arXiv:2603.09487v1 [math.PR] 10 Mar 2026

α-次指数随机算子的一致集中

TIANKUN DIAO、XUANANG HU、VLADIMIR V. ULYANOV 和 HANCHAO WAN G

ABSTRACT. 作用于结构化集合的随机矩阵，在高维几何、压缩感知和随机化算法中发挥着基础性作用。现有研究主要关注次高斯模型，在这类模型中，随机矩阵作用于集合时表现为近等距映射，并具有最优的尾部行为。然而，在实际应用中，我们经常需要处理具有重尾的分布；这类分布并非次高斯，但至少具有指数型尾部。

在本文中，我们研究随机矩阵 A，其行（或列）具有 α-次指数尾部分布，且 $\alpha \in (0, 2]$。因此，次高斯模型和次指数模型都包含在内，作为特殊情形。我们针对 Ax建立了集中型不等式，其中 x 属于 $\mathbb{R}^n$ 的有界子集，结果表明，其几何畸变由集合的 Talagrand 泛函所控制，并取决于尾参数 α。我们的结果推广了次高斯情形下已知的最优不等式（$\alpha = 2$），并为尾部更重但仍指数可积的随机矩阵提供了新的保证。

这些发现将随机矩阵理论拓展到了次高斯框架之外。此外，它们得出了适用于降维的近等距嵌入结果，并使我们能够在非高斯测量下进行稳健的高维推断。

1. INTRODUCTION

随机矩阵凭借其保持几何结构的卓越能力，在高维概率、压缩感知和随机数值线性代数中发挥着核心作用。该领域的一个基本问题是，理解随机线性映射 $A \in \mathbb{R}^{m\times n}$ 何时能在给定子集 $T \subset \mathbb{R}^n$上近似保持等距，也就是说， T 中的向量的欧几里得范数经过映射 $x \mapsto Ax$后近似不变。这种性质是 Johnson–Lindenstrauss引理和受限等距性质等经典结果的基础。

当随机矩阵 A 具有高斯元素时， A的近似等距性质如今已得到充分理解，且已知其由测度的几何复杂度所决定，例如高斯宽度或高斯复杂度；例如，参见 [8, 10]。将这些结果推广到高斯情形之外，因理论考量以及应用中高斯随机性可能不切实际或不受欢迎而受到广泛关注。这方面的早期进展聚焦于次高斯随机矩阵，并有效运用了集中不等式和链式方法；例如，参见 [4, 5]。也有研究聚焦于非次高斯情形。例如，Adamczak等人

2020年 数学主题分类。 60B11, 41A46, 46B20.关键词与短语。 α-次指数随机矩阵；随机测量算子；一般链；Talagrand γ_α 泛函。

[1] 获得了独立对数凹随机向量之和的投影范数的尾部估计，以及这些估计的统一版本，即在对数凹系综情形下，对矩阵及其子矩阵的算子范数进行尾部估计。

此后，出现了两种影响尤为深远、彼此互补的非高斯框架。第一种研究具有独立、各向同性、次高斯行的矩阵。在这种按行模型中，人们建立了随机矩阵作用于集合时的精确偏差界，其对次高斯范数的依赖达到最优，尾部行为也十分紧确；见 [3]。第二种框架关注具有独立次高斯列且列的欧几里得范数固定的矩阵。这种列式模型由 [6], 提出，刻画了稀疏和结构化随机矩阵中出现的关键现象，并揭示出与行独立或元素独立系综截然不同的行为，包括对分布参数依赖性的改善。

所有这些框架都严重依赖次高斯假设，该假设要求数据具有轻尾和强集中性质。然而，在许多应用中，随机数据呈现出偏离次高斯情形的重尾行为，但仍具有次指数尾。稳健统计、脉冲噪声环境下的信号处理，以及基于非高斯草图的随机化算法中，都自然会出现这类例子。这引出了以下问题：

> 随机矩阵作用于集合时所具有的近似等距性质，在将次高斯假设放宽为具有指数尾部的分布后，能在多大程度上得以保留？

现在我们给出主要结果。

1.1. 按行 α次指数模型。 我们的第一个主要结果推广了 Jeong 等人 [3] 引入的按行框架，适用于 α 次指数随机矩阵。

定理 1.1 (按行模型)。 设 $B \in \mathbb{R}^{l\times m}$ 为固定矩阵， $A \in \mathbb{R}^{m\times n}$ 为零均值随机矩阵。假设 $\alpha \in (0, 2]$。设 A 的各行相互独立且各向同性，并且其 ψ_α 范数（当 $\alpha < 1$ 时为拟范数）一致地有界于 K 。设 $T \subset \mathbb{R}^n$ 为有界集。那么

$$\mathbb{E}\sup_{x\in T} \left| \|BAx\|_2 - \|B\|_{HS}\|x\|_2 \right| \le C(\alpha)K^{4/\alpha}\|B\|_{op} \left(\gamma_\alpha(T) + rad(T)\right).$$

此外，对任意 $u \ge 0$，以至少 $1 - C\exp(-u^\alpha)$ 的概率，

$$\sup_{x\in T} \left| \|BAx\|_2 - \|B\|_{HS}\|x\|_2 \right| \le C(\alpha)K^{4/\alpha}\|B\|_{op} \left(\gamma_\alpha(T) + u \cdot rad(T)\right),$$

其中， $C(\alpha)$ 是仅依赖于 α的绝对常数； $\|B\|_{HS}$ 和 $\|B\|_{op}$ 分别表示 B 的 *Frobenius* 范数和算子范数； $\gamma_\alpha(T)$ 是第 *2* 节中定义的 *Talagrand* 泛函，而 $rad(T)$ 由$\sup_{t\in T} \|t\|_2$给出； ψ_α 范数（当 $\alpha < 1$时为拟范数）定义为

$$\|\xi\|_{\psi_\alpha} := \inf\left\{ t > 0 : \mathbb{E}\exp\left(\left|\frac{\xi}{t}\right|^\alpha \right) \le 2 \right\}.$$

该定理将适用于集合上次高斯矩阵的现有偏差不等式推广到了更一般的 α-次指数分布类，同时保留了对 T几何结构的正确依赖关系。

当 B 是单位矩阵时，我们有如下推论。

推论 1.1。 设 $A \in \mathbb{R}^{m\times n}$ 为零均值随机矩阵。假设 $\alpha \in (0,2]$。设 A 的各行相互独立且各向同性，并且其 ψ_α-范数（当 $\alpha < 1$ 时为拟范数）一致地有界于 K 。设 $T \subset \mathbb{R}^n$ 为有界集。则

$$E\sup_{x\in T}\left|\|Ax\|_2 - \sqrt{m}\|x\|_2\right| \le C(\alpha)K^{4/\alpha}\left(\gamma_\alpha(T) + rad(T)\right).$$

此外，对于任意 $u \ge 0$，以至少 $1 - Ce^{-u^\alpha}$ 的概率，

$$\sup_{x\in T}\left|\|Ax\|_2 - \sqrt{m}\|x\|_2\right| \le C(\alpha)K^{4/\alpha}\left(\gamma_\alpha(T) + u\cdot rad(T)\right).$$

1.2.**列式 α-次指数模型**。我们的第二个主要结果将 Plan 和 Vershynin 引入的列式框架推广 [6]至 α-次指数随机矩阵。

定理 1.2 (列式模型**)**。 设 $A \in \mathbb{R}^{m\times n}$ 为随机矩阵，其列 A_i 是 $\mathbb{R}^m$ 中相互独立的零均值随机向量。假设 $\|A_i\|_2 = 1$ 几乎必然成立，且对于某个 $\alpha \in (0,2]$，各列的 ψ_α-范数（当 $\alpha < 1$ 时为拟范数）有界于 K 。那么，

$$E\sup_{x\in T}\left|\|Ax\|_2 - \|x\|_2\right| \le C(\alpha)K\left(\gamma_\alpha(T) + rad(T)\right).$$

此外，对于任意 $u \ge 0$，以至少 $1 - Ce^{-u^\alpha}$ 的概率，有

$$\sup_{x\in T}\left|\|Ax\|_2 - \|x\|_2\right| \le C(\alpha)K\left(\gamma_\alpha(T) + u\cdot rad(T)\right).$$

