arXiv:2304.11495v2 [cs.CC] 3 Jul 2024

# 显式定向仿射提取器与线性分支程序的改进困难性

Xin Li * Yan Zhong †

2024年7月4日

摘要

仿射提取器为多种计算模型（例如 $AC^0$ 电路、奇偶决策树和一般布尔电路）提供了一些已知的最佳下界。然而，目前尚不清楚它们能否为读一次分支程序（ROBP）提供强下界。在近期的一项工作中，Gryaznov、Pudl´ak 和 Talebanfard（CCC' 22）引入了称为方向仿射提取器的更强版本，同时提出了读一次分支程序的一个推广，其中每个节点都可以进行线性查询，并证明前者蕴含了对后者中一种称为强只读一次线性分支程序（SROLBP）的特定类型的强下界。他们的主要结果给出了针对熵 $k > 2n/3$的方向仿射提取器的显式构造，从而蕴含了对 SROLBP 的平均情形复杂度 $2^{n/3-o(n)}$，且相关性呈指数级小。Chattopadhyay 和 Liao（CCC' 23）的后续工作通过与 Chattopadhyay 和 Li（STOC' 16）引入的和集提取器建立新的联系，并利用 Chattopadhyay 和 Liao（STOC' 22）对这类提取器的显式构造，将困难性提高到$2^{n-o(n)}$，但代价是将相关性增大到多项式级别。两项工作都留下了以下问题：如何构造更好的方向仿射提取器，以及如何在小相关性的情形下改进针对 SROLBP 的平均情形复杂度。

本文对方向仿射提取器、SROLBP 和 ROBP 进行了更深入的研究。主要结果包括：

- 显式构造了具有 $k = o(n)$ 和指数级小误差的方向仿射提取器，由此得到针对 SROLBP、具有指数级小相关性的平均情形复杂度 $2^{n-o_{(n)}}$，从而回答了先前工作提出的两个开放问题。
- 给出了一个 $AC^0$ 中的显式函数，对于任意常数 $\delta > 0$，该函数针对具有可忽略相关性的 ROBP，实现平均情形复杂度 $2_{(1-\delta)^n}$。此前尚无此类平均情形困难性的已知结果，而针对 ROBP 的任何 $AC^0$ 函数所能达到的最佳规模下界为 $2^{\Omega(n)}$。

我们构造中的关键要素之一，是一种基于维度扩展器、适用于仿射源的新型线性某处压缩器。该压缩器还无条件地改进了显式仿射提取器在具有可忽略误差时的熵要求。我们进一步证明，在 $F_2^n$ 中的多项式 Freiman-Ruzsa 定理成立的条件下，该压缩器同样适用于一般的弱随机源；该定理最近由 Gowers、Green、Manners 和 Tao 证明（arXiv' 23）。

*约翰斯·霍普金斯大学计算机科学系，lixints@cs.jhu.edu。研究由 NSF CAREER奖 CCF-1845349 和 NSF奖 CCF-2127575 资助。

†约翰斯·霍普金斯大学计算机科学系，yzhong36@jhu.edu。研究由 NSF CAREER奖 CCF-1845349 资助。

# 目录

1 引言 1
1.1 我们的结果 .3 1.2 技术概述 .5 1.3 论文组织结构 .10

2 预备知识 10
2.1 概率分布与熵 .10 2.2 某处随机源与提取器 .11 2.3 仿射源的结构 .11 2.4 平均条件最小熵与平均情形有种子提取器 .12 2.5 交替提取与独立性合并 .12 2.6 $\varepsilon$-偏差空间与异或引理 .13

3 仿射源的线性某处凝聚器 14

4 一般弱源的线性某处凝聚器 17
4.1 一些有用的结果 .17 4.2 构造 .18

5 方向仿射提取器 22
5.1 低次仿射相关性破坏器 .22 5.2 线性熵的方向仿射提取器 .29 5.3 次线性熵源的方向仿射弥散器和提取器 . . . . . 38

6 平均情形$AC^0$ 读一次分支程序的困难性 40 6.1 $AC^0$-可计算的 $t$-仿射相关性破坏器 .40 6.2 $AC^0$-可计算的读一次分支程序源提取器 . . . . . 48

7 开放问题 50

深度为 3 的 $AC^0[\oplus]$ 电路可以计算最优方向仿射提取器 54

B 缺失的证明 57 B.1 引理 8 的证明 .57 B.2 引理 35 的证明 .57

# 1 引言

随机性提取器是一类函数，能够从质量较差的弱随机源中提取出近乎均匀的随机比特。尽管随机性提取器最初的动机是弥合典型应用所需的随机性质量与实践中可获得的随机性质量之间的差距，但作为伪随机对象，它们后来在计算机科学中展现出广泛的应用。例如，已知有一类提取器 仿射提取器 与复杂性理论密切相关。事实上，通过标准的切换引理，它们为 $\mathsf{AC}^0$ 电路（常深度电路，包含 NOT 门以及无界扇入的 AND、OR 门）提供了强有力的规模下界 [H˚as86], ，并且已证明，它们为底层含有奇偶门的 DNF 电路给出了指数规模下界，同时使奇偶决策树具有强平均情形困难性 [CS16]。通过复杂的门消除技术，它们还为一般布尔电路给出了目前已知最好的规模下界 [DK11，FGHK16，LY22]。下面我们定义仿射提取器。

定义 1 （仿射提取器）。一个 （$n$， $k$） 仿射源是某个维数为 $k$的仿射子空间上的均匀分布，该仿射子空间属于向量空间 $\mathsf{F}_2^n$。[1] 函数 $\mathsf{Ext}: \{0,1\}^n \to \{0,1\}^m$ 是熵为 $k$ 、误差为 $\varepsilon$ 的仿射提取器，（$n$， $k$） 仿射源 $X$ 均有

$$\mathsf{Ext}(X) \approx_\varepsilon U_m,$$

其中， $U_m$ 表示在 $\{0,1\}^m$上的均匀分布， $\approx_\varepsilon$ 表示在统计距离意义下与 $\varepsilon$ 接近。若 Ext 可由多项式时间算法计算，则称其为显式的。

然而，已知的仿射提取器尚不能对衡量空间复杂度的计算模型推出强下界。例如，这一背景下的一个自然模型是分支程序，它是一个具有一个源和两个汇点的有向无环图，并且每个非汇点节点的出度均为 2。为了定义分支程序的计算过程，需要为每个非汇点节点标记一个输入比特的索引，并分别用 0 和 1 标记两条出边。此外，一个汇点标记为 1，另一个标记为 0。现在，程序通过从源出发，读取相应的输入比特并沿着相应的边，按照自然路径到达某个汇点，从而计算任意输入。当且仅当路径终止于标记为 1 的汇点时，程序接受该输入；分支程序的大小定义为其节点数，这一规模大致对应于 $2^{O(s)}$ 其中 $s$ 是计算的空间复杂度。

事实证明，对于一般分支程序，证明显式函数的非平凡下界是一个颇具挑战性的问题。经过数十年的努力，目前已知的最佳界为 $\Omega(n^2 \log^2 n)$ [Nec66]，但~~这还~~不足以将 P 与 LOGSPACE 分离。因此，大多数关于分支程序下界的研究都集中于受限模型，其中研究最为深入的是只读一次分支程序模型，即在任意计算路径上，任何输入比特至多被读取一次。在该模型中，已知指数级下界 [Weg88, Z´ak84, Dun85, Juk88, KMW91, SS92, Pon98, G´al97, BW98, ABCR99, Kab03], ； 然而，尚不清楚仿射提取器是否能在此处蕴含强下界。例如，内积对于任意熵$k > n/2$都是良好的仿射提取器，但它可以由大小为 $O(n)$ 的只读一次分支程序计算。

在一项近期研究中，[GPT22], Gryaznov, Pudl´ak 和 Talebanfard 提出了一种称为方向仿射提取器的仿射提取器推广，以及一种称为读一次线性分支程序的标准读一次分支程序推广，并证明，前者的显式构造可为后者的某些情形带来强下界。我们在下文定义这两种推广。

[1] 更一般地，仿射源和仿射提取器可以定义在任意有限域上，但本文聚焦于二元域 $\mathsf{F}_2$。

定义 2（方向仿射提取器）。若函数 DAExt : $\{0,1\}^n \to \{0,1\}^m$ 对熵 $k$ 的误差为 $\varepsilon$，且对于每个 $(n,k)$ 仿射源 $X$ 及每个非零向量 $a \in \mathbf{F}_2^n$，都有

$$(\mathsf{DAExt}(X), \mathsf{DAExt}(X+a)) \approx_\varepsilon (U_m, \mathsf{DAExt}(X+a)).$$

如果存在某个 $b \in \{0,1\}^m$，使得

$$\left|\mathsf{Supp}\left(\mathsf{DAExt}(X) \mid \mathsf{DAExt}(X+a) = b\right)\right| = 2^m$$

.

注 1. 我们的定义比 *[GPT22],* 中的定义稍微更一般，因为我们允许提取器输出多个比特。对于 $m = 1$ 这一特殊情形，我们的定义意味着，在 *[GPT22],* 中，反过来也成立，只是参数会有少量损失，如 *[CL23]* 所示。

定义 3 (线性分支程序[GPT22])。 在 $\mathbf{F}_2^n$ 上的线性分支程序是一个有向无环图 $P$，满足以下性质：

- 在 $P$ 中只有一个源 $s$ 。
- $P$ 中有两个汇点，分别标记为 0 和 1。
- 每个非汇点节点 $v$ 都标记有一个线性函数 $\ell_v : \mathbf{F}_2^n \to \mathbf{F}_2$。此外，从 $v$ 恰有两条出边，分别标记为 1 和 0。

$P$的大小是 $P$中的非汇节点数量。 $P$按如下方式计算布尔函数 $f$： $\{0,1\}^n \to \{0,1\}$ 对于每个输入 $x \in \mathbf{F}_2^n$， $P$从 $s$开始沿计算路径前进；当位于非汇节点 $v$时，沿标签为$\ell_v(x) \in \{0,1\}$的边移动到下一个节点。当路径到达汇节点时，计算结束，并将 $f(x)$定义为该汇节点上的标签。

[GPT22] 定义了两类读一次线性分支程序（简称ROLBP ）。具体而言，给定任意线性分支程序 $P$ 和其中任意节点 $v$ ，对于 $P$中的 $v$ ，令$\mathsf{Pre}_v$ 表示从源到 $v$的任意路径上出现的所有线性查询的张成空间，但不包括查询 $\ell_v$。令$\mathsf{Post}_v$表示从 $v$开始的子程序中所有线性查询的张成空间。

定义 4（弱读一次线性分支程序）。线性分支程序 $P$ 是弱读一次的，当且仅当对于 $P$的每个内部节点 $v$ ，都有 $\ell_v \notin \mathsf{Pre}_v$。

定义 5（强读一次线性分支程序）。如果对于 $P$ 的每个内部节点 $v$ ，均满足 $\mathsf{Pre}_v \cap \mathsf{Post}_v = \{0\}$，则线性分支程序 $P$ 是强读一次的。

本文将重点研究强只读一次线性分支程序，并以 SROLBP 作为简称。正如 [GPT22] 和 [CL 23], 所指出的，即使是限制更严格的 SROLBP，也能推广若干重要且得到深入研究的计算模型，例如决策树、奇偶决策树和标准读一次分支程序。这些模型在诸多领域都有应用，例如学习理论、流式算法、通信复杂度和查询复杂度。因此，正如从 $\mathsf{AC}^0$ 电路到 $\mathsf{AC}^0[\oplus]$ 电路（即带奇偶门的 $\mathsf{AC}^0$）的自然推广，以及从决策树到奇偶决策树的推广一样，研究从 ROBP 到 ROLBP 的推广也是一个自然方向。此外，正如 [GPT22], 所指出的，奇偶决策树是 $\mathsf{AC}^0[\oplus]$ 中唯一具有强平均情形下界的情形，而且它们与树状归结反驳证明系统密切相关。因此，将 ROLBP 作为奇偶决策树的推广来研究尤为重要（事实上，这正是 [GPT22]中的初衷）。下面我们定义 SROLBP 的两个复杂度度量。

定义 6. 对于布尔函数 $f\colon \{0,1\}^n \to \{0,1\}$，令 SROLBP($f$) 表示计算 $f$ 的强只读一次线性分支程序的最小可能规模，并令 SROLBP$_\varepsilon$($f$) 表示满足以下条件的强只读一次线性分支程序 $P$ 的最小可能规模

$$\mathbf{Pr}_{x\leftarrow_U \mathsf{F}_2^n}[P(x) = f(X)] \geq \frac{1}{2} + \varepsilon.$$

该定义可以自然地适用于只读分支程序。

GPT 的主要贡献是[GPT22] 证明了方向仿射提取器能够为SROLBP提供强平均情形困难性。具体而言，他们证明，对于任意熵为DAExt $k$ 且误差为 $\varepsilon$的方向仿射提取器，有SROLBP$^{\sqrt{\varepsilon/2}}$(DAExt) $\geq \varepsilon 2^{n-k-1}$。此外，他们还给出了一个针对 $k \geq \frac{2n}{3} + c$ 且满足 $\varepsilon \leq 2^{-c}$的方向仿射提取器的显式构造，该构造还意味着对于规模至多为SROLBP的 $2^{\frac{n}{3}-o(n)}$，其具有指数级平均情形困难性。因此，方向仿射提取器确实强于标准仿射提取器，并且能够在更多计算模型中给出强下界。 [GPT22] 留下了针对 $k = o(n)$显式构造方向仿射提取器的问题。

在后续工作中，Chattopadhyay 和 Liao [CL23] 证明了另一类被称为和集提取器的提取器同样能够为SROLBP提供强平均情形困难性。这些提取器由 Chattopadhyay 和 Li [CL16b], 引入，适用于两个或更多个独立弱随机源之和的提取器。通过使用 [CL22], 中的这类提取器的现有构造，他们给出了一个显式函数Ext，使得SROLBP$_{n^{-\Omega(1)}}$(Ext) $\geq 2^{n-\log^{O(1)} n}$，也就是说，分支程序规模下界已接近最优，但相关性从指数级小增大到多项式级大。类似地，[CL23] 留下了在小相关性范围内针对SROLBP获得改进的平均情形困难性的问题。

我们注意到，方向仿射提取器是仿射非篡改提取器的一个特例，该提取器由 Chattopadhyay 和 Li 定义 [CL17]。粗略地说，仿射非篡改提取器是一种仿射提取器：即使在输入仿射源经过任意无不动点的仿射函数修改后，对应提取器输出的条件下，输出仍然接近均匀分布。在这一背景下，方向仿射提取器恰好对应于篡改函数向源加入非零仿射平移的情形。此前，由 Li 提出的最佳仿射非篡改提取器 [Li23] 适用于熵为 $k \geq (1-\gamma)n$ 的情形，其中某个较小的常数为 $\gamma < 1/3$，误差为$2^{-\Omega(n)}$。因此，这并没有给出更好的方向仿射提取器构造。然而，[Li23] 确实给出了一个改进的和集提取器，从而得到一个显式函数 Ext ，使得对于任意常数SROLBP$_\varepsilon$(Ext) $\geq 2^{n-O(\log n)}$， $\varepsilon > 0$即线性分支程序的规模下界在相差 $O$(。)中的常数范围内达到最优，但相关性增加到任意常数。

## 1.1 我们的结果

在本文中，我们对方向仿射提取器、仿射不可篡改提取器、SROLBP 和标准 ROBP 进行了更为深入的研究。首先，我们注意到，SROLBP 是否比标准 ROBP 更强大，并非先验上显而易见。事实上，很容易看出，$AC^0[\oplus]$ 和奇偶决策树分别比 $AC^0$ 电路和标准决策树强大指数倍，因为奇偶性需要指数规模的 $AC^0$ 电路和决策树。然而，任意奇偶函数都可以由规模为 $O(n)$ 的 ROBP 计算。尽管如此，已有工作 [Oko93、Juk95、GI17] 表明，计算某些仿射子空间的显式特征函数需要规模为 $2^{\Omega(n)}$ 的 ROBP（例如，[GI17]中的可满足Tseitin公式）。由于这类函数可以由规模为 $O(n)$ 的 SROLBP 轻松计算，这就提供了一个

SROLBP 与 ROBP 之间的分离，并表明 SROLBP 的确比 ROBP 强大指数倍。

这进一步表明，方向仿射提取器比标准仿射提取器具有更强的性质，因为它们能推出SROLBP的强下界。接下来，我们给出方向仿射提取器的显式构造，其参数远优于[GPT22]中的构造。我们的方法适用于任意线性熵，且误差为指数级小。

定理 1. 对任意常数 $0 < \delta \leq 1$，存在一族显式方向仿射提取器 DAExt : $\{0,1\}^n \to \{0,1\}^m$，其熵为 $k \geq \delta n$，误差为 $\varepsilon = 2^{-\Omega(n)}$，输出长度为$m = \Omega(n)$。

事实上，我们的构造也适用于略低于线性的熵。

定理 2. 存在一个常数 $c > 1$ 以及一族显式方向仿射提取器 DAExt : $\{0,1\}^n \to \{0,1\}^m$，其熵为 $k \geq cn(\log\log\log n)^2/\log\log n$，误差为 $\varepsilon = 2^{-n^{\Omega(1)}}$，输出长度为 $m = n^{\Omega(1)}$；此外，还存在一族显式方向仿射分散器，其熵为$k \geq cn(\log\log n)^2/\log n$，且具有 $m = n^{\Omega(1)}$。

This 该定理立即带来了 SROLBP 平均情形困难性的显著提升 s.

~定理 3：存在一个显式函数 DAExt，使得 SROLBP$_{2^{-n^{\Omega(1)}}}$(DAExt) $\geq 2^{n-O(\frac{n}{\log\log n})}$，~~其中~~ $\tilde{O}(.)$ 隐去了 $(\log\log\log n)^2$ 因子。

特别是，我们可以在获得 $2^{n-o(n)}$ 规模下界的同时，将相关性降至指数级小；该下界针对SROLBP，且几乎是最优的。这显著改进了 $2^{n/3-o(n)}$ 在 [GPT22] 中的规模下界，以及 [CL23]中的多项式级相关性。因此，定理 2 和 3 为此前 [GPT22] 和 [CL23] 中提到的两个开放问题给出了肯定答案。

我们指出，在我们的新定义下，方向仿射提取器严格强于标准仿射提取器。因此，定理 2 还改进了可忽略误差仿射提取器的熵要求：此前已知的最佳结果由 $\sqrt{\frac{n}{\log\log n}}$ [Yeh11、Li11] 给出，而我们将其改进为 $\frac{cn}{(\log\log\log n)^2\log\log n}$。

我们还重新审视了标准ROBP的困难性结果。如前所述，在这一模型中，显式函数已知具有指数级、甚至接近最优的规模下界；目前最好的结果是一个显式函数，它需要规模为 $2^{n-O(\log n)}$[Li23] 的ROBP（事实上是SROLBP）。不过，人们也一直对寻找低复杂度类中的函数感兴趣，因为这些函数能为ROBP带来强下界。显然，$NC^0$ 类并不足够。因此，下一个可能的类是$AC^0$。事实上，先前已有工作给出了显式$AC^0$ 函数，这些函数要求ROBP的规模为 $2^{\Omega(\sqrt{n})}$[Juk88、KMW91、G´al97、BW98]，甚至 $2^{\Omega(n)}$ [GI17],；然而，据我们所知，目前尚无平均情形困难性结果。本文通过给出一个显式$AC^0$ 函数来同时改进规模下界和平均情形困难性：对于任意常数 $\delta > 0$，该函数与规模为 $2^{(1-\delta)n}$ 的ROBP之间的相关性都可忽略。

定理 4. 对于任意常数 $\delta > 0$，存在一个显式函数 $AC^0$-Ext，属于 $AC^0$，使得 ROBP$_{2-\text{poly}\,\log n}$($AC^0$-Ext) $\geq 2^{(1-\delta)n}$。

我们构造中的关键组成部分之一，是一种针对仿射源的新型线性某处压缩器。具体而言，我们有

定义 7. 对于任意 $0 < \delta < \gamma < 1$，若函数 SCond : $\mathbb{F}_2^n \to (\mathbb{F}_2^m)^\ell$ 满足以下性质，则称其为一个（$\delta$, $\gamma$）仿射某处凝聚器：对于任意在 $\mathbb{F}_2^n$ 上、熵为 $\delta n$ 的仿射源 $X$，令（$Y_1$, $\cdots$, $Y_\ell$）= SCond($X$) $\in (\mathbb{F}_2^m)^\ell$，则至少存在一个 $i \in [\ell]$，使得 $Y_i$ 是 $\mathbb{F}_2^m$ 上熵至少为 $\gamma m$ 的仿射源。

定理 5. 存在一个常数 $\beta > 0$，使得对于任意 $0 < \delta \leq 1/2$，存在一个显式的（$\delta$，$1/2 + \beta$）仿射某处凝聚器 SCond : $\mathbb{F}_2^n \to (\mathbb{F}_2^m)^t$，其中 $t = \text{poly}(1/\delta)$ 且 $m = n/\text{poly}(1/\delta)$。此外，SCond 是一个线性函数。

我们进一步证明，在 $\mathbb{F}_2^n$ 上著名的多项式 Freiman-Ruzsa 定理成立的前提下，对这一压缩器稍作修改后，它也适用于一般的弱随机源。该定理是加性组合学中最重要的猜想之一，最近由 Gowers、Green、Manners 和 Tao 证明 [GGMT23]。详情请见第 4 节。

此前，此类压缩器都基于和积定理，而其函数是次数为 多项式$(1/\delta)$ 的多项式 [BKS+05, Raz05，Zuc07]。与此相对，已有线性含种子提取器的构造：若列出提取器在所有可能种子下的输出，就会得到一个某处随机源，其中至少有一个输出接近均匀，且其函数是线性函数。然而，在许多应用中（包括我们的应用），需要使用某处压缩器，而不是简单地列出提取器的所有输出，因为前者只产生少量（例如常数个）输出，而提取器会产生 多项式$(n)$ 个输出。因此，我们的线性某处压缩器补充了现有基于和积定理的某处压缩器。此外，我们的压缩器构造基于维度扩展器，这类代数伪随机对象此前因其自身的研究价值而受到关注；据我们所知，它们在计算机科学中尚无明确应用。因此，我们的构造可视为维度扩展器在计算机科学中的首批应用之一。

最后，我们研究这样一个问题：方向仿射提取器能否以黑盒方式为$\mathsf{AC}^0[\oplus]$ 类给出强下界。Cohen 和 Tal [CT15] 通过概率方法证明，标准仿射提取器并不足够，因为深度为 3 的 $\mathsf{AC}^0[\oplus]$ 电路可以计算最优仿射提取器。我们采用了对 [CT15], 中论证的略微修改版本，证明即便是更强版本的方向仿射提取器也不足够。具体而言，深度为 3 的$\mathsf{AC}^0[\oplus]$电路也可以计算最优方向仿射提取器。这进而证明了 $\mathsf{AC}^0[\oplus]$ 与SROLBP之间存在强分离。

定理 6. 存在一个函数 $f\colon \{0,1\}^n \to \{0,1\}$，它是熵为 $k$、误差为 $\varepsilon$的方向仿射提取器，其中 $k = \log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + O(1)$，并满足以下性质。

1. $f$ 是一个次数为 $\log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + O(1)$ 的多项式。
2. $f$ 可以由大小为 $O((n/\varepsilon)^2 \cdot \log^3(n/\varepsilon))$ 的异或-与-异或电路实现。
3. $f$ 可以由大小为 $O((n^5/\varepsilon^2) \cdot \log^3(n/\varepsilon))$ 的德摩根公式实现。

## 1.2 技术概述

这里概述本文所采用的主要思想。为清晰起见，我们在某些地方将采用非正式的表述，并忽略一些技术细节。

方向仿射提取器。 我们的出发点是 Li 构造的仿射提取器[Li11], ，该提取器适用于次线性熵，并具有指数级小误差。我们首先简要回顾其中的构造。将一个熵率为 $X$ 的仿射源 $\delta$ 划分为 $O(1/\delta)$ 个块。通过适当选择块的大小，可以证明存在一个熵率为 $\Omega(\delta)$ 的“良好”块 $X_g$，且在条件为定值时 $X$，源 $X_g$ 仍具有大量熵（即，我们得到一个仿射 块源）。如果我们知道 $X_g$的位置，随机性提取就很容易：我们应用某处压缩器（例如 [BKS+05、Raz05、Zuc07]中的压缩器）将其压缩 $X_g$ 为一个具有常数行数的矩阵，使得至少有一行具有熵率 $1 - \delta/2$。此时，我们可以应用

线性双源提取器（例如内积函数）到矩阵的每一行以及源 $X$，从而得到一个仿射 某处随机 源，在固定 $X_g$的条件下。这是另一个具有常数行数的矩阵，其中至少有一行是均匀的，因此可以应用现有技术，从该源中确定性地提取随机比特 [Rao09]。

然而，当 $\delta$ 较小时，我们不知道哪个块 $X_g$ 是良块。因此在 [Li11], 中，构造会尝试所有块，然后将它们组合起来。为了使这一过程能够奏效，该构造关键性地维持以下性质：**(*)** 对于每个块 $X_i$，由该块产生的输出比特是输入比特的常数次数多项式，并且次数从第一个块到最后一个块呈几何级数递减。利用这一性质，分析将聚焦于第一个良块 $X_g$。注意，我们可以固定 $X_g$之前各块产生的所有输出，而 $X_g$ 之后各块产生的所有输出的次数都低于 $X_g$中的次数。因此，如果我们将所有这些输出进行异或，多项式异或引理 [VW08、BKS+10] 保证最终输出仍接近均匀分布。我们注意到，多项式异或引理只适用于次数至多为 $\log n$的情形。因此，使每个块的输出次数 $c$ 尽可能小非常重要。粗略地说，我们需要 $c^{O(1/\delta)} < \log n$。

我们现在的策略是将这一构造改编为方向仿射提取器。为此，我们采用不可篡改提取器构造中的技术，因为正如前面所指出的，方向仿射提取器是仿射抗篡改提取器的一种特殊情况。近期的不可篡改提取器构造通常分为两个步骤：首先生成一段辅助信息，使其以高概率不同于篡改版本；然后将这段辅助信息与其他工具（例如相关性破坏器）结合起来，实现不可篡改性。因此，我们的目标是将这两个步骤改编到方向仿射提取器中，同时仍然保持性质（*），这对于实现线性熵或略低于线性的熵至关重要。下面我们将分别介绍这两个步骤。

和之前一样，对于每个块 $X_i$，我们都会得到一个输出 $U_i$；如果 $X_i$ 是良块，该输出就接近均匀分布。将 $U_i$ 分成两部分 $U_i = U_{i1} \circ U_{i2}$。我们将使用 $U_{i1}$ 生成辅助信息，并使用 $U_{i2}$完成构造的其余部分。注意，从篡改后的输入 $X' = X + a$ 中，我们也有一个篡改版本 $U'_i = U'_{i1} \circ U'_{i2}$。下文中，我们始终用带撇号的字母表示由篡改输入产生的相应随机变量。如果 $U_{i1} \neq U'_{i1}$，那么我们就完成了；否则，我们使用 $U_{i1} = U'_{i1}$ 从 $X$的编码中采样 $\Omega(\delta^2 n)$ 个比特 $H_i$，该编码采用渐近优良二元线性码。由于 $X' = X + a$，我们有 $H_i + H'_i$ 基本上对应于从 $a$的编码中采样得到的比特。因此，根据线性码的距离性质，$H_i \neq H'_i$ 以高概率成立。然而，我们不能直接进行朴素采样，因为我们需要保持次数为常数。因此，我们还将 $U_{i1}$ 和 $X$ 的编码分别划分为 $\Omega(\delta^2 n)$ 个块，每个块包含常数个比特，并使用 $U_{i1}$ 的每个块从 $X$编码中对应的块采样一个比特。根据该码的距离性质，$X$ 的编码与 $X'$ 的编码至少有 $\Omega(\delta^2 n)$ 个块不同。因此，$H_i \neq H'_i$仍以高概率成立；现在，$H_i$ 的每个比特都是 $U_{i1}$ 和$X$中比特的常数次数多项式。现在，辅助信息串为 $U_{i1} \circ H_i$。

有了辅助信息后，我们可以使用线性有种子提取器将它附加到另一个从 $X$ 中提取出的字符串上，并以 $U_{i2}$ 作为种子。现在注意到，以高概率，从 $X$ 生成的字符串不同于从 $X'$ 生成的字符串；并且在固定 $(U_i, U'_i$ 的条件下，它们呈线性相关。因此，我们可以应用例如一个已知的仿射抗篡改提取器（最先进的、具有可忽略误差的仿射抗篡改提取器仅适用于高熵情形）。然而，[CL17] 中的已知仿射抗篡改提取器构造具有超常数次数。事实上，即使只应用一次该提取器，得到的多项式次数也会大于 $\log n$，这已经无法实现我们构造方向仿射提取器的目的（不过，我们仍然可以构造方向仿射弥散器）。

为解决这个问题 在这个问题中，我们提出了利用 $X' = X + a$特殊结构的新思路 .

