Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 125 (2015)

Improved Two-Source Extractors, and Affine Extractors for Polylogarithmic Entropy

Xin Li
Department of Computer Science
Johns Hopkins University
Baltimore, MD 21218, U.S.A.
lixints@cs.jhu.edu *

Abstract

In a recent breakthrough [CZ16], Chattopadhyay and Zuckerman gave an explicit two-source extractor for min-entropy $k \geq \log^C n$ for some large enough constant C, where n is the length of the source. However, their extractor only outputs one bit. In this paper, we improve the output of the two-source extractor to $k^{\Omega(1)}$, while the error remains $n^{-\Omega(1)}$ and the extractor remains strong in the second source. In the non-strong case, the output can be increased to k. Our improvement is obtained by giving a better extractor for (q, t, γ) non-oblivious bit-fixing sources, which can output $t^{\Omega(1)}$ bits instead of one bit as in [CZ16].

We also give the first explicit construction of deterministic extractors for affine sources over $\mathbb{F}_2$, with entropy $k \geq \log^C n$ for some large enough constant C, where n is the length of the source. Previously the best known results are by Bourgain [Bou07], Yehudayoff [Yeh11] and Li [Li11b], which require the affine source to have entropy at least $\Omega(n/\sqrt{\log \log n})$. Our extractor outputs $k^{\Omega(1)}$ bits with error $n^{-\Omega(1)}$. This is done by reducing an affine source to a non-oblivious bit-fixing source, where we adapt the alternating extraction based approach in previous work on independent source extractors [Li13a] to the affine setting. Our affine extractors also imply improved extractors for circuit sources studied in [Vio11].

We further extend our results to the case of zero-error dispersers, and give two applications in data structures that rely crucially on the fact that our two-source or affine extractors have large output size.

*Partially supported by NSF Grant CCF-1617713.

ISSN 1433-8092

1 Introduction

Randomness extraction is a broad area that studies the problem of converting biased random sources into nearly uniform random bits. The natural motivation comes from the wide application of randomness in computation, such as in algorithms, distributed computing and cryptography, and the requirement that the random bits used should be uniformly distributed. In reality, however, natural random sources almost always have serious biases, and can leak information to an adversary because of side channel attacks. These defective random sources are known as weak random sources. Therefore, intuitively, a randomness extractor takes as input one or more weak random sources, and outputs a distribution that is statistically close to uniform.

Formally, a weak random source is modeled as a probability distribution over n bit strings with some entropy k. In the context of randomness extraction, the standard measure of entropy is the so called *min-entropy*, which is defined as follows.

Definition 1.1. The *min-entropy* of a random variable X is

$$H_\infty(X) = \min_{x \in \mathsf{supp}(X)} \log_2(1/\Pr[X = x]).$$

For $X \in \{0,1\}^n$, we call X an $(n, H_\infty(X))$-source, and we say X has *entropy rate* $H_\infty(X)/n$.

However, one can easily show that it is impossible to construct deterministic randomness extractors for one (n,k) source, even if k is as large as $n-1$. Thus, the study of randomness extractors has been pursued in two different directions. The first one is to allow the extractor itself to be randomized. In this case one ends up with the notion of *seeded extractors* [NZ96], where the extractor is given a short independent uniform random seed (typically of length say $O(\log n)$). It is now possible to construct such extractors for all weak random sources. Typically, one also requires the output of the extractor to be close to uniform even given the seed. Such extractors are known as *strong seeded extractors*. Seeded extractors have a lot of applications in theoretical computer science and have been studied extensively, resulting in almost optimal constructions [LRVW03, GUV09, DKSS09].

Another direction is to impose some special structure on the weak source, and thereby allows the construction of deterministic randomness extractors. This is the focus of this paper. Formally, we have the following definition.

Definition 1.2. (Deterministic extractors for structured sources) Let $\mathcal{C}$ be a class of distributions on a finite set Ω. A function $E : \Omega \to \{0,1\}^m$ is an extractor for $\mathcal{C}$ with entropy threshold k and error ϵ, if for any weak source $X \in \mathcal{C}$ with entropy at least k, we have

$$|E(X) - U_m| \leq \epsilon.$$

Here U_m is the uniform distribution over $\{0,1\}^m$ and $|\cdot|$ stands for the statistical distance.

In this paper we will study the following classes of weak sources.

Independent Sources Here, the extractor is given as input more than one general weak random sources, and the sources are independent of each other. Using the probabilistic method, one can show that there exists a deterministic extractor for just two independent sources with each having logarithmic min-entropy, which is optimal since extractors for one weak source do not exist. In fact, the probabilistic method shows that with high probability a random function is such a two-source

extractor. However, the most interesting and important part is to give explicit constructions of such functions, which turns out to be highly challenging.

The first explicit construction of a two-source extractor appeared in [CG88], where Chor and Goldreich showed that the well known Lindsey's lemma gives an extractor for two independent (n, k) sources with $k > n/2$. Since then there has been essentially no progress on two-source extractors until in 2005 Bourgain [Bou05] gave a construction that breaks the entropy rate $1/2$ barrier, and works for two independent $(n, 0.49n)$ sources. In a different work, Raz [Raz05] gave an incomparable result of two source extractors which requires one source to have min-entropy larger than $n/2$, while the other source can have min-entropy $O(\log n)$.

Given the difficulty of constructing explicit two-source extractors, much research has been focusing on a slightly more general model, where the extractor is allowed to have more then two independent sources as the input. Starting from [BIW04], there has been a long line of fruitful results [BIW04, Raz05, Bou05, Rao06, BRSW06, Li11a, Li13b, Li13a, Li15, Coh15], which introduced many new techniques and culminated in the three source extractor of exponentially small error by the author [Li15]. However, in the two source case the situation has not been improved.

Recently, Chattopadhyay and Zuckerman [CZ16] made a breakthrough to this problem by giving the first explicit two-source extractors for (n, k) sources with $k \geq \log^C n$ for some large enough constant C. This dramatically improves the situation of two-source extractors and is actually near optimal. However, their construction only outputs one bit and thus a natural question is whether one can achieve a significantly larger output length.

Affine Sources An affine source is the uniform distribution over some unknown subspace of a vector space, and affine extractors are deterministic extractors for such sources.

Definition 1.3. (affine source) Let $\mathbb{F}_q$ be the finite field with q elements. Denote by $\mathbb{F}_q^n$ the n-dimensional vector space over $\mathbb{F}_q$. A distribution X over $\mathbb{F}_q^n$ is an $(n, k)_q$ affine source if there exist linearly independent vectors $a_1, \cdots, a_k \in \mathbb{F}_q^n$ and another vector $b \in \mathbb{F}_q^n$ s.t. X is sampled by choosing $x_1, \cdots, x_k \in \mathbb{F}$ uniformly and independently and computing

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

In this paper we focus on the case where $q = 2$. Using the probabilistic method, it is not hard to show that there exists a deterministic affine extractor, as long as $k > 2 \log n$ and the output length $m < k - O(1)$. The problem is to given an explicit construction of such a function.

There has also been a lot of work studying affine extractors and dispersers. For example, Gabizon and Raz [GR05] constructed explicit extractors for affine sources even with entropy 1. However, their constructions require the field size to be much larger than n, i.e., $q > n^{\Omega(1)}$, in order to use Weil's theorem. DeVos and Gabizon [DG10] constructed explicit extractors for $(n, k)_q$ affine sources when $q = \Omega((n/k)^2)$ and the characteristic of the field $\mathbb{F}_q$ is $\Omega(n/k)$. As the field size gets smaller, constructing explicit affine extractors becomes significantly harder.

The extreme and hardest case where the field is $\mathbb{F} = \mathrm{GF}(2)$, is the focus of the rest of the paper. Note that in this case the min-entropy $H_{\infty(X)}$ is the same as the standard Shannon entropy $H(X)$. Here, it is well known how to construct extractors for affine sources with entropy rate greater than $1/2$. However the problem becomes much harder as the entropy rate drops to $1/2$ and below $1/2$. Bourgain [Bou07] used sophisticated character sum estimates to give an extractor for affine sources

with entropy $k = \delta n$ for any constant $\delta > 0$. This was later slightly improved to $k = \Omega(n/\sqrt{\log \log n})$ by Yehudayoff [Yeh11] and the author [Li11b], which have remained the best known results. Rao [Rao09] constructed extractors for affine sources with entropy as small as polylog(n), as long as the subspace of X has a basis of low-weight vectors. In the case where one only wishes to output one bit with support $\{0,1\}$ (i.e., a *disperser*), Ben-Sasson and Kopparty [BSK09] gave constructions for entropy $\Omega(n^{4/5})$, and Shaltiel [Sha11] gave a construction for entropy $2^{\log^{0.9} n}$.

Circuit Sources Trevisan and Vadhan [TV00] considered the question of extracting random bits from *samplable sources*, which are n-bit distributions generated by some small circuit from ℓ uniform bits. They showed that such extractors imply circuit lower bounds for related circuits. They also constructed explicit extractors for such sources with min-entropy $k = \Omega(n)$ under some necessary computational assumptions (such as the existence of a function computable in time $2^{O(n)}$ which requires $2^{\Omega(n)}$ size Σ_5 circuits). Subsequently, Viola [Vio11] constructed unconditional extractors for sources generated by local circuits and circuits of small depth (e.g., NC^0 and AC^0 circuits). However, for both these sources, the extractors in [Vio11] require min-entropy at least $n^{2/3+\Omega(1)}$.

Non-oblivious bit-fixing Sources As in [CZ16], an intermediate class of sources we use in our construction is a special kind of non-oblivious bit-fixing source. Non-oblivious bit-fixing sources are sources on n bits which are uniform except some unknown q bits. However the q bits can depend arbitrarily on the $n - q$ uniform bits. Extractors for non-oblivious bit-fixing sources are equivalent to resilient functions, and were studied in [BOL78, KKL88, KZ07, Vio11]. What we need here is a special kind of non-oblivious bit-fixing sources which only requires bounded independence in the "good" bits. Such sources were first defined by Viola in [Vio11], where he constructed an extractor that extracts one bit from a non-oblivious bit-fixing source with $q \approx \sqrt{n}$ bad bits, and the "good" bits are polylog(n)-wise independent. Subsequently, [CZ16] gave an improved one-bit extractor that can handle $q = n^{1-\delta}$ for any constant $\delta > 0$. We now formally define such sources.

Definition 1.4. A distribution $\mathcal{D}$ on n bits is t-wise independent if the restriction of $\mathcal{D}$ to any t bits is uniform. Further $\mathcal{D}$ is a (t, ϵ)-wise independent distribution if the distribution obtained by restricting $\mathcal{D}$ to any t coordinates is ϵ-close to uniform.

Definition 1.5. A source X on $\{0,1\}^n$ is called a (q,t)-non-oblivious bit-fixing source if there exists a subset of coordinates $Q \subseteq [n]$ of size at most q such that the joint distribution of the bits indexed by $\overline{Q} = [n] \setminus Q$ is t-wise independent. The bits in the coordinates indexed by Q are allowed to arbitrarily depend on the bits in the coordinates indexed by $\overline{Q}$.

If the joint distribution of the bits indexed by $\overline{Q}$ is (t,γ)-wise independent then X is said to be a (q,t,γ)-non-oblivious bit-fixing source.

1.1 Our Results

In this paper, we improve the output length of the two-source extractor in [CZ16] to $k^{\Omega(1)}$.

Theorem 1.6. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$, there exists a polynomial time computable function $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ with $m = k^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$ satisfying the following: if X, Y are two independent (n,k) sources, then*

$$|(\mathsf{2Ext}(X,Y), Y) - (U_m, Y)| \leq \epsilon.$$

Since the extractor is strong in Y, if we don't need a strong two-source extractor, then we can use the output of $\mathsf{2Ext}$ to extract from Y and output almost all the min-entropy. For example, by using a strong seeded extractor from [RRV02] that uses $O(\log^2 n \log k)$ bits to extract all the min-entropy (and requiring that say $m \geq \log^3 n$), we have the following theorem.

Theorem 1.7. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$, there exists a polynomial time computable function $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ with $m = k$ and $\epsilon = n^{-\Omega(1)}$ satisfying the following: if X, Y are two independent (n, k) sources, then*

$$|\mathsf{2Ext}(X, Y) - U_m| \leq \epsilon.$$

Next we give an affine extractor over $\mathbb{F}_2$ that works for entropy $k \geq \text{polylog}(n)$, thus significantly improving all previous results in terms of the entropy requirement (even in the disperser case). Our extractor outputs $k^{\Omega(1)}$ bit and has error $n^{-\Omega(1)}$. Specifically, we have

Theorem 1.8. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$, there exists a polynomial time computable function $\mathsf{AExt} : \{0,1\}^n \to \{0,1\}^m$ with $m = k^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$ satisfying the following: for any (n, k) affine source X, we have that*

$$|\mathsf{AExt}(X) - U_m| \leq \epsilon.$$

Table 1 summarizes our result compared to previous constructions of extractors and dispersers for affine sources over $\mathbb{F}_2$.

Construction	Entropy	Output	Error
[Bou07]	$k \geq \delta n$, any constant δ	$\Theta(n)$	$2^{-\Omega(n)}$
[Yeh11, Li11b]	$k = \Omega(n/\sqrt{\log \log n})$	$n^{\Omega(1)}$	$2^{-n^{\Omega(1)}}$
[BSK09]	$k = \Omega(n^{4/5})$	1	Disperser
[Sha11]	$k \geq 2^{\log^{0.9} n}$	1	Disperser
This work	$k \geq \log^C n$ for some large constant C.	$k^{\Omega(1)}$	$n^{-\Omega(1)}$

Table 1: **Summary of Affine Extractors and Dispersers over $\mathbb{F}_2$.**

By using a reduction from NC^0 and AC^0 sources to affine sources in [Vio11], we also obtain improved extractors for NC^0 and AC^0 sources, which only require min-entropy $n^{1/2+\Omega(1)}$.

Theorem 1.9. *For any constant $\alpha > 0, d = O(1)$ and any $n, k \in \mathbb{N}$ with $k \geq n^{1/2+\alpha}$, there is an explicit extractor $\mathsf{acExt} : \{0,1\}^n \to \{0,1\}^m$ with $m = k^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$ such that if X is an (n, k) source generated by a depth-d AC^0 circuit of size n^d, then*

$$|\mathsf{acExt}(X) - U_m| \leq \epsilon.$$

All the above extractors are based on an extractor for (q, t, γ)-non-oblivious bit-fixing sources. In particular, we have the following theorem which improves the output length of such an extractor in [CZ16] from one bit to $t^{\Omega(1)}$.

Theorem 1.10. *There exists a constant $c > 0$ such that for any constant $\delta > 0$ and all $n, q, t \in \mathbb{N}$ with $q \leq n^{1-\delta}$, $t \geq c \log^{21} n$ and any $\gamma \leq 1/n^{t+1}$, there exists an explicit extractor $\mathsf{BFExt} : \{0,1\}^n \to \{0,1\}^m$ with $m = t^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$ such that for any (q, t, γ) non-oblivious bit-fixing source X on n bits, we have that*

$$|\mathsf{BFExt}(X) - U_m| \leq \epsilon.$$

Subsequent work. In a subsequent work, Meka [Mek15] constructed explicit resilient functions matching the randomized construction by Ajtai and Linial [AL93]. His construction also gives improved two-source extractors which require min-entropy $\log^C n$ for a smaller constant C. Similarly, his construction can also be used to obtain affine extractors for entropy $\log^C n$ with a smaller constant C. In another subsequent work by Chattopadhyay and the author [CL16], the techniques of this paper were used to construct explicit extractors for a new model of weak sources called *sumset sources*, which also give improved extractors for sources generated by small space computation.

