arXiv:2305.04363v1 [cs.CC] 7 May 2023

对称函数的采样与认证

Yuval Filmus[1], Itai Leigh[2], Artur Riazanov[3], 和 Dmitry Sokolov[4]

[1]Technion — 以色列理工学院，yuvalfi@cs.technion.ac.il[2]特拉维夫大学；哥本哈根大学；itai.leigh@mail.huji.ac.il[3]洛桑联邦理工学院，tunyash@gmail.com[4]洛桑联邦理工学院，sokolov.dmt@gmail.com

2022年5月

摘要

如果电路 C 对均匀输入上的输出与 X 之间的统计距离为 ε，则称其以误差 ε 采样分布 X。我们研究决策森林采样 U_k^n 的难度，其中 U_k^n 表示由汉明重量为 k 的 n位字符串组成的集合；决策森林是指每个输出位都由输入的一个决策树计算得到。对于每个 k，都存在一个深度为 $(\log\ n)$ 的O决策森林，以逆多项式误差 [采样 U_k^n Vio12b; Czu15]。我们证明，对于每个 $\varepsilon > 0$，都存在 τ，使得当决策深度为 $\tau \log(n/k)/\log\log(n/k)$ 时，采样 U_k^n的误差至少为 $1-\varepsilon$。我们的结果基于近期的稳健向日葵引理 [Alw+21; Rao19]。

我们的第二个结果涉及将一个 n位字符串集合与一个 d-局部 电路的像相匹配，即每个输出位至多依赖 d 个输入位。我们研究所有汉明重量至少为 $n/2$的 n位字符串组成的集合。我们将此前已知的局部性下界从 $\Omega(\log^* n)$ [Bey+13] 提高到 $\Omega(\sqrt{\log n})$，与最佳上界 $O(\log^2 n)$ 之间仅相差四次方因子。

1 引言

研究采样的困难性这一方向由论文 [Vio12b]，提出，并由此推动了一系列活跃的研究 [Vio12a; LV11; BIL12; Vio14; Vio16; Vio20; GW20; Vio21; Bog+22; CGZ22]。基本设定如下：给定无限供应的相互独立的均匀随机比特作为输入，我们的目标是设计一个具有多个输出位的电路，使其输出与给定目标分布之间的统计距离足够小。采样电路的构造已应用于密码学 [IN96; Bog+16] 和算法 [Hag91]。采样困难性结果[1] 在启发和改进双源提取器构造方面发挥了重要作用 [Vio14; CZ16; CS16] （详见 [Vio20] 的讨论），并给出了简洁数据结构的下界 [Vio12b; Vio20; Vio21]。

采样难度结果比计算难度结果更具挑战性。例如，自 Smolensky 的经典工作以来，人们已经知道 [Smo87] 奇偶校验需要指数级数量的门才能由一个$\mathbf{AC}^0[3]$ 电路计算，但对于电路类$\mathbf{AC}^0[p]$，对于任意 p，目前尚无已知的困难分布；而且$\mathbf{AC}^0$ 需要指数级数量的门才能计算奇偶校验[H˚, 同时87], 奇偶校验为 0 的随机向量可以由一个$\mathbf{NC}^0$ 电路采样。此外，这种分布还可以通过一个 2-局部电路采样

[1]也就是说，显 证明某一类中的任何电路所产生的分布都与目标相差甚远。

其含义是，每个输出位仅依赖于两个输入位 [Bab87; Kil88]。一种非常简单的映射即可实现这一点：$(x_1,\dots,x_n) \mapsto (x_1, x_1 \oplus x_2, x_2 \oplus x_3,\dots,x_{n-1} \oplus x_n, x_n)$。一个更一般且引人注目的事实是，$\mathsf{AC}^0$ 可以采样随机排列，进而采样所有形如 $(\boldsymbol{X}, f(\boldsymbol{X}))$ 的分布，其中 $\boldsymbol{X}$ 在 $\{0,1\}^n$ 上均匀分布，且 f 是对称的，即其取值仅取决于输入 [的汉明重量]Vio12b。

我们猜想，NC^0 在采样对称分布方面的能力，本质上受限于上述奇偶校验示例。注意，当且仅当一个函数可由 NC^0 电路计算时，它才是 $O(1)$-局部的，也就是说，它的每个输出位至多依赖常数个输入位。为简单起见，我们只关注支撑集对称的均匀分布：令 $\boldsymbol{U}_S^n$ 为 $\{0,1\}^n$ 中汉明重量属于集合 $S \subseteq \{0,\dots,n\}$ 的字符串上的均匀分布。

猜想 1.1. 对每个 $d \in \mathbb{N}$，当 n 足够大时，若 $\boldsymbol{X}$ 可由一个 d-局部函数采样，且与某个 $\boldsymbol{U}_S^n$ 的距离为 ε，其中 $S \subseteq \{0,\dots,n\}$，那么 $\boldsymbol{X}$ 与 $\boldsymbol{U}_T^n$ 的距离为 $O(\varepsilon)$，其中 T 属于以下之一：$\{0\}$、$\{n\}$、$\{0,n\}$、$\{0,2,4,\dots\}$、$\{1,3,5,\dots\}$、$[n]$。

我们关于切片采样的结果（即定理 1.2）表明，对于所有仅包含较小数值的集合 S，这一猜想均成立，其中较小数值的含义是 $\max_{x\in S} x = o(n)$。

量子分离。猜想 1.1 的一个更强版本将确定这样一个集合族 S，使得每个可由 NC^0 采样的分布都与 $\boldsymbol{U}_S^n$ 相距 $1 - o(1)$。存在一个集合 S，对于该集合，这意味着在采样方面 NC^0 与 QNC^0 之间存在分离。这源于近期在[WP23]中取得的部分分离结果：他们证明，存在一个对称函数 f，使得 $(\boldsymbol{X}, f(\boldsymbol{X}))$ 的均匀 $\boldsymbol{X} \sim \{0,1\}^n$ 可以通过一个 QNC^0 电路进行采样。然而，请注意，如果一个可由 NC^0 采样的分布 $\boldsymbol{Y}$ 与 $(\boldsymbol{X}, f(\boldsymbol{X}))$ 的距离为 η，那么由于 [WP23] 中使用的函数 f 几乎是平衡的（即 $|f^{-1}(0)| = (1 + o(1)) \cdot |f^{-1}(1)|$），因此 $\boldsymbol{Y}$ 的前 n 位与 $f^{-1}(0)$ 上的均匀分布之间的距离为 $1/2 + \eta + o(1)$。现在，如果 $f^{-1}(0)$ 上的均匀分布无法在距离 $1 - \Omega(1)$ 以内由 NC^0 采样，我们就得到了这一分离。猜想 1.1 意味着一个较弱的下界，即该距离的下界为一个常数，而不是一个趋近于 1 的函数；但对于分布采样而言，大多数已知的下界都具有非常强的距离保证。

局部证书。 还有一类有趣的集合，我们以及先前的结果也未能涵盖：$M_a = \{x \in \{0,\dots,n\} \mid x \text{ 模 } a = 0\}$，其中 a 是常数。不过，用于采样奇偶校验为 0 的向量的简单构造可以调整为匹配任意的支撑集 $\boldsymbol{U}_{M_a}^n$，即生成一个（不一定均匀的）分布，其支撑集恰好是汉明重量属于 M_a 的字符串集合。该构造见[Bey+13，命题 3。1]，其中它被表述为一个证明系统。其思路是将输入位视为一份证书，证明输出属于目标语言（在本例中，即所有 n 位字符串，其汉明重量可被 a 整除）。这种联系促使我们研究证明系统背景下的局部性。我们大幅简化并改进了 [Kre+16] 关于如下语言的局部性下界：汉明重量至少为 n 的 $n/2$ 位字符串（换言之，即多数函数的 1 输入）。

1.1 切片采样

该 k-切片是所有 n 位汉明重量为 k 的字符串组成的集合。我们用 k-切片上的均匀分布表示为 $\boldsymbol{U}_k^n$。简单计算（见第3.5节）表明，当 $\max \boldsymbol{U}_S^n$ 和 $\boldsymbol{U}_{\max S}^n$ 的统计距离接近于零时 $S = o(n)$。这意味着，为了证明其困难性

对于从次线性汉明重量字符串上的对称分布中进行采样，只需研究 $\boldsymbol{U}_k^n$ 在 $k = o(n)$ 时的情况即可。

尽管在猜想 1.1 的背景下，只需给出采样器局部性的下界即可，但对于切片而言，更自然的复杂度度量是决策深度。若函数$f\colon \{0,1\}^m \to \{0,1\}^n$ 可由深度为 d 的决策森林计算，即 f 的每个输出位都可由深度为 d 的决策树计算，也就是说，最多对输入进行 d 次自适应查询（相较之下， d-局部函数由 d 次非自适应查询计算）。Viola [Vio12b 的引理 6.4]表明，可以从深度为 d 的交换网络得到针对 $\boldsymbol{U}_k^n$ 的决策深度为 d 的采样器。Czumaj [Czu15 的定理 3.7] 证明了存在此类交换网络，其深度为 $d = O(\log n)$。下面的定理表明，对于 $k = o(n)$，这一构造几乎是紧的。

定理 1.2. 假设 $\boldsymbol{U}_k^n$ 可以通过决策深度 d 和变差距离误差 η 进行采样。

1. 对于每个 $\varepsilon > 0$，存在一个常数 τ，使得 $d \le \tau \log(n/k)/\log\log(n/k)$ 蕴含 $\eta \ge 1 - \varepsilon$。
2. 存在一个常数 τ，使得以下结论成立。对于每个 $\varepsilon \in (0,1)$，如果 $k \in [\log_2 n、 2^{\log^{1-\varepsilon} n}]$ 且 $d \le \tau \log^{\varepsilon} n$，那么 $\eta = 1 - n^{-\Omega(k)}$。对于 $k \in [1, \log_2 n)$ 和 $d \le \tau \log^{\varepsilon} n/\log\log n$，同样的界也成立。

此外，条目 1 对任意 $\boldsymbol{U}_S^n$（最大值为$_{x\in S} x = k$ 或最小值为$_{x\in S} x = n - k$）均成立。

我们证明定理 1.2 时的第一个关键观察是，只需证明 $k = 1$ 的情形，即：

定理 1.3。存在一个常数 $\tau > 0$，使得任何以决策深度 $\tau \log n/\log\log n$ 采样的分布都与 $\boldsymbol{U}_1^n$ 相差 $(1 - n^{-\Omega(1)})$。

要理解这为何蕴含结论 1，注意，前 n/k 位的$\boldsymbol{U}_k^n$边缘分布与$1 - 1/e + o(1)$ 的距离至多为 $\boldsymbol{U}_1^{n/k}$。定理 1.3进而给出采样 $1/e - o$时的距离下界 $\boldsymbol{U}_k^n$，其深度为 $\tau \log(n/k)/\log\log(n/k)$。对于每个 $1 - \varepsilon$，距离下界 ε可通过推广定理 1.3 来达到，使其直接适用于形如“ $\Theta(n/k$位的 $\boldsymbol{U}_k^n$”的分布。

结论 2 可由定理 1.3 的另一种归约推出。假设我们有一个深度为 d 的采样器，用于 $\boldsymbol{U}_k^n$。利用这个采样器，我们可以如下构造一个深度为 kd 的采样器，用于 $\boldsymbol{U}\binom{n}{k}_1$。将每个输出位与 $[n]$ 中一个大小为 k 的唯一子集对应起来，并将输出中相应的k 位的合取赋给该输出位。容易看出，所得分布与 $\boldsymbol{U}\binom{n}{k}_1$ 的距离，与初始采样器的分布和 $\boldsymbol{U}_k^n$ 的距离相同。

证明定理1.3 的技巧

采样器的局部性是一个输出位所依赖的输入位的最大数量。局部性总是以决策深度为上界。作为热身，我们来讨论一个 $\Omega(\log\log n)$ 的局部性下界，用于采样 $\boldsymbol{U}_1^n$，其思路与 [Vio20，定理 3]相近。

局部性下界的主要思想是击中集与独立集之间的二分法：对于一个 d局部源，很容易看出，要么存在 $\tau 2^d$ 个相互独立的输出位，要么存在一个大小为 $\tau d 2^d$ 的输入位击中集，使得每个输出位都依赖于该集合中的某个位。在前一种情况下，我们可以证明，至少有两个独立比特的概率非常大

