Electronic Colloquium on Computational Complexity, Report No. 27 (2017)

A reduction from efficient non-malleable extractors to low-error two-source extractors with arbitrary constant rate

Avraham Ben-Aroya* Eshan Chattopadhyay† Dean Doron‡ Xin Li§
Amnon Ta-Shma¶

Abstract

We show a reduction from the existence of explicit t-non-malleable extractors with a small seed length, to the construction of explicit two-source extractors with small error for sources with arbitrarily small constant rate. Previously, such a reduction was known either when one source had entropy rate above half [Raz05] or for general entropy rates but only for large error [CZ16].

As in previous reductions we start with the Chattopadhyay and Zuckerman approach [CZ16], where an extractor is applied on one source to create a table, and the second source is used to sample a sub-table. In previous work, a resilient function was applied on the sub-table and the use of resilient functions immediately implied large error. In this work we replace the resilient function with the parity function (that is not resilient). We prove correctness by showing that doing the sampling properly, the sample size can be made so small that it is smaller then the non-malleability parameter t of the first extractor, which is enough for the correctness.

The parameters we require from the non-malleable construction hold (quite comfortably) in a non-explicit construction, but currently it is not known how to explicitly construct such graphs. As a result we do not give an unconditional construction of an explicit low-error two-source extractor. However, the reduction shows a new connection between non-malleable and two-source extractors, and also shows resilient functions do not play a crucial role in the two-source construction framework suggested in [CZ16]. Furthermore, the reduction highlights a barrier in constructing non-malleable extractors, and reveals its importance. We hope this work would lead to further developments in explicit constructions of both non-malleable and two-source extractors.

*The Blavatnik School of Computer Science, Tel-Aviv University, Tel Aviv 69978, Israel. Supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120.

†Simons Institute for the Theory of Computing, UC Berkeley, Berkeley, CA 94720, USA. Email: eshanc@ias.edu. Part of this work was done when the author was a graduate student in UT Austin under NSF grant CCF-1526952

‡The Blavatnik School of Computer Science, Tel-Aviv University, Tel Aviv 69978, Israel. Email: deandoron@mail.tau.ac.il. Supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120. This work was done in part while visiting the Simons Institute for the Theory of Computing at UC Berkeley.

§Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA. Email: lixints@cs.jhu.edu. Supported by NSF Grant CCF-1617713.

¶The Blavatnik School of Computer Science, Tel-Aviv University, Tel Aviv 69978, Israel. Email: amnon@tau.ac.il. Supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120. This work was done in part while visiting the Simons Institute for the Theory of Computing at UC Berkeley.

ISSN 1433-8092

1 Introduction

A *randomness extractor* is a function that purifies crude randomness sources. A crude randomness source is usually modeled as an (n,k)-source, i.e., a distribution over $\{0,1\}^n$ with min-entropy at least k. It is easy to see that extracting from a general (n,k)-source is impossible and thus the requirement from the extractor needs to be relaxed. Different relaxations lead to different flavors of extractors. In this paper we focus on *two-source extractors*. An $((n_1,k_1),(n_2,k_2),\varepsilon)$ two-source extractor is a function $E:\{0,1\}^{n_1}\times\{0,1\}^{n_2}\to\{0,1\}^m$ that maps any pair of independent (n_1,k_1) and (n_2,k_2) sources X,Y to a distribution $E(X,Y)$ which is ε-close to U_m, the uniform distribution over $\{0,1\}^m$. Two-source extractors for small min-entropies having *small error* is the focal point of this paper.

A *Ramsey graph* is a graph that contains neither a large clique nor an independent set. This notion can be generalized to bipartite graphs; a *bipartite Ramsey graph* is a bipartite graph without large bipartite cliques or bipartite independent sets. It is not hard to see that bipartite Ramsey graphs correspond to two-source extractors that output one bit, with some nontrivial error.

A long line of research was devoted to explicitly constructing Ramsey graphs [Abb72, Nag75, Fra77, Chu81, FW81, Nao92, Alo98, Gro01, Bar06], bipartite Ramsey graphs [BKS+10, BRSW12, Coh15], and two-source extractors [CG88, Raz05, Bou05] culminating in the work of Chattopadhyay and Zuckerman [CZ16] who constructed a two-source extractor for $k=\mathrm{polylog}(n)$. Several improvements on the [CZ16] construction followed, including [Mek15, Li15a]. Currently, following [BADTS16, Coh16d, Li16], the best explicit construction achieves $k=O(\log n\log\log n)$ which is pretty close to the optimal $\Omega(\log n)$ bound.

The recent explicit two-source constructions mentioned above (i.e., [CZ16, BADTS16, Coh16d, Li16]) can only handle high error, in the sense that the running time of the extractor is polynomial in $\frac{1}{\varepsilon}$. While this suffices for Ramsey graph constructions, non-explicit constructions may have exponentially small error in the entropy k of the two sources. Similarly, these constructions usually output few close-to-uniform bits, while non-explicitly, almost all of the entropy can be extracted.

