\Omega(\varepsilon^2 k^{-Ck} n)}.$$

我们将在$R^2 \leqslant \|\xi\|^2 \leqslant T := n \cdot k^{Ck}/\varepsilon$这一范围内使用上述界。

$$\begin{aligned} \int_{R \leqslant \|\xi\| \leqslant \sqrt{T}} |\widehat{\pi}'_{k,\varepsilon}(\xi)| \, d\xi &\leqslant \int_{\|\xi\| \geqslant R} e^{-\Omega(\varepsilon\|\xi\|^2)} \, d\xi + \int_{\|\xi\| \leqslant \sqrt{T}} e^{-\Omega(\varepsilon^2 k^{-Ck} n)} \, d\xi \\ &\leqslant \Omega(\varepsilon)^{-k/2} \cdot e^{-\Omega(\varepsilon R^2)} + O(T^{k/2}) \cdot e^{-\Omega(\varepsilon^2 k^{-Ck} n)}. \end{aligned} \tag{15}$$

这里，第二个不等式用到了定理 5.8 和实数域 $\gg \sqrt{k}/\varepsilon$。此外，对于$k \leqslant \log(\varepsilon n)^{C'} \log\log n$，若常数$C' > C$足够大，则有 $\varepsilon^2 k^{-Ck} n \gg \sqrt{n}$。因此，我们可以将方程（15）中的第二项简化为$e^{-\sqrt{n}}$。

现在，我们关注 $\|\xi\|^2 > T$ 这一情形。由方程（13）和（14），可得 $\mathbb{E}_{j\sim[n]}[v(y_j)] \geqslant \frac{\varepsilon}{2}\|\xi\|^2$ 和 $\mathbb{E}_{j\sim[n]}[v(y_j)^2] \leqslant \frac{3}{2}9^k\|\xi\|^4$。因此，根据 Paley-Zygmund不等式，可得 $\mathbb{P}_{j\sim[n]}[v(y_j) \geqslant \frac{\varepsilon}{4}\|\xi\|^2] \geqslant \Omega(\varepsilon^2 9^{-k})$。换言之，坐标 j 中有 $\Omega(\varepsilon^2 9^{-k} n)$ 个满足 $\in [n]$ have $v(y_j)/n \geqslant \frac{\varepsilon}{4}\|\xi\|^2$，该值大于 k^{Ck}，这意味着 $\rho(v(y_j)/n) \geqslant \Omega(\frac{1}{k}) \cdot \log(\frac{\varepsilon\|\xi\|^2}{4} n)$。因此，

$$\exp\left(-\sum_{j=1}^{n}\rho(v(y_j)/n)\right) \leqslant \left(\frac{\varepsilon\|\xi\|^2}{4n}\right)^{-\Omega(\varepsilon^2 9^{-k} n/k)}.$$

在区域 $\|\xi\| > \sqrt{T}$ 上对其积分，得到

$$\int_{\|\xi\|>\sqrt{T}} |\widehat{\pi}'_{k,\varepsilon}(\xi)|\, d\xi \leqslant \int_{\sqrt{T}}^{\infty}\left(\frac{\varepsilon r^2}{4n}\right)^{-\Omega(\varepsilon^2 9^{-k} n/k)} O(r^{k-1})\, dr \leqslant \left(\frac{4n}{\varepsilon T}\right)^{\Omega(\varepsilon^2 9^{-k} n/k)} \cdot O(T)^{k/2}.$$

对于我们选取的 $T = nk^{Ck}/\varepsilon$ 且 $k \ll \log n$，上式为 $o(e^{-\sqrt{n}})$。结合公式（15），可得所需的界 $\int_{\|\xi\|>R} |\widehat{\pi}'_{k,\varepsilon}(\xi)|\, d\xi \leqslant \Omega(\varepsilon)^{-k/2} \cdot e^{-\Omega(\varepsilon R^2)} + e^{-\sqrt{n}}$。

如前所述，$\widehat{\nu}_k(\xi)$ 的界可直接通过令 $\varepsilon = 1$ 得到。证明至此完成。 □

5.1.2 频率匹配

在本节中，我们将证明定理 5.6，其重述如下。

引理（重述定理 5.6）。令 δ，实数域 $\geqslant 0$ 且 k，$D \in \mathbb{N}$ 满足 $R^2 \leqslant \min\left\{c \cdot Dk, \exp(\log(cn)/k)\right\}$，其中 c 足够小。那么

$$\sup_{\|\xi\|\leqslant R}\left|\widehat{\pi}'_{k,\varepsilon}(\xi) - \widehat{\nu}_k(\xi)\right| \leqslant 2^{-\Omega(R^2)} + O(\delta) + (1+\delta)\cdot\frac{2^{O(k)}}{n}.$$

我们将通过对傅里叶变换中的指数函数作泰勒展开来证明这一点。利用矩匹配，我们可以论证展开的前几项近似匹配，并且误差项很小，因为两者 ν_k 和 $\pi_{k,\varepsilon}$ 都具有次高斯矩。

定理 *5.6* 的证明定理 *5.6*。令 $T = CR^2$ 其中常数 C 足够大，并为简单起见，假设它是整数。固定一个 ξ 其范数至多为 R。我们需要给出以下量的界

$$\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\exp(i\langle\xi,z\rangle)] - \mathop{\mathbb{E}}_{z\sim\pi'_{k,\varepsilon}}[\exp(i\langle\xi,z\rangle)]\right|.$$

我们首先注意到 $\exp(i\langle\xi,\cdot\rangle)$ 是一个幅值至多为 1 的函数。因此，由于 $d_{\mathrm{TV}}(\pi_k,\ \varepsilon,\ \pi'_k,\ \varepsilon) \leqslant (1+\delta)^{2^{O(k)}} n$根据定理 5.2，只需证明

$$\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\exp(i\langle\xi,z\rangle)] - \mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[\exp(i\langle\xi,z\rangle)]\right| \leqslant 2^{-\Omega(R^2)} + O(\delta).$$

由于 $e^{iy} = \text{余弦}(y) + i \cdot \text{正弦}(y)$ 因此只需求出界

$$\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\cos(\langle\xi,z\rangle)] - \mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[\cos(\langle\xi,z\rangle)]\right| + \left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\sin(\langle\xi,z\rangle)] - \mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[\sin(\langle\xi,z\rangle)]\right|.$$

我们重点讨论第一项，第二项完全类似。考虑一个固定向量 z。由于 $\cos(\cdot)$ 的所有阶导数的绝对值都不超过 1，由泰勒展开可知，存在一个绝对值不超过 1 的 r_z，使得

$$\cos(\langle\xi,z\rangle) = \underbrace{\sum_{\ell=0}^{T-1}\frac{(-1)^\ell}{(2\ell)!}\langle\xi,z\rangle^{2\ell}}_{:=p_T(\langle\xi,z\rangle)} + r_z\cdot\frac{\langle\xi,z\rangle^{2T}}{(2T)!}.$$

因此，根据小低次似然比和三角不等式，我们得到

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\cos(\langle\xi,z\rangle)]-\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[\cos(\langle\xi,z\rangle)]\right| \\
&\leqslant\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[p_T(\langle\xi,z\rangle)]-\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[p_T(\langle\xi,z\rangle)]\right|+\left|\mathop{\mathbb{E}}_{z\sim\nu_k}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right|+\left|\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right| \\
&\leqslant\delta\cdot\sqrt{\mathop{\mathrm{Var}}_{\nu_k}[p_T(\langle\xi,z\rangle)]}+\left|\mathop{\mathbb{E}}_{z\sim\nu_k}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right|+\left|\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right|.
\end{aligned}$$

注意，当 c 足够小时，$T\cdot k=CR^2\leqslant D$ 且 $T=CR^2\leqslant\exp(\log(c\cdot n)/k)$。因此，$\langle\xi,z\rangle$ 在 π_k、ε 和 ν_k 下均具有阶为 T 的次高斯矩（参见定理 5.3），我们可以将最后两项界定为

$$\left|\mathop{\mathbb{E}}_{z\sim\nu_k}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right|+\left|\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}\left[r_z\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right]\right|\leqslant(1+\delta)\cdot\frac{O\left(\|\xi\|^2T\right)^T}{(2T)!}=(1+\delta)\cdot O\left(\frac{R^2}{T}\right)^T\leqslant(1+\delta)\cdot2^{-\Omega(R^2)}.$$

剩下的任务是界定方差项。再次利用 $(a+b)^2\leqslant2a^2+2b^2$ 和次高斯矩，我们得到

$$\begin{aligned}
\mathop{\mathrm{Var}}_{z\sim\nu_k}[p_T(\langle\xi,z\rangle)]&\leqslant\mathop{\mathbb{E}}_{z\sim\nu_k}\left[p_T(\langle\xi,z\rangle)^2\right]\leqslant2\mathop{\mathbb{E}}_{z\sim\nu_k}\left[(\cos(\langle\xi,z\rangle)-p_T(\langle\xi,z\rangle))^2\right]+2\mathop{\mathbb{E}}_{z\sim\nu_k}\left[\cos(\langle\xi,z\rangle)^2\right] \\
&\leqslant2\mathop{\mathbb{E}}_{z\sim\nu_k}\left[\left(\frac{\langle\xi,z\rangle^{2T}}{(2T)!}\right)^2\right]+2=O\left(\frac{R^2}{T}\right)^{2T}+2=O(1).
\end{aligned}$$

综上，我们得到

$$\left|\mathop{\mathbb{E}}_{z\sim\nu_k}[\cos(\langle\xi,z\rangle)]-\mathop{\mathbb{E}}_{z\sim\pi_{k,\varepsilon}}[\cos(\langle\xi,z\rangle)]\right|\leqslant\delta\cdot O(1)+(1+\delta)\cdot2^{-\Omega(R^2)}=O(\delta)+2^{-\Omega(R^2)}.$$

□

5.2 小引理的证明

本节给出定理 5.2 （高概率正则点）和定理 5.3 （高斯矩界）的证明。

高概率正则性。我们首先证明定理 5.2，其重述如下。

引理 (重述 定理 5.2)。 令 D，$\ell\in\mathbb{N}$ 和 $c>0$ 为充分小的常数，使得 $D=4\ell$和 $\ell<c\log n$。令 P 为一个分布，使得 $\chi^2_D(\mathsf{P}\,\|\,\mathsf{N})\leqslant\delta$。则$\boldsymbol{y}\sim\mathsf{P}$ 以至少$1-(1+\delta)^{2^{O(\ell)}}n$ 的概率是 ℓ-正则的。

证明。 我们通过一个 ε-网论证来证明所需结论。令 $\mathcal{P}_\ell$ 为 $\mathbb{R}$ 上的次数为 $\leqslant\ell$ 的多项式集合，并在 $\mathcal{P}_\ell$上定义以下两个范数：

$$\|p\|^2=\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[p(g)^2]\ ,\ \|p\|_y^2=\frac{1}{n}\sum_{i=1}^n p(y_i)^2.$$

我们需要证明

$$\mathop{\mathbb{P}}_{y\sim\mathsf{P}}\left[\forall p\neq0\in\mathcal{P}_k\colon\frac{\|p\|_y^2}{\|p\|^2}\in[0.5,1.5]\right]\geqslant1-\frac{(1+\delta)\cdot2^{O(\ell)}}{n}.$$

平方和多项式的更一般情形可以直接推出。

化归为网。 令 $\mathcal{C}$ 为单位球在 $\|\cdot\|$ 下由单位范数元素构成的 η-网。注意，由于 $\|\cdot\|$ 是一个 ℓ维欧几里得度量，$\frac{1}{\eta^{O(\ell)}}$ 个元素上存在这样的网。我们先证明：

对于固定的 y，假设对所有 $p \in \mathcal{C}$, $\frac{\|p\|_y}{\|p\|} \in [1-\eta, 1+\eta]$，其中 $\eta \leqslant 0.1$，则 y 是 ℓ-正则的。

注意，任何满足 $\|p\| = 1$ 的 p 都可以写成 $p = p_0 + \eta p'$，其中 $p_0 \in \mathcal{C}$，且 $\|p'\| \leqslant 1$。对 p' 反复应用这一点，可得

$$p = \sum_{i=0}^{\infty} \eta^i c_i p_i ,$$

其中所有 $p_i \in \mathcal{C}$，且所有 $|c_i| \leqslant 1$。由三角不等式可得：

$$\|p_0\|_y - \sum_{i \geqslant 1} \eta^i \|p_i\|_y \leqslant \|p\|_y \leqslant \|p_0\|_y + \sum_{i \geqslant 1} \eta^i \|p_i\|_y$$

利用关于 $p_i \in \mathcal{C}$ 和 $\|p_i\| = 1$ 的假设，我们得到：

$$\|p\|_y \leqslant (1+\eta)\left(\|p_0\| + \sum_{i \geqslant 1} \eta^i \|p_i\|\right) = \frac{1+\eta}{1-\eta} .$$
$$\|p\|_y \geqslant (1-\eta)\|p_0\| - (1+\eta)\sum_{i \geqslant 1} \eta^i \|p_i\| \geqslant \frac{1-3\eta}{1-\eta}$$

对网应用并集界。我们还需证明，对于所有 $p \in \mathcal{C}$，都有 $\frac{\|p\|_y}{\|p\|} \in [0.9, 1.1]$，且其成立概率至少为 $1 - (1+\delta)^{2^{O(\ell)}} n$。我们通过二阶矩方法和并集界来证明。具体而言，由于 $\mathcal{C}$ 的大小至多为 $2^{O(\ell)}$，且对于某个充分小的常数 c，$\ell = c\log n$，只需证明对于固定的 $p \in \mathcal{C}$

$$\mathbb{P}\left[\frac{\|p\|_{\boldsymbol{y}}}{\|p\|} \notin [0.9, 1.1]\right] \leqslant \frac{(1+\delta)2^{O(\ell)}}{n} .$$

注意，只需证明 $\left|\|p\|^2{}_{\boldsymbol{y}} - \|p\|^2\right| \geqslant (0.1)^2 \|p\|^2$ 以不超过相同的概率成立，即可推出这一结论。我们将证明

$$\mathop{\mathbb{E}}_{\boldsymbol{y} \sim \mathsf{P}}\left[\left(\|p\|_{\boldsymbol{y}}^2 - \|p\|^2\right)^2\right] \leqslant \frac{(1+\delta)\cdot 2^{O(\ell)}}{n} \cdot \|p\|^4 ,$$

由此可通过马尔可夫不等式得到所述结论。注意：

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{y} \sim \mathsf{P}}\left[\left(\|p\|_{\boldsymbol{y}}^2 - \|p\|^2\right)^2\right] &\leqslant \mathop{\mathbb{E}}_{g \sim \mathsf{N}}\left[\left(\|p\|_g^2 - \|p\|^2\right)^2\right] + \delta\sqrt{\mathop{\mathbb{E}}_{g \sim \mathsf{N}}\left[\left(\|p\|_g^2 - \|p\|^2\right)^4\right]} \\
&\leqslant \left(1 + \delta \cdot 2^{O(\ell)}\right) \cdot \mathop{\mathbb{E}}_{g \sim \mathsf{N}}\left[\left(\|p\|_g^2 - \|p\|^2\right)^2\right] \\
&= \left(1 + \delta \cdot 2^{O(\ell)}\right) \cdot \frac{1}{n} \mathop{\mathrm{Var}}_{g \sim \mathcal{N}(0,1)}\left[p(g)^2\right] \\
&\leqslant \frac{(1+\delta)\cdot 2^{O(\ell)}}{n} \cdot \mathop{\mathbb{E}}_{g \sim \mathcal{N}(0,1)}\left[p(g)^2\right]^2 = \frac{(1+\delta)\cdot 2^{O(\ell)}}{n} \cdot \|p\|^4 ,
\end{aligned}$$

其中第一行使用了关于 $\chi^2_D(\mathsf{P} \,\|\mathsf{N})$ 的界，第二行对高斯空间上的多项式使用了 2 到 4 超收缩性（定理 3.10），第三行展开了方差，第四行再次使用了超收缩性。

□

高斯矩界。 接下来我们给出 定理 5.3的证明，定理重述如下 .

引理 （重述 定理 5.3）。 设 $0 < \varepsilon \leqslant 1$，$0 < \delta$。设 T、k、$D \in \mathbb{N}$，使得 $kT \leqslant D$ 且 $T \leqslant \exp(\log(nC)/k)$，其申 C 为充分大的常数。对于满足 $\chi^2_D(\mathrm{P} \parallel \mathrm{N}) \leqslant \delta$ 且为单位 $\xi \in \mathbb{R}^k$ 的 P，我们有：

1. $\mathbb{E}_{z\sim\nu_k}\left[\langle z,\xi\rangle^T\right] \leqslant O(T)^{T/2}$,
2. $\mathbb{E}_{z\sim\pi_{k,\varepsilon}}\left[\langle z,\xi\rangle^T\right] \leqslant (1+\delta)\cdot O(T)^{T/2}$.