求值为 1，因为对于每个输出位，其取值为 1 的概率至少为 2^{-d}。[2] 在后一种情况下，通过对击中集中的比特逐一固定所有可能的取值，我们可以看到，我们的源是 $2^{\tau d 2^d}$ 个（$d-1$）-局部源的混合分布（凸组合）。如果我们证明（$d-1$）-局部源必定与（$1-\varepsilon$）目标分布相距至少 $1-\varepsilon 2^{\tau d 2^d}$，那么根据 [Vio20，推论 18]，我们的源与目标分布相距至少 $d=\delta$。取足够小的 n，令 δ，即可得到与目标分布距离的 $1-o(1)$ 下界。

为了将这个下界从 $\log\log n$ 提高到 $\log n/\log\log n$，我们引入了若干新思路。

单调化。 我们注意到，在某种意义上，只需处理每个位都是输入位上的一个 单调项 的源。直观原因是，输出位中取值为 1 的期望数量为 $2^{-d}\cdot n$[3]，因此存在一种赋值，使得这么多个输出位取值为 1。我们聚焦于这些输出位，并用与满足赋值对应的项替换它们，从而证明这些项中很可能至少有两项取值为 1。

向日葵。假设所有输出位都是单调项。对于 $i\in[n]$，令 N_i 为第 i 个输出位对应项中提及的输入所组成的集合。我们找到一个向日葵 $\mathcal{S}\subseteq\{N_1,\ldots,N_n\}$，也就是说，存在一个核 K，使得 $\mathcal{S}$ 中任意两个集合的交集都是 K。如果将核固定为 1，向日葵中的输出位就相互独立，因此，只要向日葵足够大，其中就很可能至少有两个输出位取值为 1。对于足够小的 $d=$，在~~任意~~ n 个集合中，总能找到一个大小为 $\Omega(\sqrt{\log n})$ 的向日葵。此外，我们可以用向日葵覆盖除 $o(n)$ 个输出位之外的所有输出位。于是得到如下二分情况：要么所有核都取值为 0，因此源很可能恒为零；要么至少有一个核取值为 1，从而对应向日葵中的输出位很可能至少有两个取值为 1。

使用鲁棒向日葵而非经典向日葵，我们得到决策深度的下界为 $\Omega(\log n/\log\log n)$。

1.2 局部证明系统

局部证明系统由 [Bey+13] 提出，并由 [Kre+16]进一步研究，其定义如下：语言 L 的局部证明系统是一个 $\mathbf{NC}^0$-电路族 C_n，满足 $L\cap\{0,1\}^n$ 恰好是 C_n所有可能输出的集合。若语言 L 具有一个 d-局部 证明系统，则对每个 n，都存在一个 d-局部函数，其像为 $L\cap\{0,1\}^n$。从采样的角度来看，可以这样理解：“采样器”必须与给定分布的支撑集完全一致，但我们不要求它匹配实际概率。

在这种设定下，困难性格局与分布采样设定有所不同。最显著的区别是，汉明重量可被p 整除的字符串所构成的语言始终具有一个 $O(1)$-局部证明系统；而我们猜想，除非 $p=1$ 或 $p=2$，否则从该语言上的均匀分布进行采样需要超常数局部性。

我们的主要贡献在于改进另一种对称语言的局部性下界： $\textsc{Maj}^{-1}(1):=\{x\in\{0,1\}^n\mid |x|\geq n/2\}$。此前，该语言的证明系统所能达到的最佳局部性下界为 $\Omega(\log^* n)$ [Kre+16]，其证明过程非常复杂，我们在附录 A 中详细展示了这一证明。我们简化了~~该~~证明，并将局部性下界提高到 $\Omega(\sqrt{\log n})$，这仅比当前最佳上界 $O(\log^2 n)$ 在多项式意义上更小。其核心思想是考虑输入位和输出位的局部性均有界的证明系统。对于这类证明系统，可以轻松推导出对局部性的非常强的要求，进而利用这些要求来计算数量

[2]需要注意的是，某些比特可能恒等于零，但这并不是真正的问题，因为这样的比特数量不会太多。[3]同样，这一点并非总是成立，因为存在恒等于零的输出。

输入影响输出的输入—输出位对的数量，由此得出输出局部性的下界。

1.3 交换网络

定理 1.2 背后的技术不适用于线性切片 $\boldsymbol{U}^n_{\alpha n}$。这是否意味着这类分布存在一种低深度采样器？目前我们唯一已知的采样器构造基于交换网络[Czu15，定理 3。7]。

深度为 d 的交换网络是一个分层图，包含 d 层，每层有 n 个节点。交换网络的边不会跨层相交，并且在每一层中，边构成一个部分匹配。交换网络定义了一个作用于 $[n]$ 的排列上的过程：我们从恒等排列开始，然后对每一层，都为该层匹配中的每条边抛一次硬币；如果硬币正面朝上，就交换该边的两个端点。

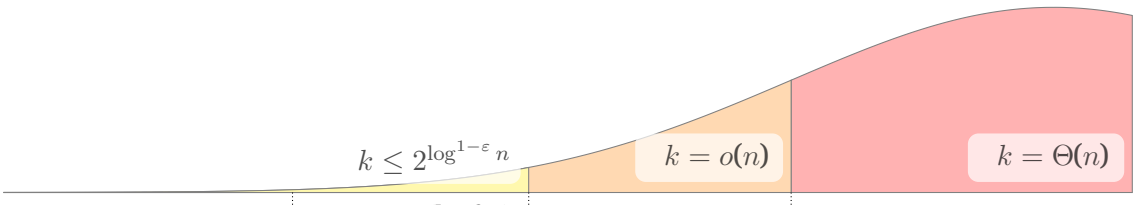
目前，生成一个接近所有排列上的均匀分布的分布的交换网络，其深度的最佳上界为 $O(\log^2 n)$ [Czu15]。如果交换网络只需打乱由 0 和 1 组成的序列，情况会更好：此时输入为 1^k0^{n-k}。 [Czu15，定理 3。7] 给出了这种情况下的 $O(\log n)$ 深度上界（该构造是随机化的）， [GTS14] 给出了一个显式的下界，用于生成 $\boldsymbol{U}^n_k$，其中$k \le \sqrt{n}$。

显而易见，要以非平凡距离内的精度采样 $\boldsymbol{U}^n_k$，交换网络的深度必须为 $\Omega(\log(n/k))$：深度为 d 的网络中，每个输入位在输出中的潜在位置最多有 2^d 个，因此，如果 $d = o(\log(n/k))$，该交换网络产生的分布中，只有 $o(n)$ 位取值为 1 的概率非零，而这与 $\boldsymbol{U}^n_k$ 相距 $(1 - o(1))$。

在第 5 节中，我们将证明，产生接近 $\boldsymbol{U}^n_{\alpha n}$ 的分布的交换网络，其深度必须为 $\Omega(\log \log n)$。我们将利用由深度为 d 的交换网络构造的采样器所具有的以下性质：第一，此类采样器的每个输入位最多会影响2^d 个输出位；第二，误差是单边的，即此类采样器绝不会输出 $\boldsymbol{U}^n_{\alpha n}$ 定义域之外的字符串。第二个性质凸显了它与局部证书的相似性；事实上，我们的下界证明也采用了类似思路。

1.4 进一步研究

我们关于使用决策森林进行切片采样的结果，结合 Viola [Vio12b]的结果，总结在图 1中。



定理 1.2	$d = \Omega(\log^{\varepsilon} n)$ $\Delta = 1 - n^{-\Omega(k)}$	$d = \tilde{\Omega}(\log(n/k))$ $\Delta = 1 - \delta$	
[Vio12b，定理 1。6]			$\Delta = 2^{-O(d)} - O(1/n)$

图 1：上表展示了定理 1.2 和 [Vio12b，定理 1。6]对采样 $\boldsymbol{U}^n_k$ 在不同 k下的含义。其上方的图展示了布尔立方体中相应比特串集合的大小。

以下是一些尚未解决的重要挑战：

- 在常数误差范围内，为线性汉明重量给出任意非平凡的决策深度（甚至是局部性！）下界。

- 将 $\boldsymbol{U}_1^n$ 的决策深度下界提升至 $\Omega(\log\ n)$。这将立即给出所有多项式 k 的紧致（相差一个常数因子）决策深度下界。

- 为 $(\boldsymbol{X}, f(\boldsymbol{X}))$ 给出任意局部性下界，其中 $\boldsymbol{X} \sim \{0,1\}^n$ 和 $f = [(x_1 + \cdots + x_n) \bmod p \geq p/2] \oplus x_1 \oplus \cdots \oplus x_n$，因为这将区分用于采样的量子与经典 $\mathbf{NC}^0$ 电路，并改进 [WP23]目前取得的部分分离。

- 对于任意 $k > 2$，找出采样一个汉明重量可被 k 整除的均匀向量的任意非平凡下界。这很可能有助于深入理解用于采样的 QNC^0 与 NC^0 之间的分离。

- 确定一种交换网络的最优深度，该网络用于采样均匀随机排列，或采样切片中的均匀向量。

2 符号与工具

我们用粗体字母表示随机变量，例如 $\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{b}, \boldsymbol{B}$。我们写作 $\boldsymbol{a} \sim A$，表示 $\boldsymbol{a}$ 服从分布 A，或者当 A 是一个集合时，表示其服从该集合上的均匀分布。对于 $x \in \{0,1\}^n$，我们用 $w(x) = |x|$：$= \{i \in [n] \mid x_i = 1\}$ 表示其汉明重量。我们记作 $e_i = 0^{i-1}10^{n-i} \in \{0,1\}^n$。对于字符串 $x \in \{0,1\}^n$ 和集合 $T \subseteq [n]$，我们写作$x_T := (x_i)_{i \in T} \in \{0,1\}^T$。我们用 $\boldsymbol{U}_k^n$表示汉明重量为 k 的向量的均匀分布。

对于定义域 $\mathcal{X}$相同的两个分布 S 和 T，统计距离定义为

$$\Delta(S,T) := \max_{A \subseteq \mathcal{X}} \left| \Pr_{\boldsymbol{x} \sim S}[\boldsymbol{x} \in A] - \Pr_{\boldsymbol{x} \sim T}[\boldsymbol{x} \in A] \right| = \frac{1}{2} \sum_{a \in \mathcal{X}} \left| \Pr_{\boldsymbol{x} \sim S}[\boldsymbol{x} = a] - \Pr_{\boldsymbol{x} \sim T}[\boldsymbol{x} = a] \right|.$$

两个随机变量之间的统计距离，定义为它们各自分布之间的统计距离。若分布 S与分布 T之间的距离满足 $\Delta(S,T) < \varepsilon$，则称分布 S与分布 T是 ε-接近的。否则，这两个分布是 ε-远的。

函数 $f\colon \{0,1\}^m \to \{0,1\}^n$ 若其每个输出位仅依赖于 d 个输入位，则称其为 d-局部函数。函数 $f\colon \{0,1\}^m \to \{0,1\}^n$ 若其每个输出位都能通过深度为 d 的决策树计算得出，即通过至多 d 次自适应输入位查询来确定，则称其决策深度为 d。如果函数是 d-局部的，那么其决策深度至多为 d。

我们将使用一种非常简单形式的 FKG 不等式：

定理 2.1 ([Har60; FKG71])。设 X 是定义在 $\{0,1\}^n$ 上的乘积分布（即，$\Pr[X = x] = \prod_{i \in [n]} \Pr[X_i = x_i]$）。设 A、$B \subseteq \{0,1\}^n$ 是两个单调事件（如果 $x \in A$ 且 $x_i \leq y_i$ 对所有 $i \in [n]$ 都成立，那么 $y \in A$，对于 B 也类似）。那么

$$\Pr_{\boldsymbol{x} \sim X}[\boldsymbol{x} \in A \cap B] \geq \Pr_{\boldsymbol{x} \sim X}[\boldsymbol{x} \in A] \cdot \Pr_{\boldsymbol{x} \sim X}[\boldsymbol{x} \in B].$$

2.1 鲁棒向日葵

本节讨论鲁棒向日葵。

定义 2.2（鲁棒向日葵）。设 $0 < \alpha, \beta < 1$ 为参数，设 $\mathcal{F}$ 为集合系统，并令 $K := \bigcap_{S \in \mathcal{F}} S$ 为 $\mathcal{F}$ 中所有集合的交集，我们称之为核。若集合族 $\mathcal{F}$ 是一个 (α, β)-鲁棒向日葵，则称其为鲁棒向日葵，如果

