arXiv:2603.19490v1 [cs.CC] 19 Mar 2026

乘积分布下不相交性问题的通信复杂度

Zach Hunter[1], Aleksa Milojevi´c[1], Benny Sudakov[1], 和 Istvan Tomon [2]

Abstract. 确定不相交性问题的随机化（或分布式）通信复杂度，是通信复杂度领域的核心问题。该问题源于 Babai、Frankl 和 Simon 在20世纪80年代的奠基性工作，并在 Kalyanasundaram-Schnitger 和 Razborov 于 1992 年的著名研究中达到高潮。然而，直到 Bottesch 、 Gavinsky 和 Klauck 最近的工作解决这一问题之前，乘积分布下如何获得紧确界的问题一直悬而未决。在这篇短文中，我们重新探讨这一经典问题，并给出一个简洁精炼的最佳界证明，同时改进了界对误差参数的定量依赖关系。

我们的方法基于一个简单的组合引理，它本身也可能具有独立价值：如果从两个分布中独立抽取的两个集合以不可忽略的概率不相交，那么就可以提取出两个测度相当大的子族，使它们彼此完全不相交（等价地说，就是得到一个关于不相交性的大单色矩形）。

1. Introduction

通信复杂度研究两个彼此分离的参与方为了计算关于双方共同输入的函数，必须交换多少信息。形式化地，给定有限集 $\mathcal{A}$， $\mathcal{B}$ 以及一个布尔函数f：$\mathcal{A}\times\mathcal{B}\to\{0,1\}$。Alice 获得 $A\in\mathcal{A}$，Bob 获得 $B\in\mathcal{B}$，双方都看不到对方的输入。双方根据一个（确定性）协议，通过双向信道交换比特进行通信，最后由其中一方输出一个值，目标是使其等于 $f(A,\ B)$。协议的代价是对所有输入 $(A,\ B)$ 取最大值，即输入 $(A,\ B)$ 时交换的比特总数；而 f 的确定性通信复杂度记为 $\mathrm{CC}(f)$，是对每个输入都正确的协议的最小代价。背景介绍、研究动机和应用可参见 [7, 10] 。

研究最多的函数之一是集合不相交性。给定一个全集 $[n]$，令 $\mathcal{A}=\mathcal{B}=2^{[n]}$ 和

$$\mathrm{DISJ}_n(A,B)=1 \iff A\cap B=\varnothing.$$

不相交性是该模型中难以计算的函数的典型例子，也是许多归约的起点。其确定性通信复杂度为 $\Theta(n)$，下界可由秩方法得到：注意到 $\mathrm{CC}(\mathrm{DISJ}_n)\geq\log_2\mathrm{rank}M$，其中$M_{A,\ B}=\mathrm{DISJ}_n(A,\ B)$是一个满秩矩阵，它编码了函数 DISJ_n （详情见 [11] ）。

本文研究分布模型下的不相交性。固定一个定义在 $\mathcal{A}\times\mathcal{B}$ 上的分布 μ，以及一个误差参数 $\varepsilon\in(0,\ 1/2)$。在该模型中，输入$(A,\ B)$按照 μ抽样，而 $D^{\mu}_{\varepsilon}(f)$表示确定性协议的最小通信代价，该协议能在占输入对比例为$(1-\varepsilon)$的样本上正确计算 f （相对于 μ而言）。请注意，协议可以依赖于 μ。

根据姚氏极小极大原理，$\max_{\mu}D^{\mu}_{\varepsilon}(f)$等于 f 在误差为 ε时的公用随机数通信复杂度。因此，确定不相交性的分布式复杂度曾是一个备受关注的问题，直到 Kalyanasundaram 和 Schnitger [6] 以及 Razborov[9] 在20世纪90年代著名地解决了这一问题。Razborov 通过构造一个困难分布 μ ，证明了不相交性的随机化通信复杂度为 $\Omega(n)$；对于该分布， $D^{\mu}_{1/3}(\mathrm{DISJ}_n)=\Theta(n)$。因此，在完全一般的分布式／随机化设定下，不相交性仍然具有线性难度。