定理 1.2 扩展了文献 [6] 中证明的定理 **1.3**。此外，我们的结果基于一种截然不同的方法。

立即可得

推论 1.2。 设 $A \in \mathbb{R}^{m\times n}$ 是一个随机矩阵，其列 A_i 是 $\mathbb{R}^m$ 中相互独立的零均值随机向量。假设 $\|A_i\|_2 = \lambda > 0$ 几乎必然成立，且这些列的 ψ_α 范数（当 $\alpha < 1$ 时为拟范数）有界于 K，其中某个 $\alpha \in (0,2]$。那么

$$E\sup_{x\in T}\left|\|Ax\|_2 - \lambda\|x\|_2\right| \le C(\alpha)K\left(\gamma_\alpha(T) + rad(T)\right).$$

此外，对于任意 $u \ge 0$，以至少 $1 - Ce^{-u^\alpha}$ 的概率，有

$$\sup_{x\in T}\left|\|Ax\|_2 - \lambda\|x\|_2\right| \le C(\alpha)K\left(\gamma_\alpha(T) + u\cdot rad(T)\right).$$

注 1.1 (关于列归一化的必要性**)**。 与推论 *1.1*不同，本结果要求 A 满足强归一化条件

$$\|A_i\|_2 = \lambda \quad a.s.\ for\ all\ i.$$

即使在一维情形 $n = 1$ 下，也不能去掉这一假设，也不能仅以各向同性取而代之。例如，令 $b \sim$ 伯努利$(1/2)$ 和 X 在 $\sqrt{m}\,\mathbb{S}^{m-1}$上服从均匀分布，且与 b独立。那么 bX 是各向同性的次高斯随机变量。此外，如果 A 的列是 bX 和 $T = \{e_1\}$的独立同分布副本，则有

$$\sup_{x\in T}\big|\|Ax\|_2 - \lambda\|x\|_2\big| = |b\sqrt{m} - \lambda| \geq \tfrac{\sqrt{m}}{2}$$

以概率 $1/2$，对于任意选择的 λ。因此，若不进行列归一化，就无法证明定理 *1.2* 中的两个不等式，因为它们要求对 $|\|Ax\|_2 - \lambda\|x\|_2|$进行与维度无关的控制 m。

1.3. 方法论评述。 我们认为，本文的主要贡献在于证明定理的方法，这种方法不同于 [6] 中采用的思路。Plan 和 Vershynin 的原始论证高度依赖次高斯随机变量的精细性质，包括精确的尾部界和矩增长，而这些性质并不能自然推广到重尾分布。因此，他们的方法无法直接推广到次高斯情形之外。

相比之下，我们的方法避免使用次高斯分布特有的工具，而是基于一种更直接的分解方法，并结合初等的集中论证。这一思路统一适用于所有 $\alpha > 0$。同时，即使在次高斯情形下，它也能带来更直接、更清晰的证明。

本文结构如下。第 2 节介绍记号和预备知识。第 3 节和第 4 节分别针对按行和按列的 α-次指数模型，推导主要的偏差不等式。第 5 节讨论其应用，包括 Johnson–Lindenstrauss 嵌入、 α-次指数随机矩阵的受限等距性质，以及列归一化结构化矩阵。最后，我们将讨论未来工作的方向。

2. Preliminaries

下文中， c, C （带或不带下标）表示常数。我们不要求这些常数每次出现时都相同。若某个常数依赖于参数 α，则记作 $c = c(\alpha)$。

全文中，我们使用以下标准记号。对于向量 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ 和 $1 \leq p < \infty$，我们用以下符号表示

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

通常的 ℓ_p-范数。对于实值随机变量 X and $1 \leq p < \infty$，其 L^p-范数定义为

$$\|X\|_{L^p} := \big(\mathsf{E}|X|^p\big)^{1/p}.$$

对于矩阵 $A=(a_{ij})$，我们记作

$$\|A\|_{\mathrm{HS}}:=\left(\sum_{i,j}a_{ij}^2\right)^{1/2}\quad\text{and}\quad\|A\|_{\mathrm{op}}:=\sup\{\|Ax\|_2:\|x\|_2=1\}$$

分别对应希尔伯特-施密特范数和算子范数。

2.1. α-次指数随机变量与各向同性向量。

定义 2.1 (α-次指数随机变量)。 设 $\alpha>0$。若存在绝对常数 $c>0$，使得对所有$t\geq 0$, 则称随机变量 ξ 为 α-次指数的,

$$P\big(|\xi-E\xi|\geq t\big)\leq 2\exp\left(-\frac{t^\alpha}{c}\right).$$

相应的 ψ_α 范数（当 $\alpha<1$ 时为拟范数）定义为

$$\|\xi\|_{\psi_\alpha}:=\inf\left\{t>0:E\exp\left(\left|\frac{\xi}{t}\right|^\alpha\right)\leq 2\right\}.$$

下面我们汇总 α-次指数随机变量的若干基本性质。这些结果都是标准且广为人知的, 因此我们省略证明。详细论述请参阅 [3]中的命题 A.1。

命题 2.1。 设 η、ξ 为随机变量，且 $\alpha>0$。
(a) 若 $|\eta|\leq|\xi|$，则 $\|\eta\|_{\psi_\alpha}\leq\|\xi\|_{\psi_\alpha}$。*(b)* 若 $\|\eta\|_{\psi_\alpha}\leq K<\infty$，则对所有 $t\geq 0$，$P(|\eta|\geq t)\leq 2exp(-t^\alpha/K^\alpha)$。*(c)* 若对所有 $t\geq 0$ 及某个 $K>0$，$P(|\eta|\geq t)\leq 2exp(-t^\alpha/K^\alpha)$，则 $\|\eta\|_{\psi_\alpha}\leq C(\alpha)K$。*(d)* 对所有 $p\geq 1$，$\|\eta^p\|_{\psi_\alpha}=\|\eta\|_{\psi_{p\alpha}}^p$。特别地，$\|\eta^2\|_{\psi_1}=\|\eta\|_{\psi_2}^2$。*(e)* 对于 p，$q\in(1,\infty)$，满足 $\frac{1}{p}+\frac{1}{q}=1$。特别地，$\|\eta\xi\|_{\psi_1}\leq\|\eta\|_{\psi_2}\|\xi\|_{\psi_2}$。*(f)* 对所有 $p\geq 1$，$E|\eta|^p\leq\left(Cp^{\frac{1}{\alpha}}\|\eta\|_{\psi_\alpha}\right)^p$，且存在某个绝对常数 $C\leq 4$。*(g)* $\|\eta-E\eta\|_{\psi_\alpha}\leq C(\alpha)\|\eta\|_{\psi_\alpha}$。*(h)* 对所有 $\beta\geq\alpha$，$\|\eta\|_{\psi_\alpha}\leq C\|\eta\|_{\psi_\beta}$，且存在某个绝对常数 C。*(i)* $\|\eta+\xi\|_{\psi_\alpha}\leq C(\alpha)(\|\eta\|_{\psi_\alpha}+\|\xi\|_{\psi_\alpha})$。

给定随机变量 ξ_1、.....、ξ_n，我们记

$$\xi:=(\xi_1,\ldots,\xi_n)$$

$\mathbb{R}^n$中对应的随机向量。ξ 的相应 ψ_α 范数（当$\alpha<1$时为拟范数）定义为

$$\|\xi\|_{\psi_\alpha}:=\sup_{x\in S^{n-1}}\|\langle\xi,x\rangle\|_{\psi_\alpha}.$$

我们称 $\mathbb{R}^n$ 中的随机向量 X 为各向同性，如果

$$\mathsf{E}\,XX^\top=I_n,$$

其中 I_n 表示 $\mathbb{R}^n$中的单位矩阵。

2.2. **一般链与Talagrand的 γ-泛函**。我们首先回顾可容许序列的概念，它在Talagrand的 γ-泛函的定义中起着核心作用。

定义 2.2 (可容许序列**)**。 设 (T, d) 是一个度量空间。 T 的可容许划分序列是一个递增序列 $(\mathcal{A}_n)_{n\geq 0}$ ， 由 T 的划分构成，并满足

$$|\mathcal{A}_0| = 1 \qquad and \qquad |\mathcal{A}_k| \leq 2^{2^k}, \quad k \geq 1.$$

利用可容许序列，我们定义 Talagrand 的 γ_2-泛函。对于度量空间 (T, d)，令

$$\gamma_2(T, d) := \inf \sup_{t\in T} \sum_{n\geq 0} 2^{n/2}\, \Delta_d(A_n(t)),$$