1. $K \notin \mathcal{F}$；
2. $\Pr_{\boldsymbol{R}}[\exists S \in \mathcal{F}\colon S \subseteq \boldsymbol{R} \cup K] \geq 1 - \beta$，其中全集中的每个元素都以概率 α独立地出现在 $\boldsymbol{R}$ 中。我们可以将此条件写成等价形式 $\Pr_{\mathbf{R}}[\exists S \in \mathcal{F}\colon \mathrm{R} \supseteq S \mid \mathrm{R} \supseteq K] \geq 1 - \beta$。

如果一个集合系统是一个具有空核的 (α, β)-鲁棒向日葵，则称其为 (α, β)-满足的。

正如 Rossman 所证明的，只要集合系统足够大，其中总会包含一个鲁棒向日葵 [Ros14]，后来的作者又对这一结果作了改进。

定理 2.3 ([Alw+21; BCW21; Rao19])。存在一个常数 $B > 0$，使得对于所有 $p, \varepsilon \in (0, 1/2]$，以下结论成立。设 $\mathcal{F}$是一个集合大小恰为 d的集合族，且满足$|\mathcal{F}| \geq (B \log(d/\varepsilon)/p)^d$。那么 $\mathcal{F}$包含一个 (p, ε)-鲁棒向日葵。

推论 2.4. 存在一个常数 $B > 0$，使得对所有 $p, \varepsilon \in (0, 1/2]$，以下结论均成立。设 $\mathcal{F}$ 是一个由非空集合构成的族，每个集合的大小至多为 d，且满足 $|\mathcal{F}| \geq d \cdot (B \log(d/\varepsilon)/p)^d$。那么 $\mathcal{F}$ 包含一个 (p, ε)-鲁棒向日葵。

证明。令 $d_0 \in [d]$ 为 $\mathcal{F}$ 中集合大小最常见的值 d_0。那么，大小为 B 的集合数量至少为 $(d/\varepsilon \log(/p)^d$，因此可以对这些集合应用定理 2.3。 □

如果从一个鲁棒向日葵中移除一个花瓣，它仍然是一个鲁棒向日葵（只是参数稍差一些）。

引理 2.5. 假设 $N_1, \ldots, N_k \subseteq X$ 是一个核为 K 的 (p, ε)-鲁棒向日葵。那么对于每个 i，集合 $N_1, \ldots, N_{i-1}, N_{i+1}, \ldots, N_k$ 构成一个核为 K 的 $(2p, 2\varepsilon)$-鲁棒向日葵。

证明。令 $\boldsymbol{\tau}$ 按照 $[3]^{X \setminus K}$ 分布，使得 $\Pr[\boldsymbol{\tau}_i = 1] = \Pr[\boldsymbol{\tau}_i = 2] = p$、$\Pr[\boldsymbol{\tau}_i = 3] = 1 - 2p$，且 $\boldsymbol{\tau}$ 的坐标相互独立。对于 $\ell \in [3]$，令 $\boldsymbol{\tau}^\ell = \{j \in X \mid \boldsymbol{\tau}_j = \ell\} \cup K$。$(p, \varepsilon)$-鲁棒向日葵的定义意味着，对每个 $\ell \in [2]$，我们都有 $\Pr[\exists j \in [k]\colon \boldsymbol{\tau}^\ell \supseteq N_j] \geq 1 - \varepsilon$。应用并合界可得

$$\Pr\left[\exists j \in [k]\colon \boldsymbol{\tau}^1 \supseteq N_j \wedge \exists j' \in [k]\colon \boldsymbol{\tau}^2 \supseteq N_{j'}\right] \geq 1 - 2\varepsilon.$$

若 $j = j'$，则由于 $\boldsymbol{\tau}^1 \cap \boldsymbol{\tau}^2 = K$，我们有 $N_j = K$，这与向日葵的定义矛盾。因此，当该事件发生时，$j \neq j'$。令 $\boldsymbol{R}$ 为 X 的子集上的一个分布，其中每个元素以概率 $2p$ 独立地出现在 $\boldsymbol{R}$ 中。由于 $(\boldsymbol{R} | \boldsymbol{R} \supseteq K)$ 与 $\boldsymbol{\tau}^1 \cup \boldsymbol{\tau}^2$ 具有相同的分布，我们有

$$\Pr_{\boldsymbol{R}}\left[\exists j \neq j' \in [k]\colon \boldsymbol{R} \supseteq N_j \wedge \boldsymbol{R} \supseteq N_{j'} \,\middle|\, \boldsymbol{R} \supseteq K\right] \geq 1 - 2\varepsilon.$$

特别地，对于每个 $i \in [k]$，我们有

$$\Pr_{\boldsymbol{R}}[\exists j \neq i \in [k]\colon \boldsymbol{R} \supseteq N_j \mid \boldsymbol{R} \supseteq K] \geq 1 - 2\varepsilon,$$

因此，集合 $N_1, \ldots, N_{i-1}, N_{i+1}, N_k$ 构成一个以 K 为核的 $(2p, 2\varepsilon)$-鲁棒向日葵。 □

我们还会用到另一个引理，它与鲁棒向日葵和经典向日葵之间的标准联系非常相似（例如，参见 [Alw+21]中的引理 1.6）：

引理 2.6. 假设 $N_1,\dots,N_m \subseteq X$ 是一个具有核 K的 $(1/(2k)\ \varepsilon)$-鲁棒向日葵。那么

$$\Pr_{\boldsymbol{R}\sim 2^X}\left[\exists I \in \binom{[m]}{k} \forall i \in I\colon \boldsymbol{R} \supseteq N_i \middle| \boldsymbol{R} \supseteq K\right] \geq 1-\varepsilon k.$$

证明。令 $\tau \sim [2k]^{X\setminus K}$，并令 $\tau^i := \{j \in X \setminus K \mid \tau_j = i\} \cup K$，其中 $i \in [2k]$。那么 $\boldsymbol{\tau}^1 \cup \cdots \cup \boldsymbol{\tau}^k$的分布等价于 $(\boldsymbol{R} \mid \boldsymbol{R} \supseteq K)$，其中 $\boldsymbol{R} \sim 2^X$。另一方面，根据 $(1/(2k)\ \varepsilon)$-鲁棒向日葵的定义，对于每个 $i \in [2k]$，我们得到

$$\Pr[\exists j \in [m]\colon \boldsymbol{\tau}^i \supseteq N_j] \geq 1-\varepsilon.$$

由于 $\boldsymbol{\tau}^i \cap \boldsymbol{\tau}^{i'} = K$ 对于任意 $i \neq i' \in [2k]$，对 $i \in [k]$应用并合界即可得到该引理。 □

3 均匀汉明重量 k 分布的采样

本节证明引言中提到的以下结果，为方便起见，我们在此重述。

定理 1.2. 假设 $\boldsymbol{U}_k^n$ 可通过决策深度为 d 的采样得到，且其变差距离误差为 η。

1. 对每个 $\varepsilon > 0$，都存在一个常数 τ，使得 $d \leq \tau \log(n/k)/\log\log(n/k)$ 蕴含 $\eta \geq 1-\varepsilon$。

2. 存在一个常数 τ，使得以下结论成立。对于每个 $\varepsilon \in (0,1)$，若 $k \in [\log_2 n、 2^{\log^{1-\varepsilon} n}]$ 且 $d \leq \tau \log^{\varepsilon} n$，则 $\eta = 1 - n^{-\Omega(k)}$。对于 $k \in [1, \log_2 n)$ 和 $d \leq \tau \log^{\varepsilon} n/\log\log n$，同样的界也成立。

此外，条目 1 适用于任意 $\boldsymbol{U}_S^n$其最大值为$_{x\in S} x = k$或最小值为$_{x\in S} x = n-k$。

定理 1.3. 存在一个常数 $\tau > 0$，使得任何通过决策深度为 $\tau \log n/\log\log n$ 的采样得到的分布，都与 $\boldsymbol{U}_1^n$ 相距 $(1-n^{-\Omega(1)})$。

我们首先在第3.1节证明定理 1.3。然后在第3.3 节证明定理 1.2 的第2 项，并在第3.4节证明该定理的第1 项。我们在第3.5节证明“此外”部分。

3.1 定理 1.3的证明

我们证明一个更一般的结果，该结果可立即推出定理 1.3。

首先，我们概述定理 1.3 关于 d-局部函数的证明。设 $\Delta(\boldsymbol{U}_1^n, \boldsymbol{X}) \leq 1-\eta$。如果一个坐标 i 是好的，则称其满足 $\Pr[\boldsymbol{X} = e_i] \geq 1/n^2$。由于 $\Pr[U_1 = e_i] = 1/n$，许多坐标都是好的：至少有 $\Omega(\eta n)$ 个。

令 $\boldsymbol{Y} \sim \{0,1\}^m$ 表示随机输入比特。每个 X_i 依赖于坐标的某个子集 $N_i \subseteq [m]$，该子集的大小至多为 $d = \tau \log n/\log\log n$，记作 $\boldsymbol{X}_i = f_i(\boldsymbol{Y}_{N_i})$。

对于每个优良坐标 i，我们选择一个赋值 $\alpha_i \in f_i^{-1}(1)$，使条件概率 $\Pr[\boldsymbol{X} = e_i \mid \boldsymbol{Y}_{N_i} = \alpha_i]$最大，也就是说，若 $\boldsymbol{Y}_{N_i} = \alpha_i$，则所有其他输出位均为 0 的概率。该概率至少为 $1/(2^d n^2)$ $=\Omega(1/n^3)$。

这些赋值 α_i 不一定彼此一致。然而，一个随机赋值ρ 对 $\boldsymbol{Y}$至少与其中 $\Omega(\eta n/2^d)$ 个 $=\Omega(\eta n^{1-o(1)})$ 一致。令 T 由以下赋值的定义域组成：

与 ρ一致的赋值 α_i。由于 $N_i = N_j$ 蕴含 $\alpha_i = \alpha_j$，因此这些定义域各不相同，并且 $\Pr[\boldsymbol{X} = e_i \mid \boldsymbol{Y}_{N_i} = \alpha_i] = 0$。选择 d 可保证 T 支撑一个 $(1/4, \varepsilon)$-鲁棒向日葵 $\mathcal{S}$，对于任意关于 n为逆多项式的 ε 均成立。令 K 为 $\mathcal{S}$的核。

如果我们从 $\mathcal{S}$ 中移除任意一个花瓣 i，那么根据引理 2.5，结果是一个（$1/2$，2ε）-鲁棒向日葵。因此，鉴于 $\boldsymbol{Y}_K$ 与 ρ一致，存在某个 $j \neq i$ 使得 $\boldsymbol{X}_j = 1$ 的概率至少为 $1 - 2\varepsilon$。如果我们将条件替换为“$\boldsymbol{Y}_{N_i}$ 与 ρ一致”（因此 α_i），那么直观上，概率只会增大；这可以通过 FKG 不等式（定理 2.1）形式化地证明。根据 α_i的定义，这意味着 $\Pr[\boldsymbol{X} = e_i] \leq |f_i^{-1}(1)| 2\varepsilon \leq 2^{d+1}\varepsilon$。取 $\varepsilon = 1/2^{d+1}n^2$ 可知 i 并不好，由此得到矛盾。

我们接下来证明定理 1.3 的推广。

定理 3.1. 设 $\boldsymbol{Y} \sim \{0,1\}^m$ 是 n 位源 $\boldsymbol{X}$ 的输入位。假设 $\boldsymbol{X}$ 的每一位都由 $\boldsymbol{Y}$ 的位构成的析取范式计算，大小至多为 s，宽度至多为 d。对于每个 $\kappa \in \mathbb{R}$，存在一个常数 τ，使得当 $d = \tau \log n / \log\log n$ 且 $s \leq \kappa n^{\kappa}$ 时，我们有 $\Delta(\boldsymbol{X}, \boldsymbol{U}_1^n) = 1 - \eta = 1 - n^{-\Omega(1)}$。

这意味着定理 1.3 成立，因为深度为 d 的决策树的输出可以表示为一个规模至多为 2^d、宽度至多为 d 的析取范式。在我们的情形中，$d = o(\log n)$，因此 $2^d \leq n$（当 n 足够大时）。

