arXiv:2603.22192v1 [math.ST] 23 Mar 2026

基于最小均方误差不连续性的稳定算法估计下界

Xifan Yu[†] 和 Ilias Zadik[◦]

摘要。平均情形复杂性领域的近期研究将稳定（对噪声稳定）算法确定为一个核心类别。具体而言，在平均情形优化中，人们猜想这一类别能够刻画许多问题的多项式时间计算能力。通过在低于猜想中的多项式时间阈值处证明重叠间隙性质（OGP）的各种相变，这一观点得到了支持。然而，Schramm 和 Li（2025）最近对此提出了挑战：他们证明，随机图中的最短路径问题具有 OGP，因此所有稳定算法都会失效，尽管该问题可以在多项式时间内求解。这一反例尤其令人好奇，因为它似乎与其他经典的“无噪声”反例截然不同，例如求解随机线性方程组。

相比之下，稳定方法在统计估计中的能力仍不明确。一个核心难点在于，缺乏一种 OGP 型现象的定义，能够统一排除所有稳定方法。相反，现有下界研究大多聚焦于相关的低次多项式类别，并局限于高斯加性模型等受限模型，这反映出控制低次估计量的最小均方误差（MMSE）在技术上难度很高。

本文证明，对于所有统计估计问题，一种自然的最小均方误差不稳定性（不连续性）条件都意味着稳定算法无法解决该问题，可视为估计任务中的重叠间隙性质。利用这一判据，我们针对以下最小均方误差不稳定任务，证明了稳定算法与多项式时间算法之间存在分离：(i) 植入最短路径，其中 Dijkstra 算法能够成功；(ii) 随机奇偶校验码，其中高斯消元能够成功；以及 (iii) 高斯子集和，其中基于格的方法能够成功。对于这三类任务，我们进一步证明，所有低次多项式都是稳定的，由此得到与低次方法之间的分离，并提出一种界定低阶最小均方误差的新方法。具体而言，我们的技术表明，最小均方误差不稳定性是最短路径、无噪声奇偶校验码和高斯子集和的共同特征。

最后，我们强调，我们的工作为长期以来的物理学观点提供了严谨的算法理论基础：一阶相变——在这一设定中对应于 MMSE 不稳定性——会对高效算法类别施加基本极限。

日期：2026年3月24日。◦统计与数据科学系，耶鲁大学；[†]计算机科学系，耶鲁大学。电子邮箱：{xifan.yu, ilias.zadik}@yale.edu。

目录

1. 引言 3 1.1. 贡献 4 2. 主要结果：最小均方误差不连续性意味着稳定算法失效 7 3. 应用 8 3.1. 植入最短路径模型 8 3.2. 随机线性码 11 3.3. 高斯子集和（GSS）的结果 12 4. 入门：证明所需的预备知识 14 5. 关键结果的证明：MMSE 不稳定性意味着稳定算法失效 14 6. 植入最短路径的证明 15 6.1. 含噪 MMSE 下界：定理 3.7 的证明 15 6.2. 对称低次多项式的稳定性：定理 3.9 的证明 22 7. 随机线性码的证明 34 7.1. 含噪 MMSE：定理 3.16 的证明 34 7.2. 低次多项式的稳定性：定理 3.19 的证明 38 8. 高斯子集和的证明 40 8.1. MMSE 不稳定性的延后证明 40 8.2. 低次多项式的稳定性 41 8.3. 定理 8.1 的延后证明 51 参考文献 54 附录 A. 稀疏张量 PCA 中稳定算法的失效 56 A.1. 证明 57

1. 引言

过去几十年来，理解平均情形问题的计算复杂性一直是研究重点，这一研究方向有别于经典的最坏情形复杂性理论。该领域一种占主导地位的方法，是针对给定的平均情形问题分析受限但强大的算法族，并将这些算法失效的临界点视为计算困难性开始显现的标志。这种方法在两大领域中最具影响力：零分布下的平均情形优化问题，以及植入分布下的统计推断问题。

对于平均情形优化，研究最广泛的算法类别是（噪声）稳定算法。这一类别通常涵盖低次多项式、低深度电路，以及在有界时间内运行的局部搜索过程（例如参见 [GJW24]）。尽管在庞大的$\mathcal{P}$类别中，这类算法看似受到诸多限制 p，近年来却出现了一个引人注目的规律：对于许多典型随机问题，稳定算法的表现与已知最佳多项式时间方法相当。显著的例子包括 [GJW24、HS25], 随机图中的最大独立集 [Wei22], ，以及随机 k-SAT [BH22]中的优化。值得指出的是，从技术角度看，支撑这些紧致稳定算法下界的关键技术工具，是一种称为重叠间隙性质（OGP）的非连通性景观性质的变体，该性质最早由 [GS14]提出（综述另见 [Gam21] ）。这些成果促使人们猜想：稳定算法的失效标志着问题真正的计算困难性阈值。

然而，至少在普遍认知层面，人们早已明白，这样的猜想不可能在完全一般的情形下成立。这类反例的典型代表是 XORSAT 的变体，其任务是求解随机线性方程组。只要存在解，高斯消元法就能成功；但人们普遍认为，在某些可满足区域中，稳定算法会失效。[1] 然而，这类反例常被视为问题无噪声结构所造成的人为现象。值得注意的是，近期的研究 [LS24]表明，即使是一个看似带有噪声的问题，也可能违背稳定算法最优性范式。他们证明，在 Erd˝os–R´enyi 随机图上，最短路径问题虽然可由 Dijkstra 算法在多项式时间内求解，却具有重叠间隙性质（OGP），这意味着稳定算法必然会失败。这一结果大幅模糊了这样一类问题的边界：对于这些问题，我们应当猜想稳定算法在随机任务上是最优的。理解这一问题正是本研究的主要动机之一。

相比之下，在统计推断领域，尤其是检测和估计任务中，对稳定算法这一统一类别的研究仍基本处于空白状态。主要原因之一是缺乏一种 OGP 类型的现象，能够针对所有稳定

[1]尽管这种看法广为流传，但据我们所知，文献中尚无形式证明。相关结果揭示了 k-XORSAT 中的碎裂或聚类现象，这些现象往往预示着稳定方法会失效，但本身并不能证明这一点 [ACOGM20，IKKM12，AM15]。

方法[2]给出一般性下界。相反，研究主要聚焦于其他受限类别，尤其是低次多项式（LDP）[3], 例如，参见 [Wei25]的最新综述。

对于检测问题，LDP 的能力已得到广泛研究，并由此提出了著名的低阶猜想 [Hop18], , 该猜想认为，对于足够“良好”的检测任务，低次多项式的能力与所有多项式时间算法相当。学界仍在讨论“良好”的确切含义，通常它包含对称性假设（关于完全对称情形下的证明，见 [HKK+26]；关于较弱对称性下更强变体的反例，见 [BHJK25]），同时也要求存在足够的噪声。或许值得指出的是，在没有噪声的情况下，低阶最优性的反例主要有两类：（a）植入随机线性方程，其中高斯消元可以成功，这与上述优化文献中的情形类似；（b）高斯子集和问题的某些变体，其中基于格的方法可以在多项式时间内成功 [Fri86, ZG18]。尽管近期研究对此提出了质疑，低阶猜想在许多典型检测问题中仍广受认可。

相比之下，低次多项式算法（LDP）在估计中的能力虽然预计与其在检测中的能力相似，但在数学上远未得到充分理解。从技术角度来看，控制低次估计量的最小均方误差（MMSE），即低阶最小均方误差，远比分析检测中的低阶似然比困难。尽管近期有一项工作 [SW22]在高斯加性模型和某些伯努利噪声设定中，发展了用于界定低阶最小均方误差的重要技术，但应用这些方法依赖于精细的论证，而且成功应用的范围仍局限于少数几类模型。因此，尽管文献中的研究者确实基于低次 MMSE 下界预测了困难性（例如，参见 [EGV25]），迄今尚未正式提出针对估计的低阶猜想。正因如此，我们认为，阐明低次多项式相较于其他多项式时间算法在估计任务中的能力，是一个重要的研究方向，也是本工作的主要动机之一。

1.1. 贡献。 在本文中，我们关注统计估计 问题，其中实例$(x,\ y) \sim \mathrm{P}_{XY}$ 被抽取，目标是设计一种算法 $\mathcal{A}$，使其在给定观测值 y时，能够以均方误差（MSE）近似信号 x。具体而言，我们旨在最小化

$$\mathbb{E}_{(x,y)\sim \mathbb{P}_{XY}}\left[\|\mathcal{A}(y)-x\|_2^2\right],$$

理想情况下达到最小均方误差 （MMSE），

$$\mathrm{MMSE} \ = \ \inf_{\mathcal{A}:\mathbb{R}^N\to\mathbb{R}^n} \mathbb{E}_{(x,y)\sim \mathbb{P}_{XY}}\left[\|\mathcal{A}(y)-x\|_2^2\right].$$

令 $T_\rho : \mathrm{R}^N \to \mathrm{R}^N, \rho \in [0,1]$ 表示作用于观测值 y 的噪声算子。以下这类算法对本文至关重要。

[2]对于植入问题，存在一种有效的 OGP 概念，但它只能给出针对马尔可夫链蒙特卡罗/局部搜索的紧下界（例如参见 [GZ22、GZ24]）。[3]需要指出的是，对于植入模型，LDP 是否实际上属于稳定方法这一问题相当棘手。

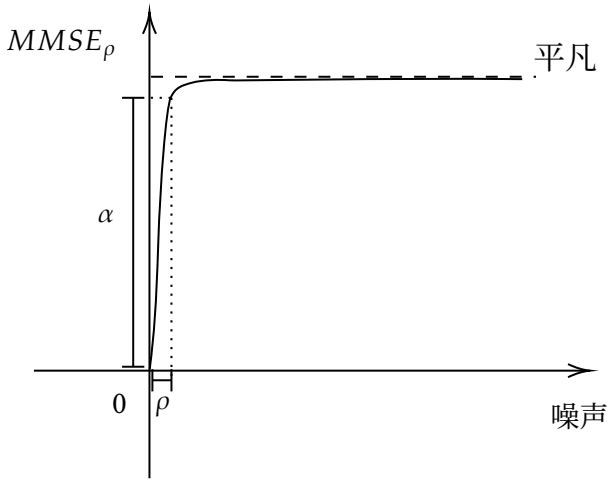


图1. 小噪声下尖锐的最小均方误差跃迁的图示，这会导致稳定算法失效，见定理1.2。

定义 1.1 （稳定算法）。算法 $\mathcal{A}: \mathrm{R}^N \to \mathrm{R}^n$ 若满足以下条件，则称为 (ρ, η)-稳定：

$$\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y) - \mathcal{A}(T_\rho(y))\|_2^2\right] \leqslant \eta \cdot \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y)\|_2^2\right].$$

最小均方误差不稳定性意味着稳定算法下界。我们的首要贡献是揭示一个普遍原理：

最小均方误差的噪声不稳定性为针对所有稳定算法的下界提供了依据。

为进一步描述这一点，定义含噪 *MMSE*，

$$\mathrm{MMSE}_\rho := \inf_{\mathcal{A}:\mathbb{R}^N\to\mathbb{R}^n} \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y)) - x\|_2^2\right], \qquad \rho \in [0,1]. \tag{1}$$

定理 1.2 （非正式表述：MMSE 跳跃意味着稳定算法失效）。 假设对于某些$\rho \in [0、 1]$ 和 $\alpha > 0$，*MMSE* 在 $(\rho、 \alpha)$意义下是不稳定的，

$$\mathrm{MMSE}_\rho - \mathrm{MMSE} \geqslant \alpha \cdot \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right].$$

那么，任何 $(\rho、 o(\alpha^2))$-稳定算法 $\mathcal{A}$ 都是 $\Omega(\alpha)$-次优的：

$$\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y) - x\|_2^2\right] \geqslant \mathrm{MMSE} + \Omega(\alpha) \cdot \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right].$$

这将估计中的稳定算法下界问题归结为理解小噪声下 MMSE 的阈值行为 ρ：MMSE 转变越陡峭（即， α越大），稳定方法因此遭遇的失效就越严重（见图 1）。这自然引出两个问题。首先，在小噪声下，MMSE 是否会出现真正的“跳跃”？ 其次，当这种跳跃发生时，它是否足够陡峭，以至于能够排除具有实际意义的算法类别？

为使上述讨论形式化，考虑归一化MMSE（NMMSE），

$$\mathrm{NMMSE}_\rho = \frac{\mathrm{MMSE}_\rho}{\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right]}, \qquad \rho \in [0,1].$$

尽管许多典型模型——例如采用独立同分布高斯先验的尖峰矩阵模型——都呈现连续的极限NMMSE曲线（例如参见 [LM17]），但已有充分记录表明，渐近NMMSE也可能是不连续的。这种现象在统计物理中称为一阶

相变，例如参见 [DMK+16]中的讨论。一个极端但广为观察到的例子是全有或全无(AoN)现象，最早见于 [RXZ19], ，其中NMMSE会从 $\rho = 0$处的近乎零，随着注入任意小的常数噪声而骤然跃升至近乎一 $\rho > 0$。如今已知许多推断模型都表现出这种行为，包括高斯加性模型 [NWZ20], 、广义线性模型 [LMB22], 、植入子图模型 [WXS22、MNWS+23]，以及一大类离散信道，例如伯努利群组测试[NWZ23]。

乍看之下，人们因此可能会得出结论：任何表现出 AoN 的模型都必然意味着所有稳定算法均会失效。虽然从原则上说这在一定程度上是正确的，但必须谨慎解读这一推论。定理1.2表明，AoN 仅对满足以下条件的算法给出 $\Omega(1)$-次优性：例如，这些算法是 $(\rho, 0。9)$-稳定的，且噪声水平为常数 ρ。然而，许多强大的稳定方法——尤其是低次多项式——通常并不满足如此强的稳定性。事实上，在典范乘积测度下（例如高斯分布或伯努利分布），次数为 D 的多项式通常是 $(\rho, O(D\rho))$-稳定的；这一性质通常可由标准的噪声算子论证得到。尽管将此类稳定性保证推广到零假设模型之外极具挑战性，但我们在本文中证明，对于若干我们所关注的植入分布，也存在类似的界限。因此，AoN 单独通常只能排除常数次数的多项式，而要排除更高次数的方法，则需要对最小均方误差转变窗口的宽度进行更精细的分析。

在应用章节中，我们开发了能够精确控制这些转变窗口的工具，从而能够证明在相关噪声尺度下低次多项式的稳定性结果，并推导出所需的分离。

最后，我们指出，将最小均方误差不连续性视为具有算法意义这一观点，在文献中有着深厚的根基。在统计物理中，一阶相变通常与计算难度相关联 [DMK+16，ZK16]。在数学研究中，我们只知道一个相关结果， [GMZ23] 该结果表明，随机布尔问题中的尖锐概率阈值意味着常数深度电路失效；然而，这种联系依赖于经典的阈值组合概念和技术，例如 H˚astad 的切换引理，而且预计无法推广到基于最小均方误差的现象，也无法推广到伯努利测度之外。我们特别指出，我们用于证明稳定算法失效的方法适用于所有参数估计任务。

应用与分离。 正如引言中所述，稳定算法或低次算法与多项式时间方法之间已知的主要分离，来自少数几个典型问题设定[4]：随机线性方程组（在优化和检测中）、最短路径问题（在优化中），以及高斯子集和问题（在检测中）。

我们证明，我们的最小均方误差不稳定性框架一致适用于估计这三种设定中的所有变体，从而得到稳定算法与多项式时间方法之间的分离——并且，作为推论，我们还证明了低次多项式与多项式时间方法之间的分离（见图2）。值得注意的是，我们的结果揭示了这些此前彼此迥异的例子背后共同的结构特征：这三个问题都表现出最小均方误差不稳定性，而这进而导致稳定方法无法达到最优。这一视角阐明了最短路径问题与“无噪声”线性系统及子集和设定的共同之处，并有望帮助学界最终确定一类恰当的任务

[4]我们将这个非常有趣的近期反例排除在此列表之外 [BHJK25]因为它将拟多项式时间方法与低次数多项式区分开来

适当的低次多项式／稳定算法最优性猜想应当适用。现在我们进一步详述具体的分离结果。

植入最短路径。考虑 Erd˝os–R´enyi 图上的植入最短路径问题$G(n,\ q\ =\ \Theta(\log\ n/n))$, 其中长度为 $L\ =\ C\log\ n/\log\log\ n$ 的路径被均匀随机地植入。对于 $C\ <\ 1$，以高概率，植入路径是唯一的最短路径，且可通过 Dijkstra 算法恢复。我们证明，最小均方误差是 $(O(1/L),\Omega(1))$-不稳定的，这意味着所有 $(O(1/L),o(1))$-稳定算法都会失败。此外，我们还证明所有次数为 $D\ =o(L)$ 的多项式都是稳定的，从而得到一个低阶最小均方误差下界，并与多项式时间方法形成分离。

随机线性码。最后，考虑随机线性码问题，其中观测到$y\ =\ Ax$，其 $A\ \in\ \{0,\ \ 1\}^{m\times n}$ 具有独立同分布的均匀分布条目，且 $x\ \in\ \{0,\ \ 1\}^{n}$ 均匀分布。对于 $m\ \geqslant n+\omega(1)$，高斯消元以高概率在多项式时间内恢复 x。我们证明，最小均方误差是 $(O((m-n)/n),\Omega(1))$-不稳定的。因此，所有 $(O((m-n)/n),o(1))$-稳定算法都会失败。我们进一步证明，所有次数为 $D=O(n/(m-n))$ 的多项式都是稳定的，因此是次优的，从而与多项式时间类形成另一种分离。

高斯子集和。设 $X_i\sim\mathcal{N}(0,1)$ 为独立同分布的 $i=1,\dots,N$。目标是根据观测值 $\sum_{i\in S}X_i$，恢复一个大小为 $k\ =\ N^{\alpha+o(1)}$、 $\alpha\ \in\ (0、\ 1)$ 的隐藏子集 $S\subset[N]$。一种基于格的方法使用 Lenstra–Lenstra–Lov´asz 算法，以高概率成功。利用 AoN 结果，我们证明最小均方误差具有 $(\exp(-\Theta(k\log(N/k))),\Omega(1))$-不稳定性。根据定理 1.2，所有 $(\exp(-\Theta(k\log(N/k))),o(1))$-稳定算法都会失败。我们进一步证明，在此噪声水平下，所有次数为 $D\ =o(\min\{k^{1/4},(n/k)^{1/5})$ 的多项式都是稳定的，从而得到低次最小均方误差下界，并证明低次多项式与多项式时间算法之间存在分离。

2. 主要结果：最小均方误差不连续性意味着稳定算法失效 E

本节将正式阐述我们主要的关联。我们假设处于任意的参数估计设置中，如第 1.1 节所述，并且噪声算子任意T_ρ：$\mathbf{R}^N\to\mathbf{R}^n$，$\rho\in[0,\ \ 1]$。

我们的主要结果如下。

定理 2.1。设 ρ，$\eta\in[0,\ \ 1]$。回顾含噪 *MMSE* 的定义，最小均方误差ρ，定义见 (1)。那么，任何 $\mathcal{A}$ 满足 $(\rho,\ \ \eta)$-稳定的

$$\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2]\geqslant MMSE_\rho-2\sqrt{2(7+4\eta)\eta}\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2].$$

特别地，若对于某些 $\varepsilon>0$, $\mathbb{E}_{(x,\ \ y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2]\leqslant$ 最小均方误差 $+\ \varepsilon\mathbb{E}_{XY}[\|x\|_2^2]$，则

$$MMSE_\rho-MMSE\leqslant(2\sqrt{2(7+4\eta)\eta}+\varepsilon)\cdot\mathbb{E}_{(x,y)\sim\mathbb{E}_{XY}}[\|x\|_2^2].$$

定理 2.1 有如下直接推论。

推论 2.2. 假设对于某些 $\rho\in[0,1]$, $\alpha>0$，满足

$$MMSE_\rho-MMSE\geqslant\alpha\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2].$$

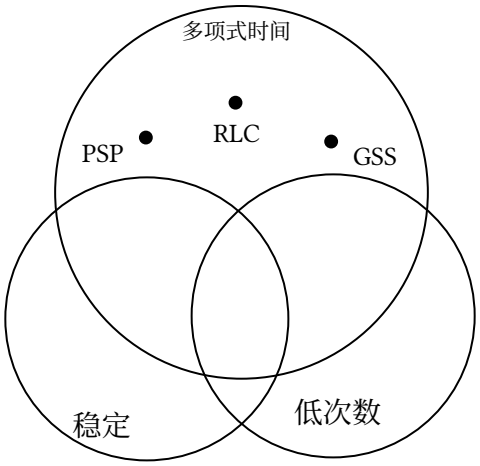


图 2. 图示说明了参数估计中，多项式时间算法可求解的类别与稳定算法/低次数多项式可求解的类别之间的分离结果。

那么，对于任意 $\eta \leqslant$ 最小值$\{\alpha^2/400, 1\},$，任何满足 (ρ, η)-稳定性的 $\mathcal{A}$ 算法都必然是 $\alpha/2$-次优的，其含义是

$$\mathbb{E}_{(x,y)\sim \mathbb{P}_{XY}}[\|\mathcal{A}(y) - x\|_2^2] \geqslant MMSE + \alpha \mathbb{E}_{XY}[\|x\|_2^2]/2. \tag{2}$$

证明。 证明依据定理最后一个展示的不等式2.1，并利用对于我们感兴趣的 η，有 $2\sqrt{2(7+4\eta)\eta} \leqslant \alpha/2$。 □

换言之，该推论表明，要证明稳定算法在估计任务上的失效，我们可以研究向输入中注入噪声后，最小均方误差上升得有多快。这正是我们在下一节的应用中采用的方法。

3. 应用

3.1. 植入最短路径模型。我们首先定义感兴趣的第一个参数化估计问题。

定义 3.1。对于 $c, C > 0$ 和 $n \in \mathbb{N}$。在定义于 n 个顶点无向图上的植入最短路径（PSP）问题中，在顶点 1 和顶点 2 之间均匀随机地抽取一条长度为 $L := (C + o(1))\log n \log\log n$ 的路径 H。观测图 G 是抽取的路径 H 与一个独立的 Erd˝os-R´enyi 随机图实例 $G(n, q)$ 的并集，其中 $q := c/n$。

统计学家的估计目标是根据观测图 G 估计植入路径 H。

先说明几点基本事项。

注记 3.2。 我们通常也会等价地用 L 和q 对植入最短路径问题进行参数化，而它们又由常数 C 和 c 隐式参数化。

注释 3.3. 我们经常用邻接矩阵表示图 G。顶点 i 和 j 之间是否存在边，由矩阵元素 $G_{ij} \in \{0, 1\}$表示。

现在，我们对这项任务的“容易程度”作如下重要说明 .

注 3.4（“PSP 的多项式时间可解性”）。 上述模型之所以称为植入最短路径问题，是因为对于任意常数 $c > 0$，一旦 $C < 1$，随着 n 增大，植入路径H 以高概率成为观测图 G 中顶点 1 和 2 之间的最短路径。特别是在这种情况下，Dijkstra 算法可以在多项式时间内以高概率从 G 准确输出 H 。

我们希望了解稳定算法在这项任务上的能力，因此定义以下自然的噪声算子。

定义 3.5 （PSP 的噪声算子）。 对于植入最短路径问题，我们考虑以下自然的噪声算子 T_ρ 。对于 $\rho \in [0, 1]$，噪声算子 T_ρ 将观测图 $\hat{}G$ 映射为其带噪版本 G ；对于每一对顶点 i, j，以概率ρ独立地从伯努利分布(q)中重新采样边连接 $\hat{} G_{ij}$ ；以剩余概率，边连接 $\hat{}$ 保持不变，且 $G_{ij} = G_{ij}$。

备注 3.6。 注意，若 $\rho = 0$、 $T_\rho(G) = G$，且若 $\rho = 1$、 $T_\rho(G)$ 的分布与一个新生成的 Erd˝os-R´enyi 随机图 $G(n, q)$ 相同。

我们的第一个结果涉及 PSP 任务的最小均方误差不稳定性。作为推论，利用定理 2.1，我们得出结论：对于这一可在多项式时间内求解的任务，稳定算法将会失败。

定理 3.7。 对于任意常数 $c > 0$ 和 $C \in (0, 1)$，若 $\rho \in (0, 1/2]$ 满足 $\rho = \omega(1/L)$,，则 *PSP* 问题的含噪 *MMSE* 至少为

$$MMSE_\rho \geqslant (1 - o(1))L.$$

特别地，当 n 足够大时，所有 $(\rho, 0。01)$-稳定算法的次优程度至少为 0。2，其含义见(2)。

PSP 的 MMSE 不稳定性结果表明，即使平均而言，从植入路径中断开任意渐增数量的边，恢复问题也会突然变得不可能。关于这一尖锐的 MMSE 跃迁的证明推迟至第 6 节。我们强调，这是对随机约束满足问题文献中所谓植入技巧的一项有趣应用 [ACO08], ，尤其体现了该技巧近期在 AoN/阈值文献中的用法[COGHK+22, MNWSZ25], 并结合了严谨的二阶矩方法。

我们的下一个结果通过定理3.7证明低次多项式无法解决 PSP 任务。我们分两步完成这一证明。首先，我们证明所有对称低次多项式对于 PSP 都是稳定的，并根据定理3.7得出它们无法解决该任务。然后，我们证明在次数为D 的多项式中，存在一个用于 PSP 任务的最优多项式是对称的，由此得出下界。

定义 3.8 （对称多项式）。 多项式 $f\colon \{0, 1\}^{\binom{[n]}{2}} \to \mathrm{R}$ 称为关于 PSP 对称的，如果对于任意固定顶点 1 和 2 的置换 $\pi\colon [n] \to [n]$，该多项式在按照置换 π 对变量 $G_{i,\ j}$ 的下标进行置换时保持不变。我们用 G_π 表示由 G 按照置换 π 对顶点进行置换后得到的图。

对称低次多项式的稳定性如下。

定理 3.9 (对称低次多项式的稳定性)。 设 $f:\ \{0,1\}^{\binom{n}{2}} \to \mathrm{R}$ 是一个次数至多为 D 的对称多项式。假设 $\frac{D(1-q)}{nq} = o(1)$, $2D < L$, $\frac{L^2 D}{n} = o(1)$, 且$q \leqslant 1/2$。那么， f对于 *PSP* 任务是 $\left(\rho, 2\left(1-(1-\rho)^D + O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\right)$-稳定的。

该定理的证明也推迟到第 6 节。我们强调，尽管其本质上直接，但它依赖于细致的计数论证，且比在零假设乘积测度下进行的任何噪声稳定性计算都复杂得多。

现在，虽然我们的稳定性结果是针对对称多项式陈述的，但我们的稳定算法下界实际上更普遍地适用于任意低次多项式。这归因于以下凸性论证。

命题 3.10。 设 $g:\ \{0,1\}^{\binom{[n]}{2}} \to \mathrm{R}$ 是一个次数至多为 D 的多项式。那么，存在一个次数至多为 D 的对称多项式 f，定义如下

$$f(G) = \frac{1}{(n-2)!} \sum_{\pi \ that \ fixes \ 1,2} g(G_\pi)$$

满足

$$\mathbb{E}_{\mathbb{P}}\left[\left(f(G) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right] \leqslant \mathbb{E}_{\mathbb{P}}\left[\left(g(G) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right].$$

证明。 这一事实是詹森不等式的直接推论:

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[\left(f(G) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right] &= \mathbb{E}_{\mathbb{P}}\left[\left(\mathop{\mathbb{E}}_{\pi \text{ that fixes } 1,2} g(G_\pi) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right] \\
&\leqslant \mathbb{E}_{\mathbb{P}}\left[\mathop{\mathbb{E}}_{\pi \text{ that fixes } 1,2}\left(g(G_\pi) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right] \\
&= \mathbb{E}_{\mathbb{P}}\left[\left(g(G) - \mathbf{1}\{\{i,j\} \in \mathbf{H}\}\right)^2\right],
\end{aligned}$$

其中，最后一个等号成立，是因为对于任意 π、 G_π 和 G，它们的分布相同，因为 P 在此类置换下保持不变。 □

结合定理 2.1、定理 3.9和命题 3.10，我们得到 PSP 的以下低次数最小均方误差下界。特别地，PSP 的 $D = O($对数 $n/$ 对数对数 $n)$ 次多项式与多项式时间算法之间存在分离。

定理 3.11 (多项式困难性). 设 $f:\ \{0,1\}^{\binom{[n]}{2}} \to \mathrm{R}^{\binom{n}{2}}$ 是一个次数至多为 D 的多项式。假设 $D < \frac{L}{2} \leqslant \frac{C}{2}\log n \log\log n$ 且 $\rho = \omega(1/L) = \omega\left(\frac{\log_{\log n} n}{\log n}\right)$，那么 f 在 *PSP* 任务中的均方误差至少为

$$\begin{aligned}
\mathbb{E}\left[\|f(G) - H\|_2^2\right] &\geqslant \left(1 - O\left(\sqrt{1-(1-\rho)^D + \sqrt{D/\log n}}\right) - o(1)\right) L \\
&\geqslant \left(1 - O\left(\sqrt{\rho D + \sqrt{D/\log n}}\right) - o(1)\right) L.
\end{aligned}$$

最后需要强调的是，我们关于 PSP 任务的低次多项式困难性几乎是紧的：存在一个次数为 L 的多项式估计器，能够以高概率实现精确恢复；而上面的定理表明，任何次数为 $o(L)$ 的多项式估计器都具有平凡的均方误差。

3.2. **随机线性码**。 我们的第二个分离结果针对随机线性码（RLC）设定。

定义 3.12。 在由 $n, m \in \mathbb{N}$ 参数化的 RLC 问题中，均匀随机抽取一个布尔矩阵 $A \in \{0,1\}^{m\times n}$，并独立地均匀随机抽取一条消息 $x \in \{0,1\}^n$。随后观测码字 $y = Ax \in \{0,1\}^m$，其中矩阵与向量的乘法在域 $\mathbb{F}_2$ 中计算。目标是根据 (y, A) 估计消息 x。

备注 3.13（“RLC 的多项式时间可解性”）。 我们注意到，在 RLC 的无噪声情形中，当 A 满列秩时，可以通过高斯消元，从 A 和 y 中精确恢复 x，且所需时间为多项式时间。也容易验证，随机布尔矩阵 $A \in \{0,1\}^{m\times n}$ 以至少 $1 - 2^{n-m}$ 的概率满列秩。因此，只要 $m - n = \omega(1)$，就存在一种多项式时间算法，能够以高概率精确恢复 x。

下面转而讨论稳定算法，并定义噪声算子。

定义 3.14 （RLC 的噪声算子）。 对于 $\rho \in [0,\ 1]$，噪声算子 T_ρ 将编码消息 $y = Ax$ 映射为其带噪版本 $\hat{}y$，其中 y 的每个坐标都以概率 ρ独立地从伯努利分布$(1/2)$中重新采样。

注 3.15。 请注意，噪声算子 T_ρ 仅作用于编码消息 y，而保持 A不变。如果 $\rho = 0, T_\rho$ 是恒等映射；如果 $\rho = 1, T_\rho$ 则将任意 y 映射为服从伯努利分布$(1/2)^{\otimes m}$的均匀随机向量。

定理 3.16。 对于任意函数 $f(n) = \omega(1)$ 和 $m = n + f(n)$，如果 $\rho \in (0,1)$ 满足 $\frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m} = \omega\left(2^{2(m-n)}\right)$，那么 *RLC* 问题的含噪 *MMSE* 至少为

$$MMSE_\rho \geqslant \left(\frac{1}{4} - o(1)\right) n.$$

特别是，当 n 足够大时，所有 $(\rho,\ 0。01)$稳定算法的性能都比最优解差 0。2，其含义见(2)。

我们现在通过直接分析后验均值来证明这个最小均方误差不稳定性定理，证明见第 7 节。

下面给出几点说明。

注 3.17。请注意，$\frac{1}{4}n$ 是通过直接输出常向量 $(1/2, 1/2, \ldots\ldots,\ 1/2)^\top \in \mathbb{R}^n$所达到的均方误差。

注记 3.18。 我们进一步详细说明上文定理 $\frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m} \gg 2^{2(m-n)}$ 中的条件 $1 \ll f(n) \ll n$。对于 ρ，其基本含义是 $\frac{m-n}{n}$ 至少需要达到 $\frac{m-n}{n}$ 的量级。

更确切地说，如果存在一个常数 $\varepsilon > 0$ 使得 $\frac{1}{2}\rho - \frac{1}{4}\rho^2 \geqslant \left(2\log_3 2 + \varepsilon\right) \cdot \frac{m-n}{m}$，那么条件 $\frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m} \gg 2^{2(m-n)}$ 就满足了。我们可以通过验证以下内容来确认这一断言

$$\frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m} = e^{-m\log\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)} \geqslant e^{3m\left(\frac{1}{2}\rho-\frac{1}{4}\rho^2\right)} \geqslant e^{3\left(\frac{2\log 2}{3}+\varepsilon\right)(m-n)} \gg 2^{2(m-n)}.$$

因此，我们的最小均方误差不稳定性结果，即定理 3.16 证明，只要我们随机化大约 $m-n$ 比例的 m 个 y 条目，恢复任务就会变得不可能。

我们的下一个结果是，对于 RLC 任务，任意 D 次多项式都具有 $(\rho, O(\rho D))$-稳定性。

定理 3.19 (RLC 的低次多项式的稳定性)。设 $f\colon \{0,1\}^{m\times n} \times \{0,1\}^m \to \mathbb{R}$ 是一个次数至多为 D 的多项式。假设 $2D \leqslant n$。那么，f 对于 *RLC* 是 $(\rho, 2(1-(1-\rho)^D))$-稳定的。

RLC 设定的 $\mathbb{F}_2$-结构事实上导致了 y 与 A 之间“几乎”独立，因此，与其他模型相比，低次稳定性的证明要简单得多。证明见第 7 节。

因此，我们得到以下关于 RLC 的低阶最小均方误差下界。特别地，我们得出结论：对于 RLC，$D = O(m/(m-n))$ 次多项式与多项式时间方法之间存在分离。

