A New Approach for Constructing Low-Error, Two-Source Extractors

Avraham Ben-Aroya[1]
The Blavatnik School of Computer Science, Tel-Aviv University
Tel Aviv 69978, Israel

Eshan Chattopadhyay[2]
Department of Computer Science, Cornell University and School of Mathematics, IAS
Ithaca, NY 14850, USA; Princeton, NJ 08540, USA
eshanc@ias.edu

Dean Doron[3]
The Blavatnik School of Computer Science, Tel-Aviv University
Tel Aviv 69978, Israel
deandoron@mail.tau.ac.il

Xin Li[4]
Department of Computer Science, Johns Hopkins University
Baltimore, MD 21218, USA
lixints@cs.jhu.edu

Amnon Ta-Shma[5]
The Blavatnik School of Computer Science, Tel-Aviv University
Tel Aviv 69978, Israel
amnon@tau.ac.il

Abstract

Our main contribution in this paper is a new reduction from explicit two-source extractors for polynomially-small entropy rate and negligible error to explicit t-non-malleable extractors with seed-length that has a good dependence on t. Our reduction is based on the Chattopadhyay and Zuckerman framework (STOC 2016), and surprisingly we dispense with the use of resilient functions which appeared to be a major ingredient there and in follow-up works. The use of resilient functions posed a fundamental barrier towards achieving negligible error, and our new reduction circumvents this bottleneck.

The parameters we require from t-non-malleable extractors for our reduction to work hold in a non-explicit construction, but currently it is not known how to explicitly construct such extractors. As a result we do not give an unconditional construction of an explicit low-error two-source extractor. Nonetheless, we believe our work gives a viable approach for solving the important problem of low-error two-source extractors. Furthermore, our work highlights an existing barrier in constructing low-error two-source extractors, and draws attention to the dependence of the parameter t in the seed-length of the non-malleable extractor. We hope this work would lead to further developments in explicit constructions of both non-malleable and two-source extractors.

[1] Supported by the Israel Science Foundation grant no. 994/14.

[2] Supported by NSF grants CCF-1526952, CCF-1412958 and the Simons Foundation. Part of this work was done when the author was a graduate student in UT Austin and while visiting the Simons Institute for the Theory of Computing at UC Berkeley.

[3] Supported by the Israel Science Foundation grant no. 994/14. This work was done in part while visiting the Simons Institute for the Theory of Computing at UC Berkeley.

[4] Supported by NSF Grant CCF-1617713.

[5] Supported by the Israel Science Foundation grant no. 994/14. This work was done in part while visiting the Simons Institute for the Theory of Computing at UC Berkeley.

33rd Computational Complexity Conference (CCC 2018).
Editor: Rocco A. Servedio; Article No. 3; pp. 3:1–3:19
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases Two-Source Extractors, Non-Malleable Extractors, Pseudorandomness, Explicit Constructions

Digital Object Identifier 10.4230/LIPIcs.CCC.2018.3

Acknowledgements We are grateful to Gil Cohen for discussions regarding [20] in the early stages of our work.

1 Introduction

A two-source extractor hashes samples from two *independent* weak sources into one output whose distribution is close to uniform. Formally, we say a distribution X is an (n, k) source if X is distributed over $\{0,1\}^n$ and its min-entropy is at least k (i.e., all strings in its support have probability mass at most 2^{-k}). An $((n_1, k_1), (n_2, k_2), \varepsilon)$ two-source extractor is a function $E\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ that maps any pair of independent (n_1, k_1) and (n_2, k_2) sources X_1, X_2 to a distribution $E(X_1, X_2)$ which is ε-close to U_m, the uniform distribution over $\{0,1\}^m$.

Non-explicitly there are $((n, k), (n, k), \varepsilon)$ two-source extractors as long as $k \geq \log n + 2\log(\frac{1}{\varepsilon}) + O(1)$. More generally,

▶ **Fact 1.** *Assume $k_1 + k_2 \geq \log(2^{k_1} n_1 + 2^{k_2} n_2) + 2\log(\frac{1}{\varepsilon}) + O(1)$. Then, there exists a (non-explicit) $((n_1, k_1), (n_2, k_2), \varepsilon)$ two-source extractor $E\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$.*

Finding such *explicit* constructions is a long-standing, important and challenging problem. A key parameter is the error ε obtained by the two-source extractor. Research in the area can be divided into three regimes:

Very large error: Finding explicit two-source extractors with any error smaller than 1 (i.e., any non-trivial error) is already very challenging and is essentially equivalent to finding an explicit *bipartite* Ramsey graph. A K *Ramsey graph* is a graph that contains no monochromatic set (i.e., a clique or an independent set) of size K; a K *bipartite Ramsey graph* is a bipartite graph with no bipartite monochromatic sets of size K. A $K = 2^k$ bipartite Ramsey graph over $2N = 2 \cdot 2^n$ vertices, is essentially equivalent to an $((n, k), (n, k), \varepsilon)$ two-source extractor, with $\varepsilon = \varepsilon(n) < 1$.

A long line of research was devoted to explicitly constructing Ramsey graphs [1, 30, 22, 12, 23, 31, 2, 24, 3], bipartite Ramsey graph [4, 5, 17], and two-source extractors [11, 34, 7]. Two years ago, Cohen [17] constructed a K bipartite Ramsey graph over $2N$ vertices with $\log K = \operatorname{polylog}(\log N)$. This corresponds to an $((n, k), (n, k), \varepsilon)$ two-source extractor, with $k = \operatorname{polylog} n$ and some non-trivial error ε. Independently, Chattopadhyay and Zuckerman [10] gave another construction that gives about the same bipartite Ramsey graphs, but with smaller error. We discuss this next.

Medium size error: Chattopadhyay and Zuckerman constructed an efficient $((n, k), (n, k), \varepsilon)$ two-source extractor, with $k = \operatorname{polylog} n$ and running time polynomial in $1/\varepsilon$. Several improvements followed, including [29, 27]. Currently, following [6, 18, 28], the best explicit construction achieves $k = O(\log n \log\log n)$ which is pretty close to the optimal $\Omega(\log n)$ bound.

All these constructions have running time which is at best polynomial in $1/\varepsilon$, and as we explain below this seems to be inherent to the approach that is taken. In contrast, 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.