其中，下确界取遍所有可容许序列 $(\mathcal{A}_k)_{k\geq 0}$。这里， $A_n(t)$ 表示满足 $t \in A$ 且满足下述条件的唯一集合 $A \in \mathcal{A}_n$

$$\Delta_d(A) := \sup_{s,t\in A} d(t, s).$$

更一般地，对于 $\alpha > 0$，(T, d) 的 γ_α-泛函定义为

$$\gamma_\alpha(T, d) := \inf \sup_{t\in T} \sum_{n\geq 0} 2^{n/\alpha}\, \Delta_d(A_n(t)),$$

其中，下确界同样取遍所有可容许序列。当 d 与欧几里得度量 d_2 重合时，为简便起见，我们记为 $\gamma_\alpha(T)$。

我们采用链方法来描述具有 α-次指数增量的随机过程的集中行为。为此，我们首先回顾 Talagrand 提出的一个基本结果。

定理 2.1 (参见 Talagrand [9, 定理 2.7。14])。 设 $(S_t)_{t\in T}$ 是一个中心化随机过程，不一定对称。对于 $n \geq 1$，定义增量度量

$$\delta_n(s, t) := \|S_s - S_t\|_{2^n}.$$

对于子集 $A \subset T$，令 $\Delta_n(A)$ 表示其关于 δ_n 的直径。令 $(\mathcal{A}_n)_{n\geq 0}$ 为 T 的一个可容许划分序列。则存在一个万有常数 $C > 0$ ，使得

$$E \sup_{s,t\in T} |S_s - S_t| \leq C \sup_{t\in T} \sum_{n\geq 0} \Delta_n(A_n(t)). \tag{2.1}$$

此外，对于任意 $u > 0$，令 k 为满足 $2^k \leq u^2$ 的最大整数。则

$$P\left(\sup_{s,t\in T} |S_s - S_t| \geq c\Delta_k(T) + \sup_{t\in T} \sum_{n\geq 0} \Delta_n(A_n(t)) \right) \leq C \exp(-u^2), \tag{2.2}$$

其中 c、$C > 0$ 为万有常数。

因此，上述定理直接给出了具有 α-次指数增量的随机过程的紧确界。

定理 2.2. 设$(Z_x)_{x\in T}$是一个以有界集$T\subset\mathbb{R}^n$为指标集的随机过程。假设该过程具有一致的α- 次指数增量，即存在 $M\geq 0$ 使得

$$\|Z_x - Z_y\|_{\psi_\alpha} \leq M\|x-y\|_2 \quad for\ all\ x,y\in T.$$

于是我们有

$$E\sup_{x,y\in T}|Z_x - Z_y| \leq C(\alpha)M\gamma_\alpha(T) \tag{2.3}$$

此外，给定 $u>0$，我们有

$$P\left(\sup_{s,t\in T}|Z_s - Z_t| \geq cM(u\cdot rad(T)+\gamma_\alpha(T))\right) \leq C\exp(-u^\alpha). \tag{2.4}$$

证明。步骤 *1*：比较 δ_n 与欧几里得距离。 根据命题2.1 (f)，对于每个具有有限 η-范数的随机变量 ψ_α以及每个 $p\geq 1$，我们有

$$\left(\mathsf{E}|\eta|^p\right)^{1/p} \leq C\ p^{1/\alpha}\ \|\eta\|_{\psi_\alpha}.$$

将其应用于 $\eta = Z_x - Z_y$ 和 $p=2^n$，我们有

$$\|Z_x - Z_y\|_{2^n} = \left(\mathsf{E}|Z_x - Z_y|^{2^n}\right)^{1/2^n} \leq C\,2^{n/\alpha}\,\|Z_x - Z_y\|_{\psi_\alpha} \leq C\,M\,2^{n/\alpha}\,\|x-y\|_2$$

对于所有 x，$y\in T$ 以及所有 $n\geq 1$。

因此，对于每个子集 $A\subset T$，

$$\Delta_n(A) = \sup_{x,y\in A}\delta_n(x,y) \leq CM\,2^{n/\alpha}\,rad(A), \qquad n\geq 1. \tag{2.5}$$

步骤 *2*：应用一般链式界。 设 $(\mathcal{A}_n)_{n\geq 0}$ 是 T 按照定义 2.2所述的任意可容许划分序列。对于每个 $t\in T$，记 $A_n(t$为 A_n 中包含 t的唯一元素。将定理 2.1 应用于中心化过程 $S_t:=Z_t$以及上文定义的距离 δ_n，得到以下界

$$\mathsf{E}\sup_{x,y\in T}|Z_x - Z_y| \leq C\sup_{t\in T}\sum_{n\geq 0}\Delta_n\big(A_n(t)\big). \tag{2.6}$$

利用（2.5），进一步得到

$$\sup_{t\in T}\sum_{n\geq 0}\Delta_n\big(A_n(t)\big) \leq CM\sup_{t\in T}\sum_{n\geq 0}2^{n/\alpha}\,rad\big(A_n(t)\big). \tag{2.7}$$

对 i 取　　下确界，并遍历所有可容许序列 $(A_n)_{n\geq 0}$，再回顾定义

$$\gamma_\alpha(T) := \inf\sup_{t\in T}\sum_{n\geq 0}2^{n/\alpha}\,rad\big(A_n(t)\big),$$

我们得出

$$\mathsf{E}\sup_{x,y\in T}|Z_x - Z_y| \leq C(\alpha)\,M\,\gamma_\alpha(T).$$

此外，注意到如果 $2^k\leq u^2$，那么

$$\Delta_k(T) \leq CM2^{k/\alpha}\,\mathrm{rad}(T) \leq CMu^{2/\alpha}\,\mathrm{rad}(T).$$

将此估计与（2.2）和（2.7）结合起来，我们得到，对于任意 $u>0$以及任意可容许序列 $(\mathcal{A}_n)_{n\geq 0}$ 在 T上，

$$\mathsf{P}\left(\sup_{s,t\in T}|Z_s-Z_t|\geq cM\left(u^{2/\alpha}\operatorname{rad}(T)+\sup_{t\in T}\sum_{n\geq 0}2^{n/\alpha}\operatorname{rad}\left(A_n(t)\right)\right)\right)\leq C\exp(-u^2).$$

通过对 u 进行适当的重新参数化，并对所有可容许序列取下确界，我们得到（2.4）。

□

3. Proof of Theorem 1.1

在本节中，我们给出定理 1.1的证明。

我们需要以下来自 Sambale [7] 的结果。

定理 3.1. *[7, 定理 1.1]* 对任意 $\alpha\in(0,2],$，设 $X_1,\ldots\ldots,X_n$ 为独立的中心化随机变量，使得对任意 i 有 $\|X_i\|_{\psi_\alpha}\leq K$，且 $A=(a_{ij})$ 为对称矩阵。则对任意 $t\geq 0$,

$$P\left(\left|X^TAX-EX^TAX\right|\geq t\right)\leq 2\exp\left(-\frac{1}{C(\alpha)}\min\left(\frac{t^2}{K^4\|A\|_{HS}^2},\left(\frac{t}{K^2\|A\|_{op}}\right)^{\frac{\alpha}{2}}\right)\right).$$

定理 3.2. *[7, 命题 2.2],* 设 B 为固定的 $m\times n$ 矩阵，并设 $X=(X_1,\ldots\ldots,X_n)\in\mathbb{R}^n$ 为具有独立坐标的中心化随机向量，满足对某个 $\alpha\in(0,2]$ 有 $EX_i^2=1$ 与 $\|X_i\|_{\psi_\alpha}\leq K$。则对任意 $t\geq 0$，我们有

$$P(|\|BX\|_2-\|B\|_{HS}|\geq tK^2\|B\|_{op})\leq 2\exp(-t^\alpha/C(\alpha)). \tag{3.1}$$

and

$$\|\|BX\|_2-\|B\|_{HS}\|_{\psi_\alpha}\leq C(\alpha)K^2\|B\|_{op}.$$

引理 3.1. 设 $B\in\mathbb{R}^{l\times m}$ 为固定矩阵，$A\in\mathbb{R}^{m\times n}$ 为均值为零、各向同性且α-次指数的矩阵，其 α-次指数参数为 K。则随机过程

$$Z_x:=\|BAx\|_2-\|B\|_{HS}\|x\|_2$$

具有 α-次指数增量，其中

$$\|Z_x-Z_y\|_{\psi_\alpha}\leq C(\alpha)K^{4/\alpha}\|B\|_{op}\|x-y\|_2,\ \forall x,y\in\mathbb{R}^n.$$