1.2 Zero-error dispersers and applications

We now extend our results to the case of zero-error dispersers. We first have the following definition.

Definition 1.11. (Zero-error dispersers and strongly hitting dispersers) [GS08] Let $\mathcal{C}$ be a class of distributions on a finite set Ω. A function $D : \Omega \to \{0,1\}^m$ is a disperser for $\mathcal{C}$ with entropy threshold k and error ϵ, if for any weak source $X \in \mathcal{C}$ with entropy at least k, we have

$$|\mathsf{Supp}(D(X))| \geq (1-\epsilon)2^m.$$

The function D is called a zero-error disperser if $\epsilon = 0$. It is called a μ-strongly hitting disperser if in addition for every $z \in \{0,1\}^m$, $\Pr[D(X) = z] \geq \mu$.

Naturally, we consider both independent sources and affine sources. Zero-error independent source dispersers are intimately connected to explicit constructions of Ramsey graphs [BRSW06, CZ16, Coh16], and zero-error affine dispersers for sub-linear entropy give explicit Boolean functions with the best known general circuit lower bounds of $3.01n$ [FGHK15]. Here we will be interested in zero-error dispersers with large output length (say $\Omega(k)$). Previously, such dispersers for independent sources are only known when $k \geq \delta n$ (for two sources) or $k \geq n^\delta$ (for $O(1/\delta)$ sources) [GS08]. The dispersers of [CZ16, Coh16], which work for $k \geq \log^C n$, only output one bit. For affine sources, such dispersers are only known when $k = \Omega(n/\sqrt{\log\log n})$. In this case Gabizon and Shaltiel [GS12] achieved output length $k - \beta n/\sqrt{\log\log n}$ for some constant $0 < \beta < 1$. All other known zero-error dispersers, such as the one by Ben-Sasson and Kopparty [BSK09] which works for entropy $\Omega(n^{4/5})$, and the one by Shaltiel [Sha11] which works for entropy $2^{\log^{0.9} n}$, only output one bit.

Note that if the error of an extractor is small enough compared to its output length, then the extractor is automatically a strongly hitting disperser and zero error disperser. For example, we have the following simple fact.

Fact 1.12. *Let $f : \Omega \to \{0,1\}^m$ be an extractor for a class $\mathcal{C}$ of sources with entropy k and error $\epsilon \leq 2^{-(m+1)}$. Then f is a $2^{-(m+1)}$-strongly hitting disperser for the same class $\mathcal{C}$ and entropy k.*

The error of our two-source extractors and affine extractors can be made n^{-C} for any constant $C > 0$, at the price of slightly increasing the running time of the extractors (but still polynomial in n). Thus we have the following direct corollary.

Corollary 1.13. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$ and any constant $C' > 0$, there exist explicit constructions of strongly $2^{-(m+1)}$-hitting dispersers for two independent (n,k) sources or affine sources on n bits with entropy k, with output length $m = C' \log n$.*

Using a generic technique to increase the output length in [GS08], we obtain the following:

Theorem 1.14. *There exist constants $C > 0, \eta > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$ and $m = \eta k$, there exists an explicit construction of a $2^{-(m+3)}$-strongly hitting disperser $D' : (\{0,1\}^n)^2 \to \{0,1\}^m$ for two independent (n,k) sources.*

Gabizon and Shaltiel [GS08] further showed that such strongly hitting dispersers can be applied to the problem of *implicit probe search*, which, informally speaking, is to search for an element in a table with few probes, while using no additional information beyond the stored elements themselves.

Definition 1.15. (Implicit probe search scheme) [GS08]. For integer parameters n, k, q, the implicit probe search problem is as follows: Store a subset $S \subseteq \{0,1\}^n$ of size 2^k in a table T of size 2^k, (where every table entry holds only a single element of S), such that given any $x \in \{0,1\}^n$ we can determine whether $x \in S$ using q queries to T. A solution to this problem is called an implicit q-probe scheme with table size 2^k and domain size 2^n.

The main goal for this problem is to give a scheme that uses as few number of queries as possible (e.g., $O(1)$ queries) and has table size as small as possible. Fiat and Naor [FN93] studied implicit $O(1)$-probe schemes, where the number of queries is a constant that does not depend on n or k. They showed that if n is large enough compared to k, then no such schemes can exist. This improves a previous bound by Yao [Yao81]. They also proved that non-explicitly, such schemes exist as long as $n \leq 2^{\text{poly}(k)}$. In addition, they gave an efficient implicit $O(1)$-probe scheme in the case where $k = \delta n$ for any constant $\delta > 0$. Gabizon and Shaltiel [GS08] showed that zero-error dispersers for a constant number of independent sources can be used to construct implicit $O(1)$-probe schemes, and they achieved $O(1/\delta)$-probe schemes in the case where $k = n^\delta$ for any constant $\delta > 0$. By using our improved zero-error disperser, we obtain the following implicit $O(1)$-probe scheme, which almost matches the non-explicit construction in [FN93].

Theorem 1.16. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k = \log^C n$ there is an efficiently computable implicit $O(1)$-probe scheme with table size 2^k and domain size 2^n.*

Note that the construction crucially relies on a strongly hitting disperser with output length $\Omega(k)$, which is made possible only because our two-source extractor has output length at least $c \log n$ for some constant $c > 1$ and error n^{-c}.

In the case of affine sources, Gabizon and Shaltiel [GS12] also gave a generic method to transform a zero-error affine disperser with $c \log n$ output bits into another zero-error affine disperser which outputs almost all entropy. Combining their transformation with our zero-error disperser, we obtain the following theorem, which improves the entropy requirement and output length of previous zero-error affine dispersers.

Theorem 1.17. *There exists a constant $C > 0$ such that for all $n, k \in \mathbb{N}$ with $k \geq \log^C n$, there is an explicit zero-error affine disperser $D : \{0,1\}^n \to \{0,1\}^m$ for entropy k, where $m = k - \log^C n$.*

Such zero-error affine dispersers can be applied to the following problem of constructing schemes for defective memory with stuck-at and adversarial errors, as shown in [GS12]. The problem was first studied by Kuznetsov and Tsybakov [KT74], where we have a memory of n cells each storing a symbol from some finite alphabet (we focus on the Boolean alphabet here). However, some subset of at most s cells are stuck (e.g., fixed to some particular string u) and cannot be modified. The goal then is to store a string $z \in \{0,1\}^m$ in the memory so that later we can read the memory and retrieve z, even without any information about which of the cells are stuck. Tsybakov [Tsy75] extended this to a more general model where besides the stuck-at errors, an adversary can choose to corrupt few cells after the string is stored. Formally, we have the following definition.

Definition 1.18. [GS12]. An (n, s, e)-stuck-at noisy memory scheme consists of

- a (possibly randomized) encoding function E such that given any $S \subset [n]$ with $|S| \leq s$, $u \in \{0,1\}^{|S|}$ and $z \in \{0,1\}^m$, E returns $x \in \{0,1\}^n$ such that $x|_S = u$, and
- a decoding function $D : \{0,1\}^n \to \{0,1\}^m$ such that for any $x \in \{0,1\}^n$ produced by E with input z (and any inputs S and u as above), and any 'noise vector' $\xi \in \{0,1\}^n$ of hamming weight at most e, $D(x + \xi) = z$.

The rate of the scheme is m/n, and the natural goal here is to tolerate as many errors as possible while making the rate (or equivalently, m) as large as possible. While one can use standard error-correcting codes for this problem, Gabizon and Shaltiel [GS12] showed that by using "invertible" zero-error dispersers, one can do much better. Specifically, they constructed an explicit stuck-at noisy memory scheme that can tolerate $s(n) = pn$ stuck-at errors for $p < 1/2$ and $e(n) = o(\sqrt{n})$ adversarial errors, with $m = n - s(n) - e(n)\log n - O(n/\sqrt{\log\log n})$ (which achieved rate $1 - p - o(1)$ for the first time) and they left open the problem of improving m to $n - s(n) - e(n)\log n - \log^{O(1)} n$.

As shown in [GS12], a longer output of the zero-error affine disperser directly translates into a larger m in the scheme. Using our improved disperser, we obtain the following improved scheme.

Theorem 1.19. *There exists a constant $C > 0$ such that for any $p < 1/2$, $e(n) = o(\sqrt{n})$ and $s(n) \leq pn$, there is an explicit $(n, s(n), e(n))$-stuck-at noisy memory scheme with $m = n - s(n) - e(n)\log n - \log^C n$.*

2 Overview of the constructions

2.1 Improved two-source extractor

Here we give a brief overview of our improved two-source extractor. Since it follows easily from the extractor for non-oblivious bit-fixing sources, we first describe our new extractor for the (q, t, γ) non-oblivious bit-fixing source on n bits with $q \leq n^{1-\delta}$ for any constant $\delta > 0$, and $\gamma \leq 1/n^{t+1}$. Our starting point is the one-bit deterministic extractor for such sources in [CZ16], which we will call BitExt. We note that from the construction of [CZ16], (by setting the parameters appropriately) this function has the following properties. First, it is a depth-4 AC^0 circuit with size $n^{O(1)}$. Second, since it's an extractor, for any (q, t, γ) non-oblivious bit-fixing source X, we have $\mathsf{BitExt}(X)$ is $n^{-\Omega(1)}$-close to uniform. Third, it's a resilient function, in the sense that any coalition of q bits has influence[1] at most $q/n^{1-\frac{\delta}{2}}$, even when the rest of the bits are only (t, γ)-wise independent.

We now describe how to extract more than one bit. One natural idea is to divide the source X into many blocks and then apply BitExt to each block. Indeed this is our first step. In the source X, we denote the "bad bits" by Q, and the "good bits" by $\overline{Q}$. To ensure that no block consists of only bad bits, we will divide X into n^α blocks for some constant $\alpha < \delta$ (it suffices to take $\alpha = \delta/4$). Thus we get $\ell = n^\alpha$ blocks $\{X_i, i \in [\ell]\}$ with each block containing $n' = n^{1-\alpha}$ bits. We now apply BitExt to each block to obtain a bit Y_i. Of course, we will set up the parameters such that BitExt is an extractor for (q, t, γ) non-oblivious bit-fixing source on $n^{1-\alpha}$ bits.

Now consider any block. Our observation is that since each block can contain at most $q \leq n^{1-\delta}$ bits from Q, the coalition of the bad bits in this block still has small influence. In particular, a

[1] Informally, the influence of a set of bits is the probability over the rest of the bits such that the function is not fixed.

simple calculation shows that $q < n'^{1-\frac{3\delta}{4}}$ and thus for each block the influence of the bad bits is bounded by $q/n'^{1-\frac{3\delta}{8}} < n^{-\frac{3\delta}{8}}$. This means that with probability at least $1 - n^{-\frac{3\delta}{8}}$ over the fixing of $X_i \cap \overline{Q}$, we have that Y_i is fixed. Thus, by a simple union bound, with probability at least $1 - n^{\alpha} n^{-\frac{3\delta}{8}} = 1 - n^{-\frac{\delta}{8}}$ over the fixing of $\overline{Q}$, we have that all $\{Y_i, i \in [\ell]\}$ are fixed.

Now consider another distribution X', which has the same distribution as X for the bits in $\overline{Q}$, while the bits in Q are fixed to 0 independent of the bits in $\overline{Q}$. We let $Y'_i, i \in [\ell]$ be the corresponding Y_i's obtained from X' instead of X. By the above argument, with probability at least $1 - n^{-\frac{\delta}{8}}$ over the fixing of $\overline{Q}$, $\{Y_i\}$ and $\{Y'_i\}$ are the same. Thus the joint distribution of $\{Y_i\}$ and $\{Y'_i\}$ are within statistical distance $n^{-\frac{\delta}{8}}$. Moreover, the bits in $\overline{Q}$ are (t, γ)-wise independent and thus they are $n^t\gamma \leq 1/n$-close to a truly t-wise independent distribution. From now on we will treat $\overline{Q}$ as being truly t-wise independent, since this only adds $1/n$ to the final error.

We will now choose a parameter $m = t^{\Omega(1)}$ for the output length. In addition, we take the generating matrix G of an asymptotically good linear binary code with message length m, codeword length $r = O(m)$ and distance $d = \Omega(m)$. It is well known how to construct such codes (and thus the generating matrix) explicitly. Note that G is an $m \times r$ matrix and any codeword can be generated by $w = vG$ for some vector $v \in \{0,1\}^m$, where all operations are in $\mathbb{F}_2$. We choose m so that $r = O(m) \leq \ell$ and now we let $Y = (Y_1, \cdots, Y_r)$ be the random vector in $\mathbb{F}_2^r$ obtained from $\{Y_i, i \in [\ell]\}$. Similarly, we have $Y' = (Y'_1, \cdots, Y'_r)$. The output of our extractor will now be $Z = (Z_1, \cdots, Z_m) = GY$, where all operations are in $\mathbb{F}_2$.

For the analysis let us consider $Z' = (Z'_1, \cdots, Z'_m) = GY'$. We will show that Z' is close to uniform and then it follows that Z is also close to uniform since they are within statistical distance $n^{-\Omega(1)}$ (as they are deterministic functions of Y and Y' respectively). To show this, we will use the XOR lemma. Consider any non-empty subset $S \subseteq [m]$ and $V'_S = \bigoplus_{i \in S} Z'_i$. Note that this is just $(\sum_{i \in S} G_i)Y'$ where G_i stands for the i'th row of G. Note that $\sum_{i \in S} G_i$ is a codeword and thus has at least $d = \Omega(m)$ 1's. On the other hand, it can have at most $r = O(m)$ 1's.

Note that the parity of up to r bits can be computed by a depth-2 AC^0 circuit of size $2^{O(r)} = 2^{O(m)}$. Recall that each input bit Y_i can be computed by a depth-4 AC^0 circuit of size $n^{O(1)}$. Thus we see that each V'_S can be computed by a depth-6 AC^0 circuit of size at most $2^{O(m)} n^{O(1)} = 2^{O(m)}$ if we choose $m > \log n$.[2] Note that all bits in Q are fixed to 0. Thus the inputs of the circuits are only from $\overline{Q}$.

Now our goal is to ensure that V'_S can be fooled by t-wise independent distributions with error $\epsilon = 2^{-m}$. By the results of Braverman [Bra10] and Tal [Tal14], it suffices to take $t = O(\log(2^{O(m)}/\epsilon)^{21}) = O(m^{21})$. Thus we can take $m = \Omega(t^{\frac{1}{21}})$ and it follows that V'_S cannot distinguish between t-wise independent distributions and uniform distribution. On the other hand, if $\overline{Q}$ is the uniform distribution, then V'_S is the XOR of at least $d = \Omega(m)$ independent random variables, with each being $n^{-\Omega(1)}$-close to uniform. Thus in this case V'_S is $(n^{-\Omega(1)})^d = 2^{-\Omega(m \log n)}$-close to uniform. Together this means that V'_S is $2^{-\Omega(m \log n)} + 2^{-m} < 2^{1-m}$ close to uniform. Since this is true for any non-empty subset S, by a standard XOR lemma it now follows that Z' is $2^{-\Omega(m)}$-close to uniform. Adding back the errors we see that Z is $n^{-\Omega(1)}$-close to uniform.