证明。 我们首先从第一个陈述推出第二个陈述（我们将适当扩展记号，证明其对 $2T$成立）。回忆一下， ν_k、π_k、ε 分别是当 $x_k(x)$ 满足 $x \sim \mathrm{N}$ 和 $x \sim \mathrm{U}_\varepsilon \mathrm{P}$ 时，$F_k(x)$ 的分布。注意，$\langle \xi, F_k(x)\rangle^T$ 是关于 x 的次数至多为 $kT \leqslant D$的多项式。由于 $\chi^2_D(\mathrm{P} \parallel \mathrm{N}) \leqslant \delta$ （因此 $\chi^2_D(\mathrm{U}_\varepsilon \mathrm{P} \parallel \mathrm{N}) \leqslant \delta$），可得

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim \mathrm{U}_\varepsilon \mathrm{P}}\left[\langle \xi, F_k(y)\rangle^T\right] &\leqslant \mathop{\mathbb{E}}_{x\sim \mathrm{N}}\left[\langle \xi, F_k(x)\rangle^T\right] + \delta\sqrt{\mathop{\mathrm{Var}}_{x\sim \mathrm{N}}\left[\langle \xi, F_k(x)\rangle^T\right]} \\
&\leqslant \mathop{\mathbb{E}}_{x\sim \mathrm{N}}\left[\langle \xi, F_k(x)\rangle^T\right] + \delta\sqrt{\mathop{\mathbb{E}}_{x\sim \mathrm{N}}\langle \xi, F_k(x)\rangle^{2T}} \\
&\leqslant (1+\delta)\cdot O(T)^{T/2}.
\end{aligned}$$

高斯情形：$\mathrm{N} = \mathcal{N}(0, \mathbb{I}_n)$。接下来我们证明第一个陈述（注意，这里我们只会用到条件 $(\mathrm{CT})^k \leqslant \mathrm{n}$，因此将 T 乘以 2 不影响结果）。固定一个单位向量 $\xi \in \mathbb{R}^k$，并定义 p^ξ：$\mathbb{R} \to \mathbb{R}$，使得 $\mathrm{p}^\xi(\mathrm{x}) = \sum_{i=1}^k \xi_i \mathrm{h}_i(\mathrm{x})$。注意，$\mathrm{p}^\xi$是一个次数至多为 k 的多项式，并满足 $\mathbb{E}_{g\sim\mathcal{N}(0,1)}[\mathrm{p}^\xi(\mathrm{g})] = 0$和 $\mathbb{E}_{g\sim\mathcal{N}(0,1)}[\mathrm{p}^\xi(\mathrm{g})^2] = 1$。此外，还有

$$\langle \xi, F_k(x)\rangle = \sum_{i=1}^k \xi_i \tilde{h}_i(x) = \tfrac{1}{\sqrt{n}}\sum_{j=1}^n\sum_{i=1}^k \xi_i h_i(x_j) = \tfrac{1}{\sqrt{n}}\sum_{j=1}^n p_\xi(x_j).$$

为简便起见，我们将省略对 ξ 的依赖。p^ξ。我们需要给出.....的界

$$n^{-T/2}\mathop{\mathbb{E}}_{g}\left[\left(\sum_{j=1}^n p(g_j)\right)^T\right].$$

我们通过展开期望中的幂来处理这个期望。首先引入一些记号。记 α 为 $\mathbb{N}^n$ 中的一个多重指标，记 $|\alpha| = \sum_{j=1}^n \alpha_j$ 为.....，记 $\|\alpha\|_0$ 为 α 中非零项的个数。于是

$$\mathop{\mathbb{E}}_{g}\left[\left(\sum_{j=1}^n p(g_j)\right)^T\right] = \sum_{|\alpha|=T}\prod_{j=1}^n \mathop{\mathbb{E}}_{g_j}\left[p(g_j)^{\alpha_j}\right].$$

注意，对于所有满足 α 且其中一个 $\alpha_j = 1$ 的情形，该乘积都为 0。此外，当 $\alpha_j = 2$ 时，对应项恰好为 1。定义集合

$$A_{m,T} = \left\{\alpha \in \mathbb{N}^n \mid \forall j : \alpha_j \neq 1, |\alpha| = T, \|\alpha\|_0 = m\right\}.$$

此外，根据高斯超收缩性（参见定理 3 10 ℓ $\mathbb{E}_g[$ $^\ell] \leqslant$ 。），我们知道对于任意整数，$p(g)$。因此，$\ell^{k\ell/2}$ 对于任意 $\alpha_j > 2$ 应用超收缩性，可得

$$\mathbb{E}\left[\left(\sum_{j=1}^n p(g_j)\right)^T\right] \leqslant \sum_{m=1}^{T/2}\sum_{\alpha\in A_{m,T}}\prod_{j=1}^n O(\alpha_j)^{k\alpha_j/2} = \sum_{m=1}^{T/2}\sum_{\alpha\in A_{m,T}}\prod_{j:\,\alpha_j>2} O(\alpha_j)^{k\alpha_j/2}.$$

我们声称该乘积至多为$(O(T)^k)^{T/2-m}$。证明将在本节末尾给出；我们先来看这如何导出最终的界。注意，$|A_{m,T}| \leqslant \binom{n}{m} \cdot m^T \leqslant (n/m)^m m^T$。因此，利用$m \leqslant T$

$$\begin{aligned} n^{-T/2}\,\mathbb{E}\left[\left(\sum_{j=1}^{n} p(\boldsymbol{g}_j)\right)^T\right] &\leqslant \sum_{m=1}^{T/2}\left(\frac{n}{m}\right)^m m^T n^{-T/2}\left(O(T)^k\right)^{T/2-m} \\ &\leqslant T^{T/2}\sum_{m=1}^{T/2}\left(\frac{m}{n}\right)^{T/2-m}\left(O(T)^k\right)^{T/2-m} \\ &= T^{T/2}\sum_{m=1}^{T/2}\left(\frac{m\cdot\left(O(T)^k\right)}{n}\right)^{T/2-m}. \end{aligned}$$

由于 $m \leqslant T/2$ 且 $(CT)^k \leqslant n$（其中常数 C 足够大），因此求和中的每一项都至多为 1，从而整个表达式至多为$O(T)^{T/2}$。

为了说明乘积至多为 $(O(T)^k)^{T/2-m}$，我们将证明，当 α 的非零系数中除一个外其余都等于 2，而剩下的一个等于 $T-2m$ 时，乘积取得最大值。注意，此时这就是乘积的值。我们不妨假设 α_j 取实数值，且都至少为 2，并且它们的和仍为 T（忽略所有为 0 的项）。注意，这是一个多面体，其顶点对应于除一个坐标外其余坐标都为 2，而剩下的一个为 $T-2m$ 的情形。对乘积取对数后，我们观察到函数 $_{k_2}\sum_j \alpha_j \log(\alpha_j)$ 是凸函数，因此在该多面体的某个顶点处取得最大值。这便得证。

□

6 高斯空间 II：连通子图计数

在本节中　　我们将证明矩阵值输入的子图计数全变差距离接近结果　　.

本节的记号。本节中我们将始终使用以下记号。n 是问题的维数，零假设分布 N 是由一个随机对称的$n \times n$ 矩阵给出的，其高斯条目相互独立；P 是一个与 N δ-低次数不可区分的分布，即 $\chi^2_D(\mathrm{P}\,\|\mathrm{N}) \leqslant \delta$。我们用 ε 表示添加到植入分布中的噪声量。不失一般性，我们可以取 $\varepsilon < 1/2$。特别地，$\mathrm{U}_\varepsilon\,\mathrm{P}$ 表示 $\sqrt{1-\varepsilon}\boldsymbol{Y} + \varepsilon \boldsymbol{G}$ 的分布，其中$\boldsymbol{Y} \sim \mathrm{P}$，$\boldsymbol{G} \sim \mathrm{N}$，且二者相互独立。　　——

我们用 ϑ 表示一个顶点数为常数的连通图。我们用 $\mathrm{v}(\vartheta)$表示 ϑ中的顶点数，用 $\mathrm{e}(\vartheta)$表示 ϑ中的边数。$\mathrm{E}(\vartheta)$表示 ϑ中的所有边组成的集合。令 $\mathcal{L}_\vartheta$ 表示从 ϑ 的顶点到 $[\mathrm{n}]$ 的所有不同单射标记的集合 π；如果存在一个 ϑ 的自同构将 π 变换为 π'，我们就认为两个标记 π、π'相同。对于边 $\mathrm{e} = (\mathrm{u}, \mathrm{v})$，我们记$\pi(\mathrm{e}) = \{\pi(\mathrm{u}), \pi(\mathrm{v})\}$。对于矩阵 M，定义

$$\chi_\vartheta(M) \coloneqq \frac{1}{\sqrt{|\mathcal{L}_\vartheta|}} \sum_{\pi\in\mathcal{L}_\vartheta} \prod_{ab\in E(\vartheta)} M_{\pi(a),\pi(b)}\,.$$

在本节中，我们将证明以下结论。

定理 6.1。设 $\delta \geqslant 0$，$D \geqslant \log n \cdot \log\log n$ 和 P，N 满足 $\chi^2_D(\mathrm{P}\,\|\,\mathrm{N}) \leqslant \delta$。设 π_ϑ 为当 $\boldsymbol{M} \sim \mathrm{U}_\varepsilon\mathrm{P}$ 时 $\chi_\vartheta(\boldsymbol{M})$ 的分布，并设 ν_ϑ 为当 $\boldsymbol{M} \sim \mathrm{N}$ 时 $\chi_\vartheta(\boldsymbol{M})$ 的分布。那么，对于充分小的常数 $\alpha > 0$，我们有：

$$d_{\mathrm{TV}}(\nu_\vartheta, \pi_\vartheta) \leqslant O\left(\delta^{\alpha\varepsilon^2} + n^{-\alpha\varepsilon^2} + \frac{\operatorname{polylog} n}{\sqrt{n}}\right),$$

其中 $O(\cdot)$ 隐藏了依赖于 ϑ 的常数。

为了证明 定理 6.1，我们使用 Chatterjee [Cha09 的定理 2.2]，它允许我们假设，向 $\boldsymbol{M}$ 添加噪声“类似于”向 $\chi(\boldsymbol{M})$ 添加噪声，也就是说，在应用子图多项式之后。

引理 6.2。 设 $F\colon \mathbb{R}^d \to \mathbb{R}$ 是一个二阶可微函数。对于以下统计量：

$$\kappa_1 := \left(\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,I_d)} \|\nabla F(g)\|_2^4 \right)^{1/4}, \quad \kappa_2 := \left(\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,I_d)} \left\|\nabla^2 F(g)\right\|_{\mathrm{op}}^4 \right)^{1/4},$$
$$\mu := \mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,I_d)} F(g), \quad \sigma^2 := \mathop{\mathrm{Var}}_{g\sim\mathcal{N}(0,I_d)} F(g),$$

我们有

$$d_{\mathrm{TV}}\Big(F(g), \mathcal{N}(\mu,\sigma^2)\Big) \leqslant \frac{2\sqrt{5}\kappa_1\kappa_2}{\sigma^2}.$$

证明定理 6.1的整体策略。 设 $\boldsymbol{G} \sim \mathsf{N}$。对于 $\boldsymbol{M} \sim \mathsf{P}$，设 $\boldsymbol{M}_\varepsilon := \sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}$。我们的整体证明策略与第 5 节相同。主要区别在于第一步，其目标是转而考虑一对分布，使得证明其高阶傅里叶系数的界更加容易。

- 我们首先利用定理 6.2 证明，$\chi_\vartheta(\mathsf{M}_\varepsilon)$ 的分布与 $\sqrt{1-\varepsilon}\mathsf{e}(\vartheta)\chi_\vartheta(\mathsf{M}) + \sigma\mathsf{M}\cdot\mathsf{g}$ 的分布之间的全变差距离很小，其中 $\mathsf{g} \sim \mathcal{N}(0,1)$ 与 M 相互独立，并且 $\sigma\mathsf{M}$ 处处具有“非平凡”的大小。更具体地，将第二个分布记为 π_ϑ，我们将证明 $\mathrm{dTV}(\pi_\vartheta, \pi_\vartheta) \lesssim \mathrm{polylog}\, n/\sqrt{n}$。类似地，我们将定义 ν_ϑ，并证明它与 ν_ϑ之间的全变差距离具有类似的界。因此，在后续论证中，只需界定 π 与 ν之间的全变差距离。这对应于下文的定理 6.3。

- 为了界定 $d\mathrm{TV}(\pi_\vartheta,\nu_\vartheta)$，我们对高概率区域 $[-\tau,\tau]$ 内的 $|\pi_\vartheta(x) - \nu_\vartheta(x)|$ 进行界定：特别地，如果我们能证明该区域内的界为 η，就能得到：

$$d_{\mathrm{TV}}(\pi_\vartheta,\nu_\vartheta) \leqslant \mathop{\mathbb{P}}_{x\sim\nu}[|x| \geqslant \tau] + 2\eta\tau.$$

  接下来我们证明 存在 η 与 τ 之间的权衡关系，使得这个界足够小。

- 令 x 取自这个高概率区域。为了控制 $|\pi_\vartheta(x) - \nu_\vartheta(x)|$，我们采用傅里叶分析方法。具体而言，我们使用以下恒等式：

$$|\pi_\vartheta(x) - \nu_\vartheta(x)| = \frac{1}{\sqrt{2\pi}}\left|\int_{-\infty}^{\infty} (\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi))\cdot\exp(-i\xi x)\right| \mathrm{d}\xi.$$

  我们将此积分分成两部分：对于较大的 ξ（量级为 $\xi \geqslant T = C\cdot\log n$），我们将证明 $\widehat{\pi}_\vartheta(\xi)$ 和 $\widehat{\nu}_\vartheta(\xi)$ 分别至多为 $1/\mathrm{poly}(n)$。

- 对于 $\xi \in [-T, T]$，我们将证明两个分布的傅里叶系数彼此接近：利用 $\widehat{\pi}_\vartheta(\xi) = \widehat{\pi}_\vartheta(\xi) \pm \mathrm{dTV}(\pi_\vartheta, \pi_\vartheta)$ 以及 $\widehat{\nu}_\vartheta(\xi) = \widehat{\nu}_\vartheta(\xi) \pm \mathrm{dTV}(\nu_\vartheta,\nu_\vartheta)$ 这一事实（如上所述，全变差距离很小），只需控制 $|\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)|$。这可通过矩匹配得出（$\chi^2\mathrm{D}(\mathsf{P}\,\|\mathsf{N})$ 很小）。精确表述见下文定理 6.5。

在本小节余下部分，我们将陈述 定理 6.3和 6.5，以及一个引理，该引理表明，对于 $\chi_\vartheta(\boldsymbol{M})$，当 $\boldsymbol{M} \sim \mathsf{N}$ 时，其矩是次高斯的（参见 定理 6.4）。随后，我们将说明如何由此推出定理 6.1。我们将在第 6.3 节 中第 6.1 节 证明定理 6.3（这构成了证明的主体部分），定理 6.5将在第 6.2 节中证明，而定理 6.4将在附录 D中证明。

转向含噪子图计数和矩界。我们先给出以下定义。

定义 6.1（含噪子图计数）。对于固定的矩阵 $M \in \mathbb{R}^{n\times n}$ 和 $\boldsymbol{G}$ 一个对称的 $n \times n$ 矩阵，其元素（除对称位置外）独立同分布于 $\mathcal{N}(0,1)$，定义

$$\widetilde{\sigma}^2(M) := \max\left\{ \operatorname*{Var}_{G\sim \mathsf{N}} \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right), \frac{\varepsilon}{2}(1-\varepsilon)^{e(\vartheta)-1} \right\}.$$

令 $\boldsymbol{M} \sim \mathsf{P}$、$g \sim \mathcal{N}(0,1)$ 相互独立。我们将 π_ϑ $\sqrt{1-\varepsilon}^{e(\vartheta)}\chi_\vartheta(\boldsymbol{M}) + \widetilde{\sigma}(\boldsymbol{M})\cdot g$ 的分布记为 $\underline{\pi}_\vartheta$。类似地，当 $\boldsymbol{M} \sim \mathsf{N}$ 时，我们将所得的分布记为 $\underline{\nu}_\vartheta$。

下面的引理分别给出了 π_ϑ 与 $\underline{\pi}_\vartheta$，以及 ν_ϑ 与 $\underline{\nu}_\vartheta$ 之间的全变差距离的界。