证明. 该结论对 B的缩放保持不变。因此，不失一般性，我们假设 B 的算子范数为 $\|B\|_{op}=1$, $K\geq 1$。

步骤 1： 证明 α-次指数增量对于 x，$y\in S^{n-1}$ 在单位球面上。不失一般性，假设 $x\neq y$ 并定义

$$p:=\mathsf{P}\left(\frac{|Z_x-Z_y|}{\|x-y\|_2}\geq s\right)=\mathsf{P}\left(\frac{\|BAx\|_2-\|BAy\|_2}{\|x-y\|_2}\geq s\right).$$

我们需要用 $\alpha-subexponentialtail$ 来界定该尾概率。考虑以下两种情形:

- $s \geq 2\|B\|_{HS}$。记作 $u := \frac{x-y}{\|x-y\|_2}$，由三角不等式可得

$$p \leq \mathsf{P}\left(\frac{\|BA(x-y)\|_2}{\|x-y\|_2} \geq s\right) = \mathsf{P}\left(\|BAu\|_2 \geq s\right) =: p_1.$$

- $0 < s < 2\|B\|_{HS}$。对于 $0 < s \leq 1$， $p \leq 1 \leq 2\exp(-cs^\alpha)$，其中 c_1 是一个足够小的常数，满足该不等式。

  对于 $1 < s \leq 2\|B\|_{HS}$，我们将 p 写作

$$p = \mathsf{P}\left(|Z| \geq s(\|BAx\|_2 + \|BAy\|_2)\right) \quad \text{where} \quad Z := \frac{\|BAx\|_2^2 - \|BAy\|_2^2}{\|x-y\|_2}.$$

于是可得

$$p \leq \mathsf{P}\left(|Z| \geq s\|BAx\|_2\right) \leq \mathsf{P}\left(\|BAx\|_2 \leq \frac{1}{2}\|B\|_{HS}\right) + \mathsf{P}\left(|Z| > \frac{s}{2}\|B\|_{HS}\right) =: p_2 + p_3,$$

其中， p_2 和 p_3 分别表示第一项和第二个求和项。

p_1 的界。

由 $s \geq 2\|B\|_{HS}$ 我们有

$$p_1 = \mathsf{P}\left(\|BAu\|_2 - \|B\|_{HS} \geq s - \|B\|_{HS}\right) \leq \mathsf{P}\left(\|BAu\|_2 - \|B\|_{HS} \geq \frac{s}{2}\right).$$

应用定理3.2 于随机向量 Au ，我们得到

$$p_1 \leq 2\exp\left(-C(\alpha)(\frac{s}{K^2})^\alpha\right).$$

p_2的界。

应用定理3.2 于随机向量 Ax ，并注意到 $\|B\|_{HS} > \frac{1}{2}s$，我们得到

$$p_2 \leq 2\exp\left(-C(\alpha)\frac{(\|B\|_{HS}/2)^\alpha}{K^{2\alpha}}\right) \leq 2\exp\left(-C(\alpha)(\frac{s}{4K^2})^\alpha\right).$$

p_3的界。

记 $u := \frac{x-y}{\|x-y\|_2}$以及 $v := x+y$，则 $\langle u, \ v\rangle = 0$ 因为 $\|x\|_2 = \|y\|_2 = 1$。我们可以将Z 写为

$$Z = \frac{\langle BA(x-y), BA(x+y)\rangle}{\|x-y\|_2} = \langle BAu, BAv\rangle.$$

注意到

$$2\langle BAu, BAv\rangle = \langle BA(u+v), BA(u+v)\rangle - \langle BAu, BAu\rangle - \langle BAv, BAv\rangle.$$

我们也将 $X_w := Aw$ 记为 $w \in \mathbb{R}^n$，那么由 $\mathsf{E}X_w X_w^T = \|w\|_2^2 I_n$ 可得

$$\mathsf{E}\|BX_w\|_2^2 = \mathsf{E}\mathrm{tr}(B^T B X_w X_w^T) = \mathrm{tr}(B^T B \mathsf{E}(X_w X_w^T)) = \|w\|_2^2\|B\|_{HS}^2.$$

因此我们可以进一步将 Z 写成

$$\begin{aligned} Z &= \frac{1}{2}\|BX_{u+v}\|_2^2 - \frac{1}{2}\|BX_u\|_2^2 - \frac{1}{2}\|BX_v\|_2^2 \\ &= \frac{1}{2}\left(\|BX_{u+v}\|_2^2 - \mathsf{E}\|BX_{u+v}\|_2^2\right) - \frac{1}{2}\left(\|BX_u\|_2^2 - \mathsf{E}\|BX_u\|_2^2\right) \\ &\quad - \frac{1}{2}\left(\|BX_v\|_2^2 - \mathsf{E}\|BX_v\|_2^2\right) \\ &= \frac{1}{2}Y_{u+v} - \frac{1}{2}Y_u - \frac{1}{2}Y_v, \end{aligned}$$

其中第二个等号利用了 Z 均值为零这一事实，而在最后一个等号中，$Y_w :$ $= \|BX_w\|_2^2 - \mathsf{E}\|BX_w\|_2^2$. 因此

$$\begin{aligned} p_3 &= \mathsf{P}(|Y_{u+v} - Y_u - Y_v| > s\|B\|_{HS}) \\ &\le \mathsf{P}(|Y_{u+v}| + |Y_u| + |Y_v| > s\|B\|_{HS}) \\ &\le \mathsf{P}\left(|Y_{u+v}| \ge \frac{s}{2}\|B\|_{HS}\right) + \mathsf{P}\left(|Y_u| + |Y_v| > \frac{s}{2}\|B\|_{HS}\right) \\ &\le \mathsf{P}\left(|Y_{u+v}| \ge \frac{s}{2}\|B\|_{HS}\right) + \mathsf{P}\left(|Y_u| \ge \left(1 - \frac{1}{8}\|v\|_2^2\right)\frac{s}{2}\|B\|_{HS}\right) \\ &\quad + \mathsf{P}\left(|Y_v| > \frac{1}{8}\|v\|_2^2 \cdot \frac{s}{2}\|B\|_{HS}\right) \\ &=: p_4 + p_5 + p_6. \end{aligned}$$

我们将给出 p_4、p_5 和 p_6 的界，所用工具为新的Hanson-Wright不等式（3.1）。

对于任意非零向量 w，定义 $\bar{w} :$ $= \frac{w}{\|w\|_2^\circ}$容易看出 $X_w = \|w\|_2 X_{\bar{w}}$与$Y_w = \|w\|_2^2 Y_{\bar{w}}$。此外还要注意

$$\|B^TB\|_{op} = \|B\|_{op}^2 = 1, \quad \|B^TB\|_{HS} \le \|B^T\|\|B\|_{HS} = \|B\|_{op}\|B\|_{HS} = \|B\|_{HS},$$

so by Theorem 1.5 we have

$$\mathsf{P}(|Y_{\bar{w}}| \ge r) \le 2\exp\left(-C(\alpha)\min\left(\frac{t^2}{K^4\|B\|_{HS}^2}, \left(\frac{t}{K^2}\right)^{\frac{\alpha}{2}}\right)\right).$$

因此对于 $0 \le t \le \|w\|_2^2\|B\|_{HS}^2$,

$$\mathsf{P}(|Y_w| \ge r) = \mathsf{P}\left(|Y_{\bar{w}}| \ge \frac{r}{\|w\|_2^2}\right) \le 2\exp\left(-C(\alpha)\left(\frac{t^2}{K^4\|B\|_{HS}^2\|w\|_2^2}\right)\right) + 2\exp\left(-C(\alpha)\left(\frac{t}{K^2\|w\|_2^2}\right)^{\frac{\alpha}{2}}\right).$$