We can now apply the reduction from two independent sources to a non-oblivious bit-fixing source, which was implicit in [Li15] and explicit in [CZ16]. This reduction reduces two independent (n, k) sources to a (q, t, γ)-non-oblivious bit-fixing source with $t = k^{\Omega(1)}$,[3] thus by applying the

[2] We can get rid of the intermediate negation gates with only a constant factor of blow-up in the circuit size, by standard tricks.

[3] The original reduction in [Li15] only gives $q = \Omega(n)$, but a simple modification can also give $q \leq n^{1-\delta}$ for any

above extractor we also get an improved two-source extractor with output length $k^{\Omega(1)}$.

2.2 Affine extractor

On the high level, our construction of affine extractors follows the framework of the recent two-source extractor construction by Chattopadhyay and Zuckerman [CZ16]. Specifically, we will first reduce an affine source to a (q, t, γ)-non-oblivious bit-fixing source, and then apply our improved deterministic extractor for such sources.

We now describe our reduction. We will mainly adapt techniques from previous work on extractors for independent sources. Specifically, by using ideas from alternating extraction (Figure 1) [DP07, DW09, Li13b, Li13a], one of the author's previous work [Li15] obtained a somewhere random source with $N = \text{poly}(n)$ rows from two independent (n, k) sources with $k \geq \text{polylog}(n)$. The somewhere random source is a random matrix, with the additional property that except for a small fraction of "bad" rows, the rest of the rows are almost t-wise independent for $t = k^{\Omega(1)}$ in the sense that any t of these rows are $\gamma = 2^{-k^{\Omega(1)}}$-close to uniform. Thus, these rows (or, say, taking one bit each row) form exactly a (q, t, γ)-non-oblivious bit-fixing source.

Now we need to adapt that construction to affine sources. Of course we now only have one affine source instead of two independent sources. However, due to the special structure of affine sources we can still apply similar ideas as in [Li13a, Li15]. Specifically, we will use a special kind of strong seeded extractors called *linear seeded extractors*. These extractors have the property that for any fixed seed, the output is a linear function of the source. We take such a seeded extractor with seed length $O(\log n)$ and error ϵ, and use every possible seed to extract from the affine source X. This gives us a matrix (or somewhere random source) of $N = \text{poly}(n)$ rows, where each row corresponds to the output of the extractor on a particular seed. A standard argument shows that if X is affine, then at least $1 - 2\epsilon$ fraction of the rows are truly uniform, although they may depend on each other in arbitrary ways. Note that this is even better for our purpose than in the case of two independent general weak random sources, since there we have to use some other ideas to reduce the error, while here the error in the somewhere random source is essentially zero. We further restrict the size of each row, so that the length is much smaller than the entropy of X.

We can now use these rows and the source X itself to do the same alternating extraction protocol as in [Li13a, Li15] to make the "good" rows almost t-wise independent for $t = k^{\Omega(1)}$, with error $\gamma = 2^{-k^{\Omega(1)}}$. To see why alternating extraction works in this case, consider one particular uniform row Y. Note that Y is a linear function of X, so Y is also an affine source. Recall that the length of Y is much smaller than the entropy of X. A standard argument shows that X can be decomposed into $X = A + B$ where both A, B are affine sources, $A = L(Y)$ for some linear bijection L, and B is independent of Y. Thus, to do the alternating extraction, we can first take a small slice of Y to be S_1, and use a linear seeded extractor Ext to compute $R_1 = \mathsf{Ext}(X, S_1)$. Note that $R_1 = \mathsf{Ext}(X, S_1) = \mathsf{Ext}(A, S_1) + \mathsf{Ext}(B, S_1)$. By the property of a strong extractor we know that with high probability over the fixing of S_1, $\mathsf{Ext}(B, S_1)$ is close to uniform (since S_1 is independent of B). Note that S_1 is a deterministic function of A and A is independent of B, thus $R_1 = \mathsf{Ext}(A, S_1) + \mathsf{Ext}(B, S_1)$ is also uniform conditioned on the fixing of S_1.

Next, suppose the length of R_1 is much smaller than the length of Y ,we can then use R_1 and apply Ext back to Y to extract $S_2 = \mathsf{Ext}(Y, R_1)$. The reason is that we can first fix $\mathsf{Ext}(A, S_1)$. Note that we have already fixed S_1 so this is a deterministic function of A (or Y). Therefore

constant $\delta > 0$.

after fixing it, $\mathsf{Ext}(B, S_1)$ is still uniform and independent of Y (since now it is a deterministic function of B), and now $R_1 = \mathsf{Ext}(A, S_1) + \mathsf{Ext}(B, S_1)$ is independent of Y. Since the length of R_1 is small, conditioned on this fixing Y still has a lot of entropy left. Therefore we can now extract $S_2 = \mathsf{Ext}(Y, R_1)$. After this we can further fix $\mathsf{Ext}(B, S_1)$ and thus also R_1. We know that with high probability over this fixing, S_2 is still close to uniform. Moreover conditioned on this fixing, B still has a lot of entropy left, and is still independent of Y. Now S_2 is a deterministic function of Y. Continue doing this, we can see that alternating extraction works as long as we always use a strong linear seeded extractor and keep the size of each R_i, S_i to be small. Intuitively, it's like alternating extraction between the two independent affine sources Y and B. Now we can use similar arguments as in [Li13a] to make the somewhere random source almost t-wise independent.[4]

However, there are a few subtle technical problems we need to deal with. First, when we generalize the above alternating extraction to run for t rows $Y^1, Y^2, \cdots, Y^t$ simultaneously, we will need to consider the concatenation $Y = Y^1 \circ Y^2 \circ \cdots \circ Y^t$ and decompose X into $X = A + B = L(Y) + B$. This ensures that we can condition on the fixing of all the intermediate random variables obtained from $Y^1, Y^2, \cdots, Y^t$ without affecting B. We can do this if we choose the parameters appropriately so that both $t = k^{\Omega(1)}$ and the size of Y^i are small compared to the entropy of X. Thus in the decomposition B still has sufficient entropy. Another subtlety arises in the analysis as follows. The alternating extraction will take some $b < \log n$ rounds, with each round consisting of some $k^{\Omega(1)}$ steps. In each round j, we start the alternating extraction using a random variable Y^{ij} obtained from Y^i, and at the end we obtain a random variable R^{ij} from X. Our goal is to show that these $\{R^{ij}\}$ will gradually become independent of each other, until at the end they become all independent, thus achieving t-wise independent. Towards this, at the end of round j, for each Y^i we need to use R^{ij} to extract $Y^{i(j+1)}$ from Y^i to start the next round. Here we would like to argue that for those $\{R^{\ell j}\}$ that have already become independent of R^{ij}, we can first fix all $\{Y^{\ell(j+1)}\}$ and all the R variables produced in round $j+1$, and $Y^{i(j+1)}$ is still uniform. This ensures that whatever is already independent will remain independent. While this is true in the case of two independent sources, it is no longer true in the case of an affine source. The reason is that, as explained above, when we fix $R = \mathsf{Ext}(A, S) + \mathsf{Ext}(B, S)$, the part of $\mathsf{Ext}(A, S)$ is a function of A (and Y). Thus this fixing may cause $Y^{i(j+1)}$ to lose entropy (note that fixing $\mathsf{Ext}(B, S)$ will not since B is independent of Y). Fortunately, we can get around this by restricting the length of the R variables to be much smaller than the length of $Y^{i(j+1)}$. We note that if we take a seeded extractor with error ϵ, and use a seed that loses ℓ bits of entropy, then the extractor still works with error increased to $2^\ell \epsilon$. Thus by appropriately choosing the parameters (making ℓ small enough compared to the seed length of the seeded extractor), we can still use $Y^{i(j+1)}$ to start the next round of alternating extraction, and the whole construction goes through.

One final point is that the extractor for non-oblivious bit-fixing source in [CZ16], as well as our improved extractor can only handle the case where $q \leq N^{1-\delta}$ for any constant $\delta > 0$. This means that to convert the affine source X into a somewhere random source in the first step, we need to take a strong linear seeded extractor with seed length $O(\log n)$ and error $\epsilon = 1/\mathrm{poly}(n)$, i.e., an extractor with optimal seed length. Previously, such a linear seeded extractor was not known. In this paper we construct such a strong linear seeded extractor by combining the lossless condenser in [GUV09] and another strong linear seeded extractor in [SU05]. We note that the condenser in

[4] We remark that we can also use the flip-flop alternating extraction developed in [Coh15], which may result in an improvement in the constants. However in this paper we do not try to optimize the constant C in our final result where $k \geq \log^C n$.

[GUV09] itself may not be linear, but can be made linear with the same parameters by a careful instantiation, following a result in [CI15]. Thus in this step we can use $O(\log n)$ bits to condense the source into a $(n' = O(k), k)$ source with error $1/\text{poly}(n)$. We then use the linear seeded extractor in [SU05], which has seed length $d = O\left(\log n' + \frac{\log n'}{\log k} \log\left(\frac{1}{\epsilon}\right)\right)$. Note that $n' = O(k)$. Thus if we take $\epsilon = 1/\text{poly}(n)$ we get $d = O(\log n)$. Altogether we get a strong seeded extractor with seed length $O(\log n)$ and error $\epsilon = 1/\text{poly}(n)$. Since both the condenser and the extractor are linear, the combined extractor is also linear.

Organization. The rest of the paper is organized as follows. We give some preliminaries in Section 3. In Section 4 we give our improved extractors for non-oblivious bit-fixing sources and two independent sources. In Section 5 we define alternating extraction, an important ingredient in our affine extractor construction. We present our construction of affine extractors in Section 6 and the improved extractors for circuit sources in Section 7. The extension to zero-error dispersers and applications are presented in Section 8. Finally we conclude with some open problems in Section 9.

3 Preliminaries

We use common notations such as $\circ$ for concatenation and $[n]$ for $\{1, 2, \cdots, n\}$. All logarithms are to the base 2. We often use capital letters for random variables and corresponding small letters for their instantiations.

3.1 Basic Definitions

Definition 3.1 (statistical distance)**.** Let D and F be two distributions on a set S. Their **statistical distance** is

$$|D - F| \stackrel{def}{=} \max_{T \subseteq S}(|D(T) - F(T)|) = \frac{1}{2}\sum_{s \in S} |D(s) - F(s)|$$

If $|D - F| \leq \epsilon$ we say that D is ϵ-*close* to F and write $D \approx_\epsilon F$.

3.2 Influence of variables

Following [CZ16], we define the influence of variables.

Definition 3.2. Let $f : \{0,1\}^n \to \{0,1\}$ be any boolean function on variables $x_1, \cdots, x_n$. The influence of a set $Q \subseteq \{x_1, \cdots, x_n\}$ on f, denoted by $\mathbf{I}_Q(f)$, is defined to be the probability that f is undetermined after fixing the variables outside Q uniformly at random. Further, for any integer q define $\mathbf{I}_q(f) = max_{Q \subseteq \{x_1, \cdots, x_n\}, |Q|=q} \mathbf{I}_Q(f)$.

More generally, let $\mathbf{I}_{Q,D}(f)$ denote the probability that f is undermined when the variables outside Q are fixed by sampling from the distribution D. We define $\mathbf{I}_{Q,t}(f) = max_{D \in D_k} \mathbf{I}_{Q,D}(f)$, where D_t is the rest of all t-wise independent distributions. Similarly, $\mathbf{I}_{Q,t,\gamma}(f) = max_{D \in D_{t,\gamma}} \mathbf{I}_{Q,D}(f)$ where $D_{t,\gamma}$ is the set of all (t,γ)-wise independent distributions. Finally, for any integer q define $\mathbf{I}_{q,t}(f) = max_{Q \subseteq \{x_1, \cdots, x_n\}, |Q|=q} \mathbf{I}_{Q,t}(f)$ and $\mathbf{I}_{q,t,\gamma}(f) = max_{Q \subseteq \{x_1, \cdots, x_n\}, |Q|=q} \mathbf{I}_{Q,t,\gamma}(f)$

3.3 Somewhere Random Sources

Definition 3.3 (Somewhere Random sources)**.** A source $X = (X_1, \cdots, X_t)$ is (r, t) *somewhere-random* (SR-source for short) if each X_i takes values in $\{0,1\}^r$ and there is an i such that X_i is uniformly distributed.

Definition 3.4. An elementary somewhere-k-source is a vector of sources $(X_1, \cdots, X_t)$, such that some X_i is a k-source. A somewhere k-source is a convex combination of elementary somewhere-k-sources.

3.4 Strong Linear Seeded Extractors

We need the following definition and property of a specific kind of extractors.

Definition 3.5. A function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ϵ) *strong seeded extractor* if for every min-entropy k source X,

$$|\mathsf{Ext}(X, R) - (U_m, R)| \leq \epsilon,$$

where U_m is the uniform distribution on m bits and R is the uniform distribution on d bits independent of X. We say that the function is a *linear strong seeded extractor* if the function $\mathsf{Ext}(\cdot, u)$ is a linear function over $\mathsf{GF}(2)$, for every $u \in \{0,1\}^d$.

We have the following simple fact.

Lemma 3.6. *If* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k, ϵ) *strong seeded extractor, then for every* (n, k) *source* X *and every independent* (d, k') *source* R*,*

$$|\mathsf{Ext}(X, R) - (U_m, R)| \leq 2^{d-k'}\epsilon.$$

Proof. Without loss of generality we can assume that R is the uniform distribution over some subset S of size $2^{k'}$. Then for any $r \in S$, we have $\Pr[R = r] = 2^{-k'}$. Thus

$$\begin{aligned}
|\mathsf{Ext}(X, R) - (U_m, R)| &= \sum_{r \in S} 2^{-k'} |\mathsf{Ext}(X, r) - U_m| = \sum_{r \in S} 2^{d-k'} \cdot 2^{-d} |\mathsf{Ext}(X, r) - U_m| \\
&\leq \sum_{r \in \{0,1\}^d} 2^{d-k'} |2^{-d}\mathsf{Ext}(X, r) - 2^{-d}U_m| = 2^{d-k'} |\mathsf{Ext}(X, R) - (U_m, R)|_{R \leftarrow U_d} \\
&\leq 2^{d-k'}\epsilon.
\end{aligned}$$

□

Lemma 3.7 ([Rao09])**.** *Let* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a linear strong seeded extractor for min-entropy* k *with error* $\epsilon < 1/2$*. Let* X *be any affine source with entropy* k*. Then,*

$$\Pr_{u \leftarrow_R U_d}[|\mathsf{Ext}(X, u) - U_m| = 0] \geq 1 - 2\epsilon$$

Definition 3.8. (condenser) A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is an $k \to_\epsilon k'$ condenser if for every X with min-entropy at least k, $C(X, Y)$ is ϵ-close to a distribution with min-entropy k', when Y is uniformly distributed on $\{0,1\}^d$. A condenser is explicit if it is computable in polynomial time. A condenser is called lossless if $k' = k + d$.

3.5 The Structure of Affine Sources

The following lemma is proved in [Li11b].