推论 3.20。对于任意函数 $f(n) = \omega(1)$ 和 $m = n + f(n)$，若存在某个常数 $\varepsilon > 0$，使得 $\frac{1}{2}\rho - \frac{1}{4}\rho^2 \geqslant \left(2^{\text{对数}\frac{2}{3}} + \varepsilon\right) \cdot \frac{m-n}{m}$，那么任何次数为 D 的多项式 $g\colon \{0,1\}^{m\times n} \times \{0,1\}^n \to \mathbb{R}^n$ 的均方误差至少为

$$\mathbb{E}\left[\|g(A,y) - x\|_2^2\right] \geqslant \left(\frac{1}{4} - O\left(\sqrt{1-(1-\rho)^D}\right) - o(1)\right) n.$$

特别地，若 $D \leqslant \frac{1}{\rho}$，则有

$$\mathbb{E}\left[\|g(A,y) - x\|_2^2\right] \geqslant \left(\frac{1}{4} - O\left(\sqrt{\rho D}\right) - o(1)\right) n.$$

3.3. **高斯子集和（GSS）的结果**。我们建立分离所用的最终模型是高斯子集和（GSS）模型。

定义 3.21. 令 $k, N \in \mathbb{N}$ 满足 $k \leqslant N$。在高斯子集和（GSS）问题中，对于某个未知的 k-子集 $S \subseteq [N]$，我们观测到 $Y = \sum_{i\in S} X_i$，其中 $(X_i)_{i\in[N]}$ 是独立同分布的 $N(0,1)$。统计学家的目标是根据 Y 和 $X_i, i = 1,2,\ldots,N$，估计 $1_S \in \{0,1\}^N$，即集合 S 的指示向量，。

注3.22（“GSS 的多项式时间可解性”）。简单计算可知，对于任意 k，$N \in \mathbb{N}$ 只要 $\sigma = 0$，就能以概率 1 恢复 S。与此不同的是，对于 PSP 或 RLC，即使在最坏情况下，相应的最坏情况任务也可解（最短路径和线性方程组属于 $\mathcal{P}$），而子集和预计不属于 $\mathcal{P}$。然而，令人惊讶的是，存在一种多项式时间算法，能够以高概率在 GSS 中精确恢复 1_S [Fri86、ZG18]。值得注意的是，该算法基于 Lenstra-Lenstra-Lovász 格基约简方案。

我们为该任务定义噪声算子。

定义 3.23 （GSS 的 Ornstein–Uhlenbeck（OU）算子）。 对于 $\rho \in [0,1]$，OU 噪声算子 T_ρ 将子集和 y 映射为带噪版本 $\hat{y} = \sqrt{1-\rho^2}y + \rho Z$，其中 Z 是独立的 $\mathcal{N}(0,1)$。

为证明这一点，我们先证明以下命题，它直接推论自 AoNresults 在 [RXZ19]中的主要结果。证明推迟到第 8节。

命题 3.24。 假设 $k \leqslant N^0$。[49]。那么对于某个 $\rho = \exp(-\Theta(k\log(N/k)))$，成立

$$MMSE_\rho \geqslant (1-o(1))k.$$

特别地，当 n 足够大时，所有 $(\rho$,0。01)-稳定的都必定是 0。4-次优的，具体含义见 (2)。

除了证明稳定算法分离之外，我们还证明了低次多项式的困难性。一如既往，我们首先证明以下关于低次多项式的稳定性定理。

定理 3.25。 设 $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ 是一个次数至多为 D 的多项式。假设 $k = o(n)$ 和$D = o\left(\min\left\{k^{1/4}, (n/k)^{1/5}\right\}\right)$。那么， f 对 *GSS* 是 $(\rho, 2(1-\overline{(\sqrt{1}-\rho^2)^D} + o(1)))$-稳定的。

这个结果的证明并非显然，它依赖一系列精细的组合下界，以及针对厄米特基的维克公式。证明推迟到第 8 节。

作为推论，我们得到以下针对 GSS 的低阶最小均方误差下界。特别地，我们得出，$D = o\left(\min\{k^{1/4}, (n/k)^{1/5}\}\right)$ 次多项式与针对 GSS 的多项式时间方法之间存在分离。

推论 3.26. 假设 $k = o(n)$, $D = o\left(\min\{k^{1/4}, (n/k)^{1/5}\}\right)$, 且 $\rho \geqslant \exp(-k\log(n/k))$。那么，任何次数为 D 的多项式 $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ 的均方误差至少为

$$\mathbb{E}\left[\|f(X,Y) - \mathbb{1}_S\|_2^2\right] \geqslant \left(1 - O\left(\sqrt{1-\left(\sqrt{1-\rho^2}\right)^D}\right) - o(1)\right)k.$$

特别地，如果 $D \leqslant \frac{1}{\rho^2}$，我们有

$$\mathbb{E}\left[\|f(X,Y) - \mathbb{1}_S\|_2^2\right] \geqslant \left(1 - O\left(\rho^2 D\right) - o(1)\right)k.$$

4. 开始：证明的预备知识

这里，我们介绍一些背景知识并讨论我们的结果，以帮助读者理解。首先，回顾一下，最小均方误差（MMSE）由后验均值实现。

事实 4.1. 后验均值达到最小均方误差：

$$\inf_{\mathcal{A}:\mathbb{R}^N\to\mathbb{R}^n} \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2] = \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathbb{E}[x|y]-x\|_2^2].$$

现在，虽然我们将上文定义的噪声算子抽象地处理，但考察自然噪声算子所满足的一些性质或许会有所启发。

(1) 当 $\rho = 0$ 时，噪声算子应为恒等映射。(2) 当 $\rho = 1$ 时，噪声算子引入最大量的噪声，且观察 $T_\rho(y)$ 不会为关于 x 的信息提供任何补充；也就是说，对于 X 支撑集中的任意 $x_1 \neq x_2$， $T_\rho(y)|x_1$ 的分布与 $T_\rho(y)|x_2$ 的分布相同。(3) 对于 $\rho_1 < \rho_2$，噪声算子 T_{ρ_1} 注入的噪声少于 T_{ρ_2}，并且存在 $\rho_3 \in [0,\ 1]$ 使得 $T_{\rho_2} = T_{\rho_3} \circ T_{\rho_1}$。

我们的噪声算子对观测值 y 的一组固定坐标大致按以下方式作用：

- （q-伯努利噪声） T_ρ 作用于 y 的一个固定子集：以概率 ρ 独立地从伯努利分布 $\mathrm{Bern}(q)$ 中重新采样该子集中的每个坐标。
- （高斯噪声） T_ρ 作用于 y 的一个固定子集：在每个坐标上独立地应用 Ornstein–Uhlenbeck 算子。换句话说，在该子集内，噪声算子用一个独立的标准高斯分布 $N(0,\ 1)$ 对 y 的每个坐标取平均，并输出 $\sqrt{1-\rho^2}y_i + \rho z_i$ 其中 $z_i \sim N(0,\ 1)$ 对于坐标 y_i。

5. 证明 关键结果：MMSE 不稳定性意味着稳定算法失败

本节给出定理2.1的（简单）证明。

证明。 令 $\mathcal{A} : \mathrm{R}^N \to \mathrm{R}^n$ 是一个 $(\rho,\ \eta)$-稳定算法。

我们首先断言，要么 $\mathbb{E}_{(x,\ y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2] \geqslant$ 最小均方误差$_\rho$，此时所需结论成立；要么 $\mathbb{E}_{(x,\ y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2] \leqslant$ 最小均方误差$_\rho \leqslant \mathbb{E}_{(x,\ y)\sim\mathbb{P}_{XY}}[\|x\|_2^2]$ 这意味着

$$\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y)\|_2^2\right] \leqslant \mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2] + 2\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2] \leqslant 4\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2]. \tag{3}$$

此外，我们的假设和（3）给出

$$\begin{aligned}
&\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))\|_2^2\right] && (4)\\
&\leqslant 2\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y)\|_2^2\right] + 2\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-\mathcal{A}(y)\|_2^2\right] && (5)\\
&\leqslant 8\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right] + 2\eta\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y)\|_2^2\right] && (6)\\
&\leqslant 8(1+\eta)\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right]。 && (7)
\end{aligned}$$

Now,

$$\begin{aligned}
&\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-x\|_2^2-\|\mathcal{A}(y)-x\|_2^2\right]\\
&=\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\langle\mathcal{A}(T_\rho(y))-\mathcal{A}(y),\mathcal{A}(T_\rho(y))+\mathcal{A}(y)-2x\rangle\right]\\
&\leqslant\sqrt{\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-\mathcal{A}(y)\|_2^2\right]\left(\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))\|_2^2+\|\mathcal{A}(y)\|_2^2+2\|x\|_2^2\right]\right)}
\end{aligned}$$

根据柯西-施瓦茨不等式。利用 $(\rho,\ \eta)$ 稳定性 $\mathcal{A}$，（3）和（7），我们有

$$\begin{aligned}
&\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-x\|_2^2-\|A(y)-x\|_2^2\right]\\
&\leqslant\sqrt{\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-\mathcal{A}(y)\|_2^2\right]\left(\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))\|_2^2+\|\mathcal{A}(y)\|_2^2+2\|x\|_2^2\right]\right)}\\
&\leqslant\sqrt{\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(T_\rho(y))-\mathcal{A}(y)\|_2^2\right]\left(\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[8(1+\eta)\|x\|_2^2+4\|x\|_2^2+2\|x\|_2^2\right]\right)}\\
&\leqslant\sqrt{2(7+4\eta)\eta\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|\mathcal{A}(y)\|_2^2\right]\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right]}\\
&\leqslant 2\sqrt{2(7+4\eta)\eta\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right]\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right]}\\
&=2\sqrt{2(7+4\eta)\eta}\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}\left[\|x\|_2^2\right].
\end{aligned}$$

重新整理不等式，可得

$$\begin{aligned}
&\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(y)-x\|_2^2]\\
&\geqslant\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|\mathcal{A}(T_\rho(y))-x\|_2^2]-2\sqrt{2(7+4\eta)\eta}\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2]\\
&\geqslant\mathrm{MMSE}_\rho-2\sqrt{2(7+4\eta)\eta}\cdot\mathbb{E}_{(x,y)\sim\mathbb{P}_{XY}}[\|x\|_2^2]
\end{aligned}$$

如所要求。 □

6. Proofs for Planted Shortest Path

6.1. **含噪 MMSE 下界：定理 3.7 的证明。** 为了证明植入最短路径问题的含噪 MMSE 下界，我们的主要思路是研究一个称为“近似路径”的自然对象。

定义 6.1（近似路径）。设 G 为一个图，$l\in\mathbb{N}$ 且 $\varepsilon\in(0,\ 1)$ 满足 $\varepsilon l\in\mathbb{N}$。若一对 $(\hat{}P,\ P)$ 是 $(1-\varepsilon)$-近似路径，则称其为 G 中连接顶点 1 和 2 的长度为l 的近似路径；其中，P 是完全图 K_n中连接顶点 1 和 2 的长度为l 的路径，$\hat{}P\subseteq P$、$\hat{}P\subseteq G$，且 $|\hat{}P|=(1-\varepsilon)l$。

换句话说，$(\hat{}P,\ P)$ 是 $(1-\varepsilon)$中的长度为l 的近似路径，当且仅当 P 是完全图 K_n 中的长度为l的路径，$\hat{}P$ 包含 P中恰好 $(1-\varepsilon)$ 比例的边，并且$\hat{}P$ 是 G的子图。

定义 6.2（近似路径的重叠）。设 $(\hat{}P_1,\ P_1)$ 和 $(\hat{}P_2,\ P_2)$ 是图 G 中两条长度为 l 的 $(1-\varepsilon)$-近似路径。如果 $|E(P_1\cap P_2)|=k$，则称 $(\hat{}P_1,\ P_1)$ 和 $(\hat{}P_2,\ P_2)$ 共享 k 条边，或重叠度为 k。

注记 6.3（近似路径与含噪植入最短路径）。 在植入最短路径问题的带噪版本中，近似路径的概念自然出现。回顾一下，在

植入最短路径问题中，抽取图 G，使其为 Erd˝os-R´enyi 随机图 $G(n,q)$ 与一条连接顶点 1 和 2、长度为 L 的均匀随机路径 H 的并集。噪声算子 T_ρ 作用于 G，以概率 ρ 独立地重新抽样每条 G_{ij}，抽样自伯努利分布(q)。因此，观测到的含噪图 $T_\rho(G)$ 是 Erd˝os-R´enyi 随机图 $G(n,q)$ 与 $\hat{}\hat{}H$ 的并集，其中 H 由 H 得到，方法是以概率ρ独立断开 H 的每条边。因此，$\hat{}H,H)$ 是 $\hat{}G$ 中连接顶点 1 和 2、长度为 L 的 $(1-\varepsilon)$-近似路径，其中 εL 服从 $\mathrm{Bin}(L,\rho)$ 分布。此外，在 H中有εL 条边被断开的条件下，$\hat{}H,H)$ 的分布是在 G 中连接顶点 1 和 2 的所有长度为 L 的$(1-\varepsilon)$-近似路径上的均匀分布。

证明的第一步是对零假设模型中顶点 1 和 2 之间长度固定的 $(1-\varepsilon)$-近似路径的二阶矩进行界定，其中我们观测到一个 Erd˝os-R´enyi 随机图 $G(n,\ q)$。我们的证明改编自 [LS24，引理 2。1]。

定理 *6.4.* 设 $N_{m,\varepsilon}$ 表示 *Erd˝os-R´enyi* 随机图 $G(n,q)$ 中顶点 *1* 和 *2* 之间长度为 m 的 $(1-\varepsilon)$-近似路径的数量；$N^{(2)}_{m},_{\varepsilon}$ 表示 *Erd˝os-R´enyi* 随机图 $G(n,q)$ 中顶点 *1* 和 *2* 之间长度为 m 的 $(1-\varepsilon)$-近似路径的非不相交对数目，也就是说，在 $G(n,q)$ 中，顶点 *1* 和 *2* 之间长度为 m 的 $(1-\varepsilon)$-近似路径对 $((\hat{}P_1,P_1),(\hat{}P_2,P_2))$ 的数量，其中 P_1 和 P_2 在完全图 K_n 中并非边不相交。若 $m=\omega(1),\ 1\leqslant \varepsilon m\leqslant m-1,\ nq\geqslant \omega(\sqrt{m}),\ m\ll n^{1/15}$，且 $mq\leqslant 1/3,,$ 则

$$E_{\mathbb{Q}}[N^{(2)}_{m,\varepsilon}]\leqslant O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right)\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2=o(\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2),$$

$$\mathbb{E}_{\mathbb{Q}}[N^2_{m,\varepsilon}]\leqslant\left(1+O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right)\right)\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2\leqslant(1+o(1))\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2,$$

and

$$\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]\geqslant\Omega\left(n^{\varepsilon m-1}\cdot(nq)\right)\geqslant\omega(1).$$

我们现在证明，定理6.4 蕴含以下 MMSE 下界。

推论 6.5。令 $x\in\{0,1\}^{\binom{n}{2}}$ 为一条从顶点 *1* 到顶点 *2*、长度为 m 的路径，并从所有此类路径中均匀随机选取。考虑植入模型 P类$_{m,\ \varepsilon,\ q}$ 其中，在一个实例 $G_0\sim G(n,q)$ 中植入一个从 x 的所有 $(1-\varepsilon)m$-子集中均匀随机选取的子集，记为 $x_\varepsilon\in\{0,1\}^{\binom{n}{2}}$。那么，若 $m=\omega(1)$，$1\leqslant\varepsilon m\leqslant m-1,nq\geqslant\omega(\sqrt{m}),m\ll n^{1/15}$，且 $mq\leqslant 1/3,$，则

$$\mathbb{E}[\|\mathbb{E}[x|G=G_0\cup x_\varepsilon]-x\|_2^2]\geqslant\mathbb{E}[\|x\|_2^2]\left(1-O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right)\right).$$

证明。该证明运用了若干精细的测度变换恒等式，这种技巧通常称为“植入技巧” [ACO08, MNWSZ25]。令 S_m 表示 K_n 中顶点 1 和顶点 2 之间长度为 m 的路径集合，以及 $M_m=|S_m|$。另令 $S_{m,\varepsilon}$ 表示 K_n 中顶点 1 和顶点 2 之间长度为m 的 $(1-\varepsilon)$近似路径集合，以及$M_{m,\varepsilon}=|S_{m,\varepsilon}|$。现在，对于任意 n 顶点 G 和 $x\in S_m$

$$Z_Y(G)=(\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}M_m)^{-1}\sum_{P\in S_{m,\varepsilon},H\in S_m:P\subseteq H}1(P\subseteq G),$$

and

$$Z_Y(x,G)=(\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}M_m)^{-1}\sum_{P\in S_{m,\varepsilon},H\in S_m:H\cap x=\varnothing,P\subseteq H}1(P\subseteq G).$$

注意，对于任意 G，以下结论成立。对于 $x_\varepsilon\in S_{m,\varepsilon}$ 的任意取值，

$$\frac{\mathbb{P}_{m,\varepsilon,q}(G|x_\varepsilon)}{\mathbb{Q}_q(G)}=1(x_\varepsilon\subseteq G)q^{-(1-\varepsilon)m}$$

因此，对于任意值的 $x\in S_m$，

$$\frac{\mathbb{P}_{m,\varepsilon,q}(G|x)}{\mathbb{Q}_q(G)}=\frac{1}{\binom{m}{\varepsilon m}q^{(1-\varepsilon)m}}\sum_{P\in S_{m,\varepsilon}:P\subseteq x}1(P\subseteq G).\tag{8}$$

因此，

$$\frac{\mathbb{P}_{m,\varepsilon,q}(G)}{\mathbb{Q}_q(G)}=M_m^{-1}\sum_{x\in S_m}\frac{\mathbb{P}_{m,\varepsilon,q}(G|x)}{\mathbb{Q}_q(G)}=Z_Y(G).\tag{9}$$

显然，对于 $\mathbb{Q}q$ 该 $G(n,\ q)$ 测度，第一步利用先验分布的均匀性，第二步利用测度变换（8），我们有

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}}Z_Y(x,G)\\
&=(\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}M_m)^{-1}M_m^{-1}\sum_{x,H\in S_m:H\cap x=\varnothing}\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}|x}\sum_{P\in S_{m,\varepsilon}:P\subseteq H}1(P\subseteq G)\\
&=(\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}M_m)^{-2}\sum_{x,H\in S_m:H\cap x=\varnothing}\mathbb{E}_{\mathbb{Q}_q}\sum_{P,P'\in S_{m,\varepsilon}:P\subseteq H,P'\subseteq x}1(P,P'\subseteq G)\\
&=(\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}M_m)^{-2}\mathbb{E}_{\mathbb{Q}_q}\sum_{x,H\in S_m:H\cap x=\varnothing}\sum_{P,P'\in S_{m,\varepsilon}:P\subseteq H,P'\subseteq x}1(P,P'\subseteq G)
\end{aligned}$$

现在令 $N_{m,\ \varepsilon}$ 为 $1-\varepsilon$-近似路径的数量。那么显然 $\mathbb{E}_{\mathbb{Q}_q}N_{m,\ \varepsilon}=\binom{m}{\varepsilon m}q^{m(1-\varepsilon)}$ 并且

$$\mathbb{E}_{\mathbb{Q}_q}N_{m,\varepsilon}^{(2)}=\mathbb{E}_{\mathbb{Q}_q}\sum_{x,H\in S_m:H\cap x=\varnothing}\sum_{P,P'\in S_{m,\varepsilon}:P\subseteq H,P'\subseteq x}1(P,P'\subseteq G).$$

结合上述结果与定理 6.4，我们得出

$$\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}}Z_Y(x,G)=\frac{\mathbb{E}_{\mathbb{Q}_q}[N_{m,\varepsilon}^{(2)}]}{\mathbb{E}_{\mathbb{Q}_q}[N_{m,\varepsilon}]^2}=O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right).$$

因此，根据定理 6.4，

$$\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}}Z_Y(x,G)=O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right).$$

此外，对于任意 $\delta>0$，由（9）可知，

$$\mathbb{P}_{m,\varepsilon,q}(Z_Y(G)\leqslant\delta)=\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}}[\frac{\mathbb{P}_{m,\varepsilon,q}(G)}{\mathbb{Q}_q(G)}1(Z_Y(G)\leqslant\delta)]\leqslant\delta.$$

因此，对于任意 $\delta > 0$,

$$\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}} \frac{Z_Y(x,G)}{Z_Y(G)} \leqslant \delta + \mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}} Z_Y(x,G)/\delta$$

对 δ 进行优化后，我们可以得出

$$\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}} \frac{Z_Y(x,G)}{Z_Y(G)} = O\left(\sqrt{\mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}} Z_Y(x,G)}\right) = O\left(\sqrt{\frac{1}{n^{\varepsilon m-1}\cdot(nq)} + \frac{\sqrt{m}}{nq}}\right).$$

但在 P类m中，ε，q 给定 x 的后验分布 G 根据（8）对所有实现 H均满足，

$$\mathbb{P}_{m,\varepsilon,q}(H|G) \propto \sum_{P\in S_{m,\varepsilon}: P\subseteq H} 1(P \subseteq G).$$

因此，

$$\mathbb{P}_{x,G\sim\mathbb{P}_{m,\varepsilon,q}|x,H\sim\mathbb{P}(x|G)}[H\cap x \neq \varnothing] = \mathbb{E}_{\mathbb{P}_{m,\varepsilon,q}} \frac{Z_Y(x,G)}{Z_Y(G)} = O\left(\sqrt{\frac{1}{n^{\varepsilon m-1}\cdot(nq)} + \frac{\sqrt{m}}{nq}}\right),$$

这意味着

$$\mathbb{E}_{x,G\sim\mathbb{P}_{m,\varepsilon,q}|x,H\sim\mathbb{P}(x|G)}[|H\cap x|] = O\left(m\sqrt{\frac{1}{n^{\varepsilon m-1}\cdot(nq)} + \frac{\sqrt{m}}{nq}}\right),$$

因此，根据西莫里恒等式（例如参见 [NWZ23，引理 2]）

$$\begin{aligned}\mathbb{E}[\|\mathbb{E}[x|G=G_0\cup x_\varepsilon]-x\|_2^2] &= \mathbb{E}[\|x\|_2^2] - \mathbb{E}_{x,\ G\sim\mathbb{P}_{m,\ \varepsilon,\ q}|x,\ H\sim\mathbb{P}类(x|G)}[|H\cap x|]\\ &= \mathbb{E}[\|x\|_2^2] - O\left(m\sqrt{\frac{1}{n^{\varepsilon m-1}\cdot(nq)} + \frac{\sqrt{m}}{nq}}\right)\\ &= \mathbb{E}[\|x\|_2^2]\left(1 - O\left(\sqrt{\frac{1}{n^{\varepsilon m-1}\cdot(nq)} + \frac{\sqrt{m}}{nq}}\right)\right).\end{aligned}$$

□

定理 3.7。现在，根据推论 6.5，我们准备证明定理 3.7中所述的含噪 MMSE 下界。在植入最短路径问题中，对于某些常数 c，$C>0$,$L=(C+o(1))\log n\log\log n$和 $q=c\log n/n$。我们验证 $\sqrt{L}=\omega(1)$，$nq=\omega(\sqrt{L})$，$L\ll n^{1/15}$和$Lq\leqslant 1/3$。

回想一下，在植入最短路径问题中，我们观测到图 G，它是$G_0\sim G(n,q)$与一条在顶点 1 和 2 之间、长度为 L 的均匀随机路径 H 的并集。如注释 6.3 所述，在植入最短路径问题的带噪版本中，观测到的是 $T_\rho(G)$，它是 Erdős-Rényi 随机图 $\hat{}G_0\sim G(n,q)$与 $\hat{}H$ 的并集，其中 $(\hat{}H,H)$ 是一条在顶点 1 和 2 之间、长度为 L 的 $(1-\varepsilon)$-近似路径，且 $\varepsilon L\sim\mathrm{Bin}(L,\rho)$。由于 $\hat{}H$ 是通过以概率 ρ 独立断开 H 中的每条边而从 H 得到的，因此我们看到，在给定 εL 的条件下，$\hat{}H$ 是 H 的一个均匀随机 $(1-\varepsilon)L$ 子集。于是，若用 $\mathbb{P}_{L,q}$ 表示植入最短路径问题中 (G_0,H) 的分布，并用 $\mathbb{P}_{L,\varepsilon,q}$ 表示

分布 $(\hat{}G_0,\ (\hat{}H,\ H))$，其中 $\hat{}G_0 \sim G(n,\ q)$、H 是一条在顶点 1 和 2 之间、长度为 L 的均匀随机路径，而 $\hat{}H$ 是 H 的一个均匀随机 $(1-\varepsilon)L$ 子集，我们有

$$\begin{aligned}\mathrm{MMSE}_\rho &= \mathbb{E}_{(G_0,H)\sim\mathbb{P}_{L,q}}\left[\left\|\mathbb{E}\left[H|T_\rho(G_0\cup H)\right]-H\right\|_2^2\right]\\ &= \mathop{\mathbb{E}}_{\varepsilon L\sim\mathrm{Bin}(L,\rho)}\mathbb{E}_{(\hat{G}_0,(\hat{H},H))\sim\mathbb{P}_{L,\varepsilon,q}}\left[\left\|\mathbb{E}\left[H|\hat{G}_0\cup\hat{H}\right]-H\right\|_2^2\right].\end{aligned}$$

由于 $L=\omega(1),\ \ nq=\omega(\sqrt{L}),\ \ L\ll n^{1/15}$ 且 $Lq\leqslant 1/3$, 若 εL 满足 $1\leqslant\varepsilon L\leqslant m-1$, 根据推论6.5，我们有

$$\mathbb{E}_{(\hat{G}_0,(\hat{H},H))\sim\mathbb{P}_{L,\varepsilon,q}}\left[\left\|\mathbb{E}\left[H|\hat{G}_0\cup\hat{H}\right]-H\right\|_2^2\right]\geqslant\mathbb{E}\left[\|H\|_2^2\right]\cdot\left(1-O\left(\frac{1}{n^{\varepsilon L-1}\cdot(nq)}+\frac{\sqrt{L}}{nq}\right)\right).$$

现在，由于 $\rho L=\omega(1)$，我们知道 $\varepsilon L\sim\mathrm{Bin}(L,\ \rho)$ 以概率$1-(1-\rho)^L\geqslant 1-\exp(-\rho L)\geqslant 1-o(1)$满足 $\varepsilon L\geqslant 1$。类似地，如果 $\rho\leqslant 1/2,$，则我们有 $\varepsilon L\leqslant L-1$，其概率至少为 $1-o(1)$。因此，

$$\begin{aligned}&\mathrm{MMSE}_\rho\\ &=\mathop{\mathbb{E}}_{\varepsilon L\sim\mathrm{Bin}(L,\rho)}\mathbb{E}_{(\hat{G}_0,(\hat{H},H))\sim\mathbb{P}_{L,\varepsilon,q}}\left[\left\|\mathbb{E}\left[H|\hat{G}_0\cup\hat{H}\right]-H\right\|_2^2\right]\\ &\geqslant\Pr(1\leqslant\varepsilon L\leqslant L-1)\cdot\mathop{\mathbb{E}}_{\varepsilon L\sim\mathrm{Bin}(L,\rho)}\left[\mathbb{E}_{(\hat{G}_0,(\hat{H},H))\sim\mathbb{P}_{L,\varepsilon,q}}\left[\left\|\mathbb{E}\left[H|\hat{G}_0\cup\hat{H}\right]-H\right\|_2^2\right]\middle|1\leqslant\varepsilon L\leqslant L-1\right]\\ &\geqslant(1-o(1))\cdot\mathbb{E}\left[\|H\|_2^2\right]\cdot\left(1-O\left(\frac{1}{nq}+\frac{\sqrt{L}}{nq}\right)\right)\\ &=(1-o(1))\mathbb{E}\left[\|H\|_2^2\right],\end{aligned}$$

由于 $nq=\omega(\sqrt{L})$。证明至此结束。 □

定理 *6.4.* 首先，我们给出 $N_{m,\ \varepsilon}$ 的期望下界。我们有

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}] &= (n-2)_{(m-1)}\binom{m}{\varepsilon m}q^{m-\varepsilon m}\\ &\geqslant(1-O(1/n))n^{m-1}\binom{m}{\varepsilon m}q^{m-\varepsilon m}\\ &=(1-O(1/n))\frac{1}{n}\binom{m}{\varepsilon m}\left(nq^{1-\varepsilon}\right)^m\\ &\geqslant(1-O(1/n))\frac{1}{n}\left((nq)^{1-\varepsilon}n^{\varepsilon}\right)^m\\ &\geqslant(1-O(1/n))\frac{n^{\varepsilon m}}{n}(nq)^{(1-\varepsilon)m}.\end{aligned}$$

特别是，如果 $1\leqslant\varepsilon m\leqslant m-1$、$nq\geqslant\omega(\sqrt{m})\geqslant\omega(1)$ 和 $mq\leqslant 1/3,$，我们有

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}] &\geqslant(1-O(1/n))\frac{n^{\varepsilon m}}{n}(nq)^{(1-\varepsilon)m}\\ &\geqslant\Omega\left(n^{\varepsilon m-1}\cdot(nq)\right)\end{aligned}$$

$$\geqslant \omega(1).$$

接下来，我们给出 $N^{(2,\,k)}_{m,\,\varepsilon}$ 上期望的上界，即在 $G(n,\ q)$ 中，顶点 1 和 2 之间长度为 m 的 $(1-\varepsilon)$-近似路径中，共享 k 条边的路径对数。回顾一下，这等于顶点 1 和 2 之间长度为 m 的 $(1-\varepsilon)$-近似路径中，满足 $|E(P_1 \cap P_2)| = k$ 的 $((\hat{}P_1,\ P_1),\ (\hat{}P_2,\ P_2))$ 对数。于是我们有

$$\mathbb{E}_{\mathbb{Q}}[N^{(2,k)}_{m,\varepsilon}] \leqslant M_{k,m-k,m-k} \sum_{j=0}^{\min\{2\varepsilon m, m-k+\varepsilon m\}} \binom{2m-k}{j} q^{2m-k-j}, \tag{10}$$

其中，$M_{k,\,m-k,\,m-k}$ 表示完全图 K_n中顶点1和2之间长度为 m 的路径对的数量，这些路径对共享 k 条边。内层求和遍历从$\hat{}\,P_1 \cup \hat{}P_2$中缺失的 $P_1 \cup P_2$的边。注意，$\hat{}\,\hat{}\,\hat{}$中的边至多有 $\min\{2\varepsilon m,\ m-k+\varepsilon m\}$条不在 $P_1 \cup P_2$中。这里，$2\varepsilon m$ 是因为对于每个 $i \in \{1,\ 2\}$，P_i 中有 εm 条边从 P_i 中缺失；而 $m-k+\varepsilon m$ 是因为$\hat{}P_1$中恰好缺失 P_1 的 εm 条边，并且$\hat{}\,P_2$ 中至多缺失 $P_2 \backslash P_1$ 的 $m-k$ 条边。

回顾一下，$N^{(2)}_{m,\,\varepsilon}$ 表示 $(1-\varepsilon)$-近似路径中非不相交对的数量，这些路径的长度为m，且位于顶点1和2之间；也就是说，$((\hat{}P_1,\ P_1),\ (\hat{}P_2,\ P_2))$ 表示 $(1-\varepsilon)$-近似路径对的数量，这些路径长度为 m，位于顶点1和2之间，并满足 $|E(P_1 \cap P_2)| > 0$。根据（10），我们有

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[N^{(2)}_{m,\varepsilon}] &= \sum_{k=1}^{m} \mathbb{E}_{\mathbb{Q}}[N^{(2,k)}_{m,\varepsilon}] \\
&= \sum_{k=1}^{m} \mathbb{E}_{\mathbb{Q}}[N^{(2,k)}_{m,\varepsilon}] \\
&\leqslant \sum_{k=1}^{m} M_{k,m-k,m-k} \sum_{j=0}^{\min\{2\varepsilon m, m-k+\varepsilon m\}} \binom{2m-k}{j} q^{2m-k-j}.
\end{aligned} \tag{11}$$

最后，我们给出 N_m 的二阶矩上界ε。注意$N^2_{m,\,\varepsilon} = \sum_{k=0}^{m} N^{(2,\,k)}_{m,\,\varepsilon} = N^{(2,\,0)}_{m,\,\varepsilon} + N^{(2)}_{m,\,\varepsilon}$，通过考虑一对 $(1-\varepsilon)$-近似路径的所有重叠模式。此外，我们有 $\mathbb{E}_{\mathbb{Q}}[N^{(2,\,0)}_{m,\,\varepsilon}] \leqslant \mathbb{E}_{\mathbb{Q}}[N_{m,\,\varepsilon}]^2$，因为对于一对不相交的路径，它们在 $(1-\varepsilon)$ 中为 $\mathbb{Q}$ -近似路径的事件是独立的。因此，由（11），我们有

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}[N^2_{m,\varepsilon}] \\
&= \mathbb{E}_{\mathbb{Q}}[N^{(2,0)}_{m,\varepsilon}] + \mathbb{E}_{\mathbb{Q}}[N^{(2)}_{m,\varepsilon}] \\
&\leqslant \mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2 + \sum_{k=1}^{m} M_{k,m-k,m-k} \sum_{j=0}^{\min\{2\varepsilon m, m-k+\varepsilon m\}} \binom{2m-k}{j} q^{2m-k-j}
\end{aligned}$$

由于 $mq \leqslant 1/3$ 和 $m = \omega(1)$，内层求和中关于 j 的相邻两项之比为 $\frac{\binom{2m-k}{j+1} q^{2m-k-j-1}}{\binom{2m-k}{j} q^{2m-k-j}} = \frac{2m-k-j}{j+1} \cdot \frac{1}{q} \geqslant \frac{1}{2m} \cdot \frac{1}{q} \geqslant \frac{3}{2}$，因此内层求和至多相差一个常数因子 3，由最后一项主导。于是我们有

$$\leqslant \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + \sum_{k=1}^{m} M_{k,m-k,m-k} \cdot 3\binom{2m-k}{t} q^{2m-k-t}$$

