Electronic Colloquium on Computational Complexity, Report No. 160 (2014)

Zero-Fixing Extractors for Sub-Logarithmic Entropy

Gil Cohen[*] Igor Shinkar[†]

November 27, 2014

Abstract

An (n, k)-bit-fixing source is a distribution on n bit strings, that is fixed on $n - k$ of the coordinates, and jointly uniform on the remaining k bits. Explicit constructions of bit-fixing extractors by Gabizon, Raz and Shaltiel [SICOMP 2006] and Rao [CCC 2009], extract $(1 - o(1)) \cdot k$ bits for $k = \operatorname{poly}\log n$, almost matching the probabilistic argument. Intriguingly, unlike other well-studied sources of randomness, a result of Kamp and Zuckerman [SICOMP 2006] shows that, for *any* k, some small portion of the entropy in an (n, k)-bit-fixing source can be extracted. Although the extractor does not extract all the entropy, it does extract $(1/2 - o(1)) \cdot \log(k)$ bits.

In this paper we prove that when the entropy k is small enough compared to n, this exponential entropy-loss is unavoidable. More precisely, one cannot extract more than $\log(k)/2 + O(1)$ bits from (n, k)-bit-fixing sources. The remaining entropy is inaccessible, information theoretically. By the Kamp-Zuckerman construction, this negative result is tight.

Our impossibility result also holds for what we call *zero-fixing* sources. These are bit-fixing sources where the fixed bits are set to 0. We complement our negative result, by giving an explicit construction of an (n, k)-zero-fixing extractor, that outputs $\Omega(k)$ bits, even for $k = \operatorname{poly}\log\log n$. Furthermore, we give a construction of an (n, k)-bit-fixing extractor, that outputs $k - O(1)$ bits, for entropy $k = (1 + o(1)) \cdot \log\log n$, with running-time $n^{O((\log\log n)^2)}$.

[*] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. Email: gil.cohen@weizmann.ac.il.

[†] Courant Institute of Mathematical Sciences, New York University. Email: ishinkar@cims.nyu.edu

ISSN 1433-8092

Contents

1 Introduction

Randomness is an invaluable resource in many areas of theoretical computer science, such as algorithm design, data structures and cryptography. For many computational tasks, the best known algorithms assume that random bits are to their disposal. In cryptography and in distributed computing, randomness is, provably, a necessity. Nevertheless, truly random bits are not always available. A source of randomness might be defective, producing random bits that are biased and correlated. Even a sample from an ideal source of randomness can suffer such defects due to information leakage. Motivated by this problem, the notion of randomness extractors was introduced.

Broadly speaking, a *randomness extractor* is a function that extracts almost truly random bits given a sample from a defective source of randomness. Well-known instantiations are seeded extractors [NZ96, Tre01, RRV99, TSZS01, SU01, GUV09, DKSS09, TSU12], two-source extractors [CG88, Raz05, Bou05, BSZ11], and more generally multi-source extractors [BIW06, Raz05, Rao09a, Li11a, Li13], and affine extractors [Bou07, GR08, DG10, Yeh11, Li11b]. Randomness extractors are central objects in pseudorandomness, with many applications beyond their original motivation. Over the last 30 years, a significant research effort was directed towards the construction of randomness extractors in different settings. We refer the reader to Shaltiel's introductory survey on randomness extractors [Sha11] for more information.

Bit-fixing extractors

A well-studied defective source of randomness is a bit-fixing source. An (n, k)-*bit-fixing source* is a distribution X over $\{0, 1\}^n$, where some $n - k$ of the bits of X are fixed, and the joint distribution of the remaining k bits is uniform. The problem of extracting randomness from bit-fixing sources was initiated in the works of [Vaz85, BBR85, CGH$^+$85], motivated by applications to fault-tolerance, cryptography and communication complexity. More recently, bit-fixing extractors have found applications to formulae lower bounds [KRT13], and for compression algorithms for "easy" Boolean functions [CKK$^+$13].

The early works on bit-fixing extractors were concentrated on positive and negative results for extracting a truly uniform string. In [CGH$^+$85], it was observed that one can efficiently extract a uniform bit even from $(n, 1)$-bit-fixing sources, simply by XOR-ing all the input bits. In a sharp contrast, it was shown that extracting two jointly uniform bits cannot be done even from $(n, n/3 - 1)$-bit-fixing sources. Given this state of affairs, early works dealt with what we call "the high-entropy regime". Using a relation to error correcting codes, Chor et al. [CGH$^+$85] showed how to efficiently extract roughly $n - t \cdot \log_2(n/t)$ truly uniform output bits from $(n, n - t)$-bit-fixing sources, with $t = o(n)$. The authors complemented this result by an almost matching upper bound of $n - (t/2) \cdot \log_2(n/t)$ on the number of truly uniform output bits one can extract. In the same paper, some results were obtained also for (n, k)-bit-fixing sources, where k is slightly below $n/2$. Further lower bounds for this regime of parameters were obtained by Friedman [Fri92].

These negative results naturally led to study the relaxation, where the output of the extractor is only required to be close to uniform, in statistical distance.[1] A simple probabilistic argument can be used to show that, computational aspects aside, one can extract $m = k - 2\log(1/\varepsilon) - O(1)$ bits that are ε-close to uniform, from any (n, k)-bit-fixing source, as long as $k \ge \log(n) + 2\log(1/\varepsilon) + O(1)$. For simplicity, in the rest of this section we think of ε as a small constant. Thus, in particular, by allowing for some small constant error $\varepsilon > 0$, one can extract almost all the entropy k from any (n, k)-bit-fixing source, even for k as low as $\log(n) + O(1)$. We call the range $\log n \le k \le o(n)$, "the low-entropy regime".

The probabilistic argument mentioned above only yields an existential proof, whereas efficiently computable extractors are far more desired. Kamp and Zuckerman [KZ06] gave the first explicit construction of an (n, k)-bit-fixing extractor, with $k = o(n)$. More precisely, for any constant $\gamma > 0$, an explicit $(n, n^{1/2+\gamma})$-bit-fixing extractor was given, with $\Omega(n^{2\gamma})$ output bits. In a subsequent work, Gabizon, Raz and Shaltiel [GRS06] obtained an explicit $(n, \log^c n)$-bit-fixing extractor, where $c > 1$ is some universal constant. Moreover, the latter extractor outputs $(1 - o(1))$-fraction of the entropy, thus getting very close to the parameters of the non-explicit construction obtained by the probabilistic method. Using different techniques, Rao [Rao09b] obtained a bit-fixing extractor with improved dependency on the error ε.

For a vast majority of randomness extraction problems, such as the problem of constructing two-source extractors and affine extractors, a naïve probabilistic argument yields (non-explicit) extractors with essentially optimal parameters. Interestingly, this is not the case for bit-fixing extractors. The first evidence for that comes from the observation mentioned above. Namely, the XOR function is an extractor for $(n, 1)$-bit-fixing sources. A result of Kamp and Zuckerman [KZ06] shows that this is not an isolated incident, and in fact, for *any* $k \ge 1$ there is an (explicit and simple) extractor for (n, k)-bit-fixing sources, that outputs $0.5 \cdot \log_2(k) - O(\log\log k)$ random bits that are close to uniform. On the other hand, one can show that, with high probability, a random function with a single output bit is constant on some bit-fixing source with entropy, say, $\log(n)/10$. Thus, in this setting, *structured* functions outperform *random* functions, in the sense that the former can extract a logarithmic amount of the entropy from bit-fixing sources with arbitrarily low entropy, whereas the latter are constant, with high probability, on some $(n, \log(n)/10)$-bit-fixing source.

1.1 Our contribution

The state of affairs discussed above leads us to the study of (n, k)-bit-fixing extractors in the "very low entropy regime", namely, for $k = o(\log n)$. More concretely, in this paper we study the following question:

> What is the number of output bits that can be extracted from (n, k)-bit-fixing sources, in terms of the dependency of k in n?

[1] Friedman [Fri92] studied other notions of closeness. Although different measures are of interest, when analyzing extractors, the gold standard measure of closeness between distributions is statistical distance. In this paper we follow the convention, and measure the error of an extractor by the statistical distance of its output to the uniform distribution.

We consider this problem both in the information-theoretic and in the computational settings. By the discussion above we see that, computational aspects aside, when $k > \log(n) + O(1)$ one can extract $k - O(1)$ random bits that are close to uniform, whereas for any k, one can extract $\Omega(\log k)$ bits. Is it possible to extract $\Omega(k)$ bits from (n, k)-bit-fixing sources for any k? Or perhaps extracting all the randomness from very low entropy bit-fixing sources is impossible information-theoretically?

