An Efficient Reduction from Two-Source to Non-malleable Extractors

Achieving Near-Logarithmic Min-entropy

Avraham Ben-Aroya
Tel-Aviv University
The Blavatnik School of Computer Science
Tel-Aviv 69978, Israel

Dean Doron
Tel-Aviv University
The Blavatnik School of Computer Science
Tel-Aviv 69978, Israel
deandoron@mail.tau.ac.il

Amnon Ta-Shma
Tel-Aviv University
The Blavatnik School of Computer Science
Tel-Aviv 69978, Israel
amnon@tau.ac.il

ABSTRACT

The breakthrough result of Chattopadhyay and Zuckerman (2016) gives a reduction from the construction of explicit two-source extractors to the construction of explicit non-malleable extractors. However, even assuming the existence of optimal explicit non-malleable extractors only gives a two-source extractor (or a Ramsey graph) for poly$(\log n)$ entropy, rather than the optimal $O(\log n)$.

In this paper we modify the construction to solve the above barrier. Using the currently best explicit non-malleable extractors we get an explicit bipartite Ramsey graphs for sets of size 2^k, for $k = O(\log n \log\log n)$. Any further improvement in the construction of non-malleable extractors would immediately yield a corresponding two-source extractor.

Intuitively, Chattopadhyay and Zuckerman use an extractor as a sampler, and we observe that one could use a weaker object – a *somewhere-random condenser* with a small entropy gap and a very short seed. We also show how to explicitly construct this weaker object using the error reduction technique of Raz, Reingold and Vadhan (1999), and the constant-degree dispersers of Zuckerman (2006) that also work against extremely small tests.

CCS CONCEPTS

• **Theory of computation → Pseudorandomness and derandomization**; **Expander graphs and randomness extractors**;

KEYWORDS

Two-source extractors, Non-malleable extractors, Condensers, Ramsey graphs

ACM Reference format:
Avraham Ben-Aroya, Dean Doron, and Amnon Ta-Shma. 2017. An Efficient Reduction from Two-Source to Non-malleable Extractors. In *Proceedings of 49th Annual ACM SIGACT Symposium on the Theory of Computing, Montreal, Canada, June 2017 (STOC'17),* 10 pages.
DOI: 10.1145/3055399.3055423

STOC'17, Montreal, Canada

DOI: 10.1145/3055399.3055423

1 INTRODUCTION

A graph G is *K-Ramsey* if it contains no clique or independent set of size K. In one of the first applications of the probabilistic method, Erdös showed [20] that there are $2\log N$-Ramsey graphs over N vertices. Erdös raised the challenge of giving an explicit description of such a graph. A related challenge is that of constructing a *K-Ramsey bipartite graph*, i.e., a bipartite graph with no bipartite clique or bipartite independent set of size K. Any explicit bipartite Ramsey graph can be translated into an explicit (non-bipartite) Ramsey graph with about the same parameters. Erdös' probabilistic argument also shows that there are $2\log N$-bipartite Ramsey graphs (where N is the number of vertices on each side) and the problem is constructing such graphs explicitly.

From a computer science point of view, bipartite Ramsey graphs are equivalent to two-source dispersers outputting one bit. More formally, a function $\mathrm{Disp} : [N] \times [N] \to \{0,1\}$ is a (zero-error) *two-source K-disperser* if for every two sets $A, B \subseteq [N]$ of cardinality at least K, $\mathrm{Disp}(A,B) = \{0,1\}$. Such a disperser gives rise to a K-Ramsey bipartite graph with N vertices on each side.

A stronger notion is that of two-source extractors. A function $\mathrm{Ext} : [N]\times[N] \to \{0,1\}$ is a *two-source k-extractor* [1] if for every two independent distributions A, B over $[N]$ with min-entropy at least k, the bit $\mathrm{Ext}(A,B)$ has small bias. One can see that every two-source k-extractor with any nontrivial bias readily implies a two-source 2^k-disperser, and thus also a bipartite 2^k-Ramsey graph.

Early research [1, 13, 21, 33] culminated in the construction of 2^k-Ramsey graphs over 2^n vertices, for $k \approx 2^{\sqrt{\log n}}$ [22] (see also [3, 5, 25, 34] and [24]). Explicitly constructing good two-source extractors was evidently more challenging. The inner product function gives a simple (and powerful) solution when $k > n/2$ [12]. Bourgain [9, 36] gave a two-source extractor construction for $k = \frac{1}{2} - \alpha$ for some small constant $\alpha > 0$. Raz [38] constructed a two-source extractor that has an unbalanced entropy requirement; the first source should have more than $n/2$ min-entropy, while the second source's min-entropy can be as low as $c \cdot \log n$ (for some constant c).

In a different line of research [7, 8] used the challenge-response mechanism for the construction of K-Ramsey graphs for smaller

[1] Throughout the paper we use uppercase letters, as K, to denote sets' cardinalities, and lowercase letters, as k, to denote the corresponding min-entropy ($K = 2^k$). Under this convention an extractor that operates on n-bit sources corresponds to a graph over $N = 2^n$ vertices.

K, culminating in explicitly constructing 2^k-Ramsey graphs for $k = \text{polylog}(n)$ [14].

Due to the difficulty of constructing good two-source extractors, another research line focused on extracting from multiple sources having low min-entropy, trying to minimize the number of sources needed. This includes [6, 27–29, 37], with the later papers using alternating extraction. Eventually, Chattopadhyay and Zuckerman [11] used non-malleable extractors to give a two-source extractor for $k = \text{polylog}(n)$. We note that the main tool in constructing non-malleable extractors is alternating extraction.

Several improvements on the [11] construction followed, including [31, 32]. Cohen and Schulman [17] observed that all the above constructions assume poly-logarithmic min-entropy, and managed to get the first multi-source construction for $k = (\log n)^{1+o(1)}$. Chattopadhyay and Li [10] reduced the number of sources in such a construction to a constant and Cohen [15] put it on five. The main result in this paper is such a construction with only two sources:

Theorem 1.1. *For every large enough n, there exists an explicit, constant-error, two-source extractor* $2\mathsf{EXT} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *for min-entropy* $k = (\log n)^{1+o(1)}$.

This immediately implies an explicit construction of K-Ramsey bipartite graphs with 2^n vertices on each side, for $K = 2^k$.

The Chattopadhyay and Zuckerman two-source extractor construction is by reduction: [11] show a reduction from the existence of explicit two-source extractors to the the construction of explicit non-malleable extractors. However, as noted by Cohen and Schulman [17], this reduction is not optimal in the sense that even if the explicit non-malleable extractor is optimal and has seed length $O(\log \frac{n}{\varepsilon})$, the resulting two-source graph is not optimal and requires both sources to have $\text{poly}(\log n)$ entropy rather than the optimal $O(\log n)$.

The reduction in this paper solves this bottleneck. Specifically, we show that if one manages to construct *non-malleable extractors* with seed length $f(n,\varepsilon)$ and entropy requirement $O(f(n,\varepsilon))$ then, essentially, this implies a constant-error *two-source extractor* with an entropy requirement of $O(f(n, \frac{1}{\text{poly}(n)}))$. This, in particular, means that if one manages to construct optimal non-malleable extractors with $f(n,\varepsilon) = O(\log \frac{n}{\varepsilon})$ then this implies a two-source extractor with entropy requirement $O(\log n)$.

Indeed, following our work, Cohen [15] and Li [30] independently constructed better explicit non-malleable extractors with the current best construction [30] having $f(n,\varepsilon) = O(\log n + \log \frac{1}{\varepsilon} \log\log \frac{1}{\varepsilon})$. Having these non-malleable extractor constructions, Cohen and Li invoked the reduction presented in this paper and concluded a corresponding two-sources extractor. The two-source extractor one gets using Li's explicit non-malleable extractor has $O(\log n \log\log n)$ required entropy. We stress, however, that the reduction presented in this paper keeps working even if, or perhaps when, optimal non-malleable extractors would be obtained. Thus, the main contribution of this paper is constructing this optimal reduction, translating any good non-malleable extractor to a corresponding two-source construction without paying the $\text{poly}(\log n)$ penalty imposed by the [11] construction.

1.1 An Overview of the Construction

Let us first recall the Chattopadhyay-Zuckerman [11] construction. We are given $x_1, x_2 \in [N]$ sampled from two independent distributions X_1 and X_2, with min-entropies k_1 and k_2, respectively. We take a t-non-malleable (k_2, ε_2)-extractor $\mathsf{nmEXT} : [N] \times [D] \to \{0,1\}$ and write a $D \times 1$ table NM, where the rows are indexed by seeds $y \in [D]$, and in row y we write $\mathsf{NM}[y] = \mathsf{nmEXT}(x_2, y)$. By the properties of non-malleable extractors there exists a large subset of the rows such that when we consider the distribution induced by these rows for a random $x_2 \sim X_2$, it is close to being t-wise independent. We deem every row in this subset "good" while the rest are "bad".