Lemma 3.9. *(Affine Conditioning). Let X be any affine source on $\{0,1\}^n$. Let $L : \{0,1\}^n \to \{0,1\}^m$ be any linear function. Then there exist independent affine sources A, B such that:*

- *$X = A + B$.*
- *For every $b \in \mathsf{Supp}(B)$, $L(b) = 0$.*
- *$H(A) = H(L(A))$ and there exists an affine function $L^{-1} : \{0,1\}^m \to \{0,1\}^n$ such that $A = L^{-1}(L(A))$.*

3.6 Previous Work that We Use

We are going to use two constructions of linear seeded extractors in this paper. The first one is for the purpose of obtaining small error. For this we use Trevisan's extractor:

Theorem 3.10 ([Tre01, RRV02])**.** *For every $n, k, m \in \mathbb{N}$ and $\epsilon > 0$ such that $m \leq k \leq n$, there is an explicit (k, ϵ) strong seeded extractor $\mathsf{TrExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O\left(\frac{\log^2(n/\epsilon)}{\log(k/m)}\right)$.*

This extractor is actually a linear seeded extractor. By setting the parameters appropriately, we get the following corollary.

Corollary 3.11 ([Tre01, RRV02])**.** *For every $n, k \in \mathbb{N}$ and $\epsilon > 0$ such that $k \leq n$, there is an explicit (k, ϵ) strong linear seeded extractor $\mathsf{TrExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\Omega(k)}$ with $d = O(\log^2(n/\epsilon))$.*

The next one is for the purpose of obtaining a short seed (i.e., $O(\log n)$). For this we need the following extractor.

Theorem 3.12 ([SU05])**.** *For every $n \in \mathbb{N}$, constant $\delta > 0$, $\epsilon \geq 2^{-k^{\delta/4}}$, and $k \geq \log^{4/\delta} n$ there is an explicit (k, ϵ) strong linear seeded extractor $\mathsf{SUExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O\left(\log n + \frac{\log n}{\log k} \log\left(\frac{1}{\epsilon}\right)\right)$ and $m = k^{1-\delta}$.*

We also note that the lossless condenser in [GUV09] can be made linear.

Theorem 3.13 ([CI15])**.** *For any constant $\alpha > 0$ and any $n \in \mathbb{N}, k \leq n, \epsilon > 0$ there is an explicit strong (k, ϵ)-lossless condenser $\mathsf{Cond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d \leq (1 + 1/\alpha)(\log(nk/\epsilon) + O(1))$ and $m \leq (1 + \alpha)k$. Moreover, Cond is a linear function for every fixed choice of the seed.*

We now have the following theorem.

Theorem 3.14. *There exists a constant $c > 1$ such that for every $n, k \in \mathbb{N}$ with $c \log^8 n \leq k \leq n$, and $\epsilon \geq n^{-2}$, there is an explicit (k, ϵ) strong linear seeded extractor $\mathsf{LExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n)$ and $m = \sqrt{k}$.*

Proof. Given any (n, k) source X, we first take $O(\log n)$ bits and use Theorem 3.13 to condense X into an (n', k) source Y with length $n' = O(k)$, and error $\epsilon/2$. We then use Theorem 3.12 to extract $m = \sqrt{k}$ bits from Y with error $\epsilon/2$ (i.e., take $\delta = 1/2$ in Theorem 3.12). One can check that the conditions of that theorem are satisfied. This will use another $O\left(\log n' + \frac{\log n'}{\log k} \log\left(\frac{2}{\epsilon}\right)\right) = O(\log n)$ bits. Since both the condenser and the extractor are linear and strong, the composed extractor is also a strong linear seeded extractor. □

Lemma 3.15 ([BIW04]). *Assume that $Y_1, Y_2, \cdots, Y_t$ are independent random variables over $\{0,1\}^n$ such that for any $i, 1 \le i \le t$, we have $|Y_i - U_n| \le \epsilon$. Let $Z = \oplus_{i=1}^t Y_i$. Then $|Z - U_n| \le \epsilon^t$.*

To prove our construction is an extractor, we need the following definition and lemma.

Definition 3.16. (ϵ-biased space) A random variable Z over $\{0,1\}$ is ϵ-biased if $|\Pr[Z=0] - \Pr[Z=1]| \le \epsilon$. A sequence of 0-1 random variables $Z_1, \cdots, Z_m$ is *ϵ-biased for linear tests* if for any nonempty set $S \subset \{1, \cdots, m\}$, the random variable $Z_S = \bigoplus_{i \in S} Z_i$ is ϵ-biased.

The following lemma is due to Vazirani. For a proof see for example [Gol95]

Lemma 3.17. *Let $Z_1, \cdots, Z_m$ be 0-1 random variables that are ϵ-biased for linear tests. Then, the distribution of $(Z_1, \cdots, Z_m)$ is $\epsilon \cdot 2^{m/2}$-close to uniform.*

4 The Two-Source Extractor

In this section we give our improved two-source extractor. First, we describe our deterministic extractor for an (q, t, γ)-non-oblivious bit-fixing source on n bits. We rely on the following result from [CZ16].

Theorem 4.1 ([CZ16]). *There exists a constant $c > 0$ such that for any $\delta > 0$ and every large enough $n \in \mathbb{N}$ the following is true. Let X be a (q, t, γ) non-oblivious bit-fixing source on n bits with $q \le n^{1-\delta}$, $t \ge c \log^{18} n$ and $\gamma \le 1/n^{t+1}$. There exists a polynomial time computable monotone boolean function $\mathsf{BitExt} : \{0,1\}^n \to \{0,1\}$ satisfying:*

- BitExt *is a depth 4 circuit in AC^0 of size $n^{O(1)}$.*
- $|\mathsf{E}_{x \leftarrow X}[\mathsf{BitExt}(x)] - \frac{1}{2}| \le \frac{1}{n^{\Omega(1)}}$.
- *For any $q > 0$, $\mathbf{I}_{q,t,\gamma}(\mathsf{BitExt}) \le q/n^{1-\frac{\delta}{2}}$.*

We need the following result by Braverman [Bra10] and Tal [Tal14] about fooling AC^0 circuits with t-wise independent distributions.

Theorem 4.2 ([Bra10, Tal14]). *Let $\mathcal{D}$ be any $t = t(m, d, \epsilon)$-wise independent distribution on $\{0,1\}^n$. Then for any circuit $\mathcal{C} \in \mathsf{AC}^0$ of depth d and size m,*

$$|E_{x \sim U_n}[\mathcal{C}(x)] - E_{x \sim \mathcal{D}}[\mathcal{C}(x)]| \le \epsilon,$$

where $t(m, d, \epsilon) = O(\log(m/\epsilon))^{3d+3}$.

Theorem 4.3 ([AGM03]). *Let $\mathcal{D}$ be a (t, γ)-wise independent distribution on $\{0,1\}^n$. Then there exists a t-wise independent distribution on $\{0,1\}^n$ that is $n^t\gamma$-close to $\mathcal{D}$.*

We also need an explicit asymptotically good binary linear codes:

Definition 4.4. A linear binary code of length n and rank k is a linear subspace C with dimension k of the vector space $\mathbb{F}_2^n$. If the distance of the code C is d we say that C is an $[n, k, d]_2$ code. C is asymptotically good if there exist constants $0 < \delta_1, \delta_2 < 1$ s.t. $k \ge \delta_1 n$ and $d \ge \delta_2 n$.

Note that every linear binary code has an associated generating matrix $G \in \mathbb{F}_2^{k \times n}$, and every codeword can be expressed as vG, for some vector $v \in \mathbb{F}_2^k$.

It is well known that we have explicit constructions of asymptotically good binary linear code. For example, the Justensen codes constructed in [Jus72].

Now we have the following construction and theorem:

We now present our extractor for a (q, t, γ) non-oblivious bit-fixing source.

Algorithm 4.5 ($\mathsf{BFExt}(X)$)**.**

Input: X— a (q, t, γ) non-oblivious bit-fixing source on n bits with $q \leq n^{1-\delta}$, $t \geq c \log^{21} n$ and $\gamma \leq 1/n^{t+1}$.
Output: Z — a string on m bits that is $n^{-\Omega(1)}$ close to uniform, with $m = t^{\Omega(1)}$.

Sub-Routines and Parameters:
Let $\alpha = \delta/4$. Let BitExt be the one-bit extractor for non-oblivious bit-fixing source in Theorem 4.1. Let G be the generating matrix of an asymptotically good $[r, m, d]_2$ code with $r = O(m) \leq n^\alpha$ and $d = \Omega(m)$. Thus G is an $m \times r$ binary matrix.

1. Divide X into $\ell = n^\alpha$ disjoint blocks, each with length $n^{1-\alpha}$.
2. For each block $X_i, i \in [\ell]$, compute $Y_i = \mathsf{BitExt}(X_i)$.
3. Let $Y = (Y_1, \cdots, Y_r)$ be the binary vector in $\mathbb{F}_2^r$. Compute $Z = GY$ where all operations are in $\mathbb{F}_2$.

We have the following theorem.

Theorem 4.6. *There exists a constant c such that for any constant $\delta > 0$ and all $n \in \mathbb{N}$, there exists an explicit extractor $\mathsf{BFExt} : \{0,1\}^n \to \{0,1\}^m$ such that for any (q, t, γ) non-oblivious bit-fixing source X on n bits with $q \leq n^{1-\delta}$, $t \geq c \log^{21} n$ and $\gamma \leq 1/n^{t+1}$, we have that*

$$|\mathsf{BFExt}(X) - U_m| \leq \epsilon,$$

where $m = t^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$.

Proof. Let the set of "bad" bits in X be Q, and the rest of the "good" bits be $\overline{Q}$. Thus $|Q| = q \leq n^{1-\delta}$. Therefore, any block X_i forms a (q, t, γ) non-oblivious bit-fixing source on $n' = n^{1-\alpha} = n^{1-\frac{\delta}{4}}$ bits.

Note that $q \leq n^{1-\delta} < n^{(1-\frac{\delta}{4})(1-\frac{3\delta}{4})} = n'^{1-\frac{3\delta}{4}}$. Thus by Theorem 4.1 we have that each Y_i is $n^{\Omega(1)}$-close to uniform. Moreover since $q \leq n^{1-\delta}$ we have that $\mathbf{I}_{q,t,\gamma}(\mathsf{BitExt}) \leq q/n'^{1-\frac{3\delta}{8}} < n^{-\frac{3\delta}{8}}$.

For any $i \in [\ell]$, the above means that with probability at least $1 - n^{-\frac{3\delta}{8}}$ over the fixing of $X_i \cap \overline{Q}$, we have that Y_i is fixed regardless of what $X_i \cap Q$ is. Thus, it is also true that with probability at least $1 - n^{-\frac{3\delta}{8}}$ over the fixing of $\overline{Q}$, we have that Y_i is fixed regardless of what Q is. By a union bound, with probability at least $1 - n^{-\frac{3\delta}{8}} n^{\frac{\delta}{4}} = 1 - n^{-\frac{\delta}{8}}$ over the fixing of $\overline{Q}$, we have that $(Y_1, \cdots, Y_\ell)$ is fixed and thus $Y = (Y_1, \cdots, Y_r)$ is also fixed.

Now consider a different distribution X' where the bits in $\overline{Q}$ have the same distribution as X, while the bits in Q are fixed to 0 independent of $\overline{Q}$. Let $Y' = (Y'_1, \cdots, Y'_r)$ and Z' be computed from X' using the same algorithm. Then, by the above argument, we have that

$$|Y - Y'| \leq n^{-\frac{\delta}{8}} \text{ and thus also } |Z - Z'| \leq n^{-\frac{\delta}{8}}.$$

Now consider X', Y', Z'. Note that by Theorem 4.3 X' is $n^t\gamma \leq 1/n$-close to a distribution where the bits in $\overline{Q}$ are truly t-wise independent, and the bits in Q are fixed to 0. Thus from now on we will think of X' as this distribution, since this only adds at most $1/n$ to the error.

Let X'' be another distribution where the bits in $\overline{Q}$ are completely uniform and independent, and the bits in Q are fixed to 0. Let Y'', Z'' be the corresponding random variables obtained from X'' instead of X'.

Take any non-empty subset $S \subseteq [m]$, and consider the random variable $V'_S = \bigoplus_{i \in S} Z'_i$.

Note that

$$V'_S = \bigoplus_{i \in S} Z'_i = \bigoplus_{i \in S} G_i Y' = (\sum_{i \in S} G_i) Y',$$

where G_i stands for the i'th row of the matrix G. Since G is the generating matrix of a $[r, m, d]_2$ code, for any non-empty subset $S \subseteq [m]$, we have that $\sum_{i \in S} G_i$ is a codeword. Thus it has at least d 1's.

Now let V''_S be the corresponding random variable obtained from X''. Note that by Theorem 4.1 each Y''_i is $n^{-\Omega(1)}$-close to uniform, and now the $\{Y''_i\}$'s are independent of each other (since they are functions applied to independent blocks of X''). Therefore by Lemma 3.15 we have that

$$|E[V''_S] - 1/2| \leq (n^{-\Omega(1)})^d = 2^{-\Omega(m \log n)}.$$

Moreover, observe that V'_S is the parity of at most $r = O(m)$ Y'_i's. Since parity on r bits can be computed by a depth-2 AC^0 circuit (i.e., a DNF or CNF) of size $2^{O(r)} = 2^{O(m)}$, and every Y'_i is computed by a depth-4 AC^0 circuit with size $n^{O(1)}$, we have that V'_S can be computed by a depth-6 AC^0 circuit with size at most $2^{O(m)}\text{poly}(n)$.

We choose $m = min\{n^{0.9\alpha}, \beta t^{\frac{1}{21}}\}$ for some small constant $0 < \beta < 1$, so that $m = t^{\Omega(1)}$ (since $t \leq n$) and $r = O(m) \leq n^\alpha$. Note that now the AC^0 circuit size is at most $2^{O(m)}\text{poly}(n) = 2^{O(m)}$ since $t^{\frac{1}{21}} = \Omega(\log n)$. Note that the bits in Q are fixed to 0, thus V'_S is computed by a depth-6 AC^0 circuit with inputs from $\overline{Q}$.

Setting $\epsilon = 2^{-m}$ in Theorem 4.2, we see that to ϵ-fool a depth-6 AC^0 circuit with size at most $2^{O(m)}$, it suffices to take $O(\log(2^{O(m)}))^{21} = O(m^{21})$-wise independent distributions. By setting β to be small enough, we can make this number less than t. Since in X', the bits in $\overline{Q}$ are t-wise independent, we have that

$$|E[V'_S] - E[V''_S]| \leq 2^{-m}.$$

Thus

$$|E[V'_S] - 1/2| \leq 2^{-\Omega(m \log n)} + 2^{-m} < 2^{1-m}.$$

Note that this holds for every non-empty subset $S \subseteq [m]$. Thus Z is ϵ'-biased for linear tests with $\epsilon' < 2 \cdot 2^{1-m} = 2^{2-m}$. By the Lemma 3.17 we have that

$$|Z' - U_m| \leq 2^{m/2} 2^{2-m} = 2^{-\Omega(m)}.$$

Adding back the errors, we have

$$|Z - U_m| \leq 2^{-\Omega(m)} + 1/n + n^{-\frac{\delta}{8}} = n^{-\Omega(1)}.$$

■

The following theorem is implicit in [Li15] and explicit in [CZ16]

Theorem 4.7 ([Li15, CZ16])**.** *There exist constants $\delta, c' > 0$ such that for every $n, t \in \mathbb{N}$ there exists a polynomial time computable function* $\mathsf{reduce} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^N$ *with $N = n^{O(1)}$ satisfying the following property: if X, Y are two independent (n, k) sources with $k \geq c' t^4 \log^2 n$, then*

$$\Pr_{y \sim Y}[\mathsf{reduce}(X, y) \text{ is a } (q, t, \gamma) \text{ non-oblivious bit-fixing source }] \geq 1 - n^{-\omega(1)},$$

where $q = N^{1-\delta}$ and $\gamma = 1/N^{t+1}$.