Exponentially small error: There are several explicit two-source extractors constructions with exponentially small error:

1. The inner-product function gives a simple construction when $k > n/2$ [11].
2. Bourgain [7] gave a two-source extractor construction for $k = \left(\frac{1}{2} - \alpha\right) n$, for some small constant $\alpha > 0$.
3. Raz [34] constructed an $((n_1, k_1), (n_2, k_2), \varepsilon)$ two-source extractor that has an unbalanced entropy requirement; the first source is long (of length n_1) and very weak (k_1 can be as small as $\log\log n_1 + O(1)$), the second source is short (of length $O(\log n_1)$) and somewhat dense with $k_2 \geq \alpha n_2$, for some constant $\alpha > \frac{1}{2}$.

On the positive side, all of these constructions have exponentially small error (in Raz's extractor, the error is exponentially small in the smaller entropy). On the negative side, 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 $\left(\frac{1}{2} - \alpha\right) n > 0.49n$.

To summarize:

- Current explicit constructions of low-error, two-source extractors require one source to have entropy rate close to half, and,
- There are explicit two-source extractors that work with astonishingly small min-entropy, but currently they only handle large error, or, more precisely, their running time is polynomial in $1/\varepsilon$.

As we shall see shortly, there is a good reason for the two barriers that are represented in the above two items. The goal of this paper is to present a new approach for bypassing these barriers.

1.1 Extractors and Entropy-Rate Half

Let us start with the rate-half barrier for low-error constructions. For that we compare two-source extractors with *strong seeded extractors.*

▶ **Definition 2.** $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong (k, ε) extractor if for every (n, k) source X, $(Y, E(X, Y))$ is ε-close to $Y \times U_m$, where Y is uniformly distributed over $\{0,1\}^d$ and is independent of X.

A seeded extractor E must have seed length $d \geq \log n + 2\log(\frac{1}{\varepsilon}) - O(1)$ [33]. In essence, the error of a *seeded* extractor has two origins:

- The fraction ε_1 of bad seeds for which $E(X, y)$ is ε_2-far from uniform, and,
- The distance ε_2 between $E(X, y)$ and U_m for good seeds.

These two errors can be very different, for example, it might be the case that for half the seeds the error is extremely small, and then ε_1 is constant and ε_2 is tiny, or vice versa. In the terminology of a seeded extractor, these two errors are unified to one parameter ε. In the two-source extractor notation these two errors are essentially *separated*, where 2^{k_2} is, roughly, the number of bad seeds making $\varepsilon_1 \approx 2^{k_2 - n_2}$, where ε of the two-source extractor represents the ε_2 above. More formally:

▶ **Fact 3.** *Suppose $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}$ is an $((n,k),(d,d'),\varepsilon_2)$ two-source extractor. Then, E is a strong $(k, \varepsilon = \varepsilon_1 + \varepsilon_2)$ extractor, for $\varepsilon_1 = 2^{d'+1-d}$, and furthermore, for every (n,k) source X,*

$$\Pr_{y \in \{0,1\}^d}[E(X,y) \not\approx_{\varepsilon_2} U_1] \ \le \ \varepsilon_1.$$

Proof. Let X be an (n,k) source and let $B \subseteq \{0,1\}^d$ so that for every $y \in B$, $E(X,y) \not\approx_{\varepsilon_2} U_1$. Partition $B = B_0 \cup B_1$ where $y \in B_z$ if the ε_2 bias is towards z. Assume towards contradiction that $|B_z| \ge 2^{d'}$ for some z and consider the flat distribution Y over the set B_z. Thus, $H_\infty(Y) \ge d'$ so $E(X,Y) \approx_{\varepsilon_2} U_1$ but by our definition, $E(X,Y)$ is biased towards z – a contradiction. Altogether, $|B| \le 2^{d'+1}$ so $\varepsilon_1 \le |B|/2^d = 2^{d'+1-d}$. ◀

The lower bound $d \ge \log n + 2\log(\frac{1}{\varepsilon}) - O(1)$ imposed on extractors, does not reveal which of the two errors forces d to be large. Stating it more precisely, define a $(k, \varepsilon_1, \varepsilon_2)$ function $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ so that for every (n,k) source X, $\Pr_{y \in \{0,1\}^d}[E(X,y) \not\approx_{\varepsilon_2} U_1] \le \varepsilon_1$. What is the dependence of d on ε_1 and ε_2?

The existence of $((n,k),(d=n, d'=O(\log n)),\varepsilon)$ two-source extractors, implies that the dependence of d on ε_1 might be very close $1 \cdot \log \frac{1}{\varepsilon_1}$. On the other hand, the dependence of d on ε_2 is larger, $d \ge d' \ge 2\log\frac{1}{\varepsilon_2}$, since we can view E as a strong (d', ε_2) extractor $\{0,1\}^d \times \{0,1\}^k \to \{0,1\}$ and $d' \ge 2\log\frac{1}{\varepsilon_2}$ is again a lower bound [33]. Thus, the two-source extractor terminology allows a finer characterization of the quality of an extractor, separating the two errors ε_1 and ε_2 above.

Looking at it that way we see why rate-half is a natural barrier: An extractor with seed length dependence $2\log(\frac{1}{\varepsilon})$ guarantees that out of the $D = 2^d$ possible seeds, at most $D^{\frac{1}{2}+\beta}$ are $D^{-\beta}$ bad. Thus, one can get an explicit two-source extractor, where the seed has some constant density $\frac{1}{2} + \beta$, and exponentially small error, by constructing an explicit strong seeded extractor with seed length dependence $(2+\gamma)\log(\frac{1}{\varepsilon})$ for some small constant γ. Constructing a two-source extractor with d'/d below half necessarily means using techniques that do not apply to strong seeded extractors. Bourgain achieves that in an ingenious way, by using additive combinatorics together with the inner product function, but, at least so far, this approach can only handle min-entropies slightly below half.

1.2 The CZ Approach

We now explain the main ideas in the construction of the two-source extractor of [10] and the bottleneck for achieving smaller error. The CZ construction builds upon two main ingredients: the existence of explicit non-malleable extractors and resilient functions, and we recall both now.