引理 6.3。假设 $D \geqslant \log n \cdot \operatorname{loglog} n$，并令 P 满足 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) \leqslant n$，则有：

$$d_{\mathrm{TV}}(\pi_\vartheta, \underline{\pi}_\vartheta) \leqslant O\left(\frac{\sqrt{\varepsilon}\cdot \operatorname{polylog} n}{\sqrt{n}} + \frac{1}{n^2}\right) \quad \textit{and} \quad d_{\mathrm{TV}}(\nu_\vartheta, \underline{\nu}_\vartheta) \leqslant O\left(\frac{\sqrt{\varepsilon}\cdot \operatorname{polylog} n}{\sqrt{n}} + \frac{1}{n^2}\right).$$

注意，第二部分只需令 $\mathsf{P} = \mathsf{N}$，即可由第一部分立即推出。

我们还将使用下面的引理，该引理表明，当 $\boldsymbol{M} \sim \mathsf{N}$ 时，$\chi_\vartheta(\boldsymbol{M})$ 的多项式数量级的矩是次高斯的。

引理 6.4。存在一个充分小的常数 $\alpha = \alpha(\vartheta)$，使得对于任意 $0 \leqslant q \leqslant n^\alpha$：

$$\mathop{\mathbb{E}}_{\boldsymbol{M}\sim \mathsf{N}} \chi_\vartheta(\boldsymbol{M})^q \leqslant \sqrt{q}^q \cdot \left(1 + \frac{1}{\operatorname{poly}(n)}\right).$$

ℓ_∞-接近度并将各部分结合起来。我们首先给出 π_ϑ 与 ν_ϑ 之间无穷距离的如下界。[6]

引理 6.5。设 $D \geqslant \log n \operatorname{loglog} n$，且 $\delta \geqslant 0$ 满足 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) \leqslant \delta$，并假设 $\varepsilon \geqslant 1/$ 多对数 n。那么存在绝对常数 $c > 0$ 使得

$$\|\underline{\pi}_\vartheta - \underline{\nu}_\vartheta\|_\infty \leqslant O(\delta^{c\varepsilon^2} + n^{-c\varepsilon^2}).$$

现在我们终于可以证明定理 6.1。

定理 *6.1* 的证明 定理 *6.1*。首先注意，当 $\varepsilon = o(1/\sqrt{\log n})$ 或 $\varepsilon = o(1/\sqrt{\log(1/\delta)})$ 时，右侧至少为 1，因此该界显然成立。故可假设上述情况均不成立。

根据三角不等式和 定理 6.3 可得：

$$d_{\mathrm{TV}}(\pi_\vartheta, \nu_\vartheta) \leqslant d_{\mathrm{TV}}(\underline{\pi}_\vartheta, \underline{\nu}_\vartheta) + O\left(\frac{\sqrt{\varepsilon}\operatorname{polylog} n}{\sqrt{n}} + \frac{1}{n^2}\right) \leqslant d_{\mathrm{TV}}(\underline{\pi}_\vartheta, \underline{\nu}_\vartheta) + O\left(\frac{\operatorname{polylog} n}{n}\right).$$

因此，只需证明 $d_{\mathrm{TV}}(\pi_\vartheta, \nu_\vartheta) \leqslant O(\delta^{\alpha\varepsilon^2} + n^{-\alpha\varepsilon^2})$，其中 α 是一个较小的绝对常数。令 $\beta = \delta^{c\varepsilon^2} + n^{-c\varepsilon^2}$，且 $\tau = \sqrt{\log(1/\beta)}$，其中 c 是定理界中的常数定理 6.5。注意，由于 $\varepsilon \geqslant \Omega(1/\sqrt{\log n})$ 且 $\varepsilon \geqslant \Omega(1/\sqrt{\log(1/\delta)})$，$\beta$ 至多为（比如说）0.001。于是，

$$\begin{aligned} d_{\mathrm{TV}}(\underline{\pi}_\vartheta, \underline{\nu}_\vartheta) &= \int_{-\infty}^{\infty} |\underline{\pi}_\vartheta(x) - \underline{\nu}_\vartheta(x)| \cdot \mathbf{1}[\underline{\nu}_\vartheta(x) \geqslant \underline{\pi}_\vartheta(x)] \,\mathrm{d}x \\ &= \int_{-\tau}^{\tau} |\underline{\pi}_\vartheta(x) - \underline{\nu}_\vartheta(x)| \,\mathrm{d}x + \int_{x\notin[-\tau,\tau]} (\underline{\nu}_\vartheta(x) - \underline{\pi}_\vartheta(x)) \,\mathrm{d}x \\ &\leqslant 2\tau \cdot \|\underline{\pi}_\vartheta - \underline{\nu}_\vartheta\|_\infty + \mathop{\mathbb{P}}_{\boldsymbol{x}\sim\underline{\nu}_\vartheta}[|\boldsymbol{x}| \geqslant \tau]. \end{aligned}$$

[6]注意，对 ε 的限制并不严格，因为当 $\varepsilon = o(1/\sqrt{\log n})$ 时，定理 6.1 是空泛成立的。

注意，由于 $\varepsilon \leqslant 1$， $\tau \leqslant O(\sqrt{\log n})$，因此根据 定理 6.4 （取 $q = 0.01\tau^2$）以及马尔可夫不等式，可得

$$\mathop{\mathbb{P}}_{\boldsymbol{x}\sim\underline{\nu}_\vartheta}[|\boldsymbol{x}| \geqslant \tau] \leqslant \left(\frac{2q}{\tau^2}\right)^{q/2} \leqslant e^{-\Omega(\tau^2)} = O(\beta)\,.$$

结合这一点，定理 6.5 可得

$$d_{\mathrm{TV}}(\underline{\pi}_\vartheta, \underline{\nu}_\vartheta) \leqslant O\left(\beta\sqrt{\log(1/\beta)}\right) \leqslant O\left(\sqrt{\beta}\right) = O(\delta^{\alpha\varepsilon^2/2} + n^{-\alpha\varepsilon^2/2})\,. \quad \square$$

6.1 含噪输入的子图计数与含噪子图计数接近

本节证明如下重述的定理 6.3 。回顾一下， π_ϑ （分别为 ν_ϑ）对应于 $\sqrt{1-\varepsilon^{e(\vartheta)}}\chi_\vartheta(\underline{\boldsymbol{M}}) + \widetilde{\sigma}(\boldsymbol{M})\cdot\boldsymbol{g}$ 的分布，其中 $\boldsymbol{g} \sim \mathcal{N}(0,1)$，而 $\boldsymbol{M} \sim \mathsf{P}$（$\boldsymbol{M} \sim \mathsf{N}$），且对于固定矩阵 $M \in \mathbb{R}^{n\times n}$

$$\widetilde{\sigma}^2(M) := \max\left\{\mathop{\mathrm{Var}}_{G\sim\mathsf{N}} \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right), \frac{\varepsilon}{2}(1-\varepsilon)^{e(\vartheta)-1}\right\},$$

其中 $\boldsymbol{G}$ 是一个对称的$n \times n$ 矩阵，其元素独立同分布为 $\mathcal{N}(0,1)$（除对称性所要求者外）。

引理 （重述定理 6.3）。 假设 $D \geqslant \log n \cdot \operatorname{loglog} n$，且令 P 满足 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) \leqslant n$，则有：

$$d_{\mathrm{TV}}(\pi_\vartheta, \underline{\pi}_\vartheta) \leqslant O\left(\frac{\sqrt{\varepsilon}\cdot\operatorname{polylog} n}{\sqrt{n}} + \frac{1}{n^2}\right) \quad \textit{and} \quad d_{\mathrm{TV}}(\nu_\vartheta, \underline{\nu}_\vartheta) \leqslant O\left(\frac{\sqrt{\varepsilon}\cdot\operatorname{polylog} n}{\sqrt{n}} + \frac{1}{n^2}\right).$$

证明。 对于一个 $n \times n$ 矩阵 M，我们将应用 定理 6.2 ，其中 $F(\boldsymbol{G}) := \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right)$。[7]沿用定理 6.2中的记号，定义：

$$\begin{aligned}
\kappa_1(M) &:= \left(\mathop{\mathbb{E}}_{\boldsymbol{G}\sim\mathsf{N}} \left\|\nabla_{\boldsymbol{G}}\, \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right)\right\|_2^4\right)^{1/4}, \\
\kappa_2(M) &:= \left(\mathop{\mathbb{E}}_{\boldsymbol{G}\sim\mathsf{N}} \left\|\nabla^2_{\boldsymbol{G}}\, \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right)\right\|_{\mathrm{op}}^4\right)^{1/4}, \\
\sigma^2(M) &:= \mathop{\mathrm{Var}}_{\boldsymbol{G}\sim\mathsf{N}} \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right).
\end{aligned}$$

只需证明，对于 $\kappa_1(\boldsymbol{M})$ 和 $\kappa_2(\boldsymbol{M})$，其高概率上界成立；并证明对于 $\sigma^2(\boldsymbol{M})$ ，其高概率下界成立，其中 $\boldsymbol{M} \sim \mathsf{P}$。对于随机变量 $\boldsymbol{X}$，我们用 $\ell_q(\boldsymbol{X})$ 表示 $(\mathbb{E}\boldsymbol{X}^q)^{1/q}$。[8] 我们将反复用到这样一个事实： $\ell_q(\cdot)$ 是一个范数。

关于 $\kappa_1(\boldsymbol{M})$ 的高概率上界。 对于 $\boldsymbol{M} \sim \mathsf{P}$，我们将证明 $\ell_q(\kappa_1(\boldsymbol{M}))$ 的一个界，其中 $q \ll D$取适当的值；随后，根据马尔可夫不等式即可得到一个足够好的高概率界。

$$\ell_q(\kappa_1(\boldsymbol{M})) = \ell_q\left(\ell_4\left(\left\|\nabla_{\boldsymbol{G}}\, \chi_\vartheta\left(\sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}\right)\right\|_2 \middle| \boldsymbol{M}\right)\right) \tag{16}$$

让我们理清上述项。为了使计算过程中需要处理的项保持简洁，我们再定义一些记号。对于 $e \in$ 期望(ϑ) 以及 $i,j \in [n],$ ，我们将 $\mathcal{L}_{\vartheta,e,\{i,j\}}$ 定义为所有满足 $\pi(e) = \{i,j\}$的 $\pi \in \mathcal{L}_\vartheta$所组成的集合。定义

$$\chi^{i,j}_{\vartheta,e}(M) := \frac{1}{\sqrt{|\mathcal{L}_\vartheta|}} \sum_{\pi\in\mathcal{L}_{\vartheta,e,\{i,j\}}} \prod_{ab\in E(\vartheta)\setminus e} M_{\pi(a),\pi(b)}\,,$$

[7]为使论证完全严谨，考虑到矩阵具有对称性，需要使用一个仅包含上三角部分的向量。这样做很容易，我们为简洁起见省略细节。

[8]采用这种记号的原因是为了避免使用类似 $\|\|\boldsymbol{X}\|_2\|_q$ 这样的记号而造成混淆：其中外层范数是随机变量范数，内层范数是向量范数。

并将 $u_{\vartheta,e}(M)$ 定义为一个$\binom{n}{2}$维向量，其中 $u_{\vartheta,e}(M)[\{i,j\}] = \chi^{i,j}_{\vartheta,e}(M)$。现在，注意到：

$$\nabla_G \chi_\vartheta\left(\sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}\right) = \sum_{e\in E(\vartheta)} \sqrt{\varepsilon}\cdot u_{\vartheta,e}\left(\sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}\right)$$

将上述内容代入方程 (16)，得到：

$$\begin{aligned}\ell_q(\kappa_1(\boldsymbol{M})) &\leqslant \sqrt{\varepsilon}\sum_{e\in E(\vartheta)} \ell_q\left(\ell_4\left(\left\|u_{\vartheta,e}\left(\sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}\right)\right|\boldsymbol{M}\right\|_2\right)\right)\\ &\leqslant \sqrt{\varepsilon}\sum_{e\in E(\vartheta)} \ell_{4q}\left(\left\|u_{\vartheta,e}\left(\sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}\right)\right\|_2\right)\end{aligned} \tag{17}$$

其中，第一个不等式由范数的三角不等式得出，第二个不等式则由詹森不等式得出。

现在我们分析单个求和项。定义 $\boldsymbol{M}_\varepsilon := \sqrt{1-\varepsilon}\boldsymbol{M} + \sqrt{\varepsilon}\boldsymbol{G}$，则有：

$$\ell_{4q}\left(\|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2\right)^{4q} = \mathop{\mathbb{E}}_{\boldsymbol{M}_\varepsilon\sim \mathsf{U}_\varepsilon\,\mathsf{P}} \|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2^{4q}\,.$$

由于这是关于 $\boldsymbol{M}_\varepsilon$ 的次数为 $C_\vartheta q \ll D$ 的多项式，其中 C_ϑ 仅取决于 ϑ，利用 $\operatorname{Var} X \leqslant \mathbb{E}X^2$ 和 $\mathbb{E}X \leqslant \sqrt{\mathbb{E}X^2}$，可得

$$\begin{aligned}\ell_{4q}\left(\|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2\right)^{4q} = \mathop{\mathbb{E}}_{\boldsymbol{M}_\varepsilon\sim \mathsf{U}_\varepsilon\,\mathsf{P}} \|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2^{4q} &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{G}\sim \mathsf{N}} \|u_{\vartheta,e}(\boldsymbol{G})\|_2^{4q} + \sqrt{\chi^2_D(\mathsf{P}\,\|\,\mathsf{N})\cdot \mathop{\operatorname{Var}}_{\boldsymbol{G}\sim\mathsf{N}} \|u_{\vartheta,e}(\boldsymbol{G})\|_2^{4q}}\\ &\leqslant (1+\sqrt{n})\cdot\sqrt{\mathop{\mathbb{E}}_{\boldsymbol{G}\sim\mathsf{N}} \|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2^{8q}}\,.\end{aligned} \tag{18}$$

总之，我们已经证明：

$$\ell_{4q}\left(\|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2\right) \leqslant O\left(n^{1/8q}\right)\cdot\left(\mathop{\mathbb{E}}_{\boldsymbol{G}\sim\mathsf{N}} \|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2^{8q}\right)^{1/8q} \tag{19}$$

现在注意，$\|u_{\vartheta,\,e}(\boldsymbol{M}_\varepsilon)\|_2^{8q} = \operatorname{Tr}\left(\left(u_{\vartheta,\,e}(\boldsymbol{G})u_{\vartheta,\,e}(\boldsymbol{G})^\top\right)^{4q}\right)$。根据 [AMP16、引理 8.7 和 8.15] （“图矩阵矩界”）：[9]

$$\begin{aligned}\mathop{\mathbb{E}}_{\boldsymbol{G}\sim\mathsf{N}} \operatorname{Tr}\left(\left(u_{\vartheta,e}(\boldsymbol{G})u_{\vartheta,e}(\boldsymbol{G})^\top\right)^{4q}\right) &\leqslant \frac{1}{|\mathcal{L}_\vartheta|^{4q}}(8q)^{8q(e(\vartheta)+v(\vartheta))}\cdot n^{4qv(\vartheta)}\\ &\leqslant \frac{v(\vartheta)^{4qv(\vartheta)}}{n^{4qv(\vartheta)}}(8q)^{8q(e(\vartheta)+v(\vartheta))}\cdot n^{4qv(\vartheta)}\\ &\leqslant (8v(\vartheta)q)^{8q(e(\vartheta)+v(\vartheta))}\,.\end{aligned}$$

将上述结果代入公式（19）可得：

$$\ell_{4q}\left(\|u_{\vartheta,e}(\boldsymbol{M}_\varepsilon)\|_2\right) \leqslant O\left(n^{1/8q}\right)\cdot(8v(\vartheta)q)^{e(\vartheta)+v(\vartheta)}\,.$$

我们明确选取$q = \log n$，并将上述结果代入公式（17）得到界：

$$\ell_{\log n}(\kappa_1(\boldsymbol{M})) \leqslant O(\sqrt{\varepsilon})\cdot e(\vartheta)\cdot(8v(\vartheta)\log n)^{e(\vartheta)+v(\vartheta)}\,. \tag{20}$$

[9] 在我们的设定中，矩阵的大小为 $\binom{n}{2}\times 1$。用图矩阵的术语来说，只有一组特殊顶点，对应于向量各项的索引。

关于 $\kappa_2(\boldsymbol{M})$ 的高概率上界。 与前一种情况类似，我们将给出 $\ell_q(\kappa_2(\boldsymbol{M}))$ 在某个较大的q下的界，而马尔可夫不等式将给出一个足够好的高概率界。按照与前一种情况类似的计算步骤，直到 公式（17），可得：

$$\ell_q(\kappa_2(\boldsymbol{M})) \leqslant \varepsilon \cdot \sum_{e \neq e' \in E(\vartheta)} \ell_{4q}\left(\left\|u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)\right\|_{\mathrm{op}}\right). \tag{21}$$

