arXiv:2304.03234v3 [math.CO] 5 Nov 2024

´关于具有随机公差的塞梅雷迪定理的阈值

¨JOP BRIET 和 DAVI CASTRO-SILVA

Abstract. 利用局部可译码理论的最新进展，我们证明，对于长度为k的等差数列，Szemer´edi 定理中随机公差的临界大小上界为 $N^{1-\frac{2}{k}+o_{(1)}}$。这改进了此前对所有奇数 k所知的最佳上界 $N^{1-\frac{1}{\lceil k/2\rceil}+o_{(1)}}$。

1. Introduction

塞梅雷´迪 [20] 证明了整数的稠密集合包含任意长的等差数列，这一结果已成为加法组合学的标志性成果。多年来，人们找到了这一结果的多种证明，运用了组合学、遍历理论以及有限阿贝尔群上的傅里叶分析等思想。

Furstenberg 的遍历论证明 [14] 开启了一系列强有力推广的大门。尤其是，它促成了 Szemer´edi 定理的一些变体，在这些变体中，等差数列的公差被限制在非常稀疏的集合内。若任意正密度集合 $A \subseteq [N]$ 都包含一个公差属于 D的（$t+1$）项等差数列，我们就称集合 $D \subseteq [N]$ t-交。Szemer´edi 定理蕴含，当 N_0足够大时，集合 $\{0,1,\dots,N_0\}$ 对所有$N \geq N_0$都是 t-交的。非平凡的例子包括 Bergelson 和 Leibman 的一项结果[3]：完全平方数（更一般地说，常数项为零的整系数多项式的像）对每个 t都是 t-交的；以及 Wooley 和 Ziegler 的一项结果的特殊情形 [23]，该结果表明素数减一也具有相同性质。

此类稀疏交集集的存在，促使人们研究随机稀疏集合是否通常具有交集性。要将这一问题定量化，则属于阈值现象的研究范畴。若由一族 $\mathcal{F} \subseteq 2^{[N]}$ 给出的 $[N]$ 的子集的某种性质满足 $A \in \mathcal{F}$ 且 $A \subseteq B \subseteq [N]$ 推出 $B \in \mathcal{F}$，我们就称该性质具有单调性。某种性质的临界大小 $m^* = m^*(N)$ 是使得从 $[N]$ 中均匀随机选取的一个 m 元子集以至少 $1/2$ 的概率具有该性质的最小 m。（如果 $\mathcal{F}$ 非空且具有单调性，则该值存在，因为这一概率会随之增加

本研究得到了荷兰研究理事会（NWO）NETWORKS计划的资助（资助编号：024.002.003）。本研究的扩展摘要发表于 EUROCOMB'23 [7] 会议论文集。

单调地随.....变化 m。）博洛巴什 ´和托马森的著名结果 [4] 断言，每个单调性质都有一个阈值函数；也就是说，概率

$$p(m) = \Pr_{A \in \binom{[N]}{m}}[A \in \mathcal{F}]$$

当 m 从 $o(m^*)$增加到 $\omega(m^*)$时，会从 $o(1)$激增至 $1 - o(1)$。一般来说，确定单调性质的临界大小是出了名的困难。

对于具有 t-交性质的集合，这个问题也仍然悬而未决。该性质显然具有单调性，我们用 $m_t^*(N)$ 表示其临界大小。Bourgain [5] 证明，1-交集合的临界大小为 $m_1^*(N) \asymp \log N$；目前，这是唯一已知精确界的情形。有人猜测 [12] 对于所有固定的 t，$\log N$ 是正确的界；事实上，尚无针对$t \geq 2$的更优下界。Frantzikinakis、Lesigne 和 Wierdl [13] 以及 Christ [11] 独立证明了：

$$m_2^*(N) \ll N^{\frac{1}{2}+o(1)}. \tag{1}$$

后来，第一作者、Dvir 和 Gopi [8] 证明，同样的上界也适用于 $m_3^*(N)$。更一般地，他们证明了

$$m_t^*(N) \ll N^{1-\frac{1}{\lceil (t+1)/2 \rceil}+o(1)}, \tag{2}$$

这一结果改进了此前对所有 $t \geq 3$已知的界。这些界中出现上取整函数，是因为可以将偶数 t的情形归约到 $t+1$的情形。这种归约源自对局部可译码纠错码的研究 [16]。文献 [8] 证明了，$(t+1)$-查询局部可译码码（LDC）的码长下界可推出 m_t^*的上界。界（2）由当时已知的最佳 LDC 界直接推出；不过，[9] 给出了界（2）的直接证明。

出于同样的原因，Alrabiah 等人 [1] 近期在 3 查询 LDC 方面取得的突破，立即意味着可以将（1）改进为

$$m_2^*(N) \ll N^{\frac{1}{3}+o(1)}.$$

由于技术原因，他们的方法无法直接推广，以改进 q查询 LDC（其中 $q \geq 4$）的界。这里，我们运用 [1] 的思想，直接证明 m_t^*的上界。由于我们的问题具有额外的算术结构，论述可以更加简洁；更重要的是，这使我们能够运用这些方法，改进此前对所有偶数 $t \geq 2$已知的最佳界。

定理 1.1。对于任意整数 $t \geq 2$，有

$$m_t^*(N) \ll N^{1-\frac{2}{t+1}+o(1)}.$$

这里给出的论证实际上适用于更一般的情形，对于大小与 t互质的任意有限加法群 G 也成立！（这样在考虑（$t+1$）项等差数列时就不会遇到整除性问题。）

令 G 为一个有限加法群，$t \geq 1$ 为一个整数，且 $\varepsilon \in (0,1)$。如果集合 $S \subseteq G$ 满足：$A \subseteq G$ 的每个大小为 $|A| \geq \varepsilon|G|$ 的子集都包含一个$(t+1)$项等差数列，且其公差属于 D，则称 $S \subseteq G$ 是(t,ε)-相交的。我们将 G 中具有 (t,ε)-相交性质的临界大小记为 $m^*_{t,\,\varepsilon}(G)$。我们的主要结果如下：

定理 1.2。对于每个 $t \geq 2$ 和 $\varepsilon \in (0,1)$，存在 $C(t,\varepsilon) > 0$ 使得

$$m^*_{t,\varepsilon}(G) \leq C(t,\varepsilon)(\log|G|)^{2t+3}|G|^{1-\frac{2}{t+1}}$$

对于每个大小与 $t!$ 互质的加法群 G 。

注意，定理 1.1 可通过将 $[N]$ 嵌入形如 $\mathbb{Z}/p\mathbb{Z}$的群中而由这一最后结果轻易得到，其中 p 是介于 $(t+1)N$ 和 $2(t+1)N$ 之间的素数。标准细节从略。