[1]瑞士苏黎世联邦理工学院数学系，瑞士。电子邮箱：{zach.hunter, aleksa.milojevic, benjamin.sudakov}@math.ethz.ch。研究部分受瑞士国家科学基金会项目 200021-228014 资助。[2]于默奥大学， $e\text{-}mail$: istvantomon@gmail.com。研究部分受瑞典研究理事会项目 VR 2023-03375 资助。

自然可以考虑受限情形，即 $\mu = \mu_A \times \mu_B$ 为乘积分布。这对应于独立地抽取集合 $A \sim \mu_A$, $B \sim \mu_B$。沿袭 Babai、Frankl 和 Simon [1], 的经典工作，我们将函数 f 的强分布复杂度 定义为

$$\mathrm{SD}_\varepsilon(f) := \sup_{\mu=\mu_A\times\mu_B} D^\mu_\varepsilon(f).$$

值得注意的是，将输入限制为相互独立，会使不相交性问题的最坏情况复杂度从线性降至约 $\sqrt{n}$。更准确地说，Babai–Frankl–Simon [1] 于 1986年证明，对每个 $\varepsilon < 1/2$,

$$\Omega(\sqrt{n}) \le \mathrm{SD}_\varepsilon(\mathrm{DISJ}_n) \le O_\varepsilon(\sqrt{n}\log_2 n).$$

上界与下界之间的对数间隙持续了数十年，直到 Bottesch、Gavinsky 和 Klauck [2], 于 2015年的工作才将其弥合。他们证明了紧确下界 $\mathrm{SD}_\varepsilon(\mathrm{DISJ}_n) = \Theta_\varepsilon(\sqrt{n})$。

他们还在更一般的背景下研究了这个问题，通过根据分布的互信息 $I(A:B)$ 设定参数，在乘积分布情形与完全一般情形之间进行插值。回顾一下，随机输入 A 和 B 从边缘分布为 μ_A, μ_B 的分布 μ 中抽样得到，它们之间的 互信息为

$$I(A:B) = \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} \mu(a,b)\log_2 \frac{\mu(a,b)}{\mu_A(a)\mu_B(b)}.$$

更一般地，他们证明，若 μ 是一个互信息至多为 k的分布，则$D^\mu_\varepsilon(\mathrm{DISJ}_n) \le O(\sqrt{n(k+1)}/\varepsilon^2)$。这些结果的证明关键依赖于 H˚astad–Wigderson 协议 [4] ： 该协议只需 $O_\varepsilon(\ell)$ 位，就能判断任意全集中的 ℓ元集合是否不相交。

本文重新探讨这一经典问题，并给出 Bottesch、Gavinsky 和 Klauck [2]结果的简洁精炼证明。此外，该证明改进了对误差参数 ε的依赖关系。

定理1. 令 μ 为 $2^{[n]}\times 2^{[n]}$ 上的一个分布，并令 $(A,B) \sim \mu$ 满足 $I(A:B) \le k$。则

$$D^\mu_\varepsilon(\mathrm{DISJ}_n) \le O\left(\sqrt{\frac{n(k+1)}{\varepsilon}}\right).$$

有趣的是，Bottesch、Gavinsky 和 Klauck 给出了一个下界，表明定理1是紧的，至少在互信息不太大的情况下是如此。具体而言，对于每个 $1 \le k \le \varepsilon n$ ，他们构造了一个测度 μ ，其互信息至多为 k，且满足 $D^\mu_\varepsilon(\mathrm{DISJ}_n) \ge \Omega(\sqrt{n(k+1)}/\varepsilon)$，由此表明在这一范围内上述界是紧的。