上述中，$u_{\vartheta,e,e'}(\boldsymbol{M})$ 定义为通过计算 $\nabla^2\chi_\vartheta(\boldsymbol{M}_\varepsilon)$得到的如下$\binom{n}{2} \times \binom{n}{2}$ 矩阵

$$u_{\vartheta,e,e'}(M)[\{i,j\},\{i',j'\}] = \frac{1}{\sqrt{|\mathcal{L}_\vartheta|}} \sum_{\pi \in \mathcal{L}_{\vartheta,e,\{i,j\}} \cap \mathcal{L}_{\vartheta,e',\{i',j'\}}} \prod_{ab \in E(\vartheta)\setminus\{e,e'\}} M_{\pi(a),\pi(b)}\,.$$

现在我们研究公式（21）中的一个求和项。对于固定的$e \neq e'$，有：

$$\begin{aligned}
\ell_{4q}\left(\left\|u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)\right\|_{\mathrm{op}}\right)^{4q} &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{M}_\varepsilon \sim \mathrm{U}_\varepsilon \mathsf{P}} \left\|u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)\right\|_{\mathrm{op}}^{4q} \\
&\leqslant \mathop{\mathbb{E}}_{\boldsymbol{M}_\varepsilon \sim \mathrm{U}_\varepsilon \mathsf{P}} \mathrm{Tr}\left(\left(u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon) u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)^\top\right)^{2q}\right) \\
&\leqslant O(n) \cdot \sqrt{\mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \mathrm{Tr}\left(\left(u_{\vartheta,e,e'}(\boldsymbol{G}) u_{\vartheta,e,e'}(\boldsymbol{G})^\top\right)^{4q}\right)},
\end{aligned}$$

其中，最终不等式的推导与公式（18）类似。由此得到：

$$\ell_{4q}\left(\left\|u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)\right\|\right) \leqslant O\left(n^{1/4q}\right) \cdot \left(\mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \mathrm{Tr}\left(\left(u_{\vartheta,e,e'}(\boldsymbol{G}) u_{\vartheta,e,e'}(\boldsymbol{G})^\top\right)^{4q}\right)\right)^{1/8q} \tag{22}$$

再次使用“图矩阵矩界”对上述表达式进行界定。为应用 [AMP16中的引理 8.7 和 8.15]，，我们需要给出e 和 e' 在 ϑ中的最小顶点分隔集大小的下界。注意，由于 ϑ 是连通的，e 和e' 在 ϑ 中的最小顶点分隔集大小至少为 1。因此：

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \mathrm{Tr}\left(\left(u_{\vartheta,e,e'}(\boldsymbol{G}) u_{\vartheta,e,e'}(\boldsymbol{G})^\top\right)^{4q}\right) &\leqslant \frac{1}{|\mathcal{L}_\vartheta|^{4q}} (8q)^{8q(e(\vartheta)+v(\vartheta))} \cdot n^{4q(v(\vartheta)-1)+1} \\
&\leqslant (8v(\vartheta)q)^{8q(e(\vartheta)+v(\vartheta))} n^{-4q+1}\,.
\end{aligned}$$

将上述结果代入公式（22）可得：

$$\ell_{4q}\left(\left\|u_{\vartheta,e,e'}(\boldsymbol{M}_\varepsilon)\right\|\right) \leqslant O(n^{3/8q}) \cdot \frac{(8v(\vartheta)q)^{e(\vartheta)+v(\vartheta)}}{\sqrt{n}}$$

我们明确选择$q = \log n$，并将上述结果代入公式（21），得到：

$$\ell_{\log n}(\kappa_2(\boldsymbol{M}))) \leqslant O(1) \cdot \frac{\varepsilon \cdot (8v(\vartheta)\log n)^{e(\vartheta)+v(\vartheta)}}{\sqrt{n}}\,. \tag{23}$$

关于 $\sigma^2(\boldsymbol{M})$ 的高概率下界。 为了以高概率给出 $\sigma^2(\boldsymbol{M})$ 的下界，我们首先通过一个在解析上更易处理的表达式，给出 $\sigma^2(\boldsymbol{M})$ 的确定性下界。然后，我们证明该表达式的一个高概率下界。

任取一个矩阵 $M \in \mathbb{R}^{n\times n}$。令 $p(\boldsymbol{G}) = \chi_\vartheta\left(\sqrt{1-\varepsilon}M + \sqrt{\varepsilon}\boldsymbol{G}\right)$。对于函数 $f\colon \mathbb{R}^d \to \mathbb{R}$，令 $D^t f$ 表示一个张量，其中包含所有阶数为 t 的偏导数，且求导方向互不重复。由于 p 是多重线性的，所有其他阶数为 t 的导数都为 0。因此，根据泰勒展开

$$p(\boldsymbol{G}) = \sum_{t=0}^{e(\vartheta)} \frac{1}{t!} \left\langle D^t\, p(0), \boldsymbol{G}^{\otimes t}\right\rangle,$$

其中，$\boldsymbol{G}^{\otimes t}$ 表示包含 t 个互不相同的$\boldsymbol{G}$的元素之所有乘积的张量。注意，对于指标$s \neq t \geqslant 1$，上方求和中的对应项互不相关且均值为 0。因此，利用常数项不会影响方差这一点，可得

$$\sigma^2(M) = \operatorname*{Var}_{\boldsymbol{G}} p(\boldsymbol{G}) = \sum_{t=1}^{e(\vartheta)} \frac{1}{(t!)^2} \mathop{\mathbb{E}}_{\boldsymbol{G}} \left\langle D^t\, p(0), \boldsymbol{G}^{\otimes t} \right\rangle^2 .$$

此外，

$$D^t\, p(0) = \sqrt{\varepsilon}^t \cdot D^t\, \chi_\vartheta\Big(\sqrt{1-\varepsilon} M\Big) = \sqrt{\varepsilon}^t \sqrt{1-\varepsilon}^{e(\vartheta)-t} \cdot D^t\, \chi_\vartheta(M) .$$

因此，我们得到

$$\begin{aligned} \sigma^2(M) &= \sum_{t \geqslant 1} \frac{\varepsilon^t (1-\varepsilon)^{e(\vartheta)-t}}{(t!)^2} \cdot \mathop{\mathbb{E}}_{\boldsymbol{G}} \left\langle \mathrm{D}^t\, \chi_\vartheta(M), \boldsymbol{G}^{\otimes t} \right\rangle^2 = \sum_{t \geqslant 1} \frac{\varepsilon^t (1-\varepsilon)^{e(\vartheta)-t}}{(t!)^2} \cdot \left\| \mathrm{D}^t\, \chi_\vartheta(M) \right\|_2^2 \\ &\geqslant \varepsilon \cdot (1-\varepsilon)^{e(\vartheta)-1} \cdot \|\nabla \chi_\vartheta(M)\|_2^2 \end{aligned} \tag{24}$$

因此，要证明 $\sigma^2(\boldsymbol{M})$ 的高概率下界，只需证明 $\|\nabla \chi_\vartheta(\boldsymbol{M})\|^2$ 的高概率下界。为此，我们将把 $\|\nabla \chi_\vartheta(\boldsymbol{M})\|_2^2$ 表示为“确定性项”和“波动项”之和。确定性部分和波动部分来自将该表达式写成子图统计量的线性组合，而这些统计量对应于$\boldsymbol{M}$。

接下来，我们还将考虑多重图的子图统计量。对于多重图 ι 和矩阵M，我们将子图统计量 $\chi_\iota(M)$ 定义为 $\sum_{\pi \in \mathcal{L}_\iota} \prod_{ab \in E(\iota)} h_{\mathrm{mult}(ab)}\Big(M_{\pi(a),\pi(b)}\Big)$，其中 $\mathrm{mult}(ab)$ 指边ab 在 ι 中的重数，而h_k 指第k 个（首一）埃尔米特多项式。

现在，我们写作：

$$\begin{aligned} \|\nabla \chi_\vartheta(\boldsymbol{M})\|_2^2 &= \sum_{\{i,j\} \in \binom{[n]}{2}} \frac{1}{|\mathcal{L}_\vartheta|} \left(\sum_{e \in E(\vartheta)} \sum_{\pi \in \mathcal{L}_{\vartheta,e,\{i,j\}}} \prod_{ab \in E(\vartheta)\setminus e} \boldsymbol{M}_{\pi(a),\pi(b)} \right)^2 \\ &= \sum_{\{i,j\} \in \binom{[n]}{2}} \frac{1}{|\mathcal{L}_\vartheta|} \sum_{e,e' \in E(\vartheta)} \sum_{\substack{\pi \in \mathcal{L}_{\vartheta,e,\{i,j\}} \\ \pi' \in \mathcal{L}_{\vartheta,e',\{i,j\}}}} \prod_{ab \in E(\vartheta)\setminus e} \boldsymbol{M}_{\pi(a),\pi(b)} \cdot \prod_{ab \in E(\vartheta)\setminus e'} \boldsymbol{M}_{\pi'(a),\pi'(b)} . \end{aligned} \tag{25}$$

现在，我们定义重叠函数的概念。重叠函数 Ov 是从顶点的一个子集$S \subseteq V(\vartheta)$ 到$V(\vartheta)$的单射。我们用 $\mathcal{O}_S$ 表示所有来自 S 的重叠函数的集合。给定重叠函数 $\mathrm{Ov} \in \mathcal{O}_S$，我们用 ϑ_{Ov} 表示通过取两个 ϑ 副本并将第一个副本中的顶点$v \in S$ 与第二个副本中的 $\mathrm{Ov}(v)$ 重叠而得到的图。

有了这个记号，我们可以将方程 (25) 写为：

$$(25) = \frac{1}{|\mathcal{L}_\vartheta|} \sum_{e,e' \in E(\vartheta)} \sum_{V(\vartheta) \supseteq S \supseteq e} \sum_{\substack{\mathrm{Ov} \in \mathcal{O}_S \\ \mathrm{Ov}(e) = e'}} \sum_{\pi \in \mathcal{L}_{\vartheta_{\mathrm{Ov}}}} \prod_{ab \in E(\vartheta_{\mathrm{Ov}})\setminus\{e,e'\}} \boldsymbol{M}^{\mathrm{mult}(ab)}_{\pi(a),\pi(b)} \tag{26}$$

在上述表达式中，$\mathrm{mult}(ab)$ 等于 1 或 2；当其等于 2 时，相应的项可以拆分为$h_2\Big(\boldsymbol{M}_{\pi(a),\pi(b)}\Big) + 1$。

将 $\mathcal{C}_{\mathrm{proto}}(\vartheta,e,e')$ 定义为所有形如 ϑ_{Ov} 的多重图所构成的多重集，其中 $\mathrm{Ov} \in \mathcal{O}_S$ 满足$S \supseteq e$ 且 $\mathrm{Ov}(e) = e'$。

We n现在，通过以下迭代方法，从 $\mathcal{C}_{\mathrm{proto}}(\vartheta,e,e')$ 构造一个多重集 $\mathcal{C}(\vartheta, e, e')$ d.

- 将 $\mathcal{C}_0$ 定义为 $\mathcal{C}$原型(ϑ, e, e')，并对每个图 $\iota \in \mathcal{C}$原型(ϑ,e,e')，将其所有重边着为红色。
- 对于时间$t \geqslant 1$，执行以下操作直至终止。

- 假设每个图 $\iota \in \mathcal{C}_{t-1}$ 都没有红色双重边，定义 $\mathcal{C}(\vartheta, e, e') := \mathcal{C}_{t-1}$ 并终止。
- 否则，选取一个 ι 含有红色双重边的图。
- 选取一条红色双重边e ，位于 ι 中，并定义以下两个新图： ι_1 删除e得到的图，以及 ι_2 将e 重新染为蓝色得到的图。
- 令 $\mathcal{C}_t := \mathcal{C}_{t-1} \cup \{\iota_1, \iota_2\} \setminus \{\iota\}$。

读者应将把 ι "分裂"为 ι_1 和 ι_2 理解为将x^2 写成$h_2(x) + 1$。最后，将 $\mathcal{C}(\vartheta)$定义为如下多重集：$\bigcup_{e,e'} \mathcal{C}(\vartheta, e, e')$。

不难验证，我们可以写成：

$$(26) = \frac{1}{|\mathcal{L}_\vartheta|} \sum_{\iota \in \mathcal{C}(\vartheta)} \chi_\iota(\boldsymbol{M}) .$$

我们可以根据 ι 为空还是非空，将上述和拆分为一个确定性项和一个波动项。具体而言：

$$(26) = \underbrace{\frac{1}{|\mathcal{L}_\vartheta|} \sum_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ empty}}} \chi_\iota(\boldsymbol{M})}_{\text{Deterministic term}} + \underbrace{\frac{1}{|\mathcal{L}_\vartheta|} \sum_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ nonempty}}} \chi_\iota(\boldsymbol{M})}_{\text{Fluctuation term}}$$

第一项不依赖于 $\boldsymbol{M}$ ，因为每个求和项都确定性地等于 $|\mathcal{L}_\iota|$。注意，确定性部分的大小至少为 1，因为对于 ι ，若其由根据恒等映射叠加两个 $\vartheta \setminus e$ 的副本而来，其中某条边为 e，则有 $|\mathcal{L}_\iota| = |\mathcal{L}_\vartheta|$。我们记录这一观察如下：

$$\text{Deterministic term} \geqslant 1 \tag{27}$$

我们将证明，以高概率，波动项的大小至多为多对数$n\sqrt{n}$ ，方法是对适当选取的偶数 q，界定其~~ℓ_q~~范数。

$$\ell_q(\text{Fluctuation term}) = \ell_q\left(\frac{1}{|\mathcal{L}_\vartheta|} \sum_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ nonempty}}} \chi_\iota(\boldsymbol{M})\right) \leqslant \frac{1}{|\mathcal{L}_\vartheta|} \sum_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ nonempty}}} \ell_q(\chi_\iota(\boldsymbol{M})) \tag{28}$$

现在我们分析上述右端中的一个求和项。

$$\begin{aligned}
\ell_q(\chi_\iota(\boldsymbol{M}))^q &= \mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathsf{P}} \chi_\iota(\boldsymbol{M})^q \\
&\leqslant \mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \chi_\iota(\boldsymbol{G})^q + \sqrt{n \cdot \mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \chi_\iota(\boldsymbol{G})^{2q}} \\
&\leqslant O(\sqrt{n}) \cdot \sqrt{\mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \chi_\iota(\boldsymbol{G})^{2q}}
\end{aligned}$$

根据图矩阵矩界 [AMP16，引理 8.7 和 8.15], 可得：

$$\mathop{\mathbb{E}}_{\boldsymbol{G} \sim \mathsf{N}} \chi_\iota(\boldsymbol{G})^{2q} \leqslant (8q)^{2qe(\iota)} \cdot (2q)^{2qv(\iota)} n^{q(v(\iota) + \mathrm{iso}(\iota))},$$

其中 $\mathrm{iso}(\iota)$ 是 ι 中孤立顶点的数量，因此我们得到：

$$\ell_q(\chi_\iota(\boldsymbol{M})) \leqslant O\left(n^{1/4q}\right) \cdot (8q)^{e(\iota) + v(\iota)} \cdot n^{(v(\iota) + \mathrm{iso}(\iota))/2}$$

因此：

$$(28) \leqslant O\left(n^{1/4q}\right) \cdot \frac{v(\vartheta)^{v(\vartheta)}}{n^{v(\vartheta)}} \cdot 2^{v(\vartheta)} v(\vartheta)^{v(\vartheta)} \cdot \max_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ nonempty}}} (8q)^{e(\iota)+v(\iota)} \cdot n^{(v(\iota)+\mathrm{iso}(\iota))/2} \tag{29}$$

任意 ι in $\mathcal{C}(\vartheta)$ 均可由两个 ϑ 副本通过某个映射 Ov $\in \mathcal{O}_{S(\iota)}$ 在某个大小至少为 2 的顶点子集$S(\iota)$ 上重叠得到，因此 $v(\iota) = 2v(\vartheta) - |S(\iota)|$。特别地，上述右端至多为

$$\max_{\substack{\iota \in \mathcal{C}(\vartheta) \\ \iota \text{ nonempty}}} O\left(n^{1/4q}\right) \cdot (16qv(\vartheta))^{2e(\vartheta)+2v(\vartheta)} \cdot \sqrt{n}^{\mathrm{iso}(\iota)-|S(\iota)|} \tag{30}$$