We further consider the problem of extracting randomness from what we call (n, k)*-zero-fixing sources*. A random variable X is an (n, k)-zero-fixing source, if it is an (n, k)-bit-fixing source, where all the fixed bits are set to zero. Clearly, extracting randomness from (n, k)-zero-fixing sources is an easier (or at least not harder) task than extracting randomness from (n, k)-bit-fixing sources. Thus, proving impossibility results for this model is more challenging.

Our first result states that when the entropy k is small enough compared to n, one cannot extract more than $0.5 \cdot \log_2(k) + O(1)$ bits from an (n, k)-zero-fixing source. This negative result is tight up to an additive factor of $O(\log \log k)$, as implied by the construction of Kamp and Zuckerman [KZ06]. In fact, the latter construction is optimal also for zero-fixing sources, when k is small enough compare to n. To state the result, we introduce the following notation. The function $\mathsf{Tower}\colon \mathbb{N} \to \mathbb{N}$ is defined as follows: $\mathsf{Tower}(0) = 1$, and for an integer $n \geq 1$, $\mathsf{Tower}(n) = 2^{\mathsf{Tower}(n-1)}$.

Theorem 1.1. *For any integers n, k such that $\mathsf{Tower}(k^{3/2}) < n$, the following holds. Let $\mathsf{Ext}\colon \{0,1\}^n \to \{0,1\}^m$ be an (n, k)-zero-fixing extractor with error ε. If $m > 0.5 \cdot \log_2(k) + O(1)$, then $\varepsilon \geq 0.99$.*

Since the impossibility result stated in Theorem 1.1 holds for zero-fixing sources, it is natural to try and complement it with feasibility results. Using a naïve probabilistic argument, one can prove the existence of an (n, k)-zero-fixing extractor, for any $k \geq \log \log n + \log \log \log n + O(1)$, with $m = k - O(1)$ output bits, where we treat the error ε as constant, for simplicity. Our second result is an almost matching explicit construction.

Theorem 1.2. *For any constant $\mu > 0$, and $n, k \in \mathbb{N}$, such that $k \geq (\log \log n)^{2+\mu}$, there exists an efficiently computable function*

$$\mathsf{ZeroBFExt}\colon \{0,1\}^n \to \{0,1\}^m,$$

where $m = \Omega(k)$, with the following property. For any (n, k)-zero-fixing source X, it holds that $\mathsf{ZeroBFExt}(X)$ is $(2^{-k^{\Omega(1)}} + (k \log n)^{-\Omega(1)})$-close to uniform.

We remark that the techniques used in [GRS06, Rao09b] for the constructions of bit-fixing extractors seem to work only for $k \geq \operatorname{poly} \log n$, even for zero-fixing sources, and new ideas are required so to exploit the extra structure of zero-fixing sources in order to extract $\Omega(k)$ bits from such sources with sub-logarithmic entropy.

Can one extract $\Omega(k)$ random bits, that are close to uniform, even from (n, k)-bit-fixing sources with $k = o(\log n)$? We show that the answer to this question is positive. Although we do not know how to construct such an extractor efficiently, the following theorem gives a semi-explicit construction. For simplicity, we state here the theorem for a constant error ε.

Theorem 1.3. *For any integers n, k, and constant $\varepsilon > 0$, such that $k > \log\log n + 2\log\log\log n + O_\varepsilon(1)$, there exists a function*

$$\mathsf{QuasiBFExt}\colon \{0,1\}^n \to \{0,1\}^m,$$

where $m = k - O_\varepsilon(1)$, with the following property. Let X be an (n,k)-bit-fixing source. Then, $\mathsf{QuasiBFExt}(X)$ *is ε-close to uniform. The running-time of evaluating* $\mathsf{QuasiBFExt}$ *is $n^{O_\varepsilon((\log\log n)^2)}$.*

On top of the semi-explicit construction in Theorem 1.3, we give a simpler existential proof for an extractor $\mathsf{QuasiBFExt}$, with parameters as in Theorem 1.3, based on the Lóvasz local lemma. See Section 5 for more details.

1.2 Proofs overview

In this section we give an overview for the proofs of Theorem 1.1 and Theorem 1.2. For the sake of clarity, in this section we allow ourselves to be informal and somewhat imprecise.

Proof overview for Theorem 1.1

To give an overview for the proof of Theorem 1.1, we start by considering a related problem. Instead of proving an upper bound on the number of output bits of an (n,k)-zero-fixing extractor, we prove an upper bound for zero-error dispersers. Generally speaking, a *zero-error disperser* for a class of sources is a function that obtains all outputs, even when restricted to any source in the class. More concretely, an (n,k)-*zero-fixing zero-error disperser* is a function $\mathsf{ZeroErrDisp}\colon \{0,1\}^n \to \{0,1\}^m$, such that for any (n,k)-zero-fixing source X, it holds that $\mathsf{supp}(\mathsf{ZeroErrDisp}(X)) = \{0,1\}^m$. We show that for any such zero-error disperser, if k is small enough compared to n, then $m \le \log_2(k+1)$. More specifically, we prove that for any integers n, k such that $\mathsf{Tower}(k^2) < n$ and $m = \lfloor \log_2(k+1) \rfloor + 1$, for any function $f\colon \{0,1\}^n \to \{0,1\}^m$, there exists an (n,k)-zero-fixing source, restricted to which f is a symmetric function, i.e., f depends only on the input's weight. In particular, f does not obtain all possible outputs.[2] This implies that if $f\colon \{0,1\}^n \to \{0,1\}^m$ is a (n,k)-zero-fixing zero-error dispersers and $\mathsf{Tower}(k^2) < n$, then $m \le \log_2(k+1)$.

Given $f\colon \{0,1\}^n \to \{0,1\}^m$, we construct the required source X in a level-by-level fashion, as follows. Trivially, f is symmetric on any $(n,1)$-zero-fixing source, regardless of the value of m. Next, we find an $(n,2)$-zero-fixing source on which f is symmetric. By the pigeonhole principle, there exists a set of indices $I_1 \subseteq [n]$, with size $|I_1| \ge n/2^m$, such that $f(e_i) = f(e_j)$ for all $i, j \in I_1$. Here, for an index $i \in [n]$, we denote by e_i the unit vector with 1 at the i^{th} coordinate. If $n > 2^m$, then $|I_1| \ge 2$, and so there exist two distinct $i, j \in I_1$. Thus, f restricted to the $(n,2)$-zero-fixing source $\{0, e_i, e_j, e_i + e_j\}$ is symmetric.

We take a further step, and find an $(n,3)$-zero-fixing source on which f is symmetric. We restrict ourselves to the index set I_1 above, and consider the complete graph with vertex

[2] If $m > \lfloor \log_2(k+1) \rfloor + 1$, then the same result can be obtained by restricting the output to the first $\lfloor \log_2(k+1) \rfloor + 1$ output bits.

set I_1, where for every two distinct vertices $i, j \in I_1$, the edge connecting them is colored by the color $f(e_i + e_j)$, where we think of $\{0,1\}^m$ as representing 2^m colors. By the multi-color variant of Ramsey theorem, there exists a set $I_2 \subseteq I_1$, of size

$$|I_2| \geq \log(|I_1|)/\mathrm{poly}(2^m),$$

such that the complete graph induced by I_2 is monochromatic. Therefore, if $n > 2^{2^{O(m)}} = 2^{\mathrm{poly}(k)}$, then $|I_2| \geq 3$, and so there exist distinct $i_1, i_2, i_3 \in I_2$ such that

$$\begin{gathered} f(e_{i_1}) = f(e_{i_2}) = f(e_{i_3}), \\ f(e_{i_1}{+}e_{i_2}) = f(e_{i_1} + e_{i_3}) = f(e_{i_2} + e_{i_3}). \end{gathered}$$

Thus, f is symmetric on the $(n,3)$-zero-fixing source spanned by $\{e_{i_1}, e_{i_2}, e_{i_3}\}$.