There are several explicit two-source constructions with exponentially-small error. The inner-product function gives a simple solution when $k>n/2$ [CG88]. Bourgain [Bou05] gave a two-source extractor construction for $k=(\frac{1}{2}-\alpha)n$, for some small constant $\alpha>0$. Raz [Raz05] 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). All of these constructions have exponentially small (in k) error. However, in all of these constructions one of the sources is required to have entropy rate close to half, i.e., the entropy of the source has to be at least $(\frac{1}{2}-\alpha)n>0.49n$.

We also mention that for *three* sources Li [Li15b] constructed an extractor with exponentially-small error for min-entropy $k=\mathrm{polylog}(n)$. Yet, achieving the same with only two sources, is a challenging and important open problem.

Central to the work of [CZ16] are *non-malleable extractors*. These are another variant of extractors, devised by Dodis and Wiches [DW09], which we now briefly describe. A *seeded extractor* is an extractor that purifies (n,k)-sources using a short truly random seed. That is, a function $E:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ is a (k,ε) strong seeded extractor if for any (n,k)-source X and for an independent $Y\sim U_d$, the distribution $(Y,E(X,Y))$ is ε-close to (Y,U_m). A (k,ε) *t-non-malleable extractor* strengthens a strong seeded extractor by requiring that under the same premise the extractor's output is close to uniform, even given the evaluation of E on t arbitrarily correlated

seeds, i.e., the distribution $(Y, E(X,Y), \{E(X, f_i(Y))\}_{i=1}^t)$ is ε-close to $(Y, U_m, \{E(X, f_i(Y))\}_{i=1}^t)$ for *any* functions $f_1, \ldots, f_t : \{0,1\}^d \to \{0,1\}^d$ with no fixed-points.

The main result in this paper is a reduction from the existence of certain non-malleable extractors to low-error two-source extractors with any constant entropy rate:

Theorem 1.1. *There exists a constant $c \geq 1$ such that the following holds. Suppose for some constant $\alpha > 0$ for every n_1, k_1, ε_1 and t there exists an explicit $E : \{0,1\}^{n_1} \times \{0,1\}^d \to \{0,1\}^m$ that is a strong (k_1, ε_1) t-n.m. extractor for $d \leq \frac{\alpha t}{c} \cdot \log(\frac{1}{\varepsilon})$. Then, for every $\varepsilon > 0$ there exists an explicit $((n_1, k_1), (n_2 = 4d/\alpha, k_2 = \alpha n_2), 2\sqrt{\varepsilon_1})$ two-source extractor that outputs m bits.*

The parameters here resemble those of Raz's extractor: one source is long with very low entropy, the other is short with constant entropy rate. The main difference is that in Raz's extractor the entropy rate has to be above half, whereas here, assuming the existence of the appropriate non-malleable extractors, the entropy rate can be an arbitrarily small constant. In particular, we have:

Corollary 1.2. *Given explicit optimal n.m. extractors, for every n_1, $k_1 \leq n_1$ and a constant $0 < \alpha < 1$, there exists an explicit $\left((n_1, k_1), (n_2 = \Omega\left(\frac{1}{\alpha}(\log n_1 + k_1)\right), k_2 = \alpha n_2), 2^{-\Omega(k_1)}\right)$ two-source extractor that outputs $m = \Omega(\alpha k_1)$ bits.*

We also mention that the requirement $d \leq \frac{\alpha t}{c} \cdot \log(\frac{1}{\varepsilon})$ is quite tight. We will show that the approach cannot work if $d \geq \frac{16t}{c} \cdot \log(\frac{1}{\varepsilon})$.

We now explain the ideas behind the reduction, starting with recalling the main ideas in the construction of the two-source extractor in [CZ16] and the bottleneck for achieving smaller error. Let X and Y be two independent (n,k)-sources. The starting point of [CZ16] is to use a t-non-malleable extractor E with error ε_1 and seed length d_1 to produce a table with $D_1 = 2^{d_1}$ entries, where the i-th entry is $E(X,i)$. Using the property of the non-malleable extractor, one can show that $(1 - \sqrt{\epsilon_1})$-fraction of the rows are uniform and almost t-wise independent. The remaining rows are, however, arbitrarily correlated with these rows. In [CZ16] they then:

- Use the second source to sample a sub-table of length D_2, such that a fraction of about ε_2 of its rows are bad, and every t good rows are $\sqrt{\varepsilon_2}$-close to uniform, and,
- Apply a *resilient function* $f : \{0,1\}^{D_2} \to \{0,1\}$ on the sub-table. f has to be resilient to $\sqrt{\varepsilon_2} D_2$ bad players, and should perform correctly even when the good players are t-wise independent.

The sub-table is $D_2^t t\sqrt{\varepsilon_1}$ close to a table where the good players are *truly* t-wise independent (as required by f) and so it is enough to choose ε_1 small enough so that $D_2^t t\sqrt{\varepsilon_1}$ is small.

While this beautiful approach does give an unbiased output bit, it seems that it is inherently bound to have high error. This is because no matter what the resilient function is, there is always one player among the D_2 players with $1/D_2$ influence over the result (in fact, [KKL88] showed there is a player with $\Omega(\frac{\log D_2}{D_2})$ influence) and so even if there is just a single bad player, this player can bias the result by $\frac{1}{D_2}$. Thus, the running time, which is at least D_2, is at least $\Omega(\frac{1}{\varepsilon})$.