At this stage we would have liked to output

$$f(\mathsf{NM}[1], \ldots, \mathsf{NM}[D]),$$

for some *resilient function*[2] f that is willing to accept a few bad players, and good players that are only close to being t-wise independent. Of course, since no one-source extractor exists – there is no such a function f. Nevertheless, Chattopadhyay and Zuckerman explore why this approach fails. Since we are trying to do the impossible (or, rather, understand why the impossible is not possible), in our examination we shall assume the underlying extractor nmEXT has optimal parameters.

Take a t-non-malleable extractor $\mathsf{nmEXT} : [N] \times [D] \to \{0,1\}$ for $t = \text{poly}(\log n)$. We get a table $\mathsf{NM}[y] = \mathsf{nmEXT}(x_2, y)$ with D rows where any t good rows are $O(t\gamma)$-close to uniform for some $\gamma \geq \varepsilon_2$. Also, the number of bad rows is βD for some $\beta \geq \varepsilon_2$. We may choose any β and γ such that $\beta\gamma = \varepsilon_2$ and in particular we may take $\beta = \gamma = \sqrt{\varepsilon_2}$. If we take a non-malleable extractor with seed length dependence $d = O(\log \frac{n}{\varepsilon_2})$ and $\varepsilon_2 \leq \frac{1}{n}$, then $D = \text{poly}(\frac{1}{\varepsilon_2})$ and $q = \beta D = \sqrt{\varepsilon_2} D \leq D^{1-\alpha}$ for some constant $\alpha > 0$. To summarize, we get a table with D rows, at most $q = D^{1-\alpha}$ bad players for some $\alpha > 0$, and every t good players are $t\sqrt{\varepsilon_2}$-close to uniform.

A function f is (q,t)-*resilient* if it is resilient even when there are q bad players and the good players are only t-wise independent. Non-explicitly it is known that there are such functions for $q = D^{1-\alpha}$ bad players out of the D players and $t = \text{polylog}(n)$. In fact, a large part of the [11] paper is devoted to explicitly constructing such a function.

The two preceding paragraphs together imply that the table NM is close to a game with D players, where the good players are t-wise independent and the number of bad players q is at most $D^{1-\alpha}$, and f is a function resilient to such a situation. Hence, it seems, $f(\mathsf{NM}[1], \ldots, \mathsf{NM}[D])$ is close to uniform yielding an impossible one-source extractor.

Yet, there are no one-source extractors, and this is because there is a gap between what we proved about the table NM, and what we require from the (q,t)-resilient function f. Specifically, we proved every t good rows are $t\sqrt{\varepsilon_2}$-close to uniform, but the function f assumes the good rows are *perfectly* t-wise independent. It is true that any distribution over D bits such that every t rows are ζ-close to uniform is $D^t\zeta$-close to a t-wise independent distribution ([4], see

[2] Roughly speaking, a resilient function is a nearly balanced Boolean function $f : \{0,1\}^D \to \{0,1\}$ whose output cannot be heavily influenced by any small set of "bad" bits. We think of the bad bits as a coalition of malicious players trying to bias the output. For a formal definition see Section 2.

Lemma 2.14) but $D^t\zeta \gg 1$ in our case because $D \geq \frac{1}{\varepsilon_2^2} \geq \frac{1}{\zeta^2}$. Thus the impossible does not happen and the one-source construction fails.

Chattopadhyay and Zuckerman use the other source to bypass the above problem. They use X_1 to sample rows from the table NM, i.e., they take a (k_1, ε_1) strong extractor $\mathsf{Ext} : [N] \times [R] \to [D]$ and output

$$2\mathsf{Ext}(x_1, x_2) = f(\mathsf{NM}[\mathsf{Ext}(x_1, 1)], \ldots, \mathsf{NM}[\mathsf{Ext}(x_1, R)])$$
$$= f(\mathsf{nmEXT}(x_2, \mathsf{Ext}(x_1, 1)), \ldots, \mathsf{nmEXT}(x_2, \mathsf{Ext}(x_1, R))).$$

In other words, x_1 samples R rows from the table NM, and these samples are fed into the resilient function. We will soon see how the sample helps solving the problem we had before.

Since extractors are good samplers, if k_1 is large enough, almost all x_1-s sample well. Namely, the fraction of the bad players in the R sampled rows is about $\sqrt{\varepsilon_2} + \varepsilon_1$ and each t good players are $t\sqrt{\varepsilon_2}$-close to uniform. We take $\varepsilon_2 \ll \varepsilon_1$, so we can just think of $\sqrt{\varepsilon_2} + \varepsilon_1$ as ε_1 fraction of bad rows. If we take a small enough ε_1 and an extractor with seed length $O(\log \frac{n}{\varepsilon_1})$ we again get that, with high probability, the sample contains at most $R^{1-\alpha}$ bad players out of the R players. Also, as before, every t good players are $t\sqrt{\varepsilon_2}$-close to uniform. Therefore, again, we may conclude that the good rows in R are $R^t \cdot t\sqrt{\varepsilon_2}$-close to being truly t-wise independent. Now, however, we may choose ε_2 smaller that R^t so than $R^t \cdot t\sqrt{\varepsilon_2} < 1$ and the argument goes through.

In a nutshell, with one source D is a function of ε_2 and $D^t\varepsilon_2$ is necessarily larger than 1; with two sources ε_2 may be chosen way smaller than R^t and the argument magically works!

As beautiful as it is, the argument has its own limitations. We first argue that the number of bad rows in the sampled table is at least $\sqrt{R}$ (out of the R rows in the table). To see this note that samplers are (almost) equivalent to extractors (see, e.g., [23]), and an extractor with error ε_1 has seed length $d_1 \geq 2\log(\frac{1}{\varepsilon_1})$. Thus, the number of rows R is at least $\frac{1}{\varepsilon_1^2}$ and the number of bad rows is at least $\varepsilon_1 R \geq \sqrt{R}$. All the currently known (q, t)-resilient functions that handle $q \geq \sqrt{R}$ bad players require t which is poly-logarithmic in R. The required entropy from X_2 is at least the entropy required by the t-non-malleable extractor nmEXT to output one bit, which is clearly at least t. Altogether, this implies the [11] construction requires $k_1 = k_2 = \text{polylog}(n)$, and this is true even if the non-malleable extractor has optimal seed length $O(\log \frac{n}{\varepsilon_2})$.

Cohen and Schulman [17] note that all previous explicit multi-source extractors (or dispersers) work with entropy at least $\log^2 n$. They were able to construct a multi-source extractor requiring only $(\log n)^{1+o(1)}$ entropy using a new primitive called independence-preserving merger. In a subsequent paper, Cohen [15] shows a five-source extractor for entropy $(\log n)^{1+o(1)}$.[3] The new ingredient that allows this lower entropy requirement is that the resilient function that is used is the Majority function, and Viola [40] showed the majority function is $(q = D^{\frac{1}{2}-\alpha}, t = O(1))$-resilient, i.e., it suffices that the good players are t-wise independent for some constant t! However, to be able to use the Majority function the number q of bad players has to be below square root the total number of players, and, informally speaking, Cohen uses the other four sources to guarantee that the number of bad rows is at most $R^{0.4}$.

The starting point of the current paper is the observation that condensers with a small entropy gap (that we soon define) are good samplers and their dependence on ε can get as small as $1 \cdot \log(\frac{1}{\varepsilon})$. Let us first discuss the entropy gap notion.

Definition 1.2. A function $\mathsf{Cond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to k', \varepsilon)$ condenser if for every (n, k)-source X, $\mathsf{Cond}(X, U_d)$ is ε-close to a distribution with min-entropy k'. The *entropy loss* of the condenser is $k + d - k'$ and the *entropy gap* is $m - k'$.

Dodis et al. [18] observe that for entropy gap 1, non-explicit constructions achieve entropy loss which is only $\log\log(\frac{1}{\varepsilon}) + O(1)$ compared with an entropy loss of $2\log(\frac{1}{\varepsilon})$ when there is no entropy gap. Furthermore, [18] show that condensers with small entropy gap are still good samplers when the test set is small (see Lemma 4.4). Dodis et al. use this property for key derivation without entropy waste. They also show that in non-explicit constructions the seed length dependence on the error is $1 \cdot \log(\frac{1}{\varepsilon})$ rather than the $2\log(\frac{1}{\varepsilon})$ in extractors.

As condensers with a small entropy gap are good samplers against small tests, it is not difficult to see that one can replace the extractor in [11] with such a condenser, and everything stays (almost) the same. Now assume there is an *explicit* construction of a small entropy gap condenser that has optimal dependence of the seed length d on the error ε. Then, we may take $d < 2\log(\frac{1}{\varepsilon})$ and therefore use the Majority function, which implies we can work with a constant t rather than a poly-logarithmic t. Assuming we also have an explicit t non-malleable extractor with optimal seed length $O(\log \frac{n}{\varepsilon})$ we get a two-source extractor requiring $O(\log n)$ as desired.

We see this observation as the main *conceptual* contribution of the paper, namely, that one can replace the sampler in the [11] construction, with a sampler against small tests, and by doing so, at least theoretically, one may reduce the independence requirement to a constant t, and the entropy requirement to the optimal $O(\log n)$. Incidentally, this framework also *simplifies* the proof, as the bulk of the work in [11] is devoted to explicitly constructing an explicit (q, t)-resilient function, which can now be replaced by the Majority function.