其中 $t = \text{最小值}\{2\varepsilon m,\ m-k+\varepsilon m\}$。现在我们调用 [LS24的断言 2。1], 该断言指出，存在某个常数 C，使得 $M_{k,\ m-k,\ m-k} \leqslant \left(\frac{k+1}{n^2} + \frac{Ck^9m^6}{n^3}\right) n^{2m-k}$ 对于 $m \ll n^{1/3}$ 和 $1 \leqslant k \leqslant m-1$。当 $k = m$ 时，我们有 $M_{m,\ 0,\ 0} = (n-2)_{(m-1)} \leqslant n^{m-1}$。因此，

$$\begin{aligned}
&\leqslant \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + 3n^{m-1}\binom{m}{\varepsilon m} q^{m-\varepsilon m} \\
&\quad + 3\sum_{k=1}^{m-1} \left(\frac{k+1}{n^2} + \frac{Ck^9m^6}{n^3}\right) n^{2m-k} \binom{2m-k}{t} q^{2m-k-t} \\
&= \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + 3n^{m-1}\binom{m}{\varepsilon m} q^{m-\varepsilon m} + 3\sum_{k=1}^{m-\varepsilon m} \left(\frac{k+1}{n^2} + \frac{Ck^9m^6}{n^3}\right) n^{2m-k} \binom{2m-k}{2\varepsilon m} q^{2(1-\varepsilon)m-k} \\
&\quad + 3\sum_{k=m-\varepsilon m+1}^{m-1} \left(\frac{k+1}{n^2} + \frac{Ck^9m^6}{n^3}\right) n^{2m-k} \binom{2m-k}{m-k+\varepsilon m} q^{m-\varepsilon m} \\
&= \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + 3\left(\frac{1}{n}\binom{m}{\varepsilon m}\left(nq^{1-\varepsilon}\right)^m\right) + 3\left(\frac{1}{n}\binom{m}{\varepsilon m}\left(nq^{1-\varepsilon}\right)^m\right)^2 \\
&\quad \cdot \left(\sum_{k=1}^{m-\varepsilon m} \frac{1}{\binom{m}{\varepsilon m}^2}\left(k+1+\frac{Ck^9m^6}{n}\right) \cdot \frac{\binom{2m-k}{2\varepsilon m}}{n^k q^k} + \sum_{k=m-\varepsilon m+1}^{m-1} \frac{1}{\binom{m}{\varepsilon m}^2}\left(k+1+\frac{Ck^9m^6}{n}\right) \cdot \frac{\binom{2m-k}{m-k+\varepsilon m}}{n^k q^{(1-\varepsilon)m}}\right) \\
&\leqslant \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + 3\mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}] + 3\mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 \cdot \left(\sum_{k=1}^{m-\varepsilon m} \frac{\binom{2m-k}{2\varepsilon m}}{\binom{m}{\varepsilon m}^2}\left(k+1+\frac{Ck^9m^6}{n}\right) \cdot \left(\frac{1}{nq}\right)^k \right. \\
&\quad \left. + \sum_{k=m-\varepsilon m+1}^{m-1} \frac{\binom{2m-k}{m-\varepsilon m}}{\binom{m}{\varepsilon m}^2}\left(k+1+\frac{Ck^9m^6}{n}\right) \cdot \left(\frac{1}{nq}\right)^{(1-\varepsilon)m} \left(\frac{1}{n}\right)^{k-(1-\varepsilon)m}\right)
\end{aligned}$$

我们可以利用斯特林近似验证，对于某个绝对常数 C_2，有 $\frac{\binom{2m-k}{2\varepsilon m}}{\binom{m}{\varepsilon m}^2} \leqslant \frac{\binom{2m}{2\varepsilon m}}{\binom{m}{\varepsilon m}^2} \leqslant C_2\sqrt{m}$；同理，对于 $k \geqslant m - \varepsilon m$，有 $\frac{\binom{2m-k}{m-\varepsilon m}}{\binom{m}{\varepsilon m}^2} \leqslant \frac{\binom{m+\varepsilon m}{m-\varepsilon m}}{\binom{m}{\varepsilon m}^2} = \frac{\binom{m+\varepsilon m}{2\varepsilon m}}{\binom{m}{\varepsilon m}^2} \leqslant \frac{\binom{2m}{2\varepsilon m}}{\binom{m}{\varepsilon m}^2} \leqslant C_2\sqrt{m}$。此外，由于 $k \leqslant m \ll n^{1/15}$，，我们有 $k+1+\frac{Ck^9m^6}{n} \leqslant k+2$。因此，我们得到

$$\leqslant \mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 + 3\mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}] + 3\mathbb{E}_{\mathbf{Q}}[N_{m,\varepsilon}]^2 \cdot C_2 \cdot \left(\sum_{k=1}^{m-\varepsilon m} \sqrt{m}\,(k+2) \cdot \left(\frac{1}{nq}\right)^k\right.$$

$$+\sum_{k=m-\varepsilon m+1}^{m-1}\sqrt{m}\,(k+2)\cdot\left(\frac{1}{nq}\right)^{(1-\varepsilon)m}\left(\frac{1}{n}\right)^{k-(1-\varepsilon)m}\Bigg)$$

$$\leqslant \mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2+O\left(\frac{1}{\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]}+\frac{\sqrt{m}}{nq}\right)\cdot\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2$$

$$\leqslant\left(1+O\left(\frac{1}{n^{\varepsilon m-1}\cdot(nq)}+\frac{\sqrt{m}}{nq}\right)\right)\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2.$$

其中，在倒数第二个不等式中，我们用到了 $nq\gg\sqrt{m}$ 以及 $n^{\varepsilon m-1}\cdot(nq)=\omega(1)$。

特别地，在假设 $nq\gg\sqrt{m}$和 $\varepsilon m\geqslant 1$下，我们有 $\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}^2]\leqslant(1+o(1))\mathbb{E}_{\mathbb{Q}}[N_{m,\varepsilon}]^2$。

□

6.2. **对称低次多项式的稳定性：定理3.9的证明。** 本部分证明植入最短路径问题中对称低次多项式的稳定性参数。回顾一下，在植入最短路径问题的背景下，如果对于任意置换 $f:\pi,\ [n]\to[n]$ 只要该置换固定顶点 1 和 2，就有 $f(G)=f(G_\pi)$，其中 G_π 是由 G 按照置换 π对顶点进行置换后得到的图。换句话说，如果将 G 视为邻接矩阵，那么 $G_\pi=\Pi^\top G\Pi$ 其中 Π是 π的置换矩阵。

在转向证 明之前，我们先定义一些记号，这些记号将在后续证明中派上用场 f.

定义 6.6 （部分标记图）。令 $K\subseteq\mathbf{N}$。我们用 $\mathcal{G}_{\leqslant D}^K$ 表示边数至多为 D 且没有孤立顶点的部分标记图集合，其中每个顶点要么未标记，要么标记为 K 中的一个元素，并且任意两个不同顶点都不会标记为 K 中的同一个顶点。

对于部分标记图 α，我们用 $v(\alpha)$ 表示其未标记顶点的数量。

在我们的证明中，集合 K 将对应于在其间植入一条随机路径的两个特殊顶点 1 和 2。我们还需要部分标记图到完全图 K_n的单射的概念。

定义 6.7。 从部分标记图到完全图 K_n 的单射，是指从部分标记图的顶点集到 $[n]$ 的单射，它将已标记顶点映射到 K中的相应元素，并将未标记顶点映射到 $[n]-K$。给定一个部分标记图 α，我们用 L_α 表示从 α 到 K_n的所有单射的集合。

我们还回顾一下植入最短路径问题中对称多项式的定义，该定义见定义 3.8，为方便读者，现重述如下。

定义 6.8。 若多项式 $f\colon\{0,\ 1\}^{\binom{[n]}{2}}\to\mathbf{R}$ 满足：对于任意固定顶点 1 和 2 的置换 $\pi:[n]\to[n]$，都有 $f(G)=f(G_\pi)$ 作为多项式成立，其中 G_π 是由图 G 按照 π置换顶点后得到的图，则称其为植入最短路径问题的对称多项式。

证明。 令 $K=\{1,2\}$ 表示特殊顶点的集合。设 $f\colon\{0,1\}^{\binom{[n]}{2}}\to\mathbf{R}$ 是一个次数至多为 D 的对称多项式。由于 $\sum_{l\in\mathrm{L}_\alpha}\chi_{l(\alpha)}(G)$ 对于 $\alpha\in\mathcal{G}_{\leqslant D}^K$ 线性独立

并张成对称多项式空间，因此我们可以将 f 唯一地写为

$$f(G) = \sum_{\alpha \in \mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha \sum_{l \in \mathcal{L}_\alpha} \chi_{l(\alpha)}(G), \tag{12}$$

其中 $\chi_S(G)$表示

$$\chi_S(G) = \prod_{\{i,j\} \in S} \frac{G_{i,j} - q}{\sqrt{q(1-q)}}.$$

我们使用以下标准事实： $\chi_S(G)$ 构成测度 $\mathbf{Q} = G(n$下的 $q)$标准正交基，即 $\mathbb{E}_{\mathbf{Q}}[\chi_S(G)\chi_{S'}(G)] = \mathbf{1}\{S = S'\}$。我们还将使用 $\mathbf{P}$ 表示植入分布，正如植入最短路径问题中所定义的那样：观测$\mathbf{H}$ 与从 $G(n$中 $q)$ 抽取的 Erd˝os-R´enyi 图的并集，其中$\mathbf{H}$是连接顶点 1 和 2、长度为 L 的均匀随机路径。在植入测度 $\mathbf{P}$ 下，有

$$\begin{aligned}
\mathbb{E}_{\mathbf{P}}[\chi_S(G)\chi_{S'}(G)] &= \mathbb{E}_{\mathbf{P}}\left[\prod_{\{i,j\} \in S \triangle S'} \frac{G_{i,j} - q}{\sqrt{q(1-q)}} \cdot \prod_{\{i,j\} \in S \cap S'} \frac{(G_{i,j} - q)^2}{q(1-q)}\right] \\
&= \mathbb{E}_{\mathbf{P}}\left[\mathbf{1}\{S \triangle S' \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|S \cap \mathbf{H}|}{2} + \frac{|S' \cap \mathbf{H}|}{2}}\right],
\end{aligned} \tag{13}$$

其中我们利用了以下事实：在给定$\mathbf{H}$的条件下，条件分布变为一个乘积分布；对于 $\{i$中的 $j\} \in \mathbf{H}$，该分布在 1 上是狄拉克测度，而在其他情况下是伯努利分布。在接下来的内容中，期望均相对于 $\mathbf{P}$ 取。

现在，我们有

$$\mathbb{E}[\big(f(G) - f(T_\rho(G))\big)^2] = \mathbb{E}[f(G)^2] + \mathbb{E}[f(T_\rho(G))^2] - 2\mathbb{E}[f(G)f(T_\rho(G))].$$

我们将分别分析这三个项。

首先，根据对称多项式 f在（12）中的展开式，我们有

$$\begin{aligned}
&\mathbb{E}[f(G)^2] \\
&= \sum_{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}[\chi_{l_1(\alpha_1)}(G)\chi_{l_2(\alpha_2)}(G)] \\
&= \sum_{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha) \cap \mathbf{H}|}{2} + \frac{|l_2(\alpha) \cap \mathbf{H}|}{2}}\right] \\
&= \sum_{\alpha \in \mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1, l_2 \in \mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha) \triangle l_2(\alpha) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha) \cap \mathbf{H}|}{2} + \frac{|l_2(\alpha) \cap \mathbf{H}|}{2}}\right] \\
&\quad + \sum_{\substack{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}: \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha) \cap \mathbf{H}|}{2} + \frac{|l_2(\alpha) \cap \mathbf{H}|}{2}}\right],
\end{aligned}$$

其中我们在倒数第二个等式中使用了（13）。

由于 $q \leqslant 1/2,$ ，我们有$\frac{1-q}{q} \geqslant 1$ ，而对 $\alpha_1 = \alpha_2$ 的第一项求和至少为

$$\begin{aligned}
&\sum_{\alpha \in \mathcal{G}^K_{\leqslant D}} \hat{f}^2_\alpha \sum_{l_1, l_2 \in \mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha) \triangle l_2(\alpha) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha)\cap \mathbf{H}|}{2}} \right] \\
\geqslant &\sum_{\alpha \in \mathcal{G}^K_{\leqslant D}} \hat{f}^2_\alpha \sum_{\substack{l_1, l_2 \in \mathcal{L}_\alpha \\ l_1(\alpha) = l_2(\alpha)}} \mathbb{E}\left[\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha)\cap \mathbf{H}|}{2}} \right] \\
\geqslant &(1 - O(1/n)) \sum_{\alpha} \hat{f}^2_\alpha n^{v(\alpha)} |\mathrm{Aut}(\alpha)|,
\end{aligned}$$

其中， $v(\alpha)$ 表示 α中的未标记顶点数。

对非同构的部分标记图 α_1, α_2 的所有配对求和，其上界可由下式给出

$$\begin{aligned}
&\left| \sum_{\substack{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}: \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha_2)\cap \mathbf{H}|}{2}} \right] \right| \\
\leqslant &\sum_{\substack{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}: \\ \alpha_1 \neq \alpha_2}} \left| \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \right| \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha_2)\cap \mathbf{H}|}{2}} \right] \\
= &\sum_{\substack{\alpha_1, \alpha_2 \in \mathcal{G}^K_{\leqslant D}: \\ \alpha_1 \neq \alpha_2}} \left| \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \right| \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq H\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap H|}{2} + \frac{|l_2(\alpha_2)\cap H|}{2}}
\end{aligned}$$

对于$\mathbf{H}$的任意单射所对应的一个固定像 H，由于 f的对称性。现在，我们通过考察像 $l_1(\alpha_1)$和 $l_2(\alpha_2)$与路径 H的交集来对该和求上界。在证明的其余部分，我们令

$$\begin{aligned}
U_H &:= (l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H, I_H := (l_1(\alpha_1) \cap l_2(\alpha_2)) \cap H, \\
s &:= |U_H|, t := |I_H|, v_s := |V(U_H) - K|, \\
v_t &:= |V(l_1(\alpha_1) \cap H) \cap V(l_2(\alpha_2) \cap H) - K|, \\
v_{s,1} &:= |V(l_1(\alpha_1)) \cap V(U_H) - K|, v_{s,2} := |V(l_2(\alpha_2)) \cap V(U_H) - K|, \\
c_1 &:= \mathbf{1}\{1 \in V(U_H)\}, c_2 := \mathbf{1}\{2 \in V(U_H)\}, c_s := cc(U_H) - c_1 - c_2,
\end{aligned}$$

其中， $cc(U_H)$表示U_H的连通分支数。现在，对于参数 $W = (s, t, v_s, v_t, v_{s,}\mathbf{1}, v_{s,}2, c_1, c_2, c_s)$的某种取值，我们用 $\mathbf{L}_{\alpha_1,\alpha_2}(W)$表示所有满足上述参数定义、其像 $l_1(\alpha_1)$和 $l_2(\alpha_2)$满足 $l_1(\alpha_1)\triangle$ $l_2(\alpha_2) \subseteq$的 (l_1, l_2)对

H。那么，我们有

$$\left|\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2}}\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}}\mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq\mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha_2)\cap\mathbf{H}|}{2}}\right]\right|$$

$$\leqslant\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2}}\left|\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\right|\sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}}\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq H\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap H|}{2}+\frac{|l_2(\alpha_2)\cap H|}{2}}$$

$$=\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2}}\left|\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\right|\sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}}\sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)}\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq H\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap H|}{2}+\frac{|l_2(\alpha_2)\cap H|}{2}}$$

$$=\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2}}\left|\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\right|\sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}}\sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)}\left(\frac{1-q}{q}\right)^{\frac{|(l_1(\alpha_1)\cup l_2(\alpha_2))\cap H|}{2}+\frac{|(l_1(\alpha_1)\cap l_2(\alpha_2))\cap H|}{2}}$$

现在回忆一下，$U_H:=(l_1(\alpha_1)\cup l_2(\alpha_2))\cap H$、$I_H:=(l_1(\alpha_1)\cap l_2(\alpha_2))\cap H$，以及 $s:=|U_H|$、$t:=|I_H|$。因此，

$$=\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2}}\left|\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\right|\sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}}\sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)}\left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

$$\leqslant 2\sum_{\alpha_1\in\mathcal{G}_{\leqslant D}^K}\left|\hat{f}_{\alpha_1}\right|\sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}}\sum_{\substack{\alpha_2\in\mathcal{G}_{\leqslant D}^K:\\ \alpha_1\neq\alpha_2,\\ |\hat{f}_{\alpha_2}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}}|\mathrm{Aut}(\alpha_2)|\\ \leqslant|\hat{f}_{\alpha_1}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}}|\mathrm{Aut}(\alpha_1)|}}\left|\hat{f}_{\alpha_2}\right|\sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)}\left(\sqrt{\frac{1-q}{q}}\right)^{s+t},$$

其中，最后一行成立是因为对于任意一对 α_1,α_2，要么 $|\hat{f}_{\alpha_2}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}}|\text{自同构群}(\alpha_2)|\leqslant|\hat{f}_{\alpha_1}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}}|\text{自同构群}(\alpha_1)|$，要么 $|\hat{f}_{\alpha_2}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}}|\text{自同构群}(\alpha_2)|>|\hat{f}_{\alpha_1}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}}|\text{自同构群}(\alpha_1)|$。

对于固定的 $\alpha_1\in\mathcal{G}_{\leqslant D}^K$ 以及 $W=(s,t,v_s,v_t,v_{s,1},v_{s,2},c_1,c_2,c_s)$ 的固定选择，下面考虑有多少个 $(\alpha_2,(l_1,l_2)\in\mathrm{L}_{\alpha_1,\alpha_2}(W))$。我们注意到 W 中这些参数之间存在一些关系。在假设 $2D<L$ 下，我们有

$$v_s=s+c_s \tag{14}$$

$$v_s+v_t\leqslant v_{s,1}+v_{s,2} \tag{15}$$

$$t\leqslant s \tag{16}$$

$$t\leqslant v_t \tag{17}$$

$$0 \leqslant c_s \leqslant s \tag{18}$$

$$1 \leqslant s. \tag{19}$$

我们可以如下验证上述六个关系

- $$\begin{aligned} v_s &= |V(U_H) - K| \\ &= \sum_{\delta \in CC(U_H)} |V(\delta)| - |V(U_H) \cap K| \end{aligned}$$

  由于 H 是一条路径， $U_H = (l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H$ 的每个连通分支 δ 也都是路径，因此 $|V(\delta)| = |\delta| + 1$。因此，

  $$\begin{aligned} &= \sum_{\delta \in CC(U_H)} (|\delta| + 1) - c_1 - c_2 \\ &= |U_H| + |CC(U_H)| - c_1 - c_2 \\ &= s + c_s, \end{aligned}$$

  其中 $CC(U_H)$表示 U_H的连通分支集合。•

$$\begin{aligned} v_s + v_t &= |V(U_H) - K| + |V(l_1(\alpha_1) \cap H) \cap V(l_2(\alpha_2) \cap H) - K| \\ &\leqslant |(V(l_1(\alpha_1)) \cup V(l_2(\alpha_2))) \cap V(U_H) - K| \\ &\quad + |(V(l_1(\alpha_1)) \cap V(l_2(\alpha_2))) \cap V(U_H) - K| \end{aligned}$$

其中不等式成立，是因为对于每个 $v \in V(U_H) = V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H)$，都有 $v \in V(l_1(\alpha_1)) \cup V(l_2(\alpha_2))$；对于 $i \in \{1, 2\}$，并且对于每个 $v \in V(l_i(\alpha_i) \cap H)$，都有 $v \in V(l_i(\alpha_i)) \cap V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) = V(l_i(\alpha_i)) \cap V(U_H)$，

$$\begin{aligned} &= |(V(l_1(\alpha_1)) \cap V(U_H) - K| + |V(l_2(\alpha_2))) \cap V(U_H) - K| \\ &= v_{s,1} + v_{s,2}. \end{aligned}$$

- $$t = |I_H| = |(l_1(\alpha_1) \cap l_2(\alpha_2)) \cap H| \leqslant |(l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H| = |U_H| = s.$$

- $$\begin{aligned} t &= |I_H| \\ &= \sum_{\delta \in CC(I_H)} |\delta| \\ &\leqslant \sum_{\delta \in CC(I_H)} |V(\delta) - K| \end{aligned}$$

因为当 $2D < L$ 时，有 $|l_1(\alpha_1) \cap l_2(\alpha_2)| < L = |H|$，因此 $I_H = (l_1(\alpha_1) \cap l_2(\alpha_2)) \cap H$ 的任何连通分量都不同时包含顶点 1 和 2，并且每个连通分量 $\delta \in \mathbb{C}((l_1(\alpha_1) \cap l_2(\alpha_2)) \cap H)$ 都满足 $|\delta| \leqslant |V(\delta) - K|$。

$$\begin{aligned} &= |V(I_H) - K| \\ &= |V((l_1(\alpha_1) \cap l_2(\alpha_2)) \cap H) - K| \\ &\leqslant |V(l_1(\alpha_1) \cap H) \cap V(l_2(\alpha_2) \cap H) - K| \\ &= v_t \end{aligned}$$

- $c_s \leqslant s$ 显然成立，因为 U_H 的连通分支数至多为 $|U_H| = s$。另一方面，$c_s \geqslant 0$ 也很容易验证，因为当 $2D < L$时，$U_H = (l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H$ 的任何连通分支都不同时包含顶点 1 和 2，而 $c_s = cc(U_H) - c_1 - c_2$ 是 U_H 中不包含顶点 1 或 2 的连通分支数。• 由于 $\alpha_1 \neq \alpha_2$，我们知道 $l_1(\alpha_1) \triangle l_2(\alpha_2) \neq \varnothing$。由于 $l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq H$，因此有$s = |U_H| = |(l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H| \geqslant 1$。

由于 H 是从 1 到 2、长度为 L 的路径，我们知道 $U_H = (l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H$ 是若干条顶点不相交路径的并。当 $c_s \geqslant 0$时，要枚举参数为W的 $U_H \subseteq H$ 的数量，我们可以枚举 U_H 中除包含 1 或 2 的连通分支之外的每个连通分支的起始顶点，共有 c_s 个分支；还可以枚举这 c_s 个分支各自的长度，以及包含顶点 1 或 2 的 $c_1 + c_2$ 个分支的长度。枚举每个分支的起始顶点，最多有 $\binom{L}{c_s}$ 种选择。根据标准插板法论证，枚举各分支的长度有 $\binom{s-1}{c_s+c_1+c_2-1}$ 种选择，因为 U_H 各连通分支的长度集合必须满足以下方程组：

$$\sum_{i=1}^{c_s} x_i + c_1 \cdot y + c_2 \cdot z = |U_H| = s,$$

$$x_i, y, z \in \mathbb{Z}^+,$$

其中，x_i 是那些不包含顶点 1 或 2 的连通分量的长度；若 $c_1 = 1$，则 y 是包含顶点 1 的连通分量的长度；若 $c_2 = 1$，则 z 是包含顶点 2 的连通分量的长度。注意，$\binom{s-1}{c_s+c_1+c_2-1}$ 是定义良好的，因为 $s \geqslant 1$，且 $c_s + c_1 + c_2 = cc(U_H) \geqslant 1$。结合连通分量起始顶点的 $\binom{L}{c_s}$ 种选择和连通分量长度的 $\binom{s-1}{c_s+c_1+c_2-1}$ 种选择，我们得出，给定 W中的参数，$L^{c_s} 2^{s-1}$ 是 U_H 的选择数的一个有效上界。

现在，对于固定的 U_H，我们继续枚举满足 W 和 $(l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H = U_H$中的参数条件的 $(l_1, l_2) \in \mathbf{L}_{\alpha_1, \alpha_2}(W)$的数量：

- 我们首先枚举满足 $S_1 = l_1(\alpha_1) \cap H$ 和 $S_2 = l_2(\alpha_2) \cap H$ 的 S_1、S_2 这一对。由于 $S_1 \cup S_2 = U_H$，我们知道 U_H 的每条边都必须属于 S_1、S_2，或同时属于两者。因此，$(S_1$、$S_2)$ 的选择至多有 $3^{|U_H|} = 3^s$ 种。• 接着，我们枚举满足 V_S、$_1 = V(l_1(\alpha_1)) \cap V(U_H) - K$ 以及 V_S、$_2 = V(l_2(\alpha_2)) \cap V(U_H) - K$ 的 V_S,1、V_S,2 这一对。由于 V_S、$_1 \cup V_S$、$_2 = V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) - K$ =$V(U_H) - K$，$V(U_H) - K$ 的每个顶点都必须属于 V_S,1、V_S,2，或同时属于两者。因此，$(V_S$,1、$V_{S2})$ 的选择至多有 $3^{|V(U_H)-K|} = 3^{v_s}$ 种。

- 接下来，对于固定的 S_1、 S_2、 $V_{S,}1$ 和 $V_{S,}2$，我们开始枚举满足 W 和 $S_1 = l_1(\alpha_1) \cap H$ 中的参数以及 $S_2 = l_2(\alpha_2) \cap H$、$V_{S,1} = V(l_1(\alpha_1)) \cap V(U_H) - K$ 和 $V_{S,2} = V(l_2(\alpha_2)) \cap V(U_H) - K$ 的 $(l_1, l_2) \in \mathrm{L}_{\alpha_1,\alpha_2}(W)$ 的数量。

  注意到，由于 $l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq H$，我们有 $l_1(\alpha_1) \cup l_2(\alpha_2) - H \subseteq l_1(\alpha_1) \cap l_2(\alpha_2)$，因此必然有 $V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) - K = V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) - K$ 中的每个顶点都同时属于 $V(l_1(\alpha_1))$ 和$V(l_2(\alpha_2))$。因此，若记 $v_o = |V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) - K|$，则有

$$\begin{aligned} v(\alpha_1) &= |V(l_1(\alpha_1)) - K| \\ &= |V(l_1(\alpha_1)) - V(U_H) - K| + |V(l_1(\alpha_1)) \cap V(U_H) - K| \\ &= |V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) - K| + |V(l_1(\alpha_1)) \cap V(U_H) - K| \\ &= v_o + v_{s,1}. \end{aligned}$$

  类似地，

$$v(\alpha_2) = v_o + v_{s,2}.$$

  因此，我们有

$$v_o = \frac{v(\alpha_1) + v(\alpha_2) - v_{s,1} - v_{s,2}}{2}. \tag{20}$$

  随后，我们可以枚举集合 $V_O \subseteq [n]$，使得 $V_O = V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) - K$，这至多给出 $\binom{n}{v_o} = \binom{n}{\frac{v(\alpha_1)+v(\alpha_2)-v_{s,\ 1}-v_{s,\ 2}}{2}}$ 种选择。

- 回忆一下 $l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq H$。因此，$l_2(\alpha_2) - l_1(\alpha_1) \subseteq (l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H$，并且

$$\begin{aligned} &V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) \\ &= V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) \\ &= V(l_1(\alpha_1)) \cup V(l_2(\alpha_2) - l_1(\alpha_1)) - V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) \\ &= V(l_1(\alpha_1)) - V((l_1(\alpha_1) \cup l_2(\alpha_2)) \cap H) \\ &= V(l_1(\alpha_1)) - V(U_H). \end{aligned}$$

  现在，对于固定的 S_1， S_2, $V_{S,}\ 1$, $V_{S,}\ 2$ 和 V_O，我们枚举满足以下条件的 $l_1(\alpha_1)$ 的数量

$$\begin{aligned} V_O \sqcup V_{S,1} &= [V(l_1(\alpha_1) \cup l_2(\alpha_2)) - V(U_H) - K] \sqcup [V(l_1(\alpha_1)) \cap V(U_H) - K] \\ &= [V(l_1(\alpha_1)) - V(U_H) - K] \sqcup [V(l_1(\alpha_1)) \cap V(U_H) - K] \\ &= V(l_1(\alpha_1)) - K. \end{aligned}$$

  换句话说，当 V_O 和 $V_{S,}\ 1$ 固定时，像 $l_1(\alpha_1)$ 的顶点集 $V(l_1(\alpha_1))$ 是固定的，而 $l_1(\alpha_1)$ 的数量有界于

$$\frac{v(\alpha_1)!}{|\mathrm{Aut}(\alpha_1)|} \leqslant \frac{v(\alpha_1)^{v_{s,1}} v_o!}{|\mathrm{Aut}(\alpha_1)|},$$

  由于 $v(\alpha_1)!$ 是从 $V(\alpha_1) - K$到 $V_O \sqcup V_S$ 的单射数量，1，且考虑到 α_1 的自同构群，给定 V_S 的不同 $l_1(\alpha_1)$ 数量，1，V_O 至多为 $\frac{v(\alpha_1)!}{\text{自同构群}(\alpha_1)|}$。在该不等式中，我们利用了 ~~$v(\alpha_1) =$~~ $v_o + v_{s,}\ 1$，因为如上所示，$V_O \sqcup V_{S,}\ _1 = V(l_1(\alpha_1)) - K$。

• 此外，我们注意到，一旦确定了 $l_1(\alpha_1)$，S_1，S_2，$l_2(\alpha_2)$ 也就确定了，因为根据条件 $l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq H$，以及 $S_1 = l_1(\alpha_1) \cap H$，$S_2 = l_2(\alpha_2) \cap H$，我们知道 $l_1(\alpha_1) - S_1 = l_2(\alpha_2) - S_2 = (l_1(\alpha_1) \cap l_2(\alpha_2)) - H$。因此，$\alpha_2$ 也由 $l_1(\alpha_1)$，S_1，S_2 确定。如果

$$|\hat{f}_{\alpha_2}| \left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_2)| > |\hat{f}_{\alpha_1}| \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}} |\mathrm{Aut}(\alpha_1)|,$$

那么，此 α_2 不参与求和，我们中止当前枚举并回溯。否则，以下结论成立：

$$|\hat{f}_{\alpha_2}| \left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_2)| \leqslant |\hat{f}_{\alpha_1}| \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}} |\mathrm{Aut}(\alpha_1)|.$$

• 对于固定的 $l_1(\alpha_1)$，$l_2(\alpha_2)$，我们完成对满足条件的 $(l_1, l_2) \in \mathrm{L}_{\alpha_1, \alpha_2}(W)$ 的数量的枚举，这些条件使其像为 $l_1(\alpha_1)$ 和 $l_2(\alpha_2)$。该数量有如下上界：

$$|\mathrm{Aut}(\alpha_1)||\mathrm{Aut}(\alpha_2)| \leqslant \frac{|\hat{f}_{\alpha_1}|}{|\hat{f}_{\alpha_2}|} \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)|^2.$$

因此，对于固定的 $l_1(\alpha_1)$，$l_2(\alpha_2)$ 以及 W 中参数的固定取值，对总和的贡献至多为

$$\begin{aligned}
&2\left|\hat{f}_{\alpha_1}\right|\left|\hat{f}_{\alpha_2}\right| |\mathrm{Aut}(\alpha_1)||\mathrm{Aut}(\alpha_2)| \left(\sqrt{\frac{1-q}{q}}\right)^{s+t} \\
\leqslant\ &2\left|\hat{f}_{\alpha_1}\right|\left|\hat{f}_{\alpha_2}\right| \frac{|\hat{f}_{\alpha_1}|}{|\hat{f}_{\alpha_2}|} \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)|^2 \left(\sqrt{\frac{1-q}{q}}\right)^{s+t} \\
=\ &2\hat{f}_{\alpha_1}^2 \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)|^2 \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}
\end{aligned}$$

总之，结合 $U_H, S_1, S_2, V_{S,1}, V_{S,2}, V_O, l_1(\alpha_1), l_2(\alpha_2)$ 的选择数量，我们得出结论：所有 $(\alpha_2, (l_1, l_2) \in \mathrm{L}_{\alpha_1,\alpha_2}(W))$ 对总和的贡献至多为

$$\begin{aligned}
&\underbrace{L^{c_s} 2^{s-1}}_{U_H} \underbrace{3^s}_{S_1,S_2} \underbrace{3^{v_s}}_{V_{S,1},V_{S,2}} \underbrace{\binom{n}{\frac{v(\alpha_1)+v(\alpha_2)-v_{s,1}-v_{s,2}}{2}}}_{V_O} \underbrace{\frac{v(\alpha_1)^{v_{s,1}} v_o!}{|\mathrm{Aut}(\alpha_1)|}}_{l_1(\alpha_1),l_2(\alpha_2)} \\
&\underbrace{\cdot\, 2\hat{f}_{\alpha_1}^2 \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)|^2 \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}}_{\substack{\text{contribution given fixed} \\ U_H,S_1,S_2,V_{S,1},V_{S,2},V_O,l_1(\alpha_1),l_2(\alpha_2)}} \\
\leqslant\ &2L^{c_s} 2^{s-1} 3^s 3^{v_s} \left(\binom{n}{\frac{v(\alpha_1)+v(\alpha_2)-v_{s,1}-v_{s,2}}{2}} v_o!\right) v(\alpha_1)^{v_{s,1}} \\
&\cdot\, \hat{f}_{\alpha_1}^2 \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)| \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}
\end{aligned}$$

$$\leqslant L^{c_s}6^s 3^{v_s} n^{\frac{v(\alpha_1)+v(\alpha_2)-v_{s,1}-v_{s,2}}{2}} v(\alpha_1)^{v_{s,1}} \hat{f}_{\alpha_1}^2 \left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)-v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_1)| \left(\sqrt{\frac{1-q}{q}}\right)^{s+t},$$

其中我们用到了 $v_o = \frac{v(\alpha_1)+v(\alpha_2)-v_{s,1}-v_{s,2}}{2}$,

$$= L^{c_s}6^s 3^{v_s} n^{v(\alpha_1)-\frac{v_{s,1}+v_{s,2}}{2}} v(\alpha_1)^{v_{s,1}} (2D)^{\frac{v(\alpha_2)-v(\alpha_1)}{2}} \hat{f}_{\alpha_1}^2 |\mathrm{Aut}(\alpha_1)| \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

$$\leqslant \hat{f}_{\alpha_1}^2 n^{v(\alpha_1)} |\mathrm{Aut}(\alpha_1)| \cdot L^{c_s}6^s 3^{v_s} n^{-\frac{v_{s,1}+v_{s,2}}{2}} (2D)^{v_{s,1}} (2D)^{\frac{v_{s,2}-v_{s,1}}{2}} \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