Together with Theorem 4.6 this immediately implies the following theorem.

Theorem 4.8. *There exists a constant $C > 0$ such that for all $n \in \mathbb{N}$, there exists a polynomial time computable function* $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ *satisfying the following: if X, Y are two independent (n, k) sources with $k \geq \log^C n$, then*

$$|(\mathsf{2Ext}(X, Y), Y) - (U_m, Y)| \leq \epsilon,$$

where $m = k^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$.

Proof. We first use Theorem 4.7 to obtain a (q, t, γ) non-oblivious bit-fixing source Z on $N = n^{O(1)}$ bits, with $q = N^{1-\delta}$ and $\gamma = 1/N^{t+1}$. We then apply the extractor for such sources in Theorem 4.6. By choosing C large enough we can ensure that $k \geq c' t^4 \log^2 n$ and $t \geq c \log^{21} n$ (e.g., take $C = 87$). Thus we see that $t = k^{\Omega(1)}$.

Therefore, by Theorem 4.6 the extractor can output $t^{\Omega(1)} = k^{\Omega(1)}$ bits with error $n^{-\Omega(1)}$. Since the reduction succeeds with probability $1 - n^{-\omega(1)}$ over the fixing of Y, the extractor is also strong in Y and the final error is $\epsilon = n^{-\Omega(1)} + n^{-\omega(1)} = n^{-\Omega(1)}$. ■

5 Alternating Extraction

An important ingredient in our construction is the following alternating extraction protocol, which has been used a lot in recent constructions of independent source extractors [Li13b, Li13a]. Here we will use it in the context of affine sources.

Alternating Extraction. Assume that we have two parties, Quentin and Wendy. Quentin has a source Q, Wendy has a source X. Also assume that Quentin has a uniform random seed S_1 (which may be correlated with Q). Let Ext_q and Ext_w be the strong linear seeded extractors in Corollary 3.11. Let ℓ be an integer parameter for the protocol. For some integer parameter $t > 0$, the *alternating extraction protocol* is an interactive process between Quentin and Wendy that runs in t steps.

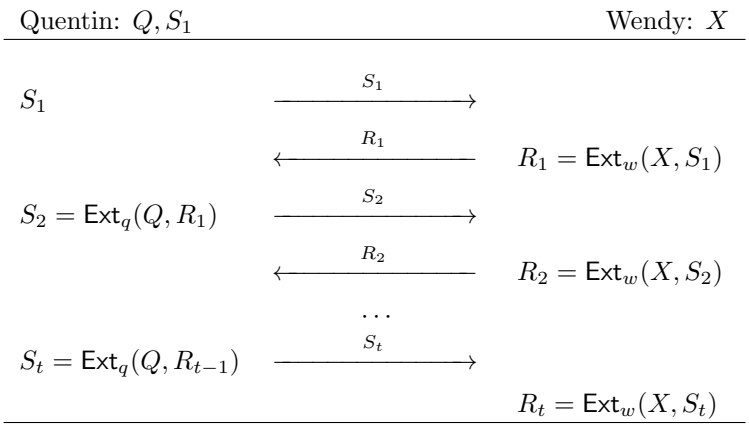

Figure 1: Alternating Extraction.

In the first step, Quentin sends S_1 to Wendy, Wendy computes $R_1 = \mathsf{Ext}_w(X, S_1)$. She sends R_1 to Quentin and Quentin computes $S_2 = \mathsf{Ext}_q(Q, R_1)$. In this step R_1, S_2 each outputs ℓ bits. In each subsequent step i, Quentin sends S_i to Wendy, Wendy computes $R_i = \mathsf{Ext}_w(X, S_i)$. She replies R_i to Quentin and Quentin computes $S_{i+1} = \mathsf{Ext}_q(Q, R_i)$. In step i, R_i, S_{i+1} each outputs ℓ bits. Therefore, this process produces the following sequence:

$$S_1, R_1 = \mathsf{Ext}_w(X, S_1), S_2 = \mathsf{Ext}_q(Q, R_1), \cdots, S_t = \mathsf{Ext}_q(Q, R_{t-1}), R_t = \mathsf{Ext}_w(X, S_t).$$

Look-Ahead Extractor. Now we can define our look-ahead extractor. Let $Y = (Q, S_1)$ be a seed, the look-ahead extractor is defined as

$$\mathsf{laExt}(X, Y) = \mathsf{laExt}(X, (Q, S_1)) \stackrel{def}{=} R_1, \cdots, R_t.$$

We first prove the following lemma.

Lemma 5.1. *Let X be an affine source on n bits, Z be a linear function of X, $Y = (Y_1 = (Q_1, S_1), Y_2 = (Q_2, S_2), \cdots, Y_h = (Q_h, S_h))$ and Y' be linear functions of Z, such that $H(Y'|Y) \geq k_1$, $H(X|Z) \geq k_2$. Assume the following hold: $\forall i$, Q_i has m bits with $m < n$, and S_i has ℓ bits; $H(Q_1) = k_q$ and S_1 is uniform; $k_q \geq 2ht\ell + 10\ell$, $k_1 \geq ht\ell + 10\ell$ and $k_2 \geq ht\ell + 10\ell$. Let Ext_q and Ext_w be strong linear seeded extractors as in Corollary 3.11, set up to use ℓ bits to extract from $(m, 10\ell)$ sources and $(n, 10\ell)$ sources respectively, with error ϵ and $\ell = O(\log^2(n/\epsilon))$. For any $i \in [h]$, let $(R_{i1}, \cdots, R_{it}) = \mathsf{laExt}(X, Y_i)$ and $\{S_{ij}, j = 1, \cdots, t\}$ denote the random variables corresponding to $\{S_j\}$ that are produced when computing $\mathsf{laExt}(X, Y_i)$. For any $j \in [t]$, let $\overline{S_{ij}} = (S_{i1}, \cdots, S_{ij})$ for $i \in [h]$ and $\overline{R_{ij}} = (R_{i1}, \cdots, R_{ij})$ for $i \in [h]$. Then for any $0 \leq j \leq t$, we have that*

$$\begin{aligned}&(R_{1j}, \{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{i(j-1)}}, i \in [h]\}, Y)\\ \approx_{(2j-1)\epsilon} &(U_\ell, \{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{i(j-1)}}, i \in [h]\}, Y).\end{aligned}$$

and

$$(S_{1(j+1)}, \{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\})$$
$$\approx_{(2j)\epsilon} (U_\ell, \{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\}).$$

Moreover, conditioned on $(\{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{i(j-1)}}, i \in [h]\})$*, we have that* X *is still an affine source,* $(\{R_{ij}, i \in [h]\})$ *are deterministic linear functions of* X*,* $H(Q_1) \geq k_q - (2j-1)h\ell$*,* $H(Y'|Y) \geq k_1 - (j-1)h\ell$ *and* $H(X|Z) \geq k_2 - (j-1)h\ell$*; conditioned on* $(\{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\})$*, we have that* X *is still an affine source,* $(\{S_{i(j+1)}, i \in [h]\})$ *are deterministic linear functions of* $\{Y_1, \cdots, Y_h\}$ *respectively,* $H(Q_1) \geq k_q - 2jh\ell$*,* $H(Y'|Y) \geq k_1 - jh\ell$ *and* $H(X|Z) \geq k_2 - jh\ell$*.*

Proof. We prove the lemma by induction on j. When $j = 0$, the statement is trivially true. Now we assume that the statement holds for some j and we prove it for $j+1$.

We first fix $(\{\overline{S_{ij}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\})$. Note that after this fixing, X is still an affine source. Since Y, Y' and Z are linear functions of X, they are still all affine sources as well. Thus by Lemma 3.9, there exist independent affine sources A, B such that $X = A + B$ and there exists a linear bijection L between A and Z. Thus B is also independent of Z. Note that we have $H(X|Z) \geq k_2 - jh\ell \geq 10\ell$, thus $H(B) \geq 10\ell$.

Since $(\{S_{i(j+1)}, i \in [h]\})$ are linear function of Y and Y is a linear function of Z, we have that $(\{S_{i(j+1)}, i \in [h]\})$ are also linear functions of Z; thus B is independent of $(A, \{S_{i(j+1)}, i \in [h]\}, Y)$. If $S_{1(j+1)}$ is uniform, then by Corollary 3.11 we have

$$(\mathsf{Ext}_w(B, S_{1(j+1)}), S_{1(j+1)}) \approx_\epsilon (U_\ell, S_{1(j+1)}).$$

Note that $R_{1(j+1)} = \mathsf{Ext}_w(X, S_{1(j+1)}) = \mathsf{Ext}_w(A, S_{1(j+1)}) + \mathsf{Ext}_w(B, S_{1(j+1)})$ since Ext_w is a linear seeded extractor. Thus for any fixing of $S_{1(j+1)}$, we have that $\mathsf{Ext}_w(B, S_{1(j+1)})$ is a deterministic linear function of B, and is thus independent of $(\mathsf{Ext}_w(A, S_{1(j+1)}), \{S_{i(j+1)}, i \in [h]\}, Y)$. Therefore, we also have that

$$(R_{1(j+1)}, \{S_{i(j+1)}, i \in [h]\}, Y) \approx_\epsilon (U_\ell, \{S_{i(j+1)}, i \in [h]\}, Y).$$

Adding back the error, we have

$$\begin{aligned} &(R_{1(j+1)}, \{\overline{S_{i(j+1)}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\}, Y) \\ \approx_{(2j+1)\epsilon} \quad &(U_\ell, \{\overline{S_{i(j+1)}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\}, Y). \end{aligned} \tag{1}$$

Moreover, note that initially $(\{S_{i(j+1)}, i \in [h]\})$ are deterministic linear functions of $\{Y_1, \cdots, Y_h\}$ respectively. Thus if we further condition on the fixing of $(\{S_{i(j+1)}, i \in [h]\})$ (i.e, conditioned on $(\{\overline{S_{i(j+1)}}, i \in [h]\}, \{\overline{R_{ij}}, i \in [h]\})$), we have that X is still an affine source, and $(\{R_{i(j+1)}, i \in [h]\})$ are deterministic linear functions of X. Furthermore $H(Q_1) \geq k_q - 2jh\ell - h\ell = k_q - (2(j+1)-1)h\ell$. On the other hand $H(Y'|Y)$ and $H(X|Z)$ will remain the same since $(\{S_{i(j+1)}, i \in [h]\})$ are deterministic linear functions of Y. So $H(Y'|Y) \geq k_1 - (j+1-1)h\ell$ and $H(X|Z) \geq k_2 - (j+1-1)h\ell$.

Now, recall that $\forall i$, we have $R_{i(j+1)} = \mathsf{Ext}_w(X, S_{i(j+1)}) = \mathsf{Ext}_w(A, S_{i(j+1)}) + \mathsf{Ext}_w(B, S_{i(j+1)})$. Let $R^A_{i(j+1)} = \mathsf{Ext}_w(A, S_{i(j+1)})$ and $R^B_{i(j+1)} = \mathsf{Ext}_w(B, S_{i(j+1)})$, thus $R_{i(j+1)} = R^A_{i(j+1)} + R^B_{i(j+1)}$. We now fix $(\{S_{i(j+1)}, i \in [h]\})$. Note that we have just shown that conditioned on this further fixing, X is still an affine source; thus Y, Y' and Z are all still affine sources as well. Note that now $R^A_{i(j+1)}$

is a deterministic linear function of A, and $R^B_{i(j+1)}$ is a deterministic linear function of B. Thus $(R^A_{i(j+1)}, i \in [h])$ is independent of $(R^B_{i(j+1)}, i \in [h])$.

We now further fix all $R^A_{i(j+1)}$. Note that these are all linear functions of A with size ℓ; thus conditioned on these fixings, we have that X is still an affine source. Since there is a bijection between A and Z, we have that now $H(Q_1) \geq k_q - (2j+1)h\ell - h\ell = k_q - 2(j+1)h\ell \geq 10\ell$. Moreover $H(Y'|Y) \geq k_1 - jh\ell - h\ell = k_1 - (j+1)h\ell$. On the other hand $H(X|Z)$ remains the same since A is a linear function of Z.

Now note that $\forall i$, $R_{i(j+1)} = R^A_{i(j+1)} + R^B_{i(j+1)}$ is a linear function of B, and is thus independent of (A, Z, Y, Y'). If we ignore the error in Equation 1, we know that conditioned on these fixings, $R_{1(j+1)} = R^A_{1(j+1)} + R^B_{1(j+1)}$ is uniform. Therefore by Corollary 3.11 we have

$$(S_{1(j+2)}, R_{1(j+1)}) \approx_\epsilon (U_\ell, R_{1(j+1)}).$$

Note that conditioned on $R_{1(j+1)}$, we have that $S_{1(j+2)}$ is a deterministic linear function of Y_1, and is thus independent of $(B, \{R_{i(j+1)}, i \in [h]\})$. Therefore we also have

$$(S_{1(j+2)}, \{R_{i(j+1)}, i \in [h]\}) \approx_\epsilon (U_\ell, \{R_{i(j+1)}, i \in [h]\}).$$

Adding back the error from Equation 1, we have

$$\begin{aligned}&(S_{1(j+2)}, \{S_{i(j+1)}, i \in [h]\}, \{R_{i(j+1)}, i \in [h]\})\\ \approx_{2(j+1)\epsilon} &(U_\ell, \{S_{i(j+1)}, i \in [h]\}, \{R_{i(j+1)}, i \in [h]\}).\end{aligned}$$

We can now further fix $(R^B_{i(j+1)}, i \in [h])$. Since $(R^A_{i(j+1)}, i \in [h])$ have already been fixed, this will fix $(R_{i(j+1)}, i \in [h])$ and thus we have fixed $(\{\overline{S_{i(j+1)}}, i \in [h]\}, \{\overline{R_{i(j+1)}}, i \in [h]\})$. Since $(R^B_{i(j+1)}, i \in [h])$ are linear functions of B, conditioned on these fixings X is still an affine source. Moreover, this fixing will not affect $H(Q_1)$ or $H(Y'|Y)$ since B is independent of (A, Z, Y, Y'). Thus we have that $H(Q_1) \geq k_q - 2(j+1)h\ell$, $H(Y'|Y) \geq k_1 - (j+1)h\ell$ and $H(X|Z) \geq k_2 - jh\ell - h\ell = k_2 - (j+1)h\ell$.

Note that $j \leq t$, thus the lemma is proved. □

6 The Affine Extractor

In this section we describe our affine extractor. First we have the following algorithm that obtains a somewhere random source.

Algorithm 6.1 ($\mathsf{SR}(X)$)**.**

Input: X— an (n,k)-affine source with $k \geq \text{polylog}(n)$.
Output: W — a source that is close to an SR-source.

Sub-Routines and Parameters:
Let $0 < \alpha < \beta < 1$ be two constants to be chosen later. Let $\ell = k^{\beta}$. Pick an integer h such that $k^{\alpha} \leq h < 2k^{\alpha}$ and $h = 2^l$ for some integer $l > 0$. Let $\mathsf{Ext}_q, \mathsf{Ext}_w$ be strong linear extractors from Corollary 3.11, set up to extract from $((h^2+12)\ell, 10\ell)$ sources and $(n, 10\ell)$ sources respectively, with seed length $\ell > \log^2 n$, error $\epsilon = 2^{-\Omega(\sqrt{\ell})}$ and output length ℓ. These will be used in IaExt. Let $\mathsf{Ext}_1, \mathsf{Ext}_2, \mathsf{Ext}_3$ be strong linear extractors from Corollary 3.11, with parameters as follows.