回顾我们的构造，对于每个块 $X_i$，我们都会得到一个 $U_{i2}$；如果 $X_i$良好，则它接近均匀，而在以 $X_i$为条件时， $X$ 仍具有足够的熵。现在，我们的想法是改用带种子非篡改提取器 snmExt。它是一种使用均匀随机种子的提取器，具有如下性质：如果对手篡改种子但不篡改源，那么在原始输入上运行提取器所得的输出，在给定篡改后输入上的输出时，接近均匀。将~ ~ ~辅助信息串附加到 $U_{i2}$ 上并得到 $U_i = U_i \circ H_i$后，以高概率有 $U_i \neq U_i'$，并且如果 $H_i$ 规模较小，种子$U_i$ 具有高熵；只要提取器是强提取器，这就足以应用带种子非篡改提取器。现在，如果带种子非篡改提取器在以种子的任意固定值为条件时也是线性的，那么我们有 $\mathsf{snmExt}(X', U_i') = \mathsf{snmExt}(X, U_i') + \mathsf{snmExt}(a, U_i')$。由于 ~ ~$\mathsf{snmExt}(X, U_i)$ 在给定 $\mathsf{snmExt}(X, U_i')$ 时接近均匀，并且该提取器是强提取器（我们可以固定种子（$\tilde{U}_i$， $\tilde{U}_i'$）），这意味着 $\mathsf{snmExt}(X, \tilde{U}_i)$ 在给定 $\mathsf{snmExt}(X', \tilde{U}_i')$ 时接近均匀。[2]

幸运的是，Li 已经给出了带种子线性不可篡改提取器的构造[Li12], ，其基于内积函数。此外，该提取器还具有以下性质：每个输出比特都是输入比特的常数次数多项式。因此，除一个问题外，一切似乎都能顺利实现： [Li12] 中的不可篡改提取器仅在源具有 $> 1/2$ 的熵率时才有效，但这里我们的目标是适用于任意线性（或略低于线性）的熵。一种自然的想法是使用某处压缩器（例如 [BKS+05、Raz05、Zuc07]）来提高 $X$ 的熵率。然而，所有已知的此类压缩器都基于和积定理，而和积定理对应的是非线性函数；应用它们会改变 $X' = X + a$ 的结构，而该结构对我们的构造十分重要。另一种想法是对 $X$ 应用线性有种子提取器，并尝试所有可能的种子。这样确实保留了 $X' = X + a$ 的结构，但会产生 $\mathsf{poly}(n)$ 个输出，将它们组合起来则会得到一个次数很大的超常数次数多项式。

这促使我们在构造中引入另一个关键要素：一种针对仿射源的新型线性某处凝聚器。简而言之，我们构造了一个线性函数，对于 $n$ 比特上熵率为 $0 < \delta \leq 1/2$ 的任意仿射源，该函数输出 $\mathsf{poly}(1/\delta)$ 行，每行包含 $n/\mathsf{poly}(1/\delta)$ 比特，并且至少有一行的熵率为 $1/2 + \beta$，其中 $\beta > 0$ 是某个绝对常数。这补充了基于和积的某处压缩器，也可视为我们工作的另一项独立贡献。我们稍后会说明该压缩器的构造；现在先假设我们已有这个线性某处凝聚器，并完成对方向仿射提取器的描述。

其余构造大致如下。我们将线性某处凝聚器~应用于源 $X$，得到常数行数，然后以 snmExt 为每一行提取，使用 $U_i$ 作为种子。这样，我们会得到常数个输出，其中至少有一个在对应的篡改输出条件下接近均匀。接下来，我们应用诸如 仿射相关性破坏器中的 [Li17、CGL22、CL22] 等仿射相关性破坏器，进一步打破不同输出之间的相关性，并通过异或将这些输出合并。相关性破坏器保证最终输出在篡改输出条件下接近均匀。为保持次数较小，我们需要将相关性破坏器中使用的所有种子提取器替换为 [Li11]中的常数次数线性有种子提取器。这样，输出比特就是输入比特的常数次数多项式，其余构造则与 [Li11]中的基本相同。

线性某处凝聚器。现在我们描述线性某处凝聚器的构造。这一构造基于另一种称为维度扩展器的伪随机对象。非正式地说，维度扩展器是从向量空间 $\mathbf{F}^n$ 到其自身的一组线性映射，使得对于任意维数 $k \leq n/2$ 较小的线性子空间 $V \subset F^n$，所有 $V$ 的像的并集所张成的张成空间

[2]实际分析涉及更多细节，因为这里的 $X$ 并不独立于（$\tilde{U}_i$、$\tilde{U}_i'$），但由于仿射结构，该性质仍然成立。我们在此省略细节。

在这组线性映射下，其维数至少为 $(1+\alpha)k$，其中某个绝对常数为 $\alpha > 0$。熟悉扩展图的读者可以看出，这是扩展图的线性代数类比。因此，我们希望给出这组线性映射的显式构造，并使映射的数量尽可能少，其中这一数量 $d$ 称为次数。维度扩展器最早由 Barak、Impagliazzo、Shpilka 和 Wigderson [BISW04], 引入，他们还证明了这类对象的存在性。后来，Bourgain 和 Yehudayoff [Bou09、BY13] 给出了任意域上次数为 $d = O(1)$ 的维度扩展器的显式构造。有趣的是，据我们所知，维度扩展器此前在计算机科学中没有任何应用，它们主要因自身的研究价值以及与其他代数伪随机对象的联系而受到研究。因此，我们的构造可以被视为维度扩展器在计算机科学中的首批应用之一。

给定一个显式的维度扩展器 $\{T_i\}_{i\in[d]}$，其中每个 $T_i$ 都是一个线性映射，对于任意熵率为 $\delta \le 1/2$的仿射源 $X$，我们首先构造如下基本的某处压缩器。将 $X$ 等分为 $X = X_1 \circ X_2$，我们的压缩器产生 $2d+2$ 个输出：$(X_1, X_2, \{X_1 + T_i(X_2)\}_{i\in[d]}, \{T_i(X_1) + X_2\}_{i\in[d]})$。我们证明，至少有一个输出的熵率为 $(1+\gamma)\delta$，其中 $\gamma > 0$为某个常数，下面给出一些直观解释。根据仿射源的结构，可以证明存在另一个与 $X_1$ 独立的仿射源 $X_3$，使得对于某个线性函数 $L$，有 $X_2 = X_3 + L(X_1)$。令 $H(X_1) = s$、$H(X_3) = r$且 $H(L(X_1)) = t$，则有 $s + r = \delta n$。如果 $s$ 或 $r$ 较小，例如 $s \ll \delta n/2$，那么必有 $r \gg \delta n/2$，从而 $H(X_2) = r + t \ge (1+\gamma)\delta n/2$。因此，$X_2$ 的熵率至少为 $(1+\gamma)\delta$。$r \ll \delta n/2$ 的情形类似。因此，我们只需考虑 $s \approx \delta n/2$ 且 $r \approx \delta n/2$的情形，并注意到必有 $s \le \delta n/2$ 或 $r \le \delta n/2$。此外，在这种情况下，$t$ 必须较小，否则我们将再次得到 $H(X_2) = r + t \ge (1+\gamma)\delta n/2$。

为简便起见，假设 $s = r = \delta n/2$，且 $t = 0$。因此，$X_1$ 和 $X_2$ 的熵率均为 $\delta \le 1/2$，并且二者相互独立。不失一般性地，假设 $X_1$和 $X_2$ 的支撑集均为线性子空间。根据维度扩展器的性质，$\mathrm{Span}(\cup_{i\in[d]} T_i(X_1))$ 的维数至少为 $(1+\alpha)\delta n/2$。现在我们论证存在一个 $i \in [d]$，使得 $T_i(X_1) + X_2$ 的支撑集维数至少为 $(1+\alpha/d)\delta n/2$，这意味着 $T_i(X_1) + X_2$的熵率至少为 $(1+\alpha/d)\delta$。为此，反设结论不成立，则对于任意 $i \in [d]$，$T_i(X_1) + X_2$ 的支撑集中的任意向量都可以表示为 $r = \delta n/2$ 个基向量与支撑集中的 $X_2$ 和 $<(\alpha/d)\delta n/2$ 个其他向量的线性组合。这意味着 $\mathrm{Span}(\cup_{i\in[d]} T_i(X_1))$ 的维数为 $< \delta n/2 + d \cdot (\alpha/d)\delta n/2 = (1+\alpha)\delta n/2$，因为 $\mathrm{Span}(\cup_{i\in[d]} T_i(X_1))$ 中的任意向量都可以表示为 $r = \delta n/2$ 个基向量与支撑集中的 $X_2$ 和 $< d \cdot (\alpha/d)\delta n/2$ 个其他向量的线性组合。这与维度扩展器的性质矛盾。

因此，在所有情况下，我们都能获得所需的熵率提升。我们的最终某处压缩器涉及反复使用基本凝聚器，这与以往的工作相同。不难看出，经过 $O(\log(1/\delta))$ 次基本凝聚器的使用后，至少有一个输出的熵率会提升至 $1/2 + \beta$，其中 $\beta > 0$ 为某个绝对常数。因此，输出的数量为 $\mathrm{poly}(1/\delta)$，且每个输出包含 $n/\mathrm{poly}(1/\delta)$ 个比特。最后，很明显，该压缩器是一个线性函数。

有了这个线性浓缩器后，我们甚至可以用新的浓缩器替换 [Li11] 中使用的某处压缩器。这会进一步降低输出比特对应多项式的次数（因为之前的某处压缩器是多项式，而不是线性函数）。因此，我们可以将方向仿射提取器的熵要求进一步提升到优于 [Li11], 的水平，从 $\sqrt[n]{\log\log n}$ 降至 $\frac{cn}{(\log\log\log n)^2 \log\log n}$。

我们证明，在多项式 Freiman-Ruzsa 定理成立的条件下，对线性浓缩器稍作修改，也适用于一般的弱随机源。粗略而言，我们的思路是仔细分析子源和碰撞概率。具体来说，已知如果某个分布的碰撞概率

较小，那么该分布接近于具有高最小熵。另一方面，如果碰撞概率较大，那么（不失一般性地）假设该分布是某个未知子集上的均匀分布，加性组合学中的已有结果表明，支撑集中存在一个较大的子集 $A$，使得 $A+A$的大小不比 $A$大太多。接着，多项式 Freiman-Ruzsa 定理推出还存在另一个较大的子集 $A' \subset A$，它与某个仿射子空间“接近”，从而大致上将分析简化为仿射源的情形。详情见第4 节。

$AC^0$ 对 ROBP 的平均情形困难性。为了证明 $AC^0$ 对 ROBP 的平均情形困难性，我们采用一个标准观察：若将某个内部节点条件化，那么该节点之前的输入比特与该节点之后的输入比特仍然相互独立。接着，我们为具有这种结构的源构造一个合适的提取器，记作 $AC^0$, 的 $^0$-Ext。具体而言，给定任意一个规模为 $s$ 的 ROBP 和任意常数 $\delta > 0$，我们可以找到一个大小为 $O(s)$ 的割集或反链（一个极大顶点子集，其中任意顶点都不是其他任何顶点的祖先），其位置大致在汇点上方深度 $\delta n$ 处。这样，在将割集中任意顶点固定后，均匀随机输入串 $X$ 就变成了两个相互独立的弱源 $A$ 和 $B$，其中 $A$ 对应程序的第一部分，$B$ 对应第二部分。由于我们不知道 ROBP 查询比特的顺序，这两个源中的比特彼此交错，因此我们将 $X = A + B$ 视为.....利用标准的平均论证，可以证明以下性质以高概率成立：(1) $A$ 和 $B$ 支撑在互不相交的输入比特子集上；(2) $A$ 的最小熵约为 $(1-\delta)n - \log s$，而 $B$ 的最小熵为 $\delta n$；(3) $B$ 是一个无关比特固定源，它是通过固定均匀随机串中的一些未知比特得到的。如果 $s \le 2^{(1-2\delta)n}$，那么 $A$ 和 $B$ 的熵率都约为 $\delta$。现在，我们的目标是为具有这种结构的源构造一个 $AC^0$ 提取器，并且该提取器对 $B$ 也是强的。这意味着，即使我们将割集中顶点的取值固定，并且 $B$，提取器的输出仍接近均匀分布。另一方面，ROBP 的输出完全由该顶点和 $B$ 决定。因此，对于规模不超过 $2^{(1-2\delta)n}$ 的 ROBP，我们的提取器具有平均情形困难性。

通常，函数 $AC^0$-Ext 将由不同的、更基础的提取器组合而成，并将它们作为构件。因此，我们需要所有这些构件都能在 $AC^0$中计算。这里，我们借鉴了先前两项关于 $AC^0$ 中提取器的研究所给出的构造：(1) Cheng 和 Li 提出的适用于比特固定源的 $AC^0$-可计算提取器$AC^0$-BFExt [CL18], ，以及 (2) Papakonstantinou、Woodruff 和 Yang 提出的 $AC^0$-可计算强线性有种子提取器 $AC^0$-LExt [PWY16]。

现在我们可以描述构造的主要思路。将 $X$ 划分为 $t = O(1/\delta)$ 个块。根据平均论证，存在一个块 $B_g$，其熵率为 $\Omega(\delta)$。对于块$X_g = A_g + B_g$，我们可以固定 $A_g$，使得 $X_g$ 是熵率为 $\Omega(\delta)$ 的无关比特固定源，并且是 $B$的确定性函数。接下来，我们固定 $B$ 中位于第 $g$块之外的比特，使得源 $X$ 中位于 $X_g$ 之外的部分是 $A$ 的确定性函数，因此与 $X_g$独立。此外，$A$和 $X$ 仍有足够的熵。

将上述用于比特固定源的提取器 $AC^0$-BFExt 应用于每个块 $X_i$，我们将 $X$ 转换为某处随机源 $Y = Y_1 \circ \cdots \circ Y_t$，其中行 $Y_g$ 是 $B_g$ 的确定性函数，且接近均匀，而其他所有行都是 $A$ 的确定性函数。此时，我们只需对各个 $Y_i$取异或，即可得到接近均匀的输出。然而，正如前面所述，我们需要提取器在 $B$ 上具有强性质，而这一简单方法并不足够。相反，我们固定 $X_i$ 对应的 $AC^0$-BFExt 产生的所有输出，其中 $i \neq g$。注意，这些输出都是 $A$的确定性函数。因此，在对此进行固定条件化后，$Y$ 就成为 $B$的确定性函数，而后者独立于 $A$。此外，只要 $AC^0$-BFExt 的输出长度不太大，$A$ 仍有足够的熵。由于 $X = A + B$，现在我们可以应用一个强的 $t$-仿射相关性

破坏器，如 [Li17 中的做法，以每个 $Y_i$ 作为种子，从 $X$ 中提取一个随机字符串，并对这些字符串取异或。相关性破坏器的性质保证，由 $Y_g$和 $X$ 生成的字符串在给定所有其他输出和 $Y$的条件下接近均匀。因此，该异或结果在给定 $B$的条件下也接近均匀。为确保相关性破坏器可在 $\mathsf{AC}^0$, 中计算，我们将已知构造中 $t$-仿射相关性破坏器所用的所有强（线性）种子提取器替换为上述 $\mathsf{AC}^0$-LExt。由于 $t = O(1/\delta)$ 是常数，相关性破坏器只涉及常数次 $\mathsf{AC}^0$-LExt 的复合运算，因此仍属于 $\mathsf{AC}^0$。

### 1.3 论文组织

本文其余部分的组织如下。在第2 节中，我们介绍一些预备知识以及已有工作中的一些基本工具。在第3 节中，我们描述针对仿射源的线性某处随机压缩器的构造。第4节基于多项式 Freiman-Ruzsa 定理，将该构造推广到一般弱源。我们在第5节给出方向仿射提取器的构造，并在第$\mathsf{AC}^0$ 节给出针对ROBP 的可计算提取器6。我们在第7节提出一些开放问题。在附录中，我们证明深度为 3 的$\mathsf{AC}^0[\oplus]$ 电路可以计算最优方向仿射提取器，并给出一些省略的证明。

## 2 预备知识

我们通常使用大写字母表示随机变量，并使用相应的小写字母表示它们的取值。设 $s$、$t$ 为两个整数，$\{V_1^1, V_1^2, \cdots, V_1^t, V_2^1, V_2^2, \cdots, V_2^t, \cdots, V_s^1, V_s^2, \cdots, V_s^t\}$ 为一组随机变量。我们用 $V_i^{[t]}$ 表示子集 $\{V_i^1, \cdots, V_i^t\}$，用 $V_{[s]}^j$ 表示子集 $\{V_1^j, \cdots, V_s^j\}$。我们使用 $V_{[s]}^{[t]}$ 作为整组随机变量的简写。我们还用 $i_{[t]}$ 表示索引集合 $\{i_1, i_2, \cdots, i_t\}$。设 $|S|$ 表示集合 $S$ 的基数。对于 $\ell$ 这一正整数，$U_\ell$ 表示 $\{0,1\}^\ell$ 上的均匀分布。当作为向量中的一个分量使用时，假设每个 $U_\ell$ 都与其他分量独立。

令 $\mathbb{F}_q$ 表示大小为 $q$的有限域。所有对数均以 2 为底。

### 2.1 概率分布与熵

定义 8（统计距离）。设 $W$ 和 $Z$ 是集合 $S$上的两个分布。它们的统计距离（变差距离）为

$$\Delta(W, Z) := \max_{T \subseteq S}(|W(T) - Z(T)|) = \frac{1}{2}\sum_{s \in S}|W(s) - Z(s)|.$$

若 $W$ 与 $Z$的统计距离为 $\varepsilon$-close，则称 $W$ 是 $\varepsilon$-close 于 $Z$的，记为 $W \approx_\varepsilon Z$，其中 $\Delta(W, Z) \leq \varepsilon$。设 $V$ 也是集合 $S$上的一个分布。有时我们将 $W \approx_\varepsilon Z \mid V$作为 $(W, V) \approx_\varepsilon (Z, V)$ 的简写。在全文中，我们将交替使用这两种记法。对于集合 $S$ 上的分布 $D$ 以及函数 $h: S \to T$，令 $h(D)$ 表示如下分布：按照 $D$选择 $x$，并输出 $h(x)$，所得结果为 $T$ 上的分布。

引理 7。对于任意函数 $\alpha$和两个随机变量 $A$、$B$，有 $\Delta(\alpha(A)\ \alpha(B)) \leq \Delta(A, B)$。定义 9（最小熵）。随机变量 $X$ 的最小熵定义为

$$H_\infty(X) = \min_{x \in \mathsf{Supp}(X)}\{-\log \mathbf{Pr}[X = x]\}.$$

对于随机变量 $X \in \{0,1\}^n$，如果 $H_\infty(X) \geq k$，则称其为 $(n, k)$-源。$X$ 的熵率定义为 $H_\infty(X)/n$。

## 2.2 某处随机源与提取器

定义 10（某处随机源）。如果源 $X = (X_1, \cdots, X_t)$ 满足：每个 $X_i$ 取值于 $\{0,1\}^r$，并且存在一个 $i$ 使得 $X_i$ 服从均匀分布，则称其为 $(t \times r)$ 某处随机源（简称 SR 源）。

定义 11. 初等某处 $k$-源是由源 $(X_1, \cdots, X_t)$ 构成的向量，其中某个 $X_i$ 是一个 $k$-源。某处 $k$-源是初等某处 k-源的凸组合。

定义 12. 如果对于每个 $k$-源 $X$，都有 $C(X, U_d)$ 与某个 $\ell$-源的距离为 $\varepsilon$-接近，则函数 $C$: $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 是一个 $(k \to \ell, \varepsilon)$-压缩器。在方便时，我们称 $C$ 为速率为 $(k/n \to \ell/m, \varepsilon)$ 的压缩器。

定义 13. 如果对于每个 $k$-源 $X$，向量 $(C(X, y)_{y \in \{0,1\}^d})$ 与某个某处-$\ell$-源的距离为 $\varepsilon$-接近，则函数 $C$: $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 是一个 $(k \to \ell, \varepsilon)$-某处压缩器。在方便时，我们称 $C$ 为速率为 $(k/n \to \ell/m, \varepsilon)$ 的某处压缩器。

定义 14（带种子提取器）。若函数 Ext : $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 满足：对于每个最小熵为 $k$ 的源 $X$，以及独立且在 $\{0,1\}^d$ 上均匀分布的 $Y$，则称其为强 $(k, \varepsilon)$- 提取器，

$$(\mathsf{Ext}(X, Y), Y) \approx_\varepsilon (U_m, Y).$$

## 2.3 仿射源的结构

本文中，仿射源涵盖线性子空间上的均匀分布；有时我们所说的仿射函数指的是仿射线性函数。

定义 15（仿射源）。设 $\mathbb{F}_q$ 是含有 $q$ 个元素的有限域。记 $\mathbb{F}_q^n$ 为 $\mathbb{F}_q$ 上的 $n$ 维向量空间。若存在 $\mathbb{F}_q^n$ 中的线性无关向量 $a_1, \cdots, a_k \in$，以及 $\mathbb{F}_1^n$ 中的另一个向量 $b \in$，使得分布 $X$ 是 $\mathbb{F}_q^n$ 上的一个 $(n, k)_q$ 仿射源，当且仅当它可通过如下方式采样得到：在 $\mathbb{F}$ 中均匀且独立地选取 $x_1, \cdots, x_k \in$，并计算

$$X = \sum_{i=1}^{k} x_i a_i + b.$$

仿射源的最小熵与其标准香农熵相同，我们简单地用$H(X)$表示仿射源 $X$的熵。

下面的引理是 [Li11]，中相应版本的略微推广，其中我们证明了 $L$可以是仿射函数，而不仅仅是线性函数。我们还证明，在 $X$ 的任意固定条件下， $L(X)$ 的熵是常数。证明请参见附录B 。

引理 8 （仿射条件化 [Li11]）。设 $X$ 是 $\{0,1\}^n$上的任意仿射源。设 $L$： $\{0,1\}^n \to \{0,1\}^m$ 是任意仿射函数。那么存在独立仿射源 $A$、$B$ 使得：

- $X = A + B$
- 存在 $c \in \{0,1\}^m$，使得对于每个 $b \in \mathsf{Supp}(B)$，都有 $L(b) = c$。
- $H(A) = H(L(A))$，并且存在仿射函数 $L^{-1}$： $\{0,1\}^m \to \{0,1\}^n$ 使得$A = L^{-1}(L(A))$。
- $H(X|_{L(X)=\ell}) = H(B)$ 对所有 $\ell \in \mathsf{Supp}(L(X))$。

以下定义是针对仿射源的条件最小熵的特例。根据引理，该定义是良定义的8。

定义 16（仿射源的条件最小熵）。设 $W$ 和 $Z$ 为两个仿射源。定义

$$H(W \mid Z) = H(W \mid_{Z=z}),\ \forall z \in \mathsf{Supp}(Z).$$

引理 9 。设 $X$、$Y$、$Z$ 为仿射源。则 $H(X \mid (Y,Z)) \geq H(X \mid Z) - \log(\mathsf{Supp}(Y))$。

当我们对仿射源的块进行顺序条件化，或论证仿射源各块的总熵时，还需要以下来自[Li11] 的引理。

引理 10 （仿射熵论证 [Li11]）。设 $X$ 为定义在 $\{0,1\}^n$上的任意仿射源。将 $X$划分为 $t$ 个任意块 $X = X_1 \circ X_2 \circ \cdots \circ X_t$。则存在正整数 $k_1$、$\cdots$、$k_t$，使得：

- $\forall j, 1 \leq j \leq t \ and \ \forall (x_1, \cdots, x_{j-1}) \in \mathsf{Supp}(X_1, \cdots, X_{j-1}),\ H(X_j \mid_{X_1=x_1,\cdots,X_{j-1}=x_{j-1}}) = k_j;$
- $\sum_{i=1}^{t} k_i = H(X).$

## 2.4 平均条件最小熵与平均情形有种子提取器

定义 17（平均条件最小熵）。平均条件最小熵定义为

$$\begin{aligned}\widetilde{H}_\infty(X \mid W) &= -\log\left(\mathbf{E}_{w \leftarrow W}\left[\max_x \mathbf{Pr}[X = x \mid W = w]\right]\right) \\ &= -\log\left(\mathbf{E}_{w \leftarrow W}\left[2^{-H_\infty(X|W=w)}\right]\right).\end{aligned}$$

引理 11（[DO RS08]）。对于任意 $s > 0$，$\mathbf{Pr}_{w \leftarrow W}[H_\infty(X \mid W = w) \geq H_\infty(X \mid W) - s] \geq 1 - 2^{-s}$。

引理 12 （[DORS08]）。如果随机变量 $B$ 至多有 $2^\ell$ 个可能值，则 $\widetilde{H}_\infty(A \mid B) \geq H_\infty(A) - \ell$。

引理 13 （[DORS08]）。对于任意 $\delta > 0$，若 Ext 是一个 $(k, \varepsilon)$ 提取器，则它也是一个 $(k+\log(1/\delta), \varepsilon+\delta)$ 平均情形提取器。

## 2.5 交替提取与独立性合并

以下技术是构造相关性破坏器的基础。

定义 18（$L$-交替提取）。设 $W$是一个 $(n_w, k_w)$-源，且 $(Q_1, \cdots, Q_L)$ 是 $L$个 $(n_q, k_q)$-源。设 $\mathsf{Ext}_q$、$\mathsf{Ext}_w$是强种子提取器，能够从最小熵为 $k$ 、误差为 $\varepsilon$ 、种子长度为 $s$的源中提取 $s$比特。设 $S_1 = \mathsf{Slice}(Q_1, d)$，其中长度$d$适当，$R_1 = \mathsf{Ext}_w(W, S_1)$ $S_2 = \mathsf{Ext}_q(Q_2, R_1)$ $\cdots$，$R_{L-1} = \mathsf{Ext}_w(W, S_{\ell-1})$ $S_L = \mathsf{Ext}_q(Q_L, R_{L-1})$，则$L$-交替提取$(Q_1, \cdots, Q_L, W) = S_L$。

引理 14（前瞻提取器[CGL16]）。设 $W$是一个 $(n_w, k_w)$-源，且 $W'$是在 $\{0,1\}^{n_w}$上的一个与 $W$任意相关的随机变量。设 $Y = (Q, S_1)$，其中 $Q$是一个 $(n_q, k_q)$-源，$S_1$是一个长度为 $s$ 比特的均匀字符串，并且 $Y' = (Q', S_1')$是一个与 $Y$任意相关的随机变量，其中 $Q'$ 和 $S_1'$ 分别是在 $n_q$ 比特和 $s$ 比特上的随机变量。设 $\mathsf{Ext}_q$、$\mathsf{Ext}_w$ 是强种子提取器，它们从最小熵为 $k$ 、误差为

$\varepsilon$ 且种子长度为 $s$的源中提取 $0 \le j \le \ell - 1$ 比特。假设（$Y$， $Y'$）与（$W$， $W'$）独立，且 $k_w$、$k_q \ge k + 2\ell s + 2\log(1/\varepsilon)$。令 laExt 为使用 Ext$_q$、Ext$_w$的 $\ell$轮前瞻提取器，并且（$R_1$， $\cdots$， $R_\ell$）$=$laExt$_\ell$（$W$， $Y$），（$R'_1$， $\cdots$， $R'_\ell$）$=$laExt$_\ell$（$W'$， $Y'$）。那么对于任意 $0 \le j \le \ell - 1$，我们有

$$R_{j+1} \approx_{O(\ell\varepsilon)} U_s \mid (Y, Y', R_0, R'_0, \cdots, R_j, R'_j).$$

以下引理概括了翻转器和非交互式概率合并器构造中的一个关键论证，而它们是相关性破坏器的组成部分。

引理 15（独立性合并引理 [CGL22]）。设 Ext：$\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 是任意一个（$k$， $\varepsilon$）强种子提取器， $X$、$X^{[t]} \in \{0,1\}^n$、 $Y$、$Y^{[t]} \in \{0,1\}^d$ 满足：对于每个 $j \in [t]$， $X$、$X^{[t]}$ 与 $Y$、$Y^{[t]} \in \{0,1\}^d$、 $W =$、Ext($X$， $Y$) 以及 $W^j =$Ext($X^j$， $Y^j$) 相互独立。假设存在 $S$、$T \subseteq [t]$ 使得

- $(Y, Y^S) \approx_\delta (U_d, Y^S)$;
- $\widetilde{H}_\infty(X \mid X^T, Z) \ge k + tm + \log(1/\varepsilon)$.

*Then*

$$W \approx_{2\varepsilon+\delta} U_m \mid (W^{S\cup T}, Y, Y^{[t]}).$$

## 2. 6 $\varepsilon$-偏差空间与异或引理

本小节中的工具用于 [Li11] 的仿射分散器和提取器构造。我们还在方向仿射分散器和提取器的构造中采用了这些技术。

定义 19（$\varepsilon$-偏差空间）。若随机变量 $Z$ 在 $\{0,1\}$ 上满足 $|\mathbf{Pr}[Z = 0] - \mathbf{Pr}[Z = 1]| \le \varepsilon$，则称其为 $\varepsilon$-偏差的。若随机变量序列 $0-1$ $Z_1, \cdots, Z_m$ 对线性测试是 $\varepsilon$-偏差的，则对于任意非空集合 $S \subset [m]$，随机变量 $Z_S = \bigoplus_{i\in S} Z_i$ 都是 $\varepsilon$-偏差的。

引理 16 ([Vaz86])。设 $Z_1$， $\cdots$， $Z_m$是 $0-1$对线性测试呈 $\varepsilon$-偏差的随机变量。因此，($Z_1$， $\cdots$， $Z_m$) 的分布 $\varepsilon \cdot 2^{m/2}$与均匀分布相差至多

定义 20。对于两个函数 $f$、$p$：$\{0,1\}^n \to \{0,1\}$，它们在均匀分布下的相关性定义为

$$\mathsf{Cor}(f, p) = \Big|\mathbf{P}\mathbf{r}_x[f(x) = p(x)] - \mathbf{P}\mathbf{r}_x[f(x) \ne p(x)]\Big|,$$

其中概率取自均匀分布。对于函数类 $C$，我们将 Cor($f$, $C$) 定义为所有定义域为 $\{0,1\}^n$ 的函数 $p \in C$ 中 Cor($f$, $p$) 的最大值。

定理17 （关于 $\mathsf{F}_2$ [VW08 上多项式的BKS+10]异或引理）。设 $P_d$ 表示次数至多为 $d$ 且定义在 $\mathsf{F}_2$上的所有多项式构成的类。设 $f$： $\{0,1\}^n \to \{0,1\}$ 是一个函数，使得$\mathsf{Cor}(f, P_d) \le 1 - 2^{-d}$并且 $f^{\oplus m}$是函数 $f$在 $m$个独立输入上的取值的异或。那么

$$\mathsf{Cor}(f^{\oplus m}, P_d) \le \exp(-\Omega(m/(4^d \cdot d))).$$

## 3 仿射源的线性某处压缩器

本节中，我们给出一种针对仿射源的线性某处压缩器的显式构造；或者更方便地说，构造一种仿射某处压缩器，其中每个输出都是输入的线性函数。我们首先给出定义。

定义 21. 对任意 $0 < \delta < \gamma < 1$，函数 $\mathsf{SCond} : \mathbb{F}_2^n \to (\mathbb{F}_2^m)^\ell$ 是一个（$\delta, \gamma$）仿射某处压缩器，若其满足以下性质：对于任意熵为 $\delta n$ 的 $\mathbb{F}_2^n$ 上的仿射源 $X$，令 $(Y_1, \cdots, Y_\ell) = \mathsf{SCond}(X) \in (\mathbb{F}_2^m)^\ell$，则至少存在一个 $i \in [\ell]$，使得 $Y_i$ 是 $\mathbb{F}_2^m$ 上熵至少为 $\gamma m$ 的仿射源。

我们将证明以下定理。

定理 18. 存在一个常数 $\beta > 0$，使得对于任意 $0 < \delta \leq 1/2$，都存在一个显式的（$\delta, 1/2 + \beta$）仿射某处压缩器 $\mathsf{SCond} : \mathbb{F}_2^n \to (\mathbb{F}_2^m)^t$，其中 $t = \mathrm{poly}(1/\delta)$ 且 $m = n/\mathrm{poly}(1/\delta)$。此外，SCond 是一个线性函数。

为证明该定理，我们将使用以下称为维度扩展器的对象。

定义 22 (维度扩展器 [BISW04、DS11])。 设 $\mathbb{F}$ 是一个域，并设 $T_1, \cdots, T_d : \mathbb{F}^n \to \mathbb{F}^n$是线性映射。集合 $T = \{T_i\}_{i=1}^d$ 是一个 $\alpha$-维度扩展器，其次数为 $d$，若对于每个子空间 $V \subset \mathbb{F}^n$ 维数至多为 $n/2$，都有

$$\dim\left(\sum_{i=1}^{d} T_i(V)\right) \geq (1+\alpha)\dim(V).$$

我们称 $T$ 是显式的，如果存在一个 $\mathrm{poly}(n)$ 时间算法，该算法以 $n$ 为输入并输出 $T$。

定理 19 ([Bou09、BY13])。 存在绝对常数 $d \in \mathbb{N}$ 和 $0 < \alpha < 1$，使得对于任意域 $\mathbb{F}$，都存在一个显式的 $\alpha$-维度扩展器，其次数为 $d$。

根据上述定理，我们首先给出一个基本仿射压缩器：

---

算法 1 BasicCond($x$)

---

输入： $x \in \mathbb{F}_2^n$——一个 $n$ 比特串。输出： $z \in (\mathbb{F}_2^m)^{2d+2}$——一个由 $2d+2$ 比特串组成的数组，每个比特串长度为 $m$，其中 $m = n/2$ 且 $d$ 是定理 19中的常数。

---

子程序和参数：
令 $T = \{T_i\}_{i=1}^d$ 为定理 $\alpha$19给出的 维数扩展器。

---

将 $x$ 分成 2 个块 $x = x_1 \circ x_2$，其中每个块包含 $n/2$ 比特。
令 $z = z_1 \circ z_2 \circ \cdots \circ z_{2d+2}$，其中 $z_1 = x_1$、 $z_2 = x_2$ 和 $z_{2i+1} = x_1 + T_i(x_2)$、 $z_{2i+2} = x_2 + T_i(x_1)$，对于任意 $i \in [d]$。

---

我们将证明以下引理。

引理 20。 对于任意 $0 < \delta \leq 1/2$， BasicCond 是一个 $(\delta, \ (1+\frac{\alpha}{4d})\delta)$ 仿射某处压缩器，其中$\alpha$， $d$ 是定理 *19*中的常数。

证明。令 $X$ 为 $\mathbf{F}_2^n$ 上熵为 $k=\delta n$ 的任意仿射源。不失一般性，假设 $X$ 的支撑集是线性子空间 $V$（如果不是，我们可以先对相应的线性子空间进行分析，然后加上仿射平移，因为这里处理的始终是线性函数）。我们首先为 $V$ 给出一组 $k$ 基向量。为此，考虑线性子空间 $W\subseteq V$，使得 $W$ 的前 $n/2$ 个比特为 $\mathbf{0}$。假设 $\dim(W)=r$，并令 $b_1$、$\cdots$、$b_r$ 为 $W$ 的一组基。接下来，将这些向量扩展为 $b_1$、$\cdots$、$b_r$、$c_1$、$\cdots$、$c_s$，它们构成 $V$ 的一组完整基，满足 $s+r=k$。

注意，由 $\{c_i\}_{i=1}^s$ 的前 $n/2$ 个比特构成的向量也线性独立，否则它们的某个线性组合会属于 $W$。令 $\{\bar{c}_i\}_{i=1}^s$ 为 $\{c_i\}_{i=1}^s$ 的前 $n/2$ 个比特，$\{\tilde{c}_i\}_{i=1}^s$ 为 $\{c_i\}_{i=1}^s$ 的第二个 $n/2$ 比特。类似地，令 $\{b_i\}_{i=1}^r$ 为 $\{b_i\}_{i=1}^r$ 的第二个 $n/2$ 比特（回忆一下，前 $n/2$ 个比特为 $\mathbf{0}$）。

现在，令 $\mathcal{Q}\subseteq[s]$ 满足 $(\{b_i\}_{i=1}^r,\{\tilde{c}_i\}_{i\in\mathcal{Q}})$ 构成 $X_2$ 的支撑线性子空间的一组基。令 $C=\{\tilde{c}_i\}_{i\in\mathcal{Q}}$ 的张成空间。源 $X$ 的采样方式是：选取一个均匀随机向量$Y=(Y_1,\cdots,Y_k)\in\mathbf{F}_2^k$，并计算

$$\sum_{i=1}^s Y_i c_i+\sum_{j=1}^r Y_{s+j}b_j=\sum_{i=1}^s Y_i(\bar{c}_i,\tilde{c}_i)+\sum_{j=1}^r Y_{s+j}(0,\tilde{b}_j).$$

因此，前 $n/2$ 比特由 $\sum_{i=1}^s Y_i\bar{c}_i$ 给出，而后 $n/2$ 比特由

$$\sum_{i=1}^s Y_i\tilde{c}_i+\sum_{j=1}^r Y_{s+j}\tilde{b}_j=\sum_{i\in\mathcal{Q}} Y_i\tilde{c}_i+\sum_{i\in[s]\setminus\mathcal{Q}} Y_i\tilde{c}_i+\sum_{j=1}^r Y_{s+j}\tilde{b}_j.$$

注意，对于任意 $i\in[s]\setminus\mathcal{Q}$，$\tilde{c}_i$ 都可以表示为 $(\{b_i\}_{i=1}^r,\{\tilde{c}_i\}_{i\in\mathcal{Q}})$ 的线性组合。令 $\overline{Y}=(\{Y_i\}_{i\in[s]\setminus\mathcal{Q}})$，则上式可写为

$$\sum_{i\in\mathcal{Q}}(Y_i+L_i(\overline{Y}))\tilde{c}_i+\sum_{j=1}^r(Y_{s+j}+L_j(\overline{Y}))\tilde{b}_j,$$