因为 $v(\alpha_1) = |V(l_1(\alpha_1)) - K| \leqslant 2D$ 并且 $v(\alpha_1) = v_o + v_{s,1}$, $v(\alpha_2) = v_o + v_{s,2}$,

$$= \hat{f}_{\alpha_1}^2 n^{v(\alpha_1)} |\mathrm{Aut}(\alpha_1)| \cdot L^{c_s}6^s 3^{v_s} n^{-\frac{v_{s,1}+v_{s,2}}{2}} (2D)^{\frac{v_{s,1}+v_{s,2}}{2}} \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

$$= \hat{f}_{\alpha_1}^2 n^{v(\alpha_1)} |\mathrm{Aut}(\alpha_1)| \cdot \left[\left(6\sqrt{\frac{1-q}{q}}\right)^s \left(3\sqrt{\frac{2D}{n}}\right)^{v_s} L^{c_s}\right] \cdot \left[\left(\sqrt{\frac{1-q}{q}}\right)^t \left(\sqrt{\frac{2D}{n}}\right)^{v_t}\right]$$
$$\cdot \left[\left(\sqrt{\frac{2D}{n}}\right)^{v_{s,1}+v_{s,2}-v_s-v_t}\right]$$

其中我们注意到 $v_{s,1} + v_{s,2} \geqslant v_s + v_t$。我们进一步使用 $v_s = s + c_s$, $v_t \geqslant t$ 并得到

$$= \hat{f}_{\alpha_1}^2 n^{v(\alpha_1)} |\mathrm{Aut}(\alpha_1)| \cdot \left[\left(18\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^s \left(3L\sqrt{\frac{2D}{n}}\right)^{c_s}\right] \cdot \left[\left(\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^t \left(\sqrt{\frac{2D}{n}}\right)^{v_t-t}\right]$$
$$\cdot \left[\left(\sqrt{\frac{2D}{n}}\right)^{v_{s,1}+v_{s,2}-v_s-v_t}\right].$$

Fin此外，对于每个固定的 α_1，我们对所有可能的参数选择求和，并且 t

$$2 \sum_{\alpha_1 \in \mathcal{G}_{\leqslant D}^K} \left|\hat{f}_{\alpha_1}\right| \sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}} \sum_{\substack{\alpha_2 \in \mathcal{G}_{\leqslant D}^K:\\ \alpha_1 \neq \alpha_2,\\ |\hat{f}_{\alpha_2}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}} |\mathrm{Aut}(\alpha_2)| \\ \leqslant |\hat{f}_{\alpha_1}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}} |\mathrm{Aut}(\alpha_1)|}} \left|\hat{f}_{\alpha_2}\right| \sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)} \left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

$$\leqslant \sum_{\alpha_1 \in \mathcal{G}_{\leqslant D}^K} \hat{f}_{\alpha_1}^2 n^{v(\alpha_1)} |\mathrm{Aut}(\alpha_1)| \cdot \sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}} \left[\left(18\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^s \left(3L\sqrt{\frac{2D}{n}}\right)^{c_s}\right]$$

$$\cdot \left[\left(\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^{t}\left(\sqrt{\frac{2D}{n}}\right)^{v_t-t}\right]\cdot\left[\left(\sqrt{\frac{2D}{n}}\right)^{v_{s,1}+v_{s,2}-v_s-v_t}\right]$$

$$\leqslant \sum_{\alpha_1\in\mathcal{G}^K_{\leqslant D}} \hat{f}^2_{\alpha_1} n^{v(\alpha_1)}|\mathrm{Aut}(\alpha_1)|\cdot\sum_{s\geqslant 1}\left(18\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^{s}\sum_{c_s=0}^{s}\left(3L\sqrt{\frac{2D}{n}}\right)^{c_s}$$

$$\cdot\sum_{t=0}^{s}\left(\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^{t}\sum_{v_t\geqslant t}\left(\sqrt{\frac{2D}{n}}\right)^{v_t-t}\sum_{i\geqslant 0}\left(\sqrt{\frac{2D}{n}}\right)^{i}$$

$$\cdot \#\{\text{feasible } W \text{ given } v(\alpha_1), s, c_s, t, v_t \text{ and } v_{s,1}+v_{s,2},\ v_{s,1}+v_{s,2}-v_s-v_t=i\}$$

其中在最后一步中，我们使用了 $v_{s,1}+v_{s,2}\geqslant v_s+v_t$，以及对于 $\alpha_1\neq\alpha_2$有 $s=|U_H|=|(l_1(\alpha_1)\cup l_2(\alpha_2))\cap H|\geqslant|l_1(\alpha_1)\triangle l_2(\alpha_2)|\geqslant 1$。给定 $v(\alpha_1)$、s、c_s、t、v_t以及$v_{s,1}+v_{s,2}-v_s-v_t=i$，可行的$W$是指在 $W=(s,t,v_s,v_t,v_{s,1},v_{s,2},c_1,c_2,c_s)$中选择参数，使得存在 α_2和 $(l_1,l_2)\in\mathrm{L}_{\alpha_1,\alpha_2}(W)$。我们有

$$c_1,c_2\in\{0,1\}$$

$$v_s=s+c_s$$

$$v_{s,1}+v_{s,2}=v_s+v_t+i.$$

因此，v_s由 s和 c_s确定；对于 c_1，c_2，最多有 4 种选择；对于$v_{s,1}$，$v_{s,2}$，最多有 v_s+v_t+i+1种选择。总而言之，给定 $v(\alpha_1)$，s，c_s，t，v_t和 i时，W的可行选择数至多为 $4(v_s+v_t+i+1)=4(s+c_s+v_t+i+1)$。因此，我们可以估计

$$2\sum_{\alpha_1\in\mathcal{G}^K_{\leqslant D}}\left|\hat{f}_{\alpha_1}\right|\sum_{\substack{W=(s,t,v_s,v_t,\\ v_{s,1},v_{s,2},c_1,c_2,c_s)}}\ \sum_{\substack{\alpha_2\in\mathcal{G}^K_{\leqslant D}:\\ \alpha_1\neq\alpha_2,\\ |\hat{f}_{\alpha_2}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_2)}{2}}|\mathrm{Aut}(\alpha_2)|\\ \leqslant|\hat{f}_{\alpha_1}|\left(\frac{n}{2D}\right)^{\frac{v(\alpha_1)}{2}}|\mathrm{Aut}(\alpha_1)|}}\left|\hat{f}_{\alpha_2}\right|\sum_{(l_1,l_2)\in\mathcal{L}_{\alpha_1,\alpha_2}(W)}\left(\sqrt{\frac{1-q}{q}}\right)^{s+t}$$

$$\leqslant\sum_{\alpha_1\in\mathcal{G}^K_{\leqslant D}}\hat{f}^2_{\alpha_1}n^{v(\alpha_1)}|\mathrm{Aut}(\alpha_1)|\cdot\sum_{s\geqslant 1}\left(18\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^{s}\sum_{c_s=0}^{s}\left(3L\sqrt{\frac{2D}{n}}\right)^{c_s}\sum_{t=0}^{s}\left(\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\right)^{t}$$

$$\cdot\sum_{v_t\geqslant t}\left(\sqrt{\frac{2D}{n}}\right)^{v_t-t}\sum_{i\geqslant 0}\left(\sqrt{\frac{2D}{n}}\right)^{i}\cdot 4(s+c_s+v_t+i+1),$$

只要满足 $18\sqrt{\frac{2D}{n}\cdot\frac{1-q}{q}}\leqslant 1-\varepsilon$ 且 $3L\sqrt{\frac{2D}{n}}\leqslant 1-\varepsilon$，对于某个常数 $\varepsilon>0$，就有 $O\left(\sqrt{\frac{D(1-q)}{nq}}\cdot\sqrt{\sum_{\alpha\in\mathcal{G}^K_{\leqslant D}}\hat{f}^2_{\alpha}n^{v(\alpha)}|\text{自同构群}(\alpha)|}\right)$。

因此，在假设 $\frac{D(1-q)}{nq}=o(1)$、$2D<L$ 和 $L^2\frac{D}{n}=o(1)$ 成立的情况下，我们有

$$\mathbb{E}[f(G)^2]$$

$$= \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$+ \sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}^K_{\leqslant D}:\\ \alpha_1\neq\alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$\leqslant \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$+ O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\cdot \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 n^{v(\alpha)}|\mathrm{Aut}(\alpha)|$$

$$\leqslant \left(1+O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\cdot \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right],$$

而且我们还有

$$\mathbb{E}[f(G)^2]$$

$$\geqslant \left(1-O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\cdot \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

类似地，在相同条件下，我们可以给交叉项的期望取一个下界：

$$\mathbb{E}[f(G)f(T_\rho(G))]$$

$$= \sum_{\alpha_1,\alpha_2\in\mathcal{G}^K_{\leqslant D}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}(1-\rho)^{|\alpha_2|} \sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}} \mathbb{E}\left[\chi_{l_1(\alpha_1)}(G)\chi_{l_2(\alpha_2)}(T_\rho(G))\right]$$

$$= \sum_{\alpha_1,\alpha_2\in\mathcal{G}^K_{\leqslant D}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}(1-\rho)^{|\alpha_2|} \sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$= \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2(1-\rho)^{|\alpha|} \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$+ \sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}^K_{\leqslant D}:\\ \alpha_1\neq\alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}(1-\rho)^{|\alpha_2|} \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$\geqslant (1-\rho)^D \sum_{\alpha\in\mathcal{G}^K_{\leqslant D}} \hat{f}_\alpha^2 \sum_{l_1,l_2\in\mathcal{L}_\alpha} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq \mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$
\begin{aligned}
&- O\left(\sqrt{\frac{D(1-q)}{nq}}\right) \cdot \sum_{\alpha \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha}^{2} n^{v(\alpha)} |\mathrm{Aut}(\alpha)| \\
\geqslant & \left((1-\rho)^{D} - O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right) \sum_{\alpha \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha}^{2} \\
&\cdot \sum_{l_1, l_2 \in \mathcal{L}_{\alpha}} \mathbb{E}\left[\mathbf{1}\{l_1(\alpha) \triangle l_2(\alpha) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha)\cap \mathbf{H}|}{2}}\right] \\
\geqslant & \left((1-\rho)^{D} - O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right) \mathbb{E}[f(G)^2]. \qquad (21)
\end{aligned}
$$

在相同条件下，且 $q \leqslant 1/2,$，我们可以对噪声算子下的二阶矩取一个上界：

$$
\begin{aligned}
&\mathbb{E}[f(T_\rho(G))^2] \\
= & \sum_{\alpha_1, \alpha_2 \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} (1-\rho)^{|\alpha_2|} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} \mathbb{E}\left[\chi_{l_1(\alpha_1)}(T_\rho(G)) \chi_{l_2(\alpha_2)}(T_\rho(G))\right] \\
= & \sum_{\alpha_1, \alpha_2 \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} (1-\rho)^{|l_1(\alpha_1) \triangle l_2(\alpha_2)|} \\
&\cdot \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1) \triangle l_2(\alpha_2)|}{2}} \left(\rho + (1-\rho)\left(\frac{1-q}{q}\right)\right)^{|l_1(\alpha_1)\cap l_2(\alpha_2)\cap \mathbf{H}|}\right] \\
= & \sum_{\alpha \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha}^{2} \sum_{l_1, l_2 \in \mathcal{L}_{\alpha}} (1-\rho)^{|l_1(\alpha) \triangle l_2(\alpha)|} \\
&\cdot \mathbb{E}\left[\mathbf{1}\{l_1(\alpha) \triangle l_2(\alpha) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha) \triangle l_2(\alpha)|}{2}} \left(\rho + (1-\rho)\left(\frac{1-q}{q}\right)\right)^{|l_1(\alpha)\cap l_2(\alpha)\cap \mathbf{H}|}\right] \\
&+ \sum_{\substack{\alpha_1, \alpha_2 \in \mathcal{G}_{\leqslant D}^{K}: \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \sum_{l_1 \in \mathcal{L}_{\alpha_1}, l_2 \in \mathcal{L}_{\alpha_2}} (1-\rho)^{|l_1(\alpha_1) \triangle l_2(\alpha_2)|} \\
&\cdot \mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1) \triangle l_2(\alpha_2) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1) \triangle l_2(\alpha_2)|}{2}} \left(\rho + (1-\rho)\left(\frac{1-q}{q}\right)\right)^{|l_1(\alpha_1)\cap l_2(\alpha_2)\cap \mathbf{H}|}\right] \\
\leqslant & \sum_{\alpha \in \mathcal{G}_{\leqslant D}^{K}} \hat{f}_{\alpha}^{2} \sum_{l_1, l_2 \in \mathcal{L}_{\alpha}} (1-\rho)^{|l_1(\alpha) \triangle l_2(\alpha)|} \\
&\cdot \mathbb{E}\left[\mathbf{1}\{l_1(\alpha) \triangle l_2(\alpha) \subseteq \mathbf{H}\} \left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap \mathbf{H}|}{2} + \frac{|l_2(\alpha)\cap \mathbf{H}|}{2}}\right]
\end{aligned}
$$

$$+\sum_{\substack{\alpha_1,\alpha_2\in\mathcal{G}^K_{\leqslant D}:\\ \alpha_1\neq\alpha_2}}|\hat{f}_{\alpha_1}\hat{f}_{\alpha_2}|\sum_{l_1\in\mathcal{L}_{\alpha_1},l_2\in\mathcal{L}_{\alpha_2}}(1-\rho)^{|l_1(\alpha_1)\triangle l_2(\alpha_2)|}$$

$$\cdot\mathbb{E}\left[\mathbf{1}\{l_1(\alpha_1)\triangle l_2(\alpha_2)\subseteq\mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha_1)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

其中我们利用了 $\rho+(1-\rho)\left(\frac{1-q}{q}\right)\leqslant\frac{1-q}{q}$ 若 $q\leqslant\frac{1}{2}$,

$$\leqslant\sum_{\alpha\in\mathcal{G}^K_{\leqslant D}}\hat{f}_\alpha^2\sum_{l_1,l_2\in\mathcal{L}_\alpha}\mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq\mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$+O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\cdot\sum_{\alpha\in\mathcal{G}^K_{\leqslant D}}\hat{f}_\alpha^2 n^{v(\alpha)}|\mathrm{Aut}(\alpha)|$$

$$\leqslant\left(1+O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\sum_{\alpha\in\mathcal{G}^K_{\leqslant D}}\hat{f}_\alpha^2\sum_{l_1,l_2\in\mathcal{L}_\alpha}\mathbb{E}\left[\mathbf{1}\{l_1(\alpha)\triangle l_2(\alpha)\subseteq\mathbf{H}\}\left(\frac{1-q}{q}\right)^{\frac{|l_1(\alpha)\cap\mathbf{H}|}{2}+\frac{|l_2(\alpha)\cap\mathbf{H}|}{2}}\right]$$

$$\leqslant\left(1+O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\mathbb{E}[f(G)^2]. \tag{22}$$

结合（21）和（22），得到

$$\begin{aligned}&\mathbb{E}[(f(G)-f(T_\rho(G)))^2]\\&=\mathbb{E}[f(G)^2]+\mathbb{E}[f(T_\rho(g))^2]-2\mathbb{E}[f(G)f(T_\rho(G))]\\&\leqslant 2\left(1-(1-\rho)^D+O\left(\sqrt{\frac{D(1-q)}{nq}}\right)\right)\mathbb{E}[f(G)^2].\end{aligned}$$

□

7. 随机线性码的证明

7.1. 含噪 MMSE：定理 3.16的证明。 本节将证明随机线性码问题的含噪 MMSE 下界。

回顾一下，在无噪声情形中，$A\in\{0,1\}^{m\times n}$ 和 $x\in\{0,1\}^n$ 均为均匀随机采样，且 $y=Ax$ 模 2。在带噪版本中，T_ρ 以概率 ρ 独立地将 y 的每个坐标重新从伯努利分布$(1/2)$中采样。

为方便起见，我们在下文重新表述定理 3.16 。

定理 7.1 (定理 3.16的重述。 对于任意函数 $f(n)=\omega(1)$ 和 $m=n+f(n)$，若 $\frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m}=\omega\left(2^{2(m-n)}\right)$，则任何用于 *RLC* 的算法 $\mathcal{A}\colon\{0,1\}^{m\times n}\times\{0,1\}^m\to\mathbb{R}^n$ 具有至少

均方误差

$$\mathbb{E}\left[\|\mathcal{A}(A, T_\rho(y)) - x\|_2^2\right] \geqslant \left(\frac{1}{4} - o(1)\right) n.$$

证明。 对于 $x \in \{0,\ 1\}^n$，我们将用 $w(x)$ 表示 x的权重，即 x中 1 的个数。记作 $\hat{}$ $y = T_\rho(y)$。注意，对于 y的每个坐标，我们都有

$$\hat{y}_i = \begin{cases} y_i & \text{with probability } 1 - \frac{\rho}{2}, \\ 1 - y_i & \text{with probability } \frac{\rho}{2}. \end{cases}$$

由于 x 的先验分布在 $\{0,\ 1\}^n$上均匀分布，给定 A 和 $\hat{}y$后， $\{0,\ 1\}^n$上的后验分布为

$$\Pr(x|A, \hat{y}) \propto \left(\frac{\frac{\rho}{2}}{1 - \frac{\rho}{2}}\right)^{w(Ax - \hat{y})} = \left(\frac{\rho}{2 - \rho}\right)^{w(Ax - \hat{y})}.$$

要使均方误差最小，最优算法应输出后验均值。要使平均汉明距离最小，最优算法应在每个坐标输出使边缘后验概率最大的 x。

我们先来理解边缘后验概率。不失一般性，考察坐标 1。定义 $S_0 := \{x \in \{0,\ 1\}^n :\ x_1 = 0\}$ 以及 $S_1 := \{x \in \{0,\ 1\}^n :\ x_1 = 1\}$。我们有

$$\Pr(x_1 = 0|A, \hat{y}) \propto \sum_{x \in S_0} \left(\frac{\rho}{2 - \rho}\right)^{w(Ax - \hat{y})},$$

$$\Pr(x_1 = 1|A, \hat{y}) \propto \sum_{x \in S_1} \left(\frac{\rho}{2 - \rho}\right)^{w(Ax - \hat{y})}.$$

为方便起见，记 $L_0 := \sum_{x \in S_0} \left(\frac{\rho}{2-\rho}\right)^{w(Ax - \hat{y})}$ 和 $L_1 := \sum_{x \in S_1} \left(\frac{\rho}{2-\rho}\right)^{w(Ax - \hat{y})}$。于是，

$$\Pr(x_1 = 0|A, \hat{y}) = \frac{L_0}{L_0 + L_1}$$
$$\Pr(x_1 = 1|A, \hat{y}) = \frac{L_1}{L_0 + L_1}.$$

我们将证明，这些边缘后验概率以高概率集中在 1 2 附近。回顾一下，$\hat{}$ - $y = T_\rho(y) = T_\rho(Ax^*)$，其中用 x^* 表示真实消息。此外，将 T_ρ 应用于 y 等价于加上一个向量 $z \in \{0,\ 1\}^m$，其各分量服从独立同分布的伯努利分布$(\rho/2)$。因此，我们有

$$\begin{aligned} &\mathbb{E}[L_0] \\ &= \sum_{x \in S_0} \mathbb{E}_{A, x^*}\left[\left(\frac{\rho}{2 - \rho}\right)^{w(Ax - T_\rho(Ax^*))}\right] \end{aligned}$$

$$= \sum_{x\in S_0} \mathop{\mathbb{E}}_{\substack{A,x^*,\\ z\sim \mathrm{Bern}(\rho/2)^{\otimes m}}} \left[\left(\frac{\rho}{2-\rho}\right)^{w(A(x-x^*)-z)} \right]$$

注意，当 $x - x^* \neq \mathbf{0}$时， $A(x-x^*)$ 是 $\mathbb{F}_2^m$中的均匀随机向量，因为 $A \in \{0,\ 1\}^{m\times n}$是均匀随机的。此外，任意随机向量与 $\mathbb{F}_2^m$中的均匀随机向量之和，仍是均匀随机向量。因此我们得到

$$\begin{aligned}
&= \sum_{x\in S_0} \left(\left(1-2^{-n}\right) \cdot \mathop{\mathbb{E}}_{u\sim \mathrm{Unif}(\mathbb{F}_2^m)} \left[\left(\frac{\rho}{2-\rho}\right)^{w(u)} \right] + 2^{-n} \cdot \mathop{\mathbb{E}}_{z\sim \mathrm{Bern}(\rho/2)^{\otimes m}} \left[\left(\frac{\rho}{2-\rho}\right)^{w(z)} \right] \right) \\
&= 2^{n-1} \cdot \left(\left(1-2^{-n}\right) \cdot \left[\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \frac{\rho}{2-\rho} \right]^m + 2^{-n} \cdot \left[\left(1-\frac{\rho}{2}\right)\cdot 1 + \frac{\rho}{2}\cdot \frac{\rho}{2-\rho} \right]^m \right) \\
&= \frac{1}{2} \cdot \left[(2^n-1) \left(\frac{1}{2-\rho}\right)^m + \left(\frac{(2-\rho)^2+\rho^2}{2(2-\rho)}\right)^m \right].
\end{aligned}$$

同理，

$$\mathbb{E}[L_1] = \mathbb{E}[L_0].$$

我们也可以计算

$$\begin{aligned}
&\mathbb{E}[L_0^2] \\
&= \sum_{x,x'\in S_0} \mathop{\mathbb{E}}_{\substack{A,x^*,\\ z\sim \mathrm{Bern}(\rho/2)^{\otimes m}}} \left[\left(\frac{\rho}{2-\rho}\right)^{w(A(x-x^*)-z)} \left(\frac{\rho}{2-\rho}\right)^{w(A(x'-x^*)-z)} \right] \\
&= \sum_{x\in S_0} \mathop{\mathbb{E}}_{\substack{A,x^*,\\ z\sim \mathrm{Bern}(\rho/2)^{\otimes m}}} \left[\left(\frac{\rho}{2-\rho}\right)^{2\cdot w(A(x-x^*)-z)} \right] \\
&\quad + \sum_{\substack{x,x'\in S_0:\\ x\neq x'}} \mathop{\mathbb{E}}_{\substack{A,x^*,\\ z\sim \mathrm{Bern}(\rho/2)^{\otimes m}}} \left[\left(\frac{\rho}{2-\rho}\right)^{w(A(x-x^*)-z)} \left(\frac{\rho}{2-\rho}\right)^{w(A(x'-x^*)-z)} \right]
\end{aligned}$$

回想一下，根据 $x = x^*$ 是否成立， $A(x-x^*)-z$ 服从$(\mathbb{F}_2)^{\otimes m}$ 均匀分布或伯努利分布$(\rho/2)^{\otimes m}$。此外，容易看出，对于 $x \neq x'$ 和固定的 x^*，当随机变量 $A \sim$服从 ($\{0$均匀分布， $1\}^{m\times n}$) 以及伯努利分布 $z \sim \mathrm{Bern}(\rho/2)^{\otimes m}$时， $A(x-x^*)-z$和$A(x'-x^*)-z$ 的分布相互独立。因此，

$$\begin{aligned}
&= 2^{n-1} \cdot \left(\left(1-2^{-n}\right) \cdot \left[\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \left(\frac{\rho}{2-\rho}\right)^2 \right]^m + 2^{-n} \cdot \left[\left(1-\frac{\rho}{2}\right)\cdot 1 + \frac{\rho}{2}\cdot \left(\frac{\rho}{2-\rho}\right)^2 \right]^m \right) \\
&\quad + 2^{n-1}(2^{n-1}-1) \cdot \left((1-2\cdot 2^{-n}) \cdot \left[\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \frac{\rho}{2-\rho} \right]^{2m} \right. \\
&\quad \left. + 2\cdot 2^{-n} \cdot \left[\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \frac{\rho}{2-\rho} \right]^m \left[\left(1-\frac{\rho}{2}\right)\cdot 1 + \frac{\rho}{2}\cdot \frac{\rho}{2-\rho} \right]^m \right)
\end{aligned}$$

$$= \frac{1}{2}\cdot\frac{2^n-1}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^2+\rho^2\right)^m + \frac{1}{2}\cdot\frac{1}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^3+\rho^3\right)^m$$
$$+ \frac{(2^{n-1}-1)^2}{(2-\rho)^{2m}} + \frac{2^{n-1}-1}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^2+\rho^2\right)^m$$
$$= \frac{(2^{n-1}-1)^2}{(2-\rho)^{2m}} + \frac{2^n-\frac{3}{2}}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^2+\rho^2\right)^m + \frac{1}{2}\cdot\frac{1}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^3+\rho^3\right)^m.$$

为证明 L_0 集中在 $\mathbb{E}[L_0]$附近，我们将证明 $\mathrm{Var}[L_0] = o(\mathbb{E}[L_0]^2)$，或等价地证明 $\mathbb{E}[L_0^2] = (1+o(1))\mathbb{E}[L_0]^2$。我们有

$$\frac{\mathbb{E}[L_0^2]}{\mathbb{E}[L_0]^2}$$
$$= \frac{\frac{(2^{n-1}-1)^2}{(2-\rho)^{2m}} + \frac{2^n-\frac{3}{2}}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^2+\rho^2\right)^m + \frac{1}{2}\cdot\frac{1}{2^m(2-\rho)^{2m}}\cdot\left((2-\rho)^3+\rho^3\right)^m}{\frac{1}{4}\cdot\left[(2^n-1)\left(\frac{1}{2-\rho}\right)^m + \left(\frac{(2-\rho)^2+\rho^2}{2(2-\rho)}\right)^m\right]^2}$$
$$= \frac{\frac{(2^{n-1}-1)^2}{2^{2m}} + \frac{2^n-\frac{3}{2}}{2^m}\cdot\left(\left(1-\frac{\rho}{2}\right)^2+\left(\frac{\rho}{2}\right)^2\right)^m + \frac{1}{2}\cdot\left(\left(1-\frac{\rho}{2}\right)^3+\left(\frac{\rho}{2}\right)^3\right)^m}{\frac{1}{4}\cdot\left(\frac{(2^n-1)^2}{2^{2m}} + 2\cdot\frac{2^n-1}{2^m}\cdot\left(\left(1-\frac{\rho}{2}\right)^2+\left(\frac{\rho}{2}\right)^2\right)^m + \left(\left(1-\frac{\rho}{2}\right)^2+\left(\frac{\rho}{2}\right)^2\right)^{2m}\right)}$$
$$= \frac{(1+O(2^{n-2m}))A^2 + (4+O(2^{-m}))AB + 2C}{(1-O(2^{n-2m}))A^2 + (2-O(2^{-m}))AB + B^2},$$

其中 $A = \frac{2^n}{2^m}$、$B = \left(\left(1-\frac{\rho}{2}\right)^2+\left(\frac{\rho}{2}\right)^2\right)^m$ 和 $C = \left(\left(1-\frac{\rho}{2}\right)^3+\left(\frac{\rho}{2}\right)^3\right)^m$，

$$= (1+o(1))\frac{1+4\frac{B}{A}+2\frac{C}{A^2}}{1+2\frac{B}{A}+\frac{B^2}{A^2}}.$$

只要 $\frac{B}{A} = o(1)$ 和 $\frac{C}{A^2} = o(1)$，我们就可以得出 $\mathrm{Var}[L_0] = o(\mathbb{E}[L_0]^2)$。我们进一步注意到，对于任意 $\rho \in [0,\ 1]$，$\left(1-\frac{\rho}{2}\right)^3+\left(\frac{\rho}{2}\right)^3 = 1-\frac{3}{2}\rho+\frac{3}{4}\rho^2 \geqslant \left(\left(1-\frac{\rho}{2}\right)^2+\left(\frac{\rho}{2}\right)^2\right)^2$。因此，$C \geqslant B^2$。只要 $\frac{C}{A^2} = o(1)$，我们也能得到 $\frac{B}{A} \leqslant \frac{C}{A}^{\frac{1}{2}} = o(1)$。

因此，如果 $\frac{1}{C} = \frac{1}{\left(1-\frac{3}{2}\rho+\frac{3}{4}\rho^2\right)^m} \gg \frac{1}{A^2} = 2^{2(m-n)}$，那么我们有 $\frac{B}{A} = o(1)$, $\frac{C}{A^2} = o(1)$并且

$$\frac{\mathrm{Var}[L_0]}{\mathbb{E}[L_0]^2} = \frac{\mathbb{E}[L_0^2]}{\mathbb{E}[L_0]^2} - 1$$
$$= (1+o(1))\frac{1+4\frac{B}{A}+2\frac{C}{A^2}}{1+2\frac{B}{A}+\frac{B^2}{A^2}} - 1$$
$$\leqslant o(1).$$

通过对称论证，我们得到 $\mathrm{Var}[L_1] = o(\mathbb{E}[L_1]^2)$。因此，根据切比雪夫不等式，对于 A 和 x^* 的随机选择，以 $1-o(1)$ 的概率，我们有

$$L_0 = (1\pm o(1))\mathbb{E}[L_0],$$

$$L_1 = (1 \pm o(1))\mathbb{E}[L_1],$$

由于 $\mathbb{E}[L_0] = \mathbb{E}[L_1]$，以 $1 - o(1)$ 的概率，我们有

$$\begin{aligned}\Pr(x_1 = 0|A, \hat{y}) &= \frac{L_0}{L_0 + L_1} \\ &= \frac{1}{2} \pm o(1), \\ \Pr(x_1 = 1|A, \hat{y}) &= \frac{L_1}{L_0 + L_1} \\ &= \frac{1}{2} \pm o(1).\end{aligned}$$

在这一高概率事件 E 的条件下，条件 MMSE 至少为

$$\begin{aligned}&\mathbb{E}[(x_1^* - \mathbb{E}[x_1|A, \hat{y}])^2|E] \\ &\geqslant \mathbb{E}\left[\left(x_1^* - \frac{1}{2} \pm o(1)\right)^2 |E\right] \\ &= \left(\frac{1}{2} - o(1)\right)^2 \\ &= \frac{1}{4} - o(1),\end{aligned}$$

因此，

$$\mathbb{E}[(x_1^* - \mathbb{E}[x_1|A, \hat{y}])^2] \geqslant \Pr(E) \cdot \mathbb{E}[(x_1^* - \mathbb{E}[x_1|A, \hat{y}])^2|E] = \frac{1}{4} - o(1).$$

由于随机线性码问题的对称性，对于任意 $i \in [n]$，我们有

$$\mathbb{E}[(x_i^* - \mathbb{E}[x_i|A, \hat{y}])^2] = \mathbb{E}[(x_1^* - \mathbb{E}[x_1|A, \hat{y}])^2] \geqslant \frac{1}{4} - o(1).$$

最后，我们得出含噪 MMSE 至少为

$$\begin{aligned}&\mathrm{MMSE}_\rho \\ &= \mathbb{E}[\|x^* - \mathbb{E}[x|A, T_\rho(y)]\|_2^2] \\ &= \mathbb{E}[\|x^* - \mathbb{E}[x|A, \hat{y}]\|_2^2] \\ &= \sum_{i=1}^n E[(x_i^* - \mathbb{E}[x_i|A, \hat{y}])^2] \\ &\geqslant \left(\frac{1}{4} - o(1)\right) n.\end{aligned}$$

□

7.2. **低次多项式的稳定性：定理 3.19 的证明。** 本节将证明随机线性码问题中多项式的稳定性参数。回顾一下，为了证明定理 3.19，我们需要证明，对于任意 D 满足 $2D \leqslant n$的多项式 $f\colon \{0, 1\}^{m \times n} \times \{0, 1\}^m \to \mathbb{R}$ 次数至多为 D 的多项式 $(\rho, 2(1 - (1 - \rho)^D))$对于随机线性码问题是稳定的。

定理 3.19。 对于多项式 $f\colon \{0,\ 1\}^{m\times n}\times\{0,\ 1\}^{m}\to\mathrm{R}$ 次数至多为 D，我们可以将其唯一地写成

$$f(A,y)=\sum_{\substack{S\subseteq[m]\times[n],\\ T\subseteq[m]:\\ |S|+|T|\leqslant D}}\hat{f}_{S,T}\chi_{S,T}(A,y),$$

其中 $\chi_{S,\ T}(A,\ y)$ 表示

$$\chi_{S,T}(A,y)=\prod_{\{i,j\}\in S}(2A_{i,j}-1)\prod_{k\in T}(2y_k-1).$$

我们将使用的一个重要事实是，对于任意 $S\subseteq[m]\times[n]$ 和 $T\subseteq[m]$，只要不存在某个索引 $i\in[m]$ 使得 $\{i\}\times[n]\subseteq S$ 和 $i\in T$，那么在随机线性码问题中采样的 $(A^S,\ y^T)$的边缘分布就是均匀伯努利分布$(1/2)^{\otimes S}\times$伯努利分布$(1/2)^{\otimes T}$，其中 $(A,\ y)$。这是因为 $y=Ax$，其中 A 和 x 均为均匀随机抽取，因此只要 A^S 不包含与 y的某个矩阵元素相对应的 A的整行，$y^T|A^S$ 就服从伯努利分布$(1/2)^{\otimes T}$。

现在，我们有

$$\mathbb{E}[(f(A,y)-f(A,T_\rho(y)))^2]=\mathbb{E}[f(A,y)^2]+\mathbb{E}[f(A,T_\rho(y))^2]-2\mathbb{E}[f(A,y)f(A,T_\rho(y))].$$

我们将分别分析这三个项。

首先，

$$\begin{aligned}\mathbb{E}[f(A,y)^2]&=\sum_{\substack{S_1,T_1,S_2,T_2:\\ |S_1|+|T_1|\leqslant D,\\ |S_2|+|T_2|\leqslant D}}\hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2}\mathbb{E}[\chi_{S_1,T_1}(A,y)\chi_{S_2,T_2}(A,y)]\\&=\sum_{\substack{S_1,T_1,S_2,T_2:\\ |S_1|+|T_1|\leqslant D,\\ |S_2|+|T_2|\leqslant D}}\hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2}\mathbb{E}[\chi_{S_1\triangle S_2,T_1\triangle T_2}(A,y)]\end{aligned}$$