- Let $d > \log^2 n$ be an integer such that when we use d uniform bits to extract from an (n,k) source as in Corollary 3.11, the error $\epsilon' = 2^{-\Omega(\sqrt{d})}$ satisfies that $2^{(2h^3+h^2+11h)\ell}\epsilon' \leq \epsilon = 2^{-\Omega(\sqrt{\ell})}$. Note that it suffices to take $d = ch^6\ell^2$ for some constant $c > 1$.
- Ext_1 uses d bits to extract from $(n, 10\ell)$ sources, with error ϵ' and output length ℓ.
- Ext_2 uses ℓ bits to extract from $(\sqrt{k}, (4h^2+20)\ell)$ sources, with error $\epsilon = 2^{-\Omega(\sqrt{\ell})}$ and output length $(2h^2+10)\ell$.
- Ext_3 uses ℓ bits to extract from from $(\sqrt{k}, 2d)$ sources, with error $\epsilon = 2^{-\Omega(\sqrt{\ell})}$ and output length d.

1. Let LExt be the linear seeded extractor in Theorem 3.14, which uses $d' = O(\log n)$ bits to extract from an (n,k) source and output $m = \sqrt{k}$ bits with error n^{-2}. Let $N = 2^{d'} = \text{poly}(n)$ and enumerate all possible choices of the seed $r_1, \cdots, r_N$. For every $i \in [N]$ compute $Y^i = \mathsf{LExt}(X, r_i)$.
2. For every $i = 1, \cdots, N$, use X and Y^i to compute W^i as follows.
   (a) Compute the binary expression of $i-1$, which consists of $d' = \log N = O(\log n)$ bits. Divide these bits sequentially from left to right into $b = \lceil \frac{d'}{l} \rceil$ blocks of size l (the last block may have less than l bits, then we add 0s at the end to make it l bits). Now from left to right, for each block $j = 1, \cdots, b$, we obtain an integer $\mathsf{Ind}_{ij} \leq 2^l$ such that the binary expression of $\mathsf{Ind}_{ij} - 1$ is the same as the bits in block j.
   (b) Let Y^{i1} be the first d bits of Y^i. Set $j = 1$. While $j < b$ do the following.
      i. Compute $R_0^{ij} = \mathsf{Ext}_1(X, Y^{ij})$ and $Y'^{ij} = \mathsf{Ext}_2(Y^i, R_0^{ij})$.
      ii. Compute $(R_1^{ij}, \cdots, R_h^{ij}) = \mathsf{IaExt}(X, Y'^{ij})$, where $Q = Y'^{ij}$ and S_1 is the first ℓ bits of Y'^{ij}.
      iii. Compute $Y^{i(j+1)} = \mathsf{Ext}_3(Y^i, R^{ij}_{\mathsf{Ind}_{ij}})$.
      iv. Set $j = j+1$.
   (c) Finally, compute $(R_1^{ib}, \cdots, R_h^{ib}) = \mathsf{IaExt}(X, Y^{ib})$ and set $W^i = R^{ib}_{\mathsf{Ind}_{ib}}$.
3. Let $W = W^1 \circ \cdots \circ W^N$.

We now introduce some notation. For any $i \in [N]$ and $j \in [b]$, we let $Y^{i(\leq j)}$ denote $(Y^{i1}, \cdots, Y^{ij})$ and similarly $Y'^{i(\leq j)}$ denote $(Y'^{i1}, \cdots, Y'^{ij})$; let $R^{i(\leq j)}_{\mathsf{Ind}_{i(\leq j)}}$ denote $(R^{i1}_{\mathsf{Ind}_{i1}}, \cdots, R^{ij}_{\mathsf{Ind}_{ij}})$ and let $f^j(i)$ denote the integer whose binary expression is the concatenation of the binary expression of $i-1$ from block 1 to block j. Recall that for any i, j, when computing $\mathsf{laExt}(X, Y^{ij})$ we will compute $S^{ij}_1, \cdots, S^{ij}_h$ and $R^{ij}_1, \cdots, R^{ij}_h$. Let $\overline{S^{ij}} = (S^{ij}_1, \cdots, S^{ij}_{\max(\mathsf{Ind}_{vj}-1, \mathsf{Ind}_{ij})})$ and similarly $\overline{R^{ij}} = (R^{ij}_1, \cdots, R^{ij}_{\max(\mathsf{Ind}_{vj}-1, \mathsf{Ind}_{ij})})$; let $\overline{S^{i(\leq j)}} = (\overline{S^{i1}}, \cdots, \overline{S^{ij}})$ and similarly $\overline{R^{i(\leq j)}} = (\overline{R^{i1}}, \cdots, \overline{R^{ij}})$. We have the following lemma.

Lemma 6.2. *Assume that $k \geq 9b^2d^2h^2$. Fix any $v \in [N]$ such that Y^v is uniform. Let $T \subset [N]$ be any subset with $|T| = h$ and $v \in T$, and let Z_T be the concatenation of all Y^i where $i \in T$. For any $j \in [b]$, define $T^j_v = \{i \in T : f^j(i) < f^j(v)\}$ and let $T_v = T^b_v$. Let $\tilde{T}^j_v = T_v \setminus T^j_v$. Then for any $j \in [b]$, we have that*

- *At the beginning of iteration j, conditioned on $(\{Y'^{i(\leq j-1)}, Y^{i(\leq j-1)}, \overline{S^{i(\leq j-1)}}, \overline{R^{i(\leq j-1)}}, i \in T_v\})$, we have*
  1. *X is still an affine source.*
  2. *$\{Y^{ij}, i \in T_v\}$ are linear functions of Z_T.*
  3. *$H(Y^v) \geq \sqrt{k} - 2(j-1)dh$, $H(X|Z_T) \geq k - h\sqrt{k} - 2(j-1)h^2\ell$.*
  4. *With probability $1 - \epsilon_{j1}$ over the fixing of $\{Y'^{i(\leq j-1)}, Y^{i(\leq j-1)}, \overline{S^{i(\leq j-1)}}, \overline{R^{i(\leq j-1)}}, i \in T_v\}, \{Y^{ij}, i \in T^{j-1}_v\}$, we have $Y^{vj} = U_d$. Here $\epsilon_{j1} = 4j(h+1)\epsilon$.*
- *At the end of iteration j, we have*

$$\begin{aligned}&(R^{vj}_{\mathsf{Ind}_{vj}}, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, i \in T_v\}, \{\overline{R^{i(\leq j)}}, i \in T^j_v\})\\ \approx_{\epsilon_{j2}} &(U_\ell, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, i \in T_v\}, \{\overline{R^{i(\leq j)}}, i \in T^j_v\}),\end{aligned}$$

  where $\epsilon_{j2} = 4j(h+1)\epsilon + (2h+1)\epsilon$.

Proof. We prove the lemma by induction on j. Note that $T_v = \{i \in T : i < v\}$ and $\tilde{T}^j_v = T_v \setminus T^j_v$ contains all the numbers in T that are less than v but have the same binary expression in the first j blocks.

We first show that properties 1, 2 and 3 in the first part of the statement hold. When $j = 1$ these are trivially true. Now suppose they hold for some j and we'll show that they hold for $j+1$. We first fix $(\{Y'^{i(\leq j-1)}, Y^{i(\leq j-1)}, \overline{S^{i(\leq j-1)}}, \overline{R^{i(\leq j-1)}}, i \in T_v\})$ and we know that conditioned on this fixing, properties 1, 2 and 3 hold. We now further fix $\{Y^{ij}, i \in T_v\}$. Note that conditioned on this fixing, X is still an affine source, and $\{R^{ij}_0, i \in T_v\}$ are deterministic linear functions of X. We then further fix $\{R^{ij}_0, i \in T_v\}$. Note that conditioned on this fixing, X is still an affine source, and $\{Y'^{ij}, i \in T_v\}$ are deterministic linear functions of $\{Y^i, i \in T_v\}$ respectively. We thus further fix $\{Y'^{ij}, i \in T_v\}$. Note that conditioned on this fixing, X is still an affine source. Now for any $i \in T_v$, in the computation of $(R^{ij}_1, \cdots, R^{ij}_h) = \mathsf{laExt}(X, Y'^{ij})$ we can fix $R^{ij}_1, R^{ij}_2, \cdots$ one by one, and we note that conditioned on the fixing of the previous one, the next one is a deterministic linear function of X. Thus we can fix $\{\overline{R^{ij}}, i \in T_v\}$ and after this fixing, X is still an affine source. Note that once $\{\overline{R^{ij}}, i \in T_v\}$ are fixed, $\{\overline{S^{ij}}, i \in T_v\}$ are also fixed since they are functions of $\{\overline{R^{ij}}, i \in T_v\}$ and $\{Y'^{ij}, i \in T_v\}$. Thus we have fixed $(\{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, \overline{R^{i(\leq j)}}, i \in T_v\})$ and conditioned on

this fixing, X is still an affine source. Furthermore now $\{Y^{i(j+1)}, i \in T_v\}$ are linear functions of $\{Y^i, i \in T_v\} = Z_T$.

Next we look at the $H(Y^v)$ and $H(X|Z_T)$. We note that since Y^v and Z_T are linear functions of X, whenever X is an affine source, they are also both affine sources. We can now repeat the argument above, and since each time we are conditioning on some linear functions of X, the conditional entropy will decrease by at most the size of the random variables being conditioned on. We further note that when we condition on $R_t^{ij}, t = 0, \cdots, h$, we may lose entropy in both $H(Y^v)$ and $H(X|Z_T)$; while when we condition on Y^{ij} and $Y'ij$, we only lose entropy in $H(Y^v)$ since they are deterministic functions of Z_T. Note that the total size of $\{R_t^{ij}, t = 0, \cdots, h, i \in T_v\}$ is at most $h(h+1)\ell < 2h^2\ell$, while the total size of $\{Y^{ij}, Y'^{ij}, R_t^{ij}, t = 0, \cdots, h, i \in T_v\}$ is at most $h(d + (2h^2+10)\ell) + h(h+1)\ell < 2dh$. Thus we know that at the beginning of iteration $j+1$, we have $H(Y^v) \geq \sqrt{k} - 2(j-1)dh - 2dh = \sqrt{k} - 2jdh$ and $H(X|Z_T) \geq k - h\sqrt{k} - 2(j-1)h^2\ell - 2h^2\ell = k - h\sqrt{k} - 2jh^2\ell$.

Now we show that property 4 in the first part and the second part hold. Again we use induction. When $j = 1$, we note that $Y^{v1} = U_\ell$. Thus property 4 holds. Now suppose property 4 holds for some j; we will show that the second part of the statement is true for iteration j, and that property 4 holds for iteration $j+1$. This will establish the lemma.

We first note that by our choice of the parameters, even if conditioned on all the $(\{Y'^{i(\leq b)}, Y^{i(\leq b)}, \overline{S^{i(\leq b)}}, \overline{R^{i(\leq b)}}, i \in T_v\})$, we have that $H(Y^v) \geq \sqrt{k} - 2bdh \geq bdh$ and $H(X|Z_T) \geq k - h\sqrt{k} - 2bh^2\ell > 0.9k$. Thus no matter when we apply Ext_1, Ext_2 or Ext_3 in the case of $i = v$, the source always has enough entropy for extraction.

We now first fix $(\{Y'^{i(\leq j-1)}, Y^{i(\leq j-1)}, \overline{S^{i(\leq j-1)}}, \overline{R^{i(\leq j-1)}}, i \in T_v\})$. Note that conditioned on this fixing, X is still an affine source. Note that Z_T is a linear function of X and $\{Y^{ij}, i \in T_v\}$ are linear functions of Z_T, thus they are all affine sources. We will now further fix $\{Y^{ij}, i \in T_v^{j-1}\}$. Note that since they are all linear functions of Z_T, conditioned on this fixing X is still an affine source. By induction hypothesis, with probability $1 - \epsilon_{j1}$ over the fixing of all these random variables, Y^{vj} is uniform.

Note that now for any $i \in T_v^{j-1}$, we have $R_0^{ij} = \mathsf{Ext}_1(X, Y^{ij})$ is a deterministic linear function of X. We can now further fix all $\{R_0^{ij}, i \in T_v^{j-1}\}$ and conditioned on this fixing, X is still an affine source. After this fixing, for any $i \in T_v^{j-1}$, we have $Y'^{ij} = \mathsf{Ext}_2(Y^i, R_0^{ij})$ is a deterministic function of Y^i. We can now further fix all $\{Y'^{ij}, i \in T_v^{j-1}\}$. Conditioned on this fixing X is still an affine source, and now we have that all the $\{R_t^{ij}, i \in T_v^{j-1}\}$ are deterministic functions of X. Moreover, for any $i \in T_v^{j-1}$, if we fix R_t^{ij}, then R_{t+1}^{ij} is a deterministic linear function of X. Thus for any $i \in T_v^{j-1}$, we can fix all $\{R_t^{ij}, t = 1, \cdots, \max(\mathsf{Ind}_{vj} - 1, \mathsf{Ind}_{ij})\}$ one by one, and conditioned on all these fixings, X is still an affine source. Note that this also fixes all $\{S_t^{ij}, t = 1, \cdots, \max(\mathsf{Ind}_{vj} - 1, \mathsf{Ind}_{ij})\}$ since they are deterministic functions of $\{R_t^{ij}\}$ and $\{Y'^{ij}\}$. Thus we have fixed all $\{Y^{ij}, Y'^{ij}, \overline{S^{ij}}, \overline{R^{ij}}, i \in T_v^{j-1}\}$.

Note that in the above process, we may lose entropy in Y^{vj} when we fix $\{R_t^{ij}, t = 0, \cdots, \max(\mathsf{Ind}_{vj} - 1, \mathsf{Ind}_{ij}), i \in T_v^{j-1}\}$ and $\{Y'^{ij}, i \in T_v^{j-1}\}$. However, note that the size of all these random variables is at most $|T_v^{j-1}|((h+1)\ell + (2h^2+10)\ell) \leq h((h+1)\ell + (2h^2+10)\ell) = (2h^3 + h^2 + 11h)\ell$, thus $H(Y^{v1}) \geq d - (2h^3 + h^2 + 11h)\ell$.