其中，每个 $L_i$ 或 $L_j$ 都是从 $\mathbf{F}_2^{s-|\mathcal{Q}|}$ 到 $\mathbf{F}_2$ 的线性函数。

容易看出，$k$ 个随机比特 $(\{Y_i\}_{i=1}^s,\{Y_{s+j}+L_j(\overline{Y})\}_{j=1}^r)$ 相互独立且服从均匀分布（特别地，这些比特的任意非平凡奇偶性都是均匀随机比特）。类似地，随机比特 $(\{Y_i+L_i(\overline{Y})\}_{i\in\mathcal{Q}},\{Y_{s+j}+L_j(\overline{Y})\}_{j=1}^r)$ 也相互独立且服从均匀分布。令 $A=$ 为 $\{\bar{c}_i\}_{i=1}^s$ 的张成空间，$B=$ 为 $\{b_i\}_{i=1}^r$ 的张成空间，且 $C=$ 为 $\{\tilde{c}_i\}_{i\in\mathcal{Q})}$ 的张成空间。因此 $\dim(A)=s$，$\dim(B)=r$，并令 $\dim(C)=|\mathcal{Q}|=t$。根据上述计算，我们有 $H(X_1)=\dim(A)=s$，$H(X_2)=\dim(B)+\dim(C)=r+t$。此外，令 $X_3=\sum_{j=1}^r(Y_{s+j}+L_j(\overline{Y}))b_j$，且 $X_4=\sum_{i\in\mathcal{Q}}(Y_i+L_i(\overline{Y}))\tilde{c}_i$，则 $X_3$ 在 $B$ 上服从均匀分布，而 $X_4$ 在 $C$ 上服从均匀分布。因此 $H(X_3)=r$ 且 $H(X_4)=t$。我们知道 $X_1=\sum_{i=1}^s Y_i\bar{c}_i$。因此 $X_1$ 与 $X_3$ 相互独立，而 $X_4$ 是 $X_1$ 的确定性函数（因此也独立于 $X_3$）。注意，$X_2=X_3+X_4$，且 $X=(X_1,X_2)=(X_1,X_3+X_4)$。

注意到 $s+r=k=\delta n$。如果 $s\geq(\frac{1}{2}+\frac{\alpha}{8d})k$，那么 $H(X_1)=s\geq(\frac{1}{2}+\frac{\alpha}{8d})k=(1+\frac{\alpha}{4d})\delta(n/2)$。类似地，如果 $r\geq(\frac{1}{2}+\frac{\alpha}{8d})k$，那么 $H(X_2)=r+t\geq r\geq(1+\frac{\alpha}{4d})\delta(n/2)$。无论哪种情况，我们都完成了证明。否则，必有 $s<(\frac{1}{2}+\frac{\alpha}{8d})k$ 以及 $r<(\frac{1}{2}+\frac{\alpha}{8d})k$，这反过来意味着 $s>(\frac{1}{2}-\frac{\alpha}{8d})k$ 以及 $r>(\frac{1}{2}-\frac{\alpha}{8d})k$。现在，如果 $t\geq\frac{\alpha}{4d}k$，那么 $H(X_2)=r+t>(\frac{1}{2}+\frac{\alpha}{8d})k=(1+\frac{\alpha}{4d})\delta(n/2)$，我们再次完成了证明。

唯一剩下的情形是 $(\frac{1}{2}-\frac{\alpha}{8d})k<s,r<(\frac{1}{2}+\frac{\alpha}{8d})k$ and $t<\frac{\alpha}{4d}k$。由于 $s+r=k$，其中一个必定不超过 $k/2=\delta n/2$。我们分两种情况讨论。

情况 1. $(\frac{1}{2}-\frac{\alpha}{8d})k<s\leq k/2$. 在这种情况下，$\dim(A)=s\leq\delta(n/2)\leq(1/2)\cdot(n/2)$。考虑由定理 19 的维度扩展器给出的 $d$线性映射 $\{T_i\}_{i=1}^d$。注意，$(1+\alpha)s>(1+\alpha)(\frac{1}{2}-\frac{\alpha}{8d})k>(\frac{1}{2}+\frac{\alpha}{8d})k>r$。我们有如下命题。

命题 21. 存在一个 $i \in [d]$，使得 $\dim(T_i(A) + B) \geq r + (1+\alpha)_d^{s-r}$。

为说明这一点，不妨反设：对于所有 $i \in [d]$，都有 $\dim(T_i(A)+B) < r + (1+\alpha)_d^{s-r}$。于是

$$\dim\left(\sum_{i=1}^{d} T_i(A)\right) < r + d \cdot \frac{(1+\alpha)s - r}{d} = (1+\alpha)s,$$

由于 $\sum_{i=1}^{d} T_i(A)$ 中的任意向量都可以表示为 $B$ 中的 $r$ 基向量与另外 $< d \cdot (1+\alpha)_d^{s-r}$ 个向量的线性组合，其中每个 ~~$T_i(A)$~~ 贡献 $< (1+\alpha)_d^{s-r}$ 个向量。现在，对于这~~个特定的~~ $i \in [d]$，由于 $X_1$ 和 $X_3$ 是独立的，我们必有

$$H(T_i(X_1) + X_3) \geq r + \frac{(1+\alpha)s - r}{d} = \frac{1+\alpha}{d}k + \frac{d-2-\alpha}{d}r \geq \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k,$$

只要 $d \geq 3$。

注意，$T_i(X_1) + X_2 = T_i(X_1) + X_3 + X_4$，且 $X_4$ 是 $X_1$ 的确定性函数。由于$H(X_4) = t < \frac{\alpha}{4d}k$，我们可以固定 $X_4$，并在任意这样的固定条件下，

$$H(T_i(X_1) + X_2) \geq \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k - t > \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k - \frac{\alpha}{4d}k = \left(\frac{1}{2} + \frac{\alpha}{4d}\right)k.$$

因此，最终我们仍有 $H(T_i(X_1) + X_2) > (\frac{1}{2} + \frac{\alpha}{4d})k = (1 + \frac{\alpha}{2d})\delta(n/2)$。

情况 2. $(\frac{1}{2} - \frac{\alpha}{8d})k < r \leq k/2$. 该情形的证明与之前类似，只需稍作修改。具体而言，我们有 $\dim(B) = r \leq \delta(n/2) \leq (1/2) \cdot (n/2)$。考虑由定理 $d$ 线性映射$\{T_i\}_{i=1}^{d}$ 19 的维度扩展器给出的 19。完全沿用之前的论证，可得以下命题。

命题 22. 存在一个 $i \in [d]$，使得 $\dim(A + T_i(B)) \geq s + (1+\alpha)_d^{r-s}$。

现在再次注意到，由于 $X_1$ 和 $X_3$ 是独立的，我们必有

$$H(X_1 + T_i(X_3)) \geq s + \frac{(1+\alpha)r - s}{d} = \frac{1+\alpha}{d}k + \frac{d-2-\alpha}{d}s \geq \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k,$$

只要 $d \geq 3$。

注意，$X_1 + T_i(X_2) = X_1 + T_i(X_3) + T_i(X_4)$，且 $X_4$ 是 $X_1$ 的确定性函数。由于 $H(X_4) = t < \frac{\alpha}{4d}k$，我们可以固定 $X_4$，并在任意此类固定条件下，

$$H(X_1 + T_i(X_2)) \geq \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k - t > \left(\frac{1}{2} + \frac{\alpha}{2d}\right)k - \frac{\alpha}{4d}k = \left(\frac{1}{2} + \frac{\alpha}{4d}\right)k.$$

因此，最终我们仍然有 $H(X_1 + T_i(X_2)) > (\frac{1}{2} + \frac{\alpha}{4d})k = (1 + \frac{\alpha}{2d})\delta(n/2)$。

□

我们可以不 下面给出我们的主压缩器，它涉及重复使用基本凝聚器 r.

**算法 2** SCond($x$)

输入： $x \in \mathrm{F}_2^n$——一个 $n$ 比特串； $0 < \delta \leq 1/2$，一个给定参数。输出： $z \in (\mathrm{F}_2^m)^\ell$ ——一个由 $\ell$ 比特串组成的矩阵，长度为 $m$，其中 $m = n/\mathrm{poly}(1/\delta)$，且$\ell = \mathrm{poly}(1/\delta)$。

子程序和参数:
令 BasicCond 为算法 1给出的基本凝聚器。

设置 $x^0 = x$ 并令 $i = 0$。初始时， $x^i$ 只有 $n_0 = 1$ 行。

1. 重复以下步骤若干 $h = O(\log(1/\delta))$ 次：对于 $x^i$ 中的每个 $j$ 和第 $j$ 行$x^i_j$，应用 BasicCond($x^i_j$)，得到 $2d + 2$ 行。将它们拼接起来，得到具有 $n_{i+1} = n_i \cdot (2d + 2)$ 行的 $x^{i+1}$。令 $i \leftarrow i + 1$。

2. 令 $z = x^h$。

现在我们可以证明主定理了。

定理 *18* 的证明。我们证明算法 2 给出了这样的仿射某处压缩器。根据引理 20，对于任意仿射源 $X$ 且满足 $H(x) = \delta n$，其中 $0 < \delta \leq 1/2$，经过 $h' = O(\log(1/\delta))$ 步后，至少有一行 $x^{h'}$ 的熵不小于 $n'/2$。不失一般性，假设该行的熵恰好为 $n'/2$ （否则，我们可以先固定支撑线性子空间中的一些基向量，从而将该行化为若干熵恰好为 $n'/2$的仿射源的凸组合）。然后再应用一次 BasicCond，其中一个输出行的熵率至少为 $1/2(1 + \frac{\alpha}{4d}) = \frac{1}{2} + \frac{\alpha}{8d}$。

很容易看出，SCond 是一个线性函数，因此最终输出中的每一行都是一个仿射源。此外，由于我们在每一步中都将每一行划分为 2 个相等的块，并从中得到$2d+2$新行，因此每一行的最终长度为 $m = n/\mathrm{poly}(1/\delta)$，总共有$\ell = \mathrm{poly}(1/\delta)$行。 □

# 4 Li一般弱源的近似某处压缩器 s

接下来我们将证明，我 们的线性某处压缩器同样适用于一般弱随机源 es.

## 4.1 一些有用的结果

定义 23。分布 $\mathcal{D}$ 的碰撞概率定义为 $\mathrm{cp}(\mathcal{D}) = \Pr_{x,\ y \leftarrow_R \mathcal{D}}[x = y]$。

定义 24. 我们称分布 $\mathcal{X}$ 是分布 $\mathcal{X}_1$、$\cdots$ 和 $\mathcal{X}_m$ 的凸组合，如果存在数值 $p_1$、$\cdots$、$p_m \in [0, 1]$，使得 $\sum_i p_i = 1$，且随机变量 $\mathcal{X}$ 等于 $\sum_i p_i \mathcal{X}_i$。

引理 23 ([BISW04])。 设 $\mathcal{X}$ 是一个满足 加法碰撞概率$(\mathcal{X}) \leq \frac{1}{KL}$ 的分布。则 $\mathcal{X}$ 与具有至少 $\frac{1}{\sqrt{L}}$ 的最小熵 $\log K$ 的分布之间的统计距离为。

我们需要以下加性组合学中的结果。

引理 24 (普拉″内克-鲁兹萨 [TV06])。 设 $A$, $B$ 是加性群中的有限子集 $G$。则

$$|A+A| \leq \frac{|A+B|^4}{|A||B|^2}.$$

引理 25（Balog-Szemeredi-Gowers [BS94, Gow98]）。令 $A$、$B$ 为加法群 $G$ 的有限子集，并令 $|A|^{1-\rho_1} \leq |B| \leq |A|^{1+\rho_1}$。若 $\mathrm{cp}(A+B) \geq |A|^{-(1+\rho_2-\rho_1)}$，则存在子集 $A' \subseteq A$、$B' \subseteq B$，使得 $|A'| \geq |A|^{1-10\rho_2}$、$|B'| \geq |B|^{1-10\rho_2}$，且 $|A'+B'| \leq |A|^{1+\rho_1+10\rho_2}$。

定理 26 （有限域 $\mathsf{F}_2^n$ [GGMT23]）。令 $A \subset \mathsf{F}_2^n$ 为满足 $|A+A| \leq M|A|$的集合。则存在一个子集 $A' \subset A$，其大小为 $|A'| \geq M^{-c}|A|$，使得 |张成空间（$A'$）| $\leq M^c|A|$，其中 $c \geq 0$ 为绝对常数。

## 4.2 构造

我们将仿射某处压缩器推广如下。

**算法 3 BasicGCond($x$)**

输入： $x \in \{0,1\}^n$ ——一个 $n$ 比特串。输出： $z \in (\{0,1\}^m)^{2d+3}$ ——一个由 $2d+3$ 个比特串组成的数组，每个比特串长度为 $m$，其中 $m = n/2$ 且$d$ 是定理19中的常数。

子程序和参数:
令 $T = \{T_i\}_{i=1}^d$ 为 $\alpha$维度扩展器，由定理 19给出。

将 $x$ 分成 2 个块 $x = x_1 \circ x_2$，每个块含有 $n/2$ 个比特。
设 $z = z_1 \circ z_2 \circ \cdots \circ z_{2d+3}$，其中 $z_1 = x_1$、$z_2 = x_2$ 和 $z_{2i+1} = x_1 + T_i(x_2)$、$z_{2i+2} = x_2 + T_i(x_1)$，对任意 $i \in [d]$ 均成立。最后设 $z_{2d+3} = x_1 + x_2$。这里所有加法都将输入视为元素属于域 $\mathrm{F}^m$。[2]

我们有如下引理。

引理27。 对于任意 $0 < \delta \leq 1/2$， BasicGCond 是一个速率为 $(\delta \to (1+\Omega(\frac{\alpha}{d}))\delta$、$2^{-\Omega(\delta n)})$ 的某处压缩器，其中 $\alpha$， $d$ 是定理 *19*中的常数。

为证明该引理，我们首先证明以下引理。

引理 28。对于任意常数 $c > 0$，存在一个常数 $\varepsilon = \Omega(\frac{\alpha}{d})$，使得以下结论成立。设 $A$、$B$是 $\mathsf{F}_2^n$的有限子集。对于任意 $K \leq 2^{n/4}$，假设 $K^{1-c\varepsilon} \leq |A|$、$|B| \leq K^{1+c\varepsilon}$。如果 $\mathrm{cp}(A+B) \geq K^{-(1+2\varepsilon)}$，那么存在子集 $\tilde{A} \subseteq A$、$\tilde{B} \subseteq B$，使得 $|\tilde{A}| \geq K^{1-O(\varepsilon)}$、$|\tilde{B}| \geq K^{1-O(\varepsilon)}$、|张成空间( $\tilde{A}$) $| \leq K^{1+O(\varepsilon)}$、|张成空间( $\tilde{B}$)$| \leq K^{1+O(\varepsilon)}$，并且 BasicCond($\mathcal{A} \circ \mathcal{B}$)的输出中至少有一行具有最小熵 $(1+\Omega(\varepsilon))\log K$，其中 $\mathcal{A}$、$\mathcal{B}$分别是关于~ ~$A$、$B$ 的均匀独立分布。

证明。 如果 加法碰撞概率$(A+B) \geq K^{-(1+2\varepsilon)}$，根据引理 25，存在子集 $A' \subseteq A$、$B' \subseteq B$ 使得$|A'| \geq |A|^{1-O(\varepsilon)}$、$|B'| \geq |B|^{1-O(\varepsilon)}$，并且 $|A'+B'| \leq |A|^{1+O(\varepsilon)} = K^{1+O(\varepsilon)}$。然后，根据引理 24，有 $|A'+A'| \leq \frac{|A'+B'|^4}{|A'||B'|^2} \leq K^{1+O(\varepsilon)}$。类似地，我们也有 $|B'+B'| \leq K^{1+O(\varepsilon)}$。

接下来，根据定理 26，存在一个大小为 $\tilde{A} \subset A'$ 的子集 $|\tilde{A}| \geq K^{-O(\varepsilon)}|A'| = K^{1-O(\varepsilon)}$，使得 |张成空间$(\tilde{A})| \leq K^{O(\varepsilon)}|A'| = K^{1+O(\varepsilon)}$。类似地，还存在这样一个子集 $\tilde{B} \subset B'$，具有相同的

~ ~性质。现在，令 $\mathcal{A}'$、$\mathcal{B}'$ 分别为 Span($A$) 和 Span($B$) 上的均匀独立分布。注意，$\mathcal{A}'$、$\mathcal{B}'$ 都是仿射源，因此 $\mathcal{A}' \circ \mathcal{B}'$ 也是一个仿射源，其熵为~ ~log $|A| + \log |B| = (1 - O(\varepsilon))2 \log K$。因此，不失一般性，我们可以将其视为一个熵恰好为 $(1 - O(\varepsilon))2 \log K < n/2$ 的仿射源。现在，根据引理 20，BasicCond($\mathcal{A}' \circ \mathcal{B}'$) 的输出中至少有一行的熵为 $(1 + \frac{\alpha}{4d})(1 - O(\varepsilon)) \log K$。注意，$|\tilde{A}||\tilde{B}| \geq (|\text{Span}(\tilde{A})||\text{Span}(\tilde{B})|)^{(1-O(\varepsilon))}$。因此，只要 $\varepsilon = \gamma \frac{\alpha}{4d}$ 对足够小的常数 $\gamma > 0$ 成立，BasicCond($\mathcal{A} \circ \mathcal{B}$) 的输出中同一行的最小熵至少为 $(1 + \frac{\alpha}{4d})(1 - O(\varepsilon)) \log K - O(\varepsilon) \log K = (1 + \Omega(\varepsilon)) \log K$。 □

引理 29. 对于任意常数 $c > 0$，存在常数 $\varepsilon = \Omega(\frac{\alpha}{d})$，使得下述结论成立。设 $A$、$B$ 是 $\mathbb{F}_2^n$ 的有限子集。对于任意 $K \leq 2^{n/4}$，假设 $K^{1-c\varepsilon} \leq |A|$、$|B| \leq K^{1+c\varepsilon}$ 和 |张成空间($B$)| $\leq K^{1+O(\varepsilon)}$。如果加法碰撞概率$(A + B) \geq K^{-(1+2\varepsilon)}$，那么存在一个子集 $\tilde{A} \subseteq A$，使得$|\tilde{A}| \geq K^{1-O(\varepsilon)}$、|张成空间( $\tilde{A}$)| $\leq K^{1+O(\varepsilon)}$，并且 BasicCond($\mathcal{A} \circ \mathcal{B}$) 的输出中至少有一行具有~最小熵 $(1 + \Omega(\varepsilon)) \log K$，其中 $\mathcal{A}$、$\mathcal{B}$ 分别是 $A$、$B$ 上的均匀独立分布。

证明。证明与前一个引理完全相同，只是在第二段中，我们~可以直接用 $B$ 替换集合 $B$ 。

□

现在我们可以证明以下引理。

引理 30. 存在一个常数 $\varepsilon = \Omega(\frac{\alpha}{d})$，使得以下结论成立。设 $A, B$ 是 $\mathbb{F}_2^n$ 的有限子集。对于任意 $K \leq 2^{n/4}$，假设 $K^{1-\varepsilon} \leq |A|, |B| \leq K^{1+\varepsilon}$。设 $X, Y$ 分别是 $A, B$ 上的均匀分布和独立分布。那么，BasicGCond($X \circ Y$) 与某个 somewhere-$(1 + \Omega(\varepsilon)) \log K$ 源的距离为 $K^{-\Omega(\varepsilon)}$。特别地，$X \circ Y$ 可以划分为互不相交的子源，使得对于每个子源，其概率质量要么 (1) 至多为 $2K^{-\varepsilon}$，要么 (2) 至少为 $K^{-O(\varepsilon)}$；并且 BasicGCond 在该子源上的输出与一个基本的 somewhere $(1 + \Omega(\varepsilon)) \log K$ 源的距离为 $K^{-\varepsilon}$。

证明。我们反复应用引理 28 和引理 29，并按如下方式将 $A \times B$ 划分为互不相交的子集。首先注意，如果 cp($A + B$) $\leq K^{-(1+2\varepsilon)}$，那么根据引理 23，$X + Y$ 与具有最小熵 $(1 + \varepsilon) \log K$ 的~ ~相差至多 $K^{-\varepsilon/2}$。否则，根据引理 28，存在子集 $A \subseteq A$、$B \subseteq B$，使得 $|\tilde{A}| \geq K^{1-O(\varepsilon)}$、$|\tilde{B}| \geq K^{1-O(\varepsilon)}$、|张成空间( $\tilde{A}$)| $\leq K^{1+O(\varepsilon)}$、|张成空间( $\tilde{B}$)| $\leq K^{1+O(\varepsilon)}$，并且 BasicCond($\mathcal{A} \circ \mathcal{B}$) 的输出中至少有一行具有最小熵 $(1 + \Omega(\varepsilon)) \log K$，其中 $\mathcal{A}$、$\mathcal{B}$ 分别是 $A$、$B$ 上的~ ~均匀独立分布。

现在考虑集合 $A^1 = A \setminus \tilde{A}$ 和 $B^1 = B \setminus \tilde{B}$。那么~ ~ $X \circ Y$ 中对应于 $(A \times B) \setminus (A \times B)$ 中元素的总概率质量至多为 $\frac{|A^1||B| + |A||B^1|}{|A||B|} \leq 2K^{-\varepsilon}$，结论得证。

否则，考虑以下三个集合：$\tilde{A} \times B^1$、$A^1 \times \tilde{B}$和 $A^1 \times B^1$。注意，这些~ ~是互不相交的子集，其并集等于 $(A \times B) \setminus (A \times B)$。我们分几种情况讨论。

情况 1。$|A^1|$ 和 $|B^1|$ 中只有一个的大小大于 $K^{1-2\varepsilon}$。不失一般性，假设$|B^1| \leq K^{1-2\varepsilon}$。注意，在这种情况下，$X \circ Y$ 中对应于 $(\tilde{A} \times B^1) \cup (A^1 \times B^1) = A \times B^1$ 中元素的总概率质量至多为 $\frac{|A||B^1|}{|A||B|} \leq K^{-\varepsilon}$。

对于 $A \times \tilde{B}$，我们反复应用引理29。初始时令 $A^* = A$。只要 $|A^*| \geq K^{1-2\varepsilon}$，如果加法碰撞概率 $(A^* + \tilde{B}) \leq K^{-(1+2\varepsilon)}$，那么再次根据引理23，对应于在 $A^*$ 和 $B$ 上的均匀独立分布的随机~变量之和的输出将 $K^{-\varepsilon/2}$-接近于具有最小熵 $(1 + \varepsilon)\log K$，此时我们停止。否则我们使用

引理29 来找到一个子集 $\tilde{A} \subseteq A^*$，使得 $|\tilde{A}| \geq K^{1-O(\varepsilon)}$，|张成空间( $\tilde{A}$)| $\leq K^{1+O(\varepsilon)}$，并且 BasicCond($\mathcal{A} \circ \mathcal{B}$) 的输出中至少有一行具有最小熵 $(1+\Omega(\varepsilon))\log K$，其中 $\mathcal{A}$，$\mathcal{B}$ ~ ~ ~分别是在 $A$和 $B$ 上的均匀独立分布。然后我们从 $A$中移除 $A^*$ 并重复此过程。当 $|A^*| < K^{1-2\varepsilon}$时，过程结束。

因此，总的来说，我们已将 $A \times B$ 划分为互不相交的子集，或等价地，将 $X \circ Y$ 划分为互不相交的子源，使得对于每个子源，其概率质量至多为 $K^{-\varepsilon}$，或者 BasicGCond 在该子源上的输出与某个 somewhere $(1+\Omega(\varepsilon))$ log $K$ 源的距离至多为 $K^{-\varepsilon}$。

情况 2。 $|A^1| > K^{1-2\varepsilon}$ 和 $|B^1| > K^{1-2\varepsilon}$。我们首先将情况 1 中的论证应用于 $\tilde{A} \times B^1$ 和 $A^1 \times \tilde{B}$。然后考虑 $A^1 \times B^1$。这与我们开始时的情况相同。也就是说，如果 cp($A^1 + B^1$) $\leq K^{-(1+2\varepsilon)}$，那么根据引理 23，我们就完成了证明。否则，根据引理 28，存在子集 $\tilde{A}^1 \subseteq A^1$、$\tilde{B}^1 \subseteq B^1$，使得 $|\tilde{A}^1| \geq K^{1-O(\varepsilon)}$、$|\tilde{B}^1| \geq K^{1-O(\varepsilon)}$、|张成空间( $\tilde{A}^1$)| $\leq K^{1+O(\varepsilon)}$，|张成空间( $\tilde{B}$)| $\leq K^{1+O(\varepsilon)}$，并且 BasicCond($\mathcal{A}^1 \circ \mathcal{B}^1$) 的输出中至少有一行具有 $(1+\Omega(\varepsilon))$ log $K$ 的最小熵，其中 $\mathcal{A}^1$、$\mathcal{B}^1$ 分别是 ~ ~上 $A^1$、$B^1$ 的均匀独立分布。因此，我们可以像之前一样继续分析。

结合这两种情况，最终我们将 $X \circ Y$ 划分为互不相交的子源，使得对于每个子源，要么（1）其概率质量至多为 $2K^{-\varepsilon}$，要么（2）该子源上的 BasicGCond 输出与某处为（$1+\Omega(\varepsilon)$) log $K$ 源的情形之间的距离至多为 $K^{-\varepsilon}$。

注意，当某个子源满足（2）时，其概率质量总是至少为 $K^{1-O(\varepsilon)} \cdot K^{1-O(\varepsilon)}/K^{2(1+\varepsilon)} = K^{-O(\varepsilon)}$。 □

现在我们可以证明引理27。

引理 27的证明。 给定一个（$n$，$\delta n$）源 $X$ 满足 $0 < \delta \leq 1/2$，且 $X = X_1 \circ X_2$，不失一般性，我们可以假设 $X$ 是集合 $S \subseteq \{0,1\}^n$ 上的均匀分布，且该集合$|S| = 2^{\delta n}$满足 $\lambda > 0$，其中常数参数将在稍后确定。对于 $i \in [2]$，定义$H_i = \{y \in \{0,1\}^m : \mathbf{Pr}[X_i = y] \geq 2^{-(1+\lambda)\delta m}\}$，这对应于 $X_i$中的重元素。注意，这意味着对于每个 $i$，$|H_i| \leq 2^{(1+\lambda)\delta m}$。令 $\tau = 2^{-\beta\delta m}$，其中常数 $\beta > 0$ 将在稍后确定。我们定义如下集合。

1. $S' = \{x \in S : \exists i,\ x_i \notin H_i\}$。
2. 对任意 $x \in S'$，定义 $I(x)$ 为满足 $x_i \notin H_i$ 的最小 $i$，并且 $T_i = \{x \in S', I(x) = i\}$。令 $B = \{i \in [2] : |T_i| < 2^{(1-\beta)\delta m}\}$，并定义 $\tilde{S} = S' \setminus (\cup_{i \in B} T_i)$。注意 $|\cup_{i \in B} T_i| \leq 2\tau|S|$。
3. $S'' = \{x \in S : \forall i,\ x_i \in H_i\} = S \setminus S'$。

~ ~ ~注意，对于任意 $x \in S$，我们有 $I(x) \notin B$。令 $X$ 为 $S$ 上的均匀分布。对于任意 $i \in [2] \setminus B$ 和任意 $y \in \{0,1\}^m$，在 $I(\tilde{X}) = i$的条件为定值时，我们有 $\Pr[\tilde{X}_i = y] \leq \frac{\mathbf{Pr}[X_i = y]}{2^{-\beta\delta m}} \leq 2^{-(1+\lambda-\beta)\delta m}$. 因此，只要 $\beta \leq \lambda/2$，$\tilde{X}_i$ 的最小熵至少为 $(1+\lambda/2)\delta m$。于是，$\tilde{X}$ 是一个初等somewhere-$(1+\lambda/2)\delta m$ 源。

现在有两种情况。

~情况 1。$\Pr[X \in S'] \geq 1 - \tau$。在这种情况下，注意到 $X$ 与 $X$ 是 $2\tau + \tau = 3\tau$-接近的，因此得证。

情况 2。$\Pr[X \in S''] \geq \tau$。在这种情况下，请注意 $|S''| \geq \tau|S| = 2^{(2-\beta)\delta m}$。此外， $S''$是 $H_1 \times H_2$ 的子集，因此

$$|H_1 \times H_2| \geq |S''| \geq 2^{(2-\beta)\delta m}.$$

然而，对于每个 $i \in [2]$，我们有 $|H_i| \leq 2^{(1+\lambda)\delta m}$，因此，对于每个 $i \in [2]$，我们也有

$$|H_i| \geq 2^{(2-\beta)\delta m}/2^{(1+\lambda)\delta m} = 2^{(1-\beta-\lambda)\delta m}.$$

现在考虑源（$Y_1$， $Y_2$），其中每个 $Y_i$ 都是在$H_i$上的独立均匀分布。我们将通过设置 $K = 2^{\delta m} \leq 2^{n/4}$和 $\varepsilon \geq 2\lambda$来应用引理 30 。注意，对于任意 $i \in [2]$，我们有 $K^{1-\varepsilon} \leq |H_i| \leq K^{1+\varepsilon}$，因为我们选择了 $\beta \leq \lambda/2$。

因此，根据引理 30，存在一个常数 $c > 0$，使得 $Y_1 \circ Y_2$ 可以被划分为互不相交的子源，使得对于每个子源，要么（1）概率质量至多为 $2K^{-\varepsilon}$，要么（2）概率质量至少为 $K^{-c\varepsilon}$，并且 BasicGCond 在该子源上的输出是 $K^{-\varepsilon}$接近于一个基本的某处（$1 + \Omega(\varepsilon)$）对数 $K$ 源。

对于 (2) 中的每个子源 $Y^j$，我们考虑其支撑集与 $S''$ 的交集。如果该交集的概率质量至多为 $K^{-c\varepsilon-4\lambda}$，则称其为坏交集；否则称其为好交集。注意，对于好交集，将交集上的均匀分布作为子源时，BasicGCond 的输出与某个基本的 somewhere（$1 + \Omega(\varepsilon) - 4\lambda$）log $K$ 源的距离为 $K^{-\varepsilon+4\lambda}$-近。另一方面，坏交集的总概率质量至多为 $K^{-4\lambda}$。

注意， $S''$ 在 $(Y_1, Y_2)$ 中的概率质量至少为 $2^{(2-\beta)\delta m}/(2^{(2+2\lambda)\delta m}) = 2^{-(\beta+2\lambda)\delta m}$。因此，如果我们将 $S''$ 上的均匀分布定义为 $X''$，那么只要取充分小的常数 $\gamma > 0$ 作为 $\lambda = \gamma\varepsilon$，BasicGCond($X''$) 就与某个 somewhere（$1 + \Omega(\varepsilon) - 4\lambda$）log $K = (1 + \lambda)\delta m$源的距离为 $(2K^{-\varepsilon} + K^{-4\lambda})/(2^{-(\beta+2\lambda)\delta m}) + K^{-\varepsilon+4\lambda} \leq 2^{-\lambda\delta m}$-近。

现在那么，BasicGCond($X''$)与某个某处（$1 + \lambda/2$）$\delta m$ 源的距离为 $2^{-\lambda\delta m}$-close。注意， $X$ 与 $X'$ 的距离为 $2\tau$-close。因此，BasicGCond($X$) 与某个某处（$1 + \lambda/2$）$\delta m$ 源的距离为$2\tau + 2^{-\lambda\delta m}$-close。

令 $\beta = \lambda/2$，我们得到，在这两种情况下，BasicGCond($X$) 与某个某处（$1 + \Omega(\frac{\alpha}{d})$)$\delta m$ 源的距离为 $2^{-\Omega(\delta n)}$-close。 □

现在，我们的主压缩器涉及反复使用基本凝聚器。

---

算法 4 SGCond($x$)

---

输入： $x \in \mathbf{F}_2^n$——一个 $n$ 比特串； $0 < \delta \leq 1/2$，一个给定参数。输出： $z \in (\mathbf{F}_2^m)^\ell$ ——一个由长度为 $m$的 $\ell$ 比特串组成的矩阵，其中 $m = n/\text{poly}(1/\delta)$ 且$\ell = \text{poly}(1/\delta)$。

---

子程序和参数:
令 BasicGCond为算法 3 给出的基本凝聚器。

---

设置 $x^0 = x$ 并令 $i = 0$。初始时， $x^i$ 仅有 $n_0 = 1$ 行。

1. 重复以下步骤 $h = O(\log(1/\delta))$ 次：对于每个 $j$ 以及 $x^i$ 中的第 $j$ 行$x^i_j$，应用 BasicGCond($x^i_j$) 得到 $2d + 3$ 行。将它们拼接起来，得到具有$n_{i+1} = n_i \cdot (2d + 3)$ 行的 $x^{i+1}$。设置 $i \leftarrow i + 1$。

2. 令 $z = x^h$。

---

通过与定理18的证明类似的论证，我们可以证明以下定理。

定理31。存在一个常数 $\beta > 0$，使得对于任意 $0 < \delta \leq 1/2$，都存在一个显式的速率为 $(\delta \to 1/2 + \beta, 2^{-\Omega(m)})$ 的某处压缩器 SGCond：$\{0,1\}^n \to (\{0,1\}^m)^t$，其中 $t = \text{poly}(1/\delta)$ 且 $m = n/\text{poly}(1/\delta)$。此外，SGCond 是一个线性函数。

# 5 方向仿射提取器

本节将介绍我们针对线性熵且误差为指数级小误差的方向仿射提取器。该构造对于次线性熵同样适用，但误差和输出长度会略有损失。

## 5.1 低次数仿射相关性破坏器

正如我们在第 1 节中所介绍的，保持方向仿射提取器的输出为 低次数 至关重要。然而，我们的构造使用了建议相关性破坏器，而现有的所有相关性破坏器的次数都高得无法满足我们的需求。为了解决这一问题，我们构造了一族低次数相关性破坏器。我们假设以下每个子程序的输入随机变量都是仿射的。该假设是成立的，因为在算法 8 中调用 37 的定理 IdACB 时，输入随机变量是仿射的。