注意，由于 $2D\leqslant n$，$(A^{S_1\triangle S_2},\ y^{T_1\triangle T_2})$ 的边缘分布是均匀分布，因此当 $(S_1,\ T_1)\neq(S_2,\ T_2)$时，$\chi_{S_1\triangle S_2,\ T_1\triangle T_2}$ 的期望为 0。

$$=\sum_{\substack{S,T:\\ |S|+|T|\leqslant D}}\hat{f}_{S,T}^2. \tag{23}$$

其次，

$$\begin{aligned}\mathbb{E}[f(A,T_\rho(y))^2]&=\sum_{\substack{S_1,T_1,S_2,T_2:\\ |S_1|+|T_1|\leqslant D,\\ |S_2|+|T_2|\leqslant D}}\hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2}\mathbb{E}[\chi_{S_1,T_1}(A,T_\rho(y))\chi_{S_2,T_2}(A,T_\rho(y))]\\&=\sum_{\substack{S_1,T_1,S_2,T_2:\\ |S_1|+|T_1|\leqslant D,\\ |S_2|+|T_2|\leqslant D}}\hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2}\mathbb{E}[\chi_{S_1\triangle S_2,T_1\triangle T_2}(A,T_\rho(y))]\end{aligned}$$

注意， T_ρ 会以概率 ρ从伯努利分布 $\mathrm{Bern}(1/2)$ 中重新采样 y 的每个坐标。因此，当$2D \leqslant n$ 时， $(A^{S_1 \triangle S_2},\ T_\rho(y)^{T_1 \triangle T_2})$ 的边缘分布是均匀分布，并且只要 $(S_1,\ T_1) \neq (S_2,\ T_2)$， $\chi_{S_1 \triangle S_2,\ T_1 \triangle T_2}$ 的期望就是 0。

$$= \sum_{\substack{S,T: \\ |S|+|T| \leqslant D}} \hat{f}_{S,T}^2. \tag{24}$$

第三，我们有

$$\mathbb{E}[f(A,y)f(A,T_\rho(y))] = \sum_{\substack{S_1,T_1,S_2,T_2: \\ |S_1|+|T_1| \leqslant D, \\ |S_2|+|T_2| \leqslant D}} \hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2}\mathbb{E}[\chi_{S_1,T_1}(A,y)\chi_{S_2,T_2}(A,T_\rho(y))]$$

注意，每当从伯努利分布$(1/2)$ 中按 T_ρ重新抽取 y^{T_2} 的某个矩阵元素时， $\chi_{S_1,T_1}(A,y)\chi_{S_2,T_2}(A,T_\rho(y))$ 的条件期望为 0。因此，

$$\begin{aligned}
&= \sum_{\substack{S_1,T_1,S_2,T_2: \\ |S_1|+|T_1| \leqslant D, \\ |S_2|+|T_2| \leqslant D}} \hat{f}_{S_1,T_1}\hat{f}_{S_2,T_2} \cdot (1-\rho)^{|T_2|} \cdot \mathbb{E}[\chi_{S_1,T_1}(A,y)\chi_{S_2,T_2}(A,y)] \\
&= \sum_{\substack{S,T: \\ |S|+|T| \leqslant D}} \hat{f}_{S,T}^2 (1-\rho)^{|T|} \\
&\geqslant (1-\rho)^D \sum_{\substack{S,T: \\ |S|+|T| \leqslant D}} \hat{f}_{S,T}^2.
\end{aligned} \tag{25}$$

结合（23）、（24）和（25），我们有

$$\begin{aligned}
\mathbb{E}[(f(A,y)-f(A,T_\rho(y)))^2] &= \mathbb{E}[f(A,y)^2] + \mathbb{E}[f(A,T_\rho(y))^2] - 2\mathbb{E}[f(A,y)f(A,T_\rho(y))] \\
&\leqslant 2(1-(1-\rho)^D) \sum_{\substack{S,T: \\ |S|+|T| \leqslant D}} \hat{f}_{S,T}^2. \\
&= 2(1-(1-\rho)^D) \cdot \mathbb{E}[f(A,y)^2].
\end{aligned}$$

由此完成证明：若 $2D \leqslant n$，则 f 是 $(\rho, 2(1-(1-\rho)^D)$-稳定的 . □

8. 高斯子集和的证明

8.1. 关于最小均方误差不稳定性的延后证明。

3.24. GSS 设定与 [RXZ19] 中的稀疏回归设定完全相同，只有一个样本，且我们处于无噪声情形（按他们的记号） $\sigma = 0$。在有噪声情形下，GSS 上的 OU 算子对应的仍是稀疏回归设定，其中 $n = 1$ 且噪声为 $\sigma = \rho/\sqrt{1-\rho^2}$。

现在，在噪声为 $\sigma = \rho/\sqrt{1-\rho^2}$ 的稀疏回归设定中，根据 [RXZ19]，AoN 的信息论阈值由样本量 $n^* = 2k\log(N/k)/\log(1+k/\sigma^2) = 2k\log(N/k)/\log(1+k(1-\rho^2)/\rho^2)$给出。特别是，如果 $n \leqslant 0.99n^*$ 且 $k/\sigma^2 = k(1-\rho^2)/\rho^2 \geqslant C$ 对于某个通用常数 $C > 0$，根据 [RXZ19（定理 3，arXiv 版本）]

最小均方误差是“平凡的”，即 $\mathrm{MMSE}_\rho \geqslant (1-o(1))k$。但很容易看出，可以选取某个 $\rho = \exp(-\Theta(k\log(N/k)))$,，从而保证 $n^* \geqslant 30$，确保对于 GSS 设定中的 $\rho > 0$，$n = 1$ 的最小均方误差是“平凡的”。□

8.2. **低次多项式的稳定性**。本节将证明 GSS 中低次多项式的稳定性参数。我们将利用以下关于厄米特多项式乘积期望的定理，其证明推迟到第 8.3节。

定理 8.1 (厄米特多项式乘积期望的图示公式)。假设x_1、……、x_k与 $N(0、1)$同分布。令 $R_{ij} = \mathbb{E}[x_i x_j]$ 表示它们的相关系数。给定$\alpha \in \mathbb{N}^k$，定义如下图 $G(\alpha) = (V、E)$，它有 $|\alpha|$ 个顶点。对于每个 $i \in [k]$，创建 α_i个顶点，分别对应于 x_i；并令 $j(v) = i$ 对应于与 x_i相关联的 α_i 个顶点中的每一个。对于$u、v \in V$，若 $j(u) \neq j(v)$，则添加一条边 $\{u、v\}$。于是，我们有

$$\mathbb{E}\left[\prod_{i=1}^{k} h_{\alpha_i}(x_i)\right] = \frac{1}{\sqrt{\alpha!}} \sum_{M\in\mathcal{M}(\alpha)} \prod_{\{a,b\}\in M} R_{j(a)j(b)},$$

其中，$\mathcal{M}(\alpha)$ 表示上述图 $G(\alpha)$ 中的完美匹配集合。

接下来，我们证明，若 $k = o(n)$ 且 $D = o(\min\{k^{1/4}, (n/k)^{1/5}\})$,，则次数至多为 D 的多项式对于 GSS 是 $(\rho, 2(1-(\sqrt{1-\rho^2})^D + o(1)))$-稳定的。

定理3.25. 设多项式 $f: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ 的次数至多为 D。我们可以将其唯一地写成

$$f(X,Y) = \sum_{\substack{\alpha\in\mathbb{N}^n, t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t} h_\alpha(X) h_t(y),$$

其中 $y = \frac{Y}{\sqrt{k}}$

注意，由于 $Y = \sum_{i\in S} X_i$ 对于 S 而言是一个均匀随机的 k-子集，且 S 属于 $[n]$，我们可以表示为$y = \langle u, X\rangle$其中 $u := \frac{1}{\sqrt{k}}\mathbb{1}_S$，而 $\mathbb{1}_S$ 是 S的指示向量。我们可以计算 f 的二阶矩，得到

$$\begin{aligned}
&\mathbb{E}\left[f(X,Y)^2\right] \\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n, t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D, |\alpha_2|+t_2\leqslant D}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right] \\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n, t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D, |\alpha_2|+t_2\leqslant D}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right].
\end{aligned}$$

现在我们应用维克公式来分析内层期望。我们有

$$\begin{aligned}
&\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right] \\
&= \mathbb{E}_S\left[\frac{1}{\sqrt{\alpha_1!\alpha_2!t_1!t_2!}} \sum_{M\in\mathcal{M}(\alpha_1,\alpha_2,t_1,t_2)} \prod_{\{a,b\}\in M} R_{j(a)j(b)}\right].
\end{aligned}$$

由此公式可知，所有内层期望均非负。接下来我们分析图 $G(\alpha_1, \alpha_2, t_1, t_2)$ 中的完美匹配。G 中有 4 类顶点：来自 $\mathbf{h}_{\alpha_1}$ 的 X-顶点、来自 $\mathbf{h}_{\alpha_2}$ 的 X-顶点、来自 $\mathbf{h}_{t_1}$ 的 y-顶点，以及来自 $\mathbf{h}_{t_2}$ 的 y-顶点。我们将来自 $\mathbf{h}_{\alpha_1}$ 的 X-顶点记作 U，将来自 $\mathbf{h}_{\alpha_2}$ 的 X-顶点记作 V，将来自 $\mathbf{h}_{t_1}$ 的 y-顶点记作 R，并将来自 $\mathbf{h}_{t_2}$ 的 y-顶点记作 T。注意，对于 S 的固定选择，$U \sqcup V$ 中对应于 X_i 的一个顶点与一个 y-顶点之间的协方差为 $\frac{1}{\sqrt{k}}\mathbf{1}\{i \in S\}$。$U \sqcup V$ 中对应于 X_i 和 X_j 的顶点之间的协方差为 $\mathbf{1}\{i = j\}$。$R \sqcup T$ 中顶点之间的协方差为 $\mathbf{1}$。

因此，对 S 取期望后，对总和有贡献的匹配由以下部分组成：U 的一个子集与 V 的一个子集之间的完美匹配，R 的一个子集与 T 的一个子集之间的完美匹配，以及 $U \sqcup V$ 的其余部分与 $R \sqcup T$ 的其余部分之间的完美匹配。此外，U 的一个子集与 V 的一个子集之间的完美匹配，在限制到对应于任意 $i \in [n]$ 的 X_i 顶点时，还必须是完美匹配，因为对于 $i \neq j$，X_i 与 X_j 不相关这一事实意味着，有贡献的匹配中不能有连接对应于不同 X_i 的顶点的边。

对于一个有贡献的匹配 $M \in \mathbf{M}(\alpha_1, \alpha_2, t_1, t_2)$，我们将连接 $U \sqcup V$ 中顶点的边 $\{a, b\} \in M$ 称为 X-边，将连接 $R \sqcup T$ 中顶点的边称为 y-边，并将连接 $U \sqcup V$ 与 $R \sqcup T$ 中顶点的边称为交叉边。我们知道，每条 X-边和每条 y-边对 M 的乘积贡献 $\mathbf{1}$。交叉边的贡献取决于两个量：（1）交叉边的数量；（2）交叉边所关联的、对应于不同变量 X_i 的顶点数量。令 m 表示交叉边的数量，z 表示交叉边所关联的不同 X_i 的数量。显然有关系 $m \geqslant z$。在对 S 取期望时，交叉边的贡献等于

$$\begin{aligned}&\Pr(\text{a fixed } z\text{-subset of } [n] \subseteq S) \cdot \left(\frac{1}{\sqrt{k}}\right)^m \\ &= (1 + O(1/n)) \left(\frac{k}{n}\right)^z \cdot \left(\frac{1}{\sqrt{k}}\right)^m.\end{aligned}$$

接下来，我们给出一个估计 $C(m, z; \alpha_1, \alpha_2, t_1, t_2)$，用于估算含有 m 条交叉边，且这些交叉边关联到 z 个不同 X_i 的贡献完美匹配数量。要枚举这样的匹配，我们可以先枚举 $U \sqcup V$ 中的 m 个顶点和 $R \sqcup T$ 中的 m 个顶点，然后在 $U \sqcup V$ 的已枚举子集与 $R \sqcup T$ 的已枚举子集之间构造一个完美匹配。这样至多得到

$$\binom{|U| + |V|}{m} \cdot \binom{|R| + |T|}{m} \cdot m! \leqslant \left(\frac{e(|U| + |V|)}{m}\right)^m \left(\frac{e(|R| + |T|)}{m}\right)^m m^m$$

选择。接下来，我们枚举 X 边，即 $U \sqcup V$ 中剩余顶点之间的完美匹配。这至多给出

$$\min\{\alpha_1!, \alpha_2!\} \leqslant \sqrt{\alpha_1!\alpha_2!}$$

选择。最后，我们枚举 y 边，即 $R \sqcup T$ 中剩余顶点之间的完美匹配。这至多给出

$$\min\{t_1!, t_2!\} \leqslant \sqrt{t_1! t_2!}$$

选择。因此，我们有

$$\begin{aligned}
&C(m, z; \alpha_1, \alpha_2, t_1, t_2) \\
&\leqslant \sqrt{\alpha_1! \alpha_2! t_1! t_2!} \left(\frac{e(|U| + |V|)}{m}\right)^m \left(\frac{e(|R| + |T|)}{m}\right)^m m^m \\
&= \sqrt{\alpha_1! \alpha_2! t_1! t_2!} \left(\frac{e^2 D^2}{m}\right)^m,
\end{aligned}$$

其中我们用到了 $(|U|+|V|)(|R|+|T|) \leqslant D^2$ 因为 $|U|+|V|+|R|+|T| = |\alpha_1|+t_1+|\alpha_2|+t_2 \leqslant 2D$.

此外，我们注意到必有 $z \geqslant |\alpha_1 \triangle \alpha_2|_0$，因为对于 i 与 α_1 不同的任何坐标 α_2，在某个有贡献的完美匹配中，都有一个对应于 X_i 的顶点与交叉边相 incident。我们还有 $m \geqslant |\alpha_1 \triangle \alpha_2|$，因为至少有 $|\alpha_1(i) - \alpha_2(i)|$ 条交叉边与对应于 X_i 的顶点相 incident。类似地，我们有 $m \geqslant |t_1 - t_2|$，因为至少有 $|t_1 - t_2|$ 条交叉边与对应于 y 的顶点相 incident。

总之，对于固定的 α_1， α_2， t_1， t_2，内层期望的总贡献至多为

$$\begin{aligned}
&\mathbb{E}_S \mathbb{E}_X \left[h_{\alpha_1}(X) h_{\alpha_2}(X) h_{t_1}(y) h_{t_2}(y)\right] \\
&= \mathbb{E}_S \left[\frac{1}{\sqrt{\alpha_1! \alpha_2! t_1! t_2!}} \sum_{M \in \mathcal{M}(\alpha_1, \alpha_2, t_1, t_2)} \prod_{\{a,b\} \in M} R_{j(a)j(b)}\right] \\
&\leqslant \frac{1}{\sqrt{\alpha_1! \alpha_2! t_1! t_2!}} \sum_{z \geqslant |\alpha_1 \triangle \alpha_2|_0} \sum_{m \geqslant \max\{|\alpha_1 \triangle \alpha_2|, |t_1 - t_2|\}} C(m, z; \alpha_1, \alpha_2, t_1, t_2) \cdot (1 + O(1/n)) \left(\frac{k}{n}\right)^z \cdot \left(\frac{1}{\sqrt{k}}\right)^m \\
&\leqslant (1 + O(1/n)) \sum_{z \geqslant |\alpha_1 \triangle \alpha_2|_0} \left(\frac{k}{n}\right)^z \sum_{m \geqslant \max\{|\alpha_1 \triangle \alpha_2|, |t_1 - t_2|\}} \left(\frac{e^2 D^2}{m\sqrt{k}}\right)^m. \qquad (26)
\end{aligned}$$

我们还注意到，对于满足 $\alpha_1 = \alpha_2 = \alpha$ 和 $t_1 = t_2 = t$的对角项，该期望（26）至少为

$$\begin{aligned}
&\mathbb{E}_S \mathbb{E}_X \left[h_\alpha(X)^2 h_t(y)^2\right] \\
&= \mathbb{E}_S \left[\frac{1}{\sqrt{\alpha! \alpha! t! t!}} \sum_{M \in \mathcal{M}(\alpha, \alpha, t, t)} \prod_{\{a,b\} \in M} R_{j(a)j(b)}\right] \\
&\geqslant \frac{1}{\alpha! t!} \cdot \alpha! t! \\
&= 1,
\end{aligned}$$

其中，在最后一个不等式中，我们对没有交叉边的完美匹配求和，得到 $\alpha!$ 种在 U 和 V 之间进行匹配的选择，再乘以 $t!$ 种在 R和 T之间进行匹配的选择。

现在我们有

$$\begin{aligned}
&\mathbb{E}[f(X,Y)^2] \\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right] \\
&= \sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)^2\right] \\
&\quad + \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right].
\end{aligned}$$

我们知道，对角项的贡献至少为

$$\begin{aligned}
&\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)^2\right] \\
\geqslant &\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2.
\end{aligned}$$

No现在来考虑非对角项。对于固定的 $\alpha_1,\alpha_2\in\mathbf{N}^n$，我们将记 e

$$\beta = \alpha_1\wedge\alpha_2 \tag{27}$$

$$a = |\alpha_1-\beta|_0 \tag{28}$$

$$b = |\alpha_2-\beta|_0 \tag{29}$$

$$c = |\alpha_1-\beta| \tag{30}$$

$$d = |\alpha_2-\beta|. \tag{31}$$

对于固定的 $\alpha_1\in\mathbf{N}^n$ 和固定的 $a,\ b,\ c,\ d\in\mathbf{N}$，我们用 $W_{\alpha_1}(a,\ b,\ c,\ d)$ 表示满足上述参数且 $\alpha_2\neq\alpha_1$ 的 $\alpha_2\in\mathbf{N}^n$ 的集合。注意，如果 $\alpha_1\neq\alpha_2$，则有$a+b\geqslant 1$。那么，非对角项的贡献至多为

$$\begin{aligned}
&\left|\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]\right| \\
\leqslant &\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}} |\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}|\cdot\left|\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]\right| \\
= &\sum_{\substack{\alpha\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha|+t_1\leqslant D,|\alpha|+t_2\leqslant D,\\ t_1\neq t_2}} |\hat{f}_{\alpha,t_1}||\hat{f}_{\alpha,t_2}|\cdot\left|\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_{t_1}(y)h_{t_2}(y)\right]\right|
\end{aligned}$$

$$+\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ \alpha_1\neq\alpha_2}}|\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}|\cdot|\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]|.$$

涉及对 $t_1\neq t_2$ 求和的第一项至多为

$$\sum_{\substack{\alpha\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha|+t_1\leqslant D,|\alpha|+t_2\leqslant D,\\ t_1\neq t_2}}|\hat{f}_{\alpha,t_1}||\hat{f}_{\alpha,t_2}|\cdot|\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_{t_1}(y)h_{t_2}(y)\right]|$$

$$\leqslant 2\sum_{\alpha\in\mathbb{N}^n,t_1\in\mathbb{N}}|\hat{f}_{\alpha,t_1}|\sum_{\substack{t_2\neq t_1,\\ |\hat{f}_{\alpha,t_2}|\leqslant|\hat{f}_{\alpha,t_1}|}}|\hat{f}_{\alpha,t_2}|\cdot\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_{t_1}(y)h_{t_2}(y)\right]$$

其中该不等式成立，是因为每一对有序对 $(t_1,\ t_2)$ 都可以重新排序，使得$\hat{}\qquad\hat{}|f_{\alpha,\ t_2}|\leqslant|f_{\alpha,\ t_1}|$。现在，我们应用（26）并得到

$$\leqslant 2(1+O(1/n))\sum_{\alpha\in\mathbb{N}^n,t_1\in\mathbb{N}}\hat{f}^2_{\alpha,t_1}\sum_{t_2\neq t_1}\sum_{z\geqslant 0}\left(\frac{k}{n}\right)^z\sum_{m\geqslant|t_1-t_2|}\left(\frac{e^2D^2}{m\sqrt{k}}\right)^m$$

$$\leqslant 2\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\alpha\in\mathbb{N}^n,t_1\in\mathbb{N}}\hat{f}^2_{\alpha,t_1}\sum_{t_2\neq t_1}\left(\frac{e^2D^2}{|t_1-t_2|\sqrt{k}}\right)^{|t_1-t_2|}$$

$$\leqslant 2\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\alpha\in\mathbb{N}^n,t_1\in\mathbb{N}}\hat{f}^2_{\alpha,t_1}\cdot 2\left(\frac{e^2D^2}{\sqrt{k}}\right)$$

$$=o\left(\sum_{\alpha\in\mathbb{N}^n,t\in\mathbb{N}}\hat{f}^2_{\alpha,t}\right),$$

前提是 $k=o(n)$和 $D=o(k^{1/4})$。

涉及对 $\alpha_1\neq\alpha_2$ 求和的第二项至多为

$$\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ \alpha_1\neq\alpha_2}}|\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}|\cdot|\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]|$$

$$\leqslant 2\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}|\hat{f}_{\alpha_1,t_1}|\sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}}\sum_{\substack{\alpha_2\in\mathbb{N}^n,t_2\in\mathbb{N}:\\ |\alpha_2|+t_2\leqslant D,\\ \alpha_2\in W_{\alpha_1}(a,b,c,d)\\ |\hat{f}_{\alpha_2,t_2}|n^{\frac{b}{2}}D^{\frac{a}{2}}\\ \leqslant|\hat{f}_{\alpha_1,t_1}|n^{\frac{a}{2}}D^{\frac{b}{2}}}}|\hat{f}_{\alpha_2,t_2}|\cdot\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]$$

其中，不等式成立是因为每个 $((\alpha_1、\ t_1)、\ (\alpha_2、\ t_2))$ 的有序对都可以重新排列，使得 $|\hat{}\ f_{\alpha_2、\ t_2}|n^{\frac{b}{2}}D^{\frac{a}{2}} \leqslant |\hat{}\ f_{\alpha_1、\ t_1}|n^{\frac{a}{2}}D^{\frac{b}{2}}$。

$$\begin{aligned}
&\leqslant 2 \sum_{\substack{\alpha_1\in\mathbb{N}^n, t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}} \sum_{\substack{\alpha_2\in\mathbb{N}^n, t_2\in\mathbb{N}:\\ |\alpha_2|+t_2\leqslant D,\\ \alpha_2\in W_{\alpha_1}(a,b,c,d)\\ |\hat{f}_{\alpha_2,t_2}|n^{\frac{b}{2}}D^{\frac{a}{2}}\\ \leqslant |\hat{f}_{\alpha_1,t_1}|n^{\frac{a}{2}}D^{\frac{b}{2}}}} \mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]\\
&\leqslant 2(1+O(1/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n, t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}}\\
&\quad \sum_{\substack{\alpha_2\in\mathbb{N}^n, t_2\in\mathbb{N}:\\ |\alpha_2|+t_2\leqslant D,\\ \alpha_2\in W_{\alpha_1}(a,b,c,d)\\ |\hat{f}_{\alpha_2,t_2}|n^{\frac{b}{2}}D^{\frac{a}{2}}\\ \leqslant |\hat{f}_{\alpha_1,t_1}|n^{\frac{a}{2}}D^{\frac{b}{2}}}} \sum_{z\geqslant|\alpha_1\triangle\alpha_2|_0} \left(\frac{k}{n}\right)^z \sum_{m\geqslant\max\{|\alpha_1\triangle\alpha_2|,|t_1-t_2|\}} \left(\frac{e^2D^2}{m\sqrt{k}}\right)^m\\
&\leqslant 2(1+O(1/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n, t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}}\\
&\quad \sum_{\substack{\alpha_2\in\mathbb{N}^n, t_2\in\mathbb{N}:\\ |\alpha_2|+t_2\leqslant D,\\ \alpha_2\in W_{\alpha_1}(a,b,c,d)}} \sum_{z\geqslant a+b} \left(\frac{k}{n}\right)^z \sum_{m\geqslant\max\{c+d,|t_1-t_2|\}} \left(\frac{e^2D^2}{m\sqrt{k}}\right)^m.
\end{aligned}$$

现在，我们继续对 $\alpha_2 \in W_{\alpha_1}(a,b,c,d)$ 的数量给出上界。为了枚举这类 α_2，我们可以先枚举满足 $\alpha_2(i) > \alpha_1(i)$ 的 b 个索引，至多有 $\binom{n}{b}$ 种选择。然后，我们枚举满足 $\alpha_1(i) > \alpha_2(i)$ 的 a 个索引，至多有 $\binom{|\alpha_1|_0}{a}$ 种选择。接下来，我们枚举坐标 $\alpha_1(i) > \alpha_2(i)$ 上 $\alpha_2(i)$ 的选择。回忆 $\beta = \alpha_1 \wedge \alpha_2$。由于

$$c = |\alpha_1 - \beta| = \sum_{i:\alpha_1(i)>\alpha_2(i)} (\alpha_1(i) - \alpha_2(i)),$$

根据标准插板法论证，当 $a \geqslant 1$ 时，在坐标 $\alpha_1(i) > \alpha_2(i)$ 上选择 α_2 的方式数至多为 $\binom{c-1}{a-1} \leqslant \binom{c}{a}$；当 $a = 0$ 时，至多为 $1 \leqslant \binom{c}{a}$。类似地，在坐标 $\alpha_1(i) < \alpha_2(i)$ 上选择 $\alpha_2(i)$ 的方式数至多为 $\binom{d}{b}$。因此，$\binom{n}{b}\binom{|\alpha_1|_0}{a}\binom{c}{a}\binom{d}{b}$ 是 $\alpha_2 \in W_{\alpha_1}(a,b,c,d)$ 数量的一个上界。回到对第二项的估计，我们有

$$\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n, t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D, |\alpha_2|+t_2\leqslant D,\\ \alpha_1\neq\alpha_2}} |\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}| \cdot \left|\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]\right|$$

$$\leqslant 2(1+O(1/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}}$$
$$\sum_{\substack{\alpha_2\in\mathbb{N}^n,t_2\in\mathbb{N}:\\ |\alpha_2|+t_2\leqslant D,\\ \alpha_2\in W_{\alpha_1}(a,b,c,d)}} \sum_{z\geqslant a+b} \left(\frac{k}{n}\right)^z \sum_{m\geqslant\max\{c+d,|t_1-t_2|\}} \left(\frac{e^2D^2}{m\sqrt{k}}\right)^m$$

$$\leqslant 2(1+O(1/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b,c,d\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}}$$
$$\sum_{t_2\in\mathbb{N}} \binom{n}{b}\binom{|\alpha_1|_0}{a}\binom{c}{a}\binom{d}{b} \sum_{z\geqslant a+b} \left(\frac{k}{n}\right)^z \sum_{m\geqslant\max\{c+d,|t_1-t_2|\}} \left(\frac{e^2D^2}{m\sqrt{k}}\right)^m$$

$$\leqslant 2(1+O(1/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1,\\ c\geqslant a,d\geqslant b}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}} n^b D^a \cdot \frac{c^a}{a!}\frac{d^b}{b!}$$
$$\sum_{z\geqslant a+b} \left(\frac{k}{n}\right)^z \sum_{t_2\in\mathbb{N}} \sum_{m\geqslant\max\{c+d,|t_1-t_2|\}} \left(\frac{e^2D^2}{m\sqrt{k}}\right)^m$$

$$\leqslant 2(1+O(k/n)) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1,\\ c\geqslant a,d\geqslant b}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}} n^b D^a \cdot \frac{c^a}{a!}\frac{d^b}{b!}$$
$$\left(\frac{k}{n}\right)^{a+b} \left[(2(c+d)+1)\sum_{m\geqslant c+d}\left(\frac{e^2D^2}{m\sqrt{k}}\right)^m + 2\sum_{i\geqslant 1}\sum_{m\geqslant c+d+i}\left(\frac{e^2D^2}{m\sqrt{k}}\right)^m\right]$$

$$\leqslant 2\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1,\\ c\geqslant a,d\geqslant b}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}} n^b D^a \cdot \frac{c^a}{a!}\frac{d^b}{b!}$$
$$\left(\frac{k}{n}\right)^{a+b} \left[(2(c+d)+1)\left(\frac{e^2D^2}{(c+d)\sqrt{k}}\right)^{c+d} + 2\sum_{i\geqslant 1}\left(\frac{e^2D^2}{(c+d+i)\sqrt{k}}\right)^{c+d+i}\right]$$

$$\leqslant 2\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1}} n^{\frac{a-b}{2}} D^{\frac{b-a}{2}} n^b D^a \left(\frac{k}{n}\right)^{a+b}$$
$$\sum_{c\geqslant a,d\geqslant b} \frac{(2(c+d)+1)c^a d^b}{a!b!(c+d)^{c+d}} \left(\frac{e^2D^2}{\sqrt{k}}\right)^{c+d}$$

$$\leqslant 3\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right) \sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}} \hat{f}^2_{\alpha_1,t_1} \sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1}} \left(\frac{k\sqrt{D}}{\sqrt{n}}\right)^{a+b}$$

$$\sum_{c\geqslant a,d\geqslant b}\frac{2(c+d)c^a d^b}{a!b!\binom{c+d}{c}c^c d^d}\left(\frac{e^2D^2}{\sqrt{k}}\right)^{c+d}$$

现在我们考察项 $\frac{2(c+d)c^a d^b}{a!\,b!\,(\binom{c+d}{c})c^c d^d}$。由于 $c\geqslant a$和 $d\geqslant b$，若 $c\geqslant 1$和 $d\geqslant 1$均成立，则有$\frac{2(c+d)c^a d^b}{a!\,b!\,(\binom{c+d}{c})c^c d^d}\leqslant 2$。另一方面，假设 $c=0$，那么必有 $a=0$和 $b\geqslant 1$，此时 $\frac{2(c+d)c^a d^b}{a!\,b!\,(\binom{c+d}{c})c^c d^d}\leqslant\frac{2d\cdot d^b}{b!d^d}\leqslant 2$。利用对称性，我们将此论证推广到 $d=0$的情形。因此，我们得出该项始终有上界 2，从而得到

$$\begin{aligned}
&\leqslant 6\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}\hat{f}^2_{\alpha_1,t_1}\sum_{\substack{a,b\geqslant 0,\\ a+b\geqslant 1}}\left(\frac{k\sqrt{D}}{\sqrt{n}}\right)^{a+b}\sum_{c\geqslant a,d\geqslant b}\left(\frac{e^2D^2}{\sqrt{k}}\right)^{c+d}\\
&\leqslant 6\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}\hat{f}^2_{\alpha_1,t_1}\sum_{i\geqslant 1}(i+1)\left(\frac{k\sqrt{D}}{\sqrt{n}}\right)^{i}\sum_{j\geqslant i}(j+1)\left(\frac{e^2D^2}{\sqrt{k}}\right)^{j}\\
&\leqslant 6\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}\hat{f}^2_{\alpha_1,t_1}\sum_{i\geqslant 1}(i+1)\left(\frac{k\sqrt{D}}{\sqrt{n}}\right)^{i}\cdot(i+1)\left(\frac{e^2D^2}{\sqrt{k}}\right)^{i}\\
&=6\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}\hat{f}^2_{\alpha_1,t_1}\sum_{i\geqslant 1}(i+1)^2\left(\frac{e^2\sqrt{k}D^{\frac{5}{2}}}{\sqrt{n}}\right)^{i}\\
&\leqslant 24\left(1+O(k/n)+O\left(D^2/\sqrt{k}\right)+O\left(D^{\frac{5}{2}}\sqrt{k/n}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^n,t_1\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D}}\hat{f}^2_{\alpha_1,t_1}\cdot\frac{e^2\sqrt{k}D^{\frac{5}{2}}}{\sqrt{n}}\\
&=o\left(\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}}\hat{f}^2_{\alpha,t}\right),
\end{aligned}$$

前提是 $k=o(n)$ 和 $D=o\left(\min\left\{k^{1/4},(n/k)^{1/5}\right\}\right)$。

结合第一项和第二项的界，非对角项的贡献至多为

$$\left|\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}}\hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]\right|$$

$$\leqslant \sum_{\substack{\alpha\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha|+t_1\leqslant D,|\alpha|+t_2\leqslant D,\\ t_1\neq t_2}} |\hat{f}_{\alpha,t_1}||\hat{f}_{\alpha,t_2}|\cdot|\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_{t_1}(y)h_{t_2}(y)\right]|$$

$$+\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ \alpha_1\neq\alpha_2}} |\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}|\cdot|\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]|$$

$$\leqslant o\left(\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}}\hat{f}_{\alpha,t}^2\right),$$

因此， f 的二阶矩为

$$\mathbb{E}\left[f(X,Y)^2\right]$$

$$=\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D}}\hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]$$

$$\geqslant\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}}\hat{f}_{\alpha,t}^2\cdot\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)^2\right]+o\left(\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}}\hat{f}_{\alpha,t}^2\right). \tag{32}$$

接下来，我们分析噪声算子 T_ρ 对 y 的影响。回顾一下， T_ρ 通过与方差为 k 的高斯变量取平均来作用于 Y 。对于 $y=\frac{1}{\sqrt{k}}Y$ ，等效算子通过与标准高斯变量 $N(0,\ 1)$ 取平均来作用于 y 。更准确地说， $T_\rho(y)=\sqrt{1-\rho^2}y+\rho z$，其中 $z\sim N(0,\ 1)$ 是一个独立高斯变量。现在，我们的目标是分析噪声模型 $\mathbb{E}[f(X,\ T_\rho(Y))^2]$ 下的二阶矩，以及相关性 $\mathbb{E}[f(X,\ Y)f(X,\ T_\rho(Y))]$。

为此，我们需要了解内层期望在噪声算子作用下的表现。我们有

$$\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(T_\rho(y))h_{t_2}(T_\rho(y))\right]$$

$$=\mathbb{E}_S\left[\frac{1}{\sqrt{\alpha_1!\alpha_2!t_1!t_2!}}\sum_{M\in\mathcal{M}(\alpha_1,\alpha_2,t_1,t_2)}\prod_{\{a,b\}\in M}R_{j(a)j(b)}\right],$$

由于应用了 T_ρ，相关系数 R_{ij} 与之前不同。现在我们有

$$\mathbb{E}_X\left[T_\rho(y)^2\right]=1$$

$$\mathbb{E}_X\left[X_iT_\rho(y)\right]=\frac{\sqrt{1-\rho^2}}{\sqrt{k}}\mathbf{1}\{i\in S\}.$$

因此，所有 X-边和 Y-边的边贡献保持不变，而交叉边的贡献则按 $\sqrt{1-\rho^2}$ 的因子缩放。特别地，这意味着

$$\left(\sqrt{1-\rho^2}\right)^D\leqslant\frac{\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(T_\rho(y))h_{t_2}(T_\rho(y))\right]}{\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]}\leqslant 1,$$