Now by Lemma 3.9, there exist independent affine sources A, B such that $X = A + B$ and there exists a linear bijection L between A and Z_T. Thus B is also independent of $(Z_T, \{Y_i, i \in T_v\})$. Note that $H(B) = H(X|Z_T) \geq 0.9k$ and $H(Y_v) \geq d - (2h^3 + h^2 + 11h)\ell$. Since Y^{vj} is a deterministic function of Y_v, by Lemma 3.6 we have

$$(\mathsf{Ext}_2(B, Y^{vj}), Y^{vj}) \approx_\epsilon (U_\ell, Y^{vj}).$$

Note that $R_0^{vj} = \mathsf{Ext}_2(A, Y^{vj}) + \mathsf{Ext}_2(B, Y^{vj})$. Let $R^A = \mathsf{Ext}_2(A, Y^{vj})$ and $R^B = \mathsf{Ext}_2(B, Y^{vj})$, thus $R_0^{vj} = R^A + R^B$. We now fix Y^{vj}. Note that after this fixing, R^B is a deterministic linear function of B, and is thus independent of $(A, Y^i, i \in T_v)$. If we ignore the error, then R^B is still uniform. Thus we can further fix all $\{Y^{ij}, i \in \tilde{T}_v^{j-1}\}$ (since they are deterministic linear functions of Y^i) and R^B is still uniform. Note that after this fixing X is still an affine source. Now, all the $\{\mathsf{Ext}_2(A, Y^{ij}), i \in \tilde{T}_v^{j-1}\}$ become deterministic linear functions of A, and are thus independent of R^B. So we can now further fix all $\{\mathsf{Ext}_2(A, Y^{ij}), i \in \tilde{T}_v^{j-1}\}$ and R^B is still uniform and independent of $\{Y^i, i \in T_v\}$. Note that conditioned on this fixing we have $R_0^{vj} = R^A + R^B$ is uniform and independent of all $\{Y^i, i \in \tilde{T}_v^{j-1}\}$. Thus by Corollary 3.11 we have

$$(Y'^{vj}, R_0^{vj}) \approx_\epsilon (U_{(2h^2+10)\ell}, R_0^{vj}).$$

We can now further fix R^B, and conditioned on this fixing Y'^{vj} is a deterministic function of Y^v, and is thus independent of all $\{\mathsf{Ext}(B, Y^{ij}), i \in \tilde{T}_v^{j-1}\}$ since they are deterministic linear functions of B. We can therefore fix all $\{\mathsf{Ext}(B, Y^{ij}), i \in \tilde{T}_v^{j-1}\}$ and Y'^{vj} is still close to uniform. Note since $\{\mathsf{Ext}(A, Y^{ij}), i \in \tilde{T}_v^{j-1}\}$ have been fixed before, now we have fixed all $\{R_0^{ij}, i \in \tilde{T}_v^{j-1}\}$. Now adding back all the error, we get that after all these fixings,

$$Y'^{vj} \approx_{2\epsilon} U_{(2h^2+10)\ell}.$$

Ignoring the error, we can now apply Lemma 5.1 (where $Y = (Y'^{ij}, i \in T_v)$, $Y' = Y^v$ and $Z = Z_T$). Thus conditioned on the further fixing of $\{S_1^{ij}, \cdots, S_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\}$ and $\{R_1^{ij}, \cdots, R_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\}$, we have that X is still an affine source, and $\{S_{\mathsf{Ind}_{vj}}^{ij}, i \in \tilde{T}_v^{j-1}\}$ are deterministic linear functions of $\{Y'^{ij}, i \in \tilde{T}_v^{j-1}\}$ respectively. Moreover $H(X|Z_T) \geq 0.9k > 10\ell$. Therefore again by Lemma 3.9, there exist independent affine sources A, B such that $X = A + B$ and there exists a linear bijection L between A and Z_T. Thus B is also independent of Z_T and $H(B) = H(X|Z_T) > 10\ell$.

Lemma 5.1 also tells us that

$$\begin{aligned}&(S_{\mathsf{Ind}_{vj}}^{vj}, \{S_1^{ij}, \cdots, S_{\mathsf{Ind}_{vj}-1}^{ij}, R_1^{ij}, \cdots, R_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\})\\ \approx_{(2(\mathsf{Ind}_{vj}-1))\epsilon}&(U_\ell, \{S_1^{ij}, \cdots, S_{\mathsf{Ind}_{vj}-1}^{ij}, R_1^{ij}, \cdots, R_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\}).\end{aligned}$$

Ignoring the error, conditioned on the further fixing of $\{S_1^{ij}, \cdots, S_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\}$ and $\{R_1^{ij}, \cdots, R_{\mathsf{Ind}_{vj}-1}^{ij}, i \in \tilde{T}_v^{j-1}\}$, we have that $S_{\mathsf{Ind}_{vj}}^{vj}$ is uniform. Since $S_{\mathsf{Ind}_{vj}}^{vj}$ is a deterministic linear function of Y^v, it is independent of B. Thus by Corollary 3.11 we have

$$(\mathsf{Ext}_w(B, S_{\mathsf{Ind}_{vj}}^{vj}), S_{\mathsf{Ind}_{vj}}^{vj}) \approx_\epsilon (U_\ell, S_{\mathsf{Ind}_{vj}}^{vj}).$$

Note that for any $i \in \tilde{T}_v^{j-1}$, we have $R_{\mathsf{Ind}_{vj}}^{ij} = \mathsf{Ext}_w(X, S_{\mathsf{Ind}_{vj}}^{ij}) = \mathsf{Ext}_w(A, S_{\mathsf{Ind}_{vj}}^{ij}) + \mathsf{Ext}_w(B, S_{\mathsf{Ind}_{vj}}^{ij})$. Thus for any fixing of $S_{\mathsf{Ind}_{vj}}^{vj}$, we have $\mathsf{Ext}_w(B, S_{\mathsf{Ind}_{vj}}^{vj})$ is a deterministic linear function of B, and is thus independent of $(A, Z_T = \{Y^i, i \in T_v\})$. Therefore it is also true that

$$(R^{vj}_{\mathsf{Ind}_{vj}}, \{S^{ij}_{\mathsf{Ind}_{vj}}, i \in \tilde{T}^{j-1}_v\}, \{Y'^{ij}, i \in \tilde{T}^{j-1}_v\}) \approx_\epsilon (U_\ell, \{S^{ij}_{\mathsf{Ind}_{vj}}, i \in \tilde{T}^{j-1}_v\}, \{Y'^{ij}, i \in \tilde{T}^{j-1}_v\}).$$

Now we add back all the error, and notice that we have already fixed all $\{Y'^{i(\leq j-1)}, Y^{i(\leq j-1)}, \overline{S^{i(\leq j-1)}}, \overline{R^{i(\leq j-1)}}, i \in T_v\}$, $\{Y^{ij}, i \in T_v\}$, $\{Y'^{ij}, i \in T^{j-1}_v\}$, $\{\overline{S^{ij}}, \overline{R^{ij}}, i \in T^{j-1}_v\}$. Furthermore notice that for any $i \in T^j_v \setminus T^{j-1}_v$, we must have $\mathsf{Ind}_{ij} < \mathsf{Ind}_{vj}$, thus $\max(\mathsf{Ind}_{vj} - 1, \mathsf{Ind}_{ij}) = \mathsf{Ind}_{vj} - 1$. Therefore for these i when we fix $\{R^{ij}_1, \cdots, R^{ij}_{\mathsf{Ind}_{vj}-1}\}$ we have fixed $\overline{R^{ij}}$. On the other hand for any $i \in T_v$ we have that $\mathsf{Ind}_{ij} \leq \mathsf{Ind}_{vj}$ so if we fix all $S^{ij}_1, \cdots, S^{ij}_{\mathsf{Ind}_{vj}}$ we have fixed $\overline{S^{ij}}$. Thus we have

$$\begin{aligned}&(R^{vj}_{\mathsf{Ind}_{vj}}, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, i \in T_v\}, \{\overline{R^{i(\leq j)}}, i \in T^j_v\})\\ \approx_{\epsilon_{j2}} &(U_\ell, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, i \in T_v\}, \{\overline{R^{i(\leq j)}}, i \in T^j_v\}),\end{aligned}$$

where $\epsilon_{j2} = \epsilon_{j1} + 2\epsilon + (2(\mathsf{Ind}_{vj} - 1))\epsilon + \epsilon \leq \epsilon_{j1} + (2h+1)\epsilon = 4j(h+1)\epsilon + (2h+1)\epsilon$. Thus the second part of the statement for iteration j is true.

Now conditioned on the fixing of $\{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, i \in T_v\}, \{\overline{R^{i(\leq j)}}, i \in T^j_v\}$, we have that $\mathsf{Ext}_w(B, S^{vj}_{\mathsf{Ind}_{vj}})$ is still uniform (ignoring the error) and is a deterministic linear function of B. We will now further fix all $\{\mathsf{Ext}_w(A, S^{ij}_{\mathsf{Ind}_{vj}}), i \in \tilde{T}^{j-1}_v\}$. Since these are all deterministic linear functions of A, conditioned on these fixings we have X is still an affine source, and $\mathsf{Ext}_w(B, S^{vj}_{\mathsf{Ind}_{vj}})$ is still uniform. Now $R^{vj}_{\mathsf{Ind}_{vj}} = \mathsf{Ext}_w(A, S^{vj}_{\mathsf{Ind}_{vj}}) + \mathsf{Ext}_w(B, S^{vj}_{\mathsf{Ind}_{vj}})$ is still uniform and is a deterministic linear function of B, and is thus independent of (Z_T, Y^v). Moreover, all $\{R^{ij}_{\mathsf{Ind}_{ij}}, i \in T^j_v\}$ have been fixed and all $\{R^{ij}_{\mathsf{Ind}_{vj}}, i \in \tilde{T}^j_v\}$ are deterministic linear functions of B. Therefore now all $\{Y^{i(j+1)} = \mathsf{Ext}_3(Y^i, R^{ij}_{\mathsf{Ind}_{ij}}), i \in T^j_v\}$ are deterministic linear functions of $\{Y^i, i \in T^j_v\}$ respectively. We can thus now further fix all these $\{Y^{i(j+1)}, i \in T^j_v\}$ and conditioned on this fixing, X is still an affine source and Y^v still has enough entropy. Now by Corollay 3.11 we have

$$(Y^{v(j+1)}, R^{vj}_{\mathsf{Ind}_{vj}}) \approx_\epsilon (U_d, R^{vj}_{\mathsf{Ind}_{vj}}).$$

Note that conditioned on $R^{vj}_{\mathsf{Ind}_{vj}}$, we have that $Y^{v(j+1)}$ is a deterministic function of Y^v, and is thus independent of all $\{R^{ij}_{\mathsf{Ind}_{vj}}, i \in \tilde{T}^j_v\}$ since they are deterministic functions of B. Note that for any $i \in \tilde{T}^j_v$, we must have $\mathsf{Ind}_{ij} = \mathsf{Ind}_{vj}$ and thus $\max(\mathsf{Ind}_{vj} - 1, \mathsf{Ind}_{ij}) = \mathsf{Ind}_{ij}$. Therefore it is also true that

$$(Y^{v(j+1)}, \{R^{ij}_{\mathsf{Ind}_{ij}}, i \in \tilde{T}^j_v\}) \approx_\epsilon (U_d, \{R^{ij}_{\mathsf{Ind}_{ij}}, i \in \tilde{T}^j_v\}).$$

Therefore we have that

$$\begin{aligned}&(Y^{v(j+1)}, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, \overline{R^{i(\leq j)}}, i \in T_v\}, \{Y^{i(j+1)}, i \in T^j_v\})\\ \approx_{\epsilon'} &(U_\ell, \{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, \overline{R^{i(\leq j)}}, i \in T_v\}, \{Y^{i(j+1)}, i \in T^j_v\}),\end{aligned}$$

where $\epsilon' = 2\epsilon + (2(\mathsf{Ind}_{vj} - 1))\epsilon + \epsilon + \epsilon \leq 2(h+1)\epsilon$. This is conditioned on the event that Y^{vj} is uniform at the beginning of iteration j, which happens with probability $1 - \epsilon_{j1}$. By Lemma 3.7, now

with all but another $2\epsilon' \leq 4(h+1)\epsilon$ probability, conditioned on $\{Y'^{i(\leq j)}, Y^{i(\leq j)}, \overline{S^{i(\leq j)}}, \overline{R^{i(\leq j)}}, i \in T_v\}, \{Y^{i(j+1)}, i \in T_v^j\}$ we have that $Y^{v(j+1)}$ is uniform. Thus property 4 in the first part holds for iteration $j+1$ with $\epsilon_{(j+1)1} = \epsilon_{j1} + 4(h+1)\epsilon = 4(j+1)(h+1)\epsilon$. □

We now have the following lemma.

Lemma 6.3. *Assume that $k \geq 9b^2d^2h^2$ and X is an (n,k)-affine source. Let $N = 2^{d'} = \text{poly}(n)$ and $W = W^1 \circ \cdots \circ W^N = \mathsf{SR}(X)$. Then there exists a subset $S \subset [N]$ with $|S| \geq (1 - \frac{2}{n^2})N$ such that for any subset $S' \subset S$ with $|S'| = h$, we have that*

$$(W^i, i \in S') \approx_\epsilon U_{h\ell},$$

where $\epsilon = 2^{-\Omega(\sqrt{\ell})}$.

Proof. First note that LExt is a strong linear seeded extractor with seed length $d' = O(\log n)$ and error n^{-2}. Thus by Lemma 3.7 there exists a subset $S \subset [N]$ with $|S| \geq (1 - \frac{2}{n^2})N$ such that $\forall i \in S$, we have that Y^i is uniform.

Now consider any subset $S' \subset S$ with $|S'| = h$. We order the elements in S' to be $i_1 < i_2 < \cdots < i_h$. Since $S' \subset S$, for any $j \in [h]$ we have that Y^{i_j} is uniform. We now apply Lemma 6.2 to the set S'. Note that $f^b(i) = i - 1$, thus for any $v \in S'$ we have $S_v'^b = \{i \in S' : i < v\}$. Also note that $W^i = R^{ib}_{\mathsf{Ind}_{ib}}$ for any $i \in [N]$. Thus by Lemma 6.2, for any $j \in [h]$ we have that

$$(W^{i_j}, W^{i_1}, \cdots, W^{i_{j-1}}) \approx_{O(bh2^{-\Omega(\sqrt{\ell})})} (U_\ell, W^{i_1}, \cdots, W^{i_{j-1}}).$$

Thus we have that

$$(W^{i_1}, \cdots, W^{i_h}) \approx_\epsilon U_{h\ell},$$

where $\epsilon = O(bh^2 2^{-\Omega(\sqrt{\ell})}) = 2^{-\Omega(\sqrt{\ell})}$ since $\ell = k^\beta > k^\alpha$, $h < 2k^\alpha$ and $b < \log n = k^{O(1)}$. □

We can now state our affine extractor.

Algorithm 6.4 ($\mathsf{AExt}(X)$)**.**

Input: X— an (n,k)-affine source with $k \geq \text{polylog}(n)$.
Output: Z — a bit that is $n^{-\Omega(1)}$-close to uniform.

Sub-Routines and Parameters:
Let SR be the function in Algorithm 6.1. Let BFExt be the extractor for non-oblivious bit-fixing sources in Theorem 4.6.

1. Let $W = W^1 \circ \cdots \circ W^N = \mathsf{SR}(X)$ where $N = \text{poly}(n)$. Take the first bit of each W^i and let V be the concatenation.
2. Compute $Z = \mathsf{BFExt}(V)$.

We now have the following theorem.

Theorem 6.5. *There exists a constant $C > 1$ such that for any (n, k) affine source X with $k \geq \log^C n$, we have that*

$$|\mathsf{AExt}(X) - U_m| \leq \epsilon,$$

where $m = k^{\Omega(1)}$ and $\epsilon = n^{-\Omega(1)}$.

Proof. By Lemma 6.3, if $k \geq 9b^2d^2h^2$ then there exists a subset $S \subset [N]$ with $|S| \geq (1 - \frac{2}{n^2})N$ such that for any subset $S' \subset S$ with $|S'| = h$, we have that

$$(W^i, i \in S') \approx_{\epsilon'} U_{h\ell},$$

where $\epsilon' = 2^{-\Omega(\sqrt{\ell})}$.