证明。我们称输出位 $i \in [n]$ 为良好位，若 $\Pr[\boldsymbol{X} = e_i] \geq 1/n^2$。令 $G \subseteq [n]$ 为所有良好位组成的集合，并令 $\overline{G} = [n] \setminus G$。我们来估计 G 的大小：$\Pr[\boldsymbol{X} \in \{e_i \mid i \notin G\}] \leq |\overline{G}|/n^2$，但 $\Pr[\boldsymbol{U}_1^n \in \{e_i \mid i \notin G\}] = |\overline{G}|/n$），因此 $|\overline{G}| \cdot (1/n - 1/n^2) \leq 1 - \eta$，由此得到 $|G| = \Omega(\eta n)$。

对于每个 $i \in G$，由于 $\boldsymbol{X}$ 的每一位都表示为析取范式，我们有 $\boldsymbol{X}_i = \bigvee_{j \in [s_i]} [\boldsymbol{Y}_{N_i^j} = \alpha_i^j]$，其中 N_i^1、……、$N_i^{s_i} \subseteq [m]$ 是集合，$\alpha_i^j \in \{0,1\}^{N_i^j}$ 是真值赋值，并且 $s_i \leq s$。根据全概率公式，我们有

$$\Pr[\boldsymbol{X} = e_i] \leq \sum_{j \in [s_i]} \Pr[\boldsymbol{X} = e_i \wedge \boldsymbol{Y}_{N_i^j} = \alpha_i^j] \leq s \Pr[\boldsymbol{X} = e_i \wedge \boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}],$$

其中 $N_i^{\max}$和 $\alpha_i^{\max}$对应于使概率 $\Pr[\boldsymbol{X} = e_i \wedge \boldsymbol{Y}_{N_i^j} = \alpha_i^j]$最大的析取范式项。

考虑满足 $\boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}$的良好输出位的期望数量：

$$\mathbb{E}\left[\sum_{i \in G} [\boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}]\right] \geq |G| 2^{-d}.$$

因此，存在一个对输入位的赋值 ρ，使得至少有 $|G|2^{-d}$ 个良好输出位满足 $\rho_{N_i^{\max}} = \alpha_i^{\max}$。令 $T \subseteq G$ 为这些输出位构成的集合。若 i，$j \in T$，则 $N_i^{\max} \neq N_j^{\max}$；否则，$\boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}$ 意味着 $\boldsymbol{Y}_{N_j^{\max}} = \alpha_j^{\max}$也成立，进而 $\boldsymbol{X} \neq e_i$，因此 $\Pr[\boldsymbol{X} = e_i] = 0$，这与 $i \in G$矛盾。此外请注意，对于 $i \in T$，集合 N_i 均非空；否则 $|T| = 1$，对于任意$d = o(\log n)$，我们便会立即得到与 G 的大小矛盾。

情况 1。$|T| < d(4B \log(d/\varepsilon))^d$。在这种情况下，我们立即得到 δ 的下界。事实上，不等式 $|G|2^{-d} \leq |T| < d(4B \log(d/\varepsilon))^d$ 蕴含 $|G| \leq d(8B \log(d/\varepsilon))^d$，结合 $|G| = \Omega(\eta n)$ 可得 $\eta \leq d(8B\log(d/\varepsilon))^d/n$。如果 ε 是 n 的多项式倒数，那么对于足够小的 τ，我们得到 $\eta = n^{-\Omega(1)}$，其中 $d = \tau \log n / \log\log n$。

情形 2。$|T| \geq d(4B\text{对数}(d/\varepsilon))^d$。根据推论 2.4，存在一个由集合 $N_{\text{最大值}t_1}$、……、$N^{\max}_{t_k}$（其中 $\{t_1,\ldots,t_k\} \subseteq T$（回顾一下，集合 $N_{\text{最大值}i}$对于 $i \in T$都互不相同，且都非空）构成的（$1/4,\varepsilon$）-鲁棒向日葵。令 K 表示该向日葵的核。考虑该向日葵的任意一个花瓣 t_i。根据引理 2.5，我们有 $\{N_{\text{最大值}i}\}_{i\in T\setminus\{t_i\}}$是一个（$1/2,2\varepsilon$）-鲁棒向日葵。令 $\boldsymbol{U}$ 为满足 $\boldsymbol{Y}_k = \rho_k$的指标集合。那么

$$\begin{aligned}
&\Pr[\boldsymbol{X}_j = 1 \text{ for some } j \in T\setminus\{t_i\} \mid \boldsymbol{Y}_{N^{\max}_{t_i}} = \rho_{N^{\max}_{t_i}}] \geq \\
&\Pr[\boldsymbol{U} \supseteq N^{\max}_j \text{ for some } j \in T\setminus\{t_i\} \mid \boldsymbol{U} \supseteq N^{\max}_{t_i}] = \\
&\frac{\Pr[\boldsymbol{U} \supseteq N^{\max}_{t_i} \text{ and } \boldsymbol{U} \supseteq N^{\max}_j \text{ for some } j \in T\setminus\{t_i\} \mid \boldsymbol{U} \supseteq K]}{\Pr[\boldsymbol{U} \supseteq N^{\max}_{t_i} \mid \boldsymbol{U} \supseteq K]} \geq \qquad \text{定理 2.1} \\
&\Pr[\boldsymbol{U} \supseteq N^{\max}_j \text{ for some } j \in T\setminus\{t_i\} \mid \boldsymbol{U} \supseteq K] \geq \\
&1 - 2\varepsilon,
\end{aligned}$$

其中，最后一个不等式源于 $(1/2,2\varepsilon)$-鲁棒向日葵的定义。回想一下，根据 T 的选取，我们有 $\rho_{N^{\max}_{t_i}} = \alpha^{\max}_{t_i}$。因此

$$\Pr[\boldsymbol{X} \neq e_{t_i} \mid \boldsymbol{Y}_{N^{\max}_{t_i}} = \alpha^{\max}_{t_i}] = \Pr[\boldsymbol{X}_j = 1 \text{ for some } j \in T\setminus\{t_i\} \mid \boldsymbol{Y}_{N^{\max}_{t_i}} = \rho_{N^{\max}_{t_i}}] \geq 1 - 2\varepsilon.$$

因此 $\Pr[\boldsymbol{X} = e_{t_i} \wedge \boldsymbol{Y}^{N^{\max}_{t_i}} = \alpha^{\max}_{t_i}] \leq 2\varepsilon$。根据 $\alpha^{\max}_{t_i}$ 的选取，$\Pr[\boldsymbol{X} = e_{t_i}] \leq s \cdot 2\varepsilon$。取 $\varepsilon < 1/(2sn^2)$，它关于 n 是多项式倒数，正如情况 1 所要求的，我们得到 $\Pr[\boldsymbol{X} = e_{t_i}] < 1/n^2$，所以 t_i 是坏的，这与 T 的选取矛盾。□

3.2 定理 3.1 的推广版本

在本节中，我们推广定理 3.1 ，使其可用于证明定理 1 中的条目 1.2。证明沿用了定理 3.1 的证明思路。为求简洁，我们决定将两个证明都列出。

定理 3.2。设 $\boldsymbol{Y} \sim \{0,1\}^m$ 为 n 位源 $\boldsymbol{X}$ 的输入位。假设 $\boldsymbol{X}$ 的每一位都由 $\boldsymbol{Y}$ 中的位构成的析取范式计算，其规模至多为 s，宽度至多为 d。设 $t \in [n]$ 为一个参数，$\alpha(n)$ 为一个函数，并设 $\boldsymbol{F}$ 为 $\{0,1\}^n$ 上的一个分布，满足以下性质：

- 对于满足 $T \geq n/2$ 的任意集合 $T \subseteq [n]$，我们有 $\Pr[\boldsymbol{F}_T = 0^T] \leq \alpha(n)$；
- $\Pr[|\boldsymbol{F}| > t] \leq \alpha(n)$。

若 $2d \cdot t \cdot (40Bt\log n)^d \leq n$，则 $\Delta(\boldsymbol{X}, \boldsymbol{F}) \geq 1 - 2\alpha(n) - 1/2n$。这里 B是推论 2.4中的常数。

证明。若 $\Pr[\boldsymbol{X}_i = 1 \wedge |\boldsymbol{X}| \leq t] \geq 1/n^2$，我们就称输出位 $i \in [n]$ 是良好的。令 G 为良好位的集合，并令 $\overline{G} := [n]\setminus G$。假设 $|G| \leq n/2$。那么根据 $\boldsymbol{F}$ 的条件，我们有 $\Pr[\boldsymbol{F}_{\overline{G}} = 0^{\overline{G}}] \leq \alpha(n)$。另一方面，$\Pr[\boldsymbol{X}_{\overline{G}} \neq 0^{\overline{G}} \wedge |\boldsymbol{X}| \leq t] < |\overline{G}|/n^2 < 1/2n$。于是 $\Delta(\boldsymbol{X},\boldsymbol{F}) \geq 1 - 2\alpha(n) - 1/2n$，正如所需。在证明的其余部分，我们将导出与 $|G| \geq n/2$ 矛盾的结论。

与定理3.1的证明类似，我们选取每个输出位的析取范式表示中概率最大的项。对于每个 $i \in G$，由于 $\boldsymbol{X}$ 的每一位都表示为析取范式，因此有 $\boldsymbol{X}_i = \bigvee_{j\in[s_i]}[\boldsymbol{Y}_{N^j_i} = \alpha^j_i]$，其中 N^1_i、……、$N^{s_i}_i \subseteq [m]$是集合，$\alpha^j_i \in \{0,1\}^{N^j_i}$是真值赋值，并且$s_i \leq s$。根据全概率公式，有

$$\Pr[\boldsymbol{X}_i = 1 \wedge |\boldsymbol{X}| \leq t] \leq \sum_{j\in[s_i]} \Pr[|\boldsymbol{X}| \leq t \wedge \boldsymbol{Y}_{N^j_i} = \alpha^j_i] \leq s\Pr[|\boldsymbol{X}| \leq t \wedge \boldsymbol{Y}_{N^{\max}_i} = \alpha^{\max}_i],$$

其中 $N_i^{\max}$和 $\alpha_i^{\max}$对应于使概率 $\Pr[|\boldsymbol{X}| \le t \wedge \boldsymbol{Y}_{N_i^j} = \alpha_i^j]$最大的析取范式中的项。满足 $\boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}$的良好输出位的期望数量至少为 $2^{-d}|G|$，因此存在一个对输入位的赋值 ρ，使得至少有 $|G|2^{-d}$ 个良好输出位满足 $\rho_{N_i^{\max}} = \alpha_i^{\max}$。令 $T \subseteq G$ 为这些输出位构成的集合。

让我们估计集合 $\mathcal{N} := \{N_i^{\max} \mid i \in T\}$ 中有多少个不同元素。假设存在 $i_1, \ldots, i_{t+1} \in T$ 使得 $N_{i_1}^{\max} = \cdots = N_{\max^{i_{t+1}}}$。那么，根据 ρ 的定义，我们也有 $\alpha_{i_1}^{\max} = \cdots = \alpha_{i_{t+1}}^{\max}$。因此，$\boldsymbol{Y}_{N_{i_1}^{\max}} = \alpha_{i_1}^{\max}$ 意味着对每个 $j \in [t+1]$，都有$\boldsymbol{Y}_{N_{i_j}^{\max}} = \alpha_{i_j}^{\max}$，进而意味着 $|\boldsymbol{X}| \ge t+1$，于是 $\Pr[\boldsymbol{Y}_{N_{i_1}^{\max}} = \alpha_{i_1}^{\max} \wedge |\boldsymbol{X}| \le t] = 0$，这与 i_1 是好的相矛盾。因此 $|\mathcal{N}| \ge |T|/t \ge |G|2^{-d}/t \ge 2^{-d}n/2t$。

令 $\varepsilon > n^{-5}$ 为一个稍后确定的参数。根据关于 n 的条件，我们有 $|\mathcal{N}| \ge 2^{-d}n/2t \ge d \cdot (4Bt \log(d/\varepsilon))^d$，因此 $\mathcal{N}$ 包含一个 $(1/(4t)\ \varepsilon)$-鲁棒向日葵 $\mathcal{S}$。令 K 表示这个向日葵的核。