Unfortunately, we are not aware of an *explicit* construction achieving a small entropy gap and seed length less than $2\log(\frac{1}{\varepsilon})$. We remark that while the seed length of the condenser of [26] is $\log(\frac{1}{\varepsilon})$, its entropy gap is big.

The main *technical* contribution of the paper is an explicit construction of a *somewhere-random* condenser with a short seed and a small entropy gap, and showing such an object also suffices for the reduction. Together with the pretty good explicit non-malleable extractor constructions we currently have, this gives an explicit two-source extractor with nearly logarithmic entropy requirement.

A somewhere-random condenser is a weaker object than a condenser. The output of a somewhere random condenser is divided into *blocks*, and, roughly speaking, the guarantee is that one of these output blocks is close to having high min-entropy. A condenser is a somewhere-random condenser with just one block.

We construct a somewhere-random condenser with a constant number of blocks, very small entropy gap, and seed length $(1 +$

[3] In fact, four sources suffice for the construction, because the fourth source is redundant as the advice correlation breaker works also with a weak dense seed.

$\alpha)\log(\frac{1}{\varepsilon})$, for any constant $\alpha < 1$. The idea is to start with an extractor that has the wrong dependence on the error, and decrease its error to the desired value ε in an efficient way that gives dependence smaller than $2\log(\frac{1}{\varepsilon})$. We use the error reduction scheme suggested by Raz et al. [39] that works by sampling (with a sampler) a constant number of seeds and outputting all the corresponding outputs of the initial extractor. They show the obtained output is a somewhere-random source with a small entropy gap and a very low error.

The sampler used in [39] is obtained by taking a random walk (of constant length) on an expander. The analysis shows that such a reduction has dependence at least $2\log(\frac{1}{\varepsilon})$ on the error. Instead, we observe that what is needed in the reduction is a disperser against very small tests (i.e., the image of any large enough set is not contained in any small set). The fact that we only need to handle small tests is again crucial, as Zuckerman [41] already constructed such dispersers having a constant degree! Usually, the degree has to be logarithmic, but for the special parameters that we need that reflect the fact we only need to handle very small tests, the degree may be constant. Using these dispersers in the error reduction scheme, we get the desired somewhere random condensers. As a by-product, we obtain a slight simplification and generalization of the [39] error reduction scheme, which we point out in Theorem 3.5.

Armed with that we go back to the [11] construction and replace the extractor Ext with a somewhere-random condenser. As we use a somewhere-random condenser rather than a condenser, when x_1 samples the rows of NM, we get an $R \times A$ table (rather than the $R \times 1$ table we had previously) where A is the constant number of blocks of our somewhere-random condenser. Let us say a row is good if one of its A blocks is good. The property of the table is that the number of bad rows is at most $R^{0.4}$, and the good rows are, informally speaking, t-wise somewhere random. Here we apply another trick from [15]: We take the parity of each row and apply the resilient function on it. The result is an almost balanced bit and we are done.

2 DEFINITIONS AND PRELIMINARIES

Throughout the paper we have the convention that lowercase variables are the logarithm (in base-2) of their corresponding uppercase variables, e.g., $n = \log N, d = \log D, a = \log A, r = \log R, r' = \log R'$, etc. The density of a set $B \subseteq [D]$ is $\rho(B) = \frac{|B|}{D}$. We denote by $[A]$ the set $\{1, \ldots, A\}$.

2.1 Random Variables, Min-Entropy

The *statistical distance* between two distributions X and Y on the same domain D is defined as $|X-Y| = \max_{A\subseteq D}(\Pr[X \in A]-\Pr[Y \in A])$. If $|X - Y| \leq \varepsilon$ we say X is ε-close to Y and denote it by $X \approx_\varepsilon Y$. We will denote by U_n the random variable distributed uniformly over $\{0,1\}^n$. We say a random variable is *flat* if it is uniform over its support.

For a function $f : D_1 \to D_2$ and a random variable X distributed over D_1, $f(X)$ is the random variable distributed over D_2 obtained by choosing x according to X and computing $f(x)$. For a set $A \subseteq D_1$, $f(A) = \{f(x) \mid x \in A\}$. For every $f : D_1 \to D_2$ and two random variables X and Y distributed over D_1, it holds that $|f(X) - f(Y)| \leq |X - Y|$.

The *min-entropy* of a random variable X is defined by

$$H_\infty(X) = \min_{x\in \mathrm{Supp}(X)} \log \frac{1}{\Pr[X = x]}.$$

A random variable X is an (n,k)-source if X is distributed over $\{0,1\}^n$ and has min-entropy at least k. When n is clear from the context we sometimes omit it and simply say that X is a k-source. Every k-source X can be expressed as a convex combination of *flat* distributions each with min-entropy at least k.

For $\varepsilon \geq 0$, the *smooth min-entropy* $H^\varepsilon_\infty(X)$ is the supremum of $H_\infty(X')$ over all distributions $X' \approx_\varepsilon X$. We have the following easy claim:

Claim 2.1. *If $H^{1/2}_\infty(X) \geq k$ then the support of X is of cardinality at least 2^{k-1}.*

2.2 Somewhere-Random Sources

Definition 2.2 (somewhere-random source). A source $X = X_1 \circ \ldots \circ X_A$ is an $(n, k, (\alpha, \beta))$ *somewhere-random (s.r.) source* if there is a random variable $I \in \{0, \ldots, A\}$ such that for every $i \in [A]$, $H^\alpha_\infty(X_i|I = i) \geq k$ and $\Pr[I = 0] \leq \beta$. The variable I is called the *indicator* of the source. If $\alpha = \beta = 0$ we say X is a k *s.r. source*. We say X is a (n, k, ζ) *s.r. source* if X is ζ-close to a k s.r. source over $\{0,1\}^n$.

Claim 2.3. *Let X be an $(n, k, (\alpha, \beta))$ s.r. source. Then, X is a $(n, k, \alpha + \beta)$ s.r. source.*

Intuitively, it is often convenient to think of a k s.r. source $X = X_1 \circ \ldots \circ X_A$ as if one of the blocks X_i is having k min-entropy, and the other blocks are arbitrarily correlated with it. Formally, X is a k s.r. source if it is a convex combination of such sources.

2.3 Extractors

Definition 2.4 (extractor). A function $\mathsf{Ext} : [N] \times [D] \to [M]$ is a (k, ε)*-strong extractor* if for every (n, k)-source X, and for Y that is uniform over $[D]$ and independent of X, it holds that $Y \circ \mathsf{Ext}(X, Y) \approx_\varepsilon Y \times U$.

Theorem 2.5 (the GUV extractor, [26]). *There exists a universal constant $c_{\mathrm{GUV}} > 0$ such that the following holds. For all positive integers n, k and $\varepsilon > 0$ there exists an efficiently-computable (k, ε)-strong extractor $\mathsf{Ext} : [N] \times [D] \to [M]$ having seed length $d = c_{\mathrm{GUV}} \log \frac{n}{\varepsilon}$ and $m = \frac{k}{2}$ output bits.*

Definition 2.6 (two-source extractor). A function $\mathsf{2Ext} : [N_1] \times [N_2] \to [M]$ is an $((n_1, k_1), (n_2, k_2), \varepsilon)$*-two-source extractor* if for every two independent sources X_1 and X_2 where X_1 is an (n_1, k_1)-source and X_2 is an (n_2, k_2)-source, it holds that $\mathsf{2Ext}(X_1, X_2) \approx_\varepsilon U_m$.

2.4 Dispersers and S.R. Condensers

Definition 2.7 (disperser). A function $\Gamma : [N] \times [D] \to [M]$ is a (K, K')*-disperser* if for every $A \subseteq [N]$ with $|A| \geq K$, it holds that $\left|\bigcup_{i\in[D]} \Gamma(A, i)\right| \geq K'$.

Definition 2.8 (s.r. condenser, condenser). A function $\mathsf{SRC} : [N] \times [D] \times [A] \to [M]$ is a $(k \to k', \zeta)$ *s.r. condenser* if for every (n, k)-source X it holds that $\mathsf{SRC}(X, U_d) = \mathsf{SRC}(X, U_d, 1) \circ \ldots \circ \mathsf{SRC}(X, U_d, A)$ is a (m, k', ζ) s.r. source. D is the *degree* of the s.r.

condenser, and A is its number of blocks. If $D = 1$ (i.e., $d = 0$) we say SRC is seedless. A condenser is a s.r. condenser with just one block.