强种子提取器的替代方案。 我们的构造将基于与以下文献中的建议相关性破坏器类似的框架[CGL22]。为保持次数较低，我们将全文使用的 GUV提取器和先压缩后哈希提取器替换为定理32中的低次强线性带种子提取器。

定理 32 （低次强线性带种子提取器 [Li11]）。 存在一个常数 $0 < \beta < 1$，使得对于任意 $0 < \delta < 1$ 以及任意 $1/\sqrt{n} < \alpha < 1$，存在一个多项式时间可计算函数 LSExt： $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 以及一个常数 $0 < \beta < 1$，满足：

- $d \leq \alpha n$， $m \geq \beta\delta\alpha n$。
- 对于任意 $(n, \delta n)$-仿射源 $X$，令 $R$ 为 $\{0,1\}^d$ 上独立于 $X$ 的均匀分布。则 $(\text{LSExt}(X, R)\ R)$ 与均匀分布的距离为 $2^{-\Omega(\delta\alpha^2 n)}$。
- 输出中的每个比特都是两个输入的比特的 4 次多项式，并且对于 $r$ 的任意固定，输出都是 $x$ 的线性函数。

低次数前瞻提取器。 第一步是构造一个低次数前瞻提取器，它是低次数建议相关性破坏器的一个组成部分。以下算法给出了这样一种构造，并采用定理 32中的低次强线性带种子提取器来实例化。由于低次强线性带种子提取器的输出长度短于最小种子长度，我们无法直接应用已有的前瞻提取器引理。因此，我们需要量身定制一组新的参数，并为低次数情形提出一个新定理。

算法5 $(k,t,\varepsilon)$-laExt$(x,y)$

输入：长度分别为 $n$、$d$ 的比特串 $x$、$y$。初始时，$x$ 的熵为 $k$。输出：长度为 $2m$ 的比特串 $(r_0,r_1)$。子程序和参数：令 $s=d/(2+2t)$，其中 $C_0\sqrt{\log(1/\varepsilon)}_k$ $n\geq s\geq C_1 n\sqrt{n}/k$，对某些常数 $C_0>0, C_1>1$ 成立。令 $\underline{\text{LSExt}}^1_w$：$\{0,1\}^n\times\{0,1\}^s\to\{0,1\}^{m_1}$ 为定理 32 中的低次强线性带种子提取器，其 $\delta_{32}=k/n$, $\alpha_{32}=d/((2t+2)n)$，误差为 $\varepsilon_0=2^{-\Omega(kd^2/((t+1)^2n^2)}$，输出长度为$m_1=\beta_{32}kd/((2t+2)n)$。令 LSExt$^1_q$：$\{0,1\}^d\times\{0,1\}^{m_1}\to\{0,1\}^{m_2}$ 为定理 32 中的低次强线性带种子提取器，其 $\delta_{32}=1/2$, $\alpha_{32}=\beta_{32}k/((2+2t)n)$，误差为 $\varepsilon_1=2^{-\Omega(dk^2/((t+1)^2n^2))}=\varepsilon_0^{\Omega(d/k)}$，输出长度为 $m_2=\beta_{32}^2kd/(4(t+1)n)$。令 LSExt$^2_w$：$\{0,1\}^n\times\{0,1\}^{m_2}\to\{0,1\}^m$ 为定理 32 中的低次强线性带种子提取器，其 $\delta_{32}=k/(2n)$ 和 $\alpha_3=\beta_{32}^2kd/((4+4t)n^2)$，误差为 $\varepsilon_2=2^{-\Omega(k^3d^2/((8+8t)^2n^4))}=\varepsilon^{\Omega(k^2/n^2)}$ 0，输出长度为 $m=\beta_{32}^3k^2d/((8+8t)n^2)$。

---

1. 令 $s_0=\text{Slice}(y,s)$ 2. 令 $\tilde{r}_0=\text{LSExt}^1_w(x,s_0)$ 3. 令 $s_1=\text{LSExt}^1_q(y,\tilde{r}_0)$ 4. 令 $r_1=\text{LSExt}^2_w(x,s_1)$ 5. 输出 $r_0=\text{Slice}(\tilde{r}_0,|r_1|)$ $r_1$

---

定理 33（二步前瞻提取器）。对于任意 $t\leq\sqrt{n}$, $t\in\mathbb{N}$ 和 $\varepsilon>0$，存在一个显式函数 laExt：$\{0,1\}^n\times\{0,1\}^d\to(\{0,1\}^m)^2$，满足以下条件。设 $X,X^{[t]}\in\{0,1\}^n$和 $Y,Y^{[t]}\in\{0,1\}^d$为随机变量，且 $(X,X^{[t]})$ 与 $(Y,Y^{[t]})$, $Y=U_d$相互独立。存在一个足够大的常数 $C>0$，使得如果

$$k=H(X)\geq C\max\left\{\left((t+1)^2\log(1/\varepsilon)n^4/d^2\right)^{1/3},(t+1)\sqrt{n}\right\};$$
$$n\geq d\geq C(t+1)\max\left\{n\sqrt{n}/k,\left(\log(1/\varepsilon)n^4/k^3\right)^{1/2}\right\},$$

then $(R_0,R_1):=\text{laExt}(X,Y)$ 及其篡改 $(R_0^{[t]},R_1^{[t]})$ 满足

$$(R_0\approx_\varepsilon U_m)\mid(Y,Y^{[t]});$$
$$(R_1\approx_\varepsilon U_m)\mid(Y,Y^{[t]},R_0,R_0^{[t]}),$$

其中 $m=\Omega(k^2d/((1+t)n^2))$。
此外，$r_0'$ 的每个比特都是输入比特的一个 4 次多项式；$r_1$ 的每个比特都是输入比特的一个 40 次多项式。

证明。我们将证明算法 5是这样一个函数。首先，我们证明算法 5 中的参数选择正确。第一个约束来自任何 LSExt 的最小种子长度要求，即我们需要保证

$$d/((2t+2)n)>1/\sqrt{n}\,;\qquad(\text{LSExt}^1_w\text{ 的种子长度要求})$$
$$\beta_{32}k/((2+2t)n)>1/\sqrt{n}\,;\qquad(\text{LSExt}^1_q\text{的种子长度要求})$$
$$\beta_{32}^2kd/((4+4t)n^2)>1/\sqrt{n}\,,\qquad(\text{LSExt}^2_w\text{ 的种子长度要求})$$

这为 $d$ 给出了一个下界：

$$d > \frac{4}{\beta_{32}^2} \cdot \frac{(t+1)n\sqrt{n}}{k}.$$

由于 $n \geq d$，我们有

$$k > \frac{4}{\beta_{32}^2} \cdot (t+1)n.$$

令 $\varepsilon \geq \varepsilon_0 + \varepsilon_1 + \varepsilon_2$，则存在一个常数 $\lambda$，使得 $\varepsilon \geq 2^{-\lambda k^3 d^2/((t+1)^2 n^4)}$。对误差取对数并将 $k$ 和 $d$ 单独移出后，我们有

$$\frac{(t+1)^2 \log(1/\varepsilon) n^4}{\lambda} \leq k^3 d^2 \iff k \geq C_0 \left((t+1)^2 \log(1/\varepsilon) n^4/d^2\right)^{1/3};$$
$$\frac{(t+1)^2 \log(1/\varepsilon) n^4}{\lambda} \leq k^3 d^2 \iff d \geq C_0'(t+1) \left(\log(1/\varepsilon) n^4/k^3\right)^{1/2},$$

对于某些足够大的常数 $C_0, C_0'$。因此，如果 $k$ 和 $d$ 满足定理 33 中的约束，它们也适用于算法 5。接下来，我们证明前瞻提取器的提取性质。

1. 由于 $Y = U_d$，$S_0$ 是均匀的。由于 $Y$ 与 $X$ 独立，$X$ 与 $S_0$ 独立。由于 $H(X) \geq k$，根据 $\mathsf{LSExt}_w^1$ 的强种子提取器性质，$\tilde{R}_0 \approx_{\varepsilon_0} U_{m_1} \mid S_0$；这也意味着，鉴于 $X$ 与 $(Y, Y^{[t]})$ 独立，$R_0 \approx_{\varepsilon_0} U_m \mid (Y, Y^{[t]})$。

2. 由于 $S_0$、$S_0^{[t]}$ 是 $Y$和 $Y^{[t]}$的线性函数，因此我们有 $H(Y \mid S_0, S_0^{[t]}) \geq d - (t+1) \cdot \frac{d}{2+2t} = d/2$。由于 $\tilde{R}_0$ 与均匀分布 $\varepsilon_0$ 接近，根据 $\mathsf{LSExt}_q^1$的强种子提取器性质，$S_1 \approx_{\varepsilon_0+\varepsilon_1} U_{m_2}$。

3. 在固定$(S_0, S_0^{[t]})$的条件下，$R_0$、$R_0^{[t]}$ 是 $X$和 $X^{[t]}$的线性函数。因此，我们有 $H(X \mid \tilde{R}_0, \tilde{R}_0^{[t]}) \geq k - (t+1)m_1 = k/2$。由于 $S_1 \approx_{\varepsilon_0+\varepsilon_1} U_{m_2}$，根据 $\mathsf{LSExt}_w^2$的强种子提取器性质，$R_1 \approx_{\varepsilon_0+\varepsilon_1+\varepsilon_2} U_m \mid (R_0, R_0^{[t]}, Y, Y^{[t]})$。

最后，输出的次数可由 $\mathsf{LSExt}$ 在定理 32中的输出次数轻易得到。这就完成了定理 33的证明。

□

低次数非篡改独立性保持合并器。 第二步是构造一个低次数非篡改独立性保持合并器。非篡改独立性保持合并器最早在 [CL16a] 中定义，用于合并某处随机源，同时保持其自身与篡改源之间的独立性。我们从定义开始。

定义 25. 一个带有误差 $\varepsilon$ 的 $(t, \ell, \varepsilon)$-NIPM：$\{0,1\}^n \times (\{0,1\}^m)^\ell \to \{0,1\}^{m_1}$，或简称 $\mathsf{NIPM}_\ell$，是满足以下性质的函数。设

- $V$、$V^{[t]}$ 是随机变量，每个随机变量都支撑在布尔 $\ell \times m$ 矩阵上，使得对于任意 $i \in [\ell]$、$V_i = U_m$；
- 对于每个 $j \in [t]$，存在一个 $h_j \in [\ell]$，使得 $(V_{h_j}, V_{h_j}^j) = (U_m, V_{h_j}^j)$；
- $X$、$X^{[t]}$ 是与 $V$、$V^{[t]}$独立的随机变量，各自的支撑集包含 $d$ 比特，且 $X$ 具有足够的熵，

*then*

$$\mathsf{NIPM}_\ell(X,V) \approx_\varepsilon U_{m_1} \mid (\mathsf{NIPM}_\ell(X^1,V^1),\cdots,\mathsf{NIPM}_\ell(X^t,V^t)).$$

**算法 6** 非交互式概率合并器$_\ell(x,v)$

输入：$x$——一个 $n$ 比特串，$v$——一个 $\ell \times m$ 比特矩阵。输出：$z$——一个 $m \cdot \prod_{i=1}^{\ell} \alpha_i$ 比特串，其中每个 $\alpha_i$ 都在下文针对 $i \in [\ell]$定义。

子程序和参数:
令 $\delta_w = k/2n$ 为定义 25中假设熵率 $x$ 的一个下界，$\delta_q = 1/2$ 为对每个 $v_{h_i}$ 的熵率下界，其中 $i \in [\ell]$。令 $\alpha_1 = m/((3+3t)n)$。
对于 $i \in [\ell-1]$：

- 令 $\delta_{32}$ 和$\alpha_{32}$ 分别由定理 32 设为 $\delta_w$ 和 $\alpha_i$；
- 令 $\delta_{32}$ 和$\alpha_{32}$ 分别由定理 32 设为 $\delta_q$ 和 $\delta_w\beta_{32}\alpha_i$；
- 令 $\alpha_{i+1} = \delta_q\delta_w\beta_{32}^2\alpha_i$,

这给出

- $\alpha_i = (\delta_w\delta_q)^{i-1}\beta_{32}^{2i-2}\alpha_1$.

对于 $i \in [\ell-1]$:

- $\mathsf{LSExt}_w^i$：$\{0,1\}^n \times \{0,1\}^{\alpha_i n} \to \{0,1\}^{\delta_w\beta_{32}\alpha_i n}$ 为定理 32 中的提取器，其误差为$\varepsilon_i^w = 2^{-\Omega(\delta_w^{2i-1}\delta_q^{2i-2}\alpha_1^2 n)}$。
- $\mathsf{LSExt}_q^i$：$\{0,1\}^m \times \{0,1\}^{\delta_w\beta_{32}\alpha_i n} \to \{0,1\}^{\alpha_{i+1}n}$ 为定理 32 中的提取器，其误差为$\varepsilon_i^q = 2^{-\Omega(\delta_q^{2i-1}\delta_w^{2i-1}\alpha_1^2 m)}$。

令 $s_1 = \mathsf{Slice}(v_1, \alpha_1 n)$。对于 $i \in [\ell-1]$：1. $r_i = \mathsf{LSExt}_w^i(x, s_i)$ 2. $s_{i+1} = \mathsf{LSExt}_q^i(v_{i+1}, r_i)$ 令 $z = s_\ell$。

定理 34（$\mathsf{NIPM}_\ell$）。对于每个 $\ell \in \mathbb{N}$、$\varepsilon > 0$，如果存在一个足够大的 $C$ 使得

- $k \geq C\max\left\{\left((t+1)^2\log(1/\varepsilon)n^{2\ell-1}/m^3\right)^{1/(2\ell-3)}, (t+1)\sqrt{n}\right\}$;
- $n \geq m \geq C\max\left\{(t+1)n\sqrt{n}/k, ((t+1)\log(1/\varepsilon)n^{2\ell-1}/k^{2\ell-3})^{1/3}\right\}$,

于是存在一个 $\mathsf{NIPM}_\ell$：$\{0,1\}^n \times (\{0,1\}^m)^\ell \to \{0,1\}^{m_1}$ 以及常数 $0 < \eta < 1$ 和 $\{\varepsilon_i^w, \varepsilon_i^q\}_{i=1}^{\ell-1}$，二者均大于 0，使得

- $\varepsilon_i^w = \varepsilon_{\Omega((n^2/k^2)^{\ell-i})}$.
- $\varepsilon_i^q = \varepsilon^{\Omega((n^2/k^2)^{\ell-i})}$。
- $m_1 \geq \eta^\ell m/(t+1)$。

- 非交互式概率合并器 NIPM 的每个输出比特都是一个关于输入比特的 $2^{\Theta}$ 次$(\ell)$ 多项式。

在分析 NIPM$_\ell$ （以及之后的 ldACB）时，我们将反复使用下面的引理。它由仿射源的引理 15 调整而来。证明见附录 B 。

引理 35（仿射源的独立性合并引理）。设 LExt: $\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ 是任意一个（$k$, $\varepsilon$）-强线性种子提取器， $X_0\in\{0,1\}^n$ 是一个仿射源， $X$、$X^{[t]}\in\{0,1\}^n$、 $Y$、$Y^{[t]}\in\{0,1\}^d$ 都是 $X_0$、 $W=$LExt$(X,Y)$ 和 $W^j=$LExt$(X^j,Y^j)$ 的线性函数，对每个 $j\in[t]$ 均成立。假设存在 $S$、$T\subseteq[t]$ 使得

- $(Y,Y^S)\approx_\delta(U_d,Y^S)$;
- $H(X\mid X^T,Y,Y^{[t]})\geq k+tm$,

*then*

$$W\approx_{\varepsilon+\delta}U_m\mid(W^{S\cup T},Y,Y^{[t]}).$$

定理 *34* 的证明。 我们将证明算法 6 就是这样的 NIPM$_\ell$。我们首先讨论每个输出比特的次数。设 $s_i$ 对所有 $d_i$ 均为 $i\in[\ell]$，则它们满足以下递推公式

$$d_i=\begin{cases}1 & \text{if } i=1\\ 3(3d_{i-1}+1)+1=9d_{i-1}+4 & \text{if } i>1\end{cases}$$

求解后得到 $d_\ell=\frac{9^{\ell-1}-1}{2}$。
现在我们用归纳法证明以下断言。令 $R^{[j]}_{h_{[j]}}:=\{R^1_{h_1},R^2_{h_2},\cdots,R^j_{h_j}\}$。

断言 36. 不失一般性，令 $1\leq h_1\leq\cdots\leq h_t\leq\ell$。对于每个 $j\in[t]$，在第 $h_j$步之后，以下结论成立

$$\begin{aligned}S_{h_j}&\approx_{\sum_{i\in[h_j-1]}\varepsilon_i^w+\sum_{i\in[h_j-1]}\varepsilon_i^q}U_{\alpha_{h_j}n}\mid(S_{[i-1]},S^{[t]}_{[i-1]},R_{[i-1]},R^{[t]}_{[i-1]}),\\ R_{h_j}&\approx_{\sum_{i\in[h_j]}\varepsilon_i^w+\sum_{i\in[h_j-1]}\varepsilon_i^q}U_{\delta_w\beta^w_{h_j}\alpha_{h_j}n}\mid(R^{[j]}_{h_{[j]}},S_{[h_j]},S^{[t]}_{[h_j]},R_{[h_j-1]},R^{[t]}_{[h_j-1]}),\end{aligned}$$

这意味着

$$S_\ell\approx_{\sum_{i\in[\ell-1]}(\varepsilon_i^q+\varepsilon_i^w)}U_{\alpha_\ell n}\mid(V^{[t]}_{h_{[t]}},S^{[t]}_\ell,S_{[\ell-1]},S^{[t]}_{[\ell-1]},R_{[\ell-1]},R^{[t]}_{[\ell-1]}).$$

证明. 下文中，若误差显然可由该断言推出，我们就不再显式写出。情形 $i\leq h_1-1$。我们用归纳法证明

$$\begin{aligned}S_i&\approx_{\sum_{j\in[i-1]}\varepsilon_j^w+\sum_{j\in[i-1]}\varepsilon_j^q}U_{\alpha_i n}\mid(S_{[i-1]},S^{[t]}_{[i-1]},R_{[i-1]},R^{[t]}_{[i-1]}),\\ R_i&\approx_{\sum_{j\in[i]}\varepsilon_j^w+\sum_{j\in[i-1]}\varepsilon_j^q}U_{\delta_w\beta_i^w\alpha_i n}\mid(R_{[i-1]},R^{[t]}_{[i-1]},S_{[i]},S^{[t]}_{[i]}).\end{aligned}\tag{1}$$

在第 1 轮中，由于 $V_1=U_m$， $S_1=U_{\alpha_1 n}$。根据引理 35，$(R_1,\ S_1,\ S_1^{[t]})\approx_\varepsilon(U_{\delta_w\beta_1^w\alpha_1 n},\ S_1,\ S_1^{[t]})$。接着，假设等式 (1) 成立 $\forall i\in[h_1-2]$。由于

$H(V_{i+1}\mid R_{[i]},\ R^{[t]}_{[i]},\ S_{[i]},\ S^{[t]}_{[i+1]})=H(V_{i+1}\mid S_{[i]},\ S^{[t]}_{[i+1]})\geq m-(t+1)(\sum_{j=1}^{i+1}\alpha_j n)\geq m/2$

根据强种子提取器的性质，公式（1）的第一部分成立。由于

$H(X \mid S_{[i]},\ S^{[t]}_{[i]},\ R_{[i]},\ R^{[t]}_{[i+1]}) = H(X \mid R_{[i]},\ R^{[t]}_{[i+1]}) \geq k-(t+1)\delta_w(\sum_{j=1}^{i+1}\beta_{32}\alpha_i n) \geq k/2$ 根据强种子提取器的性质，公式（1）的第二部分成立。情形 $i = h_j$。我们通过归纳法证明

$$\begin{aligned} S_{h_j} &\approx U_{\delta_w\beta^w_{h_j}\alpha_{h_j}n} \mid (V^{[j]}_{h_{[j]}}, S^{[j]}_{h_{[j]}}, S_{[h_j-1]}, S^{[t]}_{[h_j-1]}, R_{[h_j-1]}, R^{[t]}_{[h_j-1]}), \\ R_{h_j} &\approx U_{\alpha_i n} \mid (R^{[j]}_{h_{[j]}}, S_{[h_j]}, S^{[t]}_{[h_j]}, R_{[h_j-1]}, R^{[t]}_{[h_j-1]}), \end{aligned} \tag{2}$$

根据引理 35，对于 $i \in [h_j+1, h_{j+1}-1]$，$j \in [t]$（我们定义了 $h_{t+1}-1 := \ell$），可得

$$\begin{aligned} S_i &\approx U_{\delta_w\beta^w_i\alpha_i n} \mid (V^{[j]}_{h_{[j]}}, S^{[j]}_i, S_{[i-1]}, S^{[t]}_{[i-1]}, R_{[i-1]}, R^{[t]}_{[i-1]}), \\ R_i &\approx U_{\alpha_i n} \mid (R^{[j]}_i, S_{[i]}, S^{[t]}_{[i]}, R_{[i-1]}, R^{[t]}_{[i-1]}). \end{aligned} \tag{3}$$

在第 $h_1-1$ 轮（$V_{h_1}$，$V^1_{h_1}$）$\approx$（$U_m$，$V^1_{h_1}$）。根据引理 35，

$S_{h_1} \approx U_{\delta_w\beta^w_{h_1}\alpha_{h_1}n} \mid (V^1_{h_1}, S^1_{h_1}, S_{[h_1-1]}, S^{[t]}_{[h_j-1]}, R_{[h_1-1]}, R^{[t]}_{[h_1-1]})$。然后，根据引理 35，$R_{h_1} \approx U_{\alpha_1 n} \mid (R^1_{h_1}, S_{[h_1]}, S^{[t]}_{[h_1]}, R_{[h_1-1]}, R^{[t]}_{[h_1-1]})$。假设对于 $i \in h_{[j-1]}$，$j \in [t]$，公式 (2) 成立。由于 $(V_{h_j}, V^j_{h_j}) \approx (U_m, V^j_{h_j})$，且公式 (3) 中关于 $i \in [h_j-1]$ 的第二个方程成立，因此根据引理 35，公式 (2) 对 $i = h_j$ 成立。 □

最后，参数的设置与定理 33 的设置类似。至此，定理 34 的证明完成。

□

低次数建议相关性破坏器 现在，我们准备给出低次数相关性破坏器的构造。我们首先给出低次数建议相关性破坏器的定义。

定义 26（ldACB）。函数 ldACB：$\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$ 是针对线性相关源的建议相关性破坏器，如果满足以下条件。设

- $A$、$A^{[t]}$、$B$、$B^{[t]}$ 是定义在 $\{0,1\}^n$ 上的随机变量，且 $Y$、$Y^{[t]}$ 是定义在 $\{0,1\}^d$ 上的随机变量，使得 $(A, A^{[t]})$ 与 $(B, B^{[t]}, Y, Y^{[t]})$ 独立。此外，$H(A) \geq k$ 且 $Y = U_d$；
- $X = A + B$、$X^i = A^i + B^i$ 对每个 $i \in [t]$ 均成立；
- $\alpha, \alpha^1, \cdots, \alpha^t$ 是 $a$ 比特串，使得 $\alpha \neq \alpha^i$ 对每个 $i \in [t]$都成立；
- 输出的每一比特都是输入 $X$ $(X^i)$ 和 $Y$ $(Y^i)$ 的常数次数多项式，

*then*

$$(\mathsf{ldACB}(X, Y, \alpha) \approx_\varepsilon U_m) \mid \left(\mathsf{ldACB}(X^1, Y^1, \alpha^1), \cdots, \mathsf{ldACB}(X^t, Y^t, \alpha^t)\right).$$

此外，若存在随机变量 $X'$、$A'$、$B'$ 和 $Y'$，使得 $X' = A' + B'$ 且 $(Y = U_d) \mid Y'$，则以下结论也成立：

$$(\mathsf{ldACB}(X, Y, \alpha) \approx_\varepsilon U_m) \mid \left(\mathsf{ldACB}(X', Y', \alpha), \mathsf{ldACB}(X^1, Y^1, \alpha^1), \cdots, \mathsf{ldACB}(X^t, Y^t, \alpha^t)\right).$$

我们指出，定义 26 与标准定义不同之处在于：只要种子相对于篡改后的种子是不可篡改的，它就允许在相同辅助信息下以篡改输出为条件。我们将在方向仿射提取器的证明中使用这一性质。

在我们的构造中，还需要以下函数。

定义 27（FFAssign [CGL22]）。令 FFAssign : $(\{0,1\}^n)^2 \times \{0,1\}^a \to (\{0,1\}^n)^{2a}$ 定义如下。令 $r_0$、$r_1 \in \{0,1\}^n$ 和 $\alpha \in \{0,1\}^a$。令 $\alpha_j$ 表示 $\alpha$的第 $j$个比特。则 FFAssign$(r_0, r_1, \alpha)$ :=$(r_{\alpha_1}, r_{1-\alpha_1}, \cdots, r_{\alpha_a}, r_{1-\alpha_a})$。

算法 7 ldACB$(x, y, id)$

输入：长度分别为 $n, d, a$ 的比特串 $x = w + z, y, id$，其中 $d < n$。
输出：长度为 $n_2$ 的比特串 $y'$。
子程序和参数：
令 LSExt : $\{0,1\}^n \times \{0,1\}^{m_1} \to \{0,1\}^{m_2}$ 来自定理 32，其中 $m_1 = d/(4+2t)$，输出长度 $m_2 = \beta_{32} kd/((8+4t)n)$，熵为 $k/2$，误差为 $\varepsilon_1$。
令 laExt : $\{0,1\}^d \times \{0,1\}^{m_2} \to (\{0,1\}^v)^2$ 来自定理 33 其中 $v = \Omega(m_1/(16+16t)) = \Omega(d/(32(1+t)^2))$，熵为 $d/3$，误差为 $\varepsilon_2$。
令 FFAssign : $(\{0,1\}^v)^2 \times \{0,1\}^a \to (\{0,1\}^v)^{2a}$ 来自定义 27。
令 NIPM$_{2a}$ : $\{0,1\}^n \times (\{0,1\}^v)^{2a} \to \{0,1\}^{n_2}$ 来自定理 34，熵为 $k/2$，误差为 $\varepsilon_3$。

1. 令 $s =$ 切片$(y, m_1)$。2. 令 $q =$ LSExt$(x, s)$。3. 令 $(r_0, r_1) =$ laExt$(y, q)$。4. 令 $(v_1, v_2, \cdots, v_{(2a-1)}, v_{2a}) =$ FFAssign$((r_0, r_1)\ \alpha)$。5. 输出 $v^* =$非交互式概率合并器$_{2a}(x, v_1 \circ \cdots \circ v_{2a})$。

定理 37（ldACB）。对于每个 $0 < \varepsilon < 1$和 $n \in \mathbb{N}$以及每个 $k$、$d$、$t$、$a$，存在一个足够大的 $C$，使得如果

- $k \geq C \max\left\{\left((t+1)^2 \log(1/\varepsilon) n^{4a-1}/m^3\right)^{1/(4a-3)}, (t+1)\sqrt{n}\right\}$;
- $d \geq C(t+1)^2 \max\left\{(t+1)n\sqrt{n}/k, ((t+1)\log(1/\varepsilon) n^{4a-1}/k^{4a-3})^{1/3}\right\}$,

那么存在一个常数 $1 > \eta > 0$和一个 ldACB : $\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$，它是针对线性相关源的低次数建议相关性破坏器，满足：

- $m = \Omega(\eta^{2a} kd/((t+1)^3 n))$;
- ldACB 的每个输出比特都是输入的一个次数为 $2^{\Theta(2a)}$ 的多项式。

*P*证明。 我们将证明，算法 7 给出了这样一个函数。
F首先，我们证明 ldACB满足定义 26。

1. 令 $Q_A :=$ LSExt$(A, S)$，$Q'_A :=$ LSExt$(A', S')$，$Q_B :=$ LSExt$(B, S)$，$Q'_B :=$ LSExt$(B', S')$ 对所有 $i \in [t]$，令 $Q^i_A :=$ LSExt$(A^i, S^i)$，并令 $Q^i_B :=$ LSExt$(B^i, S^i)$。2. 由于 $Y = U_d \mid Y'$，因此 $S = U_{m_1} \mid Y'$。由于 $H(X \mid Y', S, S^{[t]}) \geq H(A) \geq k \geq k/2+(t+2)m_2$，根据引理 35 $Q \approx_{\varepsilon_1} U_{m_2} \mid (Q', Y', S, S^{[t]})$。3. 首先注意，在 $S$ 的条件下，由于 LSExt 是线性函数，$Q = Q_A + Q_B$。此外，进一步在 $Q_B$ 的条件下，$Y$ 与 $Q_A$ 相互独立。由于 $H(Y \mid Y', S, S^{[t]}, Q_B, Q_B^{[t]}, Q') \geq d-(t+1)(m_1+m_2) \geq d/3$，并且 $Q_A \approx_{\varepsilon_1} U_{m_2} \mid (Q', Y', S, S^{[t]}, Q_B, Q_B^{[t]})$，

$$R_0 \approx_{\varepsilon_1+\varepsilon_2} U_v \mid (S, S^{[t]}, Q, Q^{[t]}, Q', Y', R'_0)$$

and

$$R_1 \approx_{\varepsilon_1+\varepsilon_2} U_v \mid (S, S^{[t]}, Q, Q^{[t]}, Q', Y', R_0, R_0', R_0^{[t]}, R_1').$$

4. 根据定义 27，$V_i \approx_{\varepsilon_1+\varepsilon_2} U_v \mid V'$，$\forall i \in [2a]$。此外，对每个 $i \in [t]$，都存在$h_i \in [2a]$ 使得 $V_{h_i} = R_1$ 且 $V_{h_i}^i = R_0^i$。因此，对每个 $i \in [t]$，都存在 $h_i \in [2a]$ 使得$V_{h_i} \approx_{\varepsilon_1+\varepsilon_2} U_v \mid V_{h_i}'$。

5. 由于 $H(X \mid Q, Q', Q^{[t]}, Y', S, S^{[t]}, V'^*) \geq k - (2+t)m_2 - 2av \geq k/2$，根据定理 34，$V^* \approx_{\varepsilon_1+\varepsilon_2+\varepsilon_3} U_{n_2} \mid (V^{[t]*}, V'^*)$。

由于 LSExt 和 laExt 使输出比特的次数增加一个常数，而 NIPM$_{2a}$使次数 $2^{\Theta(2a)}$ 增加。因此，ldACB 的每个输出比特都是输入的次数为 $2^{\Theta(\ell)}$ 的多项式。

最后，参数约束由定理 34的参数约束推出。这就完成了定理 37的证明。

□

## 5.2 线性熵的方向仿射提取器

除低次数相关性破坏器之外，我们还需要以下提取器作为构建模块。

定理 38 ([CG88])。对于每个常数 $\delta > 0$，存在一个多项式时间算法 内积 $:(\{0,1\}^n)^2 \to \{0,1\}^m$，使得如果 $X$是一个 $(n, k_1)$ 源，且 $Y$是一个独立的 $(n, k_2)$ 源，并且$k_1 + k_2 \geq (1+\delta)n$，那么

$$\mathsf{IP}(X, Y) \approx_\varepsilon U_m \mid Y,$$

其中 $\varepsilon = 2^{-\frac{\delta n - m - 1}{2}}$。

定理 39（[Li11]）。对于任意仿射 $t \times r$ 某处随机源 $X$，存在一个函数 AffineSRExt，使得 AffineSRExt($X$) 输出 $m = r/t^{O(\log t)}$ </style}比特，这些比特与均匀分布 $2^{-\Omega(r/t^{O(\log t)})}$</style}相差至多 $2^{-\Omega(r/t^{O(\log t)})}$。此外，输出中的每个比特都是输入比特的 $t^{O(1)}$ 次多项式。

定理 40（带种子非篡改提取器 [Li12]）。对于任意常数 $1 > \delta > 0$，设 $X$ 是一个 $(n,k)$-源，满足 $k = (1/2+\delta)n$，且 $Y$是定义在 $\{0,1\}^{n/2-1}$上、与 $X$独立的均匀分布。设 $b_1$、$\cdots$、$b_{n/2}$是 $\mathbb{F}_{2^{n/2}}$ 的基，将其视为 $\mathbb{F}_2$ 上的向量空间。对于每个 $b_i$，令$\overline{Y}_i = (b_i Y, b_i Y^3)$，其中 $Y$被视为 $\mathbb{F}_{2^{n/2}}^*$ 中的元素，并定义一个比特 $Z_i = \mathsf{IP}(X, \overline{Y}_i)$，其中 IP 是 $\mathbb{F}_2^n$上的内积函数。从 $\{Z_i\}$中选择 $m = \Omega(n)$ 个比特，令 $\mathsf{snmExt}(X, Y) = (Z_{i_1}, \cdots, Z_{i_m})$。令 $\mathcal{A}$：$\{0,1\}^n \to \{0,1\}^n$为任意无不动点函数，则

$$|\mathsf{snmExt}(X, Y), \mathsf{snmExt}(X, \mathcal{A}(Y)), Y - U_n, \mathsf{snmExt}(X, \mathcal{A}(Y)), Y| \leq 2^{-\Omega(n)}.$$

以下关于强线性种子提取器和仿射源的命题对我们很有用。

命题 41 （[Rao09]）。设 $\mathsf{Ext}:$ $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$是一个针对最小熵 $k$、误差为 $\varepsilon < 1/2$的线性强种子提取器。设 $X$ 是任意一个熵为 $k$的仿射源。那么

$$\mathbf{Pr}_{u \leftarrow_U U_d}[\mathsf{Ext}(X, u) = U_m] \geq 1 - \varepsilon.$$

我们使用以下 在论证带有缺陷种子的强种子提取器时所需的引理。

引理 42 ([CGL16])。 设 $\mathsf{Ext}\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 是一个最小熵为 $k$、误差为 $\varepsilon$的强种子提取器。设 $X$是一个 $(n,k)$-源，并设 $Y$是定义在 $\{0,1\}^d$上的一个源，其最小熵为$d-\lambda$。那么

$$\mathsf{Ext}(X,Y) \approx_{2^\lambda \varepsilon} U_m \mid Y.$$

现在我们介绍方向仿射提取器的构造。

---

算法 8 DAExt($x$)

---

输入： $x$ —一个 $n$ 比特串。输出： $z$ —
一个 $m$ 比特串，其长度为 $\Omega(n)$。

---

子程序和参数:
令 $\ell_1 = \mathsf{poly}(2/\delta)$ $\ell_1' = (n/(2(m'+k+1)))^{\log(2d19^{+2})}$ 其中 $k$定义如下， $\ell_2 = \mathsf{poly}(4/\delta) = \ell_3$。令 BasicCond 为算法 1中的基本凝聚器。
令 $\mathsf{SCond}_i : \{0,1\}^n \to \left(\{0,1\}^{n/\ell_i^{1/\log(2d19^{+2})}}\right)^{\ell_i}$ 对于 $i \in \{1,3\}$， $\mathsf{SCond}_2 : \{0,1\}^{n/t} \to \left(\{0,1\}^{n/\left(t\ell_2^{1/\log(2d19^{+2})}\right)}\right)^{\ell_2}$ ，为定理 18中的线性仿射压缩器。
令 IP: $\left(\{0,1\}^{n/\left(t\ell_2^{1/\log(2d19^{+2})}\right)}\right)^2 \to \{0,1\}^{\Omega(n)}$ 为定理中的双源提取器 38
误差为 $\varepsilon_1 = 2^{-\Omega(n)}$，设置为从熵率之和 乐彩
最高可达 $1+2\beta_{18}$ 以上。
设 AffineSRExt 为定理 39 中的仿射某处随机源提取器，其
误差 $\varepsilon_2 = 2^{-\Omega(n)}$。
令 $\mathsf{LSExt}: \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{m'}$ 为定理 32 中的强线性种子提取器
其参数设定为从熵 $\delta n/2$ 中提取，误差为 $\varepsilon_3 = 2^{-\Omega(n)}$。
令 $\mathsf{Enc}: \{0,1\}^n \to \{0,1\}^{\lambda n}$ 为一个渐近良好的线性二元码的编码函数
其具有常数相对码率 $1/\lambda$ 和常数相对距离 $\beta$。
令 $\mathsf{snmExt}: \{0,1\}^{2(m'+k+1)} \times \{0,1\}^{m'+k} \to \{0,1\}^{n_1}$ 为来自
定理 40 中误差为 $\varepsilon_4 = 2^{-\Omega(n)}$的带种子非篡改提取器。取 $m'$ 和 $k$ 使得 $\log(n/(m'+k+1)) \in \mathbb{N}$ 其中

- $m' \le \beta_{18}\delta^2 n/(300t\ell_2\ell_3')$, $k = \Omega(n) \le \frac{n_1}{20\log(\lambda n/n_1)}$.