当 $k = 0$ 时，这一特殊情形对应于乘积分布的情况；在这种情况下，我们可以对不相交性的强分布复杂度界给出一个非常简短且完全自包含的证明。

定理 2。对于任意 $\varepsilon < 1/2$,

$$\mathrm{SD}_\varepsilon(\mathrm{DISJ}_n) \le O\left(\sqrt{n\log_2(1/\varepsilon)}\right).$$

我们构造的协议是确定性的，并基于一个我们认为本身也值得关注的组合引理。粗略地说，该引理指出：如果两个独立集合 $A \sim \mu_A$ 和 $B \sim \mu_B$ 以不可忽略的概率不相交，那么就可以提取出两个测度足够大的子族 $\mathcal{A}$, $\mathcal{B} \subseteq 2^{[n]}$ ，使得对于每一对 $(A,\ B) \in \mathcal{A}\times\mathcal{B}$ 都成立（等价地说，DISJ_n 在一个较大的组合矩形上恒等于 1）。利用该引理及 Nisan-Wigderson 归约的一个变体，我们可以借助较大的单色组合矩形，构造一个用于不相交性的高效协议（详情见第 3 节）。

这个引理的证明是自成一体的，其灵感来自我们近期关于具有许多不相交对的集合族的 Singer–Sudan 猜想的研究 [3]。具体而言，受到

寻找单色矩形以及对数秩猜想的启发，Singer 和 Sudan [12] 提出了以下问题：给定具有如下性质的集合族 $\mathcal{A}$，$\mathcal{B} \subseteq 2^{[n]}$至少有 $\varepsilon|\mathcal{A}||\mathcal{B}|$对 $(A,\ B) \in \mathcal{A}\times\mathcal{B}$彼此不相交，那么我们能找到多大的集合族 $\mathcal{F} \subseteq \mathcal{A}$，$\mathcal{G} \subseteq \mathcal{B}$，使得其中每一对集合 $(F,\ G) \in \mathcal{F} \times \mathcal{G}$都彼此不相交？在 [3], 中，我们证明可以取 $|\mathcal{F}||\mathcal{G}| \geq 2^{-O(\sqrt{n \log_2 1/\varepsilon})}|\mathcal{A}||\mathcal{B}|$。我们的主要引理是这一结论的推广，其中集合族被集合上的分布所取代。

论文结构。在第 2 节中，我们陈述并证明主要引理。接着，在第 3 节中，我们利用它证明定理 2。最后，在第 4 节中，我们利用定理 2 推导出定理1。

2. Distributions with many disjoint sets

引理 3. 设 μ_A, μ_B 是 $2^{[n]}$上的概率分布，并设 $\varepsilon \in (2^{-n}, 1/2)$ 是一个参数，使得

$$\mathbb{P}_{A\sim\mu_A, B\sim\mu_B}[A \cap B = \varnothing] \geq \varepsilon.$$

那么，存在集合族 $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$，使得 $\mu_A(\mathcal{A})\ \mu_B(\mathcal{B})\ \geq 2^{-O(\sqrt{n \log_2 1/\varepsilon})}$，且所有的数对 $(A, B) \in \mathcal{A} \times \mathcal{B}$ 都互不相交。

证明。证明的主要思路是根据 μ_A 随机选取 $\ell = \lfloor \sqrt{n/\log_2 \varepsilon^{-1}} \rfloor$ 个集合 $A_1, \ldots, A_\ell$，并考虑它们的并集 $U = \bigcup_{i=1}^{\ell} A_i$。然后令 $\mathcal{A}$ 为 $2^{[n]}$ 中包含于 U 的集合族，并令 $\mathcal{B}$ 为 $2^{[n]}$ 中包含于 $[n]\backslash U$ 的集合族。此构造保证，对于每一对集合 $(A, B) \in \mathcal{A} \times \mathcal{B}$，$A$ 和 B 都互不相交，因此我们的主要目标是证明，以正概率，$\mu_A(\mathcal{A})$ 和 $\mu_B(\mathcal{B})$ 都足够大。