任取一个花瓣 $p \in \mathcal{N}$。那么 $\mathcal{N} \setminus \{p\}$ 是一个（$1/(2t)\ 2\varepsilon$-鲁棒向日葵，由引理2.5可知。现在根据引理2.6 可得

$$\Pr_{\boldsymbol{R} \sim 2^{[m]}}[\text{There are } t \text{ distinct petals of } \mathcal{N} \setminus \{p\} \text{ contained in } \boldsymbol{R} \mid \boldsymbol{R} \supseteq K] \ge 1 - 2t\varepsilon.$$

令 $P \subseteq T$ 为与 $\mathcal{N} \setminus \{p\}$中的元素对应的输出位的索引，令 i 为与花瓣 p对应的输出位的索引，并令 $\boldsymbol{U}$ 为满足 $\boldsymbol{Y}_t = \rho_t$的输入位索引集合。那么

$$\begin{aligned}
&\Pr[|\boldsymbol{X}| > t \mid \boldsymbol{Y}_{N_i^{\max}} = \rho_{N_i^{\max}}] = \\
&\Pr[|\boldsymbol{X}_{[n]\setminus\{i\}}| \ge t \mid \boldsymbol{Y}_{N_i^{\max}} = \rho_{N_i^{\max}}] \ge \\
&\Pr\left[\sum_{j \in P}[\boldsymbol{U} \supseteq N_j^{\max}] \ge t \,\middle|\, \boldsymbol{U} \supseteq N_i^{\max}\right] = \\
&\frac{\Pr[\sum_{j \in P \cup \{i\}}[\boldsymbol{U} \supseteq N_j^{\max}] \ge t \mid \boldsymbol{U} \supseteq K]}{\Pr[\boldsymbol{U} \supseteq N_i^{\max} \mid \boldsymbol{U} \supseteq K]} \ge \text{ Theorem 2.1} \\
&\Pr\left[\sum_{j \in P}[\boldsymbol{U} \supseteq N_j^{\max}] \ge t \,\middle|\, \boldsymbol{U} \supseteq K\right] \ge \\
&1 - 2t\varepsilon.
\end{aligned}$$

回顾一下，根据 T的选取，我们有 $\rho_{N_i^{\max}} = \alpha_i^{\max}$，因此 $\Pr[|\boldsymbol{X}| > t \mid \boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}] \ge 1 - 2t\varepsilon$。于是

$$\Pr[|\boldsymbol{X}| \le t \wedge \boldsymbol{X}_i = 1] \le s \cdot \Pr[|\boldsymbol{X}| \le t \wedge \boldsymbol{Y}_{N_i^{\max}} = \alpha_i^{\max}] \le 2st \cdot \varepsilon.$$

选取 $\varepsilon = 1/(4stn^2) \ge n^{-5}$，这与 i 是好的相矛盾。 □

3.3 次多项式重量

尽管定理3.2 蕴含定理2 的第1.2项，但我们通过从$\boldsymbol{U}_1^n$出发的归约给出一个更简单的证明。

引理 3.3。设 $S \subseteq \{0,1\}^n$ 且设 $\boldsymbol{S} \sim S$。假设 $\boldsymbol{S}$ 可以通过误差为 η、深度为 d 的决策森林进行采样。此外，假设对于每个 $s \in S$，都存在一棵深度为 k 的决策树 T_s，它接受 s，且不接受任何 $S \setminus \{s\}$。那么存在一个针对 $\boldsymbol{U}_1^{|S|}$、误差为 η、决策深度为 kd 的采样器。

证明。设 $\boldsymbol{Y}$ 为采样器针对 $\boldsymbol{S}$ 采样得到的分布。对于我们的采样器针对 $\boldsymbol{U}_1^{|S|}$ 的每个输出位，我们取一个唯一的元素 $s \in S$，并通过查询 $\boldsymbol{Y}$ 的各个位来实现 T_s 的每个查询（这至多需要对输入位进行 d 次查询）。由此得到一棵深度为 kd 的决策树 T'_s。设 $\boldsymbol{X}$ 为采样得到的分布。那么

$$\begin{aligned}\Delta(\boldsymbol{X}, \boldsymbol{U}_1^{|S|}) &= \frac{1}{2}\left(\Pr[w(\boldsymbol{X}) \neq 1] + \sum_{s\in S}\left|\Pr[\boldsymbol{X} = e_s] - 1/|S|\right|\right) \\ &= \frac{1}{2}\left(\Pr[\boldsymbol{Y} \notin S] + \sum_{s\in S}\left|\Pr[\boldsymbol{Y} = s] - 1/|S|\right|\right) = \Delta(\boldsymbol{Y}, \boldsymbol{S}) = \eta.\end{aligned}$$

□

推论 3.4. 对某个常数 $\tau' > 0$，以及任意 $\varepsilon \in (0, 1)$，每个 $(\tau' \log^{\varepsilon} n)$-决策深度采样器都会输出一个与 $\boldsymbol{U}_k^n$ 相距 $(1 - n^{-\Omega(k)})$ 的分布，其中 $k \in [\log n, 2^{\log^{1-\varepsilon} n}]$。若 $k < \log_2 n$，则这对每个 $(\tau' \log^{\varepsilon} n / \log \log n)$-局部采样器都成立。

证明。查询一个 k 大小集合中的所有元素，并且当且仅当它们全为 1 时才接受的决策树，满足引理 3.3 中的条件。若我们有一个针对 $\boldsymbol{U}_k^n$ 的 $(1 - \delta)$-误差采样器，就能得到一个针对 $\boldsymbol{U}\binom{n}{k}$ 1 的 $(1 - \delta)$-误差采样器，其决策深度为 kd；由定理 1.3 可知，当 $kd \leq \tau \log \binom{n}{k} / \log \log \binom{n}{k}$ 时，$\delta = \binom{n}{k}^{-\Omega(1)} = n^{-\Omega(k)}$。由于 $\log \binom{n}{k} = \Theta(k \log(n/k))$，我们得到$d = \Omega(\log(n/k)/(\log k + \log \log n))$。

□

3.4 次线性权重

在本节中，我们证明定理1 的第 1.2项。

引理 3.5. 假设 $\boldsymbol{X}$ 是一个决策深度为 d的采样器。那么，对于任意足够小的 $\varepsilon > 0$，都存在一个常数 τ，使得如果 $d \leq \tau \log(n/k)/\log \log(n/k)$，那么 $\Delta(\boldsymbol{X}, \boldsymbol{U}_k^n) \geq 1 - 2\varepsilon - \frac{1}{2n}$。

证明。考虑采样器的前 $\ell := \varepsilon^{-1} \cdot n/k$ 位：$\boldsymbol{X}_{\leq \ell} := \boldsymbol{X}_1, \ldots\ldots, \boldsymbol{X}_\ell$。设 $\boldsymbol{Y}$为分布ℓ 的前 $\boldsymbol{U}_k^n$位。我们证明 $\boldsymbol{Y}$满足定理3.2 关于 $t = \varepsilon^{-2}$和 $\alpha(n) = \varepsilon$的条件。首先，我们有

$$\Pr[|\boldsymbol{Y}| > t] = \Pr\left[|\boldsymbol{Y}| > \varepsilon^{-1}\,\mathbb{E}[|\boldsymbol{Y}|]\right] < \varepsilon.$$

现在，任取 T 作为 $[\ell]$ 的一个子集，其大小至少为 $\ell/2$。那么

$$\Pr[\boldsymbol{Y}_T = 0^T] = \binom{n - \ell/2}{k}\binom{n}{k}^{-1} = \prod_{i=0}^{k-1}\frac{n - i - \ell/2}{n - i} \leq \left(1 - \frac{\ell}{2n}\right)^k = \left(1 - \frac{1}{2\varepsilon k}\right)^k \leq e^{-2\varepsilon^{-1}} < \varepsilon.$$

这里我们假设 $k < n/2$，否则该引理显然成立。

应用定理 3.2，对于足够小的 τ（其取值取决于 ε），我们有 $\Delta(\boldsymbol{X}_{\leq \ell}, \boldsymbol{Y}) = 1 - 2\varepsilon - 1/2n$。为了完成证明，注意到 $\Delta(\boldsymbol{X}, \boldsymbol{U}_k^n) \geq \Delta(\boldsymbol{X}_{\leq \ell}, \boldsymbol{Y})$，因为右端的随机变量是左端变量的边缘分布。

□

3.5 切片的并集

在本节中，我们通过观察到分布$\boldsymbol{U}_S^n$接近于 $\boldsymbol{U}^n_{\max_{x\in S} x}$ 只要 $\max_{x\in S} x = o(n)$，便证明定理 1.2 的“此外”部分。

命题 3.6. 设 $k = o(n)$，并假设 $S \subseteq \{0, 1, \dots, k\}$ 满足 $k \in S$。那么我们有 $\Delta(\boldsymbol{U}_k^n, \boldsymbol{U}_S^n) = o(1)$。

证明。我们使用记号 $\binom{n}{S} := \sum_{i\in S}\binom{n}{i}$。

$$\begin{aligned}
\Delta(\boldsymbol{U}_k^n, \boldsymbol{U}_S^n) &= \frac{1}{2}\binom{n}{S\setminus\{k\}}\binom{n}{S}^{-1} + \frac{1}{2}\left(1 - \binom{n}{k}\binom{n}{S}^{-1}\right) \\
&= \binom{n}{S\setminus\{k\}}\binom{n}{S}^{-1} \\
&\le \binom{n}{k}^{-1} \sum_{i\in S\setminus\{k\}} \binom{n}{i} \\
&\le \sum_{i=0}^{k-1} \left(\frac{k}{n-i}\right)^{i-k} = \Theta(k/n).
\end{aligned}$$

□

通过观察到，翻转所有输出位不会增加采样器的决策深度，可将 S 满足 $n - \min_{x\in S} x = o(n)$ 的情形归约为 $\max_{x\in S} x = o(n)$ 的情形。

4 局部证书

本节探讨局部证明系统的能力。第4.1 节给出了一个需要局部性 $\Omega(n)$ 的语言示例，其灵感来自采样领域中一个类似的下界 [LV12]。第 4.2 节接着给出了我们的主要结果，即证明系统在 $\textsc{Maj}^{-1}(1)$ 上局部性的一个下界。

如果存在输入 x、$x' \in \{0,1\}^m$，它们仅在第 i 位上不同，且 $f(x)_j \neq f(x')_j$，我们就说输入位 $i \in [m]$ 影响输出位 $j \in [n]$；等价地，输出位 j 依赖于输入位 i。如果函数 $f\colon \{0,1\}^m \to \{0,1\}^n$ 的每个输出位至多依赖于 d 个输入位，则称该函数是 d-局部的。

4.1 纠错码

本节将说明，良好的纠错码需要一个局部性为线性的证明系统。这展示了简单的计数技巧，我们也将用它来证明多数函数的下界。

命题 4.1. 设 $C \subseteq \{0,1\}^n$ 是一个良好的码，也就是说，$|C| \ge 2^{\alpha n}$，并且对于每个 $x \neq y \in C$，x 与 y 之间的汉明距离至少为 βn，其中 α 和 β 是 (0, 1) 中的常数。

若 $f\colon \{0,1\}^m \to \{0,1\}^n$ 是 d-局部函数，且 $f(\{0,1\}^m) = C$，则 $d \ge \alpha\beta n$

证明。我们不妨假设 f 的所有输入位都会影响某些输出位。任取一个输入位 $i \in [m]$ 和一个依赖于 i 的输出位 $j \in [n]$。接着取 x、$x' \in \{0,1\}^m$，它们仅在第 i 个坐标上不同，并且满足 $f(x)_j \neq f(x')_j$。由于 $f(x) \neq f(x') \in C$，它们的汉明距离必须至少为 βn，因此 i 必须影响至少 βn 个输出位，因为 x 和 x' 仅在 i 上不同。由此，每一位都至少影响 βn 个输出位。

至少有 αn 个输入位，因为 $|C| = 2^{\alpha n}$。因此，存在 $\alpha\beta n^2$ 个输入输出对，其中输入位会影响输出位。另一方面，此类输入输出对至多有 dn 个，因此 $d \ge \alpha\beta n$。

□

4.2 多数

设 $\textsc{Maj}_n$ 为集合 $\{x \in \{0,1\}^n \mid |x| \geq n/2\}$。首先，我们给出一个简单的局部性上界，该上界隐含于 [Kre+16]。