In our solution we completely dispense with the use of resilient functions. We also do not try to achieve a sub-table that is close to a *truly* t-wise independent distribution, and we strive to work with the much weaker guarantee that every t good rows are close to uniform. This leads us to the following two-source extractor:

Let X and Y be two independent (n, k)-sources. As in [CZ16] we use a t-non-malleable extractor E with error ϵ_1 and seed length d_1 to produce a table with $D_1 = 2^{d_1}$ entries, where the i-th entry is $E(X, i)$. Using the property of the non-malleable extractor, one can show that $(1 - \sqrt{\epsilon_1})$-fraction of the rows are uniform and almost t-wise independent. The remaining rows are arbitrarily correlated with these rows. Then:

- We use the second source to sample t rows from that table, with the property that with high probability (over the choice of $x_2 \in X_2$) at least one of the t samples is a good row (in the table with D_1 rows).
- We then take the parity of the t bits written in the t rows we sampled.

Conceptually, what happened is that we take a *dramatically smaller* sample set than before. Specifically, in [CZ16, BADTS16] the sample set is much larger than t, whereas in our algorithm the sample size is t. Accordingly, we replace the requirement that the fraction of bad players in the sample set is small, with the weaker requirement that not all of the players in the sample set are bad. If the sample size is t and not all the players in the sample are bad, then the good player is almost independent of the other $t - 1$ players, and therefore we can just apply the parity function on the t bits in the sample. Thus, we can also replace the resilient function f with the parity function (which is not resilient at all, but is now good for us).

Notice, that by doing so we also get rid of the annoying (and expensive) requirement that $D_2^t \varepsilon_1 < 1$, because we no longer need to convert a table where every t rows are close to uniform, to a table that is close to being perfectly t-wise independent.

The only question that is left is whether we can find a sampler with such a small sample size, that except for very few x_2-s, always sees at least one good row. This question readily translates to the existence (or the explicit existence) of dispersers of constant degree that are good against small tests. Remarkably, Zuckerman [Zuc06] gave a beautiful explicit construction with nearly optimal bounds.

We are still left with the question of whether these graphs (with nearly optimal parameters) can support such a construction. The question is non-trivial because there is a circular dependency in the construction: the sample size of the sampler determines the required t-non-malleability of the extractor, which then affects the parameters of the extractor, and in particular the number of bad rows, which, in turn, affects the required degree of the sampler. It is therefore, offhand, not clear at all whether such a construction is possible at all even assuming the best possible non-malleable extractors.

The above discussion raises the question of what is the dependence of the seed length of non-malleable extractors on the non-malleability parameter t. This question was considered before by several people. In particular, Cohen and Shinkar [CS17] independently investigated this question. We are grateful to Cohen for discussions regarding this during early stages of our work. It turns out that in non-explicit constructions the dependence is very mild, and such an approach can be easily supported.

In the paper we analyze what is the threshold beyond which such an approach cannot work. Roughly speaking, extractors with seed length below $t \log(\frac{n}{\varepsilon})$ work well, while extractors with seed length above it do not. Interestingly, the current constructions of non-malleable extractors [Coh16a, CGL15, Coh16c, Coh16b, CL16, Coh16d, Li16] for small entropies are above this threshold. This is mainly due to the use of alternating extraction techniques which treat the seed and the source symmetrically. One consequence of our work is identifying this bottleneck, and the

importance of overcoming it. We hope this work would lead to further developments in explicit constructions of both non-malleable and two-source extractors.

2 Preliminaries

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

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 that X is ε-close to Y and denote it by $X \approx_\varepsilon Y$. We will denote by U_n a random variable distributed uniformly over $\{0,1\}^n$ and which is independent of all other variables. We also say that 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, which is obtained by choosing x according to X and computing $f(x)$. For a set $A \subseteq D_1$, we simply denote $f(A) = \{f(x) \mid x \in A\}$. It is well-known that 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

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

A random variable X distributed over $\{0,1\}^n$ with min-entropy at least k is called an (n,k)-*source*. Every distribution X with $H_\infty(X) \geq k$ can be expressed as a convex combination of flat distributions, each with min-entropy at least k.

2.2 Extractors

Definition 2.1. *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* $\mathsf{2Ext}(X_1, X_2) \approx_\varepsilon U_m$.

Definition 2.2. $E : [N] \times [D] \to [M]$ *is a strong* (K, ε) t-non-malleable extractor, *if for every* (N, K)*-source* X *and every functions* $f_1, \ldots, f_t : [D] \to [D]$ *with no fixed-points*[1] *it holds that,*

$$\left|(U, E(X,U), \{E(X, f_i(U))\}_{i=1}^t) - (U, U_m, \{E(X, f_i(U))\}_{i=1}^t)\right| \leq \varepsilon,$$

where U *is the uniform distribution over* $[D]$ *and* U_m *is the uniform distribution over* $[M]$.