A s.r. condenser implies a disperser with a large error (see, e.g., [41]), i.e., where $K' \ll M$. Concretely,

Lemma 2.9. *Suppose* $\mathrm{SRC} : [N] \times [D] \times [A] \to [M]$ *is a* $(k \to k', \zeta = \frac{1}{2})$ *s.r. condenser. Define* $\Gamma : [N] \times [D \cdot A] \to [M]$ *by* $\Gamma(x;(y,a)) = \mathrm{SRC}(x,y,a)$. *Then,* Γ *is a* $(K, \frac{K'}{2})$*-disperser.*

Proof. Let $B \subseteq [N]$ be an arbitrary set such that $|B| \geq K = 2^k$. Then, $\mathrm{SRC}(B, U_d) = \mathrm{SRC}(B, U_d, 1) \circ \ldots \circ \mathrm{SRC}(B, U_d, A)$ is $\zeta = 1/2$-close to a k' s.r. source, with some indicator random variable I. Pick any index $i \neq 0$ in the support of I. Then, conditioned on $I = i$, we have that $\mathrm{SRC}(B, U_d, i)$ is $1/2$-close to a k' source. Thus, by Claim 2.1, conditioned on $I = i$, $\mathrm{SRC}(B, U_d, i)$ covers at least $2^{k'-1}$ vertices from $[M]$. But then, even without the conditioning, $\mathrm{SRC}(B, U_d, i)$ covers at least $2^{k'-1} = K'/2$ vertices from $[M]$. □

Zuckerman [41], using additive number theory and extending earlier results, showed:

Theorem 2.10 ([41], Theorem 8.3). *There exist two constants* $0 < c_1 < c_2 < 1$, *a constant* $\gamma > 0$ *and a* $(c_1 n \to c_2 m, N^{-\gamma})$ *s.r. seedless condenser with just two blocks*

$$\mathrm{SRC} : [N] \times [1] \times [A = 2] \to [M = N^{2/3}].$$

With that Zuckerman constructs dispersers with very large error, but constant degree. Specifically,

Theorem 2.11 ([41], Theorem 1.9). *Fix any constants* $0 < c_1, c_2 < 1$. *Set* $K = N^{c_1}$, $M \leq K^{1-c_2}$ *and* $K' < M$. *Then there exists an efficient family of* (K, K')*-dispersers*

$$\Gamma : [N] \times [D] \to [M]$$

with degree $D = O\left(\frac{\log \frac{N}{K}}{\log \frac{M}{K'}}\right) = O\left(\frac{n}{\log \frac{M}{K'}}\right)$.

We will use the following property of Zuckerman's disperser, which is straightforward from the construction in [41]:

Claim 2.12. *Let* $\Gamma : [N] \times [D] \to [M]$ *be the disperser of Theorem 2.11. Then, for every* $i \in [D]$ *we have that* $\Gamma(U_n, i) = U_m$.

2.5 Limited Independence and Non-Oblivious Bit-Fixing Sources

Definition 2.13. A distribution X over $\{0,1\}^n$ is called (t,γ)*-wise independent* if the restriction of X to every t coordinates is γ-close to U_t.

Lemma 2.14 ([4]). *Let* $X = X_1, \ldots, X_n$ *be a distribution over* $\{0,1\}^n$ *that is* (t,γ)*-wise independent. Then,* X *is* $(n^t\gamma)$*-close to a* t*-wise independent distribution.*

Definition 2.15. A source X over $\{0,1\}^n$ is called a (q,t,γ) *non-oblivious bit-fixing source* if there exists a subset $Q \subseteq [n]$ of size at most q such that the joint distribution of the bits in $[n] \setminus Q$ is (t,γ)-wise independent. The bits in Q are allowed to arbitrarily depend on the bits in $[n] \setminus Q$.

Definition 2.16. Let $f : \{0,1\}^n \to \{0,1\}$, $\mathcal{D}$ a distribution over $\{0,1\}^n$ and $Q \subseteq [n]$. Let $I_{Q,\mathcal{D}}(f)$ denote the probability that f is undetermined when the variables outside Q are sampled from $\mathcal{D}$. We define $I_{q,t,\gamma}(f)$ to be the maximum of $I_{Q,\mathcal{D}}(f)$ over all $Q \subseteq [n]$ of size q and all $\mathcal{D}$ that is a (t,γ)-independent distribution.

We say that f is (t,γ)*-independent* (q,ε)*-resilient* if $I_{q,t,\gamma}(f) \leq \varepsilon$.

Balanced resilient functions are deterministic extractors against non-oblivious bit-fixing sources outputting one bit. Chattopadhyay and Zuckerman [11] derandomized the Ajtai-Linial function [2] and constructed a (monotone) resilient function that handles $q = n^{1-\alpha}$ for any constant α. Their construction was later improved in [32].

Following [15, 17] we work hard to be able to use the majority function as the resilient function. The work of Viola [40] shows that for every $\alpha > 0$, the majority function over n bits is $(t,0)$-independent $(n^{1/2-\alpha}, O(\frac{\log t}{t} + n^{-\alpha}))$-resilient. Combining this with Lemma 2.14, we conclude:

Lemma 2.17. *There exists a constant* c_{Maj} *such that for every* $\alpha > 0$ *and a* $(q = n^{\frac{1}{2}-\alpha}, t, \gamma)$ *non-oblivious bit-fixing source* X *on* n *bits,*

$$\left|\Pr[\mathrm{Maj}(X_1, \ldots, X_n) = 1] - \frac{1}{2}\right| \leq c_{\mathrm{Maj}} \cdot \left(\frac{\log t}{t} + n^{-\alpha} + \gamma n^t\right).$$

2.6 Non-Malleable Extractors

Definition 2.18. A function $\mathsf{nmEXT} : [N] \times [D] \to [M]$ is a (k,ε) *t-non-malleable extractor*, if for every (n,k)-source X, for every Y that is uniform over $[D]$ and every functions $f_1, \ldots, f_t : [D] \to [D]$ with no fixed-points[4] it holds that:

$$(\mathsf{nmEXT}(X,Y), \mathsf{nmEXT}(X, f_1(Y)), \ldots, \mathsf{nmEXT}(X, f_t(Y), Y) \approx_\varepsilon (U_m, \mathsf{nmEXT}(X, f_1(Y)), \ldots, \mathsf{nmEXT}(X, f_t(Y), Y).$$

We will need the following lemma concerning the existence of a set of good seeds of a non-malleable extractor, given in [11].

Lemma 2.19 ([11], Lemma 3.4). *Let* $\mathsf{nmEXT} : [N] \times [D] \to \{0,1\}$ *be a* (k,ε) *t-non-malleable extractor. Let* X *be any* (n,k)*-source. Let* BAD *be the set defined by*

$$\mathrm{BAD} = \{r \in [D] \mid \exists \text{ distinct } r_1, \ldots, r_t \in [D], \forall i \in [t]\ r_i \neq r, |(\mathsf{nmEXT}(X,r), \mathsf{nmEXT}(X,r_1), \ldots, \mathsf{nmEXT}(X,r_t)) - (U_1, \mathsf{nmEXT}(X,r_1), \ldots, \mathsf{nmEXT}(X,r_t))| > \sqrt{\varepsilon}\}. \quad (1)$$

Then, $\rho(\mathrm{BAD}) \leq \sqrt{\varepsilon}$. *In particular,* $R = [D] \setminus \mathrm{BAD}$ *is large,* $|R| \geq (1 - \sqrt{\varepsilon})D$ *and for any* $r_1, \ldots, r_t \in R$,

$$(\mathsf{nmEXT}(X,r_1), \ldots, \mathsf{nmEXT}(X,r_t)) \approx_{5t\sqrt{\varepsilon}} U_t. \quad (2)$$

We remark that the property in Equation (1) is stronger than the one in Equation (2). The second one says rows in R are almost t-wise independent. The first one says every row in R is independent from any other t rows – good or bad. This will be important for us later on.

We also use the following lemma, which is a simple generalization of one given in [11].

[4] That is, for every i and every x, we have $f_i(x) \neq x$.

Lemma 2.20 ([11], Lemma 2.9). *Let $X_1, \ldots, X_t$ and $Y_1, \ldots, Y_k$ be Boolean random variables. Further suppose that for any $i \in [t]$,*

$$\left(X_i, \left\{X_j\right\}_{j\neq i}, Y_1, \ldots, Y_k\right) \approx_\varepsilon \left(U_1, \left\{X_j\right\}_{j\neq i}, Y_1, \ldots, Y_k\right).$$

Then, $(X_1, \ldots, X_t, Y_1, \ldots, Y_k) \approx_{5t\varepsilon} (U_t, Y_1, \ldots, Y_k)$.

Finally, good explicit constructions of t-non-malleable codes exist. Prior to the first version of our paper the best explicit construction was:

Theorem 2.21 ([15], Theorem 12.1, generalized to a constant t). *For any constant t there exist constants $c_1, c_2 \geq 1$ such that the following holds. For any integer n and for any $\varepsilon > 0$, there exists an efficiently-computable $(k = c_2 d, \varepsilon)$ t-non-malleable extractor* $\mathsf{nmExt} : [N] \times [D] \to \{0,1\}$ *with seed length $d = c_1 \log n + \log(1/\varepsilon) \cdot c_1^{\sqrt{\log\log(1/\varepsilon)}} = O(\log n) + (\log(1/\varepsilon))^{1+o(1)}$.*

The current best explicit construction is:

Theorem 2.22 ([30], Theorem 8.7). *There exists a constant c_1 such that for any integer t the following holds. For any integer n and for any $\varepsilon > 0$, there exists an efficiently-computable $(k = d, t\varepsilon)$ t-non-malleable extractor* $\mathsf{nmExt} : [N] \times [D] \to \{0,1\}$ *with seed length $d = c_1 t^2 (\log n + \log\frac{1}{\varepsilon} \cdot \log\log\frac{1}{\varepsilon})$.*

We note that independent of Li's work, Cohen [16] obtained a similar yet slightly weaker result, where the dependence on ε is $\log(\frac{1}{\varepsilon}) \cdot \mathrm{poly}\log\log(\frac{1}{\varepsilon})$.