首先，我们证明

$$\mathbb{P}[\mu_A(\mathcal{A}) \geq 2^{-3n/\ell}] \geq 1 - 2^{-2n}.$$

如果集合 U满足 $\mathbb{P}_{A\sim\mu_A}[A \subseteq U] = \mu_A(2^U) \leq 2^{-3n/\ell}$，则称其为坏的。那么，从 μ_A 中随机选取的 ℓ个集合包含于一个坏集合 U 中的概率至多为 $(2^{-3n/\ell})^\ell = 2^{-3n}$。因此，对至多 2^n 个坏集合应用并集界，可将 $U = \bigcup_{i=1}^{\ell} A_i$ 为坏集合的概率上界估计为 2^{-2n}。所以，以至少 $1 - 2^{-2n}$的概率，有 $\mu_A(\mathcal{A}) \geq 2^{-3n/\ell}$。

接下来，我们证明

$$\mathbb{P}[\mu_B(\mathcal{B}) \geq \varepsilon^\ell/2] > 2^{-n}.$$

为此，我们将通过对所有集合 $B \subseteq [n]$ 求和来计算 $\mathcal{B}$ 的期望测度：求出 B 落入 $\mathcal{B}$ 中的概率，再乘以 B 在 μ_B下的权重。注意到

$$\mathbb{P}\big[B \in \mathcal{B}\big] = \mathbb{P}\big[B \cap \bigcup_{i=1}^{\ell} A_i = \varnothing\big] = \mathbb{P}\big[A_i \subseteq [n]\backslash B \text{ for all } i \in \{1, \ldots, \ell\}\big] = \mu_A(2^{[n]\backslash B})^\ell,$$

我们得到

$$\mathbb{E}\big[\mu_B(\mathcal{B})\big] = \sum_{B\subseteq[n]} \mu_B(B) \cdot \mathbb{P}\big[B \in \mathcal{B}\big] = \sum_{B\subseteq[n]} \mu_B(B) \cdot \mu_A(2^{[n]\backslash B})^\ell.$$

将詹森不等式应用于凸函数 $x \mapsto x^\ell$ 和测度 μ_B，我们得到

$$\mathbb{E}\big[\mu_B(\mathcal{B})\big] = \sum_{B\subseteq[n]} \mu_B(B) \cdot \mu_A(2^{[n]\backslash B})^\ell \geq \Big(\sum_{B\subseteq[n]} \mu_B(B) \cdot \mu_A(2^{[n]\backslash B})\Big)^\ell \geq \varepsilon^\ell.$$

最后一个不等式由以下事实得出：$\sum_{B\subseteq[n]} \mu_B(B)\ \cdot \mu_A(2^{[n]\backslash B}) = \mathbb{P}_{A\sim\mu_A,\ B\sim\mu_B}[A \cap B = \varnothing] \geq \varepsilon$。利用 $0 \leq \mu_B(\mathcal{B})\ \leq 1$，我们可以写成

$$\varepsilon^\ell \leq \mathbb{E}\big[\mu_B(\mathcal{B})\big] \leq 1 \cdot \mathbb{P}\Big[\mu_B(\mathcal{B}) \geq \varepsilon^\ell/2\Big] + \frac{\varepsilon^\ell}{2} \cdot \mathbb{P}\Big[\mu_B(\mathcal{B}) < \varepsilon^\ell/2\Big] < \mathbb{P}\Big[\mu_B(\mathcal{B}) \geq \varepsilon^\ell/2\Big] + \frac{\varepsilon^\ell}{2}. \tag{1}$$

注意到 $\varepsilon^\ell/2 \geq \varepsilon^{\sqrt{n/\log^2 \varepsilon^{-1}/2}} = 2^{-\sqrt{n \log_2 \varepsilon^{-1} - 1}} > 2^{-n-1}$，其中最后一个不等式由我们的假设 $\varepsilon > 2^{-n}$ 得出。因此，比较公式 1 的两边可得 $\mathbb{P}[\mu_B(\mathcal{B})\ \geq \varepsilon^\ell/2] > \varepsilon^\ell/2 \geq 2^{-n-1}$。