Therefore by definition V is a (q, t, γ)-non-oblivious bit-fixing source with $q = 2N/n^2$, $t = h$ and $\gamma = \epsilon' = 2^{-\Omega(\sqrt{\ell})}$. We now apply Theorem 4.6. Note that $q = 2N/n^2 \leq N^{1-\delta}$ for some constant $\delta > 0$ since $N = \text{poly}(n)$. We also need that $t = h \geq O(\log^{21}(N)) = O(\log^{21} n)$and $\gamma \leq 1/N^{t+1} = 1/n^{O(t)}$.

Note that $b < \log n$, $d = O(h^6\ell^2)$, $\ell = k^\beta$, and $k^\alpha \leq h < 2k^\alpha$. Thus altogether all conditions are satisfied if

$$k \geq ck^{14\alpha+4\beta} \log^2 n, k^\alpha \geq c_1 \log^{21} n, \text{ and } \sqrt{\ell} = k^{\beta/2} \geq c_2 k^\alpha \log n,$$

for some constants c, c_1, c_2.

It is now easy to check that if we take α, β to be small enough with $\alpha < \beta/2$ and $k \geq \log^C n$ for a big enough constant $C > 1$, then all the above conditions are satisfied. Thus by Theorem 4.1 the output of AExt is ϵ-close to uniform with $\epsilon = N^{-\Omega(1)} = n^{-\Omega(1)}$. Note that $N = \text{poly}(n)$ so the extractor can be computed in polynomial time. ■

7 Improved Extractors for Circuit Sources

The following theorems are proved in [Vio11].

Theorem 7.1. *[Vio11] For every $d = O(1), \gamma > 0$, any (n, k) source generated by an ac circuit of depth d and size n^d is $1/n^{\omega(1)}$-close to a convex combination of affine sources with entropy $k^2/n^{1+\gamma}$.*

Using our affine extractor, this immediately implies the following theorem.

Theorem 7.2. *For any constant $\alpha > 0, d = O(1)$ and $k \geq n^{1/2+\alpha}$, there is an explicit extractor $\mathsf{acExt} : \{0,1\}^n \to \{0,1\}^m$ with $m = k^{\Omega(1)}$ such that if X is an (n, k) source generated by a depth-d AC^0 circuit of size n^d, then*

$$|\mathsf{acExt}(X) - U_m| \leq \epsilon,$$

where $\epsilon = n^{-\Omega(1)}$.

Proof. Let $\gamma = \alpha$ and apply Lemma 7.1, we see that X is $1/n^{\omega(1)}$-close to a convex combination of affine sources with entropy $k^2/n^{1+\gamma} = n^\alpha$. Thus we can apply the affine extractor from Theorem 1.8 and output $m = k^{\Omega(1)}$ bits with error $1/n^{\omega(1)} + n^{-\Omega(1)} = n^{-\Omega(1)}$. ■

8 Zero-error dispersers and applications

Gabizon and Shaltiel [GS08] showed that a two-source strongly hitting disperser with output length $m \geq c \log n$ can be transformed into another strongly hitting disperser with output length $m = \Omega(k)$.

Theorem 8.1. *[GS08] There exist constants $\eta > 0, c > 0$ such that for every sufficiently large $k, n \in \mathbb{N}$ and $m = \eta k$, any efficiently computable two-source μ-strongly hitting disperser $D' : (\{0,1\}^n)^2 \to \{0,1\}^{c \log n}$ for entropy threshold k can be transformed into an efficiently computable two-source $\mu 2^{c \log n - m - 2}$-strongly hitting disperser $D' : (\{0,1\}^n)^2 \to \{0,1\}^m$ for entropy threshold $2k$.*

Combined with Corollary 1.13, this immediately Theorem 1.14.

To construct implicit $O(1)$-probe scheme, we note that a key combinatorial object in the construction by Fiat and Naor [FN93] is the following so called *rainbow*.

Definition 8.2. [FN93] A t-sequence over a set U is a sequence of length t without repetitions, of elements in U. An (n, k, t)-rainbow is a coloring of all t-sequences over $\{0,1\}^n$ with 2^k colors such that for any $S \subseteq \{0,1\}^n$ of size 2^k, the t-sequences over S are colored in all colors.

In [FN93], Fiat and Naor showed that rainbows give implicit probe schemes.

Theorem 8.3. *[FN93] Fix any integers n, k with $\log n \leq k \leq n$. Given a polynomial time computable (n, k, t)-rainbow we can construct a polynomial time computable implicit $O(t)$-probe scheme with table size 2^k and domain size 2^n. In particular, when t is constant we get an implicit polynomial time computable $O(1)$-probe scheme.*

Gabizon and Shaltiel [GS08] showed that strongly hitting dispersers for independent sources can be used to construct rainbows.

Theorem 8.4. *[GS08] Let $0 < \eta < 1$ be any constant, and let n, k and t be integers with $\log n \leq k \leq n$. Let $m = \eta k$ and let $G : \{0,1\}^{tn} \to \{0,1\}^m$ be a polynomial time computable $t^2/2^k$-strongly hitting disperser for t independent sources with entropy threshold tk, then there is a polynomial time computable $(n, k, O(t/\eta))$-rainbow.*

Combining the above two theorems with our strongly hitting disperser (Theorem 1.14), and note that we can set $t = 2$, and $m = \eta k$ such that $2^{-(m+3)} \geq 4/2^k$, we immediately obtain Theorem 1.16.

We now turn to zero-error affine dispersers. In [GS12], Gabizon and Shaltiel proved the following theorem.

Theorem 8.5. *[GS12] There exists a constant $c > 1$ such that for every sufficiently large $n, k \geq \log^4 n$ and $m \leq k - \log^4 n$, if there is an explicit and efficiently invertible zero-error affine disperser $D' : \{0,1\}^n \to \{0,1\}^{c \log n}$ for entropy $k - m - O(\log^3 n)$, then there is another explicit and efficiently invertible zero-error affine disperser $D : \{0,1\}^n \to \{0,1\}^m$ for entropy k.*

Here a disperser $D : \{0,1\}^n \to \{0,1\}^m$ is said to be *efficiently invertible* if there is a randomized poly(n) time algorithm which given $z \in \{0,1\}^m$ and the specification of an affine source X, returns $x \in \mathsf{Supp}(X)$ such that $D(x) = z$. In [GS12], Gabizon and Shaltiel showed that affine extractors with output length $m = O(\log n)$ and error $\epsilon \leq 2^{-(m+1)}$ are efficiently invertible. Combined with Corollary 1.13, this immediately implies Theorem 1.17.

In terms of stuck-at noisy memory schemes, Gabizon and Shaltiel [GS12] proved the following theorem.

Theorem 8.6. *[GS12] Let $p < 1/2$ and let $e(n) = o(\sqrt{n})$ and $s(n) \leq pn$. For sufficiently large n, if there is an explicit and efficiently invertible zero-error affine disperser $D : 0,1^n?? \to 0,1^m$ for affine sources with entropy $n - s(n) - e(n)\log n - 1$, then there is an explicit $(n, s(n), e(n))$-stuck-at noisy memory scheme with rate m/n.*

Combined with our improved zero error disperser (Theorem 1.17), this immediately gives Theorem 1.19.

9 Conclusions and Open Problems

Constructing explicit two-source extractors and affine extractors are two related challenging problems. Through a long line of research, the recent breakthrough result of Chattopadhyay and Zuckerman [CZ16] has finally brought us close to the optimal two-source extractor. In this paper we managed to improve the output length of the two-source extractors in [CZ16] from 1 to $k^{\Omega(1)}$ in the strong case, and to k in the non-strong case. We also construct the first explicit affine extractor for poly-logarithmic entropy, thus bringing affine extractors close to optimal. Our affine extractor has output length $k^{\Omega(1)}$. However, in both two-source extractors and affine extractors the error remains $n^{-\Omega(1)}$. The most obvious open problem is to improve the error (say to exponentially small).

This seems challenging and requiring new ideas. Specifically, the current approach is to first reduce the sources to a (q, t, γ) non-oblivious bit-fixing source, and then apply a deterministic extractor for such sources. However, the extractor for non-oblivious bit-fixing sources crucially depends on resilient functions, where the analysis is done by bounding the influence of a coalition of variables. If the non-oblivious bit-fixing source has length $n^{O(1)}$ (to ensure polynomial time computability), then even one bit can have influence $\Omega(\log n/n^{O(1)})$ by the result of Kahn, Kalai and Linial [KKL88]. Therefore we cannot hope to get error $n^{-\omega(1)}$ through this approach alone.

Another related open problem left here is to increase the output length of our affine extractor. Currently our affine extractor only outputs $k^{\Omega(1)}$ bits. We note that we can use a general technique by Shaltiel [Sha06] to try to improve the output length. However for that purpose we need to use a linear seeded extractor with seed length $O(\log n)$, since the error will be increased by a factor of 2^d where d is the seed length. A linear seeded extractor with such seed length can possibly achieve output length $k^{0.9}$ [SU05], but we are not aware of any construction with output length $\Omega(k)$. On the other hand, if we can make the error smaller, then we can afford a larger seed length (such as $O(\log^2 n)$), which is enough to output almost all the entropy.

In the case of NC^0 and AC^0 sources, there is still much room for improvement. Currently it is not known how to extract from sources with min-entropy smaller than $n^{1/2}$, even if the source is generated by an NC^0 circuit where each output bit depends on at most 2 input bits.

Finally, an interesting observation of our work is that actually the bias of the one bit extractor in [CZ16] is not very important (in [CZ16] it has bias $n^{-\Omega(1)}$). Indeed, even if it only has constant bias, after the step of using the generating matrix G, we can see that the XOR of $\Omega(m)$ copies will have bias $2^{-\Omega(m)}$. However, at this moment this observation doesn't seem to help improve the parameters.

References

[AGM03] Noga Alon, Oded Goldreich, and Yishay Mansour. Almost k-wise independence versus k-wise independence. *Inf. Process. Lett.*, 88(3):107–110, 2003.

[AL93] M. Ajtai and N. Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[BIW04] Boaz Barak, R. Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, 2004.

[BOL78] Michael Ben-Or and Nathan Linial. Collective coin flipping. *Randomness and Computation*, 1978.

[Bou05] Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1:1–32, 2005.

[Bou07] Jean Bourgain. On the construction of affine-source extractors. *Geometric and Functional Analysis*, 1:33–57, 2007.

[Bra10] Mark Braverman. Polylogarithmic independence fools ac0 circuits. *Journal of the ACM*, 57(5), 2010.

[BRSW06] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2 source dispersers for $n^{o(1)}$ entropy and Ramsey graphs beating the Frankl-Wilson construction. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[BSK09] Eli Ben-Sasson and Swastik Kopparty. Affine dispersers from subspace polynomials. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, 2009.

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[CI15] Mahdi Cheraghchi and Piotr Indyk. Nearly optimal deterministic algorithm for sparse walsh-hadamard transforms. Technical Report TR15-076, Electronic Colloquium on Computational Complexity, 2015.

[CL16] Eshan Chattopadhyay and Xin Li. Extractors for sumset sources. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, 2016.

[Coh15] Gil Cohen. Local correlation breakers and applications to three-source extractors and mergers. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science*, 2015.

[Coh16] Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved Ramsey graphs. In *STOC*, 2016.

[CZ16] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *STOC*, 2016.

[DG10] Matt DeVos and Ariel Gabizon. Simple affine extractors using dimension expansion. In *Proc. of the 25th CCC*, 2010.

[DKSS09] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan. Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 181–190, 2009.

[DP07] Stefan Dziembowski and Krzysztof Pietrzak. Intrusion-resilient secret sharing. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 227–237, Washington, DC, USA, 2007. IEEE Computer Society.

[DW09] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 601–610, 2009.

[FGHK15] Magnus Gausdal Find, Alexander Golovnev, Edward Hirsch, and Alexander Kulikov. A better-than-$3n$ lower bound for the circuit complexity of an explicit function. Technical Report TR15-166, ECCC, 2015.

[FN93] A. Fiat and M. Naor. Implicit $O(1)$ probe search. *SIAM Journal on Computing*, 22:1–10, 1993.

[Gol95] Oded Goldreich. Three xor-lemmas - an exposition. *Electronic Colloquium on Computational Complexity (ECCC)*, 2(56), 1995.

[GR05] Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.

[GS08] Ariel Gabizon and Ronen Shaltiel. Increasing the output length of zero-error dispersers. In *Random 2008*, 2008.

[GS12] Ariel Gabizon and Ronen Shaltiel. Invertible zero-error dispersers and defective memory with stuck-at errors. In *Randomization, Approximation, and Combinatorial Optimization. Proceedings of RANDOM-APPROX '12*, 2012.

[GUV09] V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes. *Journal of the ACM*, 56:1–34, 2009.

[Jus72] J. Justensen. A class of constructive asymptotically good algebraic codes. *IEEE Trans. Info. Theory.*, 18:652656, 1972.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions (extended abstract). In *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, 1988.

[KT74] A. V. Kuznetsov and B. S. Tsybakov. Coding in a memory with defective cells. *Probl. Peredachi Inf.*, 10, 1974.

[KZ07] Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2007.

[Li11a] Xin Li. Improved constructions of three source extractors. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity*, pages 126–136, 2011.

[Li11b] Xin Li. A new approach to affine extractors and dispersers. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity*, pages 137–147, 2011.

[Li13a] Xin Li. Extractors for a constant number of independent sources with polylogarithmic min-entropy. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science*, pages 100–109, 2013.

[Li13b] Xin Li. New independent source extractors with exponential improvement. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, pages 783–792, 2013.

[Li15] Xin Li. Three source extractors for polylogarithmic min-entropy. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science*, 2015.

[LRVW03] C. J. Lu, Omer Reingold, Salil Vadhan, and Avi Wigderson. Extractors: Optimal up to constant factors. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 602–611, 2003.

[Mek15] Raghu Meka. Explicit resilient functions matching Ajtai-Linial. *CoRR*, abs/1509.00092, 2015.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

[Rao06] Anup Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[Rao09] Anup Rao. Extractors for low-weight affine sources. In *Proc. of the 24th CCC*, 2009.

[Raz05] Ran Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 11–20, 2005.

[RRV02] Ran Raz, Omer Reingold, and Salil Vadhan. Extracting all the randomness and reducing the error in trevisan's extractors. *JCSS*, 65(1):97–128, 2002.

[Sha06] Ronen Shaltiel. How to get more mileage from randomness extractors. In *Proceedings of the 21th Annual IEEE Conference on Computational Complexity*, pages 49–60, 2006.

[Sha11] Ronen Shaltiel. Dispersers for affine sources with sub-polynomial entropy. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, 2011.

[SU05] Ronen Shaltiel and Chris Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *Journal of the ACM*, 52:172–216, 2005.

[Tal14] Avishay Tal. Tight bounds on the fourier spectrum of ac0. Technical Report TR14-174, ECCC: Electronic Colloquium on Computational Complexity, 2014.

[Tre01] Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, pages 860–879, 2001.

[Tsy75] B. S. Tsybakov. Defect and error correction. *Probl. Peredachi Inf.*, 11, 1975.

[TV00] Luca Trevisan and Salil Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000.

[Vio11] Emanuele Viola. Extractors for circuit sources. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, 2011.

[Yao81] A. C.-C. Yao. Should tables be sorted? *Journal of the ACM*, 28:615–628, 1981.

[Yeh11] Amir Yehudayoff. Affine extractors over prime fields. *Combinatorica*, 2011.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il