现在我们将式（16）应用于 p_4、p_5 和 p_6。

- 对于 p_4。由于 $s < 2\|B\|_{HS}$和$\|u+v\|_2 = \sqrt{1 + \|v\|_2^2} \in [1, \sqrt{5})$，我们可以得出

$$\frac{s}{2}\|B\|_{HS} < \|B\|_{HS}^2 \le \|u+v\|_2^2\|B\|_{HS}^2,$$

因此

$$p_4 \le 2\exp\left(-C(\alpha)\frac{s^2}{4\|u+v\|_2^4 K^4}\right) + 2\exp\left(-C(\alpha)\left(\frac{s\|B\|_{HS}}{2K^2\|u+v\|_2^2}\right)^{\frac{\alpha}{2}}\right).$$

由于 $\|B\|_{HS} \ge \|B^TB\|_{HS} \ge \|B^TB\|_{op} = \|B\|_{op}^2 = 1$，且 $\|u+v\|_2 \in [1,\sqrt{5}]$

$$p_4 \le 2\exp\left(-C(\alpha)\frac{s^2}{K^4}\right) + 2\exp\left(-C(\alpha)\frac{s^\alpha}{K^\alpha}\right).$$

- 对于 p_5。注意到 $\|u\|_2 = 1$ 且 $1 - \frac{1}{8}\|v\|_2^2 \in (\frac{1}{2}, 1]$，因此

$$p_5 \le 2\exp\left(-C(\alpha)\frac{s^2}{K^4}\right) + 2\exp\left(-C(\alpha)\frac{s^\alpha}{K^\alpha}\right).$$

- 对于 p_6。若 $v = 0$（即 $x = -y$），则 $p_6 = \mathbf{P}(0 > 0) = 0$。现假设 $v \ne 0$，由于$\|v\|_2 \le 2$，$\|B\|_{HS} \ge 1$

则类似于 p_4 我们有

$$p_6 \le 2\exp\left(-C(\alpha)\frac{s^2}{K^4}\right) + 2\exp\left(-C(\alpha)\frac{s^\alpha}{K^\alpha}\right).$$

综合前述不等式，我们得到 p 的如下界。至此，我们已经证明

$$p \le \max\{p_1, p_2 + p_3\}, \text{ and } p_3 \le p_4 + p_5 + p_6.$$

其中 $p_i \le 4\exp\left(\frac{-cs^\alpha}{K^4}\right)$ 对某个绝对常数 c 以及 $1 \le i \le 6$。注意到 $p \le 1$ 以及不等式 $\min\{1, 16e^{-x}\} \le 2e^{-x/4}$。因此我们得到

$$p \le \min\left\{1, 16\exp\left(\frac{-cs^\alpha}{K^4}\right)\right\} \le 2\exp\left(\frac{-cs^\alpha}{4K^4}\right).$$

步骤 2： 对所有 x 和 y展示 $\alpha - subexponential$ 增量

不失一般性，我们可以假设 $\|x\|_2 = 1$ 和 $\|y\|_2 \ge 1$。令 $\bar{y} := \frac{y}{\|y\|_2}$ 为 y 在单位球上的投影，则根据三角不等式，

$$\|Z_x - Z_y\|_{\psi_\alpha} \le C(\alpha)\left(\|Z_x - Z_{\bar{y}}\|_{\psi_\alpha} + \|Z_y - Z_{\bar{y}}\|_{\psi_\alpha}\right) =: R_1 + R_2.$$

这里 R_1 它被 $C(\alpha)K^{\frac{\alpha}{4}}\|x - \bar{y}\|_2$ 所界定，因为 $x,\ \bar{y} \in S^{n-1}$，并且

$$R_2 = \|(\|y\|_2 - 1)Z_{\bar{y}}\|_{\psi_\alpha} = \|y - \bar{y}\|_2\|Z_{\bar{y}}\|_{\psi_\alpha} \le C(\alpha)K^{\frac{\alpha}{4}}\|x - \bar{y}\|_2,$$

其中第一个等号用到了 $Z_y = \|y\|_2 Z_{\bar{y}}$。第二个等号成立，因为 $\|y\|_2 - 1 = \|y - \bar{y}\|_2$，最后一个不等式由定理 3.2得出。结合这些界，我们得到

$$\|Z_x - Z_y\|_{\psi_\alpha} \le C(\alpha)K^{\frac{\alpha}{4}}(\|x - \bar{y}\|_2 + \|y - \bar{y}\|_2).$$

最后，注意到 $\|x\|_2 = 1$，因此由投影的非扩张性可得 $\|x - \bar{y}\|_2 \le \|x - y\|_2$，再根据投影的定义可得 $\|y - \bar{y}\|_2 \le \|y - x\|_2$。证毕。□

定理的证明 1.1。 设 $Z_x := \|BAx\|_2 - \|B\|_{HS}\|x\|_2$.

对于期望界，任取一个 $y \in T$，由三角不等式可得

$$\mathsf{E}\sup_{t\in T}|Z_t| \le \mathsf{E}\sup_{t\in T}|Z_t - Z_y| + \mathsf{E}|Z_y|.$$

利用引理 3.1 和定理 2.2，可得

$$\mathsf{E}\sup_{t\in T}|Z_t - Z_y| \le \mathsf{E}\sup_{t,y\in T}|Z_t - Z_y| \lesssim_\alpha M\gamma_\alpha(T, d_2).$$

利用引理 3.1和命题 2.1中的性质 (f)，可得

$$\mathsf{E}|Z_y| \lesssim \|Z_y\|_{\psi_\alpha} = \|Z_y - Z_0\|_{\psi_\alpha} \lesssim K^{4/\alpha}\|B\|_{op}\|y\|_2.$$

因此 $\mathsf{E}\sup_{t\in T}|Z_t| \le C(\alpha)K^{4/\alpha}\|B\|_{op}(\gamma_\alpha(T, d_2) + \|y\|_2)$。

对于高概率界，注意到当 $u < 1$. 时结论显然。当 $u \ge 1$，任取一个 $y \in T$，再次使用三角不等式可得

$$\sup_{t\in T}|Z_t| \le \sup_{t\in T}|Z_t - Z_y| + |Z_y| \le \sup_{t,t'\in T}|Z_t - Z_{t'}| + |Z_y|.$$

Let

$$p_1 = \mathsf{P}\left(\sup_{t,t'\in T}|Z_t - Z_{t'}| \ge c(\alpha)K^{4/\alpha}(u\cdot rad(T) + \gamma_\alpha(T))\right),$$

$$p_2 = \mathsf{P}\left(|Z_y| \ge c(\alpha)K^{4/\alpha}u\cdot rad(T)\right).$$

由定理 2.2，可得估计

$$p_1 \le C\exp(-u^\alpha).$$

另一方面，命题2.1 中的性质 (b) 得出

$$p_2 \le 2\exp(-u^\alpha).$$

因此，我们有

$$\mathsf{P}\left(\sup_{x\in T}\big|\|BAx\|_2 - \|B\|_{HS}\|x\|_2\big| \ge c(\alpha)K^{4/\alpha}\|B\|_{op}\big(\gamma_\alpha(T) + u\ \mathrm{rad}(T)\big)\right) \le p_1 + p_2 \le C\exp(-u^\alpha).$$

□

4. Proof of Theorem 1.2

与定理1.1的证明类似，我们的目标是证明该过程 ss

$$\big\{\|Ax\|_2 - \|x\|_2\big\}_{x\in T}$$

具有一致的 α次指数增量。为此，我们先考虑一个更简单的特殊情形。

引理 4.1。 设 A是一个 $m\times n$矩阵，其列向量 A_i为独立的 α-次指数随机向量，且在 $\mathbb{R}^m$ 中几乎处处满足 $\|A_i\|_2 = 1$ ，并且对所有$x \in S^{m-1}$都有 $\|\langle A_i, x\rangle\|_{\psi_\alpha} \le K$ 。则有

$$\|\|Ax\|_2 - 1\|_{\psi_\alpha} \le C(\alpha)K.$$

证明。 不失一般性，我们可以假设 K 足够大。由于

$$\|Ax\|_2^2 = \sum_{i=1}^n x_i^2 \|A_i\|_2^2 + \sum_{i \neq j} x_i x_j \langle A_i, A_j \rangle,$$

我们有

$$\|Ax\|_2^2 - 1 = \sum_{i \neq j} x_i x_j \langle A_i, A_j \rangle.$$

对于任意 $p \geq 1$,

$$\begin{aligned}
\left(\mathsf{E}\left|\|Ax\|_2^2 - 1\right|^p\right)^{\frac{1}{p}} &= \left(\mathsf{E}\left|\sum_{i \neq j} x_i x_j \langle A_i, A_j \rangle\right|^p\right)^{\frac{1}{p}} \\
&\leq \left(\mathsf{E}\left|4\sum_{i \neq j} x_i x_j \langle A_i, A_j' \rangle\right|^p\right)^{\frac{1}{p}} \\
&= \left(\mathsf{E}\left|4\langle Ax, A'x \rangle\right|^p\right)^{\frac{1}{p}}. \qquad (4.1)
\end{aligned}$$