我们将证明，对于任意 $\iota \in \mathcal{C}(\vartheta)$，上述项中 n 的幂始终至多为 $-1/2$，方法是证明 ι 中的孤立顶点集合I 是 $S = S(\iota)$的真子集。

假设I 为空，则由于S 非空，结论成立。以下我们假设I 非空。定义 $\tilde{\iota}$ 为通过映射 Ov 将两个重叠的 ϑ 副本取（多重集）并集而得到的图（不删除任何边）。注意，$\tilde{\iota}$是连通图，因为 ϑ 是连通图；而$\tilde{\iota}$是通过在一个非空集合上重叠两个 ϑ 副本得到的。由于 ι 不是空图，其非孤立顶点集合T 必定非空。由于$I = V(\tilde{\iota}) \setminus T$，且$\tilde{\iota}$是连通图，因此$I$ 与T之间必定存在一个非空的割。注意，对于 ι中的孤立顶点v，$\tilde{\iota}$中与v 关联的每条边都是重边；此外，如果顶点v 与一条重边关联，则它必定属于S。I与T 之间割中的每条边都是重边，因此必定存在某个顶点$u \in T$ 属于S，故I 是T的真子集，由此可得：

$$(30) \leqslant O\left(n^{1/4q}\right) \frac{(16qv(\vartheta))^{2(e(\vartheta)+v(\vartheta))}}{\sqrt{n}}$$

明确选择 $q = \log n$，并结合上述内容以及（28）$\leqslant$（29）$\leqslant$（30），可得：

$$\ell_{\log n}(\text{Fluctuation term}) \leqslant \frac{(32v(\vartheta)\log n)^{2(e(\vartheta)+v(\vartheta))}}{\sqrt{n}} \tag{31}$$

全变差距离界。根据方程 (20)、方程 (23)、方程 (24)、方程 (27)和方程 (31)，以及马尔可夫不等式，我们有：关于 $\boldsymbol{M} \sim \mathrm{P}$，至少以 $1 - {}^{1}{}_{n^2}$ 的概率成立：

$$\begin{aligned}
\kappa_1(\boldsymbol{M}) &\leqslant \sqrt{\varepsilon} \cdot \mathrm{polylog}\, n \\
\kappa_2(\boldsymbol{M}) &\leqslant \varepsilon \cdot \frac{\mathrm{polylog}\, n}{\sqrt{n}} \\
\sigma^2(\boldsymbol{M}) &\geqslant \varepsilon(1-\varepsilon)^{v(\vartheta)-1} \cdot \left(1 - \frac{\mathrm{polylog}\, n}{\sqrt{n}}\right).
\end{aligned}$$

我们将使用 $\mathcal{E}$ 来指代上述界成立的事件。对于 M 的固定取值，我们将 $\pi_{\vartheta M}$ 定义为 $\chi_\vartheta(\sqrt{1-\varepsilon}\mathrm{M} + \sqrt{\varepsilon}\mathrm{G})$ 的分布，其中 $\mathrm{G} \sim \mathrm{N}$；并将 $\pi_{\vartheta M}$ 定义为 $\chi_\vartheta(\sqrt{1-\varepsilon}\mathrm{M}) + \tilde{\sigma}(\mathrm{M})_ \cdot \mathrm{g}$ 的分布，其中 $\mathrm{g} \sim \mathcal{N}(0,1)$。根据定理 6.2，当 $\mathrm{M} \in \mathcal{E}$ 时，我们有：

$$d_{\mathrm{TV}}\left(\pi_{\vartheta,M}, \underline{\pi}_{\vartheta,M}\right) \leqslant \frac{\varepsilon^{3/2} \cdot \mathrm{polylog}\, n}{\Omega(\varepsilon) \cdot \sqrt{n}} = \frac{\sqrt{\varepsilon} \cdot \mathrm{polylog}\, n}{\sqrt{n}}.$$

现在，观察 $\pi_\vartheta = \mathbb{E}_{\boldsymbol{M} \sim \mathrm{P}} \pi_{\vartheta、\boldsymbol{M}}$ 和 $\pi_\vartheta = \mathbb{E}_{\underline{\boldsymbol{M}} \sim \mathrm{P}} \pi_{\vartheta、\boldsymbol{M}。_}$

最后，我们可以通过以下一系列不等式证明所需的结论。

$$\begin{aligned}
d_{\mathrm{TV}}(\pi_\vartheta, \underline{\pi}_\vartheta) &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathrm{P}} d_{\mathrm{TV}}\left(\pi_{\vartheta,\boldsymbol{M}}, \underline{\pi}_{\vartheta,\boldsymbol{M}}\right) \\
&= \mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathrm{P}} \mathbf{1}[\mathcal{E}] \cdot d_{\mathrm{TV}}\left(\pi_{\vartheta,\boldsymbol{M}}, \underline{\pi}_{\vartheta,\boldsymbol{M}}\right) + \mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathrm{P}} \mathbf{1}[\overline{\mathcal{E}}] \cdot d_{\mathrm{TV}}\left(\pi_{\vartheta,\boldsymbol{M}}, \underline{\pi}_{\vartheta,\boldsymbol{M}}\right) \\
&\leqslant \frac{\sqrt{\varepsilon} \cdot \mathrm{polylog}\, n}{\sqrt{n}} + \frac{1}{n^2}
\end{aligned}$$

□

6.2 无穷范数界

在本节中，我们证明定理 6.5，其重述如下。

引理 （重述 定理 6.5）。 令 $D \geqslant$ 对数 n 对数 对数 n，并令 $\delta \geqslant 0$ 满足 $\chi^2_D(\mathsf{P} \,\|\, \mathsf{N}) \leqslant \delta$，并假设 $\varepsilon \geqslant 1/$ 多对数 n。则存在一个绝对常数 $c > 0$ 使得

$$\|\underline{\pi}_\vartheta - \underline{\nu}_\vartheta\|_\infty \leqslant O(\delta^{c\varepsilon^2} + n^{-c\varepsilon^2}) .$$

证明。 对于任意 $x \in \mathbb{R}$，根据傅里叶反演可写为：

$$\begin{aligned}\sqrt{2\pi}|\underline{\pi}_\vartheta(x) - \underline{\nu}_\vartheta(x)| &\leqslant \int_{-\infty}^{\infty} |\widehat{\underline{\pi}}_\vartheta(\xi) - \widehat{\underline{\nu}}_\vartheta(\xi)| \,\mathrm{d}\xi \\ &= \int_{[-T,T]} |\widehat{\underline{\pi}}_\vartheta(\xi) - \widehat{\underline{\nu}}_\vartheta(\xi)| \,\mathrm{d}\xi + \int_{\mathbb{R}\setminus[-T,T]} |\widehat{\underline{\pi}}_\vartheta(\xi) - \widehat{\underline{\nu}}_\vartheta(\xi)| \,\mathrm{d}\xi \end{aligned} \tag{32}$$

对于稍后选取的某个 $T \in \mathbb{R}$。我们通过矩匹配（有界 LDLR）来界定第一项，并利用加入的噪声会衰减较大的傅里叶系数这一事实来界定第二项。

情形 $|\xi| > T$。 回顾 $\tilde{\sigma}(\boldsymbol{M}) \geqslant \frac{\varepsilon}{2}(1-\varepsilon)^{e(\vartheta)-1}$ 并注意到

$$\begin{aligned}|\widehat{\underline{\pi}}_\vartheta(\xi)| &= \left| \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{P}} \mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)} \exp\left(i\xi \left(\sqrt{1-\varepsilon}^{\,e(\vartheta)} \chi_\vartheta(\boldsymbol{M}) + \tilde{\sigma}(\boldsymbol{M}) \cdot g \right) \right) \right| \\ &\leqslant \left| \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{P}} \exp\left(i\xi \left(\sqrt{1-\varepsilon}^{\,e(\vartheta)} \chi_\vartheta(\boldsymbol{M}) \right) \right) \cdot \exp\left(-\xi^2 \tilde{\sigma}^2(\boldsymbol{M})/2 \right) \right| \\ &\leqslant \exp\left(-\xi^2 \varepsilon^2 (1-\varepsilon)^{2e(\vartheta)-2}/8 \right) .\end{aligned}$$

当 $\mathsf{P} = \mathsf{N}$时，上述结果为 $|\widehat{\nu}_\vartheta(\xi)|$建立了相同的界。将此界代入方程 (32) 的第二项并进行积分，得到：

$$\int_{\mathbb{R}\setminus[-T,T]} |\widehat{\underline{\pi}}_\vartheta(\xi) - \widehat{\underline{\nu}}_\vartheta(\xi)| \,\mathrm{d}\xi \leqslant \sqrt{\frac{4\pi}{\varepsilon^2(1-\varepsilon)^{2e(\vartheta)-2}}} \cdot \exp\left(-\frac{T^2\varepsilon^2(1-\varepsilon)^{2e(\vartheta)-2}}{8} \right) \tag{33}$$

现在我们可以对第一项进行界定了。

情形 $|\xi| \leqslant T$. 在这种情况下，注意到：

$$\begin{aligned}|\widehat{\underline{\pi}}_\vartheta(\xi) - \widehat{\underline{\nu}}_\vartheta(\xi)| &\leqslant |\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)| + d_{\mathrm{TV}}(\underline{\pi}_\vartheta, \pi_\vartheta) + d_{\mathrm{TV}}(\underline{\nu}_\vartheta, \nu_\vartheta) \\ &\leqslant |\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)| + \frac{\operatorname{polylog} n}{\sqrt{n}} .\end{aligned} \tag{34}$$

为了控制 $|\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)|$，我们将利用这样一个事实： π_ϑ 和 ν_ϑ 的矩近似匹配。事实上，根据泰勒定理，对于一个我们将在证明后续部分选取的$k \ll D$，可以写成：

$$\begin{aligned}|\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)| &\leqslant \left| \sum_{t=0}^{k-1} \mathop{\mathbb{E}}_{\boldsymbol{M}\sim \mathrm{U}_\varepsilon \mathsf{P}} \frac{(i\xi\chi_\vartheta(\boldsymbol{M}))^t}{t!} - \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{N}} \frac{(i\xi\chi_\vartheta(\boldsymbol{M}))^t}{t!} \right| + \mathop{\mathbb{E}}_{\boldsymbol{M}\sim \mathrm{U}_\varepsilon \mathsf{P}} \frac{|\xi\chi_\vartheta(\boldsymbol{M})|^k}{k!} + \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{N}} \frac{|\xi\chi_\vartheta(\boldsymbol{M})|^k}{k!} \\ &\leqslant \sum_{t=0}^{k-1} \frac{|\xi|^t}{t!} \sqrt{\chi^2_D(\mathsf{P}\,\|\,\mathsf{N}) \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{N}} \chi_\vartheta(\boldsymbol{M})^{2t}} + 2 \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{N}} \frac{|\xi\chi_\vartheta(\boldsymbol{M})|^k}{k!} + \frac{|\xi|^k}{k!} \sqrt{\chi^2_D(\mathsf{P}\,\|\,\mathsf{N}) \mathop{\mathbb{E}}_{\boldsymbol{M}\sim\mathsf{N}} \chi_\vartheta(\boldsymbol{M})^{2k}} .\end{aligned}$$

令 C 为一个充分大的绝对常数。根据 定理 6.4，我们知道 $\mathbb{E}_{\boldsymbol{M}\sim\mathsf{N}} \chi_\vartheta(\boldsymbol{M})^q \leqslant O(q)^{q/2}$，因此最后两项至多为 $(1+\sqrt{\delta})^{C^k|\xi|^k} k^{k/2}$。类似地，第一项至多为 $\sqrt{\delta} \sum_{t=0}^{k-1} C^t|\xi|^t t^{t/2}$。综上，我们得到

$$|\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)| \leqslant \sqrt{\delta} \sum_{t\geqslant 0} \frac{C^t|\xi|^t}{t^{t/2}} + \frac{C^k|\xi|^k}{k^{k/2}} \leqslant \sqrt{\delta}(1 + C|\xi|) \exp\left(C^2|\xi|^2 \right) + \frac{C^k|\xi|^k}{k^{k/2}}$$

$$\leqslant \sqrt{\delta}(1+CT)\exp(C^2T^2) + \frac{C^kT^k}{k^{k/2}}. \tag{35}$$

根据方程 (34)和方程 (35)，我们有：

$$\int_{[-T,T]} |\widehat{\pi}_\vartheta(\xi) - \widehat{\nu}_\vartheta(\xi)|\,\mathrm{d}\xi \leqslant \frac{T\,\mathrm{polylog}\,n}{\sqrt{n}} + \frac{C^kT^{k+1}}{k^{k/2}} + \sqrt{\delta}(CT^2+T)\exp(C^2T^2). \tag{36}$$

我们取 $k = \log n$，并取 $T = \alpha \min\left\{\sqrt{\log n},\ \sqrt{\log \frac{1}{\delta}}\right\}$ 其中常数 $\alpha > 0$ 足够小。将 方程 (33) 和方程 (36) 代入方程 (32)，对于一个（新的）足够大的常数 C 和足够小的常数 c，我们得到：

$$\|\pi_\vartheta - \nu_\vartheta\|_\infty \leqslant O\left(\frac{1}{\varepsilon}\right)\cdot\exp\left(-c\varepsilon^2T^2\right) + \frac{T\,\mathrm{polylog}\,n}{\sqrt{n}} + \frac{C^kT^{k+1}}{k^{k/2}} + C\sqrt{\delta}T^2\exp(C^2T^2).$$

利用 $\varepsilon \geqslant 1/$ 多对数 n，对于一个（新的）足够小的常数 α，我们可以看到：

$$\|\pi_\vartheta - \nu_\vartheta\|_\infty \leqslant O\left(n^{-\alpha\varepsilon^2} + \delta^{\alpha\varepsilon^2}\right). \qquad \square$$

致谢

我们感谢 Sam Hopkins 在低次启发式方面投入无数时间进行富有启发性的交流。S.M. 感谢 Sam Hopkins 和 Siqi Liu 就本文所研究问题的原型版本进行讨论，也感谢 Prasad Raghavendra 和 Tselil Schramm 进行许多鼓舞人心的交流。我们感谢 Shivam Nadimpalli 向我们指出 [Cha09]。我们还要感谢 Rares Darius Buhai 提供有益的交流，并感谢匿名审稿人对本文早期版本提出的宝贵意见。

最后，我们感谢美国数学研究所在“平均情形复杂性中的低次多项式方法”研讨会期间提供的热情接待与支持。本次合作正是在该研讨会上开始的；我们也感谢组织者 Sam Hopkins、Tselil Schramm 和 Alex Wein 为促成此次活动所做的工作。

参考文献

[ABI+23] Damiano Abram、Amos Beimel、Yuval Ishai、Eyal Kushilevitz 和 Varun Narayanan。《基于植入图的密码学：使用对数规模消息的安全性》。载于 Guy N. Rothblum 和 Hoeteck Wee 编，《密码学理论——第21届国际会议，TCC 2023》，中国台湾台北，2023年11月29日至12月2日，论文集第一部分，《计算机科学讲义》第14369卷，第286–315页。施普林格，2023年。2[AIM24] 平均情形复杂性低次多项式方法研讨会。http://admin. aimath.org/resources/lowdegreecomplexity/participantlist/，2024年。2[AMP16] Kwangjun Ahn、Dhruv Medarametla 和 Aaron Potechin。图矩阵：范数界及其应用。arXiv预印本 arXiv:1604.03423，2016年。5，9，31，32，34[AW85] Miklos Ajtai 和 Avi Wigderson。概率常深度电路的确定性模拟。载于第26届计算机科学基础年会（sfcs 1985），第11–19页。IEEE，1985年。3[BB20] Matthew Brennan 和 Guy Bresler。秘密泄露带来的可约性与统计-计算差距。载于学习理论会议，第648–847页。PMLR，2020年。1