2. Preliminaries

2.1. 记号。我们用 $\biguplus$ 表示不交并。我们的（标准）渐近记号定义如下。给定一个趋于无穷大的参数 n 和一个函数 $f\colon \mathbb{R}_+ \to \mathbb{R}_+$，我们记：$g(n) = o(f(n))$，表示 $g(n)/f(n) \to 0$；$g(n) = \omega(f(n))$，表示 $g(n)/f(n) \to \infty$；$g(n) \ll f(n)$，表示对某个常数 $C > 0$ 及所有 n，$g(n) \leq Cf(n)$ 成立；而 $g(n) \asymp f(n)$，表示 $g(n) \ll f(n)$ 和 $f(n) \ll g(n)$ 均成立。当渐近式中的隐含常数依赖于某个参数（例如 ε）时，我们在渐近记号中将该参数写作下标以示区别（例如将 $\ll$ 替换为 $\ll_\varepsilon$）。

2.2. 矩阵范数与不等式。我们的论证将高度依赖于对高维矩阵的分析。这里我们回顾所需的矩阵不等式。

如果 $M \in \mathbb{R}^{d\times d}$ 是一个矩阵，我们定义其算子范数

$$\begin{aligned}\|M\|_2 &= \max\left\{u^TMv:\ \|u\|_2 = \|v\|_2 = 1\right\}\\ \|M\|_{\infty\to 1} &= \max\left\{u^TMv:\ \|u\|_\infty = \|v\|_\infty = 1\right\}\\ \|M\|_{1\to 1} &= \max\left\{u^TMv:\ \|u\|_\infty = \|v\|_1 = 1\right\}.\end{aligned}$$

我们将使用以下简单的不等式：

$$\|M\|_{\infty\to 1} \leq d\|M\|_2, \quad \|M\|_{\infty\to 1} \leq \sum_{i=1}^{d}\|M(i,\cdot)\|_1$$

并且，当 M 对称时，

$$\|M\|_2 \leq \|M\|_{1\to 1}.$$

我们还将使用以下非交换版的 Khintchine 不等式，它可以从 Tomczak-Jaegermann 的一项结果中推得 [21]:

定理 2.1. 设 $n, d \geq 1$ 为整数，并设 $A_1, \ldots, A_n$ 为任意一列 $d \times d$实矩阵。则

$$\mathbb{E}_{\sigma \in \{-1,1\}^n} \left\| \sum_{i=1}^n \sigma_i A_i \right\|_2 \leq 10\sqrt{\log d} \left(\sum_{i=1}^n \|A_i\|_2^2 \right)^{1/2}.$$

2.3. 多项式集中界。我们将需要 Kim 和 Vu 提出的一个著名的多项式集中不等式 [17], ，这需要引入一些额外记号。设 $H = (V, E)$ 是一个超图，其中允许重复的边（因此 E 可以是多重集），并设 $f\colon \{0,1\}^V \to \mathbb{R}$ 为如下多项式

$$f(x) = \sum_{e \in E} \prod_{v \in e} x_v. \tag{3}$$

对于集合 $A \subseteq V$，定义

$$f_A(x) = \sum_{e \in E\colon A \subseteq e} \prod_{v \in e \setminus A} x_v,$$

其中，与空集对应的单项式定义为 1。对于 $p \in (0,1)$，若 X 的各个坐标相互独立，且每个坐标以概率 p 取值 1（以概率 $1-p$取值 0），则称 X 为 p-伯努利随机变量，定义在 $\{0,1\}^V$上，记作 $X \sim \mathrm{Bern}(p)^V$。对于每个 $i \in \{0, 1, \ldots, |V|\}$，定义

$$\mu_i = \max_{A \in \binom{V}{i}} \mathbb{E}_{X \sim \mathrm{Bern}(p)^V} f_A(X).$$

注意， μ_0 正是 $f(X)$的期望。另定义以下量

$$\mu = \max_{i \in \{0,1,\ldots,|V|\}} \mu_i \qquad \text{and} \qquad \mu' = \max_{i \in \{1,2,\ldots,|V|\}} \mu_i.$$

Kim 和 Vu 的多项式集中不等式如下:

定理 2.2。 对于任意 $k \in \mathbb{N}$，存在常数 $C, C' > 0$ 使得以下结论成立。设 $H = (V, E)$ 是一个 n顶点超图，其边的大小至多为 k，并设 f 由 (3)给出。则对于任意 $\lambda > 1$，有

$$\Pr\big[|f(X) - \mu_0| > C\lambda^{k-\frac{1}{2}}\sqrt{\mu\mu'}\big] \leq C' \exp\big(-\lambda + (k-1)\log n\big).$$

为了满足我们的需要，我们将使用这一结果的一个轻微变体；该变体可由此结果和以下基本命题轻易推出。

命题 2.3. 设 $f\colon \{0,1\}^n \to \mathbb{R}_+$ 是一个单调递增函数，且 $p \in (\frac{16}{n}, 1)$。那么，对于任意整数 $0 \leq t \leq pn/2$,

$$\mathbb{E}_{S \in \binom{[n]}{t}} f(1_S) \leq \frac{1}{2} \mathbb{E}_{X \sim \mathrm{Bern}(p)^n} f(X).$$

证明：直接计算可得，

$$\begin{aligned}\mathbb{E}_{X\sim\mathrm{Bern}(p)^n} f(X) &= \sum_{i=0}^{n} p^i(1-p)^{n-i} \sum_{S\in\binom{[n]}{i}} f(1_S)\\ &= \sum_{i=0}^{n} \binom{n}{i} p^i(1-p)^{n-i}\, \mathbb{E}_{S\in\binom{[n]}{i}} f(1_S)\\ &\geq \sum_{i\geq t} \binom{n}{i} p^i(1-p)^{n-i}\, \mathbb{E}_{S\in\binom{[n]}{t}} f(1_S)\\ &\geq \frac{1}{2}\,\mathbb{E}_{S\in\binom{[n]}{t}} f(1_S),\end{aligned}$$

其中第三行使用了 f 的单调性，第四行由 Chernoff 界得出。□

推论 2.4。对于每个 $k\in\mathbb{N}$，存在常数 C，$C'>0$ 使得以下结论成立。设 $H=(V,\ E)$ 是一个 n-顶点超图，其边的大小至多为 k，设 f如 (3) 所示给定，并设 $p\in(\frac{16}{n},\ 1)$。那么，对于任意整数 $0\leq t\leq pn/2$，有

$$\Pr_{S\in\binom{V}{t}}\big[f(1_S)\geq C(\log n)^{k-\frac{1}{2}}\mu\big]\leq\frac{C'}{n^4}.$$