令 IdACB： $\{0,1\}^n \times \{0,1\}^{n_1} \times \{0,1\}^{\log \ell_1'} \to \{0,1\}^{n_2}$ 为定理 37 中输出长度为 $n_2 = O$（$\delta n_1$/多项式（$\ell_1'$））且误差为 $\varepsilon_5 = 2^{-\Omega(n)}$。

---

令 G 为某个具有$m_1$码字长度和 $\gamma$常数相对距离的渐近良好的线性二元码的生成矩阵。因此，对于某个常数 $\alpha > 0$，G 是一个 $\alpha m_1 \times m_1$ 矩阵。令 $\mathsf{G}_i$ 表示该矩阵的第 $i$ 行。

令 $sc_1 \circ sc_2 \circ \cdots \circ sc_{\ell_1'} = \mathsf{BasicCond}^r \circ \mathsf{SCond}_1(x)$，其中 $r = \log(n/(m'+k+1)) - 1 - \ell^{1/\log(2d19^{+2})}$。将 $x$ 划分为 $t$ 个块 $x = x_1 \circ \cdots \circ x_t$，其中 $t = 2^{\lceil \log(10/\delta) \rceil} \geq \delta/10$，且每个块包含 $n/t$ 个比特。对于每个 $i$，$1 \leq i \leq t$ 执行以下操作。

1. 令 $y_{i1} \circ \cdots \circ y_{i\ell_2} = \mathsf{SCond}_2(x_i)$，其中 $y_{ij}$ 是将 $\mathsf{SCond}_2$ 应用于 $x_i$后得到的矩阵的第 $j$行。注意，$\ell_2 = O(1)$，且每个 $y_{ij}$ 包含 $\Omega(n)$ 个比特。
2. 在 $x$上应用 $\mathsf{BasicCond}^{\log t} \circ \mathsf{SCond}_3$。也就是说，先在 $X$上应用 $\mathsf{SCond}_3$，然后在输出上应用 BasicCond log $t$ 次，从而得到 $\ell_3'$ 个大小相等的块 $\mathsf{BasicCond}^{\log t} \circ \mathsf{SCond}_3(x) = x_1' \circ \cdots \circ x_{\ell_3'}'$，且每个块包含与 $y_{ij}$相同数量的比特。注意，$\ell_3' = O(1)$。
3. 对 $x_{j_1}'$ 和 $y_{ij_2}$ 的每一对应用内积，并输出 $\beta_{18}\delta^2 n/(300t\ell_2\ell_3')$ 比特。令 $sr_i$ 为串接所有内积 $\mathsf{IP}(x_{j_1}', y_{ij_2})$ 的输出所得到的矩阵，即，对于一对 $(x_{j_1}', y_{ij_2})$，$sr_i$ 的每一行都是 $\mathsf{IP}(x_{j_1}', y_{ij_2})$。
4. 令 $r_i =$ 为 $\mathsf{AffineSRExt}(sr_i)$。
5. 令 $u_i =$ 为 $\mathsf{LSExt}(x, r_i)$，将其设置为输出 $m'$ 比特。
6. 将 $u_i$ 划分为 $u_{i1} \circ u_{i2}$，其中 $u_{i1}$ 具有 $k\log(\lambda n/k) \leq n_1/10$ 比特，而 $u_{i2}$ 具有 $\geq m' - n_1/10$ 比特。
7. 将 $\mathsf{Enc}(x)$ 划分为大小相等的 $k$ 个块，使得 $\mathsf{Enc}(x) = \tilde{x}^1 \circ \tilde{x}^2 \circ \cdots \circ \tilde{x}^k$，其中每个块具有 $O(1)$ 比特。将 $u_{i1}$ 划分为 $k$ 个相等的块 $u_{i1}^{(1)} \circ \cdots \circ u_{i1}^{(k)}$。令 $h_i = \tilde{x}^1_{|u_{i1}^{(1)}} \circ \tilde{x}^2_{|u_{i1}^{(2)}} \circ \cdots \circ \tilde{x}^k_{|u_{i1}^{(k)}}$ 和 $\tilde{u}_i = u_i \circ h_i$。
8. 令 $sn_{ij_3}$ 为 snmExt，分别应用于每个 $sc_{j_3}$ 和 $\tilde{u}_i$，并输出 $n_1 \leq m'/100$ 个比特。令 $sn_i$ 为通过串接 $sn_{ij_3}$（对于 $j_3 \in [\ell_1']$）得到的 $\ell_1' \times n_1$ 矩阵，即 $sn_i$ 的第 $j$ 行为 $sn_{ij}$。
9. 令 $\tilde{y}_i = \bigoplus_{j=1}^{\ell_1'}$ 为 $\mathsf{ldACB}(x, sn_{ij}, j)$，并输出 $n_2 \leq m'/10000$ 个比特。
10. 令 $w_i =$ 为 $\mathsf{LSExt}(x, \tilde{y}_i)$，将其设置为输出 $n_3 \leq m'/1000000$ 个比特。
11. 将 $w_i$ 的比特划分为 $s_i =$ 个大小相等的 $\Omega(n)$ 块，其中每个块包含 $c_i$ 个比特，$c_i$ 为稍后选定的某个常数。对于每个 $j = 1$、$\cdots$、$s_i$，通过将 $j$ 的第 $s_i$ 个块中的所有比特相乘，计算一个比特 $v_{ij}$，即 $v_{ij} = \prod^{jc_i}_{(j-1)c_i+1} w_{i\ell}$。

输出 $m_1 = \Omega(n)$ 个比特 $\{z_j = \bigoplus_{i=1}^{t} v_{ij}\}$。

从弥散器到提取器。

对于每个码字 $\mathsf{G}_i$，令 $S_i = \{j \in [m_1] : \mathsf{G}_{ij} = 1\}$ 为满足以下条件的索引集合：码字 $\mathsf{G}_i$ 在这些索引处的比特为 1。定义

$$o_i = \bigoplus z_{j:j\in S_i}$$

为与 $\mathsf{G}_i$ 相关联的比特，即，当码字 $\mathsf{G}_i$ 的第 $j$ 个索引处为 1 时，$o_i$ 是各个 $z_j$ 的异或。

取一个常数 $0 < \beta' \leq \alpha$，其中 $\beta'$稍后选定。输出 $o = (o_1, \cdots, o_{\beta' m_1})$。

定理 43. 对任意常数 $0 < \delta \leq 1$，存在一族函数 $\mathsf{DAExt} : \{0,1\}^n \to \{0,1\}^m$，其中 $m = \Omega(n)$，使得对于任意最小熵至少为 $\delta n$ 的仿射源 $X$，以及任意非零$a \in \{0,1\}^n$，都有

$$(\mathsf{DAExt}(X), \mathsf{DAExt}(X+a)) \approx_\varepsilon (U_m, \mathsf{DAExt}(X+a)),$$

其中 $\varepsilon = 2^{-\Omega(n)}$。

证明。以下证明中，我们考虑两个并行运行的 DAExt，一个的输入为 $X$，另一个的输入为 $X + a =: X'$。我们用 $\{X_i\}_{i\in[t]}, \{SR_i\}_{i\in[t]}, \{R_i\}_{i\in[t]}, \{U_i\}_{i\in[t]}, \{U_i\}_{i\in[t]}, \{SN_i\}_{i\in[t]}, \{SC_i\}_{i\in[\ell_1]}, \{Y_i\}_{i\in[t]}, \{W_i\}_{i\in[t]}$ 表示 $\mathsf{DAExt}(X)$ 生成的随机变量，并用 $\{X'_i\}_{i\in[t]}, \{SR'_i\}_{i\in[t]}, \{R'_i\}_{i\in[t]}, \{U'_i\}_{i\in[t]}, \{U'_i\}_{i\in[t]}, \{SN'_i\}_{i\in[t]}, \{SC'_i\}_{i\in[\ell_1]}, \{Y'_i\}_{i\in[t]}, \{W'_i\}_{i\in[t]}$ 表示 $\mathsf{DAExt}(X+a)$ 生成的随机变量。

在整个证明过程中，我们维护一个随机变量 $Z$。每当一组随机变量被固定时，我们就更新 $Z$，使其表示所有已被固定的随机变量。每次修订 $Z$ 时，我们都会明确说明。初始时，$Z = 0$。

现在我们证明，算法 8构成这样一族函数，并且是高效的。我们首先论证，存在某次迭代 $g$ 使得在前面各次迭代生成的所有随机变量均为条件定值时，$X$ 和 $X_g$ 的熵率均为 $\Omega(\delta)$。

引理 44. 存在 $1 \leq g \leq t$，使得在对以下变量进行任意固定时

$$(X_i, X'_i, SR_i, SR'_i, R_i, R'_i, \tilde{U}_i, \tilde{U}'_i, SN_i, SN'_i, \tilde{Y}_i, \tilde{Y}'_i, W_i, W'_i)_{i\in[g-1]}$$

按顺序，$X$ 是一个仿射源，其 $H(X_g)$ 为 $\geq \delta n/(4t)$，且 $H(X)$ 为 $\geq 3\delta n/5 + \delta n/(3t)$。

证明。根据引理 10，将 $X$ 划分为 $t$ 个块时，存在总和为 $\delta n$ 的正整数 $k_1$、$\cdots$、$k_t$，使得对于任意 $i \in [t]$，在固定 $X_1$、$\cdots$、$X_{i-1}$、$H(X_i)$ 时，$= k_i$。因此，必然存在一个 $i$ 使得 $k_i \geq \delta n/(3t)$。令 $g$ 为满足 $H(X_g) = k_g \geq \delta n/(3t)$ 的最小索引。

1. 考虑仿射源 $X$ 和 $X' = X + a$。一旦固定 $(X_i = x_i)_{i\in[g-1]}$，$(X'_i := X_i + a_i = x_i + a_i)_{i\in[g-1]}$ 也就固定了。由于 $X_i$ 是 $X$的仿射函数，固定之后，$X$ 和 $X'$仍然是仿射源。由引理 10，固定之后 $H(X_g) = H(X'_g) \geq k_g \geq \delta n/(3t)$ 和 $H(X) = H(X') = \sum_{i=g}^{t} k_i \geq \delta n - (t-1) \cdot \delta n/(3t) \geq 2\delta n/3 + \delta n/(3t)$。现在令 $Z = \{X_i, X'_i\}_{i\in[g-1]}$。

2. 注意，在固定 $(X_i = x_i)_{i\in[g-1]}$（因此 $(X'_i = x'_i)_{i\in[g-1]}$）的条件下，$(SR_i)_{i\in[g-1]}$和 $(SR'_i)_{i\in[g-1]}$ 都是 $X$的仿射函数。一般而言，固定 $(SR_i = sr_i)_{i\in[g-1]}$ 并不一定会固定 $(SR'_i = sr'_i)_{i\in[g-1]}$，而在最坏情况下，$SR'_i = sr'_i$ 可能与 $SR_i = sr_i$线性无关。令 $\overline{SR} = SR_1 \circ \cdots \circ SR_{g-1}$ 和 $\overline{SR'} = SR'_1 \circ \cdots \circ SR'_{g-1}$。由引理 9，由于 $\overline{SR} \circ \overline{SR'}$ 至多有 $(\beta_{18}\delta^2 n/(300t)) \cdot t \cdot 2 = \beta_{18}\delta^2 n/150$ 比特，$H(X \mid Z, \overline{SR} \circ \overline{SR'}) \geq 2\delta n/3 + \delta n/(3t) - H(\overline{SR} \circ \overline{SR'}(X)) \geq 2\delta n/3 + \delta n/(3t) - \beta_{18}\delta^2 n/150$。

请注意，固定 $\overline{SR} \circ \overline{SR'}$也会固定 $\{R_i, R'_i\}_{i\in[g-1]}$。现在令 $Z = Z \cup \{SR_i, SR'_i, R_i, R'_i\}_{i\in[g-1]}$。

3. 令 $U \circ U' = U_1 \circ \cdots \circ U_{g-1} \circ U'_1 \cdots \circ U'_{g-1}$，则在对 $Z$ 进行任意固定后，$U$ 是 $X$ 的仿射函数，并且至多有 $(\beta_{18}\delta^2 n/(300t\ell_2\ell'_3)) \cdot t \cdot 2 = \beta_{18}\delta^2 n/(150\ell_2\ell'_3)$ 个比特。因此，根据引理 9 $H(X \mid Z, \tilde{U} \circ \tilde{U'}), \geq 2\delta n/3 + \delta n/(3t) - \beta_{18}\delta^2 n/150 - \beta_{18}\delta^2 n/(150\ell_2\ell'_3)$。现在，$Z = Z \cup \{U_i, U_i\}_{i\in[g-1]}$。

4. 令 $\overline{SN} \circ \overline{SN'} = SN_1 \circ \cdots \circ SN_{g-1} \circ SN'_1 \cdots \circ SN'_{g-1}$，则在对 $Z$ 进行任意固定后，$\overline{SN} \circ \overline{SN'}$ 是 $X$ 的仿射函数，并且至多有 $\beta_{18}\delta^2 n/$（$15000\ell_2\ell'_3$）个比特。因此，根据引理 9 $H$（$X \mid Z$, $\overline{SN} \circ \overline{SN'}$） $\geq 2\delta n/3 - \beta_{18}\delta^2 n/150 - 1.01 \cdot \beta_{18}\delta^2 n/$（$150\ell_2\ell'_3$）。现在，$Z = Z \cup \{\overline{SN}, \overline{SN'}\}$。

5. 令 $Y \circ Y' = Y_1 \circ \cdots \circ Y_{g-1} \circ Y'_1 \cdots \circ Y'_{g-1}$，则在 $Z$取任意固定值的条件下， $Y \circ Y'$是关于 $X$ 的仿射函数，且至多有 $\beta_{18}\delta^2 n/(1500000\ell_2\ell'_3)$比特。因此，根据引理9 $H(X \mid Z, \overline{Y\,Y'})$, $\geq 2\delta n/3 + \delta n/(3t) - \beta_{18}\delta^2 n/150 - 1.0101 \cdot \beta_{18}\delta^2 n/(150\ell_2\ell'_3)$。现在，$Z = Z \cup \{YY'\}$。

6. 令$\overline{W} \circ \overline{W'} = W_1 \circ \cdots \circ W_{g-1} \circ W'_1 \cdots \circ W'_{g-1}$，则在 $Z$取任意固定值的条件下， $\overline{W} \circ \overline{W'}$是关于 $X$ 的仿射函数，且至多有$\beta_{18}\delta^2 n/(150000000\ell_2\ell'_3)$比特。因此，根据引理 9 $H(X \mid Z, W, W')$, $\geq 2\delta n/3 + \delta n/(3t) - \delta^2 n/150 - 1.010101 \cdot \beta_{18}\delta^2 n/(150\ell_2\ell'_3)$。现在，$Z = Z \cup \{W, W'\}$。

7. 因此， $H(X_g \mid Z) \geq \delta n/(3t) - \beta_{18}(\delta^2 n/150 + \delta^2 n/(150\ell_2\ell'_3) + \delta^2 n/(15000\ell_2\ell'_3) + \delta^2 n/(1500000\ell_2\ell'_3) + \delta^2 n/(150000000\ell_2\ell'_3)) = \delta n/(3t) - \beta_{18}\delta^2 n/150 - 1.010101 \cdot \beta_{18} \cdot \delta^2 n/(150\ell_2\ell'_3) > \delta n/(4t)$ 且$H(X \mid Z) \geq 2\delta n/3 + \delta n/(3t) - \beta_{18}\delta^2 n/150 - 1.010101 \cdot \beta_{18}\delta^2 n/(150\ell_2\ell'_3) > 3\delta n/5 + \delta n/(3t)$。

□

引理 45. 对于 $X_g$ 的进一步固定，以 $1 - 2^{-\Omega(n)}$ 的概率， $R_g$ 是 $2^{-\Omega(n)}$-接近均匀的。

证明。我们检查 DAExt 在良块 $X_g$ 上的执行过程，直至步骤 4。

1. 根据引理 44， $H(X_g \mid Z) \geq \delta n/(4t)$ 定理 $\Longrightarrow^{18}$ $Y_{g1} \circ \cdots \circ Y_{g\ell_2} := \mathsf{SCond}(X_g) =$ 某处随机-率-$(1/2 + \beta_{18})$源。不失一般性，假设 $Y_{gi}$ 的熵率为 $1/2 + \beta_{18}$。 2. 根据引理 8， $\exists A_g$， $B_g$ 使得 $X = A_g + B_g$， $X_g(X) = X_g(A_g)$ $H(X_g) = H(A_g)$，且 $X_g$ 与 $B_g$ 独立。 3. 在固定 $X_g$ 之后， $B_g$（因此 $X$）的最小熵至少为 $3\delta n/5 + \delta n/(3t) - \delta n/t \geq \delta n/4$。现在，令 $Z = Z \cup \{X_g\}$。 4. 由于 $\ell'_3$ 块 $X = \tilde{X}_1 \circ \cdots \circ \tilde{X}_{\ell'_3}$ 是通过在 $X$ 上应用 $\mathsf{BasicCond}^{\log t} \circ \mathsf{SCond}_3$得到的，因此每个 $X_i$ 都是 X 的线性函数。于是，根据引理 8 可得 $X_i(X) = X_i(A_g) + X_i(B_g)$。令 $X_i(A_g) = A_{gi}$ 以及 $X_i(B_g) = B_{gi}$。根据定理 18，存在一个熵率至少为$1/2 + \beta_{18}$的块，令 $B_{gj}$ 为这样一个块。5. 注意，$\mathrm{IP}(X_j, Y_{gi}) = \mathrm{IP}(A_{gj}, Y_{gi}) + \mathrm{IP}(B_{gj}, Y_{gi})$。由于 $Y_{gi}$ 的熵率为 $\geq 1/2 + \beta_{18}$，而 $B_{gj}$ 的熵率为 $\geq 1/2 + \beta_{18}$，根据引理38，对于 $Y_{gi}$的固定，以概率 $(1 - \varepsilon_1)$，$\mathrm{IP}(B_{gj}, Y_{gi})$ 接近均匀分布。由于 $A_g$（因此 $A_{gj}$）已被固定，随机变量 $\mathrm{IP}(A_{gj}, Y_{gi})$ 也随之固定。因此，$\mathrm{IP}(B_{gj}, Y_{gi})$ $\varepsilon_1$接近均匀分布意味着 $\mathrm{IP}(Y_{gi}, X_j)$ $\varepsilon_1$ 接近均匀分布。因此，对于 $X_g$ 的固定，以概率 $(1 - \varepsilon_1 - \varepsilon_2)$， $SR_g$ 是与某处随机源 $(\varepsilon_1 + \varepsilon_2)$ 接近的。

6. $SR_g \approx_{\varepsilon_1+\varepsilon_2}$ 某处随机源 定理$\Longrightarrow^{39}$ $R_g \approx_{\varepsilon_1+\varepsilon_2}$ 均匀。

□

引理 46. 以概率 $1-2^{-\Omega(n)}$ 在进一步固定 $\left(SR_g, SR'_g, R_g, R'_g\right)$, $U_g$ 后是均匀的。

证明。由于 $X_g$ 是 $X$的线性函数，在将其固定为任意值后， $X$仍然是仿射源。此外，在将 $X_g$ 固定为某个值后（因此 $X'_g$ 也随之固定）， $SR_g$ 和 $SR'_g$ 是 $X$的线性函数。根据引理 8，存在相互独立的仿射源 $A_g$ 和 $B_g$，使得 $X = A_g + B_g$, $SR_g \circ SR'_g(X) = SR_g \circ SR'_g(A_g)$ and $H(SR_g \circ SR'_g) = H(A_g)$。因此 $H(B_g) = H(X) - H(A_g) = H(X) - H(SR_g \circ SR'_g) \geq \delta n/2$。

接下来注意到， $R_g$ 是 $SR_g$ 的确定性函数，因此与 $\tilde{B}_g$独立。此外，根据引理 45， $R_g$ 与均匀分布的距离为 $2^{-\Omega(n)}$。现在，根据定理 32 和命题 41，在固定 $R_g$的各种取值后（因此在固定 $SR_g$ 的各种取值后），以 $(1-\varepsilon_3)$ 的概率，LSExt$(B_g, R_g)$ 是均匀的。由于 LSExt 是线性函数，且 LSExt$(A_g, R_g)$ 已固定，因此在固定 $R_g$的各种取值后，以 $(1-\varepsilon_3)$ 的概率，LSExt$(X, R_g)$ 是均匀的。□

此时，令 $Z = Z \cup \{SR_g, SR'_g, R_g, R'_g\}$。我们已经证明，对于 $Z$ 的固定，在高概率意义下， $U_g$ 是均匀的。现在，我们希望证明 $U_g$ 和 $U'_g$（即附加了辅助信息的 $U_g$ 和 $U'_g$）是线性相关的，即存在一个不含 $\tilde{}$ 不动点的仿射函数 $\mathcal{A}$，使得 $\mathcal{A}(U_g) = U'_g$。我们将在下面的两个引理中证明这一点。

引理 47。在 $Z$ 的条件为定值时，存在 $A_g$、 $B_g$，使得 $X = A_g + B_g$、 $U_g(X) = U_g(A_g)$, 且 $U_g(B_g) = 0$。此外，在 $U'_g(B_g)$ 的任意固定条件下， $U'_g$ 与 $U_g$ 线性相关。

证明。根据引理 8，存在 $A_g$、 $B_g$，使得 $X = A_g + B_g$、 $U_g(X) = A_g$，且 $U_g(B_g) = 0$。此外，存在一个仿射函数 $L$，使得 $A_g = L(U_g)$。现在，在 $U'_g(B_g)$ 被固定的条件下，$U'_g$ 是 $A_g$ 的仿射函数，因此也是 $U_g$ 的仿射函数。□

提醒一下，在这一阶段，我们有

$$Z = \{X_i, X'_i, SR_i, SR'_i, R_i, R'_i, \tilde{U}_i, \tilde{U}'_i, SN_i, SN'_i, \tilde{Y}_i, \tilde{Y}'_i, W_i, W_i\}_{i\in[g-1]} \cup \{X_g, X'_g, SR_g, SR'_g, U'_g(\tilde{B}_g)\}.$$

引理 48. 在 $U_g$ 均匀的事件条件下，对于固定 $(Z, U_{g1}, U'_{g1}, H_g, H'_g)$ 且概率至少为 $1-2^{-\Omega(n)}$（在 上），存在一个仿射映射 $\mathcal{A}\colon \{0,1\}^{|U_g|} \to \{0,1\}^{|U_g|}$，其没有固定点，使得 $U'_g = \mathcal{A}(U_g)$。

证明。我们首先证明，以高概率 $U_g \neq U'_g$。

- 如果 $U_{g1} \neq U'_{g1}$，那么 $U_g(X) \neq U'_g(X)$ 始终成立。
- 如果 $U_{g1} = U'_{g1}$，我们将以概率 $1-2^{-\Omega(n)}$证明 $H_g \neq H'_g$。注意，这等价于 $H_g \oplus H'_g \neq 0 \iff (\tilde{X}^1 \oplus \tilde{X}'^1)|_{U_{g1}} \circ (\tilde{X}^2 \oplus \tilde{X}'^2)|_{U_{g1}} \circ \cdots \circ (\tilde{X}^k \oplus \tilde{X}'^k)|_{U_{g1}} \iff \tilde{a}^1|_U \circ \tilde{a}^2|_{U_{g1}} \circ \cdots \circ \tilde{a}^k|_{U_{g1}}$，概率为 $1-2^{-\Omega(n)}$，其中 $\tilde{a}^j$是通过将 Enc$(a)$（即源 $X$ 与其篡改版本 $X' = X + a$之间编码后的移位）划分为 $k$个等长块而得到的，使得 Enc$(a)$ $= \tilde{a}^1 \circ \tilde{a}^2 \circ \cdots \circ \tilde{a}^k$。令 $\ell_1$、 $\cdots$、 $\ell_k$ 分别表示每个

块中非零比特的数量。由于 $\mathsf{Enc}(a)$ 中至少有 $\beta$ 比例的比特不同于码字 $0 = \mathsf{Enc}(0)$ 中的比特，因此 $\sum_{i=1}^{k} \ell_i \geq \beta\lambda n$。所以我们有

$$\begin{aligned}
\mathbf{Pr}\left[H_g \neq H'_g \mid U_{g1} = U'_{g1}\right] &= 1 - \prod_{i=1}^{k}\left(1 - \frac{\ell_i}{\lambda n/k}\right) \\
&\geq 1 - \left(\frac{\sum_{i=1}^{k}\left(1 - \frac{\ell_i}{\lambda n/k}\right)}{k}\right)^k \\
&\geq 1 - (1-\beta)^k \\
&\geq 1 - 2^{-\Omega(n)}. \qquad (k = \Omega(n))
\end{aligned}$$

因此，总体而言，以概率 $1 - 2^{-\Omega(n)}$，$\tilde{U}_g \neq \tilde{U}'_g$。根据引理 47，$U'_g$与$\tilde{}$在 $U_g$ 条件为定值时 $Z$呈线性相关。现在注意到，$U_g$ 是由 $U_g$ 与 $H_g$ 组合而成，$U_{g1}$，$H_g$ 固定。$\tilde{}$ $\tilde{}$对于$U'_g$，同样成立。因此，存在某个仿射映射 $\mathcal{A}$，使得 $\mathcal{A}(U_g) = U'_g$。 □

$\tilde{}$ $\tilde{}$引理49. 在进一步固定$(U_g, U'_g)$的条件下，存在一个常数 $\beta > 0$，使得$SC$ 是一个$(1/2 + \beta)$仿射某处随机源。

$\tilde{}$ $\tilde{}$证明。令 $Z = Z \cup \{U_g, U'_g\}$。容易看出，界 $H(X \mid Z) \geq \delta n/2$ 成立。因此，根据定理 18，$SC$ 是一个$(1/2 + \beta_{18})$仿射某处随机源。 □

$\tilde{}$ $\tilde{}$引理50. 在 $U_g$ 为均匀分布，且 $U_g$ 与$U'_g$ 线性相关的条件下，$SN_g = \mathsf{snmExt}(SC_1, \tilde{U}_g) \circ \cdots \circ \mathsf{snmExt}(SC_{\ell'_1}, \tilde{U}_g)$ 与某个仿射某处随机源 $2^{-\Omega(n)}$ 接近。此外，存在 $h \in [\ell'_1]$ 使得

$$SN_{gh} \approx_{2^{-\Omega(n)}} U_{n_1}) \mid SN'_{gh}.$$

证明。由于 $SC$ 是一个（$1/2 + \beta$）仿射某处随机源，因此存在一个 $h \in [\ell_3]$ 使得$H$（$SC_h$）$\geq 1/2 + \beta$。对于这些种子，在对（$U_{g1}$、$U'_{g1}$）进行条件化时，$2^{0.22n_1}$相较于 $\tilde{U}_g$均匀分布的偏差至多为40根据定理42，在（$\tilde{U}_g$、$\tilde{U}'_g$）为条件时，我们有

$$(\mathsf{snmExt}(SC_h, \tilde{U}_g) \approx_{2^{0.22n_1}\varepsilon_4} U_{n_1}) \mid \mathsf{snmExt}(SC'_h, \tilde{U}'_g).$$

由于 $2^{0.22\,n_1}\varepsilon_4 = 2^{-\Omega(n)}$，snmExt 是在 $Z$ 条件为定值时关于 $SC$ 的线性函数，成立的是 t

$$(SN_{gh} \approx_{2^{-\Omega(n)}} U_{n_1}) \mid SN'_{gh},$$

并且 $SN_g$ 与仿射某处随机源 $2^{-\Omega(n)}$ 接近。 □

引理51. 以 $1 - 2^{-\Omega(n)}$ 的概率，在进一步固定 $SN'_{gh} = \mathsf{snmExt}(SC'_h, \tilde{U}'_g)$ 后，$\tilde{Y}_g \oplus \tilde{Y}'_g$ 是均匀的。

证明。首先注意到，二者 $X$ 和 $SN_g$（同样地，$X'$ 和 $SN'_g$）都是仿射源。根据定理 37，我们有

$$\mathsf{IdACB}(X, SN_{gh}, h) \approx_{\varepsilon_5} U_{n_2} \mid \underbrace{\mathsf{IdACB}(X', SN'_{gh}, h)}_{\substack{\text{same advice but} \\ SN_{gh} \approx_{2^{-\Omega(n)}} U_{n_1} \mid SN'_{gh}}}, \underbrace{\left\{\mathsf{IdACB}(X, SN_{gj}, j), \mathsf{IdACB}(X', SN'_{gj}, j)\right\}}_{\substack{\text{the set contains the output } \forall j \in [\ell'_1]\setminus\{h\} \\ \text{different advice}}}.$$

因此，可以得出的是

$$\begin{aligned}\tilde{Y}_g \oplus \tilde{Y}'_g &= \bigoplus_{j\in[\ell'_1]} \mathsf{IdACB}(X, SN_{gj}, j) \oplus \bigoplus_{j\in[\ell'_1]} \mathsf{IdACB}(X', SN'_{gj}, j) \\ &= \mathsf{IdACB}(X, SN_{gh}, h) \oplus \left( \bigoplus_{j\in[\ell'_1]\setminus\{h\}} \mathsf{IdACB}(X, SN_{gj}, j) \oplus \bigoplus_{j\in[\ell'_1]} \mathsf{IdACB}(X', SN'_{gj}, j) \right) \\ &\approx_{2^{-\Omega(n)}} U_{n_2}.\end{aligned}$$

□

令 $Z = Z \cup \{SN'_{gh}\}$.