[BBH+21] Matthew S Brennan、Guy Bresler、Sam Hopkins、Jerry Li 和 Tselil Schramm。统计查询算法与低次数测试几乎等价。载于学习理论会议，第774–774页。PMLR，2021年。1[BC90] Stanford S Bonan 和 Dean S Clark。Hermite多项式与Freud多项式的估计。《逼近理论杂志》，63(2):210–224，1990年。12[BDF+22] Andrej Bogdanov、Krishnamoorthy Dinesh、Yuval Filmus、Yuval Ishai、Avi Kaplan 和 Aksha-yaram Srinivasan。简单源的有界不可区分性。载于第13届理论计算机科学创新会议（ITCS 2022），第26–1页。Schloss Dagstuhl–Leibniz-Zentrum f¨ur Informatik，2022年。3[伯努利分布18] Ross Berkowitz。G(n, p)中团的局部极限定理。arXiv预印本 arXiv:1811.03527，2018年。5，6[BH25] Guy Bresler 和 Alina Harbuzova。通过Gram-Schmidt扰动实现尖峰协方差模型与尖峰Wigner模型的计算等价性。arXiv预印本 arXiv:2503.02802，2025年。1[BHJK25] Rares-Darius Buhai、Jun-Ting Hsieh、Aayush Jain 和 Pravesh K Kothari。拟多项式低次猜想是错误的。载于2025年IEEE第66届计算机科学基础年会（FOCS）。IEEE，2025年。2[BHK+16] Boaz Barak、Samuel B. Hopkins、Jonathan A. Kelner、Pravesh Kothari、Ankur Moitra 和 Aaron Potechin。植入团问题近乎紧的平方和下界。载于 Irit Dinur 编，《IEEE第57届计算机科学基础年会，FOCS 2016》，2016年10月9日至11日，美国新泽西州新不伦瑞克市Hyatt Regency，第428–437页。IEEE计算机学会，2016年。1，5[BIVW16] Andrej Bogdanov、Yuval Ishai、Emanuele Viola 和 Christopher Williamson。有界不可区分性与恢复秘密的复杂性。载于国际密码学年会，第593–618页。施普林格，2016年。3[BJRZ24] Andrej Bogdanov、Chris Jones、Alon Rosen 和 Ilias Zadik。植入随机子图问题的低次数安全性。载于 Elette Boyle 和 Mohammad Mahmoody 编，《密码学理论——第22届国际会议，TCC 2024》，意大利米兰，2024年12月2日至6日，论文集第二部分，《计算机科学讲义》第15365卷，第255–275页。施普林格，2024年。2[BKR23] Andrej Bogdanov、Pravesh K Kothari 和 Alon Rosen。公钥加密、局部伪随机生成器与低次数方法。载于密码学理论会议，第268–285页。施普林格，2023年。2[BMV+18] Jess Banks、Cristopher Moore、Roman Vershynin、Nicolas Verzelen 和 Jiaming Xu。聚类、稀疏主成分分析和子矩阵定位中的信息论界与相变。《IEEE信息论汇刊》，64(7):4872–4894，2018年。1[BR13] Quentin Berthet 和 Philippe Rigollet。稀疏主成分检测的复杂性理论下界。载于学习理论会议，第1046–1066页。PMLR，2013年。1[Bra08] Mark Braverman。多对数独立性可愚弄 AC^0 电路。《ACM学报》（JACM），57(5):1–10，2008年。3[BT06] Andrej Bogdanov 和 Luca Trevisan。平均情形复杂性。《理论计算机科学基础与趋势》，2(1):1–106，2006年。1

[Cha09] Sourav Chatterjee。特征值的涨落与二阶 Poincar´é 不等式。《概率论及相关领域》，143(1):1–40，2009年。5，28，37[CMRW23] Zongchen Chen、Dan Mikulincer、Daniel Reichman 和 Alexander S Wein。Metropolis 过程与模拟退火的时间下界。arXiv 预印本 arXiv:2312.13554，2023年。1[CMZ25] Zongchen Chen、Elchanan Mossel 和 Ilias Zadik。近线性规模的植入团难以被 Metropolis 过程发现。《Random Structures & Algorithms》，66(2):e21274，2025年。1[CSZ24] Zongchen Chen、Conor Sheehan 和 Ilias Zadik。低温 MCMC 阈值：稀疏张量主成分分析、稀疏回归和一种几何规则的情形。arXiv 预印本 arXiv:2408.00746，2024年。1[CW01] Anthony Carbery 和 James Wright。凸体上多项式的分布不等式与 L^q 范数不等式。《数学研究快报》，8(3):233–248，2001年。11[DGJ+10] Ilias Diakonikolas、Parikshit Gopalan、Ragesh Jaiswal、Rocco A Servedio 和 Emanuele Viola。有界独立性能够骗过半空间。《SIAM 计算杂志》，39(8):3441–3462，2010年。3[DILV24] Harm Derksen、Peter Ivanov、Chin Ho Lee 和 Emanuele Viola。伪随机性、对称性、平滑：I。发表于第39届计算复杂性会议（CCC 2024），第18–1页。Schloss Dagstuhl–Leibniz 信息学中心，2024年。12，43[DK22] Ilias Diakonikolas 和 Daniel Kane。通过格基约化进行非高斯分量分析。发表于学习理论会议，第4535–4547页。PMLR，2022年。2[DKN10] Ilias Diakonikolas、Daniel M Kane 和 Jelani Nelson。有界独立性能够骗过二次阈值函数。发表于2010 IEEE第51届计算机科学基础年会，第11–20页。IEEE，2010年。3[DKS17] Ilias Diakonikolas、Daniel M Kane 和 Alistair Stewart。高维高斯分布及高斯混合分布鲁棒估计的统计查询下界。发表于2017 IEEE第58届计算机科学基础年会（FOCS），第73–84页。IEEE，2017年。1[DKWB24] Yunzi Ding、Dmitriy Kunisky、Alexander S. Wein 和 Afonso S. Bandeira。求解稀疏主成分分析的亚指数时间算法。《计算数学基础》，24(3):865–914，2024年。5[FGR+17] Vitaly Feldman、Elena Grigorescu、Lev Reyzin、Santosh S Vempala 和 Ying Xiao。统计推断植入团的统计算法与下界。《美国计算机协会学报》（JACM），64(2):1–37，2017年。1[FK18] Michael A Forbes 和 Zander Kelley。任意顺序下用于只读分支程序的伪随机生成器。发表于2018 IEEE第59届计算机科学基础年会（FOCS），第946–955页。IEEE，2018年。3[GJ21] David Gamarnik 和 Aukosh Jagannath。重叠间隙性质与适用于 p 自旋模型的近似消息传递算法。《概率年刊》，49(1):180–205，2021年。1[GJJ+20] Mrinalkanti Ghosh、Fernando Granha Jeronimo、Chris Jones、Aaron Potechin 和 Goutham Rajendran。通过植入仿射平面得到的 Sherrington–Kirkpatrick哈密顿量平方和下界。发表于2020 IEEE第61届计算机科学基础年会（FOCS），第954–965页。IEEE，2020年。1[GJW24] David Gamarnik、Aukosh Jagannath 和 Alexander S Wein。布尔电路、低次多项式与朗之万动力学下随机优化问题的困难性。《SIAM 计算杂志》，53(1):1–46，2024年。1

[GK16] Justin Gilmer 和 Swastik Kopparty。随机图中三角形的局部中心极限定理。《Random Structures & Algorithms》，48(4):732–750，2016年。5，6[Gol11] Oded Goldreich。关于 Levin 平均情形复杂性理论的札记。《Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation: In Collaboration with Lidor Avigad, Mihir Bellare, Zvika Brakerski, Shafi Goldwasser, Shai Halevi, Tali Kaufman, Leonid Levin, Noam Nisan, Dana Ron, Madhu Sudan, Luca Trevisan, Salil Vadhan, Avi Wigderson, David Zuckerman》，第233–247页，2011年。1[HKP+17] Samuel B Hopkins、Pravesh K Kothari、Aaron Potechin、Prasad Raghavendra、Tselil Schramm 和 David Steurer。平方和方法在检测隐藏结构方面的威力。发表于2017 IEEE第58届计算机科学基础年会（FOCS），第720–731页。IEEE，2017年。1[HLV18] Elad Haramaty、Chin Ho Lee 和 Emanuele Viola。有界独立性加噪声能够骗过乘积。《SIAM 计算杂志》，47(2):493–523，2018年。3[Hop18] Samuel Hopkins。统计推断与平方和方法。博士论文，康奈尔大学，2018年。1，2[HS17] Samuel B Hopkins 和 David Steurer。基于少量样本的高效贝叶斯估计：社区检测及相关问题。发表于2017 IEEE第58届计算机科学基础年会（FOCS），第379–390页。IEEE，2017年。1[HW21] Justin Holmgren 和 Alexander S Wein。低次猜想的反例。发表于第12届理论计算机科学创新会议（ITCS 2021）。Schloss-Dagstuhl-Leibniz 信息学中心，2021年。2[Jer92] Mark Jerrum。大型团难以被 Metropolis 过程发现。《Random Structures & Algorithms》，3(4):347–359，1992年。1[JL09] Iain M Johnstone 和 Arthur Yu Lu。高维主成分分析的一致性与稀疏性。《Journal of the American Statistical Association》，104(486):682–693，2009年。5[JPR+21] Chris Jones、Aaron Potechin、Goutham Rajendran、Madhur Tulsiani 和 Jeff Xu。稀疏独立集的平方和下界。发表于第62届 IEEE 计算机科学基础年会（FOCS 2021），美国科罗拉多州丹佛，2022年2月7日至10日，第406–416页。IEEE，2021年。1[JPRX23] Chris Jones、Aaron Potechin、Goutham Rajendran 和 Jeff Xu。最稠密 k-子图的平方和下界。载于 Barna Saha 和 Rocco A. Servedio 编，《第55届美国计算机协会计算理论研讨会论文集》（STOC 2023），美国佛罗里达州奥兰多，2023年6月20日至23日，第84–95页。美国计算机协会，2023年。1[Kan11] Daniel M Kane。k-独立高斯变量能够骗过多项式阈值函数。发表于2011 IEEE第26届计算复杂性会议，第252–261页。IEEE，2011年。3[KB21] Dmitriy Kunisky 和 Afonso S Bandeira。Sherrington–Kirkpatrick哈密顿量的紧致4次平方和下界。《Mathematical Programming》，190:721–759，2021年。1[Kea98] Michael Kearns。基于统计查询的高效抗噪学习。《美国计算机协会学报》（JACM），45(6):983–1006，1998年。1[KPX24] Pravesh K. Kothari、Aaron Potechin 和 Jeff Xu。超稀疏随机图上独立集的平方和下界。载于 Bojan Mohar、Igor Shinkar 和 Ryan O’Donnell 编，《第56届美国计算机协会计算理论研讨会论文集》（STOC 2024），加拿大不列颠哥伦比亚省温哥华，2024年6月24日至28日，第1923–1934页。美国计算机协会，2024年。1

[Kra29] 米哈伊尔·克拉夫丘克。关于厄米特多项式的一种推广。《科学院报告》，189(620-622):5–3，1929年。12[Kra01] 伊利亚·克拉西科夫。非负二次型与正交多项式的界。《逼近理论杂志》，111(1):31–49，2001年。12[KWB19] Dmitriy Kunisky、Alexander S Wein 和 Afonso S Bandeira。《假设检验的计算困难性札记：基于低次似然比的预测》。载于 ISAAC 大会（国际分析、应用与计算学会），第 1–50 页。施普林格，2019年。1, 2[Lev86] 列昂尼德·A·列文。平均情形完全问题。《SIAM 计算杂志》，15(1):285–286，1986年。1[LV20] Chin Ho Lee 和 Emanuele Viola。关于有界独立性加噪声的更多研究：用于读一次多项式的伪随机生成器。《计算理论》，16(1):1–50，2020年。3[MRX20] Sidhanth Mohanty、Prasad Raghavendra 和 Jeff Xu。提升平方和下界：从 2 次到 4 次。载于 Konstantin Makarychev、Yury Makarychev、Madhur Tulsiani、Gautam Kamath 和 Julia Chuzhoy（编），第 52 届年度 ACM SIGACT 计算理论研讨会论文集，STOC 2020，美国伊利诺伊州芝加哥，2020年6月22—26日，第 840–853 页。美国计算机协会，2020年。1[MRZ16] 安德烈亚·蒙塔纳里、Daniel Reichman 和奥弗·泽伊图尼。论谱方法的局限性：从高斯隐藏团问题到高斯张量的秩一扰动。《IEEE信息论汇刊》，63(3):1572–1579，2016年。1[Nel73] 爱德华·纳尔逊。自由马尔可夫场。《泛函分析杂志》，12(2):211–227，1973年。11[O’D14] Ryan O’Donnell。《布尔函数分析》。剑桥大学出版社，2014年。10, 11, 45[Pet12] 瓦伦丁·V·彼得罗夫。《独立随机变量之和》，第 82 卷。施普林格科学与商业媒体，2012年。4, 6[Pol19] 尤里·波利安斯基。汉明空间中球面平均的超收缩性。《SIAM 离散数学杂志》，33(2):731–754，2019年。12, 43[SM62] SH·西拉兹迪诺夫和 M·马马托夫。关于密度的均值收敛。《概率论与应用》，7:433–437，1962年。6[SS22] 阿什温·萨赫和梅塔布·索尼。子图计数的局部极限定理。《伦敦数学会杂志》，105(2):950–1011，2022年。5, 6, 8[Sze75] Gabor Szeg˝o。《正交多项式》，第 23 卷。美国数学会，1975年。11, 12[Wei25] Alexander S Wein。《统计学的计算复杂性：低次多项式带来的新见解》。arXiv 预印本 arXiv:2506.10748，2025年。2[ZSWB22] Ilias Zadik、宋旻宰、Alexander S Wein 和琼·布鲁纳。基于格的方法在聚类中超越平方和。载于学习理论会议，第 1247–1248 页。PMLR，2022年。2

二项分布

本节将给出我们省略的证明，详见第 3.2 节。

引理 （定理 3.6的重述）。 设 $\gamma \leqslant 1/2$。对于任意 $y \in \mathbb{R}$ 满足 $|y| \leqslant o(\sqrt{\gamma n})$ 且 $n\gamma + y\sqrt{n\gamma(1-\gamma)} \in \{0, 1, \ldots\ldots, n\}$，

$$\mathbb{P}\left[\mathrm{Bin}(n,\gamma) = n\gamma + y\sqrt{n\gamma(1-\gamma)}\right] \geqslant \frac{1}{\sqrt{2n}} \exp\left(-\frac{y^2}{2} - O(|y|^3/\sqrt{\gamma n})\right).$$

证明。 令 $\beta := \gamma + y\sqrt{\gamma(1-\gamma)/n}$。于是，

$$\mathbb{P}[\mathrm{Bin}(n,\gamma) = \beta n] = \binom{n}{\beta n} \gamma^{\beta n} (1-\gamma)^{(1-\beta)n}.$$

首先，我们有

$$\binom{n}{\beta n} \geqslant \sqrt{\frac{1}{8n\beta(1-\beta)}} e^{nh(\beta)} \geqslant \frac{1}{\sqrt{2n}} e^{nh(\beta)},$$

其中 $h(\beta) = -\beta \log \beta - (1-\beta)\log(1-\beta)$。接下来，简单计算可得

$$e^{nh(\beta)} \gamma^{\beta n} (1-\gamma)^{(1-\beta)n} = e^{-nD(\beta\|\gamma)},$$

其中 $D(\beta\|\gamma) = \beta \log \frac{\beta}{\gamma} + (1-\beta)\log\left(\frac{1-\beta}{1-\gamma}\right)$ 是 $\mathrm{Ber}(\beta)$ 与 $\mathrm{Ber}(\gamma)$ 之间的 KL 散度。

根据界 $D(\gamma+\varepsilon\|\gamma) \leqslant \frac{\varepsilon^2}{2\gamma(1-\gamma)} + O(|\varepsilon|^3/\gamma^2)$，由定理 3.5可得，假设 $\gamma \leqslant 1/2$ 且 $\varepsilon \leqslant o(\gamma)$。令

$$|\varepsilon| = |y|\sqrt{\gamma(1-\gamma)/n} \leqslant |y|\sqrt{\gamma/n} \leqslant o(\gamma)$$

（根据假设），我们得到

$$n \cdot D(\beta\|\gamma) \leqslant y^2/2 + O(|y|^3/\sqrt{\gamma n}).$$

证明完毕。 □

引理 （重述定理 3.7）。 设 $\gamma \leqslant 1/2$，并令 $x \sim$ 伯努利分布$(\gamma)^n$ 以及 $y = \frac{1}{2\sqrt{n\gamma(1-\gamma)}} \sum_{i=1}^n (x_i - (2\gamma - 1))$。则对于任意 $k \in \mathbb{N}$，有 $\mathbb{E}[y^{2k}] \leqslant (2k-1)!! \cdot (1 + O(k^3/\gamma n))$。

证明。 令 $z_i = \frac{1}{2\sqrt{\gamma(1-\gamma)}}(x_i - (2\gamma-1))$。直接计算可得，$\mathbb{E}[z_i] = 0$、$\mathbb{E}[z_i^2] = 1$以及$0 < \mathbb{E}[z_i^k] \leqslant \left(\frac{1-\gamma}{\gamma}\right)^{k/2-1} \leqslant \gamma^{-(k-2)/2}$对于$k \geqslant 3$成立（这里假设 $\gamma \leqslant 1/2$）。