证明：对于充分大的常数 $C=C(k)>0$，令 $g:\{0,1\}^n\to\{0,1\}$ 为指示函数

$$g(1_S)=\mathbf{1}\big[f(1_S)\geq C(\log n)^{k-\frac{1}{2}}\mu\big].$$

由于 f 是单调的，因此 g。令 $\lambda=(3+k)\log n$，由定理 2.2 可得

$$\mathbb{E}_{X\sim\mathrm{Bern}(p)^n} g(X)\leq\frac{C'}{n^4}.$$

该结果现在由命题 2.3得出。 □

3. The main argument

在本节中，我们将证明定理1.2，这是我们的主要结果。将关注点从相交性次数 t 转移到相关等差数列的长度 $k:=t+1$。

于是固定一个整数 $k\geq 3$ 作为等差数列的长度，以及一个正参数 $\varepsilon>0$。令 G 为一个含有 N个元素的加法群，其中 N与 $(k-1)!$ 互素，并假设相对于 k 和 ε 足够大，以使我们的论证成立。回顾一下，我们希望证明 $m^*_{k-1,\ \varepsilon}(G)\ll_{k,\ \varepsilon}(\log\ N)^{2k+1}N^{1-\frac{2}{k}}$。

与其考虑随机交集集合，不如考虑随机交集序列。对于 G^m 中的一个序列，若其不同元素构成的集合是 $(k-1,\varepsilon)$-交集的，则称该序列是 $(k-1,\varepsilon)$-交集的。显然，均匀随机的 m 元素

序列是 $(k-1,\varepsilon)$-交集的概率，至多等于均匀随机的 m 元素集合具有该性质的概率。由于我们旨在证明临界大小的上界，只需界定最小的 m，使得 G^m 中的随机序列以至少 $1/2$ 的概率是 $(k-1,\varepsilon)$-交集的。

3.1. 化简为关于符号平均值的不等式。给定一个差值序列 $D=(d_1,\dots,d_m)\in G^m$ 以及某个集合 $A\subseteq G$，令 $\Lambda_D(A)$ 表示公差属于 D 且包含于 A中的 k-项等差数列的归一化计数：

$$\Lambda_D(A)=\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in G}\prod_{\ell=0}^{k-1}A(x+\ell d_i).$$

类似地，我们用 $\Lambda_G(A)$ 表示所有包含于 A中的 k-项等差数列所占的比例：

$$\Lambda_G(A)=\mathbb{E}_{d\in G}\mathbb{E}_{x\in G}\prod_{\ell=0}^{k-1}A(x+\ell d).$$

在证明的其余部分，我们假设 $m\in[N]$ 是一个满足以下条件的整数

$$\Pr_{D\in G^m}\big(\exists A\subseteq G:\ |A|\geq\varepsilon N,\ \Lambda_D(A)=0\big)>1/2, \tag{4}$$

这意味着随机的 m项序列不太可能是（$k-1$，ε）交相容的。注意，只要满足 $m<m^*_{k-1,\ \varepsilon}(G)$，此不等式就成立。我们的首要任务是将方程（4）化简为关于符号平均值的不等式（下文的方程（7）），对此我们有更多可用的工具。

通过对 Szemer´edi 定理的适当推广，我们知道

$$\Lambda_G(A)\gg_{k,\varepsilon}1\quad\text{for all } A\subseteq G \text{ with } |A|\geq\varepsilon N. \tag{5}$$

例如，可以利用Gowers [15]和纳格尔、罗¨德尔、沙赫特和斯科坎 [19, 18]的超图移除引理来证明。注意到 $\mathbb{E}_{D'\in G^m}\Lambda_{D'}(A)=\Lambda_G(A)$,通过合并不等式（4）和（5），我们得到

$$\mathbb{E}_{D\in G^m}\max_{A\subseteq G:\ |A|\geq\varepsilon N}\big|\Lambda_D(A)-\mathbb{E}_{D'\in G^m}\Lambda_{D'}(A)\big|\gg_{k,\varepsilon}1. \tag{6}$$

下文中，我们将不再需要条件 $|A|\geq\varepsilon N$ 在对 $A\subseteq G$取最大值时成立。这个正密度假设仅通过（5）使用。

接下来，我们应用 [9, 第 8690] 页中给出的一个简单对称化论证，将其改写成更便于使用的形式：

引理 1（对称化）。设 $c>0$，并假设

$$\mathbb{E}_{D\in G^m}\max_{A\subseteq G}\big|\Lambda_D(A)-\mathbb{E}_{D'\in G^m}\Lambda_{D'}(A)\big|\geq c.$$

Then

$$\mathbb{E}_{D\in G^m}\mathbb{E}_{\sigma\in\{-1,1\}^m}\max_{A\subseteq G}\left|\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in G}\,\sigma_i\prod_{\ell=0}^{k-1}A(x+\ell d_i)\right|\geq\frac{c}{2}.$$

证明：稍微滥用记号，对于每个 $d \in G$ 和 $A \subseteq G$，记作

$$\Lambda_d(A) = \mathbb{E}_{x \in G} \prod_{\ell=0}^{k-1} A(x + \ell d).$$

于是，根据三角不等式，引理 1 中的第一个表达式至多为

$$\mathbb{E}_{D,D' \in G^m} \max_{A \subseteq G} \Big| \frac{1}{m} \sum_{i=1}^{m} \big(\Lambda_{d_i}(A) - \Lambda_{d'_i}(A) \big) \Big|,$$

其中 D 和 D' 相互独立且服从均匀分布。由于项 $\Lambda_{d_i}(A)$ $-\Lambda_{d'_i}(A)$ 相互独立且对称分布，因此如果将这些项分别乘以任意符号，期望值保持不变。特别地，该期望等于

$$\mathbb{E}_{D,D' \in G^m} \mathbb{E}_{\sigma \in \{-1,1\}^m} \max_{A \subseteq G} \Big| \frac{1}{m} \sum_{i=1}^{m} \sigma_i \big(\Lambda_{d_i}(A) - \Lambda_{d'_i}(A) \big) \Big|.$$

再次使用三角不等式，并利用 D 和 D' 具有相同的分布这一事实，可得该式至多为

$$2\mathbb{E}_{D \in G^m} \mathbb{E}_{\sigma \in \{-1,1\}^m} \max_{A \subseteq G} \Big| \frac{1}{m} \sum_{i=1}^{m} \sigma_i \Lambda_{d_i}(A) \Big|.$$

这就证明了该结论。 □

对符号取期望 $\sigma \in \{-1,1\}^m$ 这一点对我们的论证至关重要。通过一个简单的多重线性论证，我们可以将集合 $A \subseteq G$ (可视为 $\{0,1\}^G$中的一个向量)替换为向量 $Z \in \{-1,1\}^G$。结合式 (6) 和引理 1，可得

$$\mathbb{E}_{D \in G^m} \mathbb{E}_{\sigma \in \{-1,1\}^m} \max_{Z \in \{-1,1\}^G} \Big| \mathbb{E}_{i \in [m]} \mathbb{E}_{x \in G} \, \sigma_i \prod_{\ell=0}^{k-1} Z(x + \ell d_i) \Big| \gg_{k,\varepsilon} 1. \tag{7}$$