A simple consequence, proved in [CZ16], is:

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

Lemma 2.3 ([CZ16], Lemma 3.4)**.** *Let $E : [N] \times [D] \to [M]$ be a strong (K, ε) t-non-malleable extractor. Let X be any (N, K)-source. Then there exists a set $BAD \subseteq [N]$ with $\rho(BAD) \leq \sqrt{\varepsilon}$ such that for every $y \notin BAD$, and every $y'_1, \ldots, y'_t \in [D] \setminus y$,*

$$\left|\left(E(X,y), \left\{E(X,y'_i)\right\}_{i\in[t]}\right) - \left(U_M, \left\{E(X,y'_i)\right\}_{i\in[t]}\right)\right| \quad \leq \quad \sqrt{\varepsilon}.$$

2.3 Dispersers

Definition 2.4. *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 $\left|\bigcup_{i\in[D]} \Gamma(A, i)\right| \geq K'$.*

Zuckerman showed the following remarkable explicit construction:

Theorem 2.5 ([Zuc06], Theorem 1.9)**.** *There exists a constant c_{disp} such that the following holds. Fix any constants $0 < a, b < 1$. Set $K = N^a$, $M \leq K^{1-b}$ and $K' < M$. Then there exists an efficient family of (K, K')-dispersers*

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

with degree $D = c_{disp} \cdot \frac{\log \frac{N}{K}}{\log \frac{M}{K'}} = c_{disp} \cdot \frac{n}{\log \frac{M}{K'}}$.

We remark that the parameters in Theorem 2.5 are tight up to a constant factor.

3 The construction

3.1 The Overall structure

Given:

$$\begin{aligned} E &: [N_1] \times [D] \to \{0,1\}^m \\ \Gamma &: [N_2] \times [t+1] \to [D] \end{aligned}$$

We define $\mathsf{2Ext} : [N_1] \times [N_2] \to \{0,1\}^m$ by

$$\mathsf{2Ext}(x_1, x_2) = \bigoplus_{y : \exists i \text{ s.t. } \Gamma(x_2, i) = y} E(x_1, y).$$

Theorem 3.1. *Assume E is a strong (K_1, ε_1) t-n.m. extractor and Γ is a $(B_2, \sqrt{\varepsilon_1}D)$-disperser. Then for every K_2, $\mathsf{2Ext}$ is a $\left((N_1, K_1), (N_2, K_2), \frac{B_2}{K_2} + \sqrt{\varepsilon_1}\right)$ two-source extractor.*

Proof: Let X_1 be an (N_1, K_1)-source and X_2 an (N_2, K_2)-source. W.l.o.g. X_1 and X_2 are flat. As E is t-n.m., by Lemma 2.3 there exists a set $BAD_1 \subseteq [D]$ with $\rho(BAD_1) \leq \sqrt{\varepsilon_1}$ such that for every $y \notin BAD_1$ and every $y'_1, \ldots, y'_t \in [D] \setminus \{y\}$,

$$\left|\left(E(X,y), \left\{E(X,y'_i)\right\}_{i\in[t]}\right) - \left(U_m, \left\{E(X,y'_i)\right\}_{i\in[t]}\right)\right| \quad \leq \quad \sqrt{\varepsilon_1}.$$

Let $BAD_2 \subseteq [N_2]$ be

$$BAD_2 \quad = \quad \{x_2 \in [N_2] \mid \Gamma(x_2) \subseteq BAD_1\}.$$

Thus, $\Gamma(BAD_2) \subseteq BAD_1$. Since $|BAD_1| \leq \sqrt{\varepsilon_1}D$ and Γ_2 is a $(B_2, \sqrt{\varepsilon_1}D)$- disperser, it follows that $|BAD_2| \leq B_2$. However, for any $x_2 \in [N_2] \setminus BAD_2$, there exists an $i \in [t+1]$ such that $y = \Gamma(x_2, i) \notin BAD_1$. Hence,

$$\left|\left(E(X,y), \{E(X,y_j)\}_{y_j \in \Gamma(x_2)\setminus\{y\}}\right) - \left(U_m, \{E(X,y_j)\}_{y_j \neq \Gamma(x_2)\setminus\{y\}}\right)\right| \leq \sqrt{\varepsilon_1}.$$

Thus,

$$\left|\bigoplus_{y:\exists i \text{ s.t. } \Gamma(x_2,i)=y} E(x_1,y) - U_m\right| \leq \sqrt{\varepsilon_1}.$$

Altogether, the error is at most $\frac{|BAD_2|}{K_2} + \sqrt{\varepsilon_1}$ and the proof is complete. ■

3.2 The activation threshold

In the previous subsection we assumed the existence of a $(B_2, \sqrt{\varepsilon_1}D)$-disperser Γ and a t-n.m. extractor E. However,

- The degree D_2 of the disperser Γ affects the non-malleability parameter t of the extractor, because the argument requires $t \geq D_2 - 1$,
- The non-malleability parameter t affects the degree D of the extractor, because intuitively, the greater t is the greater the degree has to be,
- The degree D determines $|BAD_1| = \sqrt{\varepsilon_1}D$, and,
- The size B_1 of the set BAD_1 determines the degree of the disperser Γ as $D_2 = O\left(\frac{\log \frac{N_2}{B_2}}{\log \frac{D}{B_1}}\right)$, and up to multiplicative factor this is also a lower bound on D_2.