我们有 $y = \frac{1}{\sqrt{n}} \sum_{i=1}^n z_i$，因此

$$\mathbb{E}[y^{2k}] = n^{-k} \sum_{i_1, i_2, \ldots, i_{2k} \in [n]} \mathbb{E}[z_{i_1} z_{i_2} \cdots z_{i_{2k}}].$$

由于 $\mathbb{E}[z_i] = 0$，如果元组 $(i_1, \ldots, i_{2k})$ 中任意一个 $u \in [n]$ 恰好出现一次，则求和项为零。令 m_u 为 u 在 $(i_1, \ldots, i_{2k})$中的重数，并令 $\ell = \sum_u (m_u - 2)\mathbf{1}(m_u \geqslant 3)$，也就是说，这些重数都大于 2。注意，$\ell$ 必须为偶数，否则某个 m_u 会等于 1。于是，

$$\mathbb{E}[z_{i_1} \cdots z_{i_{2k}}] = \prod_{u \in [n]} \mathbb{E}[z_u^{m_u}] \leqslant \prod_{u \in [n]} \gamma^{-(m_u-2)/2 \cdot \mathbf{1}(m_u \geqslant 3)} = \gamma^{-\ell/2}.$$

主导项出现在 $\ell = 0$ 时，即每个指标出现 0 次或 2 次。在这种情况下，可以先在 1、2、. . .、$2k$之间选取一个完美匹配（用于重复模式），再按顺序从 $[n]$ 中选取k 个不同的指标，从而构造出这些元组。匹配的数量为$(2k-1)!!$，而k 个指标的选法至多有n^k种。

对于 $\ell > 0$，我们可以按如下方式选取元组：(1) 从 $\{1,2,\dots,2k\}$ 中选出 ℓ 个位置作为“高重数”指标；(2) 将每个位置与 $\{1,2,\dots,2k\}$ 中的另一个位置匹配；(3) 对剩余的 $2k-\ell$ 个位置（数量为偶数）选取一个完美匹配；(4) 从 $[n]$ 中选取$k-\ell/2$ 个不同的指标。因此，

$$\begin{aligned}\mathbb{E}[y^{2k}] &\leqslant n^{-k}\sum_{\ell=0,2,4,\dots}\gamma^{-\ell/2}(2k)^{2\ell}(2k-\ell-1)!!\cdot n^{k-\ell/2}\\ &\leqslant (2k-1)!!\sum_{\ell=0,2,4,\dots}\gamma^{-\ell/2}(2k)^{2\ell}\left(\frac{2}{k}\right)^{\ell/2}n^{-\ell/2}\\ &\leqslant (2k-1)!!\cdot\left(1+O\left(\frac{k^3}{\gamma n}\right)\right).\end{aligned}$$

至此，证明完成。 □

B Krawtchouk 多项式的上界

在本节中，我们给出 定理 3.12 的证明，该证明与 [Pol19、DILV24] 中给出的证明非常相似。

引理 （定理 3.12的重述）。固定常数 $\gamma\in(0,1)$。对于任意 $k\leqslant n/2$ 以及 $|y|\leqslant o(n^{1/6})$，有$|\mathrm{Kr}_k(y)|\leqslant O(k^{1/4})\cdot e^{y^2/4}$。

证明。首先注意，根据定义，

$$\prod_{i=1}^{n}(1+\chi_i(x)z)=\sum_{k=0}^{n}s_k(x)z^k. \tag{37}$$

令 $w:=\sum_{i=1}^{n}\frac{x_i+1}{2}\in\{0,1,\dots,n\}$，它等于 x 中 1 的个数，且服从 $\mathrm{Bin}(n,\gamma)$ 分布。记 $\tilde{\gamma}:=1-\gamma$，有

$$y=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{x_i-(2\gamma-1)}{2\sqrt{\gamma\tilde{\gamma}}}=\frac{1}{\sqrt{n\gamma\tilde{\gamma}}}(w-\gamma n)\implies w=\gamma n+\sqrt{n\gamma\tilde{\gamma}}y. \tag{38}$$

设 $K_k(w)$ 为表示 $s_k(x)$的 k 次单变量多项式。那么，我们有 $\chi_i(x)=\sqrt{\tilde{\gamma}/\gamma}$ 若 $x_i=1$且 $-\sqrt{\gamma/\tilde{\gamma}}$若 $x_i=-1$，因此 方程 (37) 等价于

$$f_w(z):=\left(1+\sqrt{\frac{\tilde{\gamma}}{\gamma}}z\right)^{w}\left(1-\sqrt{\frac{\gamma}{\tilde{\gamma}}}z\right)^{n-w}=\sum_{k=0}^{n}K_k(w)z^k.$$

这是通常克拉夫丘克多项式的生成函数。特别地，我们有 $K_k(w)=\frac{1}{k!}f_{(k)w}(0)$，因此根据柯西积分公式，

$$K_k(w)=\frac{1}{2\pi i}\oint_C f_w(z)z^{-k}\frac{dz}{z},$$

其中 C 是复平面上以$z=0$为中心的任意圆，方向为逆时针。我们将选择适当的半径 r，并对 $|f_w(z)z^{-k}|$ 在 $|z|=r$时的值作界。

为简便起见，记 $\alpha=w/n$ 和 $\beta=k/n$。首先，我们将 $f_w(z)$ 改写为

$$f_w(z)=\gamma^{-\alpha n/2}\tilde{\gamma}^{-(1-\alpha)n/2}\left(\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z\right)^{\alpha n}\left(\sqrt{\tilde{\gamma}}-\sqrt{\gamma}z\right)^{(1-\alpha)n}.$$

那么，

$$\log|f_w(z)z^{-k}|=\frac{n}{2}\Big(-\alpha\log\gamma-(1-\alpha)\log\tilde{\gamma}$$

$$+\alpha\log|\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z|^2+(1-\alpha)\log|\sqrt{\tilde{\gamma}}-\sqrt{\gamma}z|^2-\beta\log|z|^2\Big)$$
$$=\frac{n}{2}\left(D(\alpha\|\gamma)+\alpha\log\frac{|\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z|^2}{\alpha}+(1-\alpha)\log\frac{|\sqrt{\tilde{\gamma}}-\sqrt{\gamma}z|^2}{1-\alpha}-\beta\log|z|^2\right).$$

这里，$D(\alpha\|\alpha')\coloneqq\alpha\log(\frac{\alpha}{\alpha'})+(1-\alpha)\log(\frac{1-\alpha}{1-\alpha'})$ 是参数为 α 和 α'的两个伯努利分布之间的 KL 散度。接下来，我们利用 log 的凹性和詹森不等式：

$$\alpha\log\frac{|\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z|^2}{\alpha}+(1-\alpha)\log\frac{|\sqrt{\tilde{\gamma}}-\sqrt{\gamma}z|^2}{1-\alpha}\leqslant\log\Big(|\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z|^2+|\sqrt{\tilde{\gamma}}-\sqrt{\gamma}z|^2\Big)$$
$$=\log(1+|z|^2).$$

等式成立是因为 $|\sqrt{\gamma}+\sqrt{\tilde{\gamma}}z|^2=\gamma+\tilde{\gamma}z^2+\sqrt{\gamma\tilde{\gamma}}(z+z)$ </style}而 $|z|^2=\tilde{\gamma}+\gamma z^2-\sqrt{\gamma\tilde{\gamma}}(z+z)$，且 $\gamma+\tilde{\gamma}=1$。

因此，我们可以选择 C 的半径 r，使 $\log(1+r^2)-\beta\log r^2$ 最小。对其关于 r^2, 求导，可得 $r^2=\frac{\beta}{1-\beta}$ 为最优值。因此，对于 $|z|=r$，

$$\log|f_w(z)z^{-k}|\leqslant\frac{n}{2}\Big(D(\alpha\|\gamma)+\log(1+r^2)-\beta\log r^2\Big)$$
$$=\frac{n}{2}\left(D(\alpha\|\gamma)+\log\frac{1}{1-\beta}-\beta\log\frac{\beta}{1-\beta}\right)$$
$$=\frac{n}{2}(D(\alpha\|\gamma)+h(\beta))$$

其中 $h(\beta)=-\beta\log\beta-(1-\beta)\log(1-\beta)$。于是，

$$|K_k(w)|\leqslant e^{\frac{n}{2}(D(\alpha\|\gamma)+h(\beta))}\cdot\frac{1}{2\pi}\oint_C|\frac{dz}{z}|\leqslant e^{\frac{n}{2}(D(\alpha\|\gamma)+h(\beta))}.$$

回忆一下，$\alpha=w/n$ 且 $\beta=k/n$。由不等式 $\binom{n}{k}\geqslant\sqrt{n}\,8k(n-k)\,e^{nh(k/n)}\geqslant\Omega(1/\sqrt{k})e^{nh(k/n)}$，当 $k\leqslant n/2$ 时成立。另一方面，根据定理 3.5，$D(\gamma+\varepsilon\|\gamma)\leqslant\frac{\varepsilon^2}{2\gamma(1-\gamma)}+O(|\varepsilon|^3)$。由方程 (38)，我们有 $\alpha=\gamma+\sqrt{\gamma\tilde{\gamma}}\,n\,y$，因此

$$D(\alpha\|\gamma)\leqslant\frac{y^2}{2n}+O_\gamma\left(\left|\frac{y}{\sqrt{n}}\right|^3\right).$$

因此，

$$|\mathrm{Kr}_k(y)|=\binom{n}{k}^{-1/2}|K_k(w)|\leqslant O(k^{1/4})e^{y^2/4+O_\gamma(y^3n^{-1/2})}\leqslant O(k^{1/4})e^{y^2/4}$$

对于 $|y|\leqslant o(n^{1/6})$. □

C 特征函数上的界

Lemma (定理 5.7 的重述定理 5.7)。 设 p 是一个次数为 $\leqslant k$ 的单变量多项式。那么：

- 若方差$_{g\sim\mathcal{N}(0,1)}[p(g)]\leqslant 9^{-}k$，则有

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right|\leqslant 1-\frac{1}{4}\mathrm{Var}_{g\sim\mathcal{N}(0,1)}[p(g)].$$

- 若方差$_{g\sim\mathcal{N}(0,1)}[p(g)] = kck$ 对任意 $c \in \mathbb{R}$ 均成立，则有

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right| \leqslant 1 - \frac{1}{16k}k^{-|c|k}.$$

- 最后，如果对于充分大的 $C > 0$，有 $\mathrm{Var}_{g\sim\mathcal{N}(0,1)}[p(g)] \geqslant kCk$，那么

$$\left|\mathop{\mathbb{E}}_{g\sim\mathcal{N}(0,1)}[\exp(ip(g))]\right| \leqslant O\left(\mathrm{Var}_{g\sim\mathcal{N}(0,1)}[p(g)]^{-\frac{1}{4k}}\right).$$

证明。令 $\mu = \mathbb{E}_{g\sim\mathcal{N}(0,1)}[p(g)]$，以及 $\sigma^2 = \mathrm{Var}_{g\sim\mathcal{N}(0,1)}[p(g)]$。我们将改为界定 $|\mathbb{E}_g[e^{i(p(g)-\mu)}]|$。根据余弦的泰勒展开，

$$\mathop{\mathbb{E}}_{g}\mathrm{Re}(e^{i(p(g)-\mu)}) \in 1 - \frac{1}{2}\mathop{\mathbb{E}}_{g}[(p(g)-\mu)^2] \pm \frac{1}{24}\mathop{\mathbb{E}}_{g}[(p(g)-\mu)^4].$$

由于 p 是次数为 k 的多项式，根据高斯空间中的 [O'D14], 2 到 4 超收缩性，有$\mathbb{E}_g[(p(g) - \mu)^4] \leqslant 9^k \cdot \mathbb{E}_g[(p(g) - \mu)^2]^2 = 9^k \cdot \mathrm{Var}_g[p(g)]^2$。对于虚部，我们有

$$\mathop{\mathbb{E}}_{g}\mathrm{Im}(e^{i(p(g)-\mu)}) \in \mathop{\mathbb{E}}_{g}[(p(g)-\mu)] \pm \frac{1}{6}\mathop{\mathbb{E}}_{g}[|p(g)-\mu|^3].$$

第一项为 0，根据 2到3超收缩性 [O'D14]，有 $\mathbb{E}_g[|p(g)-\mu|^3] \leqslant 2^{3k/2} \cdot \mathrm{Var}_g[p(g)]^{3/2}$。因此，如果 $\sigma^2 = \mathrm{Var}_g[p(g)] \leqslant 9^{-k}$，那么

$$\left|\mathop{\mathbb{E}}_{g}[e^{i(p(g)-\mu)}]\right| \leqslant 1 - \frac{1}{2}\sigma^2 + \frac{1}{24}9^k\sigma^4 + \frac{1}{6}2^{3k/2}\sigma^3 \leqslant 1 - \frac{1}{4}\sigma^2.$$

对于第二种情况，将p 改变一个常数不会影响估计，因此我们可以假设 $\mathbb{E}[ip(g)]$ 为正实数，只需对该期望的实部给出上界即可。设$\boldsymbol{x}$ 和$\boldsymbol{x}'$为相互独立的高斯随机变量，并令$\boldsymbol{x}_\theta = $余弦$(\theta)\boldsymbol{x}$ +正弦$(\theta)\boldsymbol{x}'$。注意$\boldsymbol{x}_\theta$ 也是一个标准高斯分布。令$f_{\boldsymbol{x},\boldsymbol{x}'}(\theta) = p(\boldsymbol{x}_\theta)$，于是

$$\mathop{\mathbb{E}}_{g}[\exp(ip(g))] = \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{x}'}\mathop{\mathbb{E}}_{\theta\sim\mathrm{Unif}[0,2\pi]}\left[\exp\left(if_{\boldsymbol{x},\boldsymbol{x}'}(\theta)\right)\right].$$

我们将证明，对于$\boldsymbol{x}$、$\boldsymbol{x}'$ 和 θ 的选择，至少有$k^{-\lfloor c\rfloor k}/4k$ 的概率使得$f_{\boldsymbol{x},\boldsymbol{x}'}(\theta)$ 至少$\left(k^{-\lfloor c\rfloor k}/2/4\right)$远离 2π 的整数倍。这将立即得到所需的估计。

首先，根据 Carbery-Wright，我们有概率为 2/3 使得$p(\boldsymbol{x})$ 与$p(\boldsymbol{x}')$ 相差至少$k^{-\lfloor c\rfloor k/2}$。若此情况发生，那么根据连续性，存在某些 θ 使$f_{\boldsymbol{x},\boldsymbol{x}'}(\theta)$ 与最近的整数倍 2π相差超过$k^{-\lfloor c\rfloor k/2}/2$。此外，我们还有

$$\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{x}'}[\mathop{\mathbb{E}}_{\theta}[(f'_{\boldsymbol{x},\boldsymbol{x}'}(\theta))^2]] = \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{x}',\theta}[(p'(\boldsymbol{x}_\theta))^2\boldsymbol{x}^2_{\theta+\pi/2}].$$

由于$\boldsymbol{x}_\theta$ 和$\boldsymbol{x}^{\theta+\pi/2}$ 是独立高斯随机变量，因此有

$$\mathbb{E}[(p'(\boldsymbol{x}_\theta))^2] \leqslant k \cdot \mathrm{Var}_{g\sim\mathcal{N}(0,1)}[p(g)].$$

因此，根据马尔可夫不等式和并集界，在 x 和 x′ 上，至少有 1/2 的概率存在 θ，使得 $f_{\boldsymbol{x},\boldsymbol{x}'}(\theta)$ 与某个 2π 的整数倍相差超过 $k^{ck/2}/2$，并且 $\mathbb{E}_\theta[(f'(x,x'(\theta))^2] < k^{ck+1}$。如果这两个条件都成立，那么在 θ 中必定存在一个长度至少为 $12k\,k^{-\lfloor c\rfloor k/2}$ 的区间，使得 $f_{\boldsymbol{x},\boldsymbol{x}'}(\theta)$ 与其最近的 2π 倍数至少相差 $(k^{-\lfloor c\rfloor k/2}/4)$。这就完成了第二种情况的证明。

对于第三个结果，我们注意到 $\mathbb{E}_{g\sim\mathcal{N}(0,1)}[\exp（ip（g））]$ 与以下内容成正比：

$$\int_{-\infty}^{\infty} e^{ip(x)}e^{-x^2/2}dx.$$

通过分部积分，我们得到：

$$\begin{aligned}\int e^{ip(x)}e^{-x^2/2}\,dx &= \int (ip'(x))e^{ip(x)}\cdot\frac{1}{ip'(x)}e^{-x^2/2}\,dx\\ &= \frac{e^{ip(x)}e^{-x^2/2}}{ip'(x)} - i\int e^{ip(x)}e^{-x^2/2}\left(\frac{x}{p'(x)}+\frac{p''(x)}{(p'(x))^2}\right)dx.\end{aligned}$$