从 $\{0,1\}^G$ 到 $\{-1,1\}^G$ 的变化是一种方便的技术处理，使我们可以忽略乘积中被平方的项。

3.2. 处理项数为奇数的情况。 最后一个不等式（7）正是我们需要用来证明 k 取偶数值时结果的工具，所采用的论证将在下文概述。然而，当 k取奇数值时，由于乘积中含有奇数个因子，这个不等式并不适用。处理这种情况的主要思路是从[1] 应用“柯西–施瓦茨技巧”，以得到一个更适用的不等式：

引理 2（柯西–施瓦茨技巧）。设 $c > 0$，并假设 $m \geq 2/c^2$ 是满足以下条件的整数

$$\mathbb{E}_{D \in G^m} \mathbb{E}_{\sigma \in \{-1,1\}^m} \max_{Z \in \{-1,1\}^G} \Big| \mathbb{E}_{i \in [m]} \mathbb{E}_{x \in G} \, \sigma_i \prod_{\ell=0}^{k-1} Z(x + \ell d_i) \Big| \geq c.$$

那么存在一个划分 $[m] = L \uplus R$，使得

$$\mathbb{E}_{D\in G^m}\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}} \max_{Z\in\{-1,1\}^G} \sum_{\substack{i\in L\\ j\in R}}\sum_{x\in G}\sigma_i\tau_j\prod_{\ell=1}^{k-1} Z(x+\ell d_i)Z(x+\ell d_j) \geq \frac{c^2m^2N}{8}.$$

证明：根据柯西-施瓦茨不等式，对于任意 $Z\in\{-1,1\}^G$，有

$$\begin{aligned}\left|\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in G}\,\sigma_i\prod_{\ell=0}^{k-1}Z(x+\ell d_i)\right|^2 &= \left|\mathbb{E}_{x\in G}\,Z(x)\cdot\left(\mathbb{E}_{i\in[m]}\sigma_i\prod_{\ell=1}^{k-1}Z(x+\ell d_i)\right)\right|^2\\ &\leq \left(\mathbb{E}_{x\in G}\,Z(x)^2\right)\mathbb{E}_{x\in G}\left(\mathbb{E}_{i\in[m]}\sigma_i\prod_{\ell=1}^{k-1}Z(x+\ell d_i)\right)^2\\ &= \mathbb{E}_{x\in G}\mathbb{E}_{i,j\in[m]}\,\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j).\end{aligned}$$

再次应用柯西-施瓦茨不等式，由我们的假设可得

$$\begin{aligned}c^2 &\leq \mathbb{E}_{D\in G^m}\mathbb{E}_{\sigma\in\{-1,1\}^m}\max_{Z\in\{-1,1\}^G}\left|\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in G}\,\sigma_i\prod_{\ell=0}^{k-1}Z(x+\ell d_i)\right|^2\\ &\leq \mathbb{E}_{D\in G^m}\mathbb{E}_{\sigma\in\{-1,1\}^m}\max_{Z\in\{-1,1\}^G}\mathbb{E}_{x\in G}\mathbb{E}_{i,j\in[m]}\,\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j).\end{aligned}$$

现在考虑均匀随机划分 $[m] = L \uplus R$，因此对于任意 i、$j\in[m]$，当 $i\neq j$ 时，我们有 $\mathbf{Pr}_{L,\ R}(i\in L,\quad j\in R) = 1/4$；于是

$$\begin{aligned}&\mathbb{E}_{i,j\in[m]}\,\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j)\\ &\quad= \frac{1}{m^2}\sum_{\substack{i,j=1\\ i\neq j}}^{m}\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j) + \frac{1}{m^2}\sum_{i=1}^{m}\sigma_i^2\prod_{\ell=1}^{k-1}Z(x+\ell d_i)^2\\ &\quad= \frac{4}{m^2}\mathbb{E}_{L,R}\sum_{i\in L,j\in R}\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j) + \frac{1}{m}.\end{aligned}$$

由此可得

$$c^2 \leq \frac{1}{m} + \frac{4}{m^2}\mathbb{E}_{L,R}\mathbb{E}_{D\in G^m}\mathbb{E}_{\sigma\in\{-1,1\}^m}\max_{Z\in\{-1,1\}^G}\mathbb{E}_{x\in G}\sum_{i\in L,j\in R}\sigma_i\sigma_j\prod_{\ell=1}^{k-1}Z(x+\ell d_i)Z(x+\ell d_j).$$

利用 $m\geq 2/c^2$，我们得出存在一个划分 $[m] = L \uplus R$满足该引理的结论。□

从现在起，我们假设 k 为奇数，并记 $k = 2r + 1$。[1] 对于 $i,\ j \in [m]$记$P_i(x)$ $= \{x + d_i, x + 2d_i, \dots, x + 2rd_i\}$和 $P_{ij}(x) = P_i(x) \cup P_j(x)$，其中为便于记号起见，我们省略了对差集 D 的依赖。由不等式（7）和引理2 我们得到

$$\mathbb{E}_{D\in G^m}\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}} \max_{Z\in\{-1,1\}^G} \sum_{\substack{i\in L\\ j\in R}} \sum_{x\in G} \sigma_i\tau_j \prod_{y\in P_{ij}(x)} Z(y) \gg_{k,\varepsilon} m^2N, \tag{8}$$

其中，(L, R) 是指标集 $[m]$ 的一个合适划分，并且我们假设（不失一般性） m 足够大，具体取决于 ε 和 k。

F由不等式（8）可知，我们可以固定一个满足以下条件的“好”集合 $D \in G^m$

$$\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}} \max_{Z\in\{-1,1\}^G} \sum_{\substack{i\in L\\ j\in R}} \sigma_i\tau_j \sum_{x\in G} \prod_{y\in P_{ij}(x)} Z(y) \gg_{k,\varepsilon} m^2N \tag{9}$$

并且该集合满足技术条件

$$\big|\big\{i \in L, j \in R :\ |P_{ij}(0)| \neq 4r\big\}\big| \ll_{k,\varepsilon} m^2/N \quad \text{and} \tag{10}$$

$$\max_{x\neq 0} \sum_{i=1}^{m} \sum_{\ell=-2r}^{2r} \mathbf{1}\{\ell d_i = x\} \ll \log N, \tag{11}$$