To construct an $(n,4)$-zero-fixing source on which f is symmetric, we consider the complete 3-uniform hypergraph on vertex set I_2 as above, where an edge $\{i_1, i_2, i_3\}$ is colored by $f(e_{i_1} + e_{i_2} + e_{i_3})$. Applying the multi-color Ramsey theorem for hypergraphs, we obtain a subset of the vertices $I_3 \subseteq I_2$, with size

$$|I_3| \geq \log\log(|I_2|)/\mathrm{poly}(2^m),$$

such that the induced complete hypergraph by the vertex set I_3 is monochromatic. Therefore, if $\log\log\log n \geq \mathrm{poly}(k)$, then $|I_3| \geq 4$, and thus there are distinct coordinates $i_1, i_2, i_3, i_4 \in I_3$ such that f is symmetric on the $(n,4)$-zero-fixing source spanned by $\{e_{i_1}, e_{i_2}, e_{i_3}, e_{i_4}\}$.

We continue this way, and find an (n,k)-zero-fixing source on which f is symmetric, by applying similar Ramsey-type arguments on r-uniform complete hypergraphs, with 2^m colors, for $r = 4, 5, \ldots, k-1$. A calculation shows that as long as $\mathsf{Tower}(k^2) < n$, such a source can be found.

To obtain the negative result for (n,k)-bit-fixing extractors, we follow a similar argument. The only difference is that in this case, it is enough to find an (n,k)-bit-fixing source X, such that f is symmetric restricted only to the $O(\sqrt{k})$ middle levels of X. Since most of the weight of X sits in these levels, an (n,k)-bit-fixing extractor cannot be symmetric restricted to these middle levels, regardless of the values obtained by the extractor in the remaining points of X.

Proof overview for Theorem 1.2

Informally speaking, the advantage one should exploit when given a sample from an (n,k)-zero-fixing source X, as apposed to a sample from a more general bit-fixing source, is that "1 hits randomness". More formally, if $X_i = 1$, then we can be certain that $i \in S$, where $S \subset [n]$ is the set of indices for which $X|_S$ is uniform. How should we exploit this advantage?

A natural attempt would be the following. Consider all (random) indices $1 \leq i_1 < i_2 < \cdots < i_W \leq n$, such that $X_{i_1} = \cdots = X_{i_W} = 1$. Note that W, the Hamming weight of the sample, is a random variable concentrated around $k/2$. Let $M = i_{W/2}$ be the median of these random indices. One can show that, with high probability with respect to the value of M, both the prefix $(X_1, X_2, \ldots, X_M)$ and the suffix $(X_{M+1}, X_{M+2}, \ldots, X_n)$ have entropy roughly $k/2$. Intuitively, this is because the "hidden" random bits, namely bits in coordinates $i \in S$ such that $X_i = 0$, must be somewhat intertwined with the "observed" random bits – bits in coordinates $i \in S$ for which $X_i = 1$. In particular, except with probability $2^{-\Omega(k)}$ over the value of M, both the prefix and the suffix have entropy at least $0.49k$. Thus, by appending these prefix and suffix with zeros, one can get two n bit sources $X_{\mathsf{left}}, X_{\mathsf{right}}$, each having entropy at least $0.49k$.

We observe that conditioned on the value of the median M, the random variables X_{left} and X_{right} preserve the zero-fixing structure. Unfortunately, however, $X_{\mathsf{left}}, X_{\mathsf{right}}$ are *dependent*. In this proof overview, we rather continue with the description of the zero-fixing extractor as if $X_{\mathsf{left}}, X_{\mathsf{right}}$ were independent, and deal with the dependencies later on.

After obtaining X_{left} and X_{right}, we apply the lossless-condenser of Rao [Rao09b] on each of these random variables. This is an efficiently computable function $\mathsf{Cond} \colon \{0,1\}^n \to \{0,1\}^{k \log n}$, that is one-to-one when restricted to any (n,k)-bit-fixing source. We compute $Y_{\mathsf{left}} = \mathsf{Cond}(X_{\mathsf{left}})$ and $Y_{\mathsf{right}} = \mathsf{Cond}(X_{\mathsf{right}})$ to obtain two $(k \log n, 0.49k)$-weak sources. Note that the one-to-one guarantee implies that no entropy is lost during the condensing, and so the entropy of $Y_{\mathsf{left}}, Y_{\mathsf{right}}$ equals the entropy of $X_{\mathsf{left}}, X_{\mathsf{right}}$, respectively.

At this point, for simplicity, assume we have an explicit optimal two-source extractor

$$\mathsf{TwoSourceExt} \colon \{0,1\}^{k \log n} \times \{0,1\}^{k \log n} \to \{0,1\}^m$$

to our disposal. The output of our zero-fixing extractor is then $\mathsf{TwoSourceExt}(Y_{\mathsf{left}}, Y_{\mathsf{right}})$. Working out the parameters, one can see that an optimal two-source extractor would yield an (n,k)-zero-fixing extractor for $k > \log\log n + O(\log\log\log n)$, error $2^{-\Omega(k)}$ and output length, say, $0.9k$.

Constructing two-source extractors for even sub-linear entropy, let alone for logarithmic entropy, as used in the last step, is a major open problem in pseudorandomness. Even for our short input length $k \log n = \tilde{O}(\log n)$, no $\mathrm{poly}(n)$-time construction is known. In this proof overview however, we choose to rely on such an assumption for the sake of clarity. In the real construction, we apply the split-in-the-median process above, recursively, to obtain c weak-sources, for any desired constant c. In a recent breakthrough, Li [Li13] gave an explicit construction of a multi-source extractor, that extracts a constant fraction of the entropy, from a constant number of weak-sources with poly-logarithmic entropy. In the actual construction, instead of using a two-source extractor, we use the extractor of Li with the appropriate constant c.

Working around the dependencies. So far we ignored the dependencies between X_{left} and X_{right}, even though their condensed images are given as inputs to a two-source extractor, and the latter expects its inputs to be independent. As we now explain, the dependencies between X_{left} and X_{right} can be worked around.

The crucial observation is the following: conditioned on the fixing of the Hamming weight W of the sample X, and conditioned on any fixing of the median M, the random variables $X_{\mathsf{left}}, X_{\mathsf{right}}$ are independent! To see this, fix $W = w$. Then, conditioned on the event

$M = m$, the value of the prefix $X_1, \ldots, X_m$ gives no information whatsoever about the suffix. More precisely, conditioned on any fixing of the prefix $X_1, \ldots, X_m$, the suffix is distributed uniformly at random over all $n - m$ bit strings, with zeros outside $S \cap \{m+1, \ldots, n\}$, and exactly $w/2$ ones in $S \cap \{m+1, \ldots, n\}$.

This observation motivates the following definition. We say that a random variable X is an (n, k, w)-*fixed-weight source*, if there exists $S \subseteq [n]$, with size $|S| = k$, such that a sample $x \sim X$ is obtained as follows. First, one samples a string $x' \in \{0,1\}^k$ of weight w, uniformly at random from all $\binom{k}{w}$ such strings, and then sets $X|_S = x'$, and $X_i = 0$ for all $i \notin S$. It is easy to see that any (n, k)-zero-fixing source is $2^{-\Omega(k)}$-close to a convex combination of (n, k, w)-fixed-weight sources, with w ranges over $k/3, \ldots, 2k/3$. Therefore, any extractor for (n, k, w)-fixed-weight sources, for such values of w, is also an extractor for (n, k)-zero-fixing sources.

We now reanalyze the algorithm described above. Since an (n, k)-zero-fixing source is $2^{-\Omega(k)}$-close to a convex combination of (n, k, w)-fixed-weight sources, with $k/3 \le w \le 2k/3$, we may assume, for the analysis sake, that the input is sampled from an (n, k, w)-fixed-weight source for some *fixed* $k/3 \le w \le 2k/3$. Fix also the median M to some value $m \in [n]$. Note that X_{left} is an $(n, k_{\mathsf{left}}(m), w/2)$-fixed-weight source[3], and X_{right} is an $(n, k_{\mathsf{right}}(m), w/2)$-fixed-weight source, with $k_{\mathsf{left}}(m)$ and $k_{\mathsf{right}}(m)$ being deterministic functions of m, satisfying $k_{\mathsf{left}}(m) + k_{\mathsf{right}}(m) = k$. Moreover, by the discussion above, we have that conditioned on the fixing $M = m$, the two random variables X_{left}, X_{right} are independent.

To summarize, conditioned on any fixing $M = m$, the two random variables $X_{\mathsf{left}}, X_{\mathsf{right}}$ are independent and preserve their fixed-weight structure. We further note that, with probability $1 - 2^{-\Omega(k)}$ over the value of M, it holds that $k_{\mathsf{left}}, k_{\mathsf{right}} \ge 0.49k$.