由于 $\mathbb{P}[\mu_A(\mathcal{A}) \geq 2^{-3n/\ell}] + \mathbb{P}[\mu_B(\mathcal{B}) \geq \varepsilon^\ell/2] > 1$，存在一种 $A_1, \ldots, A_\ell$ 的选择，使得

$$\mu_A(\mathcal{A}) \geq 2^{-3n/\ell} \geq 2^{-O(\sqrt{n \log_2 \varepsilon^{-1}})},$$
$$\mu_B(\mathcal{B}) \geq \varepsilon^\ell/2 = 2^{-\ell \log_2 \varepsilon^{-1} - 1} \geq 2^{-O(\sqrt{n \log_2 \varepsilon^{-1}})},$$

其中我们使用了 $\ell = \lfloor \sqrt{n / \log_2 \varepsilon^{-1}} \rfloor$。证明完毕。 □

3. Distributional Complexity for Product Distributions

本节将使用 Nisan-Wigderson 归约 [8] 的一个变体来证明定理 2，即证明 $D_\varepsilon^{\mu_A \times \mu_B}(\mathrm{DISJ}_n) \leq O(\sqrt{n \log_2 1/\varepsilon})$。。我们称测度 ν 为近乎空，如果 $\mathbb{P}_{(A,B)\sim\nu}[A \cap B = \varnothing] \leq \varepsilon/2$。此外，如果所有对 $(F, G) \in \mathcal{F} \times \mathcal{G}$ 均不相交，则称组合矩形 $R = \mathcal{F} \times \mathcal{G} \subseteq 2^{[n]} \times 2^{[n]}$ 为满 。

定理 2 的证明。 我们给出一个高效协议，在给定测度 μ 下，该协议能对占全部输入 $(1-\varepsilon)$ 的比例的输入正确计算 DISJ_n。设 $A, B \in 2^{[n]}$ 分别为 Alice 和 Bob 已知的输入。

我们首先定义一个子协议，其目标如下。假设存在一个集合 $X \subset [n]$，Alice 和 Bob 都知道该集合，并且 $A \setminus X$ 与 $B \setminus X$ 不相交，或者等价地，$A \cap B = \varnothing \Leftrightarrow (A \cap X) \cap (B \cap X) = \varnothing$。子协议结束时，Alice 和 Bob 要么宣布 A 与 B 是否不相交，要么就某个大小为 $|Y| \leq |X|/2$ 的集合 Y 达成一致，使得 $A \setminus Y$ 与 $B \setminus Y$ 不相交，从而有效地将全集的大小减半。此外，若 T 是子协议输出关于 $A \cap B = \varnothing$ 的答案的输入对集合，而 S 是答案错误的输入对集合，我们将保证 $\mu(S) \leq \frac{\varepsilon}{2}\left(\mu(T) + \frac{1}{|X|}\right)$。

我们声称，这样的子协议可以用 $O(\sqrt{|X| \log_2 1/\varepsilon})$ 个通信比特来执行。这个结论蕴含了我们的定理，因为我们可以从 $X = [n]$ 开始，不断执行该子协议，直到全集的大小缩减至 $|X| \leq 4 \log_2 1/\varepsilon$，并在每一步将全集的大小减半。如果到这一步 Alice 和 Bob 仍未确定 $A \cap B = \varnothing$，Alice 可以用 $|X| = O(\sqrt{|X| \log_2 1/\varepsilon})$ 个通信比特传送 $A \cap X$ 中的所有元素，从而解决该问题。该协议中通信的比特总数至多为

$$\sum_{i=0}^{\lfloor \log_2 n \rfloor - 1} O\left(\sqrt{\frac{n}{2^i} \log_2 1/\varepsilon}\right) = O(\sqrt{n \log_2 1/\varepsilon}).$$