这些条件对于之后界定某些坏事件的概率是必需的。事实上：

- 对于给定序列$D \in G^m$，记 $X(D)$ 为（9）左侧的表达式。则 $X(D) \leq m^2N$ 总是成立；而由等式（8）可得$\mathbb{E}_{D\in G^m} X(D) \geq c_{k,\ \varepsilon} m^2N$ 其中 c_k 是某个常数，$\varepsilon > 0$。因此

$$\Pr\big[X(D) \geq c_{k,\varepsilon} m^2N/2\big] > c_{k,\varepsilon}/2.$$

- 对于 $\ell, \ell' \in [2r]$ 以及相互独立且均匀分布的 $d_i, d_j \in G$，我们有

$$\Pr[\ell d_i = \ell' d_j] = 1/N.$$

(10) 左边（关于 $D \in G^m$）的期望为

$$\sum_{\substack{i\in L\\ j\in R}} \Pr\big[|P_{ij}(0)| \neq 4r\big] \leq \sum_{\substack{i\in L\\ j\in R}} \sum_{\ell,\ell'=1}^{2r} \Pr[\ell d_i = \ell' d_j] \leq \frac{r^2m^2}{N},$$

因此，根据马尔可夫不等式

$$\Pr\left[\big|\big\{i \in L, j \in R :\ |P_{ij}(0)| \neq 4r\big\}\big| \leq \frac{4}{c_{k,\varepsilon}} \frac{r^2m^2}{N}\right] \geq 1 - \frac{c_{k,\varepsilon}}{4}.$$

[1]偶数情形类似，但更简单。这里我们聚焦于奇数情形，因为新界限正是在这种情况下得到的。

• 对于固定的 $x \neq 0$，式（11）中的内层求和是一个指示随机变量，其以概率 $4r/N$ 取值为 1。由于不同 $i \in [m]$对应的这些随机变量相互独立，Chernoff 界意味着

$$\Pr\left[\sum_{i=1}^{m}\sum_{\ell=-2r}^{2r}\mathbf{1}\{\ell d_i = x\} \geq (1+\delta)\frac{4rm}{N}\right] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{4rm/N}.$$

令 $1+\delta = N\log N/(rm)$，对所有 $x \neq 0$ 取并集上界可得变态另类

$$\Pr\left[\max_{x\neq 0}\sum_{i=1}^{m}\sum_{\ell=-2r}^{2r}\mathbf{1}\{\ell d_i = x\} \leq 4\log N\right] \geq 1 - \frac{1}{N^3} > 1 - \frac{c_{k,\varepsilon}}{4}.$$

根据并集界，上述三个条件将以正概率同时成立，正如所愿。

3.3.化归为矩阵不等式问题。下一个关键想法是构造矩阵 M_{ij}，使得该量

$$\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}}\left\|\sum_{i\in L, j\in R}\sigma_i\tau_j M_{ij}\right\|_{\infty\to 1} \tag{12}$$

与不等式（9）左侧的表达式相关。这样做的原因是，这使我们能够使用强大的 矩阵集中不等式，从而可以得到期望（12）的良好上界；这进而转化为 m 作为 N的函数的上界，这正是我们的目标。此类矩阵不等式的用法可追溯至 Ben-Aroya、Regev 和 de Wolf 的工作 [2], ，其工作又受 Kerenidis 和 de Wolf 的研究启发 [16] （另见 [6]）。

我们将构造的矩阵以大小为 s 的集合为指标，其中（事后看来）我们选取 $s = \lfloor N^{1-2/k}\rfloor$。回顾一下， $k = 2r+1$。对于 $i \in L$, $j \in R$，定义矩阵 $M_{ij} \in \mathbb{R}^{\binom{G}{s}\times\binom{G}{s}}$ 为

$$M_{ij}(S,T) = \sum_{x\in G}\mathbf{1}\big\{|S\cap P_i(x)| = |S\cap P_j(x)| = r,\, S\triangle T = P_{ij}(x)\big\}$$

如果 $|P_{ij}(0)| = 4r$，且 $M_{ij}(S,\ T) = 0$如果 $|P_{ij}(0)| \neq 4r$；请注意，尽管这些矩阵的定义并不对称，但它们实际上是对称的。接下来，我们将由不等式（9）推出期望（12）的下界。

对于一个 向量 $Z \in \{-1,1\}^G$，用 $Z^{\odot s} \in \{-1,1\}^{\binom{G}{s}}$ 表示由以下给出的“提升”向量

$$Z^{\odot s}(S) = \prod_{y\in S} Z(y) \quad \text{for all } S \in \binom{G}{s}.$$

如果 $|P_{ij}(0)| = 4r$，那么对于所有 $Z \in \{-1,1\}^G$，我们有

$$\begin{aligned}\sum_{S,T\in\binom{G}{s}} M_{ij}(S,T)Z^{\odot s}(S)Z^{\odot s}(T) &= \sum_{S,T\in\binom{G}{s}} M_{ij}(S,T)\prod_{y\in S\triangle T} Z(y)\\ &= \sum_{x\in G}\sum_{S\in\binom{G}{s}} \mathbf{1}\{|S\cap P_i(x)| = |S\cap P_j(x)| = r\}\prod_{y\in P_{ij}(x)} Z(y)\\ &= \binom{2r}{r}^2\binom{N-4r}{s-2r}\sum_{x\in G}\prod_{y\in P_{ij}(x)} Z(y), \qquad (13)\end{aligned}$$

由于有 $\binom{2r}{r}^2\binom{N-4r}{s-2r}$ 种选择集合 $S \in \binom{G}{s}$ 使其满足 $|S\cap P_i(x)| = |S\cap P_j(x)| = r$，并且一旦选定这样的集合 S，就只有一个集合 $T \in \binom{G}{s}$ 满足$S\triangle T = P_{ij}(x)$。因此

$$\begin{aligned}&\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}}\Big\|\sum_{i\in L,j\in R}\sigma_i\tau_j M_{ij}\Big\|_{\infty\to 1}\\ &\qquad\geq \mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}}\max_{Z\in\{-1,1\}^G}\sum_{S,T\in\binom{G}{s}}\sum_{i\in L,j\in R}\sigma_i\tau_j M_{ij}(S,T)Z^{\odot s}(S)Z^{\odot s}(T)\\ &\qquad= \mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}}\max_{Z\in\{-1,1\}^G}\binom{2r}{r}^2\binom{N-4r}{s-2r}\sum_{\substack{i\in L,j\in R\\ |P_{ij}(0)|=4r}}\sigma_i\tau_j\sum_{x\in G}\prod_{y\in P_{ij}(x)} Z(y);\end{aligned}$$

将其与不等式（9）和（10）结合，我们得到下界

$$\mathbb{E}_{\substack{\sigma\in\{-1,1\}^L\\ \tau\in\{-1,1\}^R}}\Big\|\sum_{i\in L,j\in R}\sigma_i\tau_j M_{ij}\Big\|_{\infty\to 1} \gg_{k,\varepsilon} \binom{N-4r}{s-2r}m^2N. \qquad (14)$$