命题 4.2 （本质上是 [Kre+16]中的定理 3.9 和推论 3.10）。存在一个$O(\log^2 n)$-局部函数 $f\colon \{0,1\}^m \to \{0,1\}^n$ 使得 $f(\{0,1\}^m) = \textsc{Maj}_n^{-1}(1)$。

证明。为简单起见，假设 n 是奇数。构造一棵二叉树，其根节点是区间 $[1,n]$，叶节点是单元素集合 $\{1\}$、……、$\{n\}$，并且每个内部节点 $[\ell,r]$ 都有两个子节点 $[\ell,c]$、$[c+1,r]$，其中 $c = \lfloor (\ell+r)/2 \rfloor$。我们可以构造这样一棵深度为 $O(\log n)$ 的树。对于树中的每个区间 $[\ell,r]$，我们都设置一个标签 $w(\ell, r)$，其取值范围为 0 到 $r-\ell+1$，用来表示 $x_\ell + \cdots + x_r$（其中 x_1、……、x_n 是输出）。我们用 $O(\log n)$ 个输入位来实现这些变量。

若内部节点 $[\ell,r]$ 的子节点为 $[\ell,c]$、$[c+1,r]$，则当且仅当 $w(\ell,r) = w(\ell,c) + w(c+1,r)$ 时，它是一致的。此外，若该内部节点是根节点 $[1,n]$，则要求 $w(1,n) \geq (n+1)/2$。每个位置 $i \in \{1,\ldots,n\}$ 对应叶节点 $w(i,i)$，它有 $O(\log n)$ 个祖先节点。若位置 i 的所有非叶祖先节点都一致，我们就称该位置是良好的。若 i 是良好的，则第 i 个输出为 $w(i,i)$；否则为 1。由于每个 $w(\ell,r)$ 都用 $O(\log n)$ 位编码，该系统的局部性为 $O(\log^2 n)$。

每个重量至少为 $(n+1)/2$ 的向量 x_1、……、x_n，都可以通过令 $w(\ell,r) = x_\ell + \cdots + x_r$ 来使用该系统生成。反过来，考虑树上任意一种权重赋值。若根节点不一致，则输出为 1、……、1，因此可以假设根节点是一致的。修剪树时，移除所有不一致节点的子节点。若 $[\ell,r]$ 是修剪后树中的任意节点，则要么 $\ell = r$ 且 $x_\ell = w(\ell,r)$，要么 $\ell < r$ 且 $x_\ell + \cdots + x_r = r - \ell + 1 \geq w(\ell,r)$。由此可得 $x_1 + \cdots + x_n \geq w(1,n) \geq (n+1)/2$。 □

$\Omega(\log^* n)$ 的局部性下界 [Bey+13] 受以下观察启发：

命题 4.3. 设 $f\colon \{0,1\}^m \to \{0,1\}^n$，使得每个输出位都是至多$c$ 个输入位的函数，并且每个输入位至多影响 d 个输出位。假设 $cd \leq (n+1)/2$。那么$f(\{0,1\}^m) \neq \textsc{Maj}^{-1}(1)$。

证明。设 $i \in [n]$ 是任意一个输出位。“邻域” $N(i)$ 中至多有 cd 个输出位，该邻域是与 i共享某个输入位的输出所构成的集合。如果 $|N(i)| \leq (n+1)/2$，那么我们可以找到一个重量为 $(n+1)/2$ 的输出 y，使得 $y_{N(i)} = 1^{N(i)}$。设 y 由输入 x 生成。必定存在一种对 i 的输入进行设置的方法，可以将其置为零（因为存在一个满足 $z_i = 0$ 的有效输出 z）。如果我们用这种设置修改 x，那么新的输出 z在 $N(i)$ 之外与 y 一致，而且 $z_i = 0$。由于 $y_{N(i)} = 1^{N(i)}$，可得 $|z| < |y|$，这不可能，因为 y 具有可能的最小重量。 □

我们证明中的一个步骤（即，引理 4.5）本质上是针对具有影响力输入位的源对命题 4.3 证明的改编。

我们在附录中对他们的证明进行了简化阐述 A。我们自己的下界包含在以下定理中。

定理 4.4. 设 $f\colon \{0,1\}^m \to \{0,1\}^n$ 是一个 d-局部函数，满足 $f(\{0,1\}^m) = \textsc{Maj}_n^{-1}(1)$，其中 n 是奇数。那么 $d = \Omega(\sqrt{\log n})$。

证明。若一个输入位 $i \in [m]$ 至少影响 f 的 k 个输出位，则称其为 k-有影响力的。我们将证明，对于每个 k，都有 $\Omega(n/(dk))$ 个 k-有影响力的输入位。因此，输出取决于输入的输入—输出位对数量至少为 $\sum_{k\in[n]} cn/(dk) =$

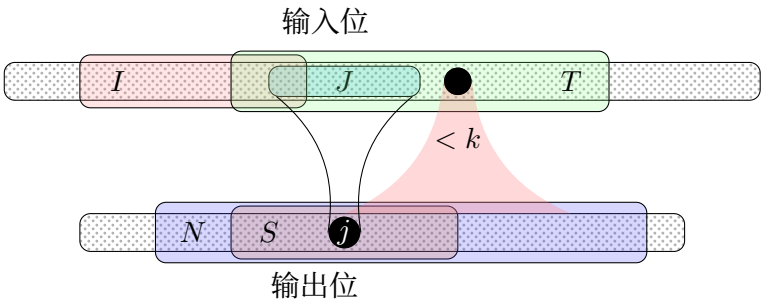


图 2： I 是 k有影响力的输入位集合，而 S 是给定集合。集合 T 由所有影响 S的输入位组成，集合 N 由所有受 $T \setminus I$影响的位组成

$\Omega(n\log n/d)$。另一方面，由于 f 是 d-局部的，这类位对至多有 nd 个。因此$d = \Omega(\sqrt{\log n})$。

接下来需要证明 k有影响力的位数量的下界。为此，将结合以下两个引理。

引理 4.5。设 I 是所有 k有影响力的输入位的集合。那么，对于大小至多为 $n/(4kd)$ 的任意输出位集合 S，存在一个从 ρ到 I 的赋值，使得

1. S 中的所有位均由 ρ固定为 1，即对于 ρ 的每个全扩展 ρ'，都有 $f(\rho')_S = 1^S$。
2. ρ 最多将 $n+1$ 个/2 输出位固定为 1。

证明。固定一个大小至多为 $n/(4kd)$ 的集合 S。

令 T 为影响 S的输入位集合，因此 $|T| \le d|S|$。令 N 为受 $T \setminus I$影响的所有位的集合。有关这些定义的示意图，请参见图2 。

由于 I 包含所有具有 k影响力的输入位，因此 $|N| \le kd|S|$。设 ρ' 为一个全赋值，使得$f(\rho')_N = 1^N$且 $|f(\rho')| = (n+1)/2$。根据引理的陈述，由于 $|N| \le n/4$，这是可能的。令 $\rho := \rho'_I$。

由于 $|f(\rho')| = (n+1)/2$，特别地，ρ至多将 $(n+1)/2$ 个输入位固定为 1。我们声称，ρ将 S 中的所有位都固定为 1。反设它没有将位 $j \in S$ 固定为 1。令 J 为影响 j 的输入位集合。令 ρ'' 为一个全赋值，除 $J \setminus I$ 外，在所有位置上都与 ρ' 一致，且满足 $f(\rho'')_j = 0$。注意到 $f(\rho'')_{[n]\setminus N} = f(\rho')_{[n]\setminus N}$，因此 $f(\rho'')$ 与 $\le f(\rho')$ 逐坐标相等。由于 $f(\rho'')_j = 0$ 且 $f(\rho')_j = 1$，我们得到 $|f(\rho'')| < n/2$，这与 f 的像为 $\mathrm{MAJ}_n^{-1}(1)$ 这一事实矛盾。 □

引理 4.6. 设 I 是一个大小至多为 $n/20$ 的任意输入位集合。那么存在一个输出位集合 S，使得 $|S| = O(|I|)$，并且对于每个将至多 $(n+1)/2$ 个输出位固定为 1 的、作用于 I 的赋值 ρ，S 中都存在一个未被 ρ 固定为 1 的位。

证明。设 $\rho_1, \ldots, \rho_K$ 是所有对 I 的赋值，且至多固定 $(n+1)/2$ 个输出位。分别用 $U_1, \ldots, U_K \subseteq [n]$ 表示 $\rho_1, \ldots, \rho_K$ 未固定的位的集合，因此$|U_1|, \ldots, |U_K| \ge (n-1)/2 \ge n/3$。于是

$$K\frac{n}{3} \le \sum_{i=1}^{K} |U_i| = \sum_{j\in[n]} |\{i \in [K] \mid U_i \ni j\}|.$$

因此，存在 j 使得 $|\{i \in [K] \mid U_i \ni j\}| \geq K/3$。令 $S_1 := \{j\}$，并对位集合 $[n] \setminus \{j\}$ 和赋值集合 $\{\rho_i : U_i \not\ni j\}$ 继续这一过程。假设前一次迭代得到一个大小为 k 的集合 $S_k \subseteq [n]$，以及一个指标集合 $T_k \subseteq [K]$。那么，对 $i \in T_k$，令 $U'_i := U_i \setminus S_k$。于是 $|U'_1|, \ldots, |U'_K| \geq (n-1)/2 - k \geq n/3$，其中最后一个不等式在 $k < n/6$ 时成立。和之前一样，存在 $j \notin S_k$ 使得 $|\{i \in T_k \mid U'_i \ni j\}| \geq |T_k|/3$。接着令 $S_{k+1} := S_k \cup \{j\}$，并令 $T_{k+1} := \{i \in T_k \mid U'_i \not\ni j\}$。

显然 $|T_k| \leq K \cdot (2/3)^{k-1}$，因此经过 $\tau = \lceil \log_{3/2} K \rceil \leq 2 \log_2 K \leq 2|I| \leq n/10$ 步后，我们可以从该集合中消除所有赋值，即 S_τ 满足：对于每个将至多 $(n+1)/2$ 个输出位固定的 I 上的赋值 ρ，都存在一个 $j \in S_\tau$，它未被 ρ 固定。（界限 $2|I| \leq n/10$ 保证条件 $k < n/6$ 成立）。 □

设 I 为所有 k-影响位构成的集合。如果 $|I| \geq n/20$，我们便立即得到所需的下界，因此不妨假设情况并非如此。接着，令 S 为由引理 4.6给出的集合，$|S| = O(|I|)$。由引理 4.5 可得 $|S| = \Omega(n/(dk))$，因此 $|I| = \Omega(n/(dk))$ 也成立。 □

5 用于采样线性切片的交换网络

本节讨论交换网络在构造决策树采样器方面的局限性。

定义 5.1. 深度为 d 的交换网络是一列 d 个匹配 $M_1, \ldots, M_d$。每个 M_i 都是由来自 $[n]$ 的元素构成的 $n/2$ 个互不相交的配对组成的集合。交换网络在切片 $\binom{[n]}{\ell}$ 上生成的分布定义如下：

- 将字符串初始化为 $1^\ell 0^{n-\ell}$；
- 对于每个 $i \in [d]$：对于 $(a, b) \in M_i$ 中的每一对，我们抛一枚公平硬币；如果结果为正面，就交换当前序列的第 a 位和第 b 位。

引理 5.2 ([Vio12b]的变体)。假设存在一个深度为 d 的交换网络，它在 $\binom{[n]}{\ell}$ 上生成变量 $\boldsymbol{X}$。那么存在一个决策深度为 d 的 $\boldsymbol{X}$ 采样器，使得每个输入位至多影响 2^d 个输出位。此外，$\boldsymbol{X}$ 的支撑集是 $\binom{[n]}{\ell}$ 的子集。

证明。输入位对应于交换网络中的抛硬币操作。每个输出位都是通过追溯其初始位置来计算的：首先，我们查询与包含该位的 M_d 中的数对相对应的硬币，然后查询与 M_{d-1} 中的数对相对应的硬币，依此类推，直到确定该位在初始序列中的位置。接着，如果该位置位于 $[\ell]$ 中，我们就输出 1，否则输出 0。很容易看出，所描述的采样器具有所需的性质。 □