对于所有满足 $|\alpha_1| + t_1 \leqslant D$ 和 $|\alpha_2| + t_2 \leqslant D$ 的 $(\alpha_1, t_1), (\alpha_2, t_2) \in \mathbf{N}^n \times \mathbf{N}$，我们得到

$$\begin{aligned}
&\mathbb{E}[f(X, T_\rho(Y))^2] \\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n, t_1,t_2\in\mathbb{N}: \\ |\alpha_1|+t_1\leqslant D, |\alpha_2|+t_2\leqslant D}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(T_\rho(y))h_{t_2}(T_\rho(y))\right] \\
&\leqslant \sum_{\substack{\alpha\in\mathbb{N}^n, t\in\mathbb{N}: \\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2 \cdot \mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2 h_t(T_\rho(y))^2\right] \\
&\quad + \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n, t_1,t_2\in\mathbb{N}: \\ |\alpha_1|+t_1\leqslant D, |\alpha_2|+t_2\leqslant D, \\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}} |\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}| \cdot \mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(T_\rho(y))h_{t_2}(T_\rho(y))\right] \\
&\leqslant \sum_{\substack{\alpha\in\mathbb{N}^n, t\in\mathbb{N}: \\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2 \cdot \mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2 h_t(y)^2\right] + o\left(\sum_{\substack{\alpha\in\mathbb{N}^n, t\in\mathbb{N}: \\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\right) \\
&\leqslant (1+o(1))\mathbb{E}\left[f(X,Y)^2\right], \qquad (33)
\end{aligned}$$

其中，最后一个不等式由 $\mathbb{E}_S\mathbb{E}_X[h_\alpha(X)^2 h_t(y)^2] \geqslant 1$ 得出。

类似地，我们继续分析相关项 $\mathbb{E}[f(X,Y)f(X,T_\rho(Y))]$。我们有

$$\begin{aligned}
&\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(T_\rho(y))\right] \\
&= \mathbb{E}_S\left[\frac{1}{\sqrt{\alpha_1!\alpha_2!t_1!t_2!}} \sum_{M\in\mathcal{M}(\alpha_1,\alpha_2,t_1,t_2)} \prod_{\{a,b\}\in M} R_{j(a)j(b)}\right],
\end{aligned}$$

其中，相关系数 R_{ij} 再次不同。现在我们有

$$\begin{aligned}
\mathbb{E}_X\left[yT_\rho(y)\right] &= \sqrt{1-\rho^2} \\
\mathbb{E}_X\left[X_i y\right] &= \frac{1}{\sqrt{k}}\mathbf{1}\{i\in S\} \\
\mathbb{E}_X\left[X_i T_\rho(y)\right] &= \frac{\sqrt{1-\rho^2}}{\sqrt{k}}\mathbf{1}\{i\in S\}.
\end{aligned}$$

因此，X-边以及对应于 X_i 和 y 的顶点之间的交叉边的所有边贡献保持不变，而 Y-边以及对应于 X_i 和 $T_\rho(y)$ 的顶点之间的交叉边的贡献则按因子 $\sqrt{1-\rho^2}$ 缩放。我们再次有

$$\left(\sqrt{1-\rho^2}\right)^D \leqslant \frac{\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(T_\rho(y))\right]}{\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(y)\right]} \leqslant 1,$$

对于所有 (α_1, t_1)，$(\alpha_2, t_2) \in \mathbf{N}^n \times \mathbf{N}$，使得 $|\alpha_1| + t_1 \leqslant D$ 且 $|\alpha_2| + t_2 \leqslant D$，因此

$$\mathbb{E}\left[f(X,Y)f(X,T_\rho(Y))\right]$$

$$
\begin{aligned}
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D}} \hat{f}_{\alpha_1,t_1}\hat{f}_{\alpha_2,t_2}\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(T_\rho(y))\right] \\
&\geqslant \sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\cdot\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)h_t(T_\rho(y))\right] \\
&\quad - \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^n,t_1,t_2\in\mathbb{N}:\\ |\alpha_1|+t_1\leqslant D,|\alpha_2|+t_2\leqslant D,\\ (\alpha_1,t_1)\neq(\alpha_2,t_2)}} |\hat{f}_{\alpha_1,t_1}||\hat{f}_{\alpha_2,t_2}|\cdot\mathbb{E}_S\mathbb{E}_X\left[h_{\alpha_1}(X)h_{\alpha_2}(X)h_{t_1}(y)h_{t_2}(T_\rho(y))\right] \\
&\geqslant \left(\sqrt{1-\rho^2}\right)^D \sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\cdot\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)^2\right] - o\left(\sum_{\substack{\alpha\in\mathbb{N}^n,t\in\mathbb{N}:\\ |\alpha|+t\leqslant D}} \hat{f}_{\alpha,t}^2\right) \\
&\geqslant \left(\left(\sqrt{1-\rho^2}\right)^D - o(1)\right)\mathbb{E}\left[f(X,Y)^2\right], \qquad (34)
\end{aligned}
$$

其中在最后一个不等式中，我们再次使用了 $\mathbb{E}_S\mathbb{E}_X\left[h_\alpha(X)^2h_t(y)^2\right]\geqslant 1$。

结合（32）、（33）和（34），可得

$$
\begin{aligned}
&\mathbb{E}\left[\left(f(X,Y)-f(X,T_\rho(Y))\right)^2\right] \\
&= \mathbb{E}\left[f(X,Y)^2\right]+\mathbb{E}\left[f(X,T_\rho(Y))^2\right]-2\mathbb{E}\left[f(X,Y)f(X,T_\rho(Y))\right] \\
&\leqslant 2\left(1-\left(\sqrt{1-\rho^2}\right)^D+o(1)\right)\mathbb{E}\left[f(X,Y)^2\right].
\end{aligned}
$$

这证实，对于 GSS，次数至多为 D 的多项式 f 是 $\left(\rho,\ 2\left(1-\left(\sqrt{1-\rho^2}\right)^D+o(1)\right)\right)$-稳定的，前提是 $D=o\left(\min\left\{k^{1/4},(n/k)^{1/5}\right\}\right)$。

□

8.3. **定理8.1的推迟证明。**

证明。 显然，若 $|\alpha|=\sum_i\alpha_i$ 为奇数，则相应的厄米特多项式乘积的期望为零。因此，从现在起我们假设 $|\alpha|$ 为偶数。回忆一下，概率论厄米特多项式的生成函数为

$$
\exp\left(tx-\frac{1}{2}t^2\right)=\sum_{n=0}^{\infty}\mathrm{He}_n(x)\frac{t^n}{n!}.
$$

设 $x_1,\ldots,x_k$ 是同分布的高斯变量，其分布与 $N(0,1)$ 相同，协方差为 $R_{ij}=\mathbb{E}[x_ix_j]$。考虑生成函数的乘积

$$
\begin{aligned}
Z(t_1,\ldots,t_k) &= \exp\left(\sum_{i=1}^k\left(t_ix_i-\frac{1}{2}t_i^2\right)\right) \\
&= \exp\left(-\frac{1}{2}\sum_{i=1}^k t_i^2\right)\exp\left(\sum_{i=1}^k t_ix_i\right).
\end{aligned}
$$

对 Z 取期望，得到

$$\begin{aligned}
&\mathbb{E}\left[Z(t_1,\dots,t_k)\right]\\
&=\exp\left(-\frac{1}{2}\sum_{i=1}^{k}t_i^2\right)\mathbb{E}\left[\exp\left(\sum_{i=1}^{k}t_ix_i\right)\right]\\
&=\exp\left(-\frac{1}{2}\sum_{i=1}^{k}t_i^2\right)\exp\left(\frac{1}{2}\mathbb{E}\left[\left(\sum_{i=1}^{k}t_ix_i\right)^2\right]\right),
\end{aligned}$$

利用高斯变量的矩母函数。使用协方差 R_{ij} 展开指数内的期望，得到

$$\begin{aligned}
&\mathbb{E}\left[Z(t_1,\dots,t_k)\right]\\
&=\exp\left(-\frac{1}{2}\sum_{i=1}^{k}t_i^2\right)\exp\left(\frac{1}{2}\mathbb{E}\left[\left(\sum_{i=1}^{k}t_ix_i\right)^2\right]\right)\\
&=\exp\left(-\frac{1}{2}\sum_{i=1}^{k}t_i^2\right)\exp\left(\frac{1}{2}\sum_{i=1}^{k}t_i^2+\sum_{1\leqslant i<j\leqslant k}t_it_jR_{ij}\right)\\
&=\exp\left(\sum_{1\leqslant i<j\leqslant k}t_it_jR_{ij}\right)\\
&=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\sum_{1\leqslant i<j\leqslant k}t_it_jR_{ij}\right)^n.
\end{aligned}$$

对期望求导并在 0 处取值，得到

$$\begin{aligned}
&\frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial t_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_k}}{\partial t_k^{\alpha_k}}\mathbb{E}\left[Z(t_1,\dots,t_k)\right]\Big|_{t_1=\cdots=t_k=0}\\
&=\prod_{i=1}^{k}\alpha_i!\cdot\frac{1}{(|\alpha|/2)!}\sum_{\substack{(a_i,b_i)_{i=1}^{|\alpha|/2}:\\ 1\leqslant a_i<b_i\leqslant k,\\ \forall j\in[n],\alpha_j=\\ \sum_{i=1}^{k}\mathbf{1}\{a_i=j\}+\mathbf{1}\{b_i=j\}}}\prod_{i=1}^{k}R_{a_i,b_i}\\
&=\sum_{M\in\mathcal{M}(\alpha)}\prod_{\{a,b\}\in M}R_{j(a),j(b)},
\end{aligned}$$

其中，$\mathbf{M}(\alpha)$ 是与 α对应的构造图中的完美匹配集合。

最后，由于我们使用以$\boldsymbol{h}_n(x)$为正交归一基、并按$\boldsymbol{h}_n(x) := \frac{1}{\sqrt{n!}}\mathrm{He}_n(x)$归一化的正交归一厄米特多项式，因此有

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{i=1}^{k} h_{\alpha_i}(x_i)\right] \\
&= \frac{1}{\sqrt{\alpha!}}\mathbb{E}\left[\prod_{i=1}^{k} \mathrm{He}_{\alpha_i}(x_i)\right] \\
&= \frac{1}{\sqrt{\alpha!}}\mathbb{E}\left[\prod_{i=1}^{k}\left[\frac{\partial^{\alpha_i}}{\partial t_i^{\alpha_i}} \exp\left(t_i x_i - \frac{1}{2}t_i^2\right)\right]\Big|_{t_i=0}\right] \\
&= \frac{1}{\sqrt{\alpha!}}\frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial t_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_k}}{\partial t_k^{\alpha_k}}\mathbb{E}\left[\prod_{i=1}^{k}\exp\left(t_i x_i - \frac{1}{2}t_i^2\right)\right]\Big|_{t_1=\cdots=t_k=0} \\
&= \frac{1}{\sqrt{\alpha!}}\frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial t_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_k}}{\partial t_k^{\alpha_k}}\mathbb{E}\left[Z(t_1,\ldots,t_k)\right]\Big|_{t_1=\cdots=t_k=0} \\
&= \frac{1}{\sqrt{\alpha!}}\sum_{M\in\mathcal{M}(\alpha)}\prod_{\{a,b\}\in M} R_{j(a),j(b)},
\end{aligned}
$$

正如所期望的那样。证明至此完成。 □

参考文献

[ACO08] Dimitris Achlioptas 和 Amin Coja-Oghlan。相变带来的算法障碍。载于《2008年第49届IEEE计算机科学基础年度研讨会》，第793–802页。IEEE，2008年。[ACOGM20] Peter Ayre、Amin Coja-Oghlan、Pu Gao 和 No¨ela M¨uller。随机线性方程的可满足性阈值。Combinatorica，40(2):179–235，2020年。[AM15] Dimitris Achlioptas 和 Michael Molloy。随机线性方程的解空间几何。Random Structures & Algorithms，46(2):197–231，2015年。[BH22] Guy Bresler 和 Brice Huang。低次多项式下随机 k-SAT 的算法相变。载于《2021年第62届IEEE计算机科学基础年度研讨会（FOCS）》，第298–309页。IEEE，2022年。[BHJK25] Rares-Darius Buhai、Jun-Ting Hsieh、Aayush Jain 和 Pravesh K Kothari。准多项式低次猜想是错误的。arXiv 预印本 arXiv:2505.17360，2025年。[COGHK+22] Amin Coja-Oghlan、Oliver Gebhard、Max Hahn-Klimroth、Alexander S Wein 和 Ilias Zadik。群组测试中的统计与计算相变。载于《学习理论会议》，第4764–4781页。PMLR，2022年。[DMK+16] Mohamad Dia、Nicolas Macris、Florent Krzakala、Thibault Lesieur、Lenka Zdeborov´a 等。对称秩一矩阵估计的互信息：副本公式的证明。《神经信息处理系统进展》，29，2016年。[EGV25] Bertrand Even、Christophe Giraud 和 NicolasVerzelen。潜变量模型的计算下界：聚类、稀疏聚类、双聚类。arXiv 预印本 arXiv:2506.13647，2025年。[Fri86] Alan M Frieze。关于用于子集和问题的 Lagarias-Odlyzko 算法。SIAM Journal on Computing，15(2):536–539，1986年。[Gam21] David Gamarnik。重叠间隙性质：优化随机结构的拓扑障碍。《美国国家科学院院刊》，118(41):e2108492118，2021年。[GJW24] David Gamarnik、Aukosh Jagannath 和 Alexander S Wein。布尔电路、低次多项式和朗之万动力学随机优化问题的困难性。SIAM Journal on Computing，53(1):1–46，2024年。[GMZ23] David Gamarnik、Elchanan Mossel 和 Ilias Zadik。尖锐阈值蕴含电路下界：从随机 2-SAT 到植入团。arXiv 预印本 arXiv:2311.04204，2023年。[GS14] David Gamarnik 和 Madhu Sudan。稀疏随机图上局部算法的极限。载于《第五届理论计算机科学创新会议》，第369–376页，2014年。[GZ22] David Gamarnik 和 Ilias Zadik。稀疏高维线性回归：平方误差估计与相变。《统计年鉴》，50(2):880–903，2022年。[GZ24] David Gamarnik 和 Ilias Zadik。植入团问题的景观：稠密子图与重叠间隙性质。《应用概率年刊》，34(4):3375–3434，2024年。[HKK+26] Jun-Ting Hsieh、Daniel M Kane、Pravesh K Kothari、Jerry Li、Sidhanth Mohanty 和 Stefan Tiegel。低次启发式的严格推论。arXiv 预印本 arXiv:2601.05850，2026年。[Hop18] Samuel Hopkins。统计推断与平方和方法。康奈尔大学，2018年。[HS25] Brice Huang 和 Mark Sellke。优化平均场自旋玻璃的紧致利普希茨困难性。Communications on Pure and Applied Mathematics，78(1):60–119，2025年。[IKKM12] Morteza Ibrahimi、Yashodhan Kanoria、Matt Kraning 和 Andrea Montanari。随机 XOR-SAT 公式的解集。载于《第二十三届ACM-SIAM离散算法年度研讨会》，第760–779页。SIAM，2012年。[LM17] Marc Lelarge 和 L´eo Miolane。对称低秩矩阵估计的基本极限。载于《学习理论会议》，第1297–1301页。PMLR，2017年。[LMB22] Clement Luneau、Nicolas Macris 和 Jean Barbier。学习稀疏规则的信息论极限。统计力学杂志：理论与实验，2022(4):044001，2022年。

[LS24] Shuangping Li 和 Tselil Schramm。某些简单的优化问题具有重叠-间隙性质。*arXiv* 预印本 *arXiv:2411.01836*，2024 年。[MNWS+23] Elchanan Mossel、Jonathan Niles-Weed、Youngtak Sohn、Nike Sun 和 Ilias Zadik。植入子图推断中的尖锐阈值。载于第三十六届学习理论年会，第 5573–5577 页。PMLR，2023 年。[MNWSZ25] Elchanan Mossel、Jonathan Niles-Weed、Nike Sun 和 Ilias Zadik。扩散引理的贝叶斯证明。*Random Structures & Algorithms*，66(4):e70008，2025 年。[NWZ20] Jonathan Niles-Weed 和 Ilias Zadik。稀疏张量主成分分析中的全有或全无现象。神经信息处理系统进展，33:17674–17684，2020 年。[NWZ23] Jonathan Niles-Weed 和 Ilias Zadik。所谓“全有”其实是“一无所有”：无噪离散信道中的尖锐相变。*IEEE* 信息论汇刊，69(8):5188–5202，2023 年。[RXZ19] Galen Reeves、Jiaming Xu 和 Ilias Zadik。稀疏线性回归中的全有或全无现象。载于学习理论会议，第 2652–2663 页。PMLR，2019 年。[SW22] Tselil Schramm 和 Alexander S Wein。低次数多项式估计的计算障碍。统计年刊，50(3):1833–1858，2022 年。[Wei22] Alexander S Wein。最大独立集的最优低次数困难性。数学统计与学习，4(3):221–251，2022 年。[Wei25] Alexander S Wein。统计学的计算复杂性：低次数多项式带来的新见解。*arXiv* 预印本 *arXiv:2506.10748*，2025 年。[WXS22] Yihong Wu、Jiaming Xu 和 H Yu Sophie。确定随机图匹配的尖锐重构阈值。*IEEE* 信息论汇刊，68(8):5391–5417，2022 年。[ZG18] Ilias Zadik 和 David Gamarnik。利用格基约化进行高维线性回归。神经信息处理系统进展，31，2018 年。[ZK16] Lenka Zdeborová 和 Florent Krzakala。推断的统计物理：阈值与算法。*Advances in Physics*，65(5):453–552，2016 年。

附录 A. 稀疏张量 PCA 中稳定算法的失效

根据我们目前的结果，自然会想到一个问题：我们的方法是否只适用于“反例”问题，即存在多项式时间方法、但稳定算法无法完成任务的问题？本节将证明事实并非如此：我们将运用该方法证明，在标准稀疏张量主成分分析设置的“计算—统计差距”区间内，稳定方法无法奏效，而在该区间内，预计不存在任何多项式时间方法能够奏效。在确立这一结果的过程中，我们还对该模型呈现的全有或全无现象的窗口大小作出了紧确改进，优于此前的工作 [NWZ20]所采用的不同方法。

模型如下。

定义 A.1。令 k，d，$n \in \mathbb{N}$ 且 $\lambda \in \mathbb{R}$ 满足 $k \leqslant n$。在 k-稀疏张量主成分分析问题中，我们观测到 $Y = \sqrt{\lambda} x^{\otimes d} + W$，其中 x 从 $\{0, 1/\sqrt{k}\}^n \subseteq \mathbb{R}^n$ 中恰有 k 个非零元素的 k- 稀疏向量集合中均匀随机抽取，而 $W \in (\mathbb{R}^n)^{\otimes d}$ 是一个具有独立同分布 $N(0, 1)$ 元素的张量。统计学家的目标是根据 Y 估计 x。

我们考虑的噪声算子仍然是 OU 算子。

注 A.2（OU 算子对稀疏张量主成分分析的影响）。容易看出，对 λ 处的稀疏张量主成分分析应用 OU 噪声算子 T_ρ，等价于观测一个新的稀疏张量主成分分析模型，其 k，d，n 保持不变，但具有新的 $\tilde{\lambda} = \lambda(1-\rho^2)$。

尽管已知在尖锐阈值 $\lambda_c = 2\log\binom{n}{k}$ [NWZ20]，处，稀疏张量主成分分析会出现全有或全无现象，但先前研究得到的临界窗口大小仅为 $o(\lambda_c)$，这不足以让我们应用自己的框架。因此，我们首先针对某种稀疏张量主成分分析模型，得出以下关于临界窗口大小的结果，其证明推迟至第A.1.2节。

命题 A.3. 固定 $\varepsilon > 0$。假设 $d \geqslant 3$ 且 $k \leqslant$ 最小值 $\left\{n^{\frac{1}{2}-\varepsilon}, n^{\frac{d-2}{d+2}-\varepsilon}\right\}$。那么，对于任意满足以下条件的 $\Delta > 0$

$$\Delta = \omega\left(\sqrt{\log\binom{n}{k}}\right),$$

在稀疏张量主成分分析模型中，均有

- 若 $\lambda = 2\log\binom{n}{k} + \Delta$，则最小均方误差 $\to 0$，
- 若 $\lambda = 2\log\binom{n}{k} - \Delta$，则最小均方误差 $\to 1$。

除了展示稳定算法的分离外，我们还通过证明低次多项式的稳定性，得出了其困难性。

定理A.4。设 $f : \mathbb{R}^{[n]^d} \to \mathbb{R}$ 是一个次数至多为 D 的多项式。假设 $D = o\left(\text{最小值}\left\{\left(\frac{k^d}{\lambda}\right)^{\frac{1}{d+\frac{1}{2}}}, \frac{n}{k^2}\right\}\right)$。那么，$f$ 对于由 n、k、d 和 λ 参数化的稀疏张量 PCA 问题是 $\left(\rho, 2\left(1-\left(\sqrt{1-\rho^2}\right)^D + o(1)\right)\right)$-稳定的。

该结果同样可由一个精巧的组合论论证以及一个关于厄米特多项式期望的图公式得出。证明推迟至第 A.1.1节。

作为推论，当 λ 接近该模型的临界窗口时，我们得到以下关于稀疏张量主成分分析的低次 MMSE 下界。

推论A.5。 固定 $\varepsilon > 0$。假设 $d \geqslant 3, k \leqslant$最小值 $\left\{n^{\frac{1}{2}-\varepsilon}, n^{\frac{d-2}{d+2}-\varepsilon}\right\}$，并且 $\Delta > 0$使得

$$\Delta = \omega\left(\sqrt{\log\binom{n-k}{k}}\right).$$

令 $\lambda = 2$对数 $\binom{n-k}{k}+\Delta$。注意，根据命题 A.3，对于该 0，*MMSE* 趋于 λ。如果 $D = o\left(\text{最小值}\left\{\binom{k^d}{\lambda^2}^{d+\frac{1}{2}}, \frac{n}{k^2}\right\}\right)$并且 $\rho \geqslant \sqrt{\frac{2\Delta}{\lambda}}$，则任意次数为 D的多项式 $f\colon \mathbf{R}^{[n]^d} \to \mathbf{R}^n$ 的均方误差至少为

$$\mathbb{E}\left[\|f(Y)-x\|_2^2\right] \geqslant 1 - O\left(\sqrt{1-\left(\sqrt{1-\rho^2}\right)^2}\right) - o(1).$$

特别地，如果 $D \leqslant \frac{1}{\rho^2}$，则有

$$\mathbb{E}\left[\|f(Y)-x\|_2^2\right] \geqslant 1 - O\left(\rho^2 D\right) - o(1).$$

A.1. **证明**。 我们需要以下定理，其证明与定理 8.1的证明几乎完全相同。

定理A.6 （厄米特多项式乘积期望的图示公式）。 假设x_1、……、 x_k是服从 $x_i \sim N(\mu_i$、 1)的高斯变量。令 $R_{ij} = \mathbb{E}[x_i x_j]$表示它们的相关系数。给定 $\alpha \in \mathbf{N}^k$，定义以下图 $G(\alpha) = (V$、 $E)$，其有 $|\alpha|$ 个顶点。对于每个$i \in [k]$，创建 α_i 个对应于 x_i的顶点，并令 $j(v) = i$ 对应于与 x_i相关联的 α_i 个顶点中的每一个。对于 u、 $v \in V$，若 $j(u) \neq j(v)$，则添加一条边 $\{u$、 $v\}$。于是，我们有

$$\mathbb{E}\left[\prod_{i=1}^{k} h_{\alpha_i}(x_i)\right] = \frac{1}{\sqrt{\alpha!}} \sum_{M \in \mathcal{M}(\alpha)} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} R_{j(a)j(b)},$$

其中，$\mathcal{M}(\alpha)$ 表示上述图 $G(\alpha)$ 中（部分）匹配的集合。

A.1.1. 低次数多项式的稳定性。

定理证明*A.4.* 考虑一个次数至多为 f的多项式 D，其唯一表示为

$$f(Y) = \sum_{\alpha \in \mathbb{N}^{n^d} : |\alpha| \leqslant D} \hat{f}_\alpha h_\alpha(Y).$$

我们将为 $\mathbb{E}[f(Y)^2]$ 和 $\mathbb{E}[f(Y)f(T_\rho(Y))]$推导下界，并为 $\mathbb{E}[f(T_\rho(Y))^2]$推导上界。

首先，我们应用图公式来计算 α_1， $\alpha_2 \in \mathbf{N}^{[n]^d}$：

$$\begin{aligned}&\mathbb{E}[h_{\alpha_1}(Y)h_{\alpha_2}(Y)]\\ &= \mathbb{E}_x \mathbb{E}[h_{\alpha_1}(Y)h_{\alpha_2}(Y)|x]\end{aligned}$$

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \sum_{M \in \mathcal{M}(\alpha_1, \alpha_2)} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} R_{j(a)j(b)} \middle| x \right].$$

从这个公式可以看出，所有内层期望都是非负的。接下来，我们分析图 $G(\alpha_1, \alpha_2)$ 中的匹配。令 $S \subset [n]$ 表示随机向量 x 的支撑集。注意，在稀疏主成分分析中， x 由 S 唯一确定，反之亦然。

对于固定的 x，我们看到每个矩阵元素 $Y = \sqrt{\lambda} x^{\otimes d} + W$ 均独立分布为 $Y_i \sim N(X_i, 1)$，其中 $X_i = \sqrt{\lambda}(x^{\otimes d})_i$ 对 $i \in [n]^d$ 成立。因此，有贡献的匹配 $M \in \mathbf{M}(\alpha_1, \alpha_2)$ 只能包含连接如下两个顶点的边：一个顶点对应于来自 $\boldsymbol{h}_{\alpha_1}$ 的 Y_i，另一个顶点对应于来自 $\boldsymbol{h}_{\alpha_2}$ 的相同 Y_i，因为对于固定的 x， Y_i 和 Y_j 在 $i \neq j$ 时不相关。因此，上述内层期望可张量化为关于每个坐标的乘积：

$$\begin{aligned}
&\mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \sum_{M \in \mathcal{M}(\alpha_1, \alpha_2)} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} R_{j(a)j(b)} \middle| x \right] \\
&= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{M \in \mathcal{M}(\alpha_1(\bar{i}), \alpha_2(\bar{i}))} \prod_{v \notin V(M)} X_{\bar{i}} \middle| x \right] \\
&= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{t=0}^{\alpha_1(\bar{i}) \wedge \alpha_2(\bar{i})} \binom{\alpha_1(\bar{i})}{t} \binom{\alpha_2(\bar{i})}{t} t! \cdot X_{\bar{i}}^{\alpha_1(\bar{i}) + \alpha_2(\bar{i}) - 2t} \middle| x \right].
\end{aligned}$$

当 $\alpha_1 = \alpha_2 = \alpha$ 时，对角项至少为

$$\begin{aligned}
&\mathbb{E}[h_\alpha(Y)^2] \\
&= \mathbb{E}_x \left[\frac{1}{\alpha!} \prod_{\bar{i} \in [n]^d} \sum_{t=0}^{\alpha(\bar{i})} \binom{\alpha(\bar{i})}{t}^2 t! \cdot X_{\bar{i}}^{2\alpha(\bar{i}) - 2t} \middle| x \right] \\
&\geqslant \mathbb{E}_x \left[\frac{1}{\alpha!} \prod_{\bar{i} \in [n]^d} \alpha(\bar{i})! \middle| x \right] \\
&= 1.
\end{aligned}$$

含有 $\alpha_1 \neq \alpha_2$ 的非对角项可由以下上界估计

$$\begin{aligned}
&\mathbb{E}[h_{\alpha_1}(Y) h_{\alpha_2}(Y)] \\
&= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{t=0}^{\alpha_1(\bar{i}) \wedge \alpha_2(\bar{i})} \binom{\alpha_1(\bar{i})}{t} \binom{\alpha_2(\bar{i})}{t} t! \cdot X_{\bar{i}}^{\alpha_1(\bar{i}) + \alpha_2(\bar{i}) - 2t} \middle| x \right]
\end{aligned}$$

现在我们观察到，t 的内层求和中相邻项的比值为 $\left(\binom{\alpha_1(i)}{t+1}\binom{\alpha_2(i)}{t+1}(t+1)!\right.$ $\left.\cdot X_i^{\alpha_1(i)+\alpha_2(i)-2(t+1)}\right)/\left(\binom{\alpha_1(i)}{t}\binom{\alpha_2(i)}{t}t!\cdot X_i^{\alpha_1(i)+\alpha_2(i)-2t}\right)=\frac{(\alpha_1(i)-t)(\alpha_2(i)-t)}{X_i^2(t+1)}\geqslant (X_i^2(t+1))^{-1}$。注意，当 X_i 非零时，$t+1\leqslant D$ 和 $X_i^2=(1/\sqrt{k})^{2d}=1/k^d$。因此，只要满足 $D=o(k^d)$，内层求和就由 $t=\alpha_1(i)\wedge\alpha_2(i)$ 处的最后一项主导，于是我们有

$$
\begin{aligned}
&\leqslant \mathbb{E}_x\left[\frac{1}{\sqrt{\alpha_1!\alpha_2!}}\prod_{\bar{i}\in[n]^d}\left(1+O\left(\frac{k^d}{D}\right)\right)\binom{\alpha_1(\bar{i})\vee\alpha_2(\bar{i})}{\alpha_1(\bar{i})\wedge\alpha_2(\bar{i})}(\alpha_1(\bar{i})\wedge\alpha_2(\bar{i}))!X_{\bar{i}}^{2|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\middle| x\right]\\
&=\mathbb{E}_x\left[\left(1+O\left(\frac{k^d}{D}\right)\right)\frac{1}{\sqrt{\alpha_1!\alpha_2!}}\prod_{\bar{i}\in[n]^d}\frac{(\alpha_1(\bar{i})\vee\alpha_2(\bar{i}))!}{|\alpha_1(\bar{i})-\alpha_2(\bar{i})|!}X_{\bar{i}}^{2|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\middle| x\right]\\
&=\mathbb{E}_x\left[\left(1+O\left(\frac{k^d}{D}\right)\right)\prod_{\bar{i}\in[n]^d}\sqrt{\frac{(\alpha_1(\bar{i})\vee\alpha_2(\bar{i}))!}{(\alpha_1(\bar{i})\wedge\alpha_2(\bar{i}))!}\frac{1}{|\alpha_1(\bar{i})-\alpha_2(\bar{i})|!}}X_{\bar{i}}^{2|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\middle| x\right]\\
&\leqslant \mathbb{E}_x\left[\left(1+O\left(\frac{k^d}{D}\right)\right)\prod_{\substack{\bar{i}\in[n]^d:\\ \alpha_1(\bar{i})\neq\alpha_2(\bar{i})}}\left(\frac{e\sqrt{\alpha_1(\bar{i})\wedge\alpha_2(\bar{i})}}{|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\cdot X_{\bar{i}}^2\right)^{|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\middle| x\right]\\
&\leqslant \mathbb{E}_x\left[\left(1+O\left(\frac{k^d}{D}\right)\right)\prod_{\substack{\bar{i}\in[n]^d:\\ \alpha_1(\bar{i})\neq\alpha_2(\bar{i})}}\left(e\sqrt{D}X_{\bar{i}}^2\right)^{|\alpha_1(\bar{i})\triangle\alpha_2(\bar{i})|}\middle| x\right]\\
&\leqslant \left(1+O\left(\frac{k^d}{D}\right)\right)(e\sqrt{D})^{|\alpha_1-\alpha_2|}\mathbb{E}_x\left[\prod_{\substack{\bar{i}\in[n]^d:\\ \alpha_1(\bar{i})\neq\alpha_2(\bar{i})}}X_{\bar{i}}^{2|\alpha_1(\bar{i})-\alpha_2(\bar{i})|}\middle| x\right]
\end{aligned}
$$

注意，只有当 $i\in S^d$ 时，$X_i=\sqrt{\lambda}(x^{\otimes d})_i$ 才非零，其中 $S\subseteq[n]$ 是 x 的支撑集，并且每当 $i\in S^d$ 时，$X_i=\sqrt{\lambda}/k^{d/2}$。因此，我们得到

$$
\begin{aligned}
&=\left(1+O\left(\frac{k^d}{D}\right)\right)(e\sqrt{D})^{|\alpha_1\triangle\alpha_2|}\left(\frac{\lambda}{k^d}\right)^{|\alpha_1\triangle\alpha_2|}\Pr(V(\alpha_1\triangle\alpha_2)\subseteq S)\\
&\leqslant\left(1+O\left(\frac{k^d}{D}\right)\right)\left(\frac{e\sqrt{D}\lambda}{k^d}\right)^{|\alpha_1\triangle\alpha_2|}\left(\frac{k}{n}\right)^{v(\alpha_1\triangle\alpha_2)}.
\end{aligned}
$$

现在，我们来求 $\mathbb{E}[f(Y)^2]$ 的下界。我们有

$$\mathbb{E}[f(Y)^2]$$

$$= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|,|\alpha_2|\leqslant D}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right]$$

$$= \sum_{\alpha\in\mathbb{N}^{[n]^d}:|\alpha|\leqslant D} \hat{f}_{\alpha}^2\mathbb{E}\left[h_{\alpha}(Y)^2\right] + \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|,|\alpha_2|\leqslant D,\\ \alpha_1\neq\alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right].$$

我们知道，对角项的贡献至少为

$$\sum_{\alpha\in\mathbb{N}^{[n]^d}:|\alpha|\leqslant D} \hat{f}_{\alpha}^2\mathbb{E}\left[h_{\alpha}(Y)^2\right]$$
$$\geqslant \sum_{\alpha\in\mathbb{N}^{[n]^d}:|\alpha|\leqslant D} \hat{f}_{\alpha}^2.$$

现在我们来考虑非对角项。对于固定的 $\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d}$，我们记

$$v_1 = |V(\alpha_1\triangle\alpha_2) - V(\alpha_2)| \tag{35}$$
$$v_2 = |V(\alpha_1\triangle\alpha_2) - V(\alpha_1)| \tag{36}$$
$$m = |\alpha_1\triangle\alpha_2|. \tag{37}$$