▶ **Definition 4.** $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong (k, ε) *t-non-malleable (n.m.) extractor*, if for every (n,k) source X and every t functions $f_1, \dots, f_t : \{0,1\}^d \to \{0,1\}^d$ with no fixed-points[6] it holds that

$$|(Y, E(X,Y), E(X,f_1(Y)), \dots, E(X,f_t(Y))) - (Y, U_m, E(X,f_1(Y)), \dots, E(X,f_t(Y)))| \ \le \ \varepsilon,$$

where Y is uniformly distributed over $\{0,1\}^d$ and is independent of X and U_m is the uniform distribution over $\{0,1\}^m$.

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

In words and roughly speaking, this means that there are many good seeds, and for a good seed y, $E(X,y)$ is close to uniform even given the value of E on t other seeds $f_1(y),\dots,f_t(y)$ maliciously chosen by an adversary. Said differently, if we build a table with $D = 2^d$ rows, and put $E(X,i)$ in the i-th row, then rows of good seeds are close to uniform, and, furthermore, those good rows are close to being t-wise independent, in the sense that every t good rows are $\approx t\varepsilon$ close to uniform (see Lemma 10).

A *resilient function* is a nearly-balanced function $f\colon \{0,1\}^D \to \{0,1\}$ whose output cannot be heavily influenced by any small set of q "bad" bits. We think of the bad bits as a coalition of malicious players trying to bias the output *after seeing* the $D-q$ coin tosses of the honest players (the honest players toss independent random coin). The function f is (q,t) resilient if it is resilient even when there are q bad players and even when the honest players are only t-wise independent.

Now, let X_1 and X_2 be two independent (n,k) sources. The starting point of [10] is to use a t-non-malleable extractor E with error ε_1 and seed length d_1 to produce a table T_1 with $D_1 = 2^{d_1}$ entries, where the i-th entry is $E(X_1,i)$. Using the property of the non-malleable extractor, one can show that $(1-\sqrt{\varepsilon_1})$-fraction of the rows are uniform and almost t-wise independent (in the sense that any t good rows are close to uniform). The remaining rows are, however, arbitrarily correlated with those rows. Then, they

- Use the second source X_2 to sample a sub-table T_2 with some D_2 rows of the table T_1, such that a fraction of at most ε_2 of its rows are bad, and every t good rows are $\sqrt{\varepsilon_1}$-close to uniform, and,
- Apply a resilient function $f : \{0,1\}^{D_2} \to \{0,1\}$ on the sub-table T_2. f has to be resilient against $\sqrt{\varepsilon_2}D_2$ bad players, and should perform correctly even when the good players are t-wise independent.

It turns out that the sub-table T_2 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, and this proves the correctness of the construction.

While this beautiful approach does give an unbiased output bit, it seems that it is inherently bound to have running time polynomial in $1/\varepsilon$. This is because no matter which resilient function we use, even if there is just a single bad player among the D_2 players, then that player alone may have $1/D_2$ influence over the result (in fact, [25] showed there is a player with $\Omega(\frac{\log D_2}{D_2})$ influence) and therefore that player can bias the result by $1/D_2$. Thus, the running time, which is at least D_2, is at least $\Omega(\frac{1}{\varepsilon})$, and this is indeed a common feature of all the constructions so far that use the CZ approach.

One could have hoped to sample a sub-table T_2 that w.h.p. avoids *all* bad players, thus dispensing with the use of the resilient function. This approach is futile: If T_2 avoids all bad players then every row y of it will do, so indeed $E(X,y)$ is close to uniform and we can compute it fast, allowing for a small error. However, this brings us back to the seeded extractors case, and we already saw this cannot handle densities above half.

1.3 Our Main Result

The main result in the paper is a reduction showing how to explicitly construct low-error two-source extractors given explicit t-non-malleable extractors with small seed length dependence on t. Formally,

▶ **Theorem 5.** *Suppose for some constant $\alpha > 0$ for every n_1, k_1, ε_1 and t there exists an explicit function*

$$E\colon \{0,1\}^{n_1} \times \{0,1\}^d \to \{0,1\}^m$$

that is a strong (k_1, ε_1) *t-non-malleable extractor with* $d \leq \alpha t \cdot \log(\frac{1}{\varepsilon_1})$.
Then, there exists an explicit function

$$F \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$$

that is a $((n_1, k_1), (n_2 = O(\frac{d}{\alpha}), k_2 = O(\alpha n_2)), 2\sqrt{\varepsilon_1})$ *two-source extractor, where the constants hidden in the big-O notation are independent of* α.

We first remark that such non-malleable extractors non-explicitly exist. In fact, much better parameters are possible:

▶ **Theorem 6.** *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* $E \colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $d \leq 2\log\frac{1}{\varepsilon} + \log(n-k) + 2\log(t+1) + 3$.

The proof of the Theorem is based on [21], where they only handle the $t = 1$ case. The Theorem was also independently proved by Cohen and Shinkar [20]. For completeness we give the proof in Appendix A.

The currently best explicit construction of t-n.m. extractors is due to Li:

▶ **Theorem 7** ([28])**.** *For any integer* n, t *and* $\varepsilon > 0$, *there exists an efficiently-computable function*

$$\mathsf{nmEXT} \colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}$$

that is a strong $(k = d, t\varepsilon)$ *t-non-malleable extractor with seed length* $d = O(t^2(\log n + \log\frac{1}{\varepsilon} \cdot \log\log\frac{1}{\varepsilon}))$.

So far the main focus in explicit constructions of t-non-malleable extractors has been getting an optimal seed length dependence on n and ε. Thus, Chattopadhyay et al. has $d = \log^2(\frac{n}{\varepsilon})$ [8] and this has been improved in [15, 16, 9, 14] with the current best construction being Theorem 7 of [28] with $d = O(\log n + \log\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon})$. However, in all these constructions t is treated as a constant. In fact, Cohen [15, Lemma 2.5] proved that if one constructs a n.m. extractor for $t = 1$ then an explicit construction for t follows at the cost of multiplying the seed by a t^2 multiplicative factor.