引理 5.3. 设 $\alpha \in (0,1)$ 是一个常数。假设 $\boldsymbol{X}$ 可以通过一个 d-局部采样器进行采样，且每个输入位最多影响 c 个输出位，$\boldsymbol{X}$ 的支撑集包含于 $\binom{[n]}{\alpha n}$，并且 $n/(cd)^2 = \omega(2^{cd})$。那么 $\Delta(\boldsymbol{X}, \boldsymbol{U}^n_{\alpha n}) = 1 - o(1)$。

证明。对于 $\boldsymbol{X}$ 的每个输出位 $i \in [n]$，令 $N(i) \subseteq [n]$ 为 i 的邻域，即与 i 共享一个影响输入位的输出位集合。根据假设，$|N(i)| \leq cd$。我们贪心地选取一组邻域互不相交的输出位：$t_1 = 1$，并且对于 $j > 1$，$t_j \in [n]$ 是一个位，使得 $N(t_j) \cap (N(t_1) \cup \cdots \cup N(t_{j-1})) = \emptyset$。对于每个输出位 i，至多有 $(cd)^2$ 个输出位 j 满足 $N(i) \cap N(j) \neq \emptyset$，因此贪心过程会得到 $\ell \geq n/(cd)^2$ 个位。

假设存在一个位 $i \in \{t_1, \ldots, t_\ell\}$ 使得 $\Pr[\boldsymbol{X}_i = 0] > 0$ 且 $\Pr[\boldsymbol{X}_{N(i)} = 1^{N(i)}] > 0$。那么我们采用命题 4.3中的方法：令 $x \in \binom{[n]}{\alpha n}$ 为 $\boldsymbol{X}$ 支撑集中的一个字符串，使得

$x_{N(i)} = 1^{N(i)}$，并令 ρ 为产生 x的输入位。由于 $\Pr[\boldsymbol{X}_i = 0] > 0$，我们可以更改 ρ 中影响第 i 个输出位的那些位，使该输出位的值变为 0。将所得输入记为 x'。于是注意到 $x_{[n]\setminus N(i)} = x'_{[n]\setminus N(i)}$，因为 $N(i)$ 是影响第 i位的输入位所影响的输出位集合。于是 $|x'| < |x|$，因此 $x' \notin \binom{[n]}{\alpha n}$，所以它不属于 $\boldsymbol{X}$的支撑集，这就产生了矛盾。

现在分析不存在满足该条件的位的情况。令 $I \subseteq \{t_1, \ldots, t_\ell\}$由满足 $\Pr[\boldsymbol{X}_i = 0] = 0$ 的那些输出位组成。如果 $|I| \geq \ell/2$，那么

$$\Delta(\boldsymbol{X}, \boldsymbol{U}^n_{\alpha n}) \geq \left|\Pr[\boldsymbol{X}_I = 1^I] - \Pr[(\boldsymbol{U}^n_{\alpha n})_I = 1^I]\right| = 1 - \binom{n-|I|}{\alpha n}\binom{n}{\alpha n}^{-1}$$
$$= 1 - \prod_{j=0}^{\alpha n - 1} \frac{n - |I| - j}{n - j} \geq 1 - \left(1 - \frac{|I|}{n - \alpha n + 1}\right)^{\alpha n} \geq 1 - 2^{-\Omega(\frac{\alpha}{1-\alpha}|I|)} = 1 - 2^{-\Omega(\ell)}.$$

由于 $\ell = \Omega(n/(cd)^2) = \omega(1)$，在这种情况下 $\Delta(\boldsymbol{X}, \boldsymbol{U}^n_{\alpha n}) = 1 - o(1)$。

现在假设 $|I| \leq \ell/2$，并令 $J = \{t_1, \ldots, t_\ell\} \setminus I$。假设 J中的所有位都满足 $\Pr[\boldsymbol{X}_{N(i)} \neq 1^{N(i)}] = 1$。我们来计算对于 $\boldsymbol{U}^n_{\alpha n}$，该事件发生的概率：

$$\Pr[(\boldsymbol{U}^n_{\alpha n})_{N(i)} \neq 1^{N(i)}] = 1 - \binom{n-|N(i)|}{\alpha n}\binom{n}{\alpha n}^{-1} = 1 - \prod_{j=0}^{\alpha n - 1} \frac{n - |N(i)| - j}{n - j}$$
$$\leq 1 - \left(1 - \frac{|N(i)|}{n}\right)^{\alpha n} \leq 1 - 2^{-\Omega(\alpha|N(i)|)} = 1 - 2^{-\Omega(cd)}.$$

因此

$$\Pr[(\boldsymbol{U}^n_{\alpha n})_{N(i)} \neq 1^{N(i)} \text{ for all } i \in J] = \prod_{i \in J}(1 - 2^{-\Omega(cd)}) \leq (1 - 2^{-\Omega(cd)})^{\ell/2}.$$

由于 $\ell/2 = n/(2(cd)^2) = \omega(2^{cd})$，我们在这种情况下也得到了所需的统计距离下界。

□

推论 5.4。设 $\alpha \in (0, 1)$ 为一个常数。任何生成与 $\boldsymbol{U}^n_{\alpha n}$ 的统计距离为 $1 - \Omega(1)$ 的分布的交换网络，其深度均为 $\Omega(\log \log n)$。

证明。考虑一个深度为 d、生成的分布与 $1 -$ 相差 $\boldsymbol{U}^n_{\alpha n}\Omega(1)$ 的交换网络。引理5.2将其转化为一个决策深度为 d、用于 $\boldsymbol{U}^n_{\alpha n}$的采样器，使得每个输入位至多影响 $c = 2^d$ 个输出位。该采样器与 $1 -$ 相差 $\boldsymbol{U}^n_{\alpha n}\Omega(1)$，且支撑集位于$\binom{[n]}{\alpha n}$之内。由引理5.3即可得到结论。

□

致谢。本项目获得了欧盟“地平线 2020”研究与创新计划的资助，资助协议编号为 802020-ERC-HARMONIC。

本研究得到瑞士国家教育、研究与创新秘书处（SERI）的资助，合同编号为 MB22.00026。

我们感谢 Mika G¨o¨os、Aleksandr Smal 和 Anastasia Sofronova 提供的富有成效的讨论与建议。

参考文献

[Alw+21]Ryan Alweiss、Shachar Lovett、Kewen Wu 和 Jiapeng Zhang。《向日葵引理的改进界》。载于：《数学年刊》194.3（2021），第 795–815 页。DOI：10.4007/annals.2021.194.3.5。URL：https://doi.org/10.4007/annals.2021.194.3.5.[Bab87] L´aszi´o Babai。《随机预言机将 PSPACE 与多项式时间层次分离》。载于：《信息处理快报》26.1（1987），第 51–53 页。ISSN：0020-0190。DOI：https://doi.org/10.1016/0020-0190(87)90036-6。URL：https://www.sciencedirect.com/science/article/pii/0020019087900366。[BCW21]Tolson Bell、Suchakree Chueluecha 和 Lutz Warnke。《关于向日葵的注记》。载于：《离散数学》344.7（2021），第 112367 页。ISSN：0012-365X。DOI：https://doi.org/10.1016/j.disc.2021.112367。URL：https://www.sciencedirect.com/science/article/pii/S0012365X21000807。[Bey+13] Olaf Beyersdorff、Samir Datta、Andreas Krebs、Meena Mahajan、Gido Scharfenberger-Fabian、Karteek Sreenivasaiah、Michael Thomas 和 Heribert Vollmer。《在常数深度下验证证明》。载于：ACM计算理论汇刊 5.1（2013），文章编号 2，共 23 页。ISSN：1942-3454。DOI：10.1145/2462896.2462898。URL：https://doi.org/10.1145/2462896.2462898。[BIL12] Chris Beck、Russell Impagliazzo 和 Shachar Lovett。《决策树的大偏差界与 AC0 电路的采样下界》。载于：2012 IEEE 第 53 届年度计算机科学基础研讨会。2012，第 101–110 页。 DOI：10.1109/FOCS.2012.82。[Bog+16] Andrej Bogdanov、Yuval Ishai、Emanuele Viola 和 Christopher Williamson。《有界不可区分性与秘密恢复复杂度》。载于：密码学进展——CRYPTO 2016：第 36 届国际密码学年会，美国加利福尼亚州圣巴巴拉，2016 年 8 月 14–18 日，论文集，第三部分。Springer。2016，第 593–618 页。[Bog+22] Andrej Bogdanov、Krishnamoorthy Dinesh、Yuval Filmus、Yuval Ishai、Avi Kaplan 和 Akshayaram Srinivasan。《简单源的有界不可区分性》。载于：第 13 届理论计算机科学创新会议，ITCS 2022，2022 年 1 月 31 日至 2 月 3 日，美国加利福尼亚州伯克利。Mark Braverman 编。第 215 卷。LIPIcs。Schloss Dagstuhl — Leibniz 信息学中心，2022，26:1–26:18。DOI：10.4230/LIPIcs.ITCS.2022.26。URL：https://doi.org/10.4230/LIPIcs.ITCS.2022.26。[CGZ22] Eshan Chattopadhyay、Jesse Goodman 和 David Zuckerman。《采样的空间复杂度》。载于：第 13 届理论计算机科学创新会议（ITCS 2022）。Mark Braverman 编。第 215 卷。莱布尼茨信息学国际论文集（LIPIcs）。德国 Dagstuhl：Schloss Dagstuhl — Leibniz 信息学中心，2022，40:1–40:23。 ISBN：978-3-95977-217-4。 DOI：10.4230/LIPIcs.ITCS.2022.40。 URL：https://drops.dagstuhl.de/opus/volltexte/2022/15636。[CS16] Gil Cohen 和 Leonard J Schulman。《近对数最小熵的提取器》。载于：2016 IEEE 第 57 届年度计算机科学基础研讨会（FOCS）。IEEE。2016，第 178–187 页。

[CZ16] Eshan Chattopadhyay 和 David Zuckerman。《显式双源提取器与鲁棒函数》。载于：第四十八届 ACM计算理论研讨会论文集。2016 年，第 670–683 页。[Czu15] Artur Czumaj。《使用交换网络的随机排列》。载于：第四十七届 ACM计算理论研讨会论文集。STOC '15。美国纽约州纽约：计算机协会，2015 年，第 703–712 页。 ISBN：9781450335362。 DOI：10.1145/2746539.2746629。 URL：https://doi.org/10.1145/2746539.2746629。[FKG71] C. M. Fortuin、P. W. Kasteleyn 和 J. Ginibre。《某些偏序集上的相关不等式》。载于：Communications in Mathematical Physics 22（1971），第 89–103 页。ISSN：0010-3616。 DOI：10.1007/bf01651330。[GTS14] Efraim Gelman 和 Amnon Ta-Shma。《Benes网络是 q $(q-1)/2n$-近似 q-集合独立的》。载于：第 34 届软件技术与理论计算机科学基础国际会议（FSTTCS 2014）。Venkatesh Raman 和 S. P. Suresh 编。第 29 卷。莱布尼茨国际信息学论文集（LIPIcs）。德国 Dagstuhl：Schloss Dagstuhl–Leibniz信息学中心，2014 年，第 327–338 页。 ISBN：978-3-939897-77-4。 DOI：10.4230/LIPIcs.FSTTCS.2014.327。 URL：http://drops.dagstuhl.de/opus/volltexte/2014/4853。[GW20] Mika G¨o¨os 和 Thomas Watson。《不相交集合采样的下界》。载于：ACM计算理论汇刊 12.3（2020），文章 20，共 13 页。 ISSN：1942-3454。 DOI：10.1145/3404858。 URL：https://doi.org/10.1145/3404858。[Hag91] Torben Hagerup。《随机排列的快速并行生成》。载于：自动机、语言与编程：第 18 届国际学术会议，西班牙马德里，1991 年 7 月 8–12 日，论文集 18。Springer。1991 年，第 405–416 页。[Har60] T. E. Harris。《某种渗流过程的临界概率下界》。载于：Proc. Cambridge Philos. Soc. 56（1960），第 13–20 页。[H˚as87] Johan H˚astad。《小深度电路的计算局限性》。载于：MIT Press（1987）。[IN96] Russell Impagliazzo 和 Moni Naor。《可证明地与子集和一样安全的高效密码方案》。载于：Journal of cryptology 9.4（1996），第 199–216 页。[Kil88] Joe Kilian。《基于不经意传输创立密码学》。载于：第二十届 ACM计算理论研讨会论文集。STOC '88。美国纽约州纽约：计算机协会，1988 年，第 20–31 页。 ISBN：0897912640。 DOI：10.1145/62212.62215。 URL：https://doi.org/10.1145/62212.62215。[Kre+16] Andreas Krebs、Nutan Limaye、Meena Mahajan 和 Karteek Sreenivasaiah。《小深度证明系统》。载于：ACM计算理论汇刊 9.1（2016），文章 2，共 26 页。 ISSN：1942-3454。 DOI：10.1145/2956229。 URL：https://doi.org/10.1145/2956229。[LV11] Shachar Lovett 和 Emanuele Viola。《有界深度电路无法采样优良编码》。载于：第 26 届 IEEE 计算复杂性年会，CCC 2011，美国加利福尼亚州圣何塞，2011 年 6 月 8–10 日。IEEE Computer Society，2011 年，第 243–251 页。 DOI：10.1109/CCC.2011.11。 URL：https://doi.org/10.1109/CCC.2011.11。