3 LOW-ERROR S.R. CONDENSERS WITH A SHORT SEED AND A SMALL ENTROPY GAP

In this section we construct s.r. condensers with ε error that have seed length $(1+\alpha)\cdot\log(\frac{1}{\varepsilon})$ and entropy gap $O(\log\frac{1}{\varepsilon})$.

Theorem 3.1. *For every constant $0 < \alpha < 1$, there exists a constant A such that for every $0 < \varepsilon = \varepsilon(n) \leq (\frac{1}{n})^{\frac{4c_{\mathrm{GUV}}}{\alpha}}$ and $m = m(n) \leq \frac{n}{2} - \log(\frac{1}{\varepsilon})$ there exists an explicit* $\mathsf{SRC} : [N] \times [R'] \times [A] \to [M]$ *that is a $(k = 2m + \log(\frac{1}{\varepsilon}) \to m - 2\log(\frac{1}{\varepsilon}) - O(a), \varepsilon)$ s.r. condenser with $r' = (1+\alpha)\log(\frac{1}{\varepsilon})$.*

Notice that the s.r. condenser achieves the small error ε using only a constant number of blocks, a small entropy gap (i.e., the min-entropy in the s.r. source is close to the block length) and seed length close to $\log(\frac{1}{\varepsilon})$.

Proof. Fix $\alpha > 0$.

Construction: The first ingredient is an extractor $\mathsf{Ext} : [N] \times [R] \to [M]$ that has error ε_0 that is too high for us, $\varepsilon_0 = \varepsilon^{1/c}$, but seed length that is still within our budget, say, $\frac{\alpha}{2}\log(\frac{1}{\varepsilon})$. Our goal is to reduce the error to $\varepsilon = \varepsilon_0^c$. Specifically, set $\varepsilon_0 = \varepsilon^{\frac{\alpha}{4c_{\mathrm{GUV}}}}$ and notice that $\varepsilon_0 \leq \frac{1}{n}$. By Theorem 2.5 there exists

$$\mathsf{Ext} : [N] \times [R] \to [M]$$

that is an explicit $(2m, \varepsilon_0)$ strong extractor with $r = c_{\mathrm{GUV}} \log\frac{n}{\varepsilon_0} \leq 2c_{\mathrm{GUV}}\log\frac{1}{\varepsilon_0} = \frac{\alpha}{2}\log\frac{1}{\varepsilon}$.

The second ingredient is a disperser, that we use as a sampler to sample many (dependent) seeds of Ext, one of which is good. Specifically, set $r' = (1+\alpha)\log(\frac{1}{\varepsilon})$ and take

$$\Gamma : [R'] \times [A] \to [R]$$

to be the $\left(K = (R')^{\frac{\alpha}{1+\alpha}}, K' = 2\varepsilon_0 R\right)$-disperser guaranteed by Theorem 2.11, when plugging-in $c_1 = \frac{\alpha}{1+\alpha}$ and $c_2 = \frac{1}{2}$. Notice that $K^{\frac{1}{2}} = (R')^{\frac{1}{2}\frac{\alpha}{1+\alpha}} = (\frac{1}{\varepsilon})^{(1+\alpha)\frac{1}{2}\frac{\alpha}{1+\alpha}} = (\frac{1}{\varepsilon})^{\frac{\alpha}{2}} \geq R$. The degree is then

$$A = O\left(\frac{r'}{\log\frac{1}{2\varepsilon_0}}\right) = O\left(\frac{(1+\alpha)\log\frac{1}{\varepsilon}}{\frac{\alpha}{4c_{\mathrm{GUV}}}\log\frac{1}{\varepsilon} - 1}\right) = O(1).$$

We define $\mathsf{SRC} : [N] \times [R'] \times [A] \to [M]$ by:

$$\mathsf{SRC}(x, y', z) = \mathsf{Ext}(x, \Gamma(y', z)).$$

Correctness: Let X be a k-source for $k = 2m + \log(\frac{1}{\varepsilon})$. W.l.o.g., X is flat (because otherwise it is a convex combination of such sources). Our goal is to prove that $\mathsf{SRC}(X, U_{r'})$ is a s.r. source with $k' = m - 2\log(\frac{1}{\varepsilon}) - O(a)$ min-entropy.

Since Ext is an extractor, the output distribution $\mathsf{Ext}(X, U_r)$ is ε_0-close to uniform. Define the set of Δ-*heavy elements* in $[M]$ by:

$$\mathbf{H} = \left\{w \in [M] \mid \Pr[\mathsf{Ext}(X, U_r) = w] \geq \frac{\Delta}{M}\right\}.$$

We claim:

Claim 3.2. *For $\Delta > 2$, $|\mathbf{H}| < \frac{2\varepsilon_0}{\Delta}M$.*

Proof. Notice that $|\mathbf{H}| \cdot \frac{\Delta}{M} \leq \Pr[\mathsf{Ext}(X, U_r) \in \mathbf{H}] \leq \frac{|\mathbf{H}|}{M} + \varepsilon_0$, where the upper bound follows because $\mathsf{Ext}(X, U_r)$ is ε_0-close to uniform. Thus $\frac{|\mathbf{H}|}{M}(\Delta - 1) \leq \varepsilon_0$ and $|\mathbf{H}| \leq \frac{\varepsilon_0}{\Delta - 1}M < \frac{2\varepsilon_0}{\Delta}M$. □

We will work with

$$\Delta = \frac{4A}{\varepsilon} \geq 4.$$

There are two possible ways $\mathbf{H}$ may get its weight in Ext, and they react differently to amplification. First, there are "typical" elements for which the fraction of seeds falling into $\mathbf{H}$ is about right (the density of $\mathbf{H}$ plus Ext's error which is ε_0). In this case the amplification with the disperser Γ guarantees that we miss $\mathbf{H}$ with one of our samples with good probability.

Moreover, there are "bad" elements $x \in X$ for which many seeds y fall into $\mathbf{H}$. Bad elements do not react well to amplification (e.g., amplification has no effect when all seeds fall into $\mathbf{H}$) but there are very few of them. Formally, we define the set of bad inputs $x \in \mathrm{Supp}(X)$ by

$$\mathrm{BadX} = \left\{x \in [N] \mid \Pr_{y\sim U_r}[\mathsf{Ext}(x, y) \in \mathbf{H}] \geq \left(1 + \frac{2}{\Delta}\right)\varepsilon_0\right\}.$$

Claim 3.3. $|\mathrm{BadX}| < 2^{2m}$.

Proof. Suppose $|\mathrm{BadX}| \geq 2^{2m}$. Let B be uniformly distributed over BadX. Then $\mathsf{Ext}(B, U_r)$ is ε_0-close to uniform, and therefore

$$\left(1 + \frac{2}{\Delta}\right)\varepsilon_0 \leq \Pr_{x\sim B, y\sim U_r}[\mathsf{Ext}(x, y) \in \mathbf{H}] \leq \frac{|\mathbf{H}|}{M} + \varepsilon_0.$$

But then $\frac{2\varepsilon_0}{\Delta} \leq \frac{|\mathbf{H}|}{M}$, which contradicts to the previous claim. □

Now consider a "typical" element, i.e. $x \in \mathrm{Supp}(X) \setminus \mathrm{BadX}$. As $x \notin \mathrm{BadX}$ there are only a few seeds y (about ε_0) such that $\mathsf{Ext}(x, y)$ falls into $\mathbf{H}$. If we sample a constant number of independent seeds, then except for probability $\varepsilon = \varepsilon_0^{O(1)}$, one of them will *not* fall into $\mathbf{H}$, and we get a s.r. source (but with a long random seed). Raz et al. [39] replace the independent samples with a good sampler (a random walk on expanders). We use a different sampler – Zuckerman's

sampler from Theorem 2.11, because we are in the small-test regime in which Zuckerman's sampler has a constant degree.

Formally, let I be a random variable defined as follows. For $x \in [N]$ and $y' \in [R']$, $I(x, y')$ is an arbitrary $z \in [A]$ such that $\mathsf{Ext}(x, \Gamma(y', z)) \notin \mathbf{H}$ if such a z exists, and 0 otherwise. Let I' be the same as I except that for all z with $\Pr[I = z] \leq \frac{4}{\Delta}$, if $I(x, y') = z$ we let $I'(x, y') = 0$.

Claim 3.4.

- $\Pr[I = 0] \leq 2\varepsilon$.
- $\Pr[I' = 0] \leq 3\varepsilon$.
- *For every* $z \in [A]$, $H_\infty(\mathsf{Ext}(X, \Gamma(U_{r'}, I'))|I' = z) \geq m - 2\log\Delta + 2$.