There is a huge gap between the dependence of the seed length on t in the non-explicit construction of Theorem 6, where t contributes an *additive* $2\log t$ factor to the seed length, and the explicit Theorem 7 where t contributes a *multiplicative* t^2 factor to the seed length.[7] Correspondingly, the quality of the two source construction we give significantly improves with a better dependence of the seed on the parameter t. In Table 1 we list the two-source extractors constructions we get for:

- The current best explicit constructions (we get nothing),
- A quadratic improvement over currently best explicit (we improve upon Raz's extractor), and,
- A further polynomial improvement.

The parameters in the second row (and Theorem 5) resemble those of Raz's extractor: one source is long with very low entropy, the other is short with constant entropy rate. The

[7] It is worth mentioning that an early construction of Cohen, Raz and Segev [19], although not explicitly stating it, does get a very good dependence of d on t with $d = O(\log\frac{n}{\varepsilon} + t)$. However, their construction only works for high min entropy and so does not imply a two-source extractor for densities below half.

Table 1 Bounds for $((n,k),(n_2,k_2),\varepsilon)$ two-source extractors assuming an explicit t n.m. extractor with various seed length d dependence on t. In all cases, the error ε is low.

Dependence on t	k	n_2	k_2	
$\omega\left(t\log\frac{1}{\varepsilon}\right)$				The approach fails
$\alpha t\log(\frac{1}{\varepsilon})$	arbitrary	$O(\frac{d}{\alpha})$	$O(\alpha)n_2$	α is any constant
$t^{\alpha}\log(\frac{1}{\varepsilon})$ or better	arbitrary	$\mathrm{poly}_{\alpha,\beta}(d)$	n_2^{β}	For some constants $\alpha,\beta<1$
$t^{\alpha}\log(\frac{1}{\varepsilon})$ or better	small enough	n	n^{β}	For some constant $\beta<1$

main difference is that in Raz's extractor the entropy rate has to be above half, whereas here, assuming the existence of the appropriate explicit non-malleable extractors, the entropy rate can be an arbitrarily small constant.

By allowing the seed-length of the n.m. extractor to have an even better dependence on t (and non-explicitly it does), we succeed in supporting polynomially-small min-entropies. More specifically, if the seed length dependence on t is $t^{\alpha}\log(\frac{1}{\varepsilon})$ for a small enough constant α, then we can support min-entropy of $k_2=n_2^{\beta}$ where $\beta=\beta(\alpha)$ is another constant.

Also, in that regime of dependence, we can set the error ε to be small enough so that $n_2=n$, in which case we get a *balanced* two-source extractor supporting some polynomially-small min-entropy (see Corollary 19).

We believe this clearly demonstrates that the dependence of the seed length on t in non-malleable extractors is directly related to the required density of the seed (i.e., second source) in low-error, two-source constructions. We believe this understanding is an important, qualitative understanding. We believe our work is the first to draw attention to this important question and we hope it will facilitate further research on achieving the correct dependence of the seed on the non-malleability parameter t.

1.4 Our Technique

In the CZ construction we have the following ingredients:

1. The use of the first source to construct a table with many good rows (every row in the table corresponds to applying an extractor on the first source, with some fixed seed).
2. The use of t-non-malleable extractors to get *local* t-wise independence, where every t good rows are close to uniform.
3. The use of the second source to sample a sub-table of the table constructed from the first source.
4. The realization that with the right choice of parameters the sub-table is *globally* close to a table where the good rows are perfectly t-wise.
5. The use of resilient functions.

In our solution we keep (1)-(3) and completely dispense with (4) and (5), i.e., we do not use resilient functions and we do not try to achieve a sub-table that is *globally* close to a *truly* t-wise independent distribution. Instead, we work with the much weaker *local* guarantee that every t good rows are close to uniform.

Thus, our construction is as follows. We are given two samples from independent sources $x_1\sim X_1$ and $x_2\sim X_2$. Then:

1. We use a t-non-malleable extractor E with error ε_1 and seed length d_1 to construct a table with $D_1=2^{d_1}$ entries, where the i-th entry is $E(X_1,i)$. Using the property of

non-malleable extractors one can show that $(1 - \sqrt{\varepsilon_1})$-fraction of the rows are good in the sense that a good row is close to uniform even conditioned on $t-1$ other rows. The remaining rows are arbitrarily correlated with the good ones. So far, everything is identical to the [10] construction.

2. We use the second sample x_2 to sample t rows from that table, with the property that with high probability (over the choice of $x_2 \sim X_2$) **at least one of the t samples is a good row** (in the table with D_1 rows).
   We note that this is very different from the [10] construction, where the requirement is that with high probability (over the choice of $x_2 \sim X_2$) the fraction of bad rows in the sub-table is about the same as the fraction of bad rows in the original table.
3. We then take the *parity* of the t strings written in the t rows we sampled.
   This is again very different from the [10] construction, where a resilient function is applied on the sub-table (and notice that the parity function is not resilient at all).

Conceptually, what happened is that we take a *dramatically smaller* sample set than before. Specifically, in [10, 6] 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 every good player (and even if there is just a single 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 dispense with the resilient function f and just use the parity function instead.

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 locally close to uniform, to a table that is globally close to being perfectly t-wise independent.

There is still a fundamental question we need to answer. Inspecting the argument, we see that 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 whether such a construction is possible at all even assuming the best possible non-malleable extractors.

The above inquiry 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 [20] independently investigated this. As we explained before, 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, non-malleable extractors with seed length below $t \log(\frac{n}{\varepsilon})$ work well, while non-malleable extractors with seed length above it do not. In Section 3 we demonstrate how the dependence of the seed length d on t affects the parameters of the two-source extractor construction.

Finally, we are left with two questions regarding *explicitness*:

- We ask whether the sampler can be made explicit, i.e., 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* that are good against small tests. Remarkably, Zuckerman [35] gave a beautiful explicit construction with nearly optimal bounds, and we show the dispersers he constructed work well for us.

- Current explicit constructions of non-malleable extractors [13, 8, 15, 14, 9, 18, 28] for small entropies are above that threshold. This is mainly due to the use of alternating extraction techniques which treat the seed and the source symmetrically. Thus, this paper raises the challenge of explicitly constructing non-malleable extractors with better seed length dependence on t.