[LV12] Shachar Lovett 和 Emanuele Viola。《有界深度电路无法采样优良编码》。载于：计算复杂性 21.2（2012），第 245–266 页。ISSN：1016-3328。DOI：10.1007/s00037-012-0039-3。URL：https://doi-org.ezlibrary.technion.ac.il/10.1007/s00037-012-0039-3。[Rao19] Anup Rao。《面向向日葵的编码》。2019 年。DOI：10.48550/ARXIV.1909.04774。URL：https://arxiv.org/abs/1909.04774。[Ros14] Benjamin Rossman。《随机图上 k-团的单调复杂度》。载于：SIAM计算杂志 43.1（2014），第 256–279 页。ISSN：0097-5397。DOI：10.1137/110839059。URL：https://doi-org.ezlibrary.technion.ac.il/10.1137/110839059。[Smo87] R. Smolensky。《布尔电路复杂度下界理论中的代数方法》。载于：第十九届 ACM计算理论研讨会论文集。STOC '87。美国纽约州纽约：计算机协会，1987 年，第 77–82 页。ISBN：0897912217。DOI：10.1145/28395.28404。URL：https://doi.org/10.1145/28395.28404。[Vio12a] Emanuele Viola。《图灵机源的提取器》。载于：近似、随机化与组合优化。算法与技术——第 15 届国际研讨会 APPROX 2012 与第 16 届国际研讨会 RANDOM 2012，美国马萨诸塞州剑桥，2012 年 8 月 15–17 日，论文集。Anupam Gupta、Klaus Jansen、Jos´e D. P. Rolim 和 Rocco A. Servedio 编。第 7408 卷。计算机科学讲义。Springer，2012 年，第 663–671 页。DOI：10.1007/978-3-642-32512-0_56。URL：https://doi.org/10.1007/978-3-642-32512-0_56。[Vio12b] Emanuele Viola。《分布的复杂度》。载于：SIAM计算杂志 41.1（2012），第 191–218 页。DOI：10.1137/100814998。URL：https://doi.org/10.1137/100814998。[Vio14] Emanuele Viola。《电路源的提取器》。载于：SIAM计算杂志 43.2（2014），第 655–672 页。DOI：10.1137/11085983X。URL：https://doi.org/10.1137/11085983X。[Vio16] Emanuele Viola。《二次映射难以采样》。载于：ACM计算理论汇刊 8.4（2016），18:1–18:4。DOI：10.1145/2934308。URL：https://doi.org/10.1145/2934308。[Vio20] Emanuele Viola。《采样下界：布尔平均情形与排列》。载于：SIAM计算杂志 49.1（2020），第 119–137 页。DOI：10.1137/18M1198405。URL：https://doi.org/10.1137/18M1198405。[Vio21] Emanuele Viola。《通过分隔器获得新的采样下界》。2021 年。[WP23] Adam Bene Watts 和 Natalie Parham。《使用浅层电路采样的无条件量子优势》。2023 年。arXiv：2301。00995 [量子物理]。

A 多数的 $\Omega(\log^* n)$ 下界的阐述

以下是对*[Bey+13*定理 5 证明的阐述。*1]*。假设 n 为奇数，并考虑一个用于判定 1 的数量多于 0 的向量（即汉明重量至少为 $(n+1)/2$）的局部性 c 证明系统。我们可以假设 $c \geq 2$。[4]

[4]或者，改变 将 $B(0)$ 或者重新定义 d-影响为有 *more than d* 输出位依赖于它 .

证明系统有输入和输出。输出的数量为 n，每个输出最多依赖 c 个输入位。若一个输入位至少被 d 个输出位依赖，则称其为 d-有影响力的输入位。此类有影响力的输入位至多有 cn/d 个。

令 $1 = B(0) < B(1) < \cdots < B(c+1)$ 为一个常数序列（取决于 c，但不取决于 n），并令 $d(\ell) = cn/B(\ell)$，使得 $cn = d(0) > d(1) > \cdots > d(c+1) = \Omega(n)$。因此，至多有 $B(\ell)$ 个输入位是 $d(\ell)$-有影响力的。

我们将构造一个集合序列 $\emptyset = R_0 \subseteq R_1 \subseteq \cdots \subseteq R_c \subseteq [n]$，使其具有以下性质：如果 ρ 是对 $d(\ell)$-有影响力变量的真值赋值，并且该赋值可以扩展为一个将 R_ℓ 中所有坐标设为零的完整真值赋值，那么对于每个坐标 $i \notin R_\ell$，ρ 也可以扩展为一个将坐标 i 设为零的完整真值赋值。[5]

给定 $R_{\ell-1}$，下面说明如何构造 R_ℓ。我们从 $R := R_{\ell-1}$ 开始，并且可能会向 R 添加更多输出坐标。在任意时刻，假设存在一个对 $d(\ell)$-有影响力变量的真值赋值 ρ，该赋值满足：(i) 可以扩展为一个将 R 中所有坐标设为零的完整真值赋值；(ii) 对于某个坐标 $i \notin R$，任何扩展 ρ 的完整真值赋值都会将坐标 i 设为一。如果发生这种情况，我们就将 i 添加到 R 中。此后，ρ 将不再出现，因为任何扩展 ρ 的完整真值赋值都不会将 i 设为一，而 i 属于 R_ℓ。最终，不再存在这样的“坏”真值赋值，于是我们设 $R_\ell := R$。

如果 $i \in R_\ell$，那么存在某个 $r \le \ell$、某个 $R \subseteq R_r$，以及对 $d(r)$-影响变量的某个真值赋值 ρ_i，使得 ρ_i 可以扩展为一个完整真值赋值，将R 中的所有坐标都设为零；而任何扩展 ρ_i 的完整真值赋值都会将 i 设为一。如果 $j \in R_\ell$ 是在i 之后添加的，那么真值赋值ρ_j 可以扩展为一个完整真值赋值，将$R \cup \{i\}$ 中的所有坐标都设为零。特别地，ρ_j 不能扩展 ρ_i（作为特例，$\rho_j \neq \rho_i$）。由此可知，如果我们将每个 ρ_i 任意扩展为对 $d(\ell)$-影响变量的真值赋值，那么所得赋值将互不相同。因此，

$$|R_\ell| \le 2^{B(\ell)}.$$

当 n足够大时，这至多为 $\frac{n-1}{2}$。

下面我们将证明，适当选择参数后，存在一个输出坐标 i，它满足以下性质：对所有 $\ell \in \{0, \ldots, c\}$，$i \notin R_\ell$，以及所有同时也是$R_\ell$ 输入的 i 输入，都是 $d(\ell+1)$-影响的。我们将证明，对每个 $\ell \in \{0, \ldots, c\}$，$i$ 都有一个输入，它是$d(\ell+1)$-影响的，但不是 $d(\ell)$-影响的。对于不同的 ℓ，这些输入各不相同（因为一个不是 $d(\ell)$-影响的输入，对于所有 $r < \ell$，也不是 $d(r)$-影响的），因此 i 依赖于 $c+1$个输入，这是不可能的。

令 $\ell \in \{0, \ldots, c\}$。下面证明，$i$ 有一个输入，该输入是 $d(\ell+1)$-有影响力的，但不是 $d(\ell)$-有影响力的。我们用反证法：假设 i 的所有 $d(\ell+1)$-有影响力的输入都是 $d(\ell)$-有影响力的。根据假设，$i \notin R_\ell$ 以及 i 中所有同时也是 R_ℓ 输入的输入，都是 $d(\ell)$-有影响力的。令 $N(i)$ 由 i 以及所有与 i 共享某个非 $d(\ell)$-有影响力位的其他输出位组成。所有这些共享的输入位实际上都是非 $d(\ell+1)$-有影响力的，因此

$$|N(i)| \le 1 + cd(\ell+1) \le 1 + cd(1) = 1 + \frac{c^2 n}{B(1)}.$$

如果 $B(1) > 2c^2$ 那么 $|N(i)| < 1 + \frac{n}{2}$，因此 $|N(i)| \le (n+1)/2$。由于 $|R_c| \le \frac{n-1}{2}$，证明系统生成了某个重量为 $(n+1)/2$ 的向量 v，其中 R_c 的所有坐标均为零，而 $N(i)$ 的所有坐标均为一。考虑任意一个生成v的完整真值赋值 α，并令 ρ 为其在 $d(\ell)$-影响坐标上的限制。由于 $i \notin R_\ell$，根据构造，我们

[5]在论文中，ρ 扩展为一个完整真值赋值，将 R_ℓ 和 i 都设为零。

知道 ρ 扩展为某个将 i 赋值为零的完整真值赋值 β。现在考虑以下完整真值赋值:

$$\gamma(j) = \begin{cases} \rho(j) \text{ 若 } j \text{ 是 } d(\ell)\text{-影响的} \\ \beta\ j \qquad j \qquad d\ \ell \qquad ( ) \text{ 若不是 } ()\text{-影响的，且影响 } N(i) \\ \alpha \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (j) \text{ 若 } j \text{ 不是 } d(\ell)\text{-影响的，且不影响 } N(i)。 \end{cases}$$

由于 α, β 都扩展了 ρ, γ 在 $d(\ell)$-影响变量上的取值与它们一致。因此， γ 生成的输出与 α 生成的输出一致，除了 $N(i)$中的坐标之外；这些坐标可能从 1 变为 0。此外， γ 的第 i个输出与 β的第 i个输出一致，也就是 0。因此， γ 生成的输出的汉明重量严格小于 $(n+1)/2$，这是不可能的。

剩下要证明的是，存在一个输出位 i 使得 $i \notin R_c$，并且对于所有 $\ell \in \{0, \ldots, c\}$， i 中同时也是 R_ℓ 输入的所有输入都是 $d(\ell+1)$-影响的。我们通过给出坏输出位数量的上界来证明这一点。若一个输出位属于 R_c，或者存在某个 $\ell \in \{0, \ldots, c\}$，使得 i 和 R_ℓ 有一个联合输入不是 $d(\ell+1)$-影响的，那么这个输出位就是坏的。如果 i 由于某个 ℓ而是坏的，那么 R_ℓ 中必定存在一个非-$d(\ell+1)$-影响输入，同时也是 i的输入。因此，坏输入的数量至多为

$$|R_c| + \sum_{\ell=0}^{c} |R_\ell| \cdot c \cdot d(\ell+1) \leq 2^{B(c)} + n \cdot c^2 \sum_{\ell=0}^{c} \frac{2^{B(\ell)}}{B(\ell+1)}.$$

经过适当选择序列 $B(\ell)$， n 的系数将严格小于 1，因此当 n 足够大时，坏输入的数量少于 n。

序列 $B(\ell)$ 的一种取法是

$$B(\ell+1) = 2^{B(\ell)} \cdot 2c^2(c+1).$$

特别地， $B(1) = 4c^2(c+1) > 2c^2$，这正是上文所需的。对于这种取法 $B(0), \ldots, B(c)$，坏输入的数量至多为

$$2^{B(c)} + \frac{n}{2},$$

如果 $2^{B(c)} \leq \frac{n-1}{2}$，那么这小于 n；而这正是证明中另一步所需的条件。

粗略地说， $B(\ell+1) \approx 2c^3 \cdot 2^{B(\ell)}$，因此 $B(c) \approx 2 \uparrow\uparrow c$。所以 $2^{B(c)} \leq \frac{n-1}{2}$，于是当 $c \leq \kappa \log^* n$ 时，该论证成立，其中 κ 是适当的常数。