设 $\gamma>0$ 为稍后确定的参数。根据定理 3.9，我们有 $|p'|_2^2 \geqslant$ 方差$_{g\sim\mathcal{N}(0,1)}$(p(g))，因此根据 Carbery-Wright，除概率为 $\gamma/3$ 的情况外，我们有 $|p'(g)| \geqslant \Omega(\gamma/k)^k|p|_2$。此外，根据定理 3.9，$|p''|_2 \leqslant k|p'|_2$。因此，根据高斯超收缩性，我们以 $1-\gamma/3$ 的概率有 $|p''(g)| \leqslant O(\log(1/\gamma))^k|p'|_2$。最后，除概率为 $\gamma/3$ 的情况外，我们有 $|g| \leqslant \log(1/\gamma)$。令 E 表示以下事件：(1) $|p'(x)| \geqslant \Omega(\gamma/k)^k|p'|_2$；(2) $|p''(x)| \leqslant O(\log(1/\gamma)^k)|p'|_2$；以及 (3) $|x| \leqslant \log(1/\gamma)$。根据上述论证和并集界，我们有 $\mathbb{P}[g \notin E] \leqslant \gamma$。于是，我们得到：

$$\begin{aligned}\left|\int_{-\infty}^{\infty} e^{ip(x)}e^{-x^2/2}dx\right| &= \left|\int_E e^{ip(x)}e^{-x^2/2}dx + \int_{E^c} e^{ip(x)}e^{-x^2/2}dx\right|\\ &\leqslant \left|\int_E e^{ip(x)}e^{-x^2/2}dx\right| + O(\gamma)\,. \end{aligned} \tag{39}$$

我们现在注意到，期望 是$O(k)$ 个区间的并，因为对于任意次数至多为k 的多项式 q 和任意阈值$c \in \mathbb{R}$，使 $q(x) \leqslant c$ 成立的x 构成至多k 个区间的并。因此，令期望 $=\cup_{\ell=1}^{O(k)} I_\ell$，其中$I_\ell=[a_\ell,b_\ell]$是一些互不相交的区间。那么，根据分部积分公式，我们有

$$\left|\int_E e^{ip(x)}e^{-x^2/2}dx\right| \leqslant \sum_{\ell=1}^{O(k)}\left(\left|\frac{e^{ip(b_\ell)}e^{-b_\ell^2/2}}{p'(b_\ell)} - \frac{e^{ip(a_\ell)}e^{-a_\ell^2/2}}{p'(a_\ell)}\right| + \int_{a_\ell}^{b_\ell}\left|e^{-x^2/2}\left(\frac{x}{p'(x)}+\frac{p''(x)}{(p'(x))^2}\right)\right|dx\right)$$

根据期望的定义，对于任意$y \in$ 期望，

$$\begin{aligned}|p'(y)| &\geqslant \Omega\left(\frac{\log(1/\gamma)}{k}\right)^k |p'|_2\\ &\geqslant \Omega\left(\frac{\log(1/\gamma)}{k}\right)^k \sqrt{\mathrm{Var}(p(g))}\\ &\geqslant \Omega(\log 1/\gamma)^k\,\mathrm{Var}(p(g))^{\Omega(1)}\,,\end{aligned}$$

其中，最后一个界由我们对 $\mathrm{Var}(p(g))$的假设得出。因此我们有

$$\sum_{\ell=1}^{O(k)}\left|\frac{e^{ip(b_\ell)}e^{-b_\ell^2/2}}{p'(b_\ell)} - \frac{e^{ip(a_\ell)}e^{-a_\ell^2/2}}{p'(a_\ell)}\right| \leqslant O(k)\cdot(\log 1/\gamma)^{-\Omega(k)}\,\mathrm{Var}(p(g))^{-\Omega(1)}\,,$$

只要 $\gamma<1/2$ 成立，这就显著强于我们的需要。

为了界定剩余项，我们注意到，对于所有$y \in$ 期望以及所有 $\gamma \leqslant 1$，都有

$$\frac{x}{p'(x)}+\frac{p''(x)}{(p'(x))^2} = O\left(\frac{(\log 1/\gamma)^k}{(\gamma/k)^{2k}\mathrm{Var}(p)^{1/2}}\right)$$

因此，若令 γ 为 $\mathrm{Var}(p(g))^{-1/(4k)}$ 的一个较小倍数，则该量至多为 $\mathrm{Var}(p(g))^{-\Omega(1)}$。因此，

$$\sum_{\ell=1}^{O(k)}\int_{a_\ell}^{b_\ell}\left|e^{-x^2/2}\left(\frac{x}{p'(x)}+\frac{p''(x)}{(p'(x))^2}\right)\right|dx \leqslant \mathrm{Var}(p(g))^{-\Omega(1)}\cdot\int_{-\infty}^{\infty}e^{-x^2/2}dx = \mathrm{Var}(p(g))^{-\Omega(1)}\,.$$

将其与（39）以及我们的选择 γ 结合，即可得到所需结论。 □

D 子图多项式的矩

令 ϑ 为一个常 规模为 nt 的连通图。回顾第6节中的记号，其中对于一个$n \times n$ 矩阵 x,

$$\chi_\vartheta(M) := \frac{1}{\sqrt{|\mathcal{L}_\vartheta|}} \sum_{\pi \in \mathcal{L}_\vartheta} \prod_{ab \in \vartheta} M_{\pi(a),\pi(b)}$$

且 $\mathcal{L}_\vartheta$ 是从 $V(\vartheta)$ 到 $[n]$ 的所有单射的集合。

We w将证明随机矩阵的子图多项式具有类似高斯分布的矩 ts.

引理 (重述 定理 6.4). 存在一个充分小的常数 $\alpha = \alpha(\vartheta)$ ，使得对于任意 $0 \leqslant q \leqslant n^\alpha$:

$$\mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathsf{N}} \chi_\vartheta(\boldsymbol{M})^q \leqslant \sqrt{q}^q \cdot \left(1 + \frac{1}{\mathrm{poly}(n)}\right).$$

证明。 我们先展开要计算的表达式。将省略 $\mathcal{L}_\vartheta$ 的下标，并记为 $\mathcal{L}$。

$$\mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathsf{N}} \chi_\vartheta(\boldsymbol{M})^q = \frac{1}{|\mathcal{L}|^{q/2}} \sum_{\pi_1,\ldots,\pi_q \in \mathcal{L}} \mathop{\mathbb{E}}_{\boldsymbol{M} \sim \mathsf{N}} \prod_{i=1}^{q} \prod_{ab \in E(\vartheta)} \boldsymbol{M}_{\pi_i(a),\pi_i(b)} \cdot \tag{40}$$

除非多重集并$H = \pi_1(\vartheta) \sqcup \cdots \sqcup \pi_q(\vartheta)$ 中的每条边都是均匀覆盖边，否则该求和项为零。

对于单个选择 π_1、.....、π_q，我们可以关联:

- $r = r(H)$ 定义为H中的连通分量数,
- 去掉标签后H 的连通分量$H1$、.....、H_r ,
- 每个连通分量 β_1、.....、β_r中 ϑ 的副本数,
- 集合$S1$、.....、$S_r \subseteq [q]$ ，其中 $|S_i| = \beta_i$ 由所有满足 $\pi_j(\vartheta)$ 属于H_i的 π_j 组成,
- 一个映射ψ_i 使得$\psi_i(j)$ 对于$j \in S_i$ 等于由以下方式得到的H_i 的子图：由$\pi_j(\vartheta$).

对于固定选择的 $\mathrm{H}_1,\ldots,\mathrm{H}_r$、$\mathrm{S}_1,\ldots,\mathrm{S}_r$ 和 ψ_1、. . .、ψ_r，求和式公式（40）中至多有 $\mathrm{n}^{|\mathrm{V(H)}|}$ 项与这些选择相对应。此外，每一项的值均为非负数，并且只取决于这些选择。将这些选择简写为 $\vec{\mathrm{H}}$、$\vec{\mathrm{S}}$、$\vec{\psi}$，将每一项的相应值记为 $\mathrm{wt}\left(\vec{\mathrm{H}}、\vec{\mathrm{S}}、\vec{\psi}\right)$，并利用 $|\mathcal{L}| = \mathrm{n}^{v(\vartheta)}$，我们可以将公式（40）界定为:

$$(40) \leqslant \sum_{\vec{H},\vec{S},\vec{\psi}} \mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right) \cdot n^{V(H) - qv(\vartheta)/2}.$$

Bef在继续之前，我们先介绍一些记号。我们用 $\models$ 表示“与.....一致”。

$$\sum_{\vec{H},\vec{S},\vec{\psi}} \mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right) \cdot n^{V(H) - qv(\vartheta)/2} = \sum_{r \geqslant 1} \sum_{\vec{\beta} \models r} \sum_{\vec{S} \models \vec{\beta}} \sum_{(\vec{H},\vec{\psi}) \models \vec{S}} \mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right) \cdot n^{\sum_i (V(H_i) - \beta_i v(\vartheta)/2)} \tag{41}$$

我们首先计算 $(\vec{H}、\vec{\psi}) \models \vec{S}$ 的方式数量。

特别地，对于每个$i \in [r],$，我们给出$(H_i, \psi_i) \models S_i$的可能方式数的一个上界。假设 $\beta_i = 2$，那么(H_i, ψ_i) 恰好只有一种选择会产生非零权重。此后，我们假设 $\beta_i \geqslant 3$。不失一般性，设 $S_i = \{1,\ldots..,\ \beta_i\}$，并且对于所有$j \in [t],$，图$\cup^j_{s=1} \psi_i(s)$ 都是连通的。给定$(\vec{H},\ \vec{\psi}) \models \vec{S}$，我们描述一种编码，然后给出有效编码数量的一个上界。

编码方案。 在我们的编码方案中，我们按顺序处理 $\psi_i(j)$，顺序为 $j = 1,\ldots,\beta_i$，并描述每一步的过程。

- 如果顶点 v 不出现在 $\bigcup_{s=1}^{j-1}\psi_i(s)$ 中，则称其为新出现的顶点。
- 若顶点 v 存在于 $\psi_i(j)$ 中，但与 $\bigcup_{s=1}^{j}\psi_i(s)$中的某条单边关联，则称其为边界顶点。
- 若顶点 v 存在于 $\psi_i(j)$ 中，且它在 $\bigcup_{s=1}^{j}\psi_i(s)$ 中的所有关联边均为双重覆盖，则称其为陈旧顶点。
- 对于 ϑ 中的每个顶点，我们标记它在 $\psi_i(j)$ 中是新出现、边界还是陈旧的；共有 $3^{v(\vartheta)}$ 种选择。
- 对于边界顶点或陈旧顶点，我们在 $\bigcup_{s=1}^{j-1}\psi_i(s)$中记下该顶点的名称，最多有 $\beta_i v(\vartheta)$ 种选择。

用 f_i、b_i、s_i 分别表示新出现、边界和陈旧顶点的数量，可知编码数的上界为：

$$\prod_{i\in[r]:\beta_i\geqslant 3} 3^{\beta_i}\cdot(\beta_i v(\vartheta))^{b_i+s_i}.$$

代入此界可得：

$$(41)\leqslant \sum_{r\geqslant 1}\sum_{\vec{\beta}\models r}\sum_{\vec{S}\models\vec{\beta}}\mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right)\cdot\prod_{i\in[r]:\beta_i\geqslant 3}3^{\beta_i}\cdot(\beta_i v(\vartheta))^{b_i+s_i}\cdot n^{V(H_i)-\beta_i v(\vartheta)/2}.$$

接下来，我们对 $\mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right)$ 进行界定。在 H_i 中，任意一条边的重数至多为 β_i，因此我们有：

$$\mathrm{wt}\left(\vec{H},\vec{S},\vec{\psi}\right)\leqslant\prod_{i\in[r]:\beta_i\geqslant 3}O\left(\sqrt{\beta_i}\right)^{\beta_i e(\vartheta)}=\prod_{i\in[r]:\beta_i\geqslant 3}O\left(\sqrt{\beta_i}^{e(\vartheta)}\right)^{\beta_i}$$

结合 $|V(H_i)|=f_i$ 和 $b_i+s_i\leqslant\beta_i v(\vartheta)$ 这一事实，我们可以得出：

$$(41)\leqslant\sum_{r\geqslant 1}\sum_{\vec{\beta}\models r}\sum_{\vec{S}\models\vec{\beta}}\prod_{i\in[r]:\beta_i\geqslant 3}O\left(\beta_i^{e(\vartheta)/2+v(\vartheta)}v(\vartheta)^{v(\vartheta)}\right)^{\beta_i}\cdot n^{f_i-\beta_i v(\vartheta)/2}$$

注意，$\mathrm{v}(\vartheta)^{v(\vartheta)}$ 是一个常数，因此我们从此将其吸收到 $\mathrm{O}(\cdot)$ 记号中。接下来，我们将证明 $\mathrm{f}_i-\beta_i\mathrm{v}(\vartheta)/2$ 是显著的负数。首先，注意 $\mathrm{f}_i+\mathrm{b}_i+\mathrm{s}_i=\beta_i\mathrm{v}(\vartheta)$。由于 $\mathrm{s}_i\geqslant\mathrm{f}_i$，我们有 $2\,\mathrm{f}_i+\mathrm{b}_i\leqslant\beta_i\mathrm{v}(\vartheta)$。现在我们将证明 b_i 的一个下界。第一步对 b_i 的边界顶点数贡献为零。由于 $\psi_i(\mathrm{j})$ 与 $\bigcup^{j-1}\mathrm{s}=1\psi_i(\mathrm{s})$ 连通，因此第一步之后的每一步必定要么至少贡献一个边界顶点，要么贡献 $\mathrm{v}(\vartheta)$ 个陈旧顶点。引入的顶点中至多一半是陈旧的，因此引入 $\mathrm{v}(\vartheta)$ 个陈旧顶点的情况至多有 $\beta_i/2$ 次。因此，边界顶点数 b_i 至少为 $\beta_i/2-1$。于是，我们得到不等式

$$2f_i\leqslant\beta_i\left(v(\vartheta)-\frac{1}{2}\right)+1,$$

这意味着

$$f_i-\frac{\beta_i v(\vartheta)}{2}\leqslant-\frac{\beta_i}{4}+\frac{1}{2}=\frac{2-\beta_i}{4}\leqslant-\frac{\beta_i}{12},$$

其中，最终不等式用到了 $\beta_i\geqslant 3$。

因此，我们得到：

$$(41)\leqslant\sum_{r\geqslant 1}\sum_{\vec{\beta}\models r}\sum_{\vec{S}\models\vec{\beta}}\prod_{i\in[r]:\beta_i\geqslant 3}O\left(\frac{\beta_i^{e(\vartheta)+v(\vartheta)}}{n^{1/12}}\right)^{\beta_i}$$

固定选择 $\vec{\beta}$后，很容易看出，$\vec{S} \models \vec{\beta}$ 的可选项数至多为：

$$q^{q/2} \cdot \prod_{i\in[r]:\beta_i\geqslant 3} q^{\beta_i}.$$

因此，

$$(41) \leqslant q^{q/2} \sum_{r\geqslant 1} \sum_{\vec{\beta}\models r} \prod_{i\in[r]:\beta_i\geqslant 3} O\left(\frac{q\beta_i^{e(\vartheta)+v(\vartheta)}}{n^{1/12}}\right)^{\beta_i}$$

我们可以把 $\alpha(\vartheta)$ 选得足够小，这样即使使用 q，$\beta \leqslant n^{\alpha(\vartheta)}$ 后，仍有：

$$\begin{aligned}
(41) &\leqslant q^{q/2} \sum_{r\geqslant 1} \sum_{\beta\models r} \prod_{i\in[r]:\beta_i\geqslant 3} O\left(n^{-1/24}\right)^{\beta_i} \\
&\leqslant q^{q/2} \sum_{\beta_1,\ldots,\beta_q\geqslant 0} \prod_{i=1}^{q} O\left(n^{-1/24}\right)^{\beta_i} \\
&\leqslant q^{q/2} \left(\sum_{\beta\geqslant 0} O\left(n^{-1/24}\right)^{\beta}\right)^{q} \\
&\leqslant q^{q/2} \left(1 + O\left(n^{-1/24}\right)\right)^{q} \\
&\leqslant q^{q/2} \cdot \left(1 + \frac{1}{\mathrm{poly}(n)}\right).
\end{aligned}$$

□