We believe identifying the connection between the seed length dependence on t and low error, two-source extractors is important on its own, and is a major contribution of the paper. We hope this work would lead to further developments in explicit constructions of both non-malleable and two-source extractors.

1.5 Related work

Li [26] showed how to build a $((n, 0.499n), (n, k), 2^{-\Omega(n)})$ two-source extractor assuming a 1-non-malleable extractor with seed-length $d = 2\log(1/\varepsilon) + o(n)$. Li's work is orthogonal to ours. First, it asks for small seed dependence on the error: the seed-length of the non-malleable extractor has to be at most 2.001, while we look on the dependence on t. Also, it achieves limited parameters (even assuming non-explicit constructions) that are close to those in Bourgain's construction, and it is also close in spirit to Bourgain's construction.

As we said before, we believe our work reveals an intrinsic connection between the dependence of the seed length of a non-malleable extractor on the non-malleability parameter t and the quality of low-error two-source extractors, and is the first work to draw attention to the important problem of the dependence of the seed length on t in explicit construction. We hope, and believe, this approach may lead to getting better explicit, low-error, two source extractors, which is a fundamental problem and a long standing barrier in TCS.

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 $f \colon D_1 \to D_2$ 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 $f \colon D_1 \to D_2$ 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 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 8.** A function $\mathsf{2Ext}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ 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 9.** $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong (k,ε) *t-non-malleable (n.m.) extractor*, if for every (n,k) source X and every functions $f_1,\ldots,f_t\colon [D] \to [D]$ with no fixed-points it holds that,

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

where Y is uniformly distributed over $\{0,1\}^d$ and is independent of X.

A simple consequence, proved in [10], is:

▶ **Lemma 10** ([10], Lemma 3.4). *Let $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ 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 $y'_1,\ldots,y'_t \in [D] \setminus y$,*

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

2.3 Dispersers

▶ **Definition 11.** A function $\Gamma\colon [N] \times [D] \to [M]$ 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 12** ([35], Theorem 1.9). *There exists a constant c_{disp} such that the following holds. For every constants $0 < a,b < 1$, every N, $K = N^a$, $M \leq K^{1-b}$ and $K' < M$ there exists an efficient family of (K,K') dispersers*

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

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

The parameters in Theorem 12 are tight up to a constant factor:

▶ **Theorem 13** ([33], Theorem 1.5). *There exists a constant c_0 such that the following holds. Let $\Gamma\colon [N] \times [D] \to [M]$ be a (K,K') disperser where $K < N$ and $K' < M/2$. Then, $D \geq c_0 \cdot \frac{\log \frac{N}{K}}{\log \frac{M}{K'}}$.*

3 The Construction

3.1 The Overall Structure

Given:

$$E\colon \{0,1\}^{n_1} \times [D] \to \{0,1\}^m$$
$$\Gamma\colon \{0,1\}^{n_2} \times [t+1] \to [D]$$

We define $\mathsf{2Ext}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ by

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

▶ **Theorem 14.** *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 10 there exists a set $BAD_1 \subseteq [D]$ with $\rho(BAD_1) \le \sqrt{\varepsilon_1}$ such that for every $y \notin BAD_1$ and every $y'_1,\dots,y'_t \in [D] \setminus \{y\}$,

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

Let $BAD_2 \subseteq [N_2]$ be

$$BAD_2 \;=\; \{x_2 \in \{0,1\}^{n_2} : \Gamma(x_2) \subseteq BAD_1\}\,.$$

Thus, $\Gamma(BAD_2) \subseteq BAD_1$. Since $|BAD_1| \le \sqrt{\varepsilon_1}D$ and Γ_2 is a $(B_2,\sqrt{\varepsilon_1}D)$ disperser, it follows that $|BAD_2| \le B_2$. However, for any $x_2 \in \{0,1\}^{n_2} \setminus BAD_2$, there exists an $i \in [t+1]$ such that $y = \Gamma(x_2,i) \notin BAD_1$. Hence,

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

Thus,

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

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 $t \ge D_2 - 1$,
- The non-malleability parameter t affects the degree $2^d = D$ 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 a 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. However, at least 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.

3.3 The analysis fails when $d \geq ct\log(\frac{1}{\varepsilon})$ for some constant c

▶ **Lemma 15.** *Suppose*

$$E\colon \{0,1\}^{n_1} \times [D] \to \{0,1\}^m$$
$$\Gamma\colon \{0,1\}^{n_2} \times [t+1] \to [D]$$

are such that E is a strong (k_1, ε_1) t-n.m. extractor and Γ is any $(B_2, B_1 = \sqrt{\varepsilon_1}D)$ disperser, as required by Theorem 14. Suppose Theorem 14 gives that

$$\mathsf{2Ext}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$$

is an $((n_1, k_1), (n_2, k_2), 2\sqrt{\varepsilon_1})$ two-source extractor with $K_2 < \sqrt{N_2}$. Then, $\log_{1/\varepsilon_1} D \leq \frac{t+1}{c_0}$, where c_0 is the constant guaranteed by Theorem 13.

Proof. We first give some easy bounds on the parameters:

- $B_2 \leq K_2$, for otherwise Theorem 12 constructs 2Ext with the trivial error 1.
- Also, $tB_2 \geq B_1$, for otherwise we can take a set $A \subseteq \{0,1\}^{n_2}$ of cardinality B_2 and the size of its neighbor set is at most $B_2 t < B_1$ violating the disperser property.
- Finally, $\frac{B_1}{t} \geq \sqrt{B_1}$ because otherwise $\sqrt{B_1} < t$ and then

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

where the last inequality follows from the assumption on ε_1. This contradicts the lower-bound for extractors [33].