Recall that at this point we apply Rao's lossless-condenser on both X_{left} and X_{right}, to obtain shorter random variables Y_{left}, Y_{right}. Rao's condenser is one-to-one when restricted to bit-fixing sources. Since X_{left} and X_{right} are fixed-weight sources, they are in particular contained in some (n, k)-bit-fixing sources, and so the random variables $Y_{\mathsf{left}}, Y_{\mathsf{right}}$ have the same entropy as $X_{\mathsf{left}}, X_{\mathsf{right}}$, respectively.

It is worth mentioning that Rao's condenser Cond is linear, and as a result, if X_{left} were a bit-fixing source, then the resulting $Y_{\mathsf{left}} = \mathsf{Cond}(X_{\mathsf{left}})$ would have been an affine source. This property was crucial for Rao's construction of bit-fixing extractors. Since we wanted to maintain independence between $X_{\mathsf{left}}, X_{\mathsf{right}}$, in our case these random variables are no longer bit-fixing sources, but rather fixed-weight sources. Thus, the resulting Y_{left}, Y_{right} are not affine sources, but only weak sources, with min-entropy $\log_2(\binom{0.49k}{w/2}) = \Omega(k)$. This is good enough for our needs, as in the next step we use a two-source extractor, and do not rely on the affine-ness.

Lastly, we apply a two-source extractor on the condensed random variables $Y_{\mathsf{left}}, Y_{\mathsf{right}}$, which is a valid application, as these sources are independent, and with probability $1 - 2^{-\Omega(k)}$, both have entropy $\Omega(k)$.

[3] To be more precise, X_{left} is not an $(n, k_{\mathsf{left}}(m), w/2)$-fixed-weight source per se, as its m^{th} bit is constantly 1. Ignoring this bit would make X_{left} a fixed-weight source.

2 Preliminaries

Throughout the paper we denote by log the logarithm to the base 2. For $n \in \mathbb{N}$, we denote the set $\{1, 2, \ldots, n\}$ by $[n]$. For $n, r \in \mathbb{N}$, we let $\log^{(r)}(n)$ be the composition of the log function with itself r times, applied to n. Formally, $\log^{(0)}(n) = n$, and for $r \geq 1$, we define $\log^{(r)}(n) = \log(\log^{(r-1)}(n))$. For an integer $h \in \mathbb{N}$, we let $\mathsf{Tower}(h)$ be a height h tower of exponents of 2. More formally, $\mathsf{Tower}(0) = 1$, and for $h \geq 1$, $\mathsf{Tower}(h) = 2^{\mathsf{Tower}(h-1)}$.

Sources of randomness

In this paper we use the following sources of randomness.

Definition 2.1 (Bit-fixing sources)**.** *Let n, k be integers such that $n \geq k$. A random variable X on n bits is called an* (n, k)-bit-fixing source, *if there exists $S \subseteq [n]$ with size $|S| = k$, such that $X|_S$ is uniformly distributed, and each X_i with $i \notin S$ is fixed.*

Definition 2.2 (Affine sources)**.** *Let n, k be integers, with $n \geq k$. A random variable X on n bits is called an* (n, k)-affine source, *if X is uniformly distributed on some affine subspace $U \subseteq \mathbb{F}_2^n$ of dimension k.*

Definition 2.3 (Weak sources)**.** *Let n, k be integers such that $n \geq k$. A random variable X on n bits is called an* (n, k)-weak source, *if for any $x \in \mathsf{supp}(X)$, it holds that $\mathbf{Pr}[X = x] \geq 2^{-k}$.*

Note that any (n, k)-bit-fixing source is an (n, k)-affine source, and any (n, k)-affine source is an (n, k)-weak source. We introduce the following two sources of randomness.

Definition 2.4 (Zero-fixing sources)**.** *Let n, k be integers such that $n \geq k$. A random variable X on n bits is called an* (n, k)-zero-fixing source, *if there exists $S \subseteq [n]$ with size $|S| = k$, such that $X|_S$ is uniformly distributed, and each X_i with $i \notin S$ is fixed to zero.*

Definition 2.5 (Fixed-weight sources)**.** *Let n, k, w be integers, with $n \geq k \geq w$. A random variable $X \subseteq \{0, 1\}^n$ is called an* (n, k, w)-fixed-weight source, *if there exists $S \subseteq [n]$, with size $|S| = k$, such that a sample from $x \sim X$ is obtained as follows. First, one samples a string $x' \in \{0, 1\}^k$ of weight w, uniformly at random from all $\binom{k}{w}$ such strings. Then, $x|_S = x'$, and $x_i = 0$ for all $i \notin S$.*

Clearly, any (n, k)-zero-fixing source is an (n, k)-bit-fixing source. For a relation between zero-fixing sources and fixed-weight sources, see Claim 4.3. We will need the following extractor and condenser.

Theorem 2.1 ([Li13])**.** *For every constant $\mu > 0$ and all integers n, k with $k \geq \log^{2+\mu} n$, there exists an explicit function $\mathsf{Li}\colon (\{0, 1\}^n)^c \to \{0, 1\}^m$, with $m = \Omega(k)$ and $c = O(1/\mu)$, such that the following holds. If $X_1, \ldots, X_c$ are independent (n, k)-weak sources, then*

$$\mathsf{Li}(X_1, \ldots, X_c) \approx_\varepsilon U_m,$$

where $\varepsilon = n^{-\Omega(1)} + 2^{-k^{\Omega(1)}}$.

Theorem 2.2 ([Rao09b])**.** *For all integers n, k, there exists an efficiently computable linear transformation $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^{k \log n}$, such that for any (n,k)-bit-fixing source X it holds that Cond restricted to X is one-to-one.*

We further use of the following well-known fact.

Fact 2.6. *For any integer n, and $0 < \alpha < 1/2$, it holds that*

$$\sum_{k=0}^{\lfloor \alpha n \rfloor} \binom{n}{k} \le 2^{H(\alpha)\cdot n},$$

where $H(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ is the binary entropy function.

3 An Impossibility Result

For the proof of Theorem 1.1 we use the following notation. Let n be an integer and let $I \subseteq [n]$. We denote by $\{0,1\}^I$ the set of binary strings of length $|I|$ indexed by the elements of I. Using this notation for a string $x \in \{0,1\}^I$, we let $x^I \in \{0,1\}^n$ be the n bit string y such that $y|_I = x$, and $y_i = 0$ for all $i \notin I$, i.e., x^I is the extension of x to an n bit string with zeros outside the coordinate set I. Also, for a function $f \colon \{0,1\}^n \to \{0,1\}^m$ and a subset $I \subseteq [n]$, we define the function $f_I \colon \{0,1\}^I \to \{0,1\}^m$ as the restriction of f obtained by fixing the coordinates outside I to zeros. That is f_I is defined as $f_I(x) = f(x^I)$

We will need the following classical result of Erdős and Rado [ER52] on Ramsey numbers of multicolored hypergraphs .

Theorem 3.1 ([ER52], Theorem 1)**.** *Let G be the complete r-uniform hypergraph with vertex set $[n]$. Assume that each edge in G is colored by some color from $[c]$. Then, there exists a subset $I \subseteq [n]$ of size $|I| \ge c^{-O(1)} \cdot \log^{(r-1)}(n)$, such that the induced complete hypergraph by I is monochromatic.*[4]

The following corollary readily follows by Theorem 3.1. Indeed, the corollary is simply a rephrasing of the theorem in a slightly different language.

Corollary 3.1. *For any function $f \colon \{0,1\}^n \to \{0,1\}^m$ and integer $r \in [n]$, there exists a set of indices $I \subseteq [n]$, with size $|I| \ge 2^{-O(m)} \cdot \log^{(r-1)}(n)$, such that f_I is constant on the r^{th} level of $\{0,1\}^I$, i.e., $f_I(x) = f_I(y)$ for all $x, y \in \{0,1\}^I$ with $|x| = |y| = r$.*

Proof. Consider the complete r-uniform hypergraph on vertex set $[n]$, where each hyperedge $S \subseteq [n]$ of size $|S| = r$ is colored with $f(1_S)$, where 1_S is the characteristic function of the

[4] We remark that Theorem 1 in [ER52] is stated somewhat differently. The theorem, as stated in the original paper, asserts that any large enough complete r-uniform hypergraph, with hyperedges colored by c colors, contains a monochromatic complete r-uniform hypergraph on N vertices. By large enough we mean that the number of vertices is some (tower) function that depends on r, c and N. For our purposes, however, it will be more convenient to apply the theorem as we state it.

set S, i.e., $(1_S)_i = 1$ if and only if $i \in S$. By Theorem 3.1, there exists a subset of the vertices $I \subseteq [n]$, of size $2^{-O(m)} \cdot \log^{(r-1)}(n)$, such that the complete hypergraph induced by the vertex set I is monochromatic. By construction, this implies that for any $x, y \in \{0,1\}^I$ with $|x| = |y| = r$, it holds that $f(x^I) = f(y^I)$. Therefore, the function $f_I : \{0,1\}^I \to \{0,1\}^m$ is as desired. □

Before proving Theorem 1.1, we prove an analogous theorem for zero-error dispersers. We do so as the proof is slightly cleaner. Moreover, we consider this to be a natural impossibility result by itself.