此外，协议给出错误答案的集合 S 的测度至多为

$$\frac{\varepsilon}{2}\left(1 + \sum_{i=0}^{\lfloor \log_2 n \rfloor - 1} \frac{1}{n/2^i}\right) \leq \frac{\varepsilon}{2} \cdot 2 = \varepsilon.$$

子协议规范： 令 M_0 为 X 的子集所对应的不相交矩阵，也就是说，对于任意 A_0，$B_0 \subseteq X$，都有 $(M_0)_{A_0, B_0} = 1 \iff A_0 \cap B_0 = \varnothing$。此外，令 $A' = X \cap A$、$B' = X \cap B$、$\mathcal{A}_0 = \mathcal{B}_0 = 2^X$，并令 $\nu = \nu_A \times \nu_B$ 为测度 μ 在 $2^X \times 2^X$ 上的重新缩放，使得 ν 是一个概率测度。

如果 $\nu_A(A') \leq \frac{\varepsilon}{2^{2|X|}}$，那么 Alice 会将此情况告知 Bob；或者，如果 $\nu_B(B') \leq \frac{\varepsilon}{2^{2|X|}}$，Bob 会用一位比特将此情况告知 Alice。如果这两种情况中任意一种发生，Alice 和 Bob 可以断定这些集合不相交。此步骤中他们出错的输入测度至多为 $\frac{\varepsilon}{2^{|X|-1}} \leq \frac{\varepsilon}{2|X|}$，因为满足 $\nu_A(A') \leq \frac{\varepsilon}{2^{2|X|}}$ 或 $\nu_B(B') \leq \frac{\varepsilon}{2^{2|X|}}$ 的配对 (A', B') 的总测度至多为 $2 \cdot 2^{|X|} \cdot \varepsilon/2^{2|X|} = \varepsilon/2^{|X|-1}$。

如果 ν 是近乎空的，那么 Alice 和 Bob 就断定他们的集合不相交，并结束协议。这样，他们的答案对于 $\mathcal{A}_0 \times \mathcal{B}_0$ 中至少 $(1 - \varepsilon/2)$ 测度的输入是正确的，因此满足 $\mu(S) \leq \frac{\varepsilon}{2}\mu(T) + \frac{\varepsilon}{2|X|}$。

如果 ν 不是近乎空的，令 $R = \mathcal{F} \times \mathcal{G}$ 为 M_0 中测度最大的完整矩形。由于 $\varepsilon/2 \geq 2^{-|X|}$（这是由 X 上的停止条件导致的），根据引理 3，我们有 $\nu(R) \geq 2^{-O(\sqrt{|X| \log_2 1/\varepsilon})}$，这

还意味着 $\nu_A(\mathcal{F})\ \nu_B(\mathcal{G}) \geq 2^{-O(\sqrt{|X|\log_2 1/\varepsilon})}$。由于所有集合 $F \in \mathcal{F}$都与所有集合 $G \in \mathcal{G}$不相交，它们的并集 $U = \bigcup_{F\in\mathcal{F}} F$和 $V = \bigcup_{G\in\mathcal{G}} G$也不相交。因此，$|U| + |V| \leq |X|$，所以 U, V中的一个大小至多为 $|X|/2$，不妨设为 U（否则，我们可以交换双方的角色）。如果 $A' \in \mathcal{F}$，那么 $A' \subseteq U$，因此 $A' \setminus U$和 $B' \setminus U$不相交。在这种情况下，Alice 和 Bob 令 $Y := U$，并停止该子协议。如果 $A' \notin \mathcal{F}$，则定义 $\mathcal{A}_1 := \mathcal{A}_0 \setminus \mathcal{F}$以及 $\mathcal{B}_1 := \mathcal{B}_0$。