显然，我们有 $v_1 + v_2 \leqslant |V(\alpha_1\triangle\alpha_2)|$，若 $\alpha_1 \neq \alpha_2$，则有 $m \geqslant 1$。令 $W_{\alpha_1}(v_1, v_2, m)$表示满足上述参数条件且满足 $\alpha_2 \neq \alpha_1$的 $\alpha_2 \in \mathbb{N}^{[n]^d}$ 集合。那么，非对角项的贡献至多为

$$\left|\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|,|\alpha_2|\leqslant D,\\ \alpha_1\neq\alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2}\mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right]\right|$$
$$\leqslant \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|,|\alpha_2|\leqslant D,\\ \alpha_1\neq\alpha_2}} |\hat{f}_{\alpha_1}||\hat{f}_{\alpha_2}| \cdot |\mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right]|$$
$$\leqslant 2\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}} |\hat{f}_{\alpha_1}| \sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}} \sum_{\substack{\alpha_2\in W_{\alpha_1}(v_1,v_2,m):\\ |\hat{f}_{\alpha_2}|n^{\frac{v_2}{2}}(dD)^{\frac{v_1}{2}}\\ \leqslant|\hat{f}_{\alpha_1}|n^{\frac{v_1}{2}}(dD)^{\frac{v_2}{2}}}} |\hat{f}_{\alpha_2}| \cdot \mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right]$$

这里的不等式成立，是因为每一对有序的 $(\alpha_1,\ \alpha_2)$ 都可以重新排序，使得$|\hat{f}_{\alpha_2}|n^{\frac{v_2}{2}}(dD)^{\frac{v_1}{2}} \leqslant |\hat{f}_{\alpha_1}|n^{\frac{v_1}{2}}(dD)^{\frac{v_2}{2}}$，

$$\leqslant 2\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}} n^{\frac{v_1-v_2}{2}}(dD)^{\frac{v_2-v_1}{2}}\sum_{\substack{\alpha_2\in W_{\alpha_1}(v_1,v_2,m):\\ |\hat{f}_{\alpha_2}|n^{\frac{v_2}{2}}(dD)^{\frac{v_1}{2}}\\ \leqslant|\hat{f}_{\alpha_1}|n^{\frac{v_1}{2}}(dD)^{\frac{v_2}{2}}}}\left(1+O\left(\frac{k^d}{D}\right)\right)\left(\frac{e\sqrt{D}\lambda}{k^d}\right)^{|\alpha_1\triangle\alpha_2|}\left(\frac{k}{n}\right)^{v(\alpha_1\triangle\alpha_2)}$$

$$\leqslant\left(2+O\left(\frac{k^d}{D}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}} n^{\frac{v_1-v_2}{2}}(dD)^{\frac{v_2-v_1}{2}}\sum_{\substack{\alpha_2\in W_{\alpha_1}(v_1,v_2,m):\\ |\hat{f}_{\alpha_2}|n^{\frac{v_2}{2}}(dD)^{\frac{v_1}{2}}\\ \leqslant|\hat{f}_{\alpha_1}|n^{\frac{v_1}{2}}(dD)^{\frac{v_2}{2}}}}\left(\frac{e\sqrt{D}\lambda}{k^d}\right)^{m}\left(\frac{k}{n}\right)^{v(\alpha_1\triangle\alpha_2)}.$$

接下来，我们对 $\alpha_2\in W_{\alpha_1}(v_1,v_2,m)$ 的数量进行界定。为枚举这类 α_2，我们首先枚举 v_2 在 $V(\alpha_1\triangle\alpha_2)-V(\alpha_1)$ 中的指标，这至多有 n^{v_2} 种选择。然后，我们枚举 v_1 在 $V(\alpha_1\triangle\alpha_2)-V(\alpha_2)$ 中的指标，这至多有 $(dD)^{v_1}$ 种选择，因为 $V(\alpha_1\triangle\alpha_2)-V(\alpha_2)\subseteq V(\alpha_1)$ 且 $|V(\alpha_1)|\leqslant d|\alpha_1|\leqslant dD$。最后，给定先前能够唯一确定 $V(\alpha_1)\cup V(\alpha_2)$ 的选择后，我们枚举 α_1 在 $V(\alpha_1)\cup V(\alpha_2)$ 上的 $\triangle\alpha_2$。注意，α_1 在 $V(\alpha_1)\cup V(\alpha_2)$ 上的 $\triangle\alpha_2$ 的选择数，上界为从 $\binom{V(\alpha_1)\cup V(\alpha_2)}{d}$ 中可重复地选取 $m=|\alpha_1\triangle\alpha_2|$ 元组的方式数，因此至多有 $\binom{|V(\alpha_1)\cup V(\alpha_2)|}{d}^m\leqslant\binom{d(|\alpha_1|+|\alpha_2|)}{d}^m\leqslant\binom{2dD}{d}^m$ 种选择。因此，$n^{v_2}(dD)^{v_1}\binom{2dD}{d}^m$ 是 $\alpha_2\in W_{\alpha_1}(v_1,v_2,m)$ 数量的一个上界。回到对非对角项的界定，我们得到

$$\left(2+O\left(\frac{k^d}{D}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}} n^{\frac{v_1-v_2}{2}}(dD)^{\frac{v_2-v_1}{2}}\sum_{\substack{\alpha_2\in W_{\alpha_1}(v_1,v_2,m):\\ |\hat{f}_{\alpha_2}|n^{\frac{v_2}{2}}(dD)^{\frac{v_1}{2}}\\ \leqslant|\hat{f}_{\alpha_1}|n^{\frac{v_1}{2}}(dD)^{\frac{v_2}{2}}}}\left(\frac{e\sqrt{D}\lambda}{k^d}\right)^{m}\left(\frac{k}{n}\right)^{v(\alpha_1\triangle\alpha_2)}$$

$$\leqslant\left(2+O\left(\frac{k^d}{D}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}} n^{\frac{v_1-v_2}{2}}(dD)^{\frac{v_2-v_1}{2}}n^{v_2}(dD)^{v_1}\binom{2dD}{d}^m\left(\frac{e\sqrt{D}\lambda}{k^d}\right)^{m}\left(\frac{k}{n}\right)^{v_1+v_2}$$

这里用到了 $\alpha_2\in W_{\alpha_1}(v_1、\ v_2、\ m)$ 的选择数量至多为 $n^{v_2}(dD)^{v_1}\binom{2dD}{d}^m$，且 $v(\alpha_1\triangle\alpha_2)\geqslant v_1+v_2$。

$$\leqslant\left(2+O\left(\frac{k^d}{D}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{\substack{v_1,v_2\geqslant 0,\\ m\geqslant 1}}\left(\frac{e\sqrt{D}\lambda}{k^d}\binom{2dD}{d}\right)^m\left(\frac{k\sqrt{dD}}{\sqrt{n}}\right)^{v_1+v_2}$$

$$=\left(2+O\left(\frac{k^d}{D}\right)\right)\sum_{\substack{\alpha_1\in\mathbb{N}^{[n]^d}:\\ |\alpha_1|\leqslant D}}\hat{f}_{\alpha_1}^2\sum_{m\geqslant 1}\left(\frac{e\sqrt{D}\lambda}{k^d}\binom{2dD}{d}\right)^m\sum_{t\geqslant 0}(t+1)\left(\frac{k\sqrt{dD}}{\sqrt{n}}\right)^t$$

$$= \left(2 + O\left(\frac{k^d}{D}\right)\right) \sum_{\substack{\alpha_1 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1| \leqslant D}} \hat{f}_{\alpha_1}^2 \cdot O\left(\frac{D^{d+\frac{1}{2}}\lambda}{k^d}\right)\left(1 + O\left(\frac{D^{d+\frac{1}{2}}\lambda}{k^d}\right) + O\left(\frac{k\sqrt{dD}}{\sqrt{n}}\right)\right)$$

$$= o\left(\sum_{\substack{\alpha \in \mathbb{N}^{[n]^d}: \\ |\alpha| \leqslant D}} \hat{f}_{\alpha}^2\right), \tag{38}$$

前提是 $D = o\left(\min\left\{\binom{k^d}{\lambda}^{\frac{1}{d+\frac{1}{2}}}, \frac{n}{k^2}\right\}\right)$。

结合对角项和非对角项的贡献，我们得到

$$\begin{aligned}
&\mathbb{E}[f(Y)^2] \\
&= \sum_{\alpha \in \mathbb{N}^{[n]^d}: |\alpha| \leqslant D} \hat{f}_{\alpha}^2 \mathbb{E}\left[h_\alpha(Y)^2\right] + \sum_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1|, |\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(Y)\right] \\
&\geqslant \sum_{\substack{\alpha \in \mathbb{N}^{[n]^d}: \\ |\alpha| \leqslant D}} \hat{f}_{\alpha}^2 \mathbb{E}\left[h_\alpha(Y)^2\right] - o\left(\sum_{\substack{\alpha \in \mathbb{N}^{[n]^d}: \\ |\alpha| \leqslant D}} \hat{f}_{\alpha}^2\right) \\
&= (1 - o(1)) \sum_{\substack{\alpha \in \mathbb{N}^{[n]^d}: \\ |\alpha| \leqslant D}} \hat{f}_{\alpha}^2 \mathbb{E}\left[h_\alpha(Y)^2\right],
\end{aligned}$$

其中我们用到了，对于任意 $\alpha \in \mathbb{N}^{[n]^d}$，$\mathbb{E}\left[h_\alpha(Y)^2\right] \geqslant 1$。

接下来，我们给出 $\mathbb{E}[f(T_\rho(Y))^2]$ 的上界。注意，在给定 x 的条件下，$T_\rho(Y)$ 的每个矩阵元素都独立服从分布 $T_\rho(Y)_i \sim N(\sqrt{1-\rho^2}X_i, 1)$，其中 $X = \sqrt{\lambda} x^{\otimes d}$。因此，我们可以展开$\mathbb{E}[f(T_\rho(Y))^2]$，并类似地使用图公式分析对角项和非对角项。

我们有

$$\begin{aligned}
&\mathbb{E}[f(T_\rho(Y))^2] \\
&= \sum_{\alpha \in \mathbb{N}^{[n]^d}: |\alpha| \leqslant D} \hat{f}_{\alpha}^2 \mathbb{E}\left[h_\alpha(T_\rho(Y))^2\right] + \sum_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1|, |\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \mathbb{E}\left[h_{\alpha_1}(T_\rho(Y)) h_{\alpha_2}(T_\rho(Y))\right].
\end{aligned} \tag{39}$$

对于 $\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d}$，图公式表明

$$\begin{aligned}
&\mathbb{E}[h_{\alpha_1}(T_\rho(Y)) h_{\alpha_2}(T_\rho(Y))] \\
&= \mathbb{E}_x \mathbb{E}[h_{\alpha_1}(T_\rho(Y)) h_{\alpha_2}(T_\rho(Y)) | x]
\end{aligned}$$

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1!\alpha_2!}} \sum_{M \in \mathcal{M}(\alpha_1,\alpha_2)} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} R_{j(a)j(b)} \middle| x \right]$$

这里，在给定 x 的条件下，与 Y相比，$T_\rho(Y)$ 的协方差结构保持不变。与 Y相比，$T_\rho(Y)$ 的均值按 $\sqrt{1-\rho^2}$ 缩放。因此，使用与之前相同的论证，我们得到

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1!\alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{M \in \mathcal{M}(\alpha_1(\bar{i}),\alpha_2(\bar{i}))} \prod_{v \notin V(M)} \sqrt{1-\rho^2} X_{\bar{i}} \middle| x \right]$$

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1!\alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{t=0}^{\alpha_1(\bar{i}) \wedge \alpha_2(\bar{i})} \binom{\alpha_1(\bar{i})}{t} \binom{\alpha_2(\bar{i})}{t} t! \cdot \left(\sqrt{1-\rho^2} X_{\bar{i}} \right)^{\alpha_1(\bar{i})+\alpha_2(\bar{i})-2t} \middle| x \right],$$

其中 $X = \sqrt{\lambda}\bar{x}^{\otimes d}$。注意，上述条件期望中的每个求和项始终非负。因此，由于与 $\mathbb{E}[h_{\alpha_1}(Y)h_{\alpha_2}(Y)]$ 的公式相比，每个内层求和项都乘上了 $\sqrt{1-\rho^2}$ 的非负次幂，我们有

$$\mathbb{E}[h_{\alpha_1}(T_\rho(Y))h_{\alpha_2}(T_\rho(Y))] \leqslant \mathbb{E}[h_{\alpha_1}(Y)h_{\alpha_2}(Y)].$$

将此不等式应用于（39）中，可得

$$\begin{aligned}
&\mathbb{E}[f(T_\rho(Y))^2] \\
&= \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(T_\rho(Y))^2\right] + \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1|,|\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \mathbb{E}\left[h_{\alpha_1}(T_\rho(Y))h_{\alpha_2}(T_\rho(Y))\right] \\
&\leqslant \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y)^2\right] + \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1|,|\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} |\hat{f}_{\alpha_1}||\hat{f}_{\alpha_2}| \mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(Y)\right] \\
&\leqslant (1+o(1)) \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y)^2\right]
\end{aligned}$$

其中我们使用了界（38）。

$$\leqslant (1+o(1))\mathbb{E}[f(Y)^2]. \tag{40}$$

最后，我们给出 $\mathbb{E}[f(Y)f(T_\rho(Y))]$ 的下界。

$$\begin{aligned}
&\mathbb{E}[f(Y)f(T_\rho(Y))] \\
&= \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y)h_\alpha(T_\rho(Y))\right] + \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{N}^{[n]^d}: \\ |\alpha_1|,|\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1}\hat{f}_{\alpha_2} \mathbb{E}\left[h_{\alpha_1}(Y)h_{\alpha_2}(T_\rho(Y))\right]. \qquad (41)
\end{aligned}$$

对于 $\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d}$，图公式给出

$$\mathbb{E}[h_{\alpha_1}(Y)h_{\alpha_2}(T_\rho(Y))]$$

$$= \mathbb{E}_x \mathbb{E}[h_{\alpha_1}(Y) h_{\alpha_2}(T_\rho(Y)) \big| x]$$

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \sum_{M \in \mathcal{M}(\alpha_1, \alpha_2)} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} R_{j(a) j(b)} \Big| x \right]$$

在给定 x 的条件下，Y_i 和 $T_{\bar{\rho}}(Y)_i$ 的协方差为 $\sqrt{1-\rho^2}$，且对于 $i \neq j$，Y_i 和 $T_\rho(Y)_j$ 不相关。与 Y 相比，$T_\rho(Y)$ 的均值缩放了 $\sqrt{1-\rho^2}$。因此，内层条件期望再次张量化为各个坐标上的乘积：

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{M \in \mathcal{M}(\alpha_1(\bar{i}), \alpha_2(\bar{i}))} \prod_{v \notin V(M)} \mu_{j(v)} \prod_{\{a,b\} \in M} \sqrt{1-\rho^2} \Big| x \right]$$

$$= \mathbb{E}_x \left[\frac{1}{\sqrt{\alpha_1! \alpha_2!}} \prod_{\bar{i} \in [n]^d} \sum_{t=0}^{\alpha_1(\bar{i}) \wedge \alpha_2(\bar{i})} \binom{\alpha_1(\bar{i})}{t} \binom{\alpha_2(\bar{i})}{t} t! \cdot \left(\sqrt{1-\rho^2}\right)^{\alpha_2(\bar{i})} X_{\bar{i}}^{\alpha_1(\bar{i}) + \alpha_2(\bar{i}) - 2t} \Big| x \right],$$

其中 $X = \sqrt{\lambda} x^{\otimes d}$，我们利用了这样一个事实：$\sqrt{1-\rho^2}$ 的指数等于匹配 $M \in \mathbf{M}(\alpha_1(i)$ 中的边数 $\alpha_2(i))$，加上与 $h^{\alpha_2(i)}$ 相对应的 $\alpha_2(\bar{i})$ 顶点中未匹配顶点的数量，而该指数恰好为 $\alpha_2(i)$。我们注意到，上述条件期望中的每个求和项始终非负，并且相较于 $\mathbb{E}[h_{\alpha_1}(Y) h_{\alpha_2}(Y)]$ 的对应项，其缩放因子为 $\prod_{i \in [n]^d} \left(\sqrt{1-\rho^2}\right)^{\alpha_2(i)} = \left(\sqrt{1-\rho^2}\right)^{|\alpha_2|}$，该因子介于 $\left(\sqrt{1-\rho^2}\right)^D$ 和 1 之间。因此，我们有

$$\left(\sqrt{1-\rho^2}\right)^D \leqslant \frac{\mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(T_\rho(Y))\right]}{\mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(Y)\right]} \leqslant 1. \tag{42}$$

将（42）的下界应用于（41）中的对角项，并将（42）的上界应用于（41）中的非对角项，可得

$$\begin{aligned}
&\mathbb{E}[f(Y) f(T_\rho(Y))] \\
&= \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y) h_\alpha(T_\rho(Y))\right] + \sum_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d} : \\ |\alpha_1|, |\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} \hat{f}_{\alpha_1} \hat{f}_{\alpha_2} \mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(T_\rho(Y))\right] \\
&\geqslant \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y) h_\alpha(T_\rho(Y))\right] - \sum_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d} : \\ |\alpha_1|, |\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} |\hat{f}_{\alpha_1}| |\hat{f}_{\alpha_2}| \mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(T_\rho(Y))\right] \\
&\geqslant \left(\sqrt{1-\rho^2}\right)^D \sum_{\alpha \in \mathbb{N}^{[n]^d} : |\alpha| \leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y)^2\right] - \sum_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^{[n]^d} : \\ |\alpha_1|, |\alpha_2| \leqslant D, \\ \alpha_1 \neq \alpha_2}} |\hat{f}_{\alpha_1}| |\hat{f}_{\alpha_2}| \mathbb{E}\left[h_{\alpha_1}(Y) h_{\alpha_2}(Y)\right]
\end{aligned}$$

$$\geqslant \left(\sqrt{1-\rho^2}\right)^D \sum_{\alpha\in\mathbb{N}^{[n]^d}:|\alpha|\leqslant D} \hat{f}_\alpha^2 \mathbb{E}\left[h_\alpha(Y)^2\right] - o\left(\sum_{\alpha\in\mathbb{N}^{[n]^d}:|\alpha|\leqslant D} \hat{f}_\alpha^2\right)$$

其中我们使用了界（38）

$$\geqslant \left(\left(\sqrt{1-\rho^2}\right)^D - o(1)\right) \mathbb{E}\left[f(Y)^2\right]. \tag{43}$$

结合（40）和（43），我们得到

$$\begin{aligned}
&\mathbb{E}\left[\left(f(Y) - f(T_\rho(Y))\right)^2\right] \\
&= \mathbb{E}[f(Y)^2] + \mathbb{E}[f(T_\rho(Y))^2] - 2\mathbb{E}\left[f(Y)f(T_\rho(Y))\right] \\
&\leqslant 2\left(1 - \left(\sqrt{1-\rho^2}\right)^D + o(1)\right) \mathbb{E}[f(Y)^2].
\end{aligned}$$

因此，对于稀疏主成分分析，只要满足 $D = o\left(\min\left\{\binom{k^d}{\lambda}^{d+\frac{1}{2}}, \frac{n}{k^2}\right\}\right)$，次数为 D 的多项式就是 $\left(\rho,\ 2\left(1 - \left(\sqrt{1-\rho^2}\right)^D + o(1)\right)\right)$-稳定的。

□

A.1.2. 稀疏张量主成分分析临界窗口/最小均方误差不稳定性的延后证明。考虑k-稀疏张量主成分分析模型，其信噪比为 λ。其最小均方误差为

$$\begin{aligned}
&\mathbb{E}\left[\|\mathbb{E}[x|Y] - x\|^2\right] \\
&= \mathbb{E}[\|x\|^2] - \mathbb{E}\left[\langle x, \mathbb{E}[x|Y]\rangle\right].
\end{aligned}$$

令 $(x,\ Y) \sim \mathrm{P}$ 表示稀疏张量主成分分析的分布，令 $x \sim \mathrm{P}(\cdot|Y)$ 表示后验分布。那么，我们有

$$\begin{aligned}
&\mathbb{E}\left[\|\mathbb{E}[x|Y] - x\|^2\right] \\
&= \mathbb{E}[\|x\|^2] - \mathbb{E}_{(x,Y)\sim\mathbb{P}}\mathbb{E}_{x'\sim\mathbb{P}(\cdot|Y)}\left[\langle x, x'\rangle\right].
\end{aligned}$$

因此，只需理解先验分布 x 与后验分布 x' 之间的期望重叠度。

令 $p_i := \Pr(\langle x,\ x'\rangle = i/k)$ 对 $0 \leqslant i \leqslant k$。于是，我们有

$$\begin{aligned}
&\mathbb{E}\left[\|\mathbb{E}[x|Y] - x\|^2\right] \\
&= \mathbb{E}[\|x\|^2] - \mathbb{E}_{(x,Y)\sim\mathbb{P}}\mathbb{E}_{x'\sim\mathbb{P}(\cdot|Y)}\left[\langle x, x'\rangle\right] \\
&= 1 - \mathbb{E}\left[\sum_{i=0}^{k} p_i \cdot \frac{i}{k}\right].
\end{aligned}$$

固定 $Y = \sqrt{\lambda}x^{\otimes d} + W$ 令 $\Omega_{n,k}$ 表示 $\{0, 1/\sqrt{k}\}^n$ 中的 k-稀疏向量集合。后验分布满足：对于 $x' \in \Omega_{n,k}$,

$$\begin{aligned}
\mathbb{P}(x'|Y) &\propto \mathbb{P}(Y|x')\mathbb{P}(x') \\
&\propto \mathbb{P}(Y|x')
\end{aligned}$$

$$
\begin{aligned}
&= \exp\left(-\|Y - \sqrt{\lambda}x'^{\otimes d}\|^2/2\right) \\
&\propto \exp\left(\sqrt{\lambda}\langle Y, x'^{\otimes d}\rangle - \lambda\|x'^{\otimes d}\|^2/2\right) \\
&= \exp\left(\sqrt{\lambda}\langle\sqrt{\lambda}x^{\otimes d} + W, x'^{\otimes d}\rangle - \lambda/2\right) \\
&\propto \exp\left(\lambda\langle x, x'\rangle^d + \sqrt{\lambda}\langle W, x'^{\otimes d}\rangle\right).
\end{aligned}
$$

因此，若定义

$$
\begin{aligned}
S_i &:= \sum_{\substack{x'\in\Omega_{n,k}:\\ \langle x,x'\rangle = i/k}} \exp\left(\lambda\langle x, x'\rangle^d + \sqrt{\lambda}\langle W, x'^{\otimes d}\rangle\right) \\
&= \sum_{\substack{x'\in\Omega_{n,k}:\\ \langle x,x'\rangle = i/k}} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \sqrt{\lambda}\langle W, x'^{\otimes d}\rangle\right),
\end{aligned}
$$

那么我们有 $p_i = \frac{S_i}{\sum_{j=0}^{k} S_j}$ 在开始分析稀疏张量主成分分析的临界窗口之前，先陈述几个有用的引理。

引理 A.7。 对于 $0 \leqslant i \leqslant k$，令 $N_i := \binom{k}{i}\binom{n-k}{k-i}$。则，

$$
\frac{N_i}{N_0} \leqslant \frac{1}{i!}\left(\frac{k^2}{n-2k}\right)^i.
$$

引理*A.7*的证明。

$$
\begin{aligned}
\frac{N_i}{N_0} &= \binom{k}{i}\frac{\binom{n-k}{k-i}}{\binom{n-k}{k}} \\
&\leqslant \frac{k^i}{i!}\cdot\frac{k^i}{(n-2k)^i} \\
&= \frac{1}{i!}\left(\frac{k^2}{n-2k}\right)^i.
\end{aligned}
$$

□

引理 A.8。 设 Z_1、……、Z_m 为高斯变量，且各自服从 $N(0,\sigma^2)$。那么，对于任意满足以下条件的 $t \geqslant 0$

$$
\sqrt{2\sigma^2(t+\log m)} \leqslant \sigma^2,
$$

我们有

$$
\Pr\left(\log\left(\sum_{i=1}^{m}\exp(Z_i)\right) \leqslant \sqrt{2\sigma^2(t+\log m)} + \log\left(1+\sqrt{2\sigma^2(t+\log m)}\right)\right) \geqslant 1 - 2\exp(-t).
$$

引理证明*A.8*. 由并集界可得

$$
\Pr(\max_i Z_i \geqslant s) \leqslant m\Pr(Z_1 \geqslant s)
$$

$$\leqslant m \exp\left(-\frac{s^2}{2\sigma^2}\right).$$

因此,

$$\begin{aligned}\Pr(\max_i Z_i \geqslant \sqrt{2\sigma^2(t+\log m)}) &\leqslant m \Pr(Z_1 \geqslant \sqrt{2\sigma^2(t+\log m)}) \\ &\leqslant m \exp\left(-\frac{2\sigma^2(t+\log m)}{2\sigma^2}\right) \\ &= \exp(-t).\end{aligned}$$

记 $E = \{\max_i Z_i \leqslant \sqrt{2\sigma^2(t+\log m)}\}$，且该事件发生的概率至少为 $1-\exp(-t)$。我们有

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^m (\exp(Z_i)-1)\cdot \mathbf{1}_E\right] &= \sum_{i=1}^m \mathbb{E}\left[(\exp(Z_i)-1)\cdot \mathbf{1}_E\right] \\ &\leqslant \sum_{i=1}^m \mathbb{E}\left[(\exp(Z_i)-1)\cdot \mathbf{1}\{Z_i \leqslant \sqrt{2\sigma^2(t+\log m)}\}\right] \\ &\leqslant \sum_{i=1}^m \int_0^{\sqrt{2\sigma^2(t+\log m)}} \exp(s)\Pr\left(Z_i \geqslant s\right) ds \\ &\leqslant \sum_{i=1}^m \int_0^{\sqrt{2\sigma^2(t+\log m)}} \exp(s)\exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &\leqslant m\sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)\exp\left(-(t+\log m)\right)\end{aligned}$$

其中我们利用了 $\sqrt{2\sigma^2(t+\log m)} \leqslant \sigma^2$，因此 $\exp(s)\exp\left(-\frac{s^2}{2\sigma^2}\right)$ 在 $s \in [0, \sqrt{2\sigma^2(t+\log m)}]$ 时递增。

$$\leqslant \sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)} - t\right).$$

根据马尔可夫不等式,

$$\begin{aligned}&\Pr\left(\sum_{i=1}^m (\exp(Z_i)-1)\cdot \mathbf{1}_E \geqslant \sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)\right) \\ \leqslant\ & \frac{\mathbb{E}\left[\sum_{i=1}^m (\exp(Z_i)-1)\cdot \mathbf{1}_E\right]}{\sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)} \\ \leqslant\ & \exp(-t).\end{aligned}$$

注意，由于 $\sqrt{2\sigma^2(t+\log m)} \leqslant \sigma^2$，我们有

$$t + \log m = \frac{\left(\sqrt{2\sigma^2(t+\log m)}\right)^2}{2\sigma^2}$$

$$\leqslant \frac{1}{2}\sqrt{2\sigma^2(t+\log m)}.$$

$$m \leqslant \exp\left(\sqrt{2\sigma^2(t+\log m)}\right)$$

因此，以事件 E and that $\sum_{i=1}^{m}(\exp(Z_i)-1)\cdot \mathbf{1}_E \leqslant \sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)$ 为条件（该事件发生的概率至少为 $1-2\exp(-t)$），我们有

$$\begin{aligned}\log\left(\sum_{i=1}^{m}\exp(Z_i)\right) &\leqslant \log\left(m+\sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)\right)\\ &\leqslant \log\left(\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)+\sqrt{2\sigma^2(t+\log m)}\exp\left(\sqrt{2\sigma^2(t+\log m)}\right)\right)\\ &= \sqrt{2\sigma^2(t+\log m)}+\log\left(1+\sqrt{2\sigma^2(t+\log m)}\right),\end{aligned}$$

由此完成证明。 □

引理 A.9. 设 X, Y 为均值为零的高斯变量，其方差为 1，协方差为 ρ。那么，存在一个绝对常数 $C>0$，使得对所有 $u \geqslant 0$,

$$\Pr(X \geqslant u, Y \geqslant u) \leqslant \frac{C}{\sqrt{1-\rho^2}}\exp\left(-\frac{u^2}{1+\rho}\right).$$

引理 *A.9* 的证明。、 $Y)=(x$、 $y)$ 的密度为

$$\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right).$$

沿 $x \geqslant u$ 和 $y \geqslant u$积分，得到

$$\begin{aligned}&\int_u^\infty\int_u^\infty \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right)dydx\\ &= \int_0^\infty\int_0^\infty \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{(u+t)^2-2\rho(u+t)(u+s)+(u+s)^2}{2(1-\rho^2)}\right)dsdt\\ &= \int_0^\infty\int_0^\infty \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{2u^2(1-\rho)+2u(1-\rho)(s+t)+s^2+t^2-2\rho st}{2(1-\rho^2)}\right)dsdt\\ &\leqslant \int_0^\infty\int_0^\infty \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{2u^2(1-\rho)+2u(1-\rho)(s+t)}{2(1-\rho^2)}\right)dsdt\\ &= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{u^2}{1+\rho}\right)\int_0^\infty\int_0^\infty \exp\left(-\frac{u(s+t)}{1+\rho}\right)dsdt\\ &= \frac{(1+\rho)^2}{2\pi u^2\sqrt{1-\rho^2}}\exp\left(-\frac{u^2}{1+\rho}\right).\end{aligned}$$

□

首先，我们证明稀疏张量主成分分析临界窗口的上边界。

命题的上边界证明 A.3。假设 $d \geqslant 3$，$k \leqslant$最小值 $\left\{n^{\frac{1}{2}-\varepsilon},\ n^{\frac{d-2}{d+2}-\varepsilon}\right\}$对于某个常数 $\varepsilon > 0$，且 $\Delta > 0$ 满足

$$\Delta = \omega\left(\sqrt{\log\binom{n-k}{k}}\right).$$

假设 $\lambda = 2\log\binom{n-k}{k} + \Delta$。

固定任意一个 $\delta > 0$。我们将证明 $\sum_{i=0}^{(1-\delta)k} p_i \to 0$。为方便起见，令 $N_i := \binom{k}{i}\binom{n-k}{k-i}$。注意，

$$N_i = \left|\{x' \in \Omega_{n,k} : \langle x, x'\rangle = i/k\}\right|.$$

我们将分析分为两部分，一部分针对较小的 i，另一部分针对较大的 i。由于 $k \leqslant n^{\frac{1}{2}-\varepsilon}$，我们知道 $\log(k^2/(n-2k)) \leqslant -c_\varepsilon \log(n/k)$，其中某个常数为 $c_\varepsilon > 0$。现在选择 $\eta > 0$，使得 $3\eta^{d-1} < \frac{c_\varepsilon}{2}$。

小重叠度分析：首先，考虑 $i \leqslant \eta k$。

回忆一下，

$$\begin{aligned} S_i &= \sum_{\substack{x' \in \Omega_{n,k}: \\ \langle x, x'\rangle = i/k}} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \sqrt{\lambda}\langle W, x'^{\otimes d}\rangle\right) \\ &= \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \sum_{\substack{x' \in \Omega_{n,k}: \\ \langle x, x'\rangle = i/k}} \exp\left(\sqrt{\lambda}\left\langle W, x'^{\otimes d} - (i/k)^d x^{\otimes d}\right\rangle\right), \end{aligned}$$

注意，$\langle W, x^{\otimes d}\rangle$和 $\left\langle W, x'^{\otimes d} - (i/k)^d x^{\otimes d}\right\rangle$相互独立，且 $\left\langle W, x'^{\otimes d} - (i/k)^d x^{\otimes d}\right\rangle$服从分布 $N(0, 1-(i/k)^{2d})$。因此，我们有

$$\begin{aligned} &\mathbb{E}\left[\frac{S_i}{S_k}\middle|\langle W, x^{\otimes d}\rangle\right] \\ &= \exp\left(\lambda\left(\left(\frac{i}{k}\right)^d - 1\right) + \left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \\ &\quad \cdot \mathbb{E}\left[\sum_{\substack{x' \in \Omega_{n,k}: \\ \langle x, x'\rangle = i/k}} \exp\left(\sqrt{\lambda}\left\langle W, x'^{\otimes d} - (i/k)^d x^{\otimes d}\right\rangle\right)\right] \\ &= \exp\left(\lambda\left(\left(\frac{i}{k}\right)^d - 1\right) + \left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) N_i \exp\left(\frac{\lambda}{2}\left(1 - \left(\frac{i}{k}\right)^{2d}\right)\right) \\ &= N_i \exp\left(\lambda\left(\left(\frac{i}{k}\right)^d - 1\right) + \left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle + \frac{\lambda}{2}\left(1 - \left(\frac{i}{k}\right)^{2d}\right)\right). \end{aligned}$$

由于 $\Delta = \omega\left(\sqrt{\log\binom{n-k}{k}}\right)$，我们可以选择 $b = \sqrt{\sqrt{\Delta \log\binom{n-k}{k}}}$，使得 $b = \omega(1)$ 和 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$。考虑事件 $E_1 = \{\langle W, x^{\otimes d}\rangle \geqslant -b\}$。我们有 $\Pr(E_1) \geqslant 1 - \exp(-b^2/2) = 1 - o(1)$。因此，在事件 E_1 成立的条件下，我们有

$$\begin{aligned}
&\mathbb{E}\left[\frac{S_i}{S_k}\middle| E_1\right] \\
&\leqslant N_i \exp\left(\lambda\left(\left(\frac{i}{k}\right)^d - 1\right) - b\left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda} + \frac{\lambda}{2}\left(1 - \left(\frac{i}{k}\right)^{2d}\right)\right) \\
&\leqslant \exp\left(\log\binom{n-k}{k} + i\log\left(\frac{k^2}{n-2k}\right) - \frac{\lambda}{2} + b\sqrt{\lambda} + \lambda\left(\frac{i}{k}\right)^d\right)
\end{aligned}$$

因为 $N_i \leqslant \binom{n-k}{k}\frac{1}{i!}\left(\frac{k^2}{n-2k}\right)^i$ 由引理 A.7，

$$= \exp\left(i\log\left(\frac{k^2}{n-2k}\right) - \frac{\Delta}{2} + b\sqrt{2\log\binom{n-k}{k} + \Delta} + \left(2\log\binom{n-k}{k} + \Delta\right)\left(\frac{i}{k}\right)^d\right)$$