Theorem 3.2. *Let* $\mathsf{ZeroErrDisp}\colon \{0,1\}^n \to \{0,1\}^m$ *be an* (n,k)*-zero-fixing zero-error disperser. If* $n > \mathsf{Tower}(k^2)$*, then* $m \le \log(k+1)$*.*

Proof. Clearly, it is enough to prove that if $m = \lfloor \log(k+1) \rfloor + 1$ then there exists an (n,k)-zero-fixing source on which $\mathsf{ZeroErrDisp}$ obtains at most $k+1$ distinct values, and thus reach a contradiction. (If m is larger, then we will obtain a contradiction by restricting the output to the first $m = \lceil \log(k+1) \rceil + 1$ bits.)

We define a sequence $[n] = I_0 \supseteq I_1 \supseteq I_2 \supseteq \ldots \supseteq I_k$, such that for each $i = 1, \ldots, k$ the restricted function f_{I_i} is symmetric on the levels $0, \ldots, i$ of the restricted hypercube $\{0,1\}^{I_i}$. Then, we shall claim that if $n > \mathsf{Tower}(k^2)$ then $|I_k| \ge k$. By taking $I \subseteq I_k$ to be any subset of size k we obtain a zero-fixing source spanned by the coordinates of I, such that the restriction of f to this zero-fixing source is a symmetric function, and in particular obtains at most $k+1$ values. By the argument above, this implies that $m \le \log(k+1)$.

We define the subsets I_i iteratively by applying Corollary 3.1 for each $i = 1, \ldots, k$ with the function $f_{i-1} : \{0,1\}^{I_{i-1}} \to \{0,1\}^m$ and with $r = i$. Letting $n_{i-1} = |I_{i-1}|$, by Corollary 3.1 we obtain a subset $I_i \subseteq I_{i-1}$ of size $n_i = |I_i| \ge 2^{-O(m)} \cdot \log^{(i-1)}(n_{i-1})$. One can show, e.g., by induction on $i = 1, \ldots, k-1$, that $n_i \ge 2^{-O(m)} \cdot \log^{(s_i)}(n)$, where $s_i = \sum_{j=1}^{i} (j-1) = i(i-1)/2$. In particular, this implies that $n_k \ge 2^{-O(m)} \cdot \log^{(k^2/2)}(n)$, and so if $n > \mathsf{Tower}(k^2)$ then $|I_k| = n_k \ge k$. This completes the proof of the theorem. □

We now turn to prove Theorem 1.1.

Proof of Theorem 1.1. The proof outline is similar to the proof of Theorem 3.2, The only difference, when considering extractors rather than zero-error dispersers, is that it is enough to find an (n,k)-zero-fixing source such that f restricted to this source is symmetric only in the middle $O(\sqrt{k})$ levels, and not on all points of X. More precisely, let $\mathsf{bottom} = k/2 - c \cdot \sqrt{k \cdot \log(1/\delta)}$ and $\mathsf{top} = k/2 + c \cdot \sqrt{k \cdot \log(1/\delta)}$, where c is a universal constant such that

$$\sum_{i=\mathsf{bottom}}^{\mathsf{top}} \binom{k}{i} \ge \left(1 - \frac{\delta}{2}\right) \cdot 2^k.$$

One can show that such a constant c exists using a Chernoff bound. Let $I_{\mathsf{bottom}-1} = [n]$ and define a sequence $I_{\mathsf{bottom}-1} \supseteq I_{\mathsf{bottom}} \supseteq I_{\mathsf{bottom}+1} \supseteq \ldots \supseteq I_{\mathsf{top}}$, such that for each $i = \mathsf{bottom}, \ldots, \mathsf{top}$, the function f_{I_i} is symmetric on the levels $\mathsf{bottom}, \ldots, i$ of the restricted hypercube $\{0,1\}^{I_i}$. For each $i = \mathsf{bottom}, \ldots, \mathsf{top}$, given I_{i-1} we apply Corollary 3.1 with

$f = f_{I_{i-1}}$ and $r = i$ to obtain $I_i \subseteq I_{i-1}$, such that the restriction f_{I_i} of $f_{I_{i-1}}$ is symmetric on level i, as well as on levels $\mathsf{bottom}, \ldots, i-1$, as f_{I_i} is a restriction of $f_{I_{i-1}}$.

By construction, it follows that if $n_{\mathsf{top}} \geq k$, then there exists an (n,k)-zero-fixing source X, such that f is symmetric restricted to the levels $\mathsf{bottom}, \ldots, \mathsf{top}$ of X. By our choice of bottom and top, with probability $1 - \delta/2$ over a uniformly random $x \sim X$, it holds that x has Hamming weight in $[\mathsf{bottom}, \mathsf{top}]$. Let C be the set of outputs obtained by f restricted to these levels of X. Note that $|C| \leq |\mathsf{top} - \mathsf{bottom} + 1| = O(\sqrt{k \cdot \log(1/\delta)})$. Thus, by considering the event $f(X) \in C$, we see that the statistical distance between the output of f on X, and the uniform distribution on m bits, is at least $1 - \delta/2 - |C|/2^m$, which is at least $1 - \delta$, whenever $m \geq 0.5 \log_2(k) + O(\log(1/\delta))$.

By Corollary 3.1, for every $i = \mathsf{bottom}, \ldots, \mathsf{top}$, we have that $|I_i| \geq 2^{-O(m)} \cdot \log^{(i-1)}(|I_{i-1}|)$ which is larger than $2^{-O(m)} \cdot \log^{(k)}(|I_{i-1}|)$. Therefore, if $n > \mathsf{Tower}(O(k^{3/2} \cdot \sqrt{\log(1/\delta)}))$, then $|I_{\mathsf{top}}| \geq k$, as required. □

4 Explicit Zero-Fixing Extractors for Double-Logarithmic Entropy

In this section we prove Theorem 1.2. We repeat the statement of the theorem here for the readers convenience.

Theorem 4.1. *For any constant $\mu > 0$, and $n, k \in \mathbb{N}$, such that $k \geq (\log\log n)^{2+\mu}$, there exists an efficiently computable function*

$$\mathsf{ZeroBFExt} \colon \{0,1\}^n \to \{0,1\}^m,$$

where $m = \Omega(k)$, with the following property. For any (n,k)-zero-fixing source X, it holds that $\mathsf{ZeroBFExt}(X)$ is $(2^{-k^{\Omega(1)}} + (k \log n)^{-\Omega(1)})$-close to uniform.

We start by proving the following lemma that, informally speaking, shows how to efficiently split one fixed-weight source to two independent fixed-weight sources, each with half the weight and roughly half the entropy of the original source.