可以迭代上述过程，从而定义子集集合族 $\mathcal{A}_0 \supseteq \mathcal{A}_1 \supseteq \cdots \supseteq \mathcal{A}_t$和 $\mathcal{B}_0 \supseteq \mathcal{B}_1 \supseteq \cdots \supseteq \mathcal{B}_t$。注意，对于每个 $1 \leq i \leq t$，我们有

$$\nu(\mathcal{A}_i \times \mathcal{B}_i) \leq \left(1 - 2^{-O(\sqrt{|X|\log_2 1/\varepsilon})}\right)\nu(\mathcal{A}_{i-1} \times \mathcal{B}_{i-1}).$$

因此，如果子协议运行 t 步，我们有

$$\nu(\mathcal{A}_t \times \mathcal{B}_t) \leq \left(1 - 2^{-O(\sqrt{|X|\log_2 1/\varepsilon})}\right)^t \leq \exp(-t2^{-O(\sqrt{|X|\log_2 1/\varepsilon})}).$$

令 $t := 2^{C\sqrt{|X|\log_2 1/\varepsilon}}(|X| + \log_2 1/\varepsilon) = 2^{O(\sqrt{|X|\log_2 1/\varepsilon})}$，其中 C为某个足够大的常数，则右侧至多为 $2^{-4|X|-2\log_2 1/\varepsilon} \leq \frac{\varepsilon^2}{2^{4|X|}}$。但这是不可能的，因为此时 $\nu(A' \times B') \leq \frac{\varepsilon^2}{2^{4|X|}}$，这意味着 $\nu_A(A') \leq \frac{\varepsilon}{2^{2|X|}}$ 或 $\nu_B(B') \leq \frac{\varepsilon}{2^{2|X|}}$。因此，该子协议至多运行 t 步，并定义了一棵具有 $O(t)$ 个叶子的协议树。

众所周知（例如参见 [7] 第 2 章引理 2.8），每棵协议树都可以重新平衡为一棵深度关于叶节点数量呈对数级的决策树。这意味着可以用 $O(\log_2 t) = O(\sqrt{|X|\log_2 1/\varepsilon})$ 通信比特执行该子协议，从而验证我们的论断。□

4. Distributions with Mutual Information

接下来，我们从定理 2 推导出定理1。Bottesch、Gavinsky 和 Klauck 指出，Jain、Radhakrishnan 和 Sen 的经典子态定理 [5]可用于将独立情形下的界转化为互信息有界情形下的界。不过，这里的关键在于，定理 2 的界对 ε 的依赖性非常好，因此我们可以应用这一归约，而不会造成太大的损失。

为陈述子态定理，我们先回顾以下两个定义。若 μ，ν 是定义在 $\mathcal{A} \times \mathcal{B}$ 上的两个概率分布，$\|\mu - \nu\|_{\mathrm{TV}}$ 表示分布 μ 与ν之间的全变差距离，其定义为 $\|\mu - \nu\|_{\mathrm{TV}} = \max_{S\subseteq\mathcal{A}\times\mathcal{B}} |\mu(S) - \nu(S)|$。此外，若 μ 的边缘分布为 μ_A，μ_B，我们定义 $I_\infty(\mu) = \max^{(a,b)\in\mathcal{A}\times\mathcal{B}} \log_2 \frac{\mu_{(a,\ b)}}{\mu_A(a)\mu_B(b)}$。

定理4 ([5])**.** 设 $\varepsilon > 0$ 为一个正参数，$\mathcal{A}$、$\mathcal{B}$ 为有限集。对于定义在 $\mathcal{A} \times \mathcal{B}$上的任意分布$\mu$，若其互信息不超过 k，则存在另一个分布 ν，使得$\|\mu - \nu\|_{\mathrm{TV}} \leq \varepsilon$ 且 $I_\infty(\nu) \leq 4(k+1)/\varepsilon$。

利用该定理，我们给出定理1的证明。

定理1的证明。 考虑测度 ν，使得 $\|\mu - \nu\|_{\mathrm{TV}} \leq \varepsilon/2$ 且 $I_\infty(\nu) \leq 8(k+1)/\varepsilon$；根据定理4，该测度存在。若 ν 的边缘分布分别为 ν_A 和 ν_B，则考虑乘积分布 $\nu_A \times \nu_B$，并使用定理2的证明中的协议，误差参数为$\varepsilon' = \varepsilon 2^{-8(k+1)/\varepsilon - 1}$，分布为 $\nu_A \times \nu_B$。