Together, $\frac{N_2}{B_2} \geq \frac{N_2}{K_2} \geq K_2 \geq B_2 \geq \frac{B_1}{t} \geq \sqrt{B_1} = \sqrt{\varepsilon_1}D$ and $\frac{D}{B_1} = \frac{1}{\sqrt{\varepsilon_1}}$. Now, $\Gamma\colon \{0,1\}^{n_2} \times [t+1] \to [D]$ is a $(B_2, B_1 = \sqrt{\varepsilon_1}D)$ disperser and therefore by Theorem 13 it has degree at least $c_0 \cdot \frac{\log \frac{N_2}{B_2}}{\log \frac{D}{B_1}}$ for some constant c_0. Therefore,

$$\begin{aligned} t+1 &\geq c_0 \cdot \frac{\log \frac{N_2}{B_2}}{\log \frac{D}{B_1}} \geq c_0 \cdot \frac{\log \sqrt{\varepsilon_1}D}{\log \frac{1}{\sqrt{\varepsilon_1}}} \\ &= 2c_0 \cdot \log_{1/\varepsilon_1}(\sqrt{\varepsilon_1}D) = 2c_0 \cdot (\log_{1/\varepsilon_1} D - 1/2) \ \geq \ c_0 \log_{1/\varepsilon_1} D. \end{aligned}$$ ◀

The analysis in the above proof is quite tight and in the next subsection we prove the converse (which also entails Theorem 5).

3.4 When $d = O(t\log(\frac{1}{\varepsilon}))$

▶ **Lemma 16.** *Let $\varepsilon_1 \leq \frac{1}{n}$. Suppose there exists an explicit*

$$E\colon \{0,1\}^{n_1} \times [D_1] \to \{0,1\}^m$$

that is a strong (k_1, ε_1) t-n.m. extractor with $\log_{1/\varepsilon_1} D_1 \leq \frac{\alpha}{8c_{disp}}t$ for some constant $\alpha > 0$, some constant t and some k_1. Then there exists an explicit

$$\mathsf{2Ext}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$$

that is a $((n_1, k_1), (n_2 = \frac{4}{\alpha}d_1, k_2 = \alpha n_2), 2\sqrt{\varepsilon_1})$ two-source extractor.

Proof. Fix t as in the hypothesis of the lemma. Set D such that $\log_{1/\varepsilon_1} D = \frac{\alpha t}{8c_{disp}}$. Let

$$\Gamma \colon [N_2 = D^{4/\alpha}] \times [D_2] \to [D]$$

be the $(B_2 = D^2, B_1 = \sqrt{\varepsilon_1}D)$ disperser promised to us by Theorem 12 for $a = \frac{\alpha}{2}$ (because $B_2 = N_2^a$) and $b = \frac{1}{2}$ (because $D = B_2^b$). By Theorem 12 the degree D_2 of Γ is

$$\begin{aligned} D_2 &= c_{disp} \cdot \frac{\log \frac{N_2}{B_2}}{\log \frac{D}{B_1}} = c_{disp} \cdot \frac{\frac{4(1-a)}{\alpha} \log D}{\log \frac{1}{\sqrt{\varepsilon_1}}} \\ &= c_{disp} \cdot \left(\frac{1}{\alpha} - \frac{1}{2}\right) \frac{8 \log D}{\log 1/\varepsilon_1} = c_{disp} \cdot \left(\frac{8}{\alpha} - 4\right) \log_{1/\varepsilon_1} D \\ &= c_{disp} \cdot \left(\frac{8}{\alpha} - 4\right) \frac{\alpha t}{8c_{disp}} = \left(1 - \frac{\alpha}{2}\right) t < t. \end{aligned}$$

Let

$$E \colon \{0,1\}^{n_1} \times [D_1] \to \{0,1\}^m$$

be the explicit, strong (k_1, ε_1) t-n.m. extractor with $\log_{1/\varepsilon_1} D_1 \le \frac{\alpha}{8c_{disp}} t = \log_{1/\varepsilon_1} D$ promised by the hypothesis of the lemma. As $\frac{1}{\varepsilon} > 1$, we see that $D_1 \le D$ and we may take D_1 larger so that it equals D.

Now let

$$\mathsf{2Ext} \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$$

be constructed from E and Γ as above. As E is a strong (k_1, ε_1) t-n.m. extractor and Γ is a $(B_2, \sqrt{\varepsilon_1} D_1)$ disperser, Theorem 14 tells us that for every k_2, 2Ext is a $((n_1, k_1), (n_2, k_2), \frac{B_2}{K_2} + \sqrt{\varepsilon_1})$ two-source extractor. Taking $k_2 = \alpha n_2$,

$$\frac{B_2}{K_2} + \sqrt{\varepsilon_1} = \frac{D_1^2}{D_1^4} + \sqrt{\varepsilon_1} = \frac{1}{D_1^2} + \sqrt{\varepsilon_1}.$$

But $D_1 \ge \frac{1}{\varepsilon_1^2}$ (this is true for any seeded extractor [33]). Altogether the error is at most $\sqrt{\varepsilon_1} + \frac{1}{\varepsilon_1^2} \le 2\sqrt{\varepsilon_1}$. ◀

3.5 When $d = O(t^\alpha \log(\frac{1}{\varepsilon}))$

A careful examination of the parameters shows that if the dependence of d_1 on t is better, our scheme yields a two-source extractor that supports even smaller min-entropies. Roughly speaking, if $\log_{1/\varepsilon_1} D_1 = t^\alpha$ for some $\alpha < 1$ we can support some polynomially-small min-entropy $k_2 = n_2^\beta$, instead of only supporting min-entropies of constant rate. Specifically:

▶ **Lemma 17.** *Let $\varepsilon_1 \le \frac{1}{n}$. There exists a constant $\beta_0 < 1$ such that for every $\beta_0 < \beta < 1$ there exist constants $\alpha < 1$ and $\gamma > 1$ so that the following holds. Suppose there exists an explicit*

$$E \colon \{0,1\}^{n_1} \times [D_1] \to \{0,1\}^m$$

that is a strong (k, ε_1) t-n.m. extractor with $\log_{1/\varepsilon_1} D_1 \le t^\alpha$ for some k_1, and t which is a large enough polynomial in $\log \frac{1}{\varepsilon_1}$. Then there exists an explicit

$$\mathsf{2Ext} \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$$

that is a $((n_1, k_1), (n_2 = d_1^\gamma, k_2 = n_2^\beta), 2\sqrt{\varepsilon_1})$ two-source extractor.

The proof is similar to the proof of Lemma 16. However, it is no longer true that K_2 is a *constant* power of N_2, so we should be more careful with the parameters of Zuckerman's disperser. Particularly, in this regime of parameters, the degree D_2 (and consequently t) is no longer constant but will be poly-logarithmic in $\frac{1}{\varepsilon}$. The following Theorem extends Theorem 12 for the more general case.