Lemma 4.1. *For every integer n, there exists an $O(n)$-time computable function*

$$\mathsf{SplitInMedian} \colon \{0,1\}^n \to \left(\{0,1\}^n\right)^2,$$

with the following property. Let X be an (n,k,w)-fixed-weight source, with $k/10 \leq w \leq 9k/10$. Denote the two n bit outputs of $\mathsf{SplitInMedian}(X)$ by X_{left} and X_{right}. Then, there exists a random variable M, and deterministic functions $k_{\mathsf{left}}, k_{\mathsf{right}}$ of M, such that conditioned on any fixing $M = m$, the following holds:

- *The random variables $X_{\mathsf{left}}, X_{\mathsf{right}}$ are independent.*
- *X_{left} is an $(n, k_{\mathsf{left}}, w/2 - 1)$-fixed-weight source.*

- X_{right} *is an* $(n, k_{\mathsf{right}}, w/2)$*-fixed-weight source, where* $k_{\mathsf{left}} + k_{\mathsf{right}} + 1 = k$*.*

Furthermore, for any $\varepsilon > 0$*, it holds that*

$$\mathop{\mathbf{Pr}}_{m \sim M}\left[\left|\frac{k_{\mathsf{left}}}{k} - \frac{1}{2}\right| \geq \varepsilon\right] \leq 2^{-\Omega(\varepsilon^2 \cdot k)} .$$

Proof. We first describe the algorithm for computing $\mathsf{SplitInMedian}(X)$, and then turn to the analysis. Let $1 \leq i_1 < i_2 < \cdots < i_w \leq n$ be the (random) indices such that $X_{i_1} = X_{i_2} = \cdots = X_{i_w} = 1$, and set $M = i_{w/2}$ to be the median coordinate. We define the n bit strings X_{left} and X_{right} as follows:

$$(X_{\mathsf{left}})_i = \begin{cases} X_i, & 1 \leq i < M; \\ 0, & M \leq i \leq n. \end{cases}$$

$$(X_{\mathsf{right}})_i = \begin{cases} 0, & 1 \leq i \leq M; \\ X_i, & M < i \leq n. \end{cases}$$

The output of $\mathsf{SplitInMedian}(X)$ is then $(X_{\mathsf{left}}, X_{\mathsf{right}})$. Clearly, the running-time of the algorithm is $O(n)$, as computing M and constructing $X_{\mathsf{left}}, X_{\mathsf{right}}$ can be carried out in linear-time.

Let $S \subseteq [n]$, with $|S| = k$, be the set of indices associated with X. That is, $X_i = 0$ for all $i \notin S$, and $X|_S$ is uniformly distributed over all k bit strings with Hamming weight w. Conditioned on the event $M = m$, it holds that conditioned on any fixing of the prefix $X_1, \ldots, X_m$, the suffix $X_{m+1}, \ldots, X_n$ is sampled uniformly at random from all $n - m$ bit strings, with Hamming weight $w/2$, and zeros outside $S \cap \{m+1, \ldots, n\}$. Since X_{left} and X_{right} are deterministic functions of these prefix and suffix, respectively, we have that conditioned on any fixing of M, the random variables X_{left} and X_{right} are independent.

As for the second and third items, we note that, for any value m, conditioned on the event $M = m$, the prefix $X_1, \ldots, X_{m-1}$ is a fixed-weight source. Indeed, $X_1, \ldots, X_{m-1}$ has the same distribution as sampling a vector $X' \in \{0,1\}^{k_{\mathsf{left}}}$, where $k_{\mathsf{left}} = |S \cap [m-1]|$, uniformly at random out of all such vectors with Hamming weight $w/2 - 1$, and setting $X|_{S \cap [m-1]} = X'$, and $X_i = 0$ for all $i \in [m-1] \setminus S$. Since X_{left} is obtained by concatenating zeros to the prefix $X_1, \ldots, X_{m-1}$, we have that X_{left} is an $(n, k_{\mathsf{left}}, w/2 - 1)$-fixed-weight source. A similar argument shows that X_{right} is an $(n, k_{\mathsf{right}}, w/2)$-fixed-weight source, where $k_{\mathsf{right}} = |S \cap \{m+1, \ldots, n\}|$. In particular, it holds that $k_{\mathsf{left}} + k_{\mathsf{right}} + 1 = k$.

For the furthermore part of the lemma, we want to bound the probability that k_{left} deviates from $k/2$, where the probability is taken with respect to the random variable M. Recall that k_{left} is a deterministic function of M, given by $k_{\mathsf{left}} = |S \cap [M-1]|$. Thus,

$$\begin{aligned}\mathop{\mathbf{Pr}}_{m \sim M}\left[k_{\mathsf{left}} \leq \left(\frac{1}{2} - \varepsilon\right) \cdot k\right] &= \mathop{\mathbf{Pr}}_{m \sim M}\left[|S \cap [m-1]| \leq \left(\frac{1}{2} - \varepsilon\right) \cdot |S|\right] \\ &= \binom{k}{w}^{-1} \cdot \sum_{t=w/2}^{(1/2-\varepsilon)\cdot k} \binom{t}{w/2}\binom{k-t}{w/2} \\ &\leq 2^{-\Omega(\varepsilon^2 \cdot k)},\end{aligned}$$

where the last inequality follows by applying Stirling's approximation (or alternatively, by approximating the binomial distribution by a uniform distribution, and applying a Chernoff bound), and our assumption that $k/10 \leq w \leq 9k/10$. By symmetry, the same upper bound holds for $\mathbf{Pr}\left[k_{\mathsf{left}} \geq \left(\frac{1}{2}+\varepsilon\right) \cdot k\right]$, which concludes the proof. $\square$

For the proof of Theorem 4.1, we need to split the source to more than 2 independent sources. The following corollary accomplishes that, based on Lemma 4.1, and a recursive argument.

Corollary 4.2. *For any integers n, c, where c is a power of 2, there exists an $O(cn)$-time computable function*

$$\mathsf{Splitter}\colon \{0,1\}^n \to \left(\{0,1\}^n\right)^c ,$$

with the following property. Let X be an (n,k,w)-fixed-weight source, with $k/3 \leq w \leq 2k/3$. Let $(Y_1,\ldots,Y_c) = \mathsf{Splitter}(X)$, with $Y_i \in \{0,1\}^n$ for all $i \in [c]$. Then, there exist random variables $M_1,\ldots,M_{c-1}$, and deterministic functions $k_1,\ldots,k_{c-1}$ of them, such that conditioned on any fixing $(M_1,\ldots,M_{c-1}) = (m_1,\ldots,m_{c-1})$, the following holds:

- *The random variables $Y_1,\ldots,Y_c$ are independent.*
- *For every $i \in [c]$, the random variable Y_i is an (n,k_i,w_i)-fixed-weight source, with $w_i \in [w/c-1, w/c]$, and $k_1+\cdots+k_c = k-c+1$.*

Furthermore, except with probability $c \cdot 2^{-\Omega(k/(c \cdot \log^2 c))}$ over the fixings of $(M_1,\ldots,M_{c-1})$, it holds that for all $i \in [c]$, $k_i \geq 0.9k/c$.

Proof. Let $d = \log_2 c$. Consider a depth d binary tree T with c leaves. With each node v of T, we associate a random variable X_v, defined recursively with respect to the depth, as follows. Let r be the root of T. We define $X_r = X$. Let v be a node in T, that is not a leaf, for which X_v was already defined. Denote by $\mathsf{leftChild}(v), \mathsf{rightChild}(v)$ the left and right children of v in T, respectively. Let $((X_v)_{\mathsf{left}}, (X_v)_{\mathsf{right}}) = \mathsf{SplitInMedian}(X_v)$. We associate the random variable $(X_v)_{\mathsf{left}}$ with the vertex $\mathsf{leftChild}(v)$, and the random variable $(X_v)_{\mathsf{right}}$ with the vertex $\mathsf{rightChild}(v)$. Namely, $X_{\mathsf{leftChild}(v)} = (X_v)_{\mathsf{left}}$ and $X_{\mathsf{rightChild}(v)} = (X_v)_{\mathsf{right}}$. Let M_v be the random variable M, in the notation of Lemma 4.1, with respect to the application of $\mathsf{SplitInMedian}$ to X_v. Let $\ell_1,\ldots,\ell_c$ be the c leaves of T. The output of $\mathsf{Splitter}$ on input X is defined by

$$\mathsf{Splitter}(X) = (X_{\ell_1},\ldots,X_{\ell_c}) .$$

We now turn to the analysis. First, clearly, $\mathsf{Splitter}$ is computable in time $O(cn)$, as it involves $c-1$ applications of $\mathsf{SplitInMedian}$. Let $h \in \{0,1,\ldots,d-1\}$, and let V_h be the set of nodes of T with depth h. Let $\varepsilon = 1/(20d)$. We prove the following, by induction on h. Conditioned on any fixing of the random variables $\{M_v \mid v \in V_0 \cup V_1 \cup \cdots \cup V_{h-1}\}$, the following holds:

- The random variables $\{X_v \mid v \in V_h\}$ are independent.

- For any $v \in V_h$, the random variable X_v is an (n, k_v, w_v)-fixed-weight source, with $w_v \in [w/2^h, w/2^h - 1]$, and where k_v is a deterministic function of the random variables $\{M_u\}_{u \in P_v}$, where P_v is the nodes on the path from the root r to v in T, not including v. Moreover, $\sum_{v \in V_h} k_v = k - h$.