这里 A' 是 A的一个独立副本。这里的第二个不等式可参考 [10] 定理 6.1.1：解耦技术。

注意， $\langle A_i, A_j' \rangle$ 是条件 α-次指数型的，并且条件 α-次指数型随机变量的范数有界于 $CK\|AX\|_2$。
因此我们有

$$\begin{aligned}
\left(\mathsf{E}\left|4\langle Ax, A'x \rangle\right|^p\right)^{\frac{1}{p}} &\leq \left(\mathsf{E}\left(\|Ax\|_2 CKp^{\frac{1}{\alpha}}\right)^p\right)^{\frac{1}{p}} \\
&= CKp^{\frac{1}{\alpha}} \|\|Ax\|_2\|_p \\
&= CKp^{\frac{1}{\alpha}} \sqrt{\|\|Ax\|_2^2\|_{\frac{2}{p}}}.
\end{aligned}$$

我们可以很容易地得到

$$\begin{aligned}
\mathsf{E}\|Ax\|_2^2 &= \mathsf{E}\sum_{i=1}^n x_i^2 \|A_i\|_2^2 + \mathsf{E}\sum_{i \neq j} x_i x_j \langle A_i, A_j \rangle \\
&= \mathsf{E}\sum_{i=1}^n x_i^2 \|A_i\|_2^2 \\
&= \sum_{i=1}^n x_i^2 \\
&= 1.
\end{aligned}$$

对于 $p > 0$， 我们定义 $f(p) := \|\|Ax\|_2^2 - 1\|_p$， 则有
(i)$f(p) \leq CKp^{\frac{1}{\alpha}} \sqrt{f(\frac{p}{2}) + 1}$， (ii)$\leq f$
(2) $\leq CK2^{\frac{1}{\alpha}} \sqrt{\mathsf{E}\|Ax\|_2^2} = CK2^{\frac{1}{\alpha}}$.
设 $CK = c_0$。 由于 K 足够大， 不失一般性， 我们可以假设 $c_0 \geq 1$.

考虑以下序列:

$$\begin{cases} a_{n+1} = c_0 \cdot 2^{\frac{n+1}{\alpha}} \sqrt{a_n + 1}, \quad n \geq 2, \quad (1) \\ a_2 = c_0 \cdot 2^{\frac{2}{\alpha}}. \quad (2) \end{cases}$$

显然，对于所有 $n \geq 2, \quad a_n \geq 1$。令 $b_n = \log_2 a_n$。

由 (1) 得:

$$\begin{aligned} b_{n+1} &= \log_2 \left(c_0 2^{\frac{n+1}{\alpha}} \sqrt{2^{b_n} + 1} \right) \\ &= \log_2 \left(\frac{b_0}{2^2} \cdot 2^{\frac{n+1}{\alpha}} \cdot c_0 \sqrt{\frac{2^{b_{n+1}}}{2^{b_n}}} \right) \\ &\leq \frac{1}{2} b_n + \frac{n+1}{\alpha} + \log_2 (2c_0) . \end{aligned}$$

由 (2) 得:

$$b_2 \leq \frac{2}{\alpha} + \log_2(2c_0).$$

因此，我们有:

$$\begin{aligned} b_n &\leq \frac{1}{2} b_{n-1} + \frac{n}{\alpha} + \log_2(2c_0) \\ &\leq \frac{1}{4} b_{n-2} + \frac{n}{\alpha} + \log_2(2c_0) + \frac{1}{2} \left(\frac{n-1}{\alpha} + \log_2(2c_0) \right) \\ &\leq \dots \\ &\leq \left(\frac{1}{2} \right)^{n-2} b_2 + \sum_{j=0}^{n-3} \left(\frac{1}{2} \right)^j \left(\frac{n-j}{\alpha} + \log_2(2c_0) \right) \\ &\leq \sum_{j=0}^{n-2} \left(\frac{1}{2} \right)^j \left(\frac{n-j}{\alpha} + \log_2(2c_0) \right) \\ &\leq 2\frac{n}{\alpha} + 2\log_2(2c_0). \end{aligned}$$

因此，相应地,

$$a_n = 2^{b_n} \leq 4c_0^2 \cdot 2^{\frac{2n}{\alpha}}.$$

结合 (i)、(ii)、(1) 和 (2)，容易看出:

$$f(2^n) \leq 4c_0^2 \cdot 2^{\frac{2n}{\alpha}}.$$

也就是说,

$$\left\| \|Ax\|_2^2 - 1 \right\|_{2^n} \leq 4c_0^2 \cdot 2^{\frac{2n}{\alpha}}$$

对于所有 $n \geq 1$。

由于

$$\|Ax\|_2^2 - 1 = (\|Ax\|_2 - 1)(\|Ax\|_2 + 1),$$

and

$$|\|Ax\|_2 - 1| \leq \|Ax\|_2 + 1,$$

由此可得

$$|\|Ax\|_2 - 1|^2 \le |\|Ax\|_2^2 - 1|.$$

Thus

$$\left(\mathsf{E}\,|\|Ax\|_2 - 1|^{2^{n+1}}\right)^{\frac{1}{2^n}} \le 4c_0^2 \cdot 2^{\frac{2n}{\alpha}}$$

$$\Rightarrow \left\|\|Ax\|_2 - 1\right\|_{2^{n+1}} \le 2c_0 \cdot 2^{\frac{n}{\alpha}} \le 2c_0 \cdot 2^{\frac{n+1}{\alpha}}$$

对所有 $n \ge 1$ 均成立。因此，对所有 $p \ge 1$，我们定义 $n_p := \min\{n : 2^n \ge 4p\}$.

于是我们有

$$\left\|\|Ax\|_2 - 1\right\|_p \le \left\|\|Ax\|_2 - 1\right\|_{2^{n_p}} \le 2c_0 \cdot 2^{\frac{n_p}{\alpha}}$$

$$\le 2c_0 \cdot (8p)^{\frac{1}{\alpha}}.$$

也就是说,

$$\left\|\|Ax\|_2 - 1\right\|_p \le C(\alpha) K p^{\frac{1}{\alpha}}.$$

（回忆一下 $c_0 = CK$）

利用命题 2.1 中的性质 (f)，我们有

$$\|\|Ax\|_2 - 1\|_{\psi_\alpha} \le C(\alpha) K.$$

□

有些出人意料的是，仅分析这一特殊情形就足以证明一般结果：

定理 4.1。 假设 $Z_x = \|Ax\|_2 - \|x\|_2$，则 $\|Z_x - Z_y\|_{\psi_\alpha} \le C(\alpha) K \|x - y\|_2$ 对所有x、 $y \in \mathbb{R}^n$ 均成立。

证明。 由于齐次性，我们可以假设 $K \ge 1.$ 。我们有

$$\left\| \frac{Z_x - Z_y}{\|x - y\|_2} \right\|_{\psi_\alpha} = \left\| \frac{(\|Ax\|_2 - \|Ay\|_2) - (\|x\|_2 - \|y\|_2)}{\|x - y\|_2} \right\|_{\psi_\alpha}$$

$$\le C(\alpha) \left(\left\| \frac{\|Ax\|_2 - \|Ay\|_2}{\|x - y\|_2} \right\|_{\psi_\alpha} + \left\| \frac{\|x\|_2 - \|y\|_2}{\|x - y\|_2} \right\|_{\psi_\alpha} \right).$$

根据三角不等式，我们有：

$$|\|Ax\|_2 - \|Ay\|_2| \le \|A(x - y)\|_2,$$

$$|\|x\|_2 - \|y\|_2| \le \|x - y\|_2.$$

因此：

$$\left\| \frac{Z_x - Z_y}{\|x - y\|_2} \right\|_{\psi_\alpha} \le C(\alpha) \left(\left\| \frac{\|A(x - y)\|_2}{\|x - y\|_2} \right\|_{\psi_\alpha} + 1 \right)$$

$$\le C(\alpha) \left(\left\| \frac{\|A(x - y)\|_2 - \|x - y\|_2}{\|x - y\|_2} \right\|_{\psi_\alpha} + 2 \right) \le C(\alpha) K. \text{ (for } K \ge 1)$$