Thus we have a circular dependence and it is not clear at all that such a construction is even possible. Indeed, as we shall see, if the seed length of E is larger than $t \log(\frac{1}{\varepsilon_1})$ such a construction is impossible. Unfortunately, this is indeed the case with all existing *explicit* constructions. However, non-explicitly better non-malleable extractors exist that comfortably suffice for the construction. Our goal in this section is to determine which dependence of the seed length on t and ε_1 suffices for the construction.

Normally, we measure the degree of a t-n.m. (K_1, ε_1) strong extractor $E : [N_1] \times [D_1] \to [M]$ as a function of N_1 and ε_1. For example there are explicit constructions of strong extractors with degree $\mathrm{poly}(\frac{\log N_1}{\varepsilon_1})$. However, since we are interested in the low-error case where $\varepsilon_1 \leq \frac{1}{\log N_1}$ we can suppress the dependence on N_1 and measure D_2 as a function of ε_1 alone. Say, $D_1 = (\frac{1}{\varepsilon_1})^{c_D(t)}$.

Lemma 3.2. *Suppose $E : [N_1] \times [D_1] \to \{0,1\}^m$ having error $\varepsilon_1 \leq \frac{1}{n_1^2}$ and $\Gamma : [N_2] \times [D_2] \to [D_1]$ satisfying the requirements of Theorem 3.1 and suppose that* 2Ext $: [N_1] \times [N_2] \to \{0,1\}^m$ *is the $((N_1, K_1), (N_2, K_2), 2\sqrt{\varepsilon_1})$ two-source extractor constructed and analyzed as above with $K_2 < \sqrt{N_2}$. Then, $D_1 \leq (\frac{1}{\varepsilon_1})^{c_D(t)}$ for $c_D(t) = \frac{t+1}{c_{disp}} + \frac{1}{2}$, where c_{disp} is the constant guaranteed by Theorem 2.5.*

Proof: Since we follow the above analysis, we have $2\sqrt{\varepsilon_1} = \frac{B_2}{K_2} + \sqrt{\varepsilon_1}$. In particular, $B_2 < K_2$. Also, $\Gamma : [N_2] \times [t+1] \to [D_1]$ is a $(B_2, B_1 = \sqrt{\varepsilon_1} D_1)$-disperser. Therefore,

$$t+1 \;\geq\; c_{disp} \cdot \frac{\log \frac{N_2}{B_2}}{\log \frac{D_1}{B_1}} \;\geq\; c_{disp} \cdot \frac{\log K_2}{\log \frac{1}{\sqrt{\varepsilon_1}}}.$$

However, by the properties of the disperser, $K_2 > B_2 \geq \frac{B_1}{t}$ (because otherwise we can take a set of size $\frac{B_1}{t}$ on the left hand size, and the size of its neighbor set is at most B_1, violating the disperser property). Also, $\frac{B_1}{t} \geq \sqrt{B_1}$ because otherwise $\sqrt{B_1} < t$ and thus

$$D_1 \;=\; \frac{B_1}{\sqrt{\varepsilon_1}} \;<\; \frac{t^2}{\sqrt{\varepsilon_1}} \leq \frac{n_1^2}{\sqrt{\varepsilon_1}} \;\leq\; \frac{1}{\varepsilon_1^2},$$

where the last inequality follows from the assumption on ε_1. This contradicts the lower-bound for extractors [RTS00]. Thus, we have $K_2 \geq \sqrt{B_1}$. Also, express $D_1 = \left(\frac{1}{\varepsilon_1}\right)^{c_D(t)}$. Then, $B_1 = \sqrt{\varepsilon_1} D_1 = \left(\frac{1}{\varepsilon_1}\right)^{c_D(t) - \frac{1}{2}}$. Hence,

$$t+1 \;\geq\; c_{disp} \left(c_D(t) - \frac{1}{2} \right).$$

It therefore follows that $c_D(t) \leq \frac{t+1}{c_{disp}} + \frac{1}{2}$. ∎

The analysis in the above proof is quite tight and we can also prove the converse:

Lemma 3.3. *For every constant $\alpha > 0$ the following holds. Suppose for every $N_1, K_1, \varepsilon_1 \leq \frac{1}{\log N_1}$ and t, there exists an explicit $E : [N_1] \times [D_1] \to \{0,1\}^m$ that is a strong (K_1, ε_1) t-n.m. extractor with $D_1 = (\frac{1}{\varepsilon_1})^{c_D(t)}$ for $\frac{1}{4} \leq c_D(t) \leq \frac{\alpha}{8 c_{disp}} t$. Then for every $\varepsilon \leq \frac{1}{n_1}$ there exists an explicit* 2Ext $: [N_1] \times [N_2] \to \{0,1\}^m$ *that is a $((N_1, K_1), (N_2 = D_1^{4/\alpha}, K_2 = N_2^{\alpha}), \varepsilon)$ two-source extractor.*