代入 $\lambda = 2\log\binom{n-k}{k} + \Delta$，

$$\leqslant \exp\left(-c_\varepsilon i\log(n/k) - \frac{\Delta}{2} + o(\Delta) + 3i\log(n/k)\left(\frac{i}{k}\right)^{d-1}\right)$$

利用 $\log(k^2/(n-2k)) \leqslant -c_\varepsilon \log(n/k), \Delta \leqslant \log\binom{n-k}{k}$，以及 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$

$$\leqslant \exp\left(3i\log(n/k)\eta^{d-1} - c_\varepsilon i\log(n/k) - \frac{\Delta}{2} + o(\Delta)\right)$$

因为 $i \leqslant \eta k$。最后，利用 $3\eta^{d-1} \leqslant c_\varepsilon$ 并得到

$$\leqslant \exp\left(-\frac{\Delta}{4}\right).$$

对 $i \leqslant \eta k$ 求和，得到

$$\mathbb{E}\left[\sum_{i=0}^{\eta k}\frac{S_i}{S_k}\middle| E\right] \leqslant \eta k \exp\left(-\frac{\Delta}{4}\right)$$

因为 $\Delta = \omega\left(\sqrt{\log\binom{n-k}{k}}\right)$。因此，在事件 E_1 条件下，根据马尔可夫不等式，我们有

$$\sum_{i=0}^{\eta k} p_i = o(1)$$

$$\leqslant \sum_{i=0}^{\eta k} \frac{S_i}{S_k}$$
$$\to 0.$$

由于 E_1 的概率为 $1-o(1)$，我们得出，当 $\lambda = 2\log\binom{n-k}{k} + \Delta$ 时，$\sum_{i=0}^{\eta k} p_i \to 0$。

大重叠分析：接下来，我们考虑 $\eta k \leqslant i \leqslant (1-\delta)k$。

回顾

$$S_i = \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \sum_{\substack{x'\in\Omega_{n,k}:\\ \langle x,x'\rangle = i/k}} \exp\left(\sqrt{\lambda}\left\langle W, x'^{\otimes d} - (i/k)^d\, x^{\otimes d}\right\rangle\right).$$

若记

$$Y_i = \sum_{\substack{x'\in\Omega_{n,k}:\\ \langle x,x'\rangle = i/k}} \exp\left(\sqrt{\lambda}\left\langle W, x'^{\otimes d} - (i/k)^d\, x^{\otimes d}\right\rangle\right),$$

则有

$$S_i = \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) Y_i.$$

取定 $i \in [\eta k,\ (1-\delta)k]$。由引理A.8，以至少 $1-2\exp(-t)$的概率，有

$$\Pr\left(\log(Y_i) \geqslant B_i + \log(1+B_i)\right) \leqslant 2\exp(-t_i),$$

其中 $B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(t_i + \log N_i)}$ 对于某个 $t_i \geqslant 0$，前提是

$$B_i \leqslant \lambda\left(1-(i/k)^{2d}\right).$$

现在我们选择 $t_i = 2\log k$。我们可以验证，对于 $i \leqslant (1-\delta)k$，

$$\begin{aligned} B_i &= \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} \\ &= \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(2\log k + \log\binom{k}{i} + \log\binom{n-k}{k-i}\right)} \\ &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + k + (k-i)\log(e(n-k)/(k-i)))} \end{aligned}$$

现在我们利用 $k-i \geqslant \delta k$ 并得到

$$\begin{aligned} &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)((k-i)\log((n-k)/k) + O(k))} \\ &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(\left(1-\frac{i}{k}\right)\log\binom{n-k}{k} + O(k)\right)} \end{aligned}$$

$$\leqslant \lambda\left(1-(i/k)^{2d}\right)\sqrt{\frac{\left(2\log\binom{n-k}{k}\right)(1-(i/k)+o(1))}{\lambda\left(1-(i/k)^{2d}\right)}}$$

$$\leqslant \lambda\left(1-(i/k)^{2d}\right),$$

由于 $\lambda \geqslant 2\log\binom{n-k}{k}$ 且 $1-(i/k)+o(1) \leqslant 1-(i/k)^{2d}$ 适用于 $\eta k \leqslant i \leqslant (1-\delta)k$。因此，对于任意 $\eta k \leqslant i \leqslant (1-\delta)k$，以至少 $1-2\exp(-2\log k) = 1-\frac{2}{k^2}$ 的概率，我们有

$$\log(Y_i) \leqslant B_i + \log(1+B_i),$$

其中 $B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)}$。对所有 $i \leqslant (1-\delta)k$ 应用并集界，我们以至少 $1-\frac{2}{k}$ 的概率有

$$\begin{aligned}
&\sum_{i=\eta k}^{(1-\delta)k} S_i \\
&= \sum_{i=\eta k}^{(1-\delta)k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) Y_i \\
&\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \exp\left(B_i + \log(1+B_i)\right) \\
&= S_k \sum_{i=\eta k}^{(1-\delta)k} \frac{1}{S_k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \exp\left(B_i + \log(1+B_i)\right)
\end{aligned}$$

现在我们使用 $S_k = \exp(\lambda + \sqrt{\lambda}\langle W\ x^{\otimes d}\rangle)$，得到

$$\begin{aligned}
&= S_k \sum_{i=\eta k}^{(1-\delta)k} \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\left[\lambda + \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right]\right) \exp\left(B_i + \log(1+B_i)\right) \\
&= S_k \sum_{i=\eta k}^{(1-\delta)k} \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\lambda + B_i + \log(1+B_i)\right).
\end{aligned} \tag{44}$$

将上述事件称为 E_2，其发生的概率至少为 $1-\frac{2}{k}$。

再次，令我们对事件 $E_1 = \{\langle W\ x^{\otimes d}\rangle \geqslant -b\}$进行条件化，该事件发生的概率为 $1-o(1)$。这里回顾一下 $b = \sqrt{\Delta\log\binom{n-k}{k}}$，因此 $b = \omega(1)$，且 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$。在给定 E_1 和 E_2的条件下，我们来分析（44）中的和。我们有

$$\sum_{i=\eta k}^{(1-\delta)k} \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\lambda + B_i + \log(1+B_i)\right)$$

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(-b\left(\left(\frac{i}{k}\right)^d-1\right)\sqrt{\lambda}+\left(\left(\frac{i}{k}\right)^d-1\right)\lambda+B_i(1+o(1))\right)$$

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(-b\left(\left(\frac{i}{k}\right)^d-1\right)\sqrt{\lambda}+\left(\left(\frac{i}{k}\right)^d-1\right)\lambda\right.$$

$$\left.+(1+o(1))\sqrt{2\lambda\left(1-\left(\frac{i}{k}\right)^{2d}\right)(2\log k+\log N_i)}\right)$$

现在我们重新使用先前的计算结果：$2\log k+\log N_i \leqslant (1-(i/k))\log\binom{n-k}{k}+O(k)$，以及界 $b\sqrt{\log\binom{n-k}{k}}=o(\Delta)$ 和 $\lambda=2\log\binom{n-k}{k}+\Delta\leqslant 3\binom{n-k}{k}$，从而得到

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(o(\Delta)+\left(\left(\frac{i}{k}\right)^d-1\right)\lambda+(1+o(1))\sqrt{2\lambda\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(1-\frac{i}{k}+o(1)\right)\log\binom{n-k}{k}}\right)$$

现在代入 $\lambda=2\log\binom{n-k}{k}+\Delta$，得到

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(o(\Delta)+2\log\binom{n-k}{k}\left(\left(\frac{i}{k}\right)^d-1+(1+o(1))\sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(1-\frac{i}{k}+o(1)\right)}\right)\right)$$

$$\cdot\exp\left(\Delta\left(\left(\frac{i}{k}\right)^d-1+\left(\frac{1}{2}+o(1)\right)\sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(1-\frac{i}{k}+o(1)\right)}\right)\right)$$

容易验证，存在某个常数 $\gamma_{\eta,\,\delta}>0$ 使得 $f(x)=x^d-1+\sqrt{(1-x)^{2d}(1-x)}<-\gamma_{\eta,\,\delta}$ 对所有 $\eta\leqslant x\leqslant 1-\delta$ 均成立。因此，

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(o(\Delta)-2(\gamma_{\eta,\delta}+o(1))\log\binom{n-k}{k}\right)$$

$$\leqslant \exp\left(\log k+o(\Delta)-2(\gamma_{\eta,\delta}+o(1))\log\binom{n-k}{k}\right)$$

$$=o(1).$$

将此界代回（44），并在 E_1 和 E_2 的条件下，我们看到，以至少 $1-o(1)$ 的概率，有

$$\sum_{i=\eta k}^{(1-\delta)k} p_i$$

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \frac{S_i}{S_k}$$

$$\leqslant \sum_{i=\eta k}^{(1-\delta)k} \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right) \exp\left(\left(\left(\frac{i}{k}\right)^d - 1\right)\lambda + B_i + \log(1+B_i)\right)$$
$$= o(1).$$

现在结合小重叠度和大重叠度的分析，我们看到，以概率$1-o(1)$，有

$$\sum_{i=0}^{k} p_i \cdot \frac{i}{k}$$
$$\geqslant \frac{(1-\delta)k}{k}\left(1 - \sum_{i=0}^{(1-\delta)k} p_i\right)$$
$$\geqslant 1 - \delta - o(1).$$

由于 $\delta > 0$ 是任意的，我们看到，对于 $\lambda = 2\log\binom{n-k}{k} + \Delta$，有

$$\mathrm{MMSE} = 1 - \mathbb{E}\left[\sum_{i=0}^{k} p_i \cdot \frac{i}{k}\right] \to 0.$$

□

接下来，我们证明临界窗口的下边缘。

命题 *A.3* 下边缘的证明。设 $d \geqslant 3$，$k \leqslant$最小值 $\left\{n^{\frac{1}{2}-\varepsilon},\ n^{\frac{d-2}{d+2}-\varepsilon}\right\}$_对于某个常数 $\varepsilon > 0$，并且 $\Delta > 0$ 满足

$$\Delta = \omega\left(\sqrt{\log\binom{n-k}{k}}\right).$$

设 $\lambda = 2\log\binom{n-k}{k} - \Delta$。

固定一个任意的 ry $\delta > 0$。我们将证明 $\sum_{i=\delta k}^{k} p_i \to 0$。回想一下，我们设定了 $N_i := \binom{k}{i}\binom{n-k}{k-i}$.
我们首先给出 S_0 的下界。回想一下，我们先前设定了 $b = \sqrt{\frac{\Delta}{\log\binom{n}{k}}}$令

$$M = \sum_{\substack{x' \in \Omega_{n,k}: \\ \langle x, x'\rangle = 0}} \mathbf{1}\{\langle W, x'^{\otimes d}\rangle \geqslant \sqrt{\lambda} + b\}.$$

于是，我们有如下确定性不等式 $S_0 \geqslant M\exp(\lambda + \sqrt{\lambda}b)$。

我们将对 M 使用二阶矩方法，以证明 M 的高概率下界，从而得到 S_0。M 的期望为

$$\mathbb{E}[M] = N_0 \cdot \Pr\left(\langle W, x'^{\otimes d}\rangle \geqslant \sqrt{\lambda} + b\right)$$
$$= \exp\left(\log\binom{n-k}{k} + \log\Phi(-\sqrt{\lambda} - b)\right)$$

其中 Φ 表示标准正态分布的累积分布函数。

$$= \exp\left(\log\binom{n-k}{k} - \frac{(\sqrt{\lambda}+b)^2}{2} - O(\log(\sqrt{\lambda}+b))\right)$$

利用渐近关系 $\Phi(x) \sim \frac{1}{\sqrt{2\pi}|x|}\exp(-x^2/2)$对于 $x \to -\infty$,

$$= \exp\left(\log\binom{n-k}{k} - \frac{\lambda}{2} - b\sqrt{\lambda} - \frac{b^2}{2} - O(\log(\sqrt{\lambda}+b))\right)$$

$$= \exp\left(\log\binom{n-k}{k} - \frac{\lambda}{2} - o(\Delta) - O(\log(\sqrt{\lambda}+b))\right)$$

因为 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$ 和 $b^2 = \frac{\Delta}{\sqrt{\log\binom{n}{k}}} = o(\Delta)$。现在代入 $\lambda = 2\log\binom{n-k}{k} - \Delta$，得到

$$= \exp\left(\frac{\Delta}{2} - o(\Delta)\right).$$

接下来，我们对 M的二阶矩给出上界。我们有

$$\begin{aligned}
&\frac{\mathbb{E}[M^2]}{\mathbb{E}[M]^2} \\
&= \frac{\sum_{\substack{x',x''\in\Omega_{n,k}:\\ \langle x,x'\rangle=\langle x,x''\rangle=0}} \Pr\left(\langle W, x'^{\otimes d}\rangle \geqslant \sqrt{\lambda}+b, \langle W, x''^{\otimes d}\rangle \geqslant \sqrt{\lambda}+b\right)}{\sum_{\substack{x',x''\in\Omega_{n,k}:\\ \langle x,x'\rangle=\langle x,x''\rangle=0}} \Pr\left(\langle W, x'^{\otimes d}\rangle \geqslant \sqrt{\lambda}+b\right)\Pr\left(\langle W, x''^{\otimes d}\rangle \geqslant \sqrt{\lambda}+b\right)} \\
&= \mathbb{E}_{r\sim\mathrm{Hypergeo}(n-k,k,k)} \frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2}
\end{aligned}$$

这里， r表示从 $\{x'$进行两次均匀随机抽样所得样本之间重叠部分的大小：$\Omega_{n,\ k}$：$\langle x,$ $x'\rangle = 0\}$，并且 Z和 Z_r是联合中心高斯变量，方差为 1，协方差为 $\left(\frac{r}{k}\right)^d$。

$$= \sum_{r=0}^{k} \Pr(r) \frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2}.$$

要证明 $\mathbb{E}[M^2] \leqslant (1+o(1))\mathbb{E}[M^2]$，只需证明

$$\sum_{r=1}^{k} \Pr(r) \frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} = o(1). \tag{45}$$

为证明（45），我们将分析分为三部分：小重叠度、中等重叠度和大重叠度。取 $\eta > 0$ 为一个足够小的常数，使得 $3\eta^{d-1} \leqslant \frac{c_\varepsilon}{2}$。

小重叠分析： 考虑 $1 \leqslant r \leqslant \eta k$。
根据引理A.7

$$\Pr(r) = \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n-k}{k}} = \frac{N_r}{N_0} \leqslant \frac{1}{r!}\left(\frac{k^2}{n-2k}\right)^r,$$

并根据引理 A.9 以及 $\Pr(Z \geqslant x) \geqslant \frac{1}{2x\sqrt{2\pi}}\exp(-x^2/2)$ 对于 $x \geqslant 1$,

$$\begin{aligned}
&\frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} \\
&\leqslant C\frac{1}{\sqrt{1-\left(\frac{r}{k}\right)^{2d}}}\exp\left(-\frac{(\sqrt{\lambda}+b)^2}{1+\left(\frac{r}{k}\right)^d}+\left(\sqrt{\lambda}+b\right)^2\right) \\
&\leqslant 2C\exp\left((\sqrt{\lambda}+b)^2\frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right).
\end{aligned}$$

因此,

$$\begin{aligned}
&\sum_{r=1}^{\eta k}\Pr(r)\frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} \\
&\leqslant \sum_{r=1}^{\eta k}\frac{2C}{r!}\left(\frac{k^2}{n-2k}\right)^r\exp\left((\sqrt{\lambda}+b)^2\frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right) \\
&= 2C\sum_{r=1}^{\eta k}\exp\left((\sqrt{\lambda}+b)^2\frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}+r\log\left(\frac{k^2}{n-2k}\right)-r\right)
\end{aligned}$$

由于 $\log\left(\frac{k^2}{n-2k}\right) \leqslant -c_\varepsilon\log(n/k)$ 且 $(\sqrt{\lambda}\pm b)^2 \leqslant 2\log\binom{n-k}{k}+o(\Delta) \leqslant 3k\log(n/k),$, 我们有

$$\begin{aligned}
&\leqslant 2C\sum_{r=1}^{\eta k}\exp\left(3r\log(n/k)\frac{\left(\frac{r}{k}\right)^{d-1}}{1+\left(\frac{r}{k}\right)^d}-c_\varepsilon r\log(n/k)-r\right) \\
&\leqslant 2C\sum_{r=1}^{\eta k}\exp\left(-\frac{c_\varepsilon}{2}r\log(n/k)-r\right) \\
&= o(1).
\end{aligned}$$

中等重叠分析： 考虑 $\eta k \leqslant r \leqslant (1-\eta)k$。我们有

$$\Pr(r) = \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n-k}{k}} \leqslant 2^k\left(\frac{e(n-k)}{k-r}\right)^{k-r}\left(\frac{k}{n-k}\right)^k \leqslant C_\eta^k\left(\frac{k}{n-k}\right)^r,$$

对于某个常数 $C_\eta > 0$，并且与小重叠分析中的界相同，

$$\frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} \leqslant 2C \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right).$$

因此，

$$\begin{aligned}
&\sum_{r=\eta k}^{(1-\eta)k} \Pr(r) \frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} \\
&\leqslant \sum_{r=\eta k}^{(1-\eta)k} 2C \cdot C_\eta^k \left(\frac{k}{n-k}\right)^r \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right) \\
&\leqslant 2C \sum_{r=\eta k}^{(1-\eta)k} \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d} + O(k) - r\log\left(\frac{n}{k}\right)\right)
\end{aligned}$$

由于 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$ 和 $b^2 = o(\Delta)$，我们现在使用 $(\sqrt{\lambda}+b)^2 \leqslant (\sqrt{2\log\binom{n-k}{k}-\Delta}+b)^2 \leqslant (\sqrt{2\log\binom{n-k}{k}}+b)^2 \leqslant 2\log\binom{n-k}{k} + o(\Delta)$，得到

$$\begin{aligned}
&\leqslant 2C \sum_{r=\eta k}^{(1-\eta)k} \exp\left(\left(2\log\binom{n-k}{k} + o(\Delta)\right) \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d} + O(k) - r\log\left(\frac{n}{k}\right)\right) \\
&\leqslant 2C \sum_{r=\eta k}^{(1-\eta)k} \exp\left(k\log\left(\frac{n}{k}\right)\left(\frac{2\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d} - \frac{r}{k} + o(1)\right)\right)
\end{aligned}$$

现在注意到，由于 $d \geqslant 3 \geqslant 2$，存在一个常数 $\gamma_\eta > 0$，使得对于所有 $\eta \leqslant x \leqslant 1-\eta$，$f(x) = \frac{2x^d}{1+x^d} - x \leqslant -\gamma_\eta$。因此，

$$\begin{aligned}
&\leqslant 2C \sum_{r=\eta k}^{(1-\eta)k} \exp\left(-\frac{\gamma_\eta}{2} k\log\left(\frac{n}{k}\right)\right) \\
&= o(1).
\end{aligned}$$

大重叠分析： 考虑 $(1-\eta)k \leqslant r \leqslant k$。我们有

$$\begin{aligned}
\Pr(r) &= \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n-k}{k}} \\
&= \frac{\binom{k}{k-r}\binom{n-k}{k-r}}{\binom{n-k}{k}} \\
&\leqslant \frac{1}{((k-r)!)^2} \cdot \frac{k^{k-r}(n-k)^{k-r}}{\binom{n-k}{k}},
\end{aligned}$$

并且我们得到与之前相同的界

$$\frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2} \leqslant 2C \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right).$$

因此，

$$\sum_{r=(1-\eta)k}^{k} \Pr(r) \frac{\Pr\left(Z \geqslant \sqrt{\lambda}+b, Z_r \geqslant \sqrt{\lambda}+b\right)}{\Pr\left(Z \geqslant \sqrt{\lambda}+b\right)^2}$$

$$\leqslant 2C \sum_{r=(1-\eta)k}^{k} \frac{1}{((k-r)!)^2} \cdot \frac{k^{k-r}(n-k)^{k-r}}{\binom{n-k}{k}} \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d}\right)$$

$$\leqslant 2C \sum_{r=(1-\eta)k}^{k} \exp\left((\sqrt{\lambda}+b)^2 \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d} - \log\binom{n-k}{k} + (k-r)\log(k(n-k)) - O(k-r)\right)$$

由于 $b\sqrt{\log\binom{n-k}{k}} = o(\Delta)$ and $b^2 = o(\Delta)$，现在我们使用 $(\sqrt{\lambda}+b)^2 \leqslant (\sqrt{2\log\binom{n-k}{k}} - \Delta + b)^2 \leqslant 2\log\binom{n-k}{k} - \Delta + o(\Delta)$ 得到

$$\leqslant 2C \sum_{r=(1-\eta)k}^{k} \exp\left(\left(2\log\binom{n-k}{k} - \Delta + o(\Delta)\right) \frac{\left(\frac{r}{k}\right)^d}{1+\left(\frac{r}{k}\right)^d} - \log\binom{n-k}{k} + (k-r)\log(k(n-k)) - O(k-r)\right)$$

现在令 $s = k - r$。我们有 $0 \leqslant s \leqslant \eta k$。然后，$(r/k)^d = 1 - d\frac{s}{k} + O(s^2/k^2)$，并且对于某个常数 $c_\eta > 0$，有

$$\leqslant 2C \sum_{r=(1-\eta)k}^{k} \exp\left(\left(2\log\binom{n-k}{k} - \Delta + o(\Delta)\right)\left(\frac{1}{2} - \left(\frac{1}{2} - c_\eta\right) d \cdot \frac{s}{k}\right) - \log\binom{n-k}{k} + s\log(k(n-k)) - O(s)\right)$$

$$\leqslant 2C \sum_{r=(1-\eta)k}^{k} \exp\left(-\frac{\Delta}{4} + o(\Delta) - sd(1-2c_\eta)\log\left(\frac{n-k}{k}\right) + s\log(k(n-k)) - O(s)\right)$$

$$= 2C \sum_{r=(1-\eta)k}^{k} \exp\left(-\frac{\Delta}{4} + o(\Delta) - s\left(\log\left(\frac{(n-k)^{d(1-2c_\eta)-1}}{k^{d(1-2c_\eta)+1}}\right) + O(1)\right)\right)$$

$$= o(1),$$

这里我们用到了 $k \leqslant n^{\frac{1}{2}-\varepsilon} \leqslant n^{\frac{d-1}{d+1}-\varepsilon}$。

综合所有重叠度的分析，我们得出 $\mathbb{E}[M^2] \leqslant (1+o(1))\mathbb{E}[M^2]$。因此，以概率 $1-o(1)$，有 $M \geqslant (1-o(1))\mathbb{E}[M]$ 且

$$\begin{aligned} S_0 &\geqslant (1-o(1))\mathbb{E}[M]\exp(\lambda+\sqrt{\lambda}b) \\ &\geqslant (1-o(1))\exp\left(\frac{\Delta}{2}-o(\Delta)+\lambda+\sqrt{\lambda}b\right) \\ &\geqslant (1-o(1))\exp\left(\frac{\Delta}{2}-o(\Delta)+\lambda\right). \end{aligned}$$

另一方面，根据引理A.8 如上边缘分析中所述，以至少 $1-\frac{2}{k}$ 的概率，对于所有 $\delta k \leqslant i \leqslant k$,

$$S_i \leqslant \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right)\exp\left(B_i+\log(1+B_i)\right),$$

其中 $B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k+\log N_i)}$。我们需要确保 B_i 满足 $B_i \leqslant \lambda\left(1-(i/k)^{2d}\right)$ 引理 A.8 的要求。我们很容易验证，对于 $\delta k \leqslant i \leqslant (1-\delta)k$

$$\begin{aligned} B_i &= \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k+\log N_i)} \\ &= \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(2\log k+\log\binom{k}{i}+\log\binom{n-k}{k-i}\right)} \\ &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k+k+(k-i)\log(e(n-k)/(k-i)))} \\ &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)((k-i)\log((n-k)/k)+O(k))} \\ &\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(\left(1-\frac{i}{k}\right)\log\binom{n-k}{k}+O(k)\right)} \\ &\leqslant \lambda\left(1-(i/k)^{2d}\right)\sqrt{\frac{\left(2\log\binom{n-k}{k}\right)(1-(i/k)+o(1))}{\lambda\left(1-(i/k)^{2d}\right)}} \\ &\leqslant \lambda\left(1-(i/k)^{2d}\right), \end{aligned}$$

因为 $\lambda = 2$对数 $\binom{n-k}{k}-\Delta = (2-o(1))$对数 $\binom{n-k}{k}$，且存在某个常数 $c_\delta > 0$，使得 $\frac{1-(i/k)+o(1)}{1-(i/k)^{2d}} \leqslant 1-c_\delta$ 对于 $\delta k \leqslant i \leqslant (1-\delta)k$。

现在，对于 $(1-\delta)k \leqslant i \leqslant k-1$，我们可以检查

$$B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k+\log N_i)}$$

$$= \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(2\log k + \log\binom{k}{k-i} + \log\binom{n-k}{k-i}\right)}$$
$$\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(2\log k + (k-i)\log(k(n-k)) + O(k-i)\right)}$$

现在，由于 $k \leqslant \min\{n^{\frac{d-2}{d+2}-\varepsilon}, n^{\frac{1}{2}-\varepsilon}\}$，我们有 $\log(k) \leqslant c_\varepsilon \log((n-k)/k)$，其中某个常数 $c_\varepsilon < \min\{\frac{d-2}{4}, 1\}$。因此，

$$\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left((k-i)\log((n-k)/k) + 2(k-i)\log k + O(k-i)\right)}$$
$$\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left((k-i)\log((n-k)/k) + 2c_\varepsilon(k-i)\log((n-k)/k) + O(k-i)\right)}$$
$$\leqslant \sqrt{2\lambda\left(1-(i/k)^{2d}\right)\left(\left(1-\frac{i}{k}\right)(1+2c_\varepsilon)\log\binom{n-k}{k} + O(k\log k)\right)}$$
$$\leqslant \lambda\left(1-(i/k)^{2d}\right)\sqrt{\frac{\left(2\log\binom{n-k}{k}\right)\left((1-(i/k))(1+2c_\varepsilon)+o(1)\right)}{\lambda\left(1-(i/k)^{2d}\right)}}$$
$$\leqslant \lambda\left(1-(i/k)^{2d}\right),$$

这里我们用到了 $\lambda = 2\log\binom{n-k}{k} - \Delta = (2-o(1))\log\binom{n-k}{k}$，并且 $1-(i/k)^{2d} = 2d\frac{k-i}{k} + O\left(\frac{(k-i)^2}{k^2}\right) \geqslant 2\cdot((1-(i/k))(1+2c_\varepsilon)+o(1)),$，因为 $d \geqslant 3$ 且 $c_\varepsilon < 1$。

最后，当 $i = k$ 时，由于 $B_k = 0$，无需进行任何检查。因此，我们已经证明，以至少 $1-\frac{2}{k}$ 的概率，

$$\sum_{i=\delta k}^{k} S_i \leqslant \sum_{i=\delta k}^{k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right)\exp\left(B_i + \log(1+B_i)\right),$$

其中 $B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)}$。现在，以事件 $E_1' = \{\{\langle W, x^{\otimes d}\rangle \leqslant b\}\}$ 为条件。回想一下，$b = \sqrt{\sqrt{\frac{\Delta}{\log\binom{n-k}{k}}}}$，并且 E_1' 以 $1-o(1)$ 的概率发生。因此，以至少 $1-\frac{k}{2}-o(1) = 1-o(1)$ 的概率，

$$\sum_{i=\delta k}^{k} S_i$$
$$\leqslant \sum_{i=\delta k}^{k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d \sqrt{\lambda}\langle W, x^{\otimes d}\rangle\right)\exp\left(B_i + \log(1+B_i)\right)$$
$$\leqslant \sum_{i=\delta k}^{k} \exp\left(\lambda\left(\frac{i}{k}\right)^d + \left(\frac{i}{k}\right)^d b\sqrt{\lambda}\right)\exp\left(B_i + \log(1+B_i)\right)$$

现在我们利用 $b\sqrt{\lambda} = o(\Delta)$ 以及 $B_i = \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)}$, 得到

$$\leqslant \sum_{i=\delta k}^{k} \exp\left((\lambda + o(\Delta))\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} + o(\Delta)\right).$$

回想一下，以至少 $1-o(1)$ 的概率,

$$S_0 \geqslant (1-o(1))\exp\left(\frac{\Delta}{2} - o(\Delta) + \lambda\right).$$

因此，以 $1-o(1)$ 的概率，我们有

$$\begin{aligned}
&\frac{1}{S_0}\sum_{i=\delta k}^{k} S_i \\
&\leqslant (1+o(1))\sum_{i=\delta k}^{k} \exp\left(-\frac{\Delta}{2} - \lambda + (\lambda + o(\Delta))\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} + o(\Delta)\right) \\
&\leqslant (1+o(1))\sum_{i=\delta k}^{k} \exp\left(-\frac{\Delta}{2} - \lambda + \lambda\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} + o(\Delta)\right) \\
&\leqslant (1+o(1))\sum_{i=\delta k}^{(1-\delta)k} \exp\left(-\frac{\Delta}{2} - \lambda + \lambda\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} + o(\Delta)\right) \\
&\quad + (1+o(1))\sum_{i=(1-\delta)k}^{k} \exp\left(-\frac{\Delta}{2} - \lambda + \lambda\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-(i/k)^{2d}\right)(2\log k + \log N_i)} + o(\Delta)\right)
\end{aligned}$$

现在我们重新使用之前的界：对所有$\delta k \leqslant i \leqslant (1-\delta)k$，有 $2\log k + \log N_i \leqslant \log\binom{n-k}{k}(1-(i/k)+o(1))$；对所有$(1-\delta)k \leqslant i \leqslant k$，有 $2\log k + \log N_i \leqslant \log\binom{n-k}{k}((1-(i/k))(1+2c_\varepsilon)+o(1))$，其中 $0 < c_\varepsilon < \min\{\frac{d-2}{4}, 1\}$ 是一个常数，得到

$$\begin{aligned}
&\leqslant \exp(o(\Delta))\sum_{i=\delta k}^{(1-\delta)k} \exp\left(-\frac{\Delta}{2} - \lambda + \lambda\left(\frac{i}{k}\right)^d + \sqrt{2\lambda\left(1-\left(\frac{i}{k}\right)^{2d}\right)\log\binom{n-k}{k}\left(1-\frac{i}{k}+o(1)\right)}\right) \\
&\quad + \exp(o(\Delta))\sum_{i=(1-\delta)k}^{k} \exp\left(-\frac{\Delta}{2} - \lambda + \lambda\left(\frac{i}{k}\right)^d\right. \\
&\quad \left. + \sqrt{2\lambda\left(1-\left(\frac{i}{k}\right)^{2d}\right)\log\binom{n-k}{k}\left(\left(1-\frac{i}{k}\right)(1+2c_\varepsilon)+o(1)\right)}\right)
\end{aligned}$$

现在代入 $\lambda = 2\log\binom{n-k}{k} - \Delta$。我们得到

$$\leqslant \exp(o(\Delta)) \sum_{i=\delta k}^{(1-\delta)k} \exp\left(\Delta\left(-\frac{1}{2} + \frac{1}{2}\sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(1-\frac{i}{k}+o(1)\right)}\right)\right)$$
$$\cdot \exp\left(\left(2\log\binom{n-k}{k} - \Delta\right)\left(-1+\left(\frac{i}{k}\right)^{d} + \sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(1-\frac{i}{k}+o(1)\right)}\right)\right)$$
$$+ \exp(o(\Delta)) \sum_{i=(1-\delta)k}^{k} \exp\left(\Delta\left(-\frac{1}{2} + \frac{1}{2}\sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(\left(1-\frac{i}{k}\right)(1+2c_\varepsilon)+o(1)\right)}\right)\right)$$
$$\cdot \exp\left(\left(2\log\binom{n-k}{k} - \Delta\right)\left(-1+\left(\frac{i}{k}\right)^{d} + \sqrt{\left(1-\left(\frac{i}{k}\right)^{2d}\right)\left(\left(1-\frac{i}{k}\right)(1+2c_\varepsilon)+o(1)\right)}\right)\right)$$

现在注意到，存在一个常数 $\gamma_\delta > 0$，使得 $f(x) = -1+x^d+\sqrt{(1-x^{2d})(1-x)} \leqslant -\gamma_\delta$ 对所有 $\delta \leqslant x \leqslant 1-\delta$ 都成立；并且对于充分小的 $\delta > 0$，存在 $\tau_\gamma > 0$ 使得 $g(1-y) = -1+(1-y)^d + \sqrt{(1-(1-y)^{2d})\,y(1+2c_\varepsilon)} \leqslant -dy + \sqrt{2d(1+2c_\varepsilon)}y^2 + O(y^2) \leqslant -\tau_\delta y$ 对所有 $0 \leqslant y \leqslant \delta$ 都成立。这里我们利用了 $c_\varepsilon \leqslant \frac{d-2}{4}$，因此 $\sqrt{2d(1+2c_\varepsilon)} < d$。于是，

$$\leqslant \exp(o(\Delta)) \sum_{i=\delta k}^{(1-\delta)k} \exp\left(-\gamma_\delta\left(2\log\binom{n-k}{k} - \Delta\right)\right)$$
$$+ \exp(o(\Delta)) \sum_{i=(1-\delta)k}^{k} \exp\left(-\tau_\delta\frac{k-i}{k}\left(2\log\binom{n-k}{k} - \Delta\right)\right)$$
$$= o(1)。$$

因此，以概率 $1-o(1)$，我们有

$$\sum_{i=\delta k}^{k} p_i \leqslant \frac{1}{S_0}\sum_{i=\delta k}^{k} S_i = o(1).$$

由于 $\delta > 0$ 是任意的，我们得出，对于 $\lambda = 2\log\binom{n-k}{k} - \Delta$，有

$$\mathrm{MMSE} = 1 - \mathbb{E}\left[\sum_{i=0}^{k} p_i \cdot \frac{i}{k}\right] \geqslant (1-\delta) - \mathbb{E}\left[\sum_{i=\delta k}^{k} p_i \cdot \frac{i}{k}\right] \geqslant 1-\delta-o(1),$$

因此，最小均方误差 $\to 1$。

□