引理 52. 在固定 $(Z, \tilde{Y}_g, \tilde{Y}'_g)$ 的过程中，以 $1 - 2^{-\Omega(n)}$ 的概率， $W_g$ 在 $W'_g$ 在条件为定值时是均匀的。此外，存在一个随机变量 $\hat{} B_g$ ，在该变量条件为定值时， $X$ 是 $W_g$ 的仿射函数。

~ ~证明. 根据引理 35， $W_g \approx_{2^{-\Omega(n)}} U_{n_3} \mid W'_g$。由于在 $(Y_g, Y'_g)$ 在条件为定值时， $W_g$ 是 $X$ 的线性函数 $\hat{}\hat{}\hat{}$，根据引理 8，存在 $A_g$ 和 $B_g$，使得 $W_g(X) = W_g(A_g)$ 且 $W_g(B_g) = 0$。此外，在固定 $\hat{} B_g$ 后， $X$ 是 $W_g$ 的仿射函数。现在，令 $Z = Z \cup \{\tilde{Y}_g, \tilde{Y}'_g, W'_g, \hat{} B_g\}$。 □

引理 53. 对所有 $1 \le j \le m_1$， $z_j$ 都是关于 $x$ 的比特的常数次数多项式。

证明。
我们来考察在算法 8 的循环中，多项式的次数如何累积。

1. 根据与 [Li11], 类似的论证， $u_i$ 的每个比特都是 $O$次数为 (1) 的多项式，关于 $x$。

2. 在步骤 7 中，只需考虑 $h_i$ 的每个比特。对于每个 $j \in [k]$， $\tilde{x}_j$ 有 $O(1)$ 个比特。由于要从 $\tilde{x}_j$ 中采样一个比特，每个 $u_{i1}^{(j)}$ 的长度只需为 $\log|\tilde{x}_j| = O(1)$，即可编码 $\tilde{x}_j$ 的所有索引。因此， $h_i$ 的第 $j$ 个比特是 $\tilde{x}_j$ 的线性函数，也是 $u_{i1}^{(j)}$ 的次数为 $\log|\tilde{x}_j| = O(1)$ 的多项式。

3. 在步骤 8 中， $sn_i$ 的每个比特都是关于 $sc$ 和 $\tilde{u}_i$ 的双线性映射。因此， $sn$ 的每个比特都是关于 $x$ 的次数为 $O(1)$ 的多项式。

4. 在步骤 9 中， $\tilde{y}_i$ 的每个比特都是 $O$的比特的次数为 (1) 的多项式，并且 $x$ 和 $sc$依据定理37。

5. 在步骤 10 中， $w_i$ 的每个比特都是输入的常数次数多项式，根据定理32。

6. 在步骤 11 中，由于每个 $c_i$ 都是常数，因此通过对 $c_i$ 的比特取乘积所得到的单项式的次数是常数。因此，对于所有 $j \in [s_i]$ ， $v_{ij}$ 的每个比特都是$x$的次数为 $O(1)$ 的多项式。

最后，显然，对于每个 $j \in [m_1]$， $z_j$ 的每个比特都是 $x$ 的常数次数多项式。 □

引理 54. 对于任意整数 $s \in [m_1]$，令 $\tilde{Z} = (\tilde{Z}_1, \cdots, \tilde{Z}_s)$，其中 $\tilde{Z}_j = \bigoplus_{i=g}^{t} V_{ij}$。那么，在固定 $W'_g$ and $\hat{} B_g = \hat{} b$ 的条件下，存在某个 $b \in \{0,1\}^s$，使得

$$\left|\mathsf{Supp}(\mathsf{DAExt}(X) \mid \mathsf{DAExt}(X') = b)\right| = 2^s.$$

证明。首先注意，对于所有 $i \in [t]$， $W_i$ 的每个比特都是 $X$.~ ~中该比特的常数次数多项式。因此，在固定 $Z$ 的条件下，对于每个 $i \geq g+1$， $W_i$ 和 $W_i'$ 的每个比特都是关于 $W_g$ 中各比特的次数为 $\leq c(\delta)$ 的多项式。因此，对于每个 $i \geq g+1$， $W_i$ 和 $W_i'$ 中该比特的次数是 $W_g$ 中该比特次数的 $c(\delta)$ 倍。因此，如果

$$c_i > c(\delta) c_{i+1}, \ \forall i$$

那么，多项式 $\bigoplus_{i=g+1}^{t} V_{ij}$ 和 $\bigoplus_{i=g+1}^{t} V'_{ij}$ 的次数小于 $V_{gj}$的次数。~ ~因此，存在一种固定 $\{V_{ij}, V'_{ij}\}_{i\in[g+1-t],\ j\in[m]}$ 使得 $Z_j \oplus Z'_j$在 $\{0,1\}$中可以取两个值~。由于 $V'_{gj}$ 已被固定， $Z'_j$ 也随之固定。这确保存在 $\{\tilde{z}'_j\}_{j\in[m]}$，使得 ~ ~$Z_j \mid (Z'_j = \tilde{z}'_j)$ 在 $\{0,1\}$中可以取两个值。

~ ~接下来，我们证明 $Z_j$ 在 $Z_S$ 条件化时可以在 $\{0,1\}$ 中取两个值，其中 $S \subseteq [m] \setminus \{j\}$，且 ~ ~ ~ ~$Z_S$ 表示 $(Z_i)_{i\in S}$。假设存在某个 $(z_i)_{i\in S}$，使得当 $(Z_i = z_i)_{i\in S}$时，将 $Z_j$ 固定为 $z_j$，则有

$$P_g = \prod_{i\in S} \left(\tilde{Z}_i + z_i + 1\right) \left(\tilde{Z}_j + z_j\right) \equiv 0.$$

然而，这不可能成立，因为 $P_g$ 有一个由 $W_g$ 中的比特构成的单项式 $V_{gj}$，该单项式不同于 $Z_i$ 中相同次数的单项式~（如果 $z_j = 1$）。由于 $W_g$ 是均匀的， $V_{gj}$ 在对 $Z_S$进行任意固定后都非零。因此， $P_g$ 不可能始终为 0。 □

从弥散器构造提取器的技术，本质上遵循与 [Li11]相同的思路。为使证明完整，我们在此重述这些技术。

引理 55。随机变量 $O_1$、 $\cdots$、 $O_{\alpha m_1}$ 构成一个 $\varepsilon$-偏差空间。

证明。令 $\varnothing \neq T \subseteq [\alpha m_1]$、 $S_i = \{j \in [m_1] : \mathsf{G}_{ij} = 1\}$、 $S_T = \{j \in [m_1] : \ \oplus_{i\in T} \mathsf{G}_{ij} = 1\}$。那么

$$\bigoplus_{i\in T} O_i = \bigoplus Z_{j:j\in S_T}.$$

因为码字的任意非零线性组合仍然是码字。集合 $S_T$ 的基数至少为 $\gamma m_1$。现在注意，在对 $Z$进行固定并以其为条件时，每个 $O_i$ 都是关于 $W_g$的次数为 $c_g$ 的多项式。此外， $\forall i \in S_T$,

$$O_i = V_{gi} \oplus \bigoplus_{j=g+1}^{t} V_{ji}.$$

由于对于所有的 $i \in S_T$， $V_{gi}$ 都是 $W_g$ 的比特中某个不相交集合（相对于 $V_{g\ell}$ 的 $\forall \ell \in S_T \setminus \{i\}$）的乘积，因此很容易看出， $\{V_{gi} : \ i \in S_T\}$ 是同一个函数的 $|S_T| \geq \gamma m_1$ 个独立副本的集合，我们将其简单地称为 $f$。由于 $P := \bigoplus_{i\in T} \left(\bigoplus_{j=g+1}^{t} V_{ji}\right)$ 的次数小于 $c_g - 1$，而 $f$ 的次数为 $c_g$,

$$\mathsf{Cor}(f, P) \leq 1 - 2^{-c_g}.$$

根据定理 17,

$$\mathsf{Cor}(f^{\oplus|S_T|}, P) \leq \exp(-\Omega(|S_T|/(4^{c_g-1} \cdot (c_g - 1)))) \leq 2^{-\Omega(\gamma m_1)}.$$

由于我们对 $T$ 的选择是任意的，且此类选择至多有 $2^{\alpha m_1}-1$ 种，因此存在绝对常数 $c_0$ 使得

$$\mathsf{Cor}(f^{\oplus|S_T|},P)\le 2^{-c_0\gamma m_1)}$$

对于任意 $\varnothing\neq T\subseteq[\alpha m_1]$。
由于 $f^{\oplus|S_T|}$ 是均匀的，因此有

$$\Delta(\bigoplus_{i\in T}O_i,U)\le 2^{-c_0\gamma m_1}.$$

我们得出 $O_1$、$\cdots$、$O_{\alpha m_1}$ 构成一个 $\varepsilon$-偏差空间。 □

现在根据引理 16，

$$\Delta(O-U_{\beta' m_1})\le 2^{\beta' m_1/2}\cdot 2^{-c_0\gamma m_1}.$$

选择 $0<\beta'\le\alpha$ 使得 $\beta'\le c_0\gamma$。于是

$$\Delta(O-U_m)\le 2^{-c_0 m/2}.$$

因此，算法 8的输出是 $m$ 个比特，与均匀分布相差 $2^{-\Omega(m)}$。 □

## 5.3 次线性熵源的方向仿射弥散器和提取器

在本小节中，我们将展示如何将算法 8 的熵要求降至次线性。我们首先研究如何对弥散器实现这一点。

定理 56. 存在一个常数 $c>1$ 和一个高效函数族 DADisp：$\{0,1\}^n\to\{0,1\}^m$，使得 $m=n^{\Omega(1)}$，并且对于每个熵为 $cn(\log\log n)^2/\log n$的仿射源 $X$，存在某个 $b\in\{0,1\}^m$，使得 $|\mathsf{Supp}(\mathsf{DAExt}(X)\mid \mathsf{DAExt}(X+a)=b)|=2^m$。

证明概要。 弥散器构造是算法 8，直到“从弥散器到提取器”阶段为止；但目前，我们不对输入的熵作任何假设。在降低熵要求时，我们主要关注两个约束量—— $W_g$的比特长度，以及 $\{W_i$ 中每个比特的次数：$i\in[g+1,t]\}$。

我们首先逐步考察 $W_g$ 的长度。
步骤 1。 在算法 8 的步骤 1 中，根据定理 18，我们得到 $\ell_2=\mathsf{poly}(1/\delta)$，且每个 $Y_{gj}$ 有$n/\mathsf{poly}(1/\delta)$ 比特。

步骤 2。我们得到 $\ell_3'=\mathsf{poly}(1/\delta)$。
步骤 3。 矩阵 $SR_g$ 的总行数为 $\ell_2\ell_3'=\mathsf{poly}(1/\delta)$，每行有$\delta^2 n/(300t\ell_2\ell_3')=n/(\mathsf{poly}(1/\delta))$ 比特。根据定理 38，误差为 $2^{-n/\mathsf{poly}(1/\delta)}$。
步骤 4。 我们应用AffineSRExt。根据定理 39，我们得到每个 $R_g$ 有 $n/\mathsf{poly}(1/\delta)^{O(\log(1/\delta))}$ 比特，误差为 $2^{-n/\mathsf{poly}(1/\delta)^{O(\log(1/\delta))}}$。
步骤 5。 根据定理 32，应用 LSExt 后，$U_g$ 具有 $n/(1/\delta)^{O(\log(1/\delta))}$ 比特，误差为 $2^{-n/(1/\delta)^{O(\log(1/\delta))}}$。
步骤 8。首先注意，根据定理 18，$SC_g$ 具有 $n/(1/\delta)^{O(\log(1/\delta))}$ 行，根据定理 40，$SN_g$ 的每一行具有 $n/(1/\delta)^{O(\log(1/\delta))}$ 比特，误差为$2^{-n/(1/\delta)^{O(\log(1/\delta))}}$。

步骤 9。 根据定理 37，$\tilde{Y}_g$ 具有 $n/(1/\delta)^{O(\log(1/\delta))}$ 比特，误差为 $2^{-n/(1/\delta)^{O(\log(1/\delta))}}$。
步骤 10。 根据定理 32，应用 LSExt 后，$W_g$ 有 $n/(1/\delta)^{O(\log(1/\delta))}$ 比特，误差为 $2^{-n/(1/\delta)^{O(\log(1/\delta))}}$。现在我们检查 步骤 11 中生成的多项式次数是否满足要求：

在定理 43 的分析中，该定理添加了约束 $c_i > c(\delta)c_{i+1}, \forall i$。首先注意，经过一系列对随机变量的固定后，$X$ 是 $W_g$ 的仿射函数。现在根据定理 18，$\mathsf{SCond}_2(X_i)$ 和 $\mathsf{BasicCond}^{\log t} \circ \mathsf{SCond}_3(x)$ 的每个比特都是输入比特的线性函数。函数 IP 是一个二次多项式。因此，$SR_i$ 的每个比特都是输入比特的二次多项式。由于 AffineSRExt 的输出的每个比特都是输入比特的次数 $\mathrm{poly}(1/\delta)$ 多项式，因此，$R_i$ 的每个比特都是 $W_g$ 的比特的次数 $\mathrm{poly}(1/\delta)$ 多项式。根据定理 32，$U_i$ 的每个比特都是输入比特的常数 ˜ 次数多项式。由于 Enc 是线性的，且 $X_i$ 有 $O(1)$ 比特，因此 $h_i$ 的每个比特都是输入的常数次数多项式。根据定理 18 和定理 40，$\tilde{SN}_{ij}$ 的每个比特都是输入比特的常数次数多项式。根据定理 37，$Y_i$ 的每个比特都是输入比特的次数 $2^{\log(\ell_1')} = 2^{\log((1/\delta)^{O(\log(1/\delta))})} = (1/\delta)^{O(\log(1/\delta))}$ 次数多项式。根据定理 32，$W_i$ 的每个比特都是输入比特的常数次数多项式。因此，我们得出结论：对于每个 $i \geq g+1$，$W_i$ 的每个比特都是 $W_g$ 的比特的次数 $(1/\delta)^{O(\log(1/\delta))}$ 多项式。因此，我们有

$$c(\delta) = (1/\delta)^{O(\log(1/\delta))}.$$

由于我们需要对每个 $1 \leq i \leq 10/\delta$ 都有 $c_i > c(\delta)c_{i+1}$，因此对于所有的 $c_i$，我们有如下上界。

$$c(\delta)^{10/\delta} = ((1/\delta)^{O(\log(1/\delta))})^{O(1/\delta)} = (1/\delta)^{O((1/\delta)\log(1/\delta))}.$$

由于每个 $W_i$ 有 $n/(1/\delta)^{O(\log(1/\delta))}$ 比特，因此只需拥有

$$n/(1/\delta)^{O(\log(1/\delta))} > (1/\delta)^{O((1/\delta)\log(1/\delta))}.$$

取某个常数 $c$，使得 $\delta = c(\log\log n)^2/\log n$ 即可。 □

现在我们讨论提取器的情形。

定理 57. 存在一个常数 $c > 1$ 和一个高效函数族 $\mathsf{DAExt} : \{0,1\}^n \to \{0,1\}^m$，使得 $m = n^{\Omega(1)}$，并且对于每个熵为 $cn(\log\log\log n)^2/\log\log n$ 的仿射源 $X$，

$$(\mathsf{DAExt}(X), \mathsf{DAExt}(X+a)) \approx_\varepsilon (U_m, \mathsf{DAExt}(X+a)) ,$$

其中 $\varepsilon = 2^{-n^{\Omega(1)}}$。

证明思路。我们继续讨论次线性熵弥散器。假设熵的设置使我们确实得到了一个弥散器。注意，弥散器输出

$$n/\left((\log(1/\delta))^{O(1/\delta)} \cdot (1/\delta)^{O((1/\delta)\log(1/\delta))}\right) = n/(1/\delta)^{O((1/\delta)\log(1/\delta))}$$

比特。从这里开始，只需再考虑一个约束，即多项式的次数。对于提取器，我们需要保证定理 17 中的异或引理能给出次指数误差。换句话说，我们需要保证

$$\frac{n/(1/\delta)^{O((1/\delta)\log(1/\delta))}}{(1/\delta)^{O((1/\delta)\log(1/\delta))} \cdot 2^{(1/\delta)^{O((1/\delta)\log(1/\delta))}}} = n^{\Omega(1)}.$$

取某个常数 $c$，使得 $\delta = c\,(\log\log\log n)^2/\log\log n$ 即可。 □

# 6 读一次分支程序的平均情形 $AC^0$ 困难性

在本节中，我们构造一个可由 $AC^0$ 计算的提取器，它能够从规模适当的任意一次读取分支程序的任意输出的原像中提取随机性。

我们使用以下两个来自先前工作的 $AC^0$ 提取器构造。

定理 58（[CL18]）。对于任意常数 $c \in \mathbb{N}$、$\delta \in (0,1]$, 存在一个显式确定性（$k = \delta n$、$\varepsilon = 2^{-\log^c n}$）-提取器 $AC^0$-BFExt：$\{0,1\}^n \to \{0,1\}^{\Omega(k)}$，对于任意（$n$、$k$）-比特固定源，该提取器可由深度为 $O(c)$ 的 $AC^0$电路计算。

定理 59 ([PWY16])。对于任意常数 $c \in \mathbb{N}$、$\delta \in (0,1]$, 存在一个显式强线性带种子 ($k = \delta n$、$\varepsilon = 2^{-\log^c n}$)-提取器 $AC^0$-LExt：$\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\Omega(k)}$ 该提取器可由 $AC^0$ 深度为 $O(c)$的电路计算，种子长度为 $d = O(\log^{c+1} n)$。

## 6.1 $AC^0$-可计算的 $t$-仿射相关性破坏器

我们构造的 $AC^0$-可计算的 $t$-仿射相关性破坏器基于 $t$- 仿射相关性破坏器的框架，该破坏器见 [CL22], ，而后者又应用了 [Li17] 中的标准相关性破坏器。为此，我们首先给出一个 $AC^0$-可计算的翻转-翻转器，作为标准相关性破坏器中的子程序。随后，我们用 $t$-仿射相关性破坏器中的强种子提取器以及标准相关性破坏器中的强种子提取器替换为 $AC^0$-LExt。

---

算法 9 $AC^0$-翻转器$(x, y, b)$

---

输入：长度分别为 $n_1$、$n_1$ 的均匀比特串 $x$、$y$，一个比特 $b$ 和一个电路深度参数 $c \in \mathbb{N}$。

输出：长度为 $n_2$ 的比特串 $\hat{x}$。
参数与子程序：令 $n_2 = \Omega(n_1) \le n_1/20$ 且 $d = \Omega(n_1) \le n_2/10$。令 $AC^0$-$LExt_1$：$\{0,1\}^{n_1} \times \{0,1\}^d \to \{0,1\}^d$ 为来自定理 59 的 $(k_1 = n_1/10, \varepsilon_1 = 2^{-\log^c n})$-强线性种子提取器，$AC^0$-$LExt_2$：$\{0,1\}^{n_2} \times \{0,1\}^d \to \{0,1\}^d$ 为来自定理 59 的 $(k_2 = n_2, \varepsilon_2 = 2^{-\log^c n})$-强线性种子提取器。令 $AC^0$-$LExt_3$：$\{0,1\}^{n_1} \times \{0,1\}^d \to \{0,1\}^{n_2}$ 为来自定理 59 的 $(k_3 = n_1/2, \varepsilon_3 = 2^{-\log^c n})$-强线性种子提取器。

设 laExt：$\{0,1\}^{n_1} \times \{0,1\}^{n_2+d} \to \{0,1\}^{2d}$ 是一个前瞻提取器，用于执行 2 轮交替提取协议，其中使用 $AC^0$-$LExt_1$,$AC^0$-$LExt_2$ 作为带种子提取器。

---

1. 设 $\tilde{y} = \text{Slice}(y, n_2)$, $s_0 = \text{Slice}(\tilde{y}, d)$, $\text{laExt}(x, (\tilde{y}, s_0)) = r_0, r_1$
2. 设 $\overline{y} = AC^0\text{-LExt}_3(y, r_b)$ 3. 设 $\overline{s_0} = \text{Slice}(\overline{y}, d)$, $\text{laExt}(x, (\overline{y}, \overline{s_0}))$ $= \overline{r_0}, \overline{r_1}$4. 设 $\hat{y} = AC^0\text{-LExt}_3(y, \overline{r_{1-b}})$ 5. 设 $y_0 = \text{Slice}(\hat{y}, d)$ 6. 输出 $\hat{x} = AC^0\text{-LExt}_3(x, y_0)$

---

定理 60（$AC^0$-翻转器）。对于任意整数 $c, n_1 > 0$ 以及任意 $\varepsilon > 0$，存在一个显式函数 $AC^0$-翻转器：$\{0,1\}^{n_1} \times \{0,1\}^{n_1} \times \{0,1\} \to \{0,1\}^m$，满足以下条件：设 $X$ 是 $n_1$ 比特上的独立均匀源，$X'$ 是一个定义在 $n_1$ 比特上、与 $X$ 任意相关的随机变量。设 $Y$ 是 $n_1$ 比特上的独立均匀源，$Y$ 是一个定义在 $n_1$ 比特上、与 $Y$ 任意相关的随机变量。假设 $(X, X')$ 与 $(Y, Y')$ 独立。如果 $k = \Omega(n_1)$，那么

$\mathsf{AC}^0$-翻转器可以由深度为 $O(c)$ 的 $\mathsf{AC}^0$ 电路计算，并且对于任意比特 $b$，都有

$$\mathsf{AC}^0\text{-flip-flop}(X,Y,b) \approx_\varepsilon U_m \mid (Y,Y').$$

此外，对于任意满足 $b \neq b'$ 的比特 $b$、$b'$，我们有

$$\mathsf{AC}^0\text{-flip-flop}(X,Y,b) \approx_\varepsilon U_m \mid (\mathsf{AC}^0\text{-flip-flop}(X',Y',b'),Y,Y').$$

其中 $m \geq \Omega(k)$，且 $\varepsilon = 6 \cdot 2^{-\log^c n_1}$。

证明。 我们证明算法 9 构造了此类函数。

1. 由于 $S_0 = U_d$ 和 $\widetilde{H}_\infty(X) \geq k_1$，根据 $\mathsf{AC}^0$-$\mathsf{LExt}_1$，在固定 $\sim \sim S_0$ 条件下的性质，$R_0 \approx_{\varepsilon_1} U_d$ 是 $X$ 的线性函数，因而也是 $(Y,Y')$ 的线性函数。由于 $H_\infty(Y \mid S_0, S_0') \geq n_2 - 2d \geq k_2$ 且 $R_0 \approx_{\varepsilon_1} U_d$，根据 $\mathsf{AC}^0$-$\mathsf{LExt}_2$，在固定 $\widetilde{R_0\ S_1} \approx_{\varepsilon_1+\varepsilon_2} U_d$ 条件下的性质，$Y'$ 是一个线性函数，因而独立于 $X$。由于 $H_\infty(X \mid R_0, R_0') \geq n_1 k - 2d \geq k_1$ 且 $S_1 \approx_{\varepsilon_1+\varepsilon_2} U_d$，根据 $\mathsf{AC}^0$-$\mathsf{LExt}_1$，在 $S_1$ 条件下，$R_1 \approx_{\varepsilon_1+\varepsilon_2+\varepsilon_1} U_d$ 是 $X$ 的线性函数，因而独立于 $(Y,Y')$。

2. 由于 $R_b \approx_{\varepsilon_1+b\ (\varepsilon_1+\varepsilon_2)} U_d$ 且 $R_b\,(R_b')$ 在条件为定值 $\widetilde{\{S_0,S_1\}}(\{S_0',S_1'\})$ 时独立于 $Y(Y')$。固定 $(R_b,R_b')$ 和 $\overline{Y'}$、$H_\infty(Y \mid S_0,S_0',S_1,S_1',\overline{Y'}) \geq n_1 - 4d - n_2 \geq k_3$，则根据 $\mathsf{AC}^0$-$\mathsf{LExt}_3$ 的性质，$\overline{Y} \approx_{\varepsilon_3+\varepsilon_1+b\ (\varepsilon_1+\varepsilon_2)} U_{n_2}$。

3. 既然 $\overline{Y'}$ 已固定，我们可以固定 $(\overline{R_0'},\overline{S_1'},\overline{R_1'})$。这至多使 $X$ 损失 $2d$ 的熵。$\widetilde{\text{由于}}\ \overline{S_0}$ 是 $\overline{Y}$ 的一个切片，因此 $\overline{S_0} \approx_{\varepsilon_3+\varepsilon_1+b(\varepsilon_1+\varepsilon_2)} U_d$。此外，由于 $H_\infty(X \mid R_0,R_0',R_1,R_1',\overline{R_0'},\overline{R_1'}) \geq n_1 - 6d \geq k_1$，根据在固定 $\overline{S_0}$、$\overline{R_0} \approx_{\varepsilon_3+2\varepsilon_1+b(\varepsilon_1+\varepsilon_2)} U_d$ 的条件下 $\mathsf{AC}^0$-$\mathsf{LExt}_1$，的性质，$U_d$ 是 $X$ 的线性函数，因此独立于 $(Y,Y')$。由于 $H_\infty(\overline{Y} \mid \overline{Y'},S_0,S_0',S_1,S_1',\overline{S_0},\overline{S_0'}) \geq n_2 - 6d \geq k_2$ 且 $\overline{R_0} \approx_{\varepsilon_3+2\varepsilon_1+b(\varepsilon_1+\varepsilon_2)} U_d$，根据在固定 $\overline{R_0}$、$\overline{S_1} \approx_{\varepsilon_3+(2+b)\varepsilon_1+(b+1)\varepsilon_2} U_d$ 的条件下 $\mathsf{AC}^0$-$\mathsf{LExt}_2$，的性质，$_{\varepsilon_2}\ U_d$ 是 $\overline{Y}$ 的线性函数，因此独立于 $X$。$\widetilde{\text{由于}}\ H_\infty(X \mid R_0,R_0',R_1,R_1',\overline{R_0},\overline{R_0'},\overline{R_1'}) \geq n_1 - 7d \geq k_1$ 且 $\overline{S_1} \approx_{\varepsilon_3+(2+b)\varepsilon_1+(b+1)\varepsilon_2} U_d$，根据在固定 $\overline{S_1}$、$\overline{R_1} \approx_{\varepsilon_3+(3+b)\varepsilon_1+(b+1)\varepsilon_2} U_d$ 的条件下 $\mathsf{AC}^0$-$\mathsf{LExt}_1$，的性质，$_{\varepsilon_2}\ U_d$ 是 $X$ 的线性函数，因此独立于 $(Y,Y')$。

4. 根据上述分析，对所有 $b' \in \{0,1\}$ $\overline{R_{1-b'}} \approx_{\varepsilon_3+3\varepsilon_1+\varepsilon_2} U_d \mid \overline{R_{1-b}}\ Y\ Y'$ $\overline{R_{1-b}} \approx_{\varepsilon_3+3\varepsilon_1+\varepsilon_2} U_d \mid \overline{R_{1-b}}\ \overline{R_{1-b}'}\ Y\ Y'$，(，) 和 (，，)。因此，在固定 (，) 的条件下，$\hat{}\,(\hat{}\,)$ 是 $Y\,(Y')$ 的线性函数，因此与 $X\,(X')$ 独立。根据引理 15，$\hat{}\ Y \approx_{2\varepsilon_3+3\varepsilon_1+\varepsilon_2} U_{n_2} \mid (\hat{}\,Y',\ \overline{R_{1-b}},\ \overline{R_{1-b}'})$。

5. 现在很容易看出 $Y_0 \approx_{2\varepsilon_3+3\varepsilon_1+\varepsilon_2} U_{n_2} \mid (Y_0,\overline{R_{1-b}},\overline{R_{1-b}'})$，且 $Y_0\,(Y_0')$ 与 $\sim X\,(X')$ 独立。我们还有 $H_\infty(X \mid R_0,R_0',R_1,R_1',\overline{R_0},\overline{R_1},\overline{R_0'},\overline{R_1'}) \geq n_1 - 8d \geq k_3$。根据引理 15，有 $\hat{}\ X \approx_{3\varepsilon_3+3\varepsilon_1+\varepsilon_2} U_{n_1} \mid (\hat{}\,X',Y,Y')$。

这就完成了定理60的证明。 □

$\mathsf{AC}^0$-可计算的标准相关性破坏器。下面的算法是对 [Li17]中的相关性破坏器所作的修改，使其能够由 $\mathsf{AC}^0$ 电路计算。

定义28（$\mathsf{AC}^0$-CB）。函数 $\mathsf{AC}^0$-CB：$\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$ 是一个熵为 $k$、误差为 $\varepsilon$ 的相关性破坏器，可由深度为 $c$ 的 $\mathsf{AC}^0$ 电路计算（简称为（$k$，$\varepsilon$，$c$）-仿射相关性破坏器），当且仅当对每个 $X$、$X' \in \{0,1\}^n$、$Y$、$Y' \in \{0,1\}^d$、$\alpha$、$\alpha' \in \{0,1\}^a$，满足以下条件：

- $X$ 是一个（$n$，$k$）源，且 $Y$ 是均匀的
- $(X, X')$ 与 $(Y, Y')$ 独立
- $\alpha \neq \alpha'$

$\mathsf{AC}^0$-CB 可由深度为 $O(c)$ 的 $\mathsf{AC}^0$ 电路计算，并且

$$\mathsf{AC}^0\text{-CB}(X, Y, \alpha) \approx_\varepsilon U_m \mid \mathsf{AC}^0\text{-AffCB}(X', Y', \alpha').$$

如果 $\mathsf{AC}^0$-CB 满足以下条件，则称其为强的

$$\mathsf{AC}^0\text{-CB}(X, Y, \alpha) \approx_\varepsilon U_m \mid (\mathsf{AC}^0\text{-AffCB}(X', Y', \alpha'), Y', Y).$$

---

算法 10 $\mathsf{AC}^0$-CB$(x, y, id)$

---

输入：长度分别为 $n$、$d$、$a$ 的比特串 $x$、$y$、$id$。
输出：长度为 $m$ 的比特串 $\hat{v}$。
子程序和参数：
固定一个常数 $c$。令 $\ell = \log(a)$，$s = d/(1000(\ell+1))$，$r = s/a$, $m = \Omega(d)$。
令 $\mathsf{AC}^0$-LExt$_0$：$\{0,1\}^n \times \{0,1\}^{s_0} \to \{0,1\}^{d_0}$ 为 $\mathsf{AC}^0$-可计算的强种子提取器
来自 59 个集合，用于从一个（$n$，$d$）源中提取，其中 $d = \Omega(n)$，种子长度为 $s_0 = O(\log^{c+1} n)$，输出 $d_0 = \Omega(d) \leq 0.3d$, $d_0 \geq 200\ell s$ 且误差为 $\varepsilon_n = 2^{-\log^c n}$。
令 内积：$\{0,1\}^{d_0} \times \{0,1\}^{d_0} \to \{0,1\}^{d_0/6}$ 为定理 38 中误差为 $\varepsilon_{\mathsf{IP}} = 2^{-\Omega(d)}$。
令 $\mathsf{AC}^0$-laExt$^{2\ell+1}$：$\{0,1\}^d \times \{0,1\}^{d_0/6} \to \left(\{0,1\}^{3s}\right)^{2\ell+1}$ 为引理14 中的前瞻提取器，其 Quentin 和 Wendy 对应的提取器如下：

- $\mathsf{AC}^0$-LExt$_q$：$\{0,1\}^{d_0/6} \times \{0,1\}^{3s} \to \{0,1\}^{3s}$ 为定理 59 中用于提取 $(d_0/6, d_0/12)$-源的 $\mathsf{AC}^0$-可计算强种子提取器，误差为 $\varepsilon_d = 2^{\neg \log^c d_0/6}$。
- $\mathsf{AC}^0$-LExt$_w$：$\{0,1\}^d \times \{0,1\}^{3s} \to \{0,1\}^{3s}$ 是定理59 中的 $\mathsf{AC}^0$ 可计算强种子提取器，用于从 $(d, d/4)$ 源中提取，误差为 $\leq \varepsilon_d = 2^{\neg \log^c d_0/6}$。

令 $\mathsf{AC}^0$-laExt$^{\ell+1}$：$\{0,1\}^n \times \{0,1\}^{d_0/6} \to \left(\{0,1\}^{3s}\right)^{\ell+1}$ 为引理 14中的前瞻提取器。

- $\mathsf{AC}^0$-LExt$_q$：$\{0,1\}^{d_0/6} \times \{0,1\}^{3s} \to \{0,1\}^{3s}$ 是定理59 中的 $\mathsf{AC}^0$ 可计算强种子提取器，用于从 $(d_0/6, d_0/12)$ 源中提取，误差为 $\varepsilon_d = 2^{\neg \log^c d_0/6}$。
- $\mathsf{AC}^0$-LExt$'_w$：$\{0,1\}^n \times \{0,1\}^{3s} \to \{0,1\}^{3s}$ 设其为定理 59 中用于从 $(n, d/4)$ 源中提取、误差为 $\varepsilon_n = 2^{\neg \log^c n}$ 的 $\mathsf{AC}^0$-可计算强种子提取器。

令 $\mathsf{AC}^0$-翻转器：$\{0,1\}^{3s} \times \{0,1\}^{3s} \times \{0,1\}^a \to \{0,1\}^r$ 为定理 60 中误差为 $6 \cdot \varepsilon_n$ 的 $\mathsf{AC}^0$-可计算翻转器。令 $\mathsf{AC}^0$-LExt：$\{0,1\}^{3s} \times \{0,1\}^r \to \{0,1\}^r$ 为定理 59 中设定为从均匀源提取、误差为 $\varepsilon_s = 2^{-\log^c(3s)}$ 的 $\mathsf{AC}^0$-可计算强种子提取器。令 $\mathsf{AC}^0$-LExt′：$\{0,1\}^{3s} \times \{0,1\}^r \to \{0,1\}^r$ 为定理 59 中设定为从均匀源提取、误差为 $\varepsilon_s$ 的 $\mathsf{AC}^0$-可计算强种子提取器。令 $\mathsf{AC}^0$-LExt″：$\{0,1\}^d \times \{0,1\}^r \to \{0,1\}^s$ 为定理 59 中设定为从 $(d, d/4)$ 源提取、误差为 $\varepsilon_d$ 的 $\mathsf{AC}^0$-可计算强种子提取器。令 $\mathsf{AC}^0$-LExt‴：$\{0,1\}^n \times \{0,1\}^s \to \{0,1\}^m$ 为定理 59 中设定为从 $(n, d/4)$ 源提取、误差为 $\varepsilon_n$ 的 $\mathsf{AC}^0$-可计算强种子提取器。