Proof: Set $\varepsilon_1 = (\varepsilon/2)^2$. Let $D_1 = (\frac{1}{\varepsilon_1})^{c_D(t)}$, $B_1 = \sqrt{\varepsilon_1} D_1$, $B_2 = D_1^2$, $K_2 = B_2^2$ and $N_2 = K_2^{1/\alpha}$. Let $\Gamma : [N_2] \times [D_2] \to [D_1]$ be the $(B_2, B_1 = \sqrt{\varepsilon_1} D_1)$-disperser promised to us by Theorem 2.5 (for $a = \alpha/2$ and $b = \frac{1}{2}$). By that theorem, the degree D_2 of Γ is

$$D_2 \;=\; c_{disp} \cdot \frac{\log \frac{N_2}{B_2}}{\log \frac{D_1}{B_1}} \;=\; c_{disp} \cdot \frac{\left(\frac{1}{\alpha} - \frac{1}{2}\right) \log K_2}{\log \frac{1}{\sqrt{\varepsilon_1}}} \;=\; c_{disp} \cdot \left(\frac{1}{\alpha} - \frac{1}{2}\right) \frac{8 \log D_1}{\log 1/\varepsilon_1} \;=\; c_{disp} \cdot \left(\frac{8}{\alpha} - 4\right) c_D(t),$$

where the constant hidden in the big-O notation is independent of all other parameters.

Set $t = D_2 - 1$ and let $E : [N_1] \times [\widetilde{D}] \to \{0,1\}^m$ be the explicit, strong (K_1, ε_1) t-n.m. extractor with $\widetilde{D} = (\frac{1}{\varepsilon_1})^{c_D(t)}$ promised by the theorem. We also see that $\widetilde{D} = D_1$. Let 2Ext $: [N_1] \times [N_2] \to \{0,1\}^m$ constructed from E and Γ as above.

We see that E is a strong (K_1, ε_1) t-n.m. extractor and Γ is a $(B_2, \sqrt{\varepsilon_1} D_1)$-disperser. Therefore, by Theorem 3.1, 2Ext is a $((N_1, K_1), (N_2, K_2), \frac{B_2}{K_2} + \sqrt{\varepsilon_1})$ two-source extractor. But $\frac{B_2}{K_2} + \sqrt{\varepsilon_1} = \frac{1}{D_1^2} + \frac{\varepsilon}{2} = \varepsilon_1^{2 c_D(t)} + \frac{\varepsilon}{2} = (\frac{\varepsilon}{2})^{4 c_D(t)} + \frac{\varepsilon}{2} \leq \varepsilon$, where the last inequality follows from the assumption $c_D(t) \geq \frac{1}{4}$. ∎

4 The dependence of the seed on the non-malleability degree

The current best explicit construction is due to Li. We cite his result for comparison with the non-explicit construction.

Theorem 4.1 ([Li16]). *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 strong $(k = d, t\varepsilon)$ t-n.m. extractor $\mathsf{nmEXT} : [N] \times [D] \to \{0,1\}$ with seed length $d = c_1 t^2 (\log n + \log \frac{1}{\varepsilon} \cdot \log\log \frac{1}{\varepsilon})$.*

Non-explicitly, we extend the [DW09] result and prove that

Theorem 4.2. *Let n, k, t and ε be such that $k \geq (t+1)m + 2\log\frac{1}{\varepsilon} + \log d + 4\log t + 3$. There exist a strong (k, ε) t-n.m. extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d \leq 2\log\frac{1}{\varepsilon} + \log(n-k) + 2\log(t+1) + 3$.*

Independently this was proved by Cohen and Shinkar [CS17].

Proof: Choose a function $E : [N] \times [D] \to [M]$ uniformly at random. Fix a flat source X (which we identify with a subset $X \subseteq [N]$ of size K), t functions $f_1, \ldots, f_t : [D] \to [D]$ with no fixed-points and a distinguisher function $\mathcal{D} : \{0,1\}^{(t+1)m+d} \to \{0,1\}$. We want to bound the probability (over E) that

$$\begin{aligned} &\Pr[\mathcal{D}(E(X,Y), E(X, f_1(Y)), \ldots, E(X, f_t(Y)), Y) = 1] - \\ &\quad \Pr[\mathcal{D}(U_m, E(X, f_1(Y)), \ldots, E(X, f_t(Y)), Y) = 1] \;>\; \varepsilon. \end{aligned}$$

For every $y \in [D]$ and $z_1, \ldots, z_t \in [M]$, define

$$\mathrm{Count}(y, z_1, \ldots, z_t) \;=\; \left|\{z \in [M] : \mathcal{D}(z, z_1, \ldots, z_t, y) = 1\}\right|.$$

For every $x \in X$ and $y \in [D]$, define the following random variables (where the randomness comes from E):

$$\begin{aligned} \mathbf{L}(x,y) &= \mathcal{D}(E(x,y), E(x, f_1(y)), \ldots, E(x, f_t(y)), y) \\ \mathbf{R}(x,y) &= \frac{1}{M} \cdot \mathrm{Count}(y, E(x, f_1(y)), \ldots, E(x, f_t(y))) \\ \mathbf{Q}(x,y) &= \mathbf{L}(x,y) - \mathbf{R}(x,y) \\ \overline{\mathbf{Q}} &= \frac{1}{KD} \sum_{x \in X, y \in [D]} \mathbf{Q}(x,y). \end{aligned}$$

As we mentioned above, we want to bound $\Pr[\overline{\mathbf{Q}} > \varepsilon]$. Notice that for every $x \in X$ and $y \in [D]$, due to the fact that $f_1, \ldots, f_t$ have no fixed points, we have that $\mathbb{E}[\mathbf{L}(x,y)] = \mathbb{E}[\mathbf{R}(x,y)]$ and thus $\mathbb{E}[\overline{\mathbf{Q}}] = 0$. However, the values of $\mathbf{Q}$ on different inputs are not independent.