Proof. For the first item,

$$\Pr[I = 0] \leq \Pr[X \in \text{BadX}] + \Pr[I = 0 | X \notin \text{BadX}].$$

- Clearly,

  $$\Pr[X \in \text{BadX}] \leq |\text{BadX}| \cdot 2^{-k} = |\text{BadX}| \cdot 2^{-(2m + \log(\frac{1}{\varepsilon}))} \leq \varepsilon.$$

  Intuitively, we "drown" the bad elements among all the elements in X. Said differently, we increase the entropy requirement to reduce the error.
- Fix an element $x \notin \text{BadX}$ and call $y \in [R]$ *bad for* x if $\mathsf{Ext}(x, y) \in \mathbf{H}$. As $x \notin \text{BadX}$, the number of seeds y that are bad for x is at most $(1 + \frac{2}{\Delta})\varepsilon_0 R \leq 2\varepsilon_0 R$. Notice that $(I = 0 | X = x)$ if $y' \in [R']$ is such that $\Gamma(y', 1), \ldots, \Gamma(y', A)$ are all bad for x. Since Γ is a $\left(K = (R')^{\frac{\alpha}{1+\alpha}}, K' = 2\varepsilon_0 R\right)$-disperser, the number of such y'-s is at most $K = (R')^{\frac{\alpha}{1+\alpha}}$. Hence,

  $$\begin{aligned} \Pr[I = 0 | X \notin \text{BadX}] &\leq \frac{(R')^{\frac{\alpha}{1+\alpha}}}{R'} = (R')^{-(1 - \frac{\alpha}{1+\alpha})} \\ &= (R')^{-(\frac{1}{1+\alpha})} = (\varepsilon^{1+\alpha})^{\frac{1}{1+\alpha}} = \varepsilon, \end{aligned}$$

  as desired.

  For the second item, $\Pr[I' = 0] \leq \Pr[I = 0] + \frac{4A}{\Delta} = 3\varepsilon$.

  For the third item, let w be in the support of

  $$(\mathsf{Ext}(X, \Gamma(U_{r'}, z)) \mid I' = z).$$

Hence $w \notin \mathbf{H}$. It follows that

$$\begin{aligned} \Pr[\mathsf{Ext}(X, \Gamma(U_{r'}, z)) = w | I' = z] &\leq \frac{\Pr[\mathsf{Ext}(X, \Gamma(U_{r'}, z)) = w]}{\Pr[I' = z]} \\ &= \frac{\Pr[\mathsf{Ext}(X, U_r) = w]}{\Pr[I' = z]} \\ &\leq \frac{\Delta}{M} \cdot \frac{1}{\Pr[I' = z]} \leq \frac{\Delta^2}{4M}, \end{aligned}$$

where the equality on the second line follows from Claim 2.12. Thus, for every $z \in [A]$ in the support of I',

$$H_\infty(\mathsf{Ext}(X, \Gamma(U_{r'}, I')) \mid I' = z) \geq m - 2\log\Delta + 2,$$

concluding our proof. □

□

If one does not care about keeping the seed length of the construction close to $1 \cdot \log\frac{1}{\varepsilon}$, a merging step can be performed, leading to a generic error reduction scheme for condensers. Also, the case of reducing the error from ε_0 to a constant power of ε_0 is not the standard one, and we are usually interested, say, in obtaining a polynomially-small error condenser from a constant error one. Using the same techniques and the curve merger (see [19, Theorem 5.5]) one obtains the following version of the [39] result:

Theorem 3.5. *Suppose* $C : [N] \times [D] \to [M]$ *is an explicit* $(k_{in} \to k_{out}, \varepsilon_0)$ *condenser,* $0 < \alpha < 1$ *and* $0 < \varepsilon < \varepsilon_0$. *Then, there exists an explicit* $C' : [N] \times [D'] \to [M]$ *that is a* $(k'_{in} \to k'_{out}, \varepsilon)$ *condenser with*

- $k'_{in} = k_{in} + \log\frac{1}{\varepsilon}$,
- $k'_{out} = (1 - \alpha)k_{out} - O\left(\log\frac{d}{\varepsilon(1-\varepsilon_0)}\right)$, *and*
- $d' = d + O(\frac{1}{\alpha}\log\frac{d}{\varepsilon})$.

The proof goes along the same lines as those of Theorem 3.1 and we omit it (and we also do not need Theorem 3.5 for our construction).

4 FROM S.R. CONDENSERS TO S.R. SAMPLERS

Extractors are good samplers in the sense that if E is a (k, ε)-extractor then for every test S we have that $\Pr[E(X, U) \in S]$ deviates from the density of S by at most ε. However, extractors are quite limited in the parameters they can achieve and in particular require seed length that is at least $2\log(\frac{1}{\varepsilon})$ [35]. Dodis, Pietrzak and Wichs [18] observed that if we are only interested in fooling sparse tests (or equivalently, if we also allow multiplicative error), it suffices to use condensers with a small entropy gap. Moreover, Dodis et al. note that such condensers can bypass the severe limitations that confine extractors. In this section we show that the Dodis et al. result carries over to s.r. condensers. We obtain explicit s.r. samplers with the parameters we need, and in particular seed length that is smaller than $2\log(\frac{1}{\varepsilon})$.

We first define samplers with both multiplicative and additive error.

Definition 4.1 (sampler). Let $S : [N] \times [R'] \to [D]$.

- We say $x \in [N]$ is (c, ε) bad for $B \subseteq [D]$ if

  $$\Pr_{y' \in [R']}[S(x, y') \in B] > c\rho(B) + \varepsilon.$$

- We say S is a $(K; c, \varepsilon)$ sampler if for every $B \subseteq [D]$,

  $$|\{x \in [N] \mid x \text{ is } (c, \varepsilon) \text{ bad for } B\}| < K.$$

Definition 4.2 (s.r. sampler). Let $S : [N] \times [R'] \times [A] \to [D]$.

- We say $x \in [N]$ is (c, ε) bad for $B \subseteq [D]$ if

  $$\Pr_{y' \in [R']}[\forall_{z \in [A]}\, S(x, y', z) \in B] > c\rho(B) + \varepsilon.$$

- We say S is a $(K; c, \varepsilon)$ s.r. sampler if for every $B \subseteq [D]$,

  $$|\{x \in [N] \mid x \text{ is } (c, \varepsilon) \text{ bad for } B\}| < K.$$

Definition 4.3. We say $S : [N] \times [R'] \times [A] \to [D]$ is *simple* if for every $x \in [N]$, and every $y'_1, y'_2 \in [R']$, $z_1, z_2 \in [A]$, if $(y'_1, z_1) \neq (y'_2, z_2)$ then $S(x, y'_1, z_1) \neq S(x, y'_2, z_2)$.

The following lemma is based on [18].

Lemma 4.4. *If* X *is a* $(d, d - g)$*-source then for every set* $B \subseteq [D]$,

$$\Pr[X \in B] \leq 2^g \cdot \rho(B).$$

Proof. If X is flat, then the probability that X is in B is bounded by the density of B inside the support of X, i.e., it is at most $\frac{|B|}{\mathrm{Supp}(X)|} \le \frac{|B|}{2^{d-g}} = 2^g \cdot \rho(B)$. Since every $(d, d-g)$-source is a convex combination of such flat sources, the lemma follows. □

Lemma 4.5. *If $X = X_1 \circ \ldots \circ X_A$ is a $(d, d-g, \zeta)$ s.r. source then for every set $B \subseteq [D]$,*

$$\Pr_{x \sim X}[\forall_{z \in [A]} X_z \in B] \le 2^g \cdot \rho(B) + \zeta.$$

Proof. X is ζ-close to a $(d, d-g)$ s.r. source X'. Let I be an indicator of X'.

$$\begin{aligned}\Pr[\forall_{z\in[A]} X'_z \in B] &\le \sum_{i=1}^{A} \Pr[I=i] \cdot \Pr_{x\sim X'}[\forall_{z\in[A]} X'_z \in B | I=i] \\ &\le \sum_{i=1}^{A} \Pr[I=i] \cdot \Pr[X'_i \in B | I=i] \\ &\le \sum_{i=1}^{A} \Pr[I=i] \cdot 2^g \cdot \rho(B) \le 2^g \cdot \rho(B),\end{aligned}$$

where the third inequality follows from Lemma 4.4. □

Theorem 4.6. *If $C : [N] \times [R'] \times [A] \to [D]$ is a $(k \to d-g, \varepsilon)$ s.r. condenser then C is a $(2^k; 2^g, \varepsilon)$ s.r. sampler.*

Proof. Let $B \subseteq [D]$, and let BAD denote the set of elements in $[N]$ that are $(2^g, \varepsilon)$ bad for B. If $|\mathrm{BAD}| \ge K$, then $C(\mathrm{BAD}, U)$ is a $(d, d-g, \varepsilon)$ s.r. source. By Lemma 4.5,

$$\Pr_{x \in \mathrm{BAD}, y' \in [R']}[\forall_{z\in[A]} C(x,y')_z \in B] \le 2^g \cdot \rho(B) + \varepsilon.$$

Therefore, there must exist at least one $x \in \mathrm{BAD}$ such that

$$\Pr_{y' \in [R']}[\forall_{z\in[A]} C(x,y')_z \in B] \le 2^g \cdot \rho(B) + \varepsilon,$$

in contradiction to the definition of BAD. Thus, $|\mathrm{BAD}| < K$, as required. □

We now instantiate Theorem 4.6 with the s.r. condenser from Theorem 3.1 to obtain:

Theorem 4.7. *For every constant $0 < \alpha < 1$ there exist constants $A = A(\alpha), b = b(\alpha)$ such that for every $d = d(n) \le n/2$ there exists an explicit*

$$S : [N] \times [R'] \times [A] \to [D]$$

that is a $(K = D^2; c = (\frac{1}{\varepsilon})^4, \varepsilon = (\frac{1}{n})^b)$ simple s.r. sampler with $r' = (1+\alpha)\log(\frac{1}{\varepsilon})$.