---

令 $\mathsf{AC}^0$-非交互式概率合并器$_2$ 构造为二交替提取 $\{0,1\}^r \times \{0,1\}^r \times \{0,1\}^{3s} \to \{0,1\}^r$，来自定义 18，并具有以下 Quentin 和 Wendy 的提取器：

- $\mathsf{AC}^0$-$\mathrm{LExt}'_q$：$\{0,1\}^r \times \{0,1\}^r \to \{0,1\}^{r/2}$ 表示定理59 中的 $\mathsf{AC}^0$-可计算强种子提取器，用于从 $(r,r)$ 源中提取，误差为 $\varepsilon_r = 2^{-\log^c r}$。
- $\mathsf{AC}^0$-$\mathrm{LExt}''_w$：$\{0,1\}^{3s} \times \{0,1\}^{r/2} \to \{0,1\}^r$ 表示 $\mathsf{AC}^0$-可计算的强种子提取器，来自定理 59，用于从 $3s$，$s$) 源中提取，误差为 $\varepsilon_s$。

---

1. 令 $y_0 \circ y_1 = \mathrm{Slice}(y, s_0 + 0.3d)$，其中 $y_0$ 长度为 $s_0$，且 $x_1 = \mathsf{AC}^0\text{-}\mathrm{LExt}_0(x, y_0)$。
2. 计算 $z =$ 内积$(x_1, y_1)$。
3. 令 $r_0$、$r_1$、$\cdots$、$r_{2\ell} = \mathsf{AC}^0\text{-}\mathrm{laExt}^{2\ell+1}(y, z)$。
4. 令 $s_0$、$s_1$、$\cdots$、$s_\ell =$ 为 $\mathsf{AC}^0\text{-}\mathrm{laExt}^{\ell+1}(x, z)$。
5. 令 $V^0$ 为一个 $a \times r$ 矩阵，其第 $i$ 行为 $V_i^0 = \mathsf{AC}^0$-翻转器$(s_0, r_0, \alpha_i)$，并且包含 $r$ 个比特。
6. 对于 $j = 1$、$\cdots$、$\ell$，执行以下操作。将矩阵 $v^{j-1}$ 按每两行合并：注意，$v^{j-1}$ 有 $a/2^{j-1}$ 行；对于 $i = 1$、$\cdots$、$a/2^j$，计算输出 $r$ 个比特的 $\overline{v_i^{j-1}} = \mathsf{AC}^0$-非交互式概率合并器$(v_{2i-1}^{j-1}, v_{2i}^{j-1}, r_{2j-1})$，以及包含 $r$ 个比特的 $\tilde{v}_i^{j-1} = \mathsf{AC}^0\text{-}\mathrm{LExt}(r_{2j}, \overline{v_i^{j-1}})$。最后计算包含 $r$ 个比特的 $v_i^j = \mathsf{AC}^0\text{-}\mathrm{LExt}'(s_j, \tilde{v}_i^{j-1})$。
7. 计算 $\hat{v} = \mathsf{AC}^0\text{-}\mathrm{LExt}'''(x, \mathsf{AC}^0\text{-}\mathrm{LExt}''(y, v^\ell))$。

---

定理 61（$\mathsf{AC}^0$-CB）。对于每个常数 $c$，都存在一个显式的强相关性破坏器 $\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^{\Omega(d)}$，其熵为 $d$，误差为 $\varepsilon = O(a \cdot 2^{-\log^c n})$，其中 $d = \Omega(n)$ 且 $a = O(d^{\text{对数}^c d})$。此外，该相关性破坏器可由深度为 $O(c)$ 的 $\mathsf{AC}^0$ 电路计算。

证明。我们证明，算法 10 给出了这样一个相关性破坏器。我们将逐步分析该算法。

步骤 1. 固定 $(Y_0, Y_0')$；在此固定条件下，$X_1$ 是 $X$ 的线性函数，并且与 $(Y_1, Y_1')$ 及 $(Y, Y')$ 独立。

步骤 2. 由于 $X_1 \approx_{\varepsilon_n} U_{0.3d}$ 和 $Y_1 = U_{0.3d}$，根据内积的定义，有 $Z \approx_{\varepsilon_{\mathrm{IP}}} U_{d_0/6}$。步骤 3。

- 进一步固定 $(Y_1, Y_1')$，在该固定条件下，有 $(Z, Z')$ 是 $(X_1, X_1')$ 的确定性函数，因此 $(Z, Z')$ 与 $(Y, Y')$ 独立。
- $Z \approx_{\varepsilon_{\text{内积}}} U_{d_0/6}$ 和 $\widetilde{H}_\infty(Y \mid Y_0, Y_0', Y_1, Y_1') \geq d - 2s_0 - 2 \cdot d_0 \geq 0.3d$。
- 根据引理 14，由于 $H_\infty(Y \mid Y_1, Y_1') \geq 0.3d \geq d/4 + 2(2\ell+1)(3s) + 2\log(1/\varepsilon_d)$ 以及 $d_0/6 \geq d_0/12 + 2(2\ell+1)(3s) + 2\log(1/\varepsilon_d)$，我们有，对于任意 $0 \leq j \leq 2\ell - 1$，均有

$$R_{j+1} \approx_{O(\ell\varepsilon_d)} U_{3s} \mid (Z, Z', R_0, R_0', \cdots, R_j, R_j').$$

根据混合论证和三角不等式，我们有：

$$(Z, Z', R_0, R_0', \cdots, R_{2\ell}, R_{2\ell}') \approx_{O(\ell^2\varepsilon_d)} (Z, Z', U_{3s}, R_0', \cdots, U_{3s}, R_{2\ell}'). \tag{4}$$

其中，每个 $U_{3s}$ 都独立于所有先前的随机变量（但可能依赖于后续的随机变量）。

• 在固定（$Z$， $Z'$）的条件下，我们有 $\{(R_i, R_i')\}_{i\in[0,2\ell]}$ 是（$Y_1$， $Y_1'$）的确定性函数，因此独立于（$X$， $X'$）。

步骤 4。

• 固定 $(X_1, X_1')$；在此固定条件下，$(Z, Z')$ 是 $(Y_1, Y_1')$ 的确定性函数，因此 $(Z, Z')$ 与 $(X, X')$ 独立。

• $Z \approx_{\varepsilon_{\text{内积}}} U_{d_0/6}$ 和 $\widetilde{H}_\infty(X \mid X_1, X_1') \geq d - 2\cdot d_0 \geq 0.4d$。

• 根据引理 14，由于 $H_\infty\widetilde{(}X \mid X_1, X_1') \geq 0.4d \geq d/4 + 2(\ell+1)(3s) + 2\log(1/\varepsilon_n)$ 以及 $d_0/6 \geq d_0/12 + 2(\ell+1)(3s) + 2\log(1/\varepsilon_d)$，我们可得，对于任意 $0 \leq j \leq \ell-1$，均有

$$S_{j+1} \approx_{O(\ell(\varepsilon_n+\varepsilon_d)/2)} U_{3s} \mid (Z, Z', \{S_0, S_0', \cdots, S_j, S_j'\}).$$

根据混合论证和三角不等式，可得

$$(Z, Z', S_0, S_0', \cdots, S_\ell, S_\ell') \approx_{O(\ell^2(\varepsilon_n+\varepsilon_d)/2)} (Z, Z', U_{3s}, S_0', \cdots, U_{3s}, S_\ell'). \tag{5}$$

其中每个 $U_{3s}$ 都独立于之前的所有随机变量（但可能依赖于之后的随机变量）。

• 在固定（$Z$， $Z'$）的条件下，我们有 $\{(S_i, S_i')\}_{i\in[0,\ell]}$ 是（$X_1$， $X_1'$）的确定性函数，因此独立于（$Y$， $Y'$）。

因此，我们可以得出，在固定（$X_1$， $X_1'$， $Y_1$， $Y_1'$， $Z$， $Z'$）的条件下，我们有 $\{(R_i, R_i')\}_{i\in[0,2\ell]}$ 是（$Y$， $Y'$）的确定性函数，且 $\{(S_i, S_i')\}_{i\in[0,\ell]}$ 是（$X$， $X'$）的确定性函数，因此它们相互独立。此外，给定之前的随机变量，每个 $R_i$ 和 $S_i$ 都接近均匀分布。从现在开始，我们将假设每个 $R_i$ 和 $S_i$ 都是均匀的（*），并在最后加回 $O(\ell^2(\varepsilon_n + \varepsilon_d))$ 的误差。由于在下面的算法和分析中，每个 $R_i$ 和 $S_i$ 至多作为种子源使用两次，这已经足够了。

步骤 5。根据定理 60，对于所有 $i \in [a]$, $V_i^0 \approx_{O(\varepsilon_n)} U_s$。此外，由于 $\alpha \neq \alpha'$，存在一个 $i \in [a]$ 使得 $V_i^0 \approx_{O(\varepsilon_n)} U_s \mid (V_i'^0, R_0, R_0')$。现在，在固定 $(R_0, R_0')$ 的条件下，$(V^0, V'^0)$ 是 $(S_0, S_0')$ 的确定性函数，因此独立于 $\{(R_i, R_i')\}_{i\in[2\ell]}$。 St

**ep 6.** First note that the followings:

1. 在固定 $(R_0, R_0')$ 的条件下，$(V^0, V'^0)$ 是 $(S_0, S_0')$ 的线性函数。

2. $V^0$ 的每一行都接近均匀分布，并且 $V^0$ 中存在一行，即使在对应的 $V'^0$ 行条件下，该行仍接近均匀分布。

在下面的分析中，我们通过归纳法证明，对于任意 $j \in [0, \ell]$,

(a) $V^j$ 的每一行都接近均匀分布，并且 $V^j$ 中存在一行，在对应的 $V_j'$ 行取定值时，该行接近均匀分布。

对于任意 $j \in [\ell]$，均有

(b) 在固定 $(R_0, R_0', \cdots, R_{2j-2}, R_{2j-2}')$ 后，$(V^{j-1}, V'^{j-1})$ 是 $(S_0, S_0', \cdots, S_{j-1}, S_{j-1}')$ 的线性函数。

(c) $\overline{V^{j-1}}(\tilde{V}^{j-1})$ 的每一行都接近均匀分布，并且 $\overline{V^{j-1}}(\tilde{V}^{j-1})$ 中存在一行，即使在对应的 $\overline{V'^{j-1}}$ $(\tilde{V}'^{j-1})$ 行取定值的条件下，也接近均匀分布。

对于每次迭代 $j \in [\ell]$，步骤 6 会生成 3 个新的某处随机矩阵：$\overline{V^{j-1}}$、$\tilde{V}^{j-1}$ 和 $V^j$，它们的大小分别为 $(a/2^j) \times r$、$(a/2^j) \times r$ 和 $(a/2^j) \times r$。它们各自都具有一些性质：矩阵 $\overline{V^{j-1}}$。在固定 $(R_0, R'_0, \cdots, R_{2j-2}, R'_{2j-2})$ 的条件下，根据我们的假设 (*)，$R_{2j-1} = U_{3s}$。现在，在 $(R_0, R'_0, \cdots, R_{2j-1}, R'_{2j-1})$ 取定值的条件下，根据引理 14，$\overline{V^{j-1}}$ 的每一行都以 $O(2^{j-1}(\varepsilon_s + \varepsilon_n))$ 接近均匀分布。由于 $V^{j-1}$ 中存在一行，即使给定 $V'^{j-1}$ 中对应的行，也接近均匀分布，根据引理 15，$\overline{V^{j-1}}$ 中存在一行，即使在 $\overline{V'^{j-1}}$ 中同一行取定值的条件下，也接近均匀分布。此外，在固定 $(R_0, R'_0, \cdots, R_{2j-1}, R'_{2j-1})$ 的条件下，$(\overline{V^{j-1}}, \overline{V'^{j-1}})$ 是 $(V^{j-1}, V'^{j-1})$ 的线性函数；根据归纳假设，它是 $(S_0, S'_0, \cdots, S_{j-1}, S'_{j-1})$ 的线性函数，因此与 $R_{2j}$ 独立。

矩阵 $\tilde{V}^{j-1}$。首先注意，对于每个 $i \in [a/2^j]$，矩阵 $\tilde{V}^{j-1}$ 的第 $i$ 行是通过使用矩阵 $\overline{V^{j-1}}$ 的第 $i$ 行从 $S_j$ 中提取得到的。此外，在 $\overline{V^{j-1}}$ 为定值时，$\tilde{V}^{j-1}$ $R_{2j}$ $R_{2j} = U_{3s} \mid R_0\ R'_0 \cdots R_{2j-1}\ R'_{2j-1}$ 是一个确定性函数。由于 (,,,,) 且 $\widetilde{H}_\infty(R_{2j} \mid \overline{V^{j-1}}, \tilde{V}^{j-1}_{[u]}) \geq 3s - ur \geq 3s - ar/2^{j-1} \geq 3s - ar \geq s + \log(1/\varepsilon_s)$，其中 $u \in [a/2^j - 1]$，根据 AC$^0$-LExt 的定义，$\tilde{V}^{j-1}$ 的每一行都是均匀的。由于 $\overline{V^{j-1}}$ 中有一行在条件为对应的 $\overline{V'^{j-1}}$ 行时，与均匀分布的距离为 $O(2^{j-1}(\varepsilon_s + \varepsilon_n))$，根据引理 15，$\tilde{V}^{j-1}$ 中也有一行即使在条件为对应的 $\tilde{V}'^{j-1}$ 行时也接近均匀分布。

矩阵 $V^j$。首先注意，对于每个 $i \in [a/2^j]$，矩阵 $V^j$ 的第 $i$ 行是通过使用矩阵 $\tilde{V}^{j-1}$ 的第 $i$ 行从 $S_j$ 中提取得到的。此外，在 $\tilde{V}^{j-1}$ 为定值时，$V^j$ 是 $S_j$ 的确定性函数。由于 $S_j = U_{3s} \mid (S_0, S'_0, \cdots, S_{j-1}, S'_{j-1})$ 和 $\widetilde{H}_\infty(S_j \mid \tilde{V}^{j-1}, V^j_{[u]}) \geq 3s - ur \geq 3s - ar/2^{j-1} \geq 3s - ar \geq s + \log(1/\varepsilon_s)$，其中 $u \in [a/2^j - 1]$，根据 AC$^0$-LExt′ 的定义，$V^j$ 的每一行都与均匀分布相差 $O(2^j(\varepsilon_s + \varepsilon_n))$。由于 $\tilde{V}^{j-1}$ 中存在一行，在对应的 $\tilde{V}'^{j-1}$ 行给定时，该行接近均匀分布，因此根据引理 15，$V^j$ 中也存在一行，即使在对应的 $V'^j$ 行给定时，该行仍接近均匀分布。

设定 $j = \ell$，我们得到 $\text{ins}^{V^\ell} \approx_{O(2^\ell(\varepsilon_s+\varepsilon_n))} U_r \mid V'^\ell$。

步骤 7。注意，由于 $V^\ell \approx_{O(a(\varepsilon_s+\varepsilon_n))} U_r \mid V'^\ell$，通过两次迭代使用引理 15，可得 $H(Y \mid \{R_i, R'_i\}_{i\in[0,2\ell]}) \geq d/4 + 2\log(1/\varepsilon_d)$ 以及 $H(X \mid \{S_i, S'_i\}_{i\in[0,\ell]}) \geq d + 2\log(1/\varepsilon_n)$，即 $\hat{V} \approx_{O(\varepsilon_n+\ell^2(\varepsilon_n+\varepsilon_d)+a(\varepsilon_s+\varepsilon_n)+\varepsilon_d+\varepsilon_n)} U_m \mid \hat{V}' \iff \hat{V} \approx_{O(a\varepsilon_n)} U_m \mid \hat{V}'$。由于 $\hat{V}$ 是在 $(Y, Y', \{S_i, S'_i\}_{i\in[0,\ell]})$ 条件为定值时 $X$ 的确定性函数，并且 AC$^0$-LExt‴ 是强的，因此 $\hat{V} \approx_{O(a\varepsilon_n)} U_m \mid (\hat{V}', Y, Y')$ 也成立。定理 61 的证明完毕。 □

AC$^0$-可计算的 $t$-仿射相关性破坏器。以下定义是对 $t$-仿射相关性破坏器[CL22] 的修改，将其转换为 AC$^0$-可计算的设置。

定义 29（AC$^0$-AffCB）。函数 AC$^0$-AffCB：$\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$ 是一个针对熵 $k$ 且误差为 $\varepsilon$ 的 $t$-仿射相关性破坏器，可以由深度为 $c$ 的 AC$^0$ 电路计算（或简称为一个（$t$，$k$，$\varepsilon$，$c$）-仿射相关性破坏器），如果对于任意 $X$、$A$、$B \in \{0,1\}^n$、$Y$、$Y^{[t]} \in \{0,1\}^d$、$Z$ 以及满足以下条件的字符串 $\alpha$、$\alpha^{[t]} \in \{0,1\}^a$：

- $X = A + B$
- $\widetilde{H}_\infty(A \mid Z) \geq k$
- $(Y, Z) = (U_d, Z)$

  给定 $Z$ 时，• $A$ 与（$B$，$Y$，$Y^{[t]}$）相互独立
- $\alpha$、$\alpha^1$、$\cdots$、$\alpha^t$ 是 $a$ 比特串，使得 $\alpha \neq \alpha^i$ 对每个 $i \in [t]$

$\mathsf{AC}^0$-AffCB 可由深度为 $O(c)$ 的 $\mathsf{AC}^0$ 电路计算，并且

$$\mathsf{AC}^0\text{-}\mathsf{AffCB}(X,Y,\alpha) \approx_\varepsilon U_m \mid \left\{\mathsf{AC}^0\text{-}\mathsf{AffCB}(X^i,Y^i,\alpha^i)\right\}_{i\in[t]}.$$

我们称 $\mathsf{AC}^0$-AffCB 是强的，当且仅当

$$\mathsf{AC}^0\text{-}\mathsf{AffCB}(X,Y,\alpha) \approx_\varepsilon U_m \mid \left(\left\{\mathsf{AC}^0\text{-}\mathsf{AffCB}(X^i,Y^i,\alpha^i)\right\}_{i\in[t]}, Y^{[t]}, Y\right).$$

算法 11 下文给出一种强 $t$、 $k$、 $\varepsilon$、 $c$-仿射相关性破坏器的构造。

**算法 11 $\mathsf{AC}^0$-AffCB$(x,y,id)$**

输入：长度分别为 $n$、 $d=\Omega(n)$ $a$ 的比特串 $x=w+z$、 $y$、 $id$。输出：长度为 $r$ 的比特串 $q_{\lceil \log t\rceil}$。子程序和参数：固定一个常数 $c$。令 $d_0' = O(\log^{c+1} n)$、 $d_0 \le \min\{k,d\}/(10t+10)$、 $d_x \le d_0/(2\log t)$、 $r = k/(10+10t)$、 $d_y = {}_{4t}{}^{r}\log t$。令 $\mathsf{AC}^0$-LExt：$\{0,1\}^n \times \{0,1\}^{d_0'} \to \{0,1\}^{d_0}$ 为定理 59 中误差为 $\varepsilon_n = 2^{-\log^c n}$ 的 $\mathsf{AC}^0$ 可计算强种子提取器。令 $\mathsf{AC}^0$-CB：$\{0,1\}^d \times \{0,1\}^{d_0} \times \{0,1\}^a \to \{0,1\}^{d_x}$ 为定理 61 中误差为 $\varepsilon' = O(a\cdot 2^{-\log^c d})$ 的 $\mathsf{AC}^0$ 可计算相关性破坏器。令 $\mathsf{AC}^0$-LExt′：$\{0,1\}^n \times \{0,1\}^{d_x} \to \{0,1\}^r$ 为定理 59 中误差为 $\varepsilon_n = 2^{-\log^c n}$ 的 $\mathsf{AC}^0$ 可计算强种子提取器。令 $\mathsf{AC}^0$-LExt$_w$：$\{0,1\}^d \times \{0,1\}^{d_y} \to \{0,1\}^{d_x}$ 为定理 59 中误差为 $\varepsilon_d = 2^{-\log^c d}$ 的 $\mathsf{AC}^0$ 可计算强种子提取器。令 $\mathsf{AC}^0$-LExt$_q$：$\{0,1\}^r \times \{0,1\}^{d_x} \to \{0,1\}^{d_y}$ 为定理 59 中误差为 $\varepsilon_r = 2^{-\log^c r}$ 的 $\mathsf{AC}^0$ 可计算强种子提取器。

---

设 $y_0 =$ 切片$(y, d_0')$
Let $x_0 = \mathsf{AC}^0\text{-}\mathsf{LExt}(x, y_0)$
令 $y_1 = \mathsf{AC}^0\text{-}\mathsf{CB}(y, x_0, \alpha)$
令 $q_0 = \mathsf{AC}^0\text{-}\mathsf{LExt}'(x, y_1)$
对于每个 $i$、 $1 \le i \le \lceil \log t\rceil$，执行以下操作。

1. 让 $s_{i-1} =$ 切片$(q_{i-1}, d_y)$
2. 令 $r_{i-1} = \mathsf{AC}^0\text{-}\mathsf{LExt}_w(y, s_{i-1})$
3. 令 $\overline{s_i} = \mathsf{AC}^0\text{-}\mathsf{LExt}_q(q_{i-1}, r_{i-1})$
4. 令 $\overline{r_i} = \mathsf{AC}^0\text{-}\mathsf{LExt}_w(y, \overline{s_i})$
5. $q_i = \mathsf{AC}^0\text{-}\mathsf{LExt}'(x, \overline{r_i})$

---

定理62（$\mathsf{AC}^0$-AffCB）。对于每个 $c\in\mathbb{N}$、常数 $0<\delta<1$ 和 $n\in\mathbb{N}$ 以及每个 $k$、 $d$、 $t$、 $a$，存在一个常数 $C$，使得如果

- $k \ge \delta n$
- $d=\Omega(n)$ 且 $d \le n$
- $t = O(1)$

- $a \le C \underset{(n)}{\overset{n}{c}}$ 对数

那么，存在一个强 AC$^0$-AffCB：$\{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$，它可由深度为 $O(c)$ 的 AC$^0$ 电路计算

- $m = \Omega(k)$
- $\varepsilon = O(2^{-\log^{c-1} k})$

证明。我们将证明，算法 11 给出了这样一个函数。
首先，我们证明 AC$^0$-AffCB满足定义 29。

1. 对所有 $i \in [t]$，令 $X^i_{0,A} :=$ AC$^0$-LExt$(A, Y^i_0)$, $X^i_{0,B} :=$ AC$^0$-LExt$(B, Y^i_0)$, $Q^i_{0,A} :=$ AC$^0$-LExt′$(A, Y^i_1)$, $Q^i_{0,B} :=$ AC$^0$-LExt′$(B, Y^i_1)$。令 $Z$ 为定义 29 中的 $Z_{29}$。

2. 根据 AC$^0$-LExt 的定义，

$$X_{0,A} \approx_{\varepsilon_n} U_{d_0} \mid (Z, Y_0, Y_0^{[t]}, X_{0,B}, X_{0,B}^{[t]}).$$

3. 由于 $\widetilde{H}_\infty(Y \mid Z, Y_0, Y_0^{[t]}, X_{0,B}, X_0^{[t],B}) \ge d - (t+1)d'_0 \ge 9d/10$, $R_{1,A}, R_1^{[t],A}$ 在给定 $Y, Y^{[t]}$ 的条件下，与 $Z, Y_0, Y_0^{[t]}, X_{0,B}, X_0^{[t],B}$ 独立，且 AC$^0$-CB 是一个强相关性破坏器，因此有 $\forall i \in [t]$：

$$Y_1 \approx_{\varepsilon_n + \varepsilon'} U_{d_x} \mid (Y_1^i, Z, Y_0, Y_0^{[t]}, X_{0,B}, X_{0,B}^{[t]}, X_0, X_0^i).$$

4. 由于在固定 $X_0$、$X_0^{[t]}{}_B$、$Y_1$ 后，$Y_1$ 是 $Y$ 的确定性函数，且与 $X_0^{[t]}$ 独立，

$$Y_1 \approx_{\varepsilon_n + \varepsilon'} U_{d_x} \mid (Y_1^i, Z, Y_0, Y_0^{[t]}, X_{0,B}, X_{0,B}^{[t]}, X_0, X_0^{[t]}).$$

5. 根据引理 15，成立 $\forall i \in [t]$：

$$Q_{0,A} \approx_{2\varepsilon_n + \varepsilon'} U_r \mid (Q^i_{0,A}, Z, Y_0, Y_0^{[t]}, X_{0,B}, X_{0,B}^{[t]}, X_0, X_0^{[t]}, Y_1, Y_1^{[t]}).$$

由于 $(Q_{0,B}, Q_0^{[t],B})$ 与 $Q_{0,A}$ 相互独立，令

$$Z_0 := (Z, Y_0, Y_0^{[t]}, X_{0,B}, X_{0,B}^{[t]}, X_0, X_0^{[t]}, Y_1, Y_1^{[t]}, Q_{0,B}, Q_{0,B}^{[t]}),$$

同样成立：

$$Q_{0,A} \approx_{2\varepsilon_n + \varepsilon'} U_r \mid (Q^i_{0,A}, Z_0),$$

这等价于

$$Q_0 \approx_{2\varepsilon_n + \varepsilon'} U_r \mid (Q^i_0, Z_0).$$

断言 63. $\overline{S_i}$、$\overline{R_i}$、$Q_i$、$R_i$ 中的每一个都接近均匀分布，且与任意 min $\{2^i, t\}$ 个篡改随机变量独立。

证明。对于每个 $i \in [\lceil \log t \rceil]$，令

$$Z_{i,1,B} := (Z_{i-1}, S_{i-1,B}, S^{[t]}_{i-1,B});\ Z_{i,2} := (Z_{i,1,B}, S_{i-1}, S^{[t]}_{i-1});\ Z_{i,3} := (Z_{i,2}, R_{i-1}, R^{[t]}_{i-1})$$

$$Z_{i,3,B} := (Z_{i,3}, \overline{S_{i,B}}, \overline{S_{i,B}}^{[t]});\ Z_{i,4} := (Z_{i,3}, \overline{S_i}, \overline{S_i}^{[t]});\ Z_i := (Z_{i,4}, \overline{R_i}, \overline{R_i}^{[t]}),$$

令 $T_i$ 为 $[t]$ 中大小为 $2^i$ 的任意子集（若 $2^i \le t$），否则令其为 $[t]$。现在我们根据要通过归纳法证明的顺序 $\mathcal{C}$ 为各项声明定义一个排列。第一个声明是带有 $i=0$ 的子步骤 5。接下来各项声明依次为轮次 $i$、子步骤 1；轮次 $i$、子步骤 2，……，轮次 $i$、子步骤 5；轮次 $i+1$、子步骤 1；轮次 $i+1$、子步骤 2，……，轮次 $\lceil \log t \rceil$、子步骤 5。首先注意，根据上述论证，下面子步骤 5 中的声明对 $i=0$ 成立。显然，$\mathcal{C}$ 中关于阶数 $\le k$ 的声明蕴含关于阶数 $k+1$ 的声明。

子步骤 1：$S_{i-1} \approx_{2\varepsilon_n+\varepsilon'+(i-1)(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_{d_y} \mid (S_{i-1}^{T_{i-1}}, Z_{i-1})$；$S_{i-1,\ A} \approx_{2\varepsilon_n+\varepsilon'} U_{d_y} \mid (S_{i-1}^{T_{i-1}},\ _A Z_{i,\ 1,\ B})$，前提是子步骤 5 中的陈述对 $i-1$ 成立。

子步骤 2：根据引理15，$R_{i-1} \approx_{2\varepsilon_n+\varepsilon'+\varepsilon_d+(i-1)(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_{d_x} \mid (R_{i-1}^{T_{i-1}}, Z_{i,\ 2})$ 成立，前提是子步骤 1 中的陈述成立，且 $H_\infty(Y \mid Z_{i,\ 2}) \ge 9d/10 - 2(i-1)(t+1)d_x \ge d/2$。

子步骤 3：根据引理 15，$\overline{S_i} \approx_{2\varepsilon_n+\varepsilon'+\varepsilon_d+\varepsilon_r+(i-1)(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_{d_x} \mid (\overline{S_i}^{T_i}, Z_{i,\ 3})$；$S_{i,\ A} \approx_{2\varepsilon_n+\varepsilon'+\varepsilon_d+\varepsilon_r+(i-1)(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_{d_x} \mid (\widetilde{S_{i,\ A}}^{T_i}, Z_{i,\ 3,\ B})$ 成立，前提是子步骤 2 中的陈述成立，且 $H_\infty(Q_{i-1} \mid Z_{i,\ 3}) \ge r - (2i-1)(t+1)d_y \ge r/2$。

子步骤 4：根据引理 15，只要子步骤 3 中的陈述成立，就有 $\overline{R_i} \approx_{2\varepsilon_n+\varepsilon'+2\varepsilon_d+\varepsilon_r+(i-1)(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_{d_y} \mid (\overline{R_i}^{T_i}, Z_{i,4})$，并且 $H_\infty(Y \mid Z_{i,4}) \ge 9d/10 - (2i-1)(t+1)d_x \ge d/2$。

子步骤 5：根据引理 15，只要子步骤 4 中的陈述成立，就有 $Q_i \approx_{2\varepsilon_n+\varepsilon'+i(2\varepsilon_d+\varepsilon_n+\varepsilon_r)} U_r \mid (Q_i^{T_i}, Z_i)$，并且 $H_\infty(X \mid Z_i) \ge k - (d_0+r)(t+1) - 2(i-1)(t+1)d_x \ge k/2$。

□

现在注意，在条件为定值 $Z_{\lceil \log t \rceil}$（其中包含 $(\overline{R_{\lceil \log t \rceil}}, \overline{R_{\lceil \log t \rceil}}^{[t]})$）时，$Q_{\lceil \log t \rceil} \approx_{O(\varepsilon'+(\log t)\varepsilon_n)} U_r \mid Q_{\lceil \log t \rceil}^{[t]}$。此外，$Q_{\lceil \log t \rceil}$、$Q_{\lceil \log t \rceil}^{[t]}$ 是 $X$ 的确定性函数，并且与 $Y$、$Y^{[t]}$ 独立。因此，我们有

$$Q_{\lceil \log t \rceil} \approx_{O((a+\log t)\cdot 2^{-\log^c n})} U_r \mid (Q_{\lceil \log t \rceil}^{[t]}, Y, Y^{[t]}).$$

这就完成了定理的证明 62。 □

## 6.2 $\mathsf{AC}^0$-C 用于只读分支程序源的可计算提取器 s

**算法12 $\mathsf{AC}^0$-Ext($x$)**

输入：$x$ —一个 $n$ 比特串。输出：$z$ —一个 $m$ 比特串，含有 $m = \Omega(n)$。

子程序和参数：
令 $\mathsf{AC}^0$-BFExt：$\{0,1\}^n \to \{0,1\}^{n_1}$ 为定理 58 中的线性有种子提取器，设置为从最小熵 $k_1 = \delta n$ 中提取，误差为 $\varepsilon_1 = 2^{-\log^c}(n/t)$。令 $\mathsf{AC}^0$-AffCB：$\{0,1\}^n \times \{0,1\}^{n_1} \times \{0,1\}^a \to \{0,1\}^m$，$a = \log(t)$，为定理 $t$中的仿射相关性破坏器，误差为 $\varepsilon_2 = O(2^{-\log^{c-1} n})$。

1. 将 $x$ 划分为 $t = 2/\delta$ 块，使得 $x = x_1 \circ \cdots \circ x_t$。2. 令 $y_1 \circ \cdots \circ y_t = \mathsf{AC}^0$-BFExt($x_1$) $\circ \cdots \circ \mathsf{AC}^0$-BFExt($x_t$)，使得每个 $y_i$ 的长度为 $n_1 < \delta^2/100n$ 比特。3. 令 $s$ 为一个矩阵，其第 $i$ 行为 $s_i$，即 $\mathsf{AC}^0$-AffCB($x, y_i, i$)。4. 输出 $z = \bigoplus_{j=1}^t s_i$。

定理 64. 对于任意常数 $0 < \delta \leq 1$，存在一个函数族 $\mathsf{AC}^0$-Ext：$\{0,1\}^n \to \{0,1\}^m$，可在 $\mathsf{AC}^0$，中计算，使得对于任意源 $X = A + B$，其中 $A$ 和 $B$ 相互独立且具有不相交的张成空间，$A$ 具有 $\delta n$ 熵，并且 $B$ 是一个熵为 $\delta n$的近似比特固定源，当 $m =$ 为 $\Omega(n)$ 且 $\varepsilon = O(2^{-\log^{c-1} n})$ 时，$\mathsf{AC}^0$-Ext$(X) \approx_\varepsilon U_m \mid B$。

证明。

引理 65。存在 $g \in [t]$，使得在将 $\{B_1, \cdots, B_{g-1}, A_g, B_{g+1}, \cdots, B_t\}$ 固定后，以下结论成立。

- $X_g$ 是一个熵率为 $\delta$、且 $Y_g \approx_{2^{-\log^c(n/t)}} U_{n_1}$ 的几乎比特固定源。
- $Y_g$ 是 $B$ 的一个确定性函数。
- $\{Y_1, \cdots, Y_{g-1}, Y_{g+1}, \cdots, Y_t\}$ 是 $A$ 的确定性函数。