□

证明的其余部分与定理1.1 的证明思路相同，因此略去。

5. Applications

5.1. **Johnson-Lindenstrauss 引理**。我们的结果直接推得，所有满足定理1.1 或定理1.2 假设的矩阵都可作为用于降维的 Johnson–Lindenstrauss 嵌入。下面给出相应的 Johnson–Lindenstrauss 引理。

引理 5.1。设 $A \in \mathbb{R}^{m\times n}$ 满足定理*1.1*中的条件。对于任意 $\delta > 0$ 和$0 \le \varepsilon \le 1$，若

$$m \ge C(\alpha)K^{\frac{8}{\alpha}}\varepsilon^{-2}(\log\frac{1}{\delta})^{\frac{2}{\alpha}},$$

则对于任意 $x, y \in \mathbb{R}^n$，以至少 $1-\delta$ 的概率，有

$$(1-\varepsilon)\|x-y\|_2 \le \frac{1}{\sqrt{m}}\|A(x-y)\|_2 \le (1+\varepsilon)\|x-y\|_2.$$

证明。通过缩放，不妨设 $\|x-y\|_2 = 1$。由定理 2.2 以及 $B = I_m$，可得

$$\|\frac{1}{\sqrt{m}}\|A(x-y)\|_2 - \|x-y\|_2\|_{\psi_\alpha} \le C(\alpha)K^{4/\alpha}m^{-1/2}.$$

结合 t 命题中的性质 (b)2.1，我们可以立即得到该结果。 □

类似地，我们有如下引理。

引理 5.2. 设 $A \in \mathbb{R}^{m\times n}$ 在 *1.2* 的条件下满足推论中的条件，其中 $\lambda = \sqrt{m}$。对任意$\delta > 0$ 和 $0 \le \varepsilon \le 1$，若

$$m \ge C(\alpha)K^2\varepsilon^{-2}(\log\frac{1}{\delta})^{\frac{2}{\alpha}},$$

则对任意 x、$y \in \mathbb{R}^n$，以至少为 $1-\delta$ 的概率，我们有

$$(1-\varepsilon)\|x-y\|_2 \le \frac{1}{\sqrt{m}}\|A(x-y)\|_2 \le (1+\varepsilon)\|x-y\|_2.$$

5.2. **受限等距性质：** **α-次指数随机矩阵**。若向量 $x \in \mathbb{R}^n$ 至多有 s 个非零分量，则称其为 s-稀疏的，即

$$\|x\|_0 := \big|\{\, l : x_l \ne 0 \,\}\big| \le s. \tag{5.1}$$

在压缩感知中，人们试图根据以下形式的线性观测重建这样的向量

$$y = \Phi x, \tag{5.2}$$

其中 $\Phi \in \mathbb{R}^{m\times n}$ 是给定的感知矩阵。由于只有当 $m = n$ 时该系统才是完全确定的，因此这一问题主要在欠采样情形 $m \ll n$ 下受到关注。

一种广受研究的重构方法基于 ℓ_1-最小化，即：

$$\min_{z\in\mathbb{R}^n} \|z\|_1 \quad \text{subject to} \quad \Phi z = y. \tag{5.3}$$

这自然引出了一个问题：何时该凸优化问题的解与原始稀疏信号一致?

一个基本的充分条件是受限等距性质（R.I.P.）。如果一个 $m \times n$矩阵 A 满足阶数为 s 、常数为 $\delta \in$ （0，1）的受限等距性质，则称其满足该性质，即

$$(1-\delta)\|x\|_2^2 \le \|Ax\|_2^2 \le (1+\delta)\|x\|_2^2 \tag{5.4}$$

对所有 s-稀疏向量 $x \in \mathbb{R}^n$均成立。受限等距常数 δ_s 定义为使（5.4）成立的最小 δ 。

Dai 等人 [2] 研究了具有 α-次指数尾的随机矩阵的受限等距性质。相比之下，本文得到的结论集中于由我们的两个主要定理产生的特定实例。

命题 5.1。 设 $A \in \mathbb{R}^{m\times n}$ 满足定理*1.1*中的条件。对于任意 $\delta \in (0,\ 1)$，若

$$m \ge C(\alpha)\delta^{-2}K^{\frac{8}{\alpha}}\Big((s\log\frac{en}{s})^{\frac{1}{\alpha}}+u\Big)^2,$$

则以至少 $1-C\exp(-u^\alpha)$的概率，有

$$\delta_s \le \delta.$$

证明。 由齐次性，只需验证（5.4）对所有向量

$$x \in T := \Sigma_s^{2n} \cap \mathbb{S}^{n-1},$$

其中 Σ_s^{2n}表示 s 中的 $\mathbb{R}^{2n}$ 稀疏向量集合。结合推论 1.1，我们得到，其概率至少为 $1-Ce^{-u^\alpha}$，

$$\sup_{x\in T}\left|\frac{1}{\sqrt{m}}\|Ax\|_2-1\right| \le C(\alpha)\,m^{-1/2}K^{4/\alpha}\big(\gamma_\alpha(T)+u\ \mathrm{rad}(T)\big). \tag{5.5}$$

因此，接下来只需估计集合 T 的 γ_α 泛函。

$\gamma_\alpha(T)$ 的一个一般上界为

$$\gamma_\alpha(T) \lesssim \int_0^\infty \big(\log N(T,d,u)\big)^{1/\alpha}\,du,$$

例如，参见 Talagrand 第 3 章 [9] （其中讨论了 $\alpha=2$ 的情形）。另一方面，标准的体积论证给出

$$N\big(T,\|\cdot\|_2,u\big) \le \Big(\frac{e\,2n}{s}\Big)^s(1+2/u)^s.$$

结合这些估计，我们得到

$$\gamma_\alpha(T) \le C(\alpha)s^{1/\alpha}\left(\log^{1/\alpha}\Big(\frac{en}{s}\Big)+\int_0^1\log^{1/\alpha}\Big(1+\frac{2}{u}\Big)\,du\right) \le C(\alpha)\big(s\log(en/s)\big)^{1/\alpha}.$$

将此界代入（5.5）即可得到所需结论。

□

同理，我们有如下命题。

命题 5.2. 设 $A \in \mathbb{R}^{m\times n}$ 满足定理1.2中的条件。则对于任意$\delta \in (0,1)$，若

$$m \geq C(\alpha)\delta^{-2}K^2\Big((s\log\frac{en}{s})^{\frac{1}{\alpha}}+u\Big)^2,$$

那么，以至少 $1-C\exp(-u^\alpha)$ 的概率，有

$$\delta_s \leq \delta.$$

5.3. **对 α-次指数矩阵的列进行归一化。** 我们转而研究列相互独立、各向同性且具有 α-次指数性质的随机矩阵。

这些假设保证各列的欧几里得范数集中在 $\sqrt{m}$ 附近。受此启发，我们将每一列重新缩放，使其落在球面 $\sqrt{m}\mathbb{S}^{n-1}$ 上。然而，为避免范数异常小的列造成过度放大，有必要将注意范围限制在所有列范数均一致地具有正下界的事件上。下面定义的事件 $\mathcal{F}$ 将这一要求形式化。

推论 5.1。 设 $A \in \mathbb{R}^{m\times n}$。假设 $\alpha \in (0,2]$。设 A 的列相互独立、对称、各向同性，且其 ψ_α-范数（当 $\alpha<1$ 时为拟范数）有界于 K。假设

$$m \geq C(\alpha)K^4(\log n)^{2/\alpha}.$$

将 A_i 记为 A 的第 i 列。定义事件

$$\mathcal{F} := \Big\{\min_{1\leq i\leq n}\|A_i\|_2 \geq \tfrac{\sqrt{m}}{2}\Big\}.$$

Then

$$P(\mathcal{F}) \geq 1-2\exp\Big(-c(\alpha)\frac{m^{\alpha/2}}{K^{2\alpha}}\Big).$$

在事件 $\mathcal{F}$ 的条件下，将列归一化矩阵 A 定义为

$$\widetilde{A}_i := \frac{\sqrt{m}}{\|A_i\|_2}A_i, \qquad i=1,2,\dots,n.$$

Then

$$E\Big[\sup_{x\in T}\big|\|\widetilde{A}x\|_2-\sqrt{m}\|x\|_2\big| \;\Big|\; \mathcal{F}\Big] \leq C(\alpha)K(\gamma_\alpha(T)+rad(T)).$$