Furthermore, except with probability $\delta_h = 2^h \cdot 2^{-\Omega(k/(c \cdot \log^2 c))}$ over the fixings of $\{M_v \mid v \in V_0 \cup V_1 \cup \cdots \cup V_{h-1}\}$, it holds

$$\forall v \in V_h \quad \left(\frac{1}{2} - \varepsilon\right)^h \leq \frac{k_v}{k} \leq \left(\frac{1}{2} + \varepsilon\right)^h .$$

These claims clearly hold for $h = 0$. We now prove that the claims hold for $h \geq 1$, assuming they hold for $1, \ldots, h-1$. By the induction hypothesis, conditioned on any fixings of $\{M_v \mid v \in V_0 \cup V_1 \cup \cdots \cup V_{h-2}\}$, the random variables $\{X_v \mid v \in V_{h-1}\}$ are independent. Since $\{X_v \mid v \in V_h\} = \{X_{\mathsf{leftChild}(v)}, X_{\mathsf{rightChild}(v)} \mid v \in V_{h-1}\}$, the independence of the random variables $\{X_v \mid v \in V_h\}$, conditioned on the further fixings of $\{M_v \mid v \in V_{h-1}\}$, follows by Lemma 4.1. The second item readily follows by the induction hypothesis and Lemma 4.1.

As for the furthermore part, by the induction hypothesis, except with probability δ_{h-1} over the fixings of $\{M_v \mid v \in V_0 \cup V_1 \cup \cdots \cup V_{h-2}\}$, it holds that

$$\forall v \in V_{h-1} \quad \left(\frac{1}{2} - \varepsilon\right)^{h-1} \leq \frac{k_v}{k} \leq \left(\frac{1}{2} + \varepsilon\right)^{h-1} .$$

One can easily verify that conditioned on this event, by our choice of ε, the hypothesis of Lemma 4.1 is met when computing $\mathsf{SplitInMedian}(X_v)$, namely,

$$\frac{k_v}{10} \leq w_v \leq \frac{9k_v}{10} .$$

Thus, by the union bound, except with probability

$$\delta_{h-1} + \sum_{v \in V_{h-1}} 2^{-\Omega(\varepsilon^2 \cdot k_v)} , \tag{1}$$

it holds that for all $v \in V_h$

$$\frac{k_v}{k} = \frac{k_v}{k_{\mathsf{parent}(v)}} \cdot \frac{k_{\mathsf{parent}(v)}}{k} \leq \left(\frac{1}{2} + \varepsilon\right) \cdot \left(\frac{1}{2} + \varepsilon\right)^{h-1} = \left(\frac{1}{2} + \varepsilon\right)^h ,$$

where $\mathsf{parent}(v)$ is the parent of v in the tree T. Similarly,

$$\frac{k_v}{k} = \frac{k_v}{k_{\mathsf{parent}(v)}} \cdot \frac{k_{\mathsf{parent}(v)}}{k} \geq \left(\frac{1}{2} - \varepsilon\right) \cdot \left(\frac{1}{2} - \varepsilon\right)^{h-1} = \left(\frac{1}{2} - \varepsilon\right)^h .$$

Since $|V_{h-1}| = 2^{h-1}$, $\varepsilon = 1/(20d) = O(1/\log c)$ and $k_v = \Omega(k/2^h) \geq \Omega(k/c)$, the error expression in Equation (1) is bounded above by

$$\delta_{h-1} + 2^{h-1} \cdot 2^{-\Omega(k/(c \cdot \log^2 c))} \leq \delta_h .$$

This concludes the inductive proof. The proof of the corollary follows by plugging $h = d$. □

For the proof of Theorem 4.1 we also need the following claim, which states that an (n,k)-zero-fixing source is close to a convex combination of fixed-weight sources, with weight roughly $k/2$.

Claim 4.3. *Let X be an (n,k)-zero-fixing source. Then, X is $2^{-\Omega(k)}$-close to a convex combination of (n,k,w)-fixed-weight sources, with $k/3 \le w \le 2k/3$.*

Proof. We first note that X can be written as the convex combination

$$X = \sum_{w=0}^{k} \lambda_w X_w,$$

where $\lambda_w = \binom{k}{w} \cdot 2^{-k}$, and X_w is an (n,k,w)-fixed-weight source. To see this, let $S \subseteq [n]$, $|S| = k$, be the set that is associated with the source X. Namely, $X|_S$ is uniformly distributed, whereas $X|_{S^c}$ is fixed to 0. Sampling $x \sim X$ can be done in two steps. In the first step, one samples a weight W according to a binomial distribution $\mathrm{Bin}(k, 1/2)$. Namely, for any $0 \le w \le k$, $\mathbf{Pr}[W = w] = \binom{k}{w} \cdot 2^{-k}$. In the second step, one samples a string $x' \in \{0,1\}^k$ uniformly at random among all strings with Hamming weight w. Lastly, we set $X|_S = x'$, and $x_i = 0$ for all $i \notin S$. It is easy to verify that this two-steps procedure yields the same distribution as sampling from the (n,k)-zero-fixing source X. Note that the sampling done in the second step, conditioned on the event $W = w$, is from an (n,k,w)-fixed-weight source, which we denote by X_w.

By Fact 2.6, we have that

$$\sum_{w=k/3}^{2k/3} \lambda_w \ge 1 - 2 \cdot 2^{(H(1/3)-1)\cdot k} = 1 - 2^{-\Omega(k)}.$$

This concludes the proof, as it shows that X is $2^{-\Omega(k)}$-close to the convex combination

$$X = \sum_{w=k/3}^{2k/3} \lambda_w X_w.$$

□

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. We first describe the construction of $\mathsf{ZeroBFExt}$ and then turn to the analysis. For the construction of $\mathsf{ZeroBFExt}$ we need the following building blocks:

- Let $\mathsf{Li}\colon \left(\{0,1\}^{k \log n}\right)^c \to \{0,1\}^\ell$ be the multi-source extractor from Theorem 2.1, set to extract $\ell = \Omega(k)$ bits from c independent $(k \log n, k)$-weak-sources, with $k \ge O(\log^{2+\mu}(k \log n))$. By Theorem 2.1, it suffices to take $c = O(1/\mu)$.
- With c as above, let $\mathsf{Splitter}\colon \{0,1\}^n \to (\{0,1\}^n)^c$ be the function from Corollary 4.2.

- Let $\mathsf{Cond}\colon \{0,1\}^n \to \{0,1\}^{k \log n}$ be the lossless-condenser of Rao from Theorem 2.2.

With these building blocks, we compute $\mathsf{ZeroBFExt}(X)$ as follows. We first compute $(Y_1,\ldots,Y_c) = \mathsf{Splitter}(X)$. Secondly, for each $i \in [c]$, we compute $Z_i = \mathsf{Cond}(Y_i)$. The output is then $\mathsf{ZeroBFExt}(X) = \mathsf{Li}(Z_1,\ldots,Z_c)$.

We now turn to the analysis. By Claim 4.3, X is $2^{-\Omega(k)}$-close to a convex combination of (n,k,w)-weight-fixing sources $\{X_w\}_{w=k/3}^{2k/3}$. Therefore, $\mathsf{Splitter}(X)$ is $2^{-\Omega(k)}$-close to a convex combination of the random variables $\{\mathsf{Splitter}(X_w)\}_{w=k/3}^{2k/3}$. We denote $((Y_w)_1,\ldots,(Y_w)_c) = \mathsf{Splitter}(X_w)$. Fix such w. By Corollary 4.2, conditioned on some carefully chosen random variables, $(Y_w)_1,\ldots,(Y_w)_c$ are independent random variables. Moreover, except with probability $2^{-\Omega(k)}$ with respect to the conditioning, it holds that for all $i \in [c]$, $(Y_w)_i$ is an $(n,k',w/c)$-weight-fixing source, with $k' \geq 0.9k/c$. Since

$$\binom{k'}{w/c} \geq \left(\frac{k'}{w/c}\right)^{w/c} \geq \left(\frac{0.9k}{w}\right)^{w/c} \geq \left(\frac{0.9}{2/3}\right)^{w/c} = 2^{\Omega(k)},$$

we have that $H_\infty((Y_w)_i) = \Omega(k)$ for all $i \in [c]$, except with probability $2^{-\Omega(k)}$.