▶ **Theorem 18** ([35], Theorem 1.9)**.** *There exist constants $c_1, c_2 > 1$ such that the following holds. For every $0 < \delta < 1$, N, $K = N^{\delta}$, $M \leq N^{\delta^{c_2}}$ and $K' < M$ there exists an efficient family of (K, K') dispersers*

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

with degree $D = \left(\frac{1}{\delta}\right)^{c_1} \cdot \frac{n}{\log \frac{M}{K'}}$.

We are now ready to prove Lemma 17.

Proof of Lemma 17. Let c_1 and c_2 be as in Theorem 18. Set $\beta_0 = 1 - \frac{1}{c_2}$ and fix some $\beta_0 < \beta < 1$. Fix t as in the hypothesis of the lemma. Set D such that $\log_{1/\varepsilon_1} D = t^{\alpha}$ for $\alpha = \alpha(\beta)$ we will soon explicitly determine. Let

$$\Gamma \colon [N_2 = D^{1/\delta^{c_2}}] \times [D_2] \to [D]$$

be the $(B_2 = N_2^{\delta}, B_1 = \sqrt{\varepsilon_1} D)$ disperser promised to us by Theorem 18, for $\delta = \frac{1}{2} n_2^{-(1-\beta)}$. Notice that $b_2 = \delta n_2 = \frac{1}{2} n_2^{\beta}$ and set $k_2 = 2b_2 = n_2^{\beta}$. Also, observe that $n_2 = \frac{1}{\delta^{c_2}} d = (2^{c_2} d)^{\gamma'}$ for

$$\gamma' \; = \; \frac{1}{1 - c_2(1-\beta)}.$$

As $\beta > \beta_0$ we see that $\gamma' > 1$. It follows that $n_2 = d^{\gamma}$ for some $\gamma' < \gamma < 2\gamma'$.

By Theorem 18, the degree D_2 of Γ is

$$\begin{aligned} D_2 &= \left(\frac{1}{\delta}\right)^{c_1} \cdot \frac{n_2}{\log \frac{D}{B_1}} = \left(\frac{1}{\delta}\right)^{c_1} \cdot \frac{2n_2}{\log(1/\varepsilon_1)} \\ &= \left(2n_2^{1-\beta}\right)^{c_1} \cdot \frac{2 \cdot n_2}{\log(1/\varepsilon_1)} = 2^{c_1+1} \cdot \frac{n_2^{1+c_1(1-\beta)}}{\log(1/\varepsilon_1)} = 2^{c_1+1} \cdot \frac{(\log D)^{\gamma(1+c_1(1-\beta))}}{\log(1/\varepsilon_1)}. \end{aligned}$$

Set $\xi = \gamma(1 + c_1(1-\beta)) > 1$ and $\alpha = \frac{1}{2\xi}$ (note that α is in fact a function of β). We get that:

$$D_2 = 2^{c_1+1} \frac{\log^{\xi} D}{\log(1/\varepsilon_1)} = 2^{c_1+1} \left(\log^{\xi-1} \frac{1}{\varepsilon_1}\right) \left(\log_{1/\varepsilon_1} D\right)^{\xi} = 2^{c_1+1} \left(\log^{\xi-1} \frac{1}{\varepsilon_1}\right) t^{\alpha\xi}.$$

Now, note that $t^{\alpha\xi} = \sqrt{t}$, so $D_2 < t$ as long as $t > 4^{c_1+1} \log^{2(\xi-1)} \frac{1}{\varepsilon_1}$.

Let

$$D \colon \{0,1\}^{n_1} \times [D_1] \to \{0,1\}^{m}$$

be the explicit, strong (k_1, ε_1) t-n.m. extractor with $\log_{1/\varepsilon_1} D_1 \leq t^{\alpha} = \log_{1/\varepsilon_1} D$ promised by the hypothesis of the lemma. Again, we can take $D_1 = D$.

Now let

$$\mathsf{2Ext} \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^{m}$$

be constructed from E and Γ as in Section 3.1. We have 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, so by Theorem 14 2Ext is a $((n_1, k_1), (n_2, k_2), \frac{B_2}{K_2} + \sqrt{\varepsilon_1})$ two-source extractor.

In our case, $\frac{B_2}{K_2} = 2^{b_2 - k_2} = 2^{-b_2}$. We stress that $b_2 \geq \frac{1}{2} \log \frac{1}{\varepsilon_1}$. To see this, note that $2b_2 = n_2^{\beta} = d_1^{\beta\gamma}$. As $\beta\gamma \geq \beta\gamma' = \frac{\beta}{1 - c_2(1-\beta)} \geq 1$, and $d_1 \geq 2 \log \frac{1}{\varepsilon_1}$ (again, this is true for any seeded extractor), we finally have that $2b_2 \geq d_1 > \log \frac{1}{\varepsilon_1}$. Overall,

$$\frac{B_2}{K_2} + \sqrt{\varepsilon_1} \leq 2\sqrt{\varepsilon_1}$$

and we are done. ◀

Next, we show that we can *balance* the above two-source extractor (i.e., $n_1 = n_2$) by choosing the error ε_1 appropriately and assuming k_1 is small enough. The resulting two-source extractor supports polynomially-small min-entropies from both sources. Formally:

▶ **Corollary 19.** *Let $\varepsilon_1 \leq \frac{1}{n}$. There exists a constant $\beta_0 < 1$ such that for every $\beta_0 < \beta < 1$ there exits a constant $\alpha < 1$ so that the following holds. Suppose there exists an explicit*

$$E \colon \{0,1\}^{n_1} \times [D_1] \to \{0,1\}^m$$

that is a strong (k_1, ε_1) t-n.m. extractor with $\log_{1/\varepsilon_1} D_1 \leq t^{\alpha}$ for some $k_1 \leq d_1$, and t which is a large enough polynomial in $\log \frac{1}{\varepsilon_1}$. Then there exists an explicit

$$\mathsf{2Ext} \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$$

that is an $((n, k = k_1), (n, k), \varepsilon)$ two-source extractor for $k = n^{\beta}$ and $\varepsilon = 2^{-n^{\Omega(1)}}$.