3.4. 应用 Khintchine 型不等式。 现在我们需要计算上述期望的上界。这里的主要思路是使用定理 2.1中给出的非交换版本的 Khintchine 不等式。直观地说，由于随机符号 σ_i 的存在，该不等式表明最后一个表达式中的求和会发生大量抵消，因此左侧的期望（14）远小于人们可能预期的值。

为了应用定理2.1，最好将矩阵 M_{ij} 分组，并且只使用一半的随机符号 σ_i （另一个想法来自[1]）。对于 $i \in L$， $\tau \in \{-1,1\}^R$，我们定义矩阵

$$M_i^\tau = \sum_{j\in R}\tau_j M_{ij}.$$

随后我们将给出该表达式的一个上界

$$\max_{\tau\in\{-1,1\}^R}\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i M_i^\tau\Big\|_{\infty\to 1}$$

而其本身是（14）中的期望的上界。

为实现这一目标，我们将通过移除所有 ℓ_1-权重显著高于平均值的行和列来对矩阵 M_i^T 进行剪枝。由这些矩阵的对称性和非负性可知，以集合 $S \in \binom{G}{s}$ 为索引的行或列的 ℓ_1-权重有如下界：

$$\begin{aligned}\sum_{T\in\binom{G}{s}} \left|\sum_{j\in R} \tau_j M_{ij}(S,T)\right| &\leq \sum_{T\in\binom{G}{s}} \sum_{j\in R} M_{ij}(S,T) \\ &= \sum_{\substack{j\in R \\ |P_{ij}(0)|=4r}} \sum_{x\in G} \mathbf{1}\big\{|S\cap P_i(x)| = |S\cap P_j(x)| = r\big\}.\end{aligned}$$

为了证明剪枝对最终界的影响很小，我们证明只有少数行和列具有较大的 ℓ_1-权重。为此，令 U为一个均匀分布的 $\binom{G}{s}$值随机变量，并对每个 $i \in L$，定义与上面最后一个表达式对应的随机变量：

$$X_i := \sum_{\substack{j\in R \\ |P_{ij}(0)|=4r}} \sum_{x\in G} \mathbf{1}\big\{|U\cap P_i(x)| = |U\cap P_j(x)| = r\big\}.$$

对（13）中的计算，取 Z 全 $\mathbf{1}$ 向量，可得

$$\mathbb{E}[X_i] = \frac{1}{\binom{N}{s}} \sum_{\substack{j\in R \\ |P_{ij}(0)|=4r}} \binom{2r}{r}^2 \binom{N-4r}{s-2r} N \ll_k \frac{1}{\binom{N}{s}} \binom{N-4r}{s-2r} mN.$$

由于 $s = \lfloor N^{1-2/k} \rfloor$，我们有 $\binom{N-4r}{s-2r}/\binom{N}{s} \ll_k (s/N)^{2r} \asymp N^{-(2-2/k)}$，因此

$$\mathbb{E}[X_i] \ll_k \frac{m}{N^{1-2/k}}. \tag{15}$$

下面的引理给出了关于 X_i 的上尾估计，前提是 m 足够大。

引理 3. 假设 $m \geq N^{1-2/k}$。那么，对于每个 $i \in L$，我们有

$$\Pr\Big[X_i \geq (\log N)^k \frac{m}{N^{1-2/k}}\Big] \leq \frac{1}{N^4}.$$

证明：固定一个 $i \in L$。考虑顶点集为 G、边集为如下内容的超图 H_i

$$E(H_i) = \biguplus_{\substack{j\in R \\ |P_{ij}(0)|=4r}} \biguplus_{x\in G} \binom{P_i(x)}{r} \times \binom{P_j(x)}{r},$$

并令 $f\colon\ \mathbb{R}^G \to \mathbb{R}$ 为与 H_i 中（3）所述的多项式，

$$f(t) = \sum_{e\in E(H_i)} \prod_{v\in e} t_v.$$

注意， $X_i = f(1_U)$，其中 U 在 $\binom{G}{s}$ 上均匀分布，而 $1_U \in \mathbb{R}^G$ 表示其（随机）指示向量。

对于每个 $0 \le \ell \le 2r$，我们希望对该量给出上界。

$$\mu_\ell := \max_{A \in \binom{G}{\ell}} \mathbb{E}_{t \sim \mathrm{Bern}(s/N)^G} f_A(t).$$

（回顾第 2 节中引入的记号。）由（15），我们有 $\mu_0 \ll_k mN^{-(1-2/k)}$。对于集合 $A \in \binom{G}{\ell}$，定义其在 H_i 中的度为

$$\deg(A) = |\{e \in E(H_i) : e \supseteq A\}|,$$

其中，我们将重复边的重数计入其中。注意，对于任意 $B \subseteq A$，我们有 $\deg(A) \le \deg(B)$。于是，

$$\mu_\ell = \max_{A \in \binom{G}{\ell}} \left(\frac{s}{N}\right)^{2r-\ell} \deg(A).$$

对于任意 $v \in G$，我们有 $\deg(v) \ll_k m$，因为 v 包含在 $O_k(1)$ 个长度为 k 且具有固定公差的等差数列中。因此，对于 $\ell \in [r]$，我们有

$$\mu_\ell \le \left(\frac{s}{N}\right)^{2r-\ell} \max_{v \in G} \deg(v) \ll_k mN^{-2r/(2r+1)} = \frac{m}{N^{1-1/k}}.$$

令 $A \subseteq G$ 为大小为 $\ell \in \{r+1, \dots, 2r\}$ 的集合，且

$$e \in \binom{P_i(x)}{r} \times \binom{P_j(x)}{r}$$

是 $E(H_i)$ 的一条边，其中包含 A。根据鸽巢原理， A 包含一个元素 $a \in P_i(x)$ 和一个元素 $b \in P_j(x)$。已知 a 将 x 限制在一个大小至多为 $2r$ 的集合中。此外，由 (11) 可知，对每个 x， $O_k(\log N)$ 个可能的 $j \in R$ 取值至多满足 $b \in P_j(x)$。因此，

$$\mu_\ell \ll_k \left(\frac{s}{N}\right)^{2r-\ell} \log N \le \log N.$$

根据我们对 m 的假设可知，对于每个 $\ell \in \{0, \dots, 2r\}$，都有 $\mu_\ell \ll_k mN^{-(1-2/k)} \log N$。结论现在可直接由推论 2.4 得出。□