Proof. We are given α. Set A to be the constant $A = A(\alpha)$ given in Theorem 3.1. Set $b = b(\alpha) = \frac{4c_{GUV}}{\alpha}$, $\varepsilon = (\frac{1}{n})^b$ and let $r' = (1+\alpha)\log(\frac{1}{\varepsilon})$. Given d, set $m = d - a - r'$. Let

$$\mathrm{SRC} : [N] \times [R'] \times [A] \to [M]$$

denote the $(2m + \log(\frac{1}{\varepsilon}) \to m - 2\log(\frac{1}{\varepsilon}) - O(a), \varepsilon)$ s.r. condenser from Theorem 3.1. Define a new condenser

$$S : [N] \times [R'] \times [A] \to [R'] \times [A] \times [M]$$

by

$$S(x, y', z) = (y', z, \mathrm{SRC}(x, y', z)).$$

It is immediate that S is simple (because the seed is part of the output). Notice that $M \cdot R' \cdot A = D$. By Theorem 4.6, S is a

$$(2^{2m+\log(\frac{1}{\varepsilon})}; c = 2^{m+r'+a-(m-2\log(\frac{1}{\varepsilon})-O(a))}, \varepsilon)$$

s.r sampler, and:

- $2^{2m+\log(\frac{1}{\varepsilon})} = \frac{M^2}{\varepsilon} \le (MR')^2 \le D^2 = K$.
- $c = 2^{m+r'+a-(m-2\log(\frac{1}{\varepsilon})-O(a))} \le 2^{r'+2\log(\frac{1}{\varepsilon})+O(a)}$
  $= 2^{(3+\alpha)\log(\frac{1}{\varepsilon})+O(a)} \le \varepsilon^{-4}$.

□

5 FROM S.R. SAMPLERS TO TWO-SOURCE EXTRACTORS: EXTENDING THE CZ APPROACH

In this section we prove:

Theorem 5.1 (Theorem 1.1 restated). *For every* constant $\varepsilon > 0$ *there exists a constant c such that for every large enough integer n, there exists an explicit $((n,k_1),(n,k_2),\varepsilon)$ two-source extractor* $\mathsf{2EXT} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *for any $k_1, k_2 \ge c \cdot \log n \cdot \log\log n$.*

The proof closely follows the intuition given in the introduction, but makes it rigorous and as a result many parameters enter the discussion. We encourage the reader to read Section 1.1 before reading this section.

Proof.

The construction: We are given n and a constant ε.

- Set t large enough so that $c_{\mathrm{Maj}}\frac{\log t}{t} \le \frac{\varepsilon}{6}$, where c_{Maj} is the constant from Lemma 2.17. A calculation shows that $t = \lceil \frac{36c_{\mathrm{Maj}}^2}{\varepsilon^2} \rceil$ suffices, so $t = O(1)$.
- We now fix parameters for the s.r. sampler. We set $\alpha = \frac{1}{2}$. Let $A = A(\alpha), b = b(\alpha)$ be the constants determined by α in Theorem 4.7. Fix $\varepsilon_1 = (\frac{1}{n})^b$ and $c = (\frac{1}{\varepsilon_1})^4$. Set $R' = (\frac{1}{\varepsilon_1})^{1+\alpha}$.
- Set ε_2 small enough so that $5c_{\mathrm{Maj}}t\sqrt{\varepsilon_2}(R')^t \le \frac{\varepsilon}{6}$ and $c\sqrt{\varepsilon_2} \le \frac{1}{2}(R')^{-0.6}$. We may take $\varepsilon_2 = \frac{\varepsilon^2}{30^2 c_{\mathrm{Maj}}^2 t^2 c^2 (R')^{2t}}$ and so $\varepsilon_2^{-1} = (cR')^{O(1)} = n^{O(1)}$.
- Set $d = c_2(tA)^2 f(n, \varepsilon_2)$ where $f(n,\varepsilon_2) = \log n + \log\frac{1}{\varepsilon_2}\log\log\frac{1}{\varepsilon_2}$ and c_2 is the constant from Theorem 2.22 (we note that everything keeps working if someone gives an improved construction with, say, $f(n,\varepsilon) = O(\log\frac{n}{\varepsilon})$).
- Let
  $$S : [N] \times [R'] \times [A] \to [D],$$
  be the $(K = D^2; c, \varepsilon_1)$ s.r. simple sampler guaranteed by Theorem 4.7 where N, R', D, A, c and ε_1 were defined before.
- Let
  $$\mathsf{nmEXT} : [N] \times [D] \to \{0,1\},$$
  be the $(t' = tA, k_2 = d, \varepsilon_2)$-non-malleable extractor guaranteed by Theorem 2.22. Notice that d was chosen to be sufficiently large for nmEXT.

After fixing the above, given $x_1, x_2 \in [N]$, the construction is as follows:

(1) For every $y' \in [R']$ and $z \in [A]$, compute
$$\mathsf{NM}(x_1, x_2; y', z) = \mathsf{nmEXT}(x_2, S(x_1, y', z)).$$

(2) For every $y' \in [R']$, compute

$$\oplus \mathsf{NM}(x_1, x_2; y') = \bigoplus_{z=1}^{A} \mathsf{NM}(x_1, x_2; y', z).$$

(3) Output

$$2\mathsf{EXT}(x_1, x_2) = \mathrm{Maj}(\oplus \mathsf{NM}(x_1, x_2; 1), \ldots, \oplus \mathsf{NM}(x_1, x_2; R')).$$

Correctness: We now prove correctness. Let X_1 be an (n, k_1)-source for $k_1 = k + \log \frac{2}{\varepsilon}$ and X_2 be an (n, k_2)-source independent from X_1.

Let $\mathrm{BAD} \subseteq [D]$ be the set of bad rows of nmEXT of density at most $\sqrt{\varepsilon_2}$ guaranteed to us by Lemma 2.19. Note that BAD depends only on X_2. We say $x_1 \in [N]$ is *bad* if x is (c, ε_1) bad for BAD (see Definition 4.2) and good otherwise.

Claim 5.2. $\Pr_{x_1 \sim X_1}[x_1 \text{ is bad}] \leq \frac{\varepsilon}{2}$.

Proof. The number of bad elements is at most K, and $H_\infty(X_1) \geq k + \log \frac{2}{\varepsilon}$ so we can conclude that $\Pr_{x_1 \sim X_1}[x_1 \text{ is bad}] \leq \frac{K}{2^{k_1}} = \frac{\varepsilon}{2}$. □

Define

$$\oplus R(x_1, x_2) = (\oplus \mathsf{NM}(x_1, x_2, 1), \ldots, \oplus \mathsf{NM}(x_1, x_2, R')).$$

Lemma 5.3. *For every good $x_1 \in \mathrm{Supp}(X_1)$, $\oplus R(x_1, X_2)$ is a (q, t, γ) non-oblivious bit-fixing source for $q = (R')^{0.4}$ and $\gamma = 5t\sqrt{\varepsilon_2}$.*

Proof. Fix any good $x_1 \in \mathrm{Supp}(X_1)$. Call $y' \in [R']$ *a bad row* if $\forall_{z \in [A]}\ S(x, y', z) \in \mathrm{BAD}$, and good otherwise. Since x_1 is good, the number of bad rows is at most $(c\rho(\mathrm{BAD}) + \varepsilon_1)R'$. However,

$$\begin{aligned}(c\rho(\mathrm{BAD}) + \varepsilon_1)R' &\leq cR'\sqrt{\varepsilon_2} + \varepsilon_1 R' \\ &\leq \frac{1}{2}(R')^{-0.6}R' + \varepsilon_1 \left(\frac{1}{\varepsilon_1}\right)^{1+\alpha} < (R')^{0.4}.\end{aligned}$$

Next, fix t distinct *good* rows $y'_1, \ldots, y'_t$. Let $z_1, \ldots, z_t \in [A]$ be s.t. $S(x_1, y'_i, z_i) \notin \mathrm{BAD}$ (the z_i-s exist because we look at t good rows). Then, for every $i \in [t]$,

$$\Big(\mathsf{NM}(x_1, X_2; y'_i, z_i), \left\{\mathsf{NM}(x_1, X_2; y'_i, z)\right\}_{z \neq z_i}, \left\{\mathsf{NM}(x_1, X_2; y'_j, z)\right\}_{j \neq i, z \in [A]}\Big)$$

is $\sqrt{\varepsilon_2}$-close to

$$\Big(U_1, \left\{\mathsf{NM}(x_1, X_2; y'_i, z)\right\}_{z \neq z_i}, \left\{\mathsf{NM}(x_1, X_2; y'_j, z)\right\}_{j \neq i, z \in [A]}\Big),$$

where we have used the fact that a good row is independent of any other tA (good or bad) rows (see Lemma 2.19), and the fact that S is simple. We remark that the property we are using is more than just the t-wise independence of the good rows.