该协议的代价为 $O(\sqrt{n\log_2 1/\varepsilon'}) = O(\sqrt{n(k+1)/\varepsilon})$。此外，若 $S \subseteq 2^{[n]} \times 2^{[n]}$ 是该协议出错的输入对集合，则 S的测度至多为 $(\nu_A \times \nu_B)(S) \leq \varepsilon'$。由于$I_\infty(\nu) \leq 8(k+1)/\varepsilon$，这意味着对于任意一对 $(A, B) \in \mathcal{A}\times\mathcal{B}$，都有 $\nu(A,\ B) \leq 2^{8(k+1)/\varepsilon}\nu_A(A)\nu_B(B)$，因此 $\nu(S) \leq 2^{8(k+1)/\varepsilon}\varepsilon' \leq \varepsilon/2$。

最后，由于 $\|\mu - \nu\|_{\mathrm{TV}} \leq \varepsilon/2$，我们有 $|\mu(S) - \nu(S)| \leq \varepsilon/2$，从而可以得出 $\mu(S) \leq \varepsilon$。因此，我们得到一个复杂度至多为 $O(\sqrt{n(k+1)/\varepsilon})$ 的协议，其在至多 ε比例的输入上出错，证明完毕。□

致谢。作者感谢 Avi Wigderson 提供的宝贵见解，并将 Babai、Frankl 和 Simon 的论文介绍给他们。

References

[1] L. Babai、P. Frankl 和 J. Simon, *Complexity classes in communication complexity theory (preliminary version)*, 载于 *Proceedings of the 27th Annual Symposium on Foundations of Computer Science (FOCS 1986)*, 1986年, 第 337–347 页。[2] R. C. Bottesch、D. Gavinsky 和 H. Klauck, *Correlation in hard distributions in communication complexity Ap-*, 载于, , *proximation Randomization and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM*, LIPIcs, 第 卷, 第 页。[3] Z. *2015)* 40, 2015, 544 572 Hunter、A. Milojevi *Disjoint pairs in set systems and combinatorics of low rank* c、B. *matrices* 2411 13510 Sudakov 和 I. Tomon, , arXiv: 。。[4] J. H 。 *The randomized communication complexity of set disjointness Theory of Computing* astad 和 A. Wigderson, , (), –。 **3** 2007 211 219 [5] R. Jain、J. Radhakrishnan 和 P. Sen, *Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states Proceedings of the 43rd Annual Symposium on Foundations*, 载于, , 第 – 页。[6] B. *of Computer Science (FOCS 2002)* 2002 429 438 Kalyanasundaram 和 G. Schnitger, *The probabilistic communication complexity of set intersection. SIAM J.* (), –。[7] E. Kushilevitz 和 *Discrete Math.* **5** 1992 545 557 N. Nisan, *Communication Complexity.* 剑桥大学出版社, 1997年。[8] N. Nisan 和 A. Wigderson, *On rank vs. communication complexity, Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS 1994)* 1994 831 836 , 第 – 页。[9] A. A. Razborov, *On the distributional complexity of disjointness Theoretical Computer Science* **106** 1992 2, , (), 第 – 期。[10] A. Rao 和 385 390 A. Yehudayoff, *Communication Complexity and Applications*, 剑桥大学出版社, 2020年。[11] A. A. Sherstov, *Communication complexity theory: thirty-five years of set disjointness Mathematical Foundations*, 载于, , 《计算机科学讲义》, *of Computer Science 2014 (MFCS 2014) Part I* 8634, 2014, 第 卷。施普林格, 第 24–43 页。[12] N. Singer 和 M. Sudan, *Point-hyperplane incidence geometry and the log-rank conjecture ACM Transactions on*, () 。 *Computation Theory* **14** 2022