Recall that $Z_i = \mathsf{Cond}(Y_i)$. With the notation above, we have that Z_i is $2^{-\Omega(k)}$-close to a convex combination of $(Z_w)_i = \mathsf{Cond}((Y_w)_i)$, where $k/3 \leq w \leq 2k/3$. Since $(Y_w)_i$ is contained in some (n,k)-bit-fixing source, Theorem 2.2 guarantees that Cond restricted to the support of $(Y_w)_i$ is one-to-one, and so $H_\infty((Z_w)_i) = H_\infty((Y_w)_i) = \Omega(k)$. Thus, except with probability $2^{-\Omega(k)}$, we have that for all $i \in [c]$, $(Z_w)_i$ is a $(k\log n, \Omega(k))$-weak source. This implies that $\mathsf{ZeroBFExt}(X_w) = \mathsf{Li}((Z_w)_1,\ldots,(Z_w)_c)$ is $(2^{-k^{\Omega(1)}} + (k\log n)^{-\Omega(1)})$-close to uniform, which completes the proof of the theorem as $\mathsf{Li}(X)$ is $2^{-\Omega(k)}$-close to a convex combination of $\{\mathsf{Li}(X_w)\}_{w=k/3}^{2k/3}$.

As for the running-time. Computing $Y_1,\ldots,Y_c$ by applying $\mathsf{Splitter}$ to X is done in time $O(n)$. Applying Rao's condenser to each Y_i can be done in poly(n)-time. Finally, Li's extractor runs in time poly$(k\log n) = o(n)$. □

A comment regarding the error. As stated, the extractor $\mathsf{ZeroBFExt}$ in Theorem 1.2 has an error of $2^{-k^{\Omega(1)}} + (k\log n)^{-\Omega(1)}$. This error is induced by the error of Li's multi-source extractor. Indeed, the error contributed by the other parts of the construction of $\mathsf{ZeroBFExt}$ is only $2^{-\Omega(k)}$. The error of Li's extractor, when applied to (n,k)-weak sources, is stated to be $n^{-\Omega(1)} + 2^{-k^{\Omega(1)}}$. However, by inspection, one can see that Li's extractor has an error of $(\delta n)^{O(1)} + 2^{-k^{\Omega(1)}}$, for any desired parameter $\delta > 0$. The running-time of the extractor is poly(n/δ). Clearly, when one is interested in poly(n) running-time, then one must take $\delta \geq 1/\text{poly}(n)$. However, in our case, the inputs to Li's extractor have length $O(k\log n)$. Thus, we can set ε to be such that the total error in our application of Li's extractor is $2^{-k^{\Omega(1)}}$, and the running-time of that application would then be poly$(2^k \cdot \log n)$, which is $o(n)$ for the parameters of interest, namely for $k = o(\log n)$. To summarize, the error in Theorem 4.1 can be reduced to $2^{-k^{\Omega(1)}}$. We choose to state Theorem 1.2 as we did so to be able to use Li's extractor in a black-box fashion.

5 Bit-Fixing Extractors for Double-Logarithmic Entropy

In this section we prove Theorem 1.3. We restate the theorem here, allowing also for non-constant error ε.

Theorem 5.1. *For any integers n, k, and $\varepsilon > 0$, such that*

$$k > \log(\log(n)/\varepsilon^2) + 2\log\log(\log(n)/\varepsilon) + O(1),$$

there exists a function

$$\mathsf{QuasiBFExt}\colon \{0,1\}^n \to \{0,1\}^m,$$

where $m = k - 2\log(1/\varepsilon) - O(1)$, with the following property. Let X be an (n,k)-bit-fixing source. Then, $\mathsf{QuasiBFExt}(X)$ is ε-close to uniform. The running time of evaluating $\mathsf{QuasiBFExt}$ is $n^{O(\log^2(\frac{\log n}{\varepsilon}))}$.

Before proving Theorem 1.3, we sketch two proofs for the existence of (n,k)-bit-fixing extractors, with double-logarithmic entropy. Our first proof relies on the Lóvasz local lemma.

Lemma 5.1 (Lóvasz local lemma [EL75, Spe77]). *Let $E_1, \ldots, E_k$ be events in a probability space, such that each event occurs with probability at most p, and such that each event is independent of all but at most d events. If $ep(d+1) \le 1$,* [5] *then*

$$\mathbf{Pr}\left[\bigcap_{i=1}^{k} \bar{E}_i\right] > 0.$$

Existential proof-sketch based on the Lóvasz Local Lemma. Let $f\colon \{0,1\}^n \to \{0,1\}^m$ be a random function. For any (n,k)-bit-fixing source X, let E_X be the event $\mathsf{SD}(f(X), U_m) > \varepsilon$ (here the randomness is taken over f). Fix an (n,k)-bit-fixing source X. By taking the union bound over all 2^{2^m} test functions, Chernoff bound implies that

$$\mathop{\mathbf{Pr}}_{f}[E_X] \le 2^{2^m} \cdot 2^{-\Omega(2^k \cdot \varepsilon^2)} = p.$$

Consider any two bit-fixing sources X, Y. We note that if there exists a coordinate $i \in [n]$, in which both X and Y are fixed, then in order for X and Y to be dependent, it must hold that $X_i = Y_i$. Indeed, if $X_i \neq Y_i$ then $X \cap Y = \emptyset$. Thus, for any (n,k)-bit-fixing source X, there are at most $d = \binom{n}{k} \cdot 2^k$ bit-fixing sources Y such that E_X depends on E_Y. One can easily verify that by taking

$$\begin{aligned} k &= \log\log(n) + 2\log(1/\varepsilon) + \log(\log\log(n) + 2\log(1/\varepsilon)) + O(1), \\ m &= k - 2\log(1/\varepsilon) - O(1), \end{aligned}$$

the hypothesis of Lemma 5.1 is met. Thus, even for k as above, there exists an (n,k)-bit-fixing extractor, with error ε, that outputs $m = k - 2\log(1/\varepsilon) - O(1)$.

[5] Here e is the base of the natural logarithm.

Existential proof-sketch based on Rao's linear lossless-condenser. Theorem 2.2 states that there exists a linear function $\mathsf{Cond}\colon \{0,1\}^n \to \{0,1\}^{k\log n}$, such that for any (n,k)-bit-fixing source X, the mapping Cond, restricted to X, is one-to-one. Since Cond is linear, this implies that $\mathsf{Cond}(X)$ is a $(k\log n, k)$-affine source. At this point, one can use a simple probabilistic argument to show the existence of (n,k)-affine extractors, with error ε, that outputs $m = k - 2\log(1/\varepsilon) - O(1)$ bits, as long as $k \geq \log(n) + 2\log(1/\varepsilon) + O(1)$. By applying the latter (implicit) extractor to the affine source $\mathsf{Cond}(X)$, we obtain (n,k)-bit-fixing extractors with parameters as in the proof-sketch based on the Lóvasz local lemma.

We note that by iterating over all $(2^M)^{2^N}$ functions $f\colon \{0,1\}^N \to \{0,1\}^M$, and checking each of them against any of the possible $\binom{2^N}{K+1} \cdot 2^{2^M}$ pairs of an (N,K)-affine source and a test function, one can find an (N,K)-affine extractor, with $K = \log(N) + \log\log(N) + O(1)$ and $M = K - O(1)$ output bits, in time $2^{O(2^N \cdot \log N)}$. After the application of Cond in the proof-sketch above, we only need $(k\log n, k)$-affine extractors, with $k = \log\log n + \log\log\log n + O(1)$. Namely, we can set $N = k\log n$, $K = k$ and $M = m$. Thus, the proof-sketch above, together with this brute-force search for affine extractors, yields a construction of an (n,k)-bit-fixing extractors, in time $2^{n^{O(\log\log n)}}$.

The proof of Theorem 1.3 follows the same argument as the second proof-sketch. The improvement in running-time, from the $2^{n^{O(\log\log n)}}$-time algorithm described above to the stated $n^{O((\log\log n)^2)}$, is obtained by using a more efficient construction of essentially-optimal affine extractors, as capture by the following lemma.

Lemma 5.2. *For every integer n and $\varepsilon > 0$, there exists an affine extractor*

$$\mathsf{QuasiAffExt}\colon \{0,1\}^n \to \{0,1\}^m,$$

for (n,k)-affine sources, with $k = \log(n/\varepsilon^2) + \log\log(n/\varepsilon^2) + O(1)$, and any $m \leq k - 2\log(1/\varepsilon) - O(1)$. The running-time of evaluating $\mathsf{QuasiAffExt}$ *at a given point $x \in \{0,1\}^n$ is $2^{2^m} \cdot 2^{O(n \cdot \log(n/\varepsilon))}$.*

The proof of Lemma 5.2 makes use of sample spaces that are almost k-wise independent, introduced by Naor and Naor [NN93].