By Lemma 2.20,

$$\Big(\mathsf{NM}(x_1, X_2; y'_1, z_1), \ldots, \mathsf{NM}(x_1, X_2; y'_t, z_t), \left\{\mathsf{NM}(x_1, X_2; y'_i, z)\right\}_{(y'_i, z) \notin \{(y'_1, z_1), \ldots, (y'_t, z_t)\}}\Big)$$

is $5t\sqrt{\varepsilon_2}$-close to

$$\Big(U_t, \left\{\mathsf{NM}(x_1, X_2; y'_i, z)\right\}_{(y'_i, z) \notin \{(y'_1, z_1), \ldots, (y'_t, z_t)\}}\Big).$$

This shows

$$(\oplus \mathsf{NM}(x_1, X_2, y'_1), \ldots, \oplus \mathsf{NM}(x_1, X_2, y'_t)) \approx_{5t\sqrt{\varepsilon_2}} U_t$$

so $\oplus R(x_1, X_2)$ is a $(q = (R')^{0.4}, t, \gamma)$ non-oblivious bit-fixing source for $\gamma = 5t\sqrt{\varepsilon_2}$ as desired. □

Therefore, by Lemma 2.17, for any good x_1,

$$\left|\Pr[\mathrm{Maj}(\oplus \mathsf{NM}(x_1, X_2, 1), \ldots, \oplus \mathsf{NM}(x_1, X_2, R')) = 1] - \frac{1}{2}\right| \leq c_{\mathrm{Maj}}\left(\frac{\log t}{t} + (R')^{-0.1} + 5t\sqrt{\varepsilon_2}(R')^t\right) \leq 3 \cdot \frac{\varepsilon}{6} = \frac{\varepsilon}{2},$$

where the probability is over X_2. Overall, we have:

$$|2\mathsf{EXT}(X_1, X_2) - U_1| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon,$$

as desired. The requested entropies are $k_2 = d$, $k_1 = O(d)$ and $d = O(f(n, \frac{1}{\mathrm{poly}(n)}))$ so $k_1, k_2 = O(\log n \cdot \log\log n)$. The explicitness follows from the fact that $R' = \mathrm{poly}(n)$ and the explicitness of the other ingredients.

□

We note that if one improves the construction of non-malleable extractors and gets $f(n, \varepsilon_2) = O(\log \frac{n}{\varepsilon_2})$, then this would immediately imply a two-source extractors with entropies $k_1 = k_2 = O(\log n)$.

ACKNOWLEDGMENTS

We thank Gil Cohen and David Zuckerman for commenting on a first version of the paper.

The work is supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120.

REFERENCES

[1] HL Abbott. 1972. Lower bounds for some Ramsey numbers. *Discrete Mathematics* 2, 4 (1972), 289–293.
[2] Miklós Ajtai and Nathan Linial. 1993. The influence of large coalitions. *Combinatorica* 13, 2 (1993), 129–145.
[3] Noga Alon. 1998. The Shannon capacity of a union. *Combinatorica* 18, 3 (1998), 301–310.
[4] Noga Alon, Oded Goldreich, and Yishay Mansour. 2003. Almost k-wise independence versus k-wise independence. *Inform. Process. Lett.* 88, 3 (2003), 107–110.
[5] Boaz Barak. 2006. A Simple Explicit Construction of an $n^{\tilde{O}(\log n)}$-Ramsey Graph. *arXiv preprint math/0601651* (2006).
[6] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. 2006. Extracting randomness using few independent sources. *SIAM J. Comput.* 36, 4 (2006), 1095–1118.
[7] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. 2010. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *Journal of the ACM (JACM)* 57, 4 (2010), 20.
[8] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2012. 2-source dispersers for $n^{o(1)}$ entropy, and Ramsey graphs beating the Frankl-Wilson construction. *Annals of Mathematics* 176, 3 (2012), 1483–1544.
[9] Jean Bourgain. 2005. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory* 1, 01 (2005), 1–32.
[10] Eshan Chattopadhyay and Xin Li. 2016. Explicit Non-malleable Extractors, Multi-source Extractors, and Almost Optimal Privacy Amplification Protocols. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*. IEEE, 158–167.
[11] Eshan Chattopadhyay and David Zuckerman. 2016. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*. ACM, 670–683.
[12] Benny Chor and Oded Goldreich. 1988. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.* 17, 2 (1988), 230–261.
[13] Fan RK Chung. 1981. A note on constructive methods for Ramsey numbers. *Journal of Graph Theory* 5, 1 (1981), 109–113.
[14] Gil Cohen. 2015. Two-Source Dispersers for Polylogarithmic Entropy and Improved Ramsey Graphs. *arXiv preprint arXiv:1506.04428* (2015).

[15] Gil Cohen. 2016. Making the most of advice: New correlation breakers and their applications. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*. IEEE, 188–196.
[16] Gil Cohen. 2016. Two-source extractors for quasi-logarithmic min-entropy and improved privacy amplification protocols. In *Electronic Colloquium on Computational Complexity (ECCC)*.
[17] Gil Cohen and Leonard J Schulman. 2016. Extractors for near logarithmic min-entropy. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*. IEEE, 178–187.
[18] Yevgeniy Dodis, Krzysztof Pietrzak, and Daniel Wichs. 2014. Key derivation without entropy waste. In *Advances in Cryptology–EUROCRYPT 2014*. Springer, 93–110.
[19] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. 2013. Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. *SIAM J. Comput.* 42, 6 (2013), 2305–2328.
[20] Paul Erdös. 1947. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* 53, 4 (1947), 292–294.
[21] Peter Frankl. 1977. A constructive lower bound for Ramsey numbers. *Ars Combinatoria* 3, 297-302 (1977), 28.
[22] Peter Frankl and Richard M. Wilson. 1981. Intersection theorems with geometric consequences. *Combinatorica* 1, 4 (1981), 357–368.
[23] Oded Goldreich. 2011. *A sample of samplers: A computational perspective on sampling*. Lecture Notes in Computer Science, Vol. 6650. Springer, 302–332.
[24] Parikshit Gopalan. 2006. Constructing ramsey graphs from boolean function representations. In *Computational Complexity, 2006. CCC 2006. Twenty-First Annual IEEE Conference on*. IEEE, 14–pp.
[25] Vince Grolmusz. 2001. Low rank co-diagonal matrices and ramsey graphs. *Journal of combinatorics* 7, 1 (2001), R15–R15.
[26] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. 2009. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Journal of the ACM (JACM)* 56, 4 (2009), 20.
[27] Xin Li. 2011. Improved constructions of three source extractors. In *Computational Complexity (CCC), 2011 IEEE 26th Annual Conference on*. IEEE, 126–136.
[28] Xin Li. 2013. Extractors for a constant number of independent sources with polylogarithmic min-entropy. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 100–109.
[29] Xin Li. 2013. New independent source extractors with exponential improvement. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. ACM, 783–792.
[30] Xin Li. 2016. Improved non-malleable extractors, non-malleable codes and independent source extractors. *arXiv preprint arXiv:1608.00127* (2016).
[31] Xin Li. 2016. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*. IEEE, 168–177.
[32] Raghu Meka. 2017. Explicit resilient functions matching Ajtai-Linial. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1132–1148.
[33] Zs Nagy. 1975. A constructive estimation of the Ramsey numbers. *Mat. Lapok* 23 (1975), 301–302.
[34] Moni Naor. 1992. Constructing Ramsey graphs from small probability spaces. *IBM Research Report RJ* 8810 (1992).
[35] Jaikumar Radhakrishnan and Amnon Ta-Shma. 2000. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics* 13, 1 (2000), 2–24.
[36] Anup Rao. 2007. An exposition of Bourgain's 2-source extractor. In *Electronic Colloquium on Computational Complexity (ECCC)*.
[37] Anup Rao. 2009. Extractors for a constant number of polynomially small min-entropy independent sources. *SIAM J. Comput.* 39, 1 (2009), 168–194.
[38] Ran Raz. 2005. Extractors with weak random seeds. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. ACM, 11–20.
[39] Ran Raz, Omer Reingold, and Salil Vadhan. 1999. Error reduction for extractors. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*. IEEE, 191–201.
[40] Emanuele Viola. 2014. Extractors for circuit sources. *SIAM J. Comput.* 43, 2 (2014), 655–672.
[41] David Zuckerman. 2006. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*. ACM, 681–690.