引理 3 表明，对于每个矩阵 $M_i^\intercal$，所有行和列中至多有 N^{-4} 的比例具有超过 $(\log N)^k mN^{-(1-2/k)}$ 的 ℓ_1-权重。现在定义 $\widetilde{M_i^\intercal}$ 为从 $M_i^\intercal$ 得到的“修剪后”矩阵，即将所有此类重行和重列置零。注意， $M_i^\intercal$ 是对称的，因此

$$\|\widetilde{M_i^\intercal}\|_2 \le \|\widetilde{M_i^\intercal}\|_{1\to 1} = \max_{S \in \binom{G}{s}} \|\widetilde{M_i^\intercal}(S, \cdot)\|_1 \le (\log N)^k \frac{m}{N^{1-2/k}};$$

这个算子范数界使得处理剪枝后的矩阵更加方便。

我们首先证明，用原矩阵的剪枝版本替换原矩阵，对我们的界影响可以忽略不计。事实上，根据 X_i 的定义可知，其最大值不超过 mN，因此

$$\begin{aligned}\big\|M_i^\tau - \widetilde{M_i^\tau}\big\|_{\infty\to 1} &\le \sum_{S\in\binom{G}{s}} \big\|M_i^\tau(S,\cdot) - \widetilde{M_i^\tau}(S,\cdot)\big\|_1 \\ &\le 2\binom{N}{s}\cdot \mathbb{E}\big[X_i\,\mathbf{1}\big\{X_i \ge (\log N)^k mN^{-(1-2/k)}\big\}\big] \\ &\le 2\binom{N}{s}\cdot mN\Pr\big[X_i \ge (\log N)^k mN^{-(1-2/k)}\big].\end{aligned}$$

（第二个不等式中的乘以 2，是因为我们必须同时考虑重行和重列。）由引理 3 我们得出

$$\big\|M_i^\tau - \widetilde{M_i^\tau}\big\|_{\infty\to 1} \le \frac{2m}{N^3}\binom{N}{s} \quad \text{for all } i\in L,\ \tau\in\{0,1\}^R. \tag{16}$$

接下来，我们将定理 2.1 中的集中不等式应用于剪枝后的矩阵 $\widetilde{\ } M_i^\tau$；由此得到

$$\begin{aligned}\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i\widetilde{M_i^\tau}\Big\|_{\infty\to 1} &\le \binom{N}{s}\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i\widetilde{M_i^\tau}\Big\|_2 \\ &\le 10\binom{N}{s}\sqrt{\log\binom{N}{s}}\Big(\sum_{i\in L}\|\widetilde{M_i^\tau}\|_2^2\Big)^{1/2} \\ &\le 10\binom{N}{s}\sqrt{\log\binom{N}{s}}\Big(\sum_{i\in L}\|\widetilde{M_i^\tau}\|_{1\to 1}^2\Big)^{1/2} \\ &\le 10\binom{N}{s}\sqrt{s\log N}\cdot m^{1/2}(\log N)^k\frac{m}{N^{1-2/k}}.\end{aligned}$$

由三角不等式和我们先前得到的界，我们得出

$$\begin{aligned}\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i M_i^\tau\Big\|_{\infty\to 1} &\le \mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i\widetilde{M_i^\tau}\Big\|_{\infty\to 1} + \sum_{i\in L}\big\|M_i^\tau - \widetilde{M_i^\tau}\big\|_{\infty\to 1} \\ &\le 10\binom{N}{s}\sqrt{s\log N}\cdot m^{1/2}(\log N)^k\frac{m}{N^{1-2/k}} + \frac{2m^2}{N^3}\binom{N}{s}.\end{aligned}$$

3.5. 完成证明。现在基本上已经完成了，只需将所得的上界和下界结合起来即可。事实上，将最后一个不等式与公式（14）结合可得

$$\binom{N-4r}{s-2r}m^2N \ll_{k,\varepsilon} \binom{N}{s}\sqrt{ms\log N}(\log N)^k\frac{m}{N^{1-2/k}}.$$

重新整理并利用 $\binom{N}{s}/\binom{N-4r}{s-2r} \ll_k (N/s)^{2r} = N^{2-2/k}$，我们得出

$$m \ll_{k,\varepsilon} s(\log N)^{2k+1} = N^{1-2/k}(\log N)^{2k+1}.$$

由于我们从假设（4）出发，这表明 $m^*_{k-1,\,\varepsilon}(G) \ll_{k,\,\varepsilon} N^{1-2/k}(\log\ N)^{2k+1}$，正如所希望的那样。

4. Discussion

我们对 $m_t^*(N)$ 的界远未达到猜测的 $\Theta_t(\log\ N)$，而且我们认为这些界并非最佳。下面简要提及几条可能有助于改进界的研究方向，为便于说明，重点讨论涉及三项等差数列的情形 $m_2^*(N)$：

- 我们论证中可能存在的低效之处在于，经过对称化步骤（引理 1）之后，$D \in G^m$ 是随机序列这一事实并未以任何重要方式加以利用。[2] 对 $m_2^*(N)$ 的界进行改进，或许可以利用本证明中考察的最坏情形 D 与问题所涉及的平均情形之间可能存在的差异。
- 另一种可能性是，对于固定序列 D，使用非交换 Khintchine 不等式的多线性版本，直接界定引理 1 中的最终表达式。赋予三线性形式（或张量）$\mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$以范数

$$\|T\| = \sup\Big\{|T(x,y,z)| \mid \|x\|_3, \|y\|_3, \|z\|_3 \le 1\Big\}.$$

对于三线性形式 $T_1, \dots, T_m$，形如

$$\mathbb{E}_{\sigma\in\{-1,1\}^m}\Big\|\sum_{i=1}^m \sigma_i T_i\Big\| \le C(N)\Big(\sum_{i=1}^m \|T_i\|^2\Big)^{\frac{1}{2}}$$

的界将意味着 $m_2^*(N) \ll C(N)^2$。

本文所用的技术为形如以下形式的置换张量建立了目前已知的最佳界

$$T(x,y,z) = \sum_{i=1}^N x_i y_{\pi_1(i)} z_{\pi_2(i)},$$

其中 $\pi_1,\ \pi_2 \in S_N$ 是排列；这种情形足以处理 Λ_D 我们证明中出现的形式。我们认为，这种方法对排列张量得到的界并非最优，而该问题上更精确的界将有助于改进 $m_2^*(N)$。但请注意，单靠这条途径不足以得到形如 $m_2^*(N) \le \operatorname{poly}\log(N)$ 的结果：存在一列排列张量（源自 LDC 构造），它们表明 $C(N) \ge (\log\ N)^{\omega(1)}$ 是必要的 [10, 6]。

[2] 它被以较弱的方式使用，以获得技术条件（10）和（11），但这些大多只是技术细节，对主要论证并非必不可少。

• 本文使用的主要技术工具是非交换 Khintchine 不等式（定理 2.1）。一般而言，该不等式是尖锐的，但当所考虑的矩阵族高度非交换时，可以对其改进；关于这一点的讨论，见 [22, 第 7] 章。不过，我们的矩阵相当接近可交换，因此一种可能的改进途径是，为方程 $(\sum_{x\in G}\prod_{y\in P_{ij}(x)}Z(y)$ 中出现的多线性形式寻找真正非交换的矩阵嵌入（9）。