证明。由于 $X = A + B$，因此对所有 $i \in [t]$，都有 $X_i = A_i + B_i$。由于 $H_\infty(B) \geq \delta n$，且每个 $B_i$ 的块长度为 $n/t$，存在 $g \in [t]$ 使得 $H_\infty(B_g) \geq \delta n/t = \delta^2 n/2$。现在根据 $\mathsf{AC}^0$-BFExt 的提取性质，我们有 $Y_g \approx_{\varepsilon_1} U_{n_1}$。 □

引理 66。在额外固定 $Y = \{Y_i\}_{i \in [t] \setminus \{g\}}$ 的条件下，$H_\infty(A) \geq \delta n/4$。

证明。由于 $X_g$ 的熵至多为 $\delta n/2$，$H_\infty(B_g) \geq \delta^2 n/2$。现在，由于 $A_g$ 与 $B_g$ 独立，$\sim H_\infty(A) \leq (1-\delta)\delta n/2$。由于 $|Y| \leq t n_1 \leq \delta n/50$，我们有 $H_\infty(A \mid A_g, Y) \geq \delta n - (1-\delta)\delta n/2 - \delta n/50 \geq \delta n/4$。 □

引理 67. 对 $(A_g, \{S_i, Y_i\}_{i \in [t]}, B)$ 进行固定时，以 $1 - \varepsilon_2$ 的概率，$Z \approx_{2\varepsilon_1 + \varepsilon_2} U_m$。

证明. 令 $Z_{29} = \{Y_i, B_i\}_{i \in [t] \setminus \{g\}} \cup \{A_g\}$。根据引理 65，以 $1 - 2^{-\log^c(n/t)}$ 的概率，$Y$ 是某处随机源。此外，由于 $A$ 和 $B$ 相互独立，我们有 $Y_g = U_{n_1} \mid Z$。根据引理 66，$\widetilde{H}_\infty(A \mid Z_{29}) \geq 4\delta/n$。根据定理62，$S_g \approx_{\varepsilon_1 + \varepsilon_2} U_m \mid (\{S_i\}_{i \in [t] \setminus \{g\}}, Y^{[t]})$。由于 $Y_g$ 是 $B_g$ 的确定性函数，且在 $Y_g$ 和 $Z_{29}$ 在条件为定值时，$S_g$ 是 $A$ 的确定性函数，因此有 $S_g \approx_{\varepsilon_1 + \varepsilon_2 + \varepsilon_1} U_m \mid (\{S_i, Y_i\}_{i \in [t] \setminus \{g\}}, B)$，这意味着 $Z \approx_{2\varepsilon_1 + \varepsilon_2} U_m \mid B$。 □

□

定理 68. 对于任意常数 $\delta > 0$，令 $\mathsf{AC}^0$-Ext 为定理 *64* 中的一个函数，对于 $\delta_{64} = \delta/3$其误差为 $\varepsilon = 2^{-\Omega(\log^{c-1} n)}$，则

$$\mathsf{ROBP}_{2\varepsilon}(\mathsf{AC}^0\text{-Ext}) > 2^{(1-\delta)n}.$$

我们分两步证明上述定理。首先，我们回顾 [CL23] 中的一个引理，并证明存在两个源的和 $X = A + B$，满足以下 3 个性质：(1) $A$ 和 $B$ 支撑在输入比特的不相交子集上；(2) $A$ 的最小熵为 $(1-\delta)n - \log s$，且 $B$ 的最小熵至少为 $\delta n$；以及 (3) $B$ 是一个无关比特固定源。然后，我们证明在以 ROBP的输出为条件时，我们的提取器输出接近均匀。

引理 69 （ [CL23]中的引理 3.1 的一个特例）。令 $X$ 为在$\mathsf{F}_2^n$上均匀分布的随机变量。对于每个只读一次分支程序 $f$：$\mathsf{F}_2^n \to \{0,1\}$，其大小为 $s$，以及每个 $d \in [n]$，存在一个随机变量 $E$，以及随机变量 $A$，$B \in \mathsf{F}_2^n$ 使得

- $E$ 的支撑大小至多为 $2s$。

- $X = A + B$。

对于每个 $e \in \mathrm{Supp}(E)$，定义 $A_e = A\,|_{E=e}$, $B_e = B\,|_{E=e}$，则有

– $A^e$ 和 $B^e$ 相互独立。 – $B^e$ 在维数为 $d$的坐标子集 $V_e^B$上均匀分布。
– 存在 $V_e^B$ 的一个补子空间 $V_e^A$，使得 $A_e \in V_e^A$。

存在一个确定性函数 $g$，使得 $g(E, B) = f(X)$。

接下来我们证明下面的命题，由此可得 ROBP 的平均情形下界。

命题 70. 对于任意常数 $\delta > 0$，令 $\mathsf{AC}^0$-Ext 为定理 64 中的函数，具有 $\delta_{64} = \delta/3$ 输出 1 比特，误差为 $\varepsilon$，并且 $f\colon \{0,1\}^n \to \{0,1\}$ 为任意一个大小为 $s = 2^{(1-\delta)n}$ 的 ROBP 。令 $X$ 为在 $\mathsf{F}_2^n$上均匀分布的随机变量。则

$$(\mathsf{AC}^0\text{-}\mathsf{Ext}(X), B, E, f(X)) \approx_\varepsilon (U, B, E, f(X)).$$

证明。 注意到 $\mathsf{AC}^0$-Ext 是一个强（$\delta n/3$, $\varepsilon$）提取器，因此根据引理 13，它是一个（$\delta n/3 + \mathrm{poly}\log n$, $2\varepsilon$）平均情形提取器。由于 $\widetilde{H}_\infty(A \mid E) = 2^{n-\delta n/3-\log(2s)} = 2\delta/3 - 1 \geq \delta n/3 + \mathrm{polylog}\ n$，我们有

$$(\mathsf{AC}^0\text{-}\mathsf{Ext}(X), B, E) \approx_{2\varepsilon} (U, B, E).$$

由于 $f(X) = g(E, B)$ 是 $E$ 和 $B$ 的确定性函数，我们可以得出

$$(\mathsf{AC}^0\text{-}\mathsf{Ext}(X), B, E, f(X)) \approx_{2\varepsilon} (U, B, E, f(X)).$$

□

## 7 开放问题

我们的工作留下了若干自然的开放问题。最显而易见的问题是进一步改进方向仿射提取器的构造，以及针对 SROLBP 的平均情形困难性。证明针对 WROLBP 的显式函数具有任何困难性也会很有意思，这似乎需要新的思路。最后，一个值得探讨的问题是，是否存在 $\mathsf{AC}^0$ 中的函数，能够对 ROBP 达到最优困难性，或对 SROLBP 达到强困难性。

## 致谢

我们感谢匿名审稿人提出的宝贵意见，并感谢一位审稿人指出了 [GI17]。

## 参考文献

[ABCR99] Alexander E. Andreev、Juri L. Baskakov、Andrea E. F. Clementi 和 Jos´e D. P. Rolim。小型伪随机集合产生困难函数：分支程序新的紧致显式下界。载于 Jir´ı Wiedermann、Peter van Emde Boas 和 Mogens Nielsen（编辑），《自动机、语言与编程》，第26届国际研讨会，ICALP'99，捷克共和国布拉格，1999年7月11日至15日，论文集，《计算机科学讲义》第1644卷，第179–189页。Springer，1999年。

[BISW04] Boaz Barak、Russel Impagliazzo、Amir Shpilka 和 Avi Wigderson。维度扩展器的定义与存在性。讨论（无书面记录），2004年。[BKS+05] Boaz Barak、Guy Kindler、Ronen Shaltiel、Benny Sudakov 和 Avi Wigderson。模拟独立性：凝缩器、拉姆齐图、分散器和提取器的新构造。载于第37届ACM计算理论研讨会年会论文集，第1–10页，2005年。[BKS+10] Arnab Bhattacharyya、Swastik Kopparty、Grant Schoenebeck、Madhu Sudan 和 David Zuckerman。Reed–Muller码的最优测试。载于第51届IEEE计算机科学基础年会，FOCS 2010，2010年10月23日至26日，美国内华达州拉斯维加斯，第488–497页。IEEE 计算机学会，2010年。[Bou09] Jean Bourgain。扩展器与维度扩张。《数学报告》，347(7)：357–362，2009年。[BS94] Antal Balog 和 Endre Szemer´edi。集合加法的一个统计定理。《组合数学》，14：263–268，1994年。[BW98] Beate Bollig 和 Ingo Wegener。一个需要指数规模读一次分支程序的极简单函数。《信息处理快报》，66(2)：53–57，1998年。[BY13] Jean Bourgain 和 Amir Yehudayoff。 $SL_2(\mathbb{R})$中的扩张与单调扩张器。《几何与泛函分析》，23(1)：1–41，2013年。[CG88] Benny Chor 和 Oded Goldreich。从弱随机性源中获得无偏比特与概率通信复杂度。《SIAM计算杂志》，17(2)：230–261，1988年。[CGL16] Eshan Chattopadhyay、Vipul Goyal 和 Xin Li。不可篡改提取器与码及其多种篡改扩展。载于第四十八届 ACM 计算理论研讨会年会论文集，STOC ’16，第285–298页，美国纽约州纽约，2016年。计算机协会。[CGL22] Eshan Chattopadhyay、Jesse Goodman 和 Jyun-Jie Liao。近似对数熵的仿射提取器。载于2021年第62届IEEE计算机科学基础年会（FOCS），第622–633页，2022年。[CL16a] Eshan Chattopadhyay 和 Xin Li。显式不可篡改提取器、多源提取器与近乎最优的隐私放大协议。载于2016年IEEE第57届计算机科学基础年会（FOCS），第158–167页，2016年。[CL16b] Eshan Chattopadhyay 和 Xin Li。和集源的提取器。载于 Daniel Wichs 和 Yishay Mansour（编辑），第48届ACM SIGACT计算理论研讨会年会论文集，STOC，美国马萨诸塞州剑桥，2016年6月18日至21日，第299–311页。ACM，2016年。[CL17] Eshan Chattopadhyay 和 Xin Li。小深度电路的不可篡改码和提取器，以及仿射函数。载于 Hamed Hatami、Pierre McKenzie 和 Valerie King（编辑），第49届ACM SIGACT计算理论研讨会年会论文集，STOC 2017，加拿大魁北克省蒙特利尔，2017年6月19日至23日，第1171–1184页。ACM，2017年。

[CL18] Kuan Cheng 和 Xin Li。《AC0 中的随机性提取与小局部性》。载于 Eric Blais、Klaus Jansen、Jos´e D. P. Rolim 和 David Steurer 编，《近似、随机化与组合优化：算法与技术》，APPROX/RANDOM 2018，2018年8月20日至22日，美国新泽西州普林斯顿，LIPIcs第116卷，第37:1–37:20页。Schloss Dagstuhl - Leibniz-Zentrum f¨ur Informatik，2018年。[CL22] Eshan Chattopadhyay 和 Jyun-Jie Liao。《两个源的提取器》。载于 Stefano Leonardi 和 Anupam Gupta 编，《STOC '22：第54届ACM SIGACT计算理论年会》，意大利罗马，2022年6月20日至24日，第1584–1597页。ACM，2022年。[CL23] Eshan Chattopadhyay 和 Jyun-Jie Liao。《针对线性分支程序的困难性及其他结果》。载于《第38届计算复杂性会议论文集》，CCC '23，德国达格施图尔，2023年。Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik。[CS16] Gil Cohen 和 Igor Shinkar。《奇偶函数析取范式的复杂度》。载于 Madhu Sudan 编，《2016年ACM理论计算机科学创新会议论文集》，美国马萨诸塞州剑桥，2016年1月14日至16日，第47–58页。ACM，2016年。[CT15] Gil Cohen 和 Avishay Tal。《低次多项式的两个结构性结果及其应用》。载于《近似、随机化与组合优化：算法与技术》，APPROX/RANDOM 2015，2015年。[DK11] Evgeny Demenkov 和 Alexander Kulikov。《仿射分散器电路复杂度的一个初等3n-o(n)下界证明》。载于《第36届国际计算机科学数学基础会议论文集》，第256–265页，2011年。[DORS08] Yevgeniy Dodis、Rafail Ostrovsky、Leonid Reyzin 和 Adam Smith。《模糊提取器：如何从生物特征识别数据及其他含噪数据中生成强密钥》。《SIAM计算杂志》，38(1)：97–139，2008年。[DS11] Zeev Dvir 和 Amir Shpilka。《关于有限域上的维度扩展器》。《Combinatorica》，31(3)：305，2011年。[Dun85] Paul E. Dunne。《单次分支程序复杂度的下界》。载于 Lothar Budach 编，《计算理论基础》，FCT '85，东德科特布斯，1985年9月9日至13日，《计算机科学讲义》第199卷，第90–99页。Springer，1985年。[FGHK16] Magnus Gausdal Find、Alexander Golovnev、Edward A. Hirsch 和 Alexander S. Ku- likov。《显式函数电路复杂度的优于3n的下界》。载于《2016年IEEE第57届计算机科学基础年会》（FOCS），第89–98页，2016年。[G´al97] Anna G´al。《一个需要指数规模读一次分支程序的简单函数》。《信息处理快报》，62(1)：13–16，1997年。[GGMT23] W. T. Gowers、Ben Green、Freddie Manners 和 Terence Tao。《关于马顿猜想》，2023年。

[GI17] Ludmila Glinskih 和 Dmitry Itsykson。《可满足的Tseitin公式对非确定性读一次分支程序而言是困难的》。载于 Kim G. Larsen、Hans L. Bod- laender 和 Jean-Francois Raskin 编，《第42届国际计算机科学数学基础研讨会》（MFCS 2017），LIPIcs第83卷，第26:1–26:12页，德国达格施图尔，2017年。Schloss Dagstuhl – Leibniz-Zentrum f¨ur Informatik。[Gow98] William T. Gowers。《关于长度为四的算术级数的Szemer´edi定理的新证明》。《几何与泛函分析》（GAFA），8：529–551，1998年。[GPT22] Svyatoslav Gryaznov、Pavel Pudl´ak 和 Navid Talebanfard。《线性分支程序与方向仿射提取器》。载于《第37届计算复杂性会议》（CCC 2022），第234卷，第4:1–4:16页，2022年。[H˚as86] Johan H˚astad。《小深度电路的近乎最优下界》。载于 Juris Hartma- nis 编，《第18届ACM计算理论研讨会年会论文集》，1986年5月28日至30日，美国加利福尼亚州伯克利，第6–20页。ACM，1986年。[Juk88] Stasys Jukna。《接触电路的熵及其复杂度下界》。《理论计算机科学》，57：113–129，1988年。[Juk95] Stasys Jukna。《关于读k次分支程序的注记》。《RAIRO——理论信息学与应用》，28：75–83，1995年1月。[Kab03] Valentine Kabanets。《近乎k-独立与困难布尔函数》。《理论计算机科学》，297(1-3)：281–295，2003年。[KMW91] Matthias Krause、Christoph Meinel 和 Stephan Waack。《区分擦除图灵机类 l e、nl e、co-nl e 和 p e》。《理论计算机科学》，86(2)：267–275，1991年。[Li11] Xin Li。《仿射提取器与分散器的新方法》。载于《第26届IEEE计算复杂性年会论文集》，CCC，2011年。[Li12] Xin Li。《不可篡改提取器、双源提取器与隐私放大》。载于《2012年IEEE第53届计算机科学基础年会》，2012年。[Li17] Xin Li。《改进的不可篡改提取器、不可篡改码与独立源提取器》。载于《第49届ACM SIGACT计算理论研讨会年会论文集》，STOC 2017，蒙特利尔，加拿大魁北克省，2017年6月19日至23日，STOC 2017，第1144–1156页，美国纽约州纽约，2017年。计算机协会。[Li23] Xin Li。《具有渐近最优熵的双源提取器及更多结果》。技术报告，Arxiv，2023年。https://arxiv.org/abs/2303.06802。[LY22] Jiatu Li 和 Tianqi Yang。《显式函数的3.1n - o(n)电路下界》。载于《第54届ACM SIGACT计算理论年会论文集》，STOC 2022，第1180–1193页，美国纽约州纽约，2022年。计算机协会。[Nec66] E. I. Nechiporuk。《关于一个布尔函数》。《苏联科学院学报》，164(4)：765–766，1966年。

[Oko93] EA Okolnishnikova. 关于分支程序下界的研究。《西伯利亚数学进展》，3:152–156，1993年1月。[Pon98] Stephen Ponzio. 使用读一次分支程序进行整数乘法的下界。《SIAM计算杂志》，28(3):798–815，1998年。[PWY16] Periklis A Papakonstantinou、David P Woodruff 和 Guang Yang. 从大数据中获取真随机性。《科学报告》，6:33740，2016年。[Rao 09] Anup Rao. 低权仿射源的提取器。载于《2009年第24届IEEE计算复杂性年度会议论文集》（CCC '09），第95–101页。IEEE 计算机学会，2009年。[Raz05] Ran Raz. 使用弱随机种子的提取器。载于第37届 ACM 计算理论研讨会论文集，第11–20页，2005年。[SS92] Janos Simon 和 Mario Szegedy. 只读一次分支程序的新下界定理及其应用。载于《计算复杂性理论进展》，1992年。[TV06] Terence Tao 和 Van H. Vu.《加性组合学》。《剑桥高等数学研究》。剑桥大学出版社，2006年。[Vaz86] Umesh Virkumar Vazirani.《随机性、对手与计算（随机多项式时间）》。博士论文，加州大学伯克利分校，1986年。AAI8718194。[VW08] Emanuele Viola 和 Avi Wigderson. 多项式与协议的范数、异或引理及下界。《计算理论》，4(7):137–168，2008年。[Weg88] Ingo Wegener. 团函数的分支程序和决策树的复杂性。《ACM杂志》，35(2):461–471，1988年。[Yeh11] Amir Yehudayoff. 素数域上的仿射提取器。《组合数学》，31(2):245–256，2011年。[Z´ak84] Stanislav Z´ak. 一次性分支程序的指数下界。载于 Michal Chytil 和 Václav Koubek 编，《1984年计算机科学数学基础：1984年9月3日至7日，捷克斯洛伐克布拉格》，论文集，《计算机科学讲义》第176卷，第562–566页。Springer，1984年。[Zuc07] David Zuckerman. 线性次数提取器与最大团和色数的不可近似性。载于《计算理论》，2007年。

## 深度 3 的 $AC^0[\oplus]$ 电路可以计算最优方向仿射提取器

在本节中，我们扩展 [CT15] 中的结果，并证明深度为 3 的 $AC^0[\oplus]$ 电路可以计算由概率方法给出的最优方向仿射提取器。

方向仿射提取器的存在性。我们首先展示最优方向仿射提取器。

命题 71. 存在普适常数 $n_0$、$c$，使得以下结论成立。对于任意 $\varepsilon > 0$ 和 $n > n_0$，都存在一个维数为 $k$、偏差为 $\varepsilon$ 的方向仿射提取器 $F\colon \mathsf{F}_2^n \to \mathsf{F}_2,$，其中 $k = \log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + c$。

证明。 为便于证明，我们使用 DAExt在 [GPT22]中的定义来处理。

定义 30. 布尔函数 $f\colon \mathsf{F}_2^n \to \mathsf{F}_2$ 是一个维数为 $d$、偏差为 $\varepsilon$ 的方向仿射提取器，当且仅当对于每个仿射子空间 $X$ 以及每个非零 $a$，都有

$$\mathsf{DAExt}(X) + \mathsf{DAExt}(X + a) \approx_\varepsilon U_1.$$

该定义与定义 2 等价，但误差会发生二次增长。证明请参见附录 $B$中的 [CL23]。令 $F\colon \mathsf{F}_2^n \to \mathsf{F}_2$为一个随机函数，即 $\{F(x)\ x \in \mathsf{F}_2^n\}$是新生成的随机比特。固定一个维数为 $U \subseteq \mathsf{F}_2^n$ 的仿射子空间 $k$，以及一个非零 $a \in \mathsf{F}_2^n$。根据 $U + a$ 是否与 $U$重合，需要考虑两种情况。

情况 1。 $U + a \neq U$。对于任意 $x_1$、$x_0 \in U$、$x_1 \neq x_0$，由于 $x_1 + x_0 \in U$，有 $x_1 + a \notin \{x_0, x_0 + a\}$。因此，$\{F(x) + F(x + a)\ x \in U\}$ 是独立随机比特，并且有

$$\mathbf{Pr}\left[\frac{1}{2^k}\left|\sum_{x\in U}(-1)^{F(x)+F(x+a)}\right| \geq \varepsilon\right] \leq 2 \cdot e^{-\frac{2^k\varepsilon^2}{2}}. \qquad \text{(Hoeffding Inequality)}$$

情况 2。 $U + a = U$。对于任意 $x_1$、$x_0 \in U$ $x_1 \neq x_0$、$x_1 \in \{x_0, x_0 + a\} \iff x_1 = x_0 + a$。如果情况如此，那么 $F(x)+F(x+a) = F(x+a)+F(x)$。因此，$\{\left((-1)^{F(x)+F(x+a)} + (-1)^{F(x+a)+F(x)}\right)/2、x \in \mathsf{F}_2^n\}$ 是支撑集为 $\{-1, 1\}$ 的独立随机变量，并且有

$$\mathbf{Pr}\left[\frac{1}{2^k}\left|\sum_{x\in U}(-1)^{F(x)+F(x+a)}\right| \geq \varepsilon\right] \leq 2 \cdot e^{-\frac{2^{k-1}\varepsilon^2}{2}}. \qquad \text{(Hoeffding Inequality)}$$

具有相同底层线性子空间的仿射子空间对的数量至多为$\binom{2^n}{2}\binom{2^n}{k} \leq 2^{(k+2)n}$。因此，对所有具有相同底层线性子空间的仿射子空间对应用并集界，若 $2^{(k+2)n} \cdot 2 \cdot e^{-\frac{2^k\varepsilon^2}{2}} = 2^{(k+2)n+1-\frac{2^k\varepsilon^2}{\ln 4}} < 1$且 $2^{(k+2)n} \cdot 2 \cdot e^{-\frac{2^{k-1}\varepsilon^2}{2}} = 2^{(k+2)n+1-\frac{2^{k-1}\varepsilon^2}{\ln 4}} < 1$，则存在一个维数为 $k$、误差为 $\varepsilon$的方向仿射提取器。可以验证，选择与 $k = \log\frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + c$ 相同的某个固定常数 $c$，即可使上述不等式成立。 [CT15]中也是如此。

和集线性注入器的存在性。下面的和集线性注入器定义略微推广了 [CT15]中的注入器概念。它们将用于构造一个更加“结构化”的随机函数，该函数是一个DAExt。

定义31. 一个大小为 $m$ 的 $(n, k_1, k_2, d)$ 和集线性注入器是一个由 $d \times n$ 个矩阵$\{A_1, \cdots, A_m\}$ 组成的族，它们定义在 $\mathsf{F}_2$ 上，并具有以下性质：对于任意一对子空间 $U$、$V \subseteq \mathsf{F}_2^n$，其维数分别为 $k_1$、$k_2$，且 $\dim(U \cap V) \leq 1$，存在一个 $i \in [m]$，使得 $\ker(A_i) \cap (U + V) = \{0\}$。

引理 72. 对于满足 $2 \leq k_1$、$k_2 \leq n$ 的任意 $n$、$k_1$、$k_2$，存在一个大小为 $m = n(k_1 + k_2)$ 的（$n$、$k_1$、$k_2$、$k_1 + k_2 + 1$）线性注入器。

证明。固定一对子空间 $U$、$V \subseteq \mathbf{F}_2^n$，其维数分别为 $k_1$、$k_2$，其中 $U \cap V = \{0\}$。令 $A$ 为一个 $d \times n$ 矩阵，使得 $A$ 的每个元素均从 $\mathbf{F}_2$ 中均匀且独立地随机抽取。对于每个 $u+v \in (U+V) \setminus \{0\}$，有 $\mathbf{Pr}[A(u+v) = 0] = 2^{-d}$。对 $U \setminus \{0\}$ 和 $V \setminus \{0\}$ 中所有元素对应用并集界，可得

$$\mathbf{Pr}\left[\ker(A) \cap (U+V) \neq \{0\}\right] \leq 2^{k_1+k_2-d}.$$

令 $A_1$、$\cdots$、$A_m$ 为 $d \times n$ 矩阵，其中每个矩阵的元素均从 $\mathbf{F}_2$ 中均匀且独立地随机采样。由上述等式可得

$$\mathbf{Pr}\left[\forall i \in [m]\ \ker(A_i) \cap (U+V) \neq \{0\}\right] \leq 2^{m(k_1+k_2-d)}.$$

维数为 $k_1$ 和 $k_2$ 的两个线性子空间之和的数量至多为 $\binom{2^n}{k_1}\binom{2^n}{k_2}$，而其上界为 $2^{n(k_1+k_2)-2}$，其中 $k \geq 2$。因此，如果 $2^{n(k_1+k_2)-2} \cdot 2^{m(k_1+k_2-d)} < 1$，则存在一个大小为 $m$ 的 $(n, k_1, k_2, d)$ 线性注入器。后一等式对 $d = k_1 + k_2 + 1$ 和 $m = n(k_1+k_2)$ 成立。

□

结构更丰富的随机函数。现在我们应用和集注入器，以减少命题 71 中使用的随机性。

引理 73. 设 $n_0, c$ 为命题 71 中的常数。设 $n > n_0$，并设 $k, \varepsilon$ 满足 $k = \log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + c$。设 $\{A_1, \cdots, A_m\}$ 是一个大小为 $m$ 的 $(n, k, 2, d)$ 线性注入器。那么，存在函数 $f_1, \cdots, f_m : \mathbf{F}_2^d \to \mathbf{F}_2$，使得由下式定义的函数 $f : \mathbf{F}_2^n \to \mathbf{F}_2$

$$f(x) = \bigoplus_{i=1}^{m} f_i(A_i x) \tag{6}$$

是一个维数为 $k$、偏差为 $\varepsilon$ 的方向仿射提取器。

证明. 证明思路是，对于证明命题 71，只要满足“$(U+V)$-路”独立性即可，其中 $U$ 是任意仿射子空间，并且对任意 $a \neq 0 \in \mathbf{F}_2$，$V = \{0, a\} = \mathsf{span}\{0, a\}$。换言之，我们只需要 $\{f(x)\}_{x \in U \cup (U+a)}$ 是独立随机比特，而不需要 $f$ 的真值表上的完全独立性。现在我们构造这样一个随机函数，并在命题 71 的证明中用这个新构造的函数替换随机函数，由此在一类受限的随机函数中找到最优方向仿射提取器。这将使我们能够分析其复杂度。设 $F_1, \cdots, F_m : \mathbf{F}_2^d \to \mathbf{F}_2$ 是相互独立的随机函数，也就是说，随机比特 $\{F_i(x) : i \in [m], x \in \mathbf{F}_2^d\}$ 是独立的。定义随机函数 $F : \mathbf{F}_2^n \to \mathbf{F}_2$ 如下

$$F(x) = \bigoplus_{i=1}^{m} F_i(A_i x).$$

设 $(U, U+a)$ 是任意一对具有相同底层线性子空间 $U'$ 的仿射子空间，令 $V = \mathsf{span}\{0, a\}$。根据定义 31，存在一个 $i \in [m]$，使得 $\ker(A_i) \cap (U'+V) = \{0\}$。这意味着，对于任意两个不同的元素 $u$，$v \in U$，$A_i(u)$ $A_i(v)$ $A_i(u+a)$ $A_i(v+a)$ 两两不同。否则便会得到矛盾：$A_i(u+v) = 0$ 或 $A_i(a) = 0$ 或 $A_i(u+v+a) = 0$，从而 $u+v$ 或 $a$ 或 $u+v+a$（一个非零向量）属于 $U'+V$，且落在 $\ker(A_i)$ 中。由于 $F_i$ 是一个随机函数，且 $A_i$ 在 $U \cup (U+a)$ 上是单射，因此随机比特 $\{F_i(u)\}_{u \in U \cup (U+a)}$ 相互独立。由于对所有 $x \in (U \cup (U+a))$，新鲜的随机硬币 $F_i(A_i x)$ 仅用于生成 $F(x)$，因此 $F(x)$ 在 $U \cup (U+a)$ 中是独立且随机的。

□

定理 74。 令 $f$为方程 (6)中的函数，其中$\{A_1, \cdots, A_{n(k+2)}\}$是引理 $(n, k, 2, k+3)$所述的和集线性注入器。则 $f$ 是一个维数为 $k$、误差为 $\varepsilon$的方向仿射提取器，其中 $k = \log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + O(1)$。此外，

1. $\deg(f) = \log \frac{n}{\varepsilon^2} + \log\log \frac{n}{\varepsilon^2} + O(1)$。2. $f$ 可以由大小为 $O((n/\varepsilon)^2 \cdot \log^3(n/\varepsilon))$ 的异或-与-异或电路实现。3. $f$ 可以由大小为 $O((n^5/\varepsilon^2) \cdot \log^3(n/\varepsilon))$ 的德摩根公式实现。

证明。 与[CT15]完全相同。 □

□

# B 缺失的证明

## B.1 引理 8 的证明

我们回顾引理8。

引理 8 (仿射条件化 [Li11])。 设 $X$ 是定义在 $\{0,1\}^n$上的任意仿射源。设 $L: \{0,1\}^n \to \{0,1\}^m$ 是任意仿射函数。那么存在独立仿射源 $A$, $B$ 使得:

- $X = A + B$
- 存在 $c \in \{0,1\}^m$, 使得对每个 $b \in \mathrm{Supp}(B)$, 都有 $L(b) = c$。
- $H(A) = H(L(A))$, 并且存在一个仿射函数 $L^{-1}: \{0,1\}^m \to \{0,1\}^n$, 使得$A = L^{-1}(L(A))$。
- $H(X|_{L(X)=\ell}) = H(B)$, 对所有 $\ell \in \mathrm{Supp}(L(X))$ 均成立。

证明。我们证明第二和第四条。
第二条。令 $L = \overline{L} + c'$, 其中 $\overline{L}: \{0,1\}^n \to \{0,1\}^m$ 是一个线性函数。考虑集合 $\mathrm{Supp}(X) \cap \mathrm{Ker}(\overline{L})$, 它是一个线性子空间；令 $B$ 为该线性子空间加上任意仿射平移 $c''$, 则有 $L(B) = L(c'') = \overline{L}(c'') + c' := c$。令 $A = X - B$。那么$L(A) = L(X) - L(B) = \overline{L}(X) + c' - c = \overline{L}(X - (B - c'')) - \overline{L}(c'') = \mathrm{Supp}(X) \cap$张成空间$(\overline{L}) - \overline{L}(c'')$。
第 4 个要点。对于任意 $\ell \in \mathrm{Supp}(L(X))$, 在固定 $L(X) = \ell$ 的条件为定值时，根据第二个要点，有 $L(X) = L(A) + L(B) = L(A) + c = \ell$。根据第三个要点，这意味着 $A = L^{-1}(L(A)) = L^{-1}(\ell - c)$。因此, $H(X|_{L(X)=\ell}) = H(L^{-1}(\ell - c) + B) = H(B)$, 从而独立于 $\ell$。 □

## B.2 引理 35 的证明

我们回顾引理35。

引理 35（仿射源的独立性合并引理）。设 $\mathrm{LExt}: \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$是任意一个（$k$, $\varepsilon$）强线性种子提取器, $X_0 \in \{0,1\}^n$ 是一个仿射源, $X$、$X^{[t]} \in \{0,1\}^n$、$Y$、$Y^{[t]} \in \{0,1\}^d$都是 $X_0$ 的线性函数，并且对于每个 $j \in [t]$, 都有 $W = \mathrm{LExt}(X, Y)$ 和 $W^j = \mathrm{LExt}(X^j, Y^j)$。假设存在 $S$、$T \subseteq [t]$ 使得

- $(Y, Y^S) \approx_\delta (U_d, Y^S)$;

- $H(X \mid X^T, Y, Y^{[t]}) \geq k + tm,$

*then*

$$W \approx_{\varepsilon+\delta} U_m \mid (W^{S\cup T}, Y, Y^{[t]}).$$

证明。 首先注意到，由于 $X$、$X^T$、$Y$、$Y^{[t]}$ 都是 $X_0$的线性函数，根据引理 8，$X$ 在给定 $X^T$、$Y$、$Y^{[t]}$ 的条件下的熵是常数。因此，只需使用香农熵 $H$ 代替平均~情况下的最小熵 $H_\infty$。

在固定 $Y^S$ 后，当条件为定值时，有 $W^S$ 是 $X^S$ 的线性函数，因此也是 $X_0$ 的线性函数。根据引理 8，存在仿射源 $A = W^S(X)$ 和 $B$，使得 $X = A + B$。根据引理 9，$H(B) = H(X \mid Y^S, W^S) \geq H(X) - |S| \cdot m$。现在进一步将 $(X^T, Y, Y^{[t]\setminus S})$ 条件化，我们有 $H(B \mid X^T, Y, Y^{[t]\setminus S}) \geq H(X \mid X^T, Y, Y^T) - |S| \cdot m \geq k$。根据命题 41，可知以概率 $1-\varepsilon$，$\mathrm{LExt}(X, Y) = \mathrm{LExt}(B, Y) + \mathrm{LExt}(A, Y) = \mathrm{LExt}(B, Y) + \mathrm{const} = U_m$。由于 $W^T$ 是 $X^T$ 和 $Y^T$ 的确定性函数，我们所证明的结果意味着

$$W \approx_{\varepsilon+\delta} U_m \mid (W^{S\cup T}, Y, Y^{[t]}).$$

□