此外，仍在给定 $\mathcal{F}$ 的条件下，对于每个 $u>0$，以至少 $1-3\exp(-u^2)$ 的概率，

$$\sup_{x\in T}\big|\|\widetilde{A}x\|_2-\sqrt{m}\|x\|_2\big| \leq C(\alpha)K\big(\gamma_\alpha(T)+u\cdot\mathrm{rad}(T)\big).$$

证明。 我们首先引入一个高概率事件，在该事件上 A 的所有列的欧几里得范数都远离零。在此事件的条件下，我们验证归一化后的矩阵 A具有独立、零均值、次高斯列，且其 ψ_2-范数受到控制，然后应用推论 1.2。

步骤 *1*：列范数的统一下界。令 $A_1,\dots,A_n$ 表示 A的各列，并定义

$$F := \Big\{\min_{1\leq i\leq n}\|A_i\|_2 \geq \tfrac{\sqrt{m}}{2}\Big\}.$$

根据定理 3.2，对于每个 $i \in \{1, \dots, n\}$，都有

$$\big\| \|A_i\|_2 - \sqrt{m} \big\|_{\psi_\alpha} \;\le\; C(\alpha) K^2,$$

因此

$$\mathsf{P}\Big(\|A_i\|_2 \le \tfrac{\sqrt{m}}{2} \Big) = \mathsf{P}\Big(\|A_i\|_2 - \sqrt{m} \le -\tfrac{\sqrt{m}}{2} \Big) \le 2 \exp\Big(-c(\alpha) \, \frac{m^{\alpha/2}}{K^{2\alpha}} \Big).$$

应用并集界可得

$$\mathsf{P}(F) \;\ge\; 1 - 2n \exp\Big(-c(\alpha) \, \frac{m^{\alpha/2}}{K^{2\alpha}} \Big) \;=\; 1 - 2 \exp\Big(\log n - c(\alpha) \, \frac{m^{\alpha/2}}{K^{2\alpha}} \Big).$$

在假设 $m \ge C(\alpha) K^4 (\log n)^{2/\alpha}$ 成立的情况下，$\log n$ 项可吸收到常数中，因此

$$\mathsf{P}(F) \;\ge\; 1 - 2 \exp\Big(-c(\alpha) \, \frac{m^{\alpha/2}}{K^{2\alpha}} \Big) \;\ge\; \tfrac{1}{2}.$$

步骤 *2*：归一化列的条件性质。在事件 F 上，列 ~ 非零，因此我们可以定义列归一化矩阵 A，定义如下：

$$\widetilde{A}_i \;:=\; \frac{\sqrt{m}}{\|A_i\|_2} \, A_i, \qquad i = 1, \dots, n.$$

~显然，对所有 i，~都有 $\|A_i\|_2 = \sqrt{m}$。此外，在给定 F 的条件下可得

$$\|\widetilde{A}_i \mid F\|_{\psi_\alpha} = \Big\| \frac{\sqrt{m}}{\|A_i\|_\alpha} A_i \;\Big|\; F \Big\|_{\psi_\alpha} \le 2 \, \|A_i \mid F\|_{\psi_\alpha} \le C \, \|A_i\|_{\psi_\alpha}. \tag{5.6}$$

第一个不等式在 F 上使用了~$\sqrt{m}/\|A_i\|_2 \le 2$。第二个不等式源于这样一个事实：在概率至少为 $1/2$ 的事件上进行条件化，至多会使 α-次指数范数扩大一个绝对常数倍。因此，根据假设 $\|A_i\|_{\psi_\alpha} \le K$，我们有

$$\|\widetilde{A}_i \mid F\|_{\psi_\alpha} \;\le\; CK.$$

步骤 *3*：条件中心化与独立性。由于 A_i 的各个元素均为对称分布，其符号与绝对值相互独立；因此，在给定 F（它仅依赖于 $\|A_i\|_2$）后，A_i 的条件分布仍然对称。特别地，

$$\mathsf{E}[\widetilde{A}_i \mid F] = 0.$$

最后，F 可按列分解（它是若干分别依赖于各个~列的事件的交集），且 A 的各列相互独立。因此，在给定 F后，A 的各列仍相互独立。

~ ~结论。给定 F后，各列 $A_1, \dots, A_n$ 相互独立，且为零均值的次高斯向量，满足

$$\|\widetilde{A}_i\|_2 = \sqrt{m}, \qquad \|\widetilde{A}_i \mid F\|_{\psi_\alpha} \le CK, \qquad \mathsf{E}[\widetilde{A}_i \mid F] = 0.$$

因此，我们恰好处于推论 1.2 所述的情形 $\lambda = \sqrt{m}$，所需结论随即得证。□

致谢 Xuanang Hu（通讯作者）和 Hanchao Wang 的研究工作得到国家重点研发计划（编号：2024YFA1013501）、国家自然科学基金（编号：12571162）以及山东省自然科学基金（编号：ZR2024MA082）的资助。Vladimir V. Ulyanov 的研究工作是在HSE大学基础研究计划以及莫斯科国立罗蒙诺索夫大学莫斯科基础与应用数学中心计划框架内开展的。

作者贡献
这些作者对本研究的贡献相同。
数据可用性
本文所述研究未使用任何数据。

References

[1] Adamczak, R. 等. 对数凹随机向量之和范数的尾部估计. Proc. Lond. Math. Soc. (3) 108(3) (2014), 600–637. DOI: 10.1112/plms/pdt031.[2] Dai, G. 等. 关于对数凹尾混沌与受限等距性质. J. Funct. Anal. 289(11) (2025), Paper No. 111130, 41. DOI: 10.1016/j.jfa.2025.111130.[3] Jeong, H. 等. 集合上的次高斯矩阵：最优尾部依赖性及其应用. Comm. Pure Appl. Math. 75(8) (2022), 1713–1754. DOI: 10.1002/cpa.22024.[4] Liaw, C. 等. 一种简单工具，用于界定随机矩阵在几何集合上的偏差. 泛函分析的几何方面. 第2169卷. 数学讲义. 施普林格，Cham，2017, 277–299. ISBN: 978-3-319-45281-4; 978-3-319-45282-1.[5] Mendelson, S.、Pajor, A. 和 Tomczak-Jaegermann, N. 伯努利系综和次高斯系综的一致不确定性原理. Constr. Approx. 28(3) (2008), 277–289. DOI: 10.1007/s00365-007-9005-8.[6] Plan, Y. 和 Vershynin, R. 作用于集合的随机矩阵：独立列. Electron. J. Probab. 30 (2025), Paper No. 143, 19. DOI: 10.1214/25-ejp1404.[7] Sambale, H. 关于 α-次指数随机变量集中性的一些笔记. 高维概率 IX——空灵之卷. 第80卷. 概率进展. Birkh¨auser/施普林格

，

Cham，[2023]©2023, 167–192. ISBN: 978-3-031-26978-3; 978-3-031-26979-0. DOI: 10.1007/978-3-031-26979-0_7.[8] Schechtman, G. 关于将欧氏空间的子集嵌入赋范空间的两点观察。《数学进展》**200**(**1**) (2006),125–135. DOI: 10.1016/j.aim.2004.11.003.[9] Talagrand, M. 随机过程的上界与下界。第60卷。数学与相关领域成果。第3辑。现代数学综述丛书。现代方法与经典问题。施普林格，海德堡，2014，xvi+626.ISBN: 978-3-642-54074-5;978-3-642-54075-2. DOI: 10.1007/978-3-642-54075-2.

[10] Vershynin, R. 高维概率。第47卷。剑桥统计与概率数学丛书。剑桥大学出版社，剑桥，2018，xiv+284.ISBN:978-1-108-41519-4. DOI: 10.1017/9781108231596.

Shandong University, Jinan, China.电子邮箱地址：tiankundiao@gmail.com

Shandong University, Jinan, China.电子邮箱地址：xuananghu7@gmail.com

Research University Higher School of Economics and Lomonosov Moscow State University, Moscow, Russia
电子邮箱地址：vulyanov@cs.msu.ru

Shandong University, Jinan, China.电子邮箱地址：wanghanchao@sdu.edu.cn