References

[1] Omar Alrabiah、Venkatesan Guruswami、Pravesh K. Kothari 和 Peter Manohar。《基于半随机 CSP 驳斥的 3 查询局部可译码码的近立方下界》。载于：STOC 2023。美国佛罗里达州奥兰多：计算机协会，2023 年，第 1438–1448 页。 ISBN：9781450399135。 DOI：10.1145/3564246.3585143。[2] Avraham Ben-Aroya、Oded Regev 和 Ronald de Wolf。《矩阵值函数的超收缩不等式及其在量子计算和 LDC 中的应用》。载于：2008 IEEE 第 49 届 IEEE 计算机科学基础年会（FOCS）。IEEE 计算机学会，2008 年，第 477–486 页。 DOI：10.1109/FOCS.2008.45。[3] V. Bergelson 和 A. Leibman。《van der Waerden 定理和 Szemer´edi 定理的多项式推广》。载于：J. Amer. Math. Soc. 9.3（1996），第 725–753 页。 ISSN：0894-0347。 DOI：10.1090/S0894-0347-96-00194-4。[4] B. Bollob´as 和 A. Thomason。《阈值函数》。载于：Combinatorica 7.1（1987），第 35–38 页。 ISSN：0209-9683。 DOI：10.1007/BF02579198。[5] J. Bourgain。《Ruzsa 关于复现集合的问题》。载于：Israel J. Math. 59.2（1987），第 150–166 页。 ISSN：0021-2172。 DOI：10.1007/BF02787258。[6] Jop Bri¨et。《局部可译码码的嵌入》。2016 年。 DOI：doi.org/10.48550/arXiv.16

arXiv: 1611.06385 [cs.CC].[7] Jop Bri¨et 和 Davi Castro-Silva。“以随机公差提升 Szemer´edi 定理的阈值上限”。载于：第 12 届欧洲组合学、图论与应用会议论文集。Daniel Kr´al’ 和 Jaroslav Neˇsetˇril 编。EUROCOMB 2023。马萨里克大学，布尔诺，2023 年，第 231–237 页。 ISBN：978-80-280-0344-9。 DOI：10.5817/CZ.MUNI.EUROCOMB23-032。[8] Jop Bri¨et、Zeev Dvir 和 Sivakanth Gopi。“局外分布与局部可译码”。载于：Theory Comput. 15 (2019)，第 12 号论文，24 页。 DOI：10.4086/toc.2019.v015a012。[9] Jop Bri¨et 和 Sivakanth Gopi。“高斯宽度界及其在随机情形下等差数列中的应用”。载于：Int. Math. Res. Not. IMRN 22 (2020)，第 8673–8696 页。 ISSN：1073-7928。 DOI：10.1093/imrn/rny238。[10] Jop Bri¨et、Assaf Naor 和 Oded Regev。“局部可译码与射影张量积的余型失效”。载于：Electron. Res. Announc. Math. Sci. 19 (2012)，第 120–130 页。 ISSN：1935-9179。 DOI：10.3934/era.2012.19.120。 URL：https://doi.org/10.3934/era.2012.19.120。

[11] Michael Christ。《随机多线性算子不等式》。2011 年。 DOI：10.48550/ARXIV.1108.5655。[12] Nikos Frantzikinakis、Emmanuel Lesigne 和 M´at´e Wierdl。“Szemer´edi 定理中的随机公差及相关结果”。载于：J. Anal. Math. 130 (2016)，第 91–133 页。 ISSN：0021-7670。 DOI：10.1007/s11854-016-0030-z。[13] Nikos Frantzikinakis、Emmanuel Lesigne 和 M´at´e Wierdl。“随机序列与多重遍历平均的逐点收敛”。载于：Indiana Univ. Math. J. 61.2 (2012)，第 585–617 页。 ISSN：0022-2518。 DOI：10.1512/iumj.2012.61.4571。[14] Harry Furstenberg。“对角测度的遍历行为与 Szemer´edi 关于等差数列的定理”。载于：J. Analyse Math. 31 (1977)，第 204–256 页。 ISSN：0021-7670。 DOI：10.1007/BF02813304。[15] W Timothy Gowers。“超图正则性与多维 Szemer´edi 定理”。载于：Annals of Mathematics 166 (2007)，第 897–946 页。 DOI：10.4007/annals.2007.166.897。[16] Iordanis Kerenidis 和 Ronald de Wolf。“通过量子论证证明 2 查询局部可译码的指数下界”。载于：J. Comput. System Sci. 69.3 (2004)。早期版本发表于 STOC’03，第 395–420 页。 ISSN：0022-0000。 DOI：10.1016/j.jcss.2004.04.007。[17] Jeong Han Kim 和 Van H. Vu。“多元多项式的集中性及其应用”。载于：Combinatorica 20.3 (2000)，第 417–434 页。 ISSN：0209-9683。DOI：10.1007/s004930070014。[18] Brendan Nagle、Vojtˇech R¨odl 和 Mathias Schacht。“正则 k一致超图的计数引理”。载于：Random Structures Algorithms 28.2 (2006)，第 113–179 页。 ISSN：1042-9832。 DOI：10.1002/rsa.20117。 URL：https://doi.org/10.1002/rsa.20117。[19] Vojtˇech R¨odl 和 Jozef Skokan。“ k一致超图的正则性引理”。载于：Random Structures Algorithms 25.1 (2004)，第 1–42 页。 ISSN：1042-9832。 DOI：10.1002/rsa.20017。 URL：https://doi.org/10.1002/rsa.20017。[20] E. Szemer´edi。“不含 k 个元素等差数列的整数集合”。载于：Acta Arith. 27 (1975)，第 199–245 页。 ISSN：0065-1036。 DOI：10.4064/aa-27-1-199-245。[21] Nicole Tomczak-Jaegermann。“光滑模数与凸性模数，以及迹类的 Rademacher 平均 $S_p(1 \leq p < \infty)$”。载于：Studia Math. 50 (1974)，第 163–182 页。 ISSN：0039-3223。[22] Joel A. Tropp。《矩阵集中不等式导论》。载于：Foundations and Trends® in Machine Learning 8.1-2 (2015)，第 1–230 页。 DOI：10.1561/2200000048。URL：http://dx.doi.org/10.1561/2200000048。[23] Trevor D. Wooley 和 Tamar D. Ziegler。“沿素数的多重回归与收敛”。载于：Amer. J. Math. 134.6 (2012)，第 1705–1732 页。 ISSN：0002-9327。 DOI：10.1353/ajm.2012.0048。

CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands

Email address: j.briet@cwi.nl

CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands
Email address: davisilva15@gmail.com