Proof. Following the notations of Lemma 17, let β_0, α, γ be the constants set according to β. Let $\mathsf{2Ext} \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ be the explicit $((n_1, k_1), (n_2 = d_1^{\gamma}, k_2 = n_2^{\beta}), 2\sqrt{\varepsilon_1})$ that is guaranteed to us.

We require $n = n_1 = n_2 = d_1^{\gamma}$, so as $d_1 = t^{\alpha} \log \frac{1}{\varepsilon_1} \leq t \log \frac{1}{\varepsilon_1}$ and t is polynomial in $\log \frac{1}{\varepsilon_1}$, denote $t \log \frac{1}{\varepsilon_1} = \log^{\eta'} \frac{1}{\varepsilon_1}$ and $n = \log^{\eta} \frac{1}{\varepsilon_1}$ for some large enough constants $\eta', \eta = \gamma\eta'$. This guarantees that $\varepsilon = 2\sqrt{\varepsilon_1} = 2^{-n^{\Omega(1)}}$.

Next, note that $k_1 \leq d_1$ and $d_1 = n^{\frac{1}{\gamma}}$. Indeed, $n^{\frac{1}{\gamma}} \leq n^{\beta}$ since we already observed in the proof of Lemma 17 that $\gamma\beta \geq 1$. Overall $k_1 \leq n^{\beta}$ for every $\beta > \beta_0$. As by construction $k_2 = n_2^{\beta}$ for every $\beta > \beta_0$ as well, the proof is concluded. ◀

References

1 H.L. Abbott. Lower bounds for some Ramsey numbers. *Discrete Mathematics*, 2(4):289–293, 1972.

2 Noga Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998.

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

4 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.

5 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.

6 Avraham Ben-Aroya, Dean Doron, and Amnon Ta-Shma. An efficient reduction from two-source to non-malleable extractors: achieving near-logarithmic min-entropy. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1185–1194. ACM, 2017.

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

8 Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 285–298. ACM, 2016.

9 Eshan Chattopadhyay and Xin Li. Explicit non-malleable extractors, multi-source extractors, and almost optimal privacy amplification protocols. In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 158–167. IEEE, 2016.

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

11 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.

12 Fan R.K. Chung. A note on constructive methods for Ramsey numbers. *Journal of Graph Theory*, 5(1):109–113, 1981.

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

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

15 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.

16 Gil Cohen. Non-malleable extractors with logarithmic seeds. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 23, page 30, 2016.

17 Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved Ramsey graphs. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 278–284. ACM, 2016.

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

19 Gil Cohen, Ran Raz, and Gil Segev. Nonmalleable extractors with short seeds and applications to privacy amplification. *SIAM Journal on Computing*, 43(2):450–476, 2014.

20 Gil Cohen and Igor Shinkar. Personal communication, 2017.

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

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

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

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

25 Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 68–80. IEEE, 1988.

26 Xin Li. Non-malleable extractors, two-source extractors and privacy amplification. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 688–697. IEEE, 2012.

27 Xin Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 168–177. IEEE, 2016.

28 Xin Li. Improved non-malleable extractors, non-malleable codes and independent source extractors. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 1144–1156. ACM, 2017. doi:10.1145/3055399.3055486.

29 Raghu Meka. Explicit resilient functions matching ajtai-linial. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1132–1148. SIAM, 2017.

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

31 Moni Naor. Constructing Ramsey graphs from small probability spaces. *IBM Research Report RJ*, 8810, 1992.

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

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

34 Ran Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 11–20. ACM, 2005.

35 David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 681–690. ACM, 2006.

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

In this section we extend the [21] result, where non-malleability was considered only in the case of $t = 1$. We repeat Theorem 6 and prove:

▶ **Theorem 20.** *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 $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d \leq 2\log\frac{1}{\varepsilon} + \log(n-k) + 2\log(t+1) + 3$.*

This was also independently proved by Cohen and Shinkar [20].

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 $f_1, \dots, f_t \colon [D] \to [D]$ with no fixed-points and a distinguisher function $\mathcal{D}\colon \{0,1\}^{(t+1)m+d} \to \{0,1\}$. We want to bound the probability (over E) that

$$\Pr[\mathcal{D}(E(X,Y), E(X,f_1(Y)), \dots, E(X,f_t(Y)), Y) = 1] - \\ \Pr[\mathcal{D}(U_m, E(X,f_1(Y)), \dots, E(X,f_t(Y)), Y) = 1] \;>\; \varepsilon.$$

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

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

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

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

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 $f_1, \dots, f_t$ 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

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

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

▶ **Lemma 21.** *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 $\mathbf{Q}_1, \dots, \mathbf{Q}_{m = K|V'|}$ such that*

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

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 $\{1, \dots, m\}$ such that if $(y,y') \in E$ and $\mathbf{Q}_j = \mathbf{Q}(x,y')$, $\mathbf{Q}_i = \mathbf{Q}(x,y)$ then $j \le i$.

Fix some $i \in [m]$ and assume $\mathbf{Q}_i = \mathbf{Q}(x,y)$. The key point is that the variables $\mathbf{Q}_1, \dots, \mathbf{Q}_{i-1}$ 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 $e_1, \dots, e_t$, we have that

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

and as we noted, $\mathbf{Q}_1, \dots, \mathbf{Q}_{i-1}$ 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 22** ([32, 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 $\left\{\mathbf{Q}_1^1, \ldots, \mathbf{Q}_{s_1}^1\right\}, \ldots, \left\{\mathbf{Q}_1^t, \ldots, \mathbf{Q}_{s_t}^t\right\}$ such that for every $k \in [t+1]$ and $i \in [s_k]$, $\mathbb{E}[\mathbf{Q}_i^k \mid \mathbf{Q}_1^k, \ldots, \mathbf{Q}_{i-1}^k] = 0$. Now, define $S_i^k = \sum_{j=1}^{i} \mathbf{Q}_j^k$ and note that every sequence $S_1^k, \ldots, S_{s_k}^k$ is a martingale. Also, $|S_i^k - S_{i-1}^k| = |\mathbf{Q}_i^k| \leq 1$ with probability 1. Thus, using Azuma's inequality,

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

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

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. ◀