To see why the $\mathbf{Q}$-s are not independent, think for example about the case where $t = 2$ and y is such that $f_2(f_1(y)) = y$. In such a scenario,

$$\begin{aligned} \mathbf{L}(x,y) &= \mathcal{D}(E(x,y), E(x, f_1(y)), E(x, f_2(y)), y) \\ \mathbf{L}(x, f_1(y)) &= \mathcal{D}(E(x, f_1(y)), E(x, f_1(f_1(y))), E(x,y), f_1(y)), \end{aligned}$$

so, depending on $\mathcal{D}$, $\mathbf{Q}(x,y)$ and $\mathbf{Q}(x, f_1(y))$ may not be independent. Luckily, it is sufficient to disregard such cycles in order to obtain sufficient "independence".

Let $G = (V = [D], E)$ be a directed graph (multiple edges allowed) such that

$$E \;=\; \{(y, f_k(y)) \mid y \in [D], k \in [t]\}\,,$$

so the out-degree of every vertex is exactly t.

Lemma 4.3. *Assume that there exists a subset $V' \subseteq V$ such that the induced subgraph $G' \subseteq G$ is acyclic. Then, the set $\{\mathbf{Q}(x,y)\}_{x\in X, y\in V'}$ can be enumerated by $\mathbf{Q}_1, \ldots, \mathbf{Q}_{m=K|V'|}$ such that*

$$\mathbb{E}[\mathbf{Q}_i \mid \mathbf{Q}_1, \ldots, \mathbf{Q}_{i-1}] \;=\; 0$$

for every $i \in [m]$.

Proof: G' is acyclic so it induces a partial order on V'. Use this partial order to induce a total order on $\{1, \ldots, m\}$ such that if $(y, y') \in E$ and $\mathbf{Q}_j = \mathbf{Q}(x, y')$, $\mathbf{Q}_i = \mathbf{Q}(x, y)$ then $j \leq i$.

Fix some $i \in [m]$ and assume $\mathbf{Q}_i = \mathbf{Q}(x, y)$. The key point is that the variables $\mathbf{Q}_1, \ldots, \mathbf{Q}_{i-1}$ never query E on the input (x, y). Conditioned on any choice of the value of E for all points other than (x, y), denote them by $e_1, \ldots, e_t$, we have that

$$\mathbb{E}[\mathbf{Q}_i] \;=\; \mathbb{E}\left[\mathcal{D}(E(x,y), e_1, \ldots, e_t, y) - \frac{1}{M} \cdot \mathrm{Count}(y, e_1, \ldots, e_t)\right] \;=\; 0,$$

and as we noted, $\mathbf{Q}_1, \ldots, \mathbf{Q}_{i-1}$ are deterministic functions of E and independent of $E(x, y)$. ∎

We now need a partition of the vertices of G into acyclic induced subgraphs. The following lemma shows that such a partition exists with a small number of sets.

Lemma 4.4 ([NL82, Corollary 4])**.** *For any directed graph $G = (V, E)$ with maximum out-degree t (multiple edges allowed), there exists a partition $V = V_1 \cup \ldots \cup V_{t+1}$ such that for every $i \in [t+1]$, the subgraph of G induced by V_i is acyclic.*

In light of the above two lemmas, there exists a partition of $\{\mathbf{Q}(x,y)\}_{x\in X, y\in[D]}$ to $t+1$ sets $\left\{\mathbf{Q}^1_1, \ldots, \mathbf{Q}^1_{s_1}\right\}, \ldots, \left\{\mathbf{Q}^t_1, \ldots, \mathbf{Q}^t_{s_t}\right\}$ such that for every $k \in [t+1]$ and $i \in [s_k]$, $\mathbb{E}[\mathbf{Q}^k_i \mid \mathbf{Q}^k_1, \ldots, \mathbf{Q}^k_{i-1}] = 0$. Now, define $S^k_i = \sum_{j=1}^{i} \mathbf{Q}^k_j$ and note that every sequence $S^k_1, \ldots, S^k_{s_k}$ is a martingale. Also, $|S^k_i - S^k_{i-1}| = |\mathbf{Q}^k_i| \leq 1$ with probability 1. Thus, using Azuma's inequality,

$$\begin{aligned}
\Pr[\overline{\mathbf{Q}} > \varepsilon] \;&=\; \Pr\left[\sum_{k=1}^{t+1} S^k_{s_k} > \varepsilon K D\right] \;\leq\; \sum_{k=1}^{t+1} \Pr\left[S^k_{s_k} > \frac{\varepsilon K D}{t+1}\right] \\
&\leq\; \sum_{k=1}^{t+1} \exp\left(-\frac{\left(\frac{\varepsilon K D}{t+1}\right)^2}{2 \cdot s_k}\right) \;\leq\; (t+1)e^{-\frac{\varepsilon^2 K D}{2(t+1)^2}},
\end{aligned}$$

where the last inequality follows from the fact that $s_k \leq KD$.

To complete our analysis, we require E to work for *any* X, $f_1, \ldots, f_t$ and $\mathcal{D}$. By the union bound, the probability for a random E to fail, denote it by p_E, is given by

$$\begin{aligned}
p_E \;&\leq\; \binom{N}{K} D^{tD} 2^{D \cdot M^{t+1}} (t+1) e^{-\frac{\varepsilon^2 K D}{2(t+1)^2}} \\
&\leq\; 2^{K \log\left(\frac{Ne}{K}\right) + tDd + DM^{t+1} + \log(t+1) - \frac{\varepsilon^2 K D \log e}{2(t+1)^2}} \\
&\leq\; 2^{K(n-k+2) + tDd + DM^{t+1} + \log(t+1) - \frac{\varepsilon^2 K D}{2(t+1)^2}}.
\end{aligned}$$

To prove that $p_E < 1$ (in fact this will show $p_E \ll 1$) it is sufficient to prove that:

1. $K(n-k+2) \leq \frac{\varepsilon^2 KD}{8(t+1)^2}$.

2. $D(td + M^{t+1}) + \log(t+1) \leq \frac{\varepsilon^2 KD}{8(t+1)^2}$, or alternatively $D(2td + M^{t+1}) \leq \frac{\varepsilon^2 KD}{8(t+1)^2}$.

Item (1) is true whenever

$$D \geq \frac{8(t+1)^2(n-k+2)}{\varepsilon^2}.$$

Item (2) is true whenever

$$K \geq \frac{8(t+1)^2(2td + M^{t+1})}{\varepsilon^2}.$$

The bounds on d and k follow from the above two inequalities. ∎

References

[Abb72] HL Abbott. Lower bounds for some ramsey numbers. *Discrete Mathematics*, 2(4):289–293, 1972.

[Alo98] Noga Alon. The shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998.

[BADTS16] Avraham Ben-Aroya, Dean Doron, and Amnon Ta-Shma. Explicit two-source extractors for near-logarithmic min-entropy. *ECCC*, 2016.

[Bar06] Boaz Barak. A simple explicit construction of an $n^{\tilde{o}(\log n)}$-ramsey graph. *arXiv preprint math/0601651*, 2006.

[BKS+10] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *Journal of the ACM (JACM)*, 57(4):20, 2010.

[Bou05] Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[BRSW12] 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. *Annals of Mathematics*, 176(3):1483–1544, 2012.

[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.

[CGL15] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. *arXiv preprint arXiv:1505.00107*, 2015.

[Chu81] Fan RK Chung. A note on constructive methods for ramsey numbers. *Journal of Graph Theory*, 5(1):109–113, 1981.

[CL16] Eshan Chattopadhyay and Xin Li. 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*, pages 158–167. IEEE, 2016.

[Coh15] Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved ramsey graphs. *arXiv preprint arXiv:1506.04428*, 2015.

[Coh16a] Gil Cohen. Local correlation breakers and applications to three-source extractors and mergers. *SIAM Journal on Computing*, 45(4):1297–1338, 2016.

[Coh16b] Gil Cohen. Making the most of advice: New correlation breakers and their applications. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 188–196. IEEE, 2016.

[Coh16c] Gil Cohen. Non-malleable extractors-new tools and improved constructions. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 50. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[Coh16d] Gil Cohen. Two-source extractors for quasi-logarithmic min-entropy and improved privacy amplification protocols. In *ECCC*, 2016.

[CS17] Gil Cohen and Igor Shinkar. Personal communication, 2017.

[CZ16] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 670–683. ACM, 2016.

[DW09] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 601–610. ACM, 2009.

[Fra77] Peter Frankl. A constructive lower bound for ramsey numbers. *Ars Combinatoria*, 3(297-302):28, 1977.

[FW81] Peter Frankl and Richard M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981.

[Gro01] Vince Grolmusz. Low rank co-diagonal matrices and Ramsey graphs. *Journal of combinatorics*, 7(1):R15–R15, 2001.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 68–80. IEEE, 1988.

[Li15a] Xin Li. Improved constructions of two-source extractors. *arXiv preprint arXiv:1508.01115*, 2015.

[Li15b] Xin Li. Three-source extractors for polylogarithmic min-entropy. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 863–882. IEEE, 2015.

[Li16] Xin Li. Improved non-malleable extractors, non-malleable codes and independent source extractors. *arXiv preprint arXiv:1608.00127*, 2016.

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

[Nag75] Zs Nagy. A constructive estimation of the ramsey numbers. *Mat. Lapok*, 23:301–302, 1975.

[Nao92] Moni Naor. Constructing ramsey graphs from small probability spaces. *IBM Research Report RJ*, 8810, 1992.

[NL82] Victor Neumann-Lara. The dichromatic number of a digraph. *Journal of Combinatorial Theory, Series B*, 33(3):265–270, 1982.

[Raz05] Ran Raz. Extractors with weak random seeds. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 11–20. ACM, 2005.

[RTS00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[Zuc06] David Zuckerman. 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*, pages 681–690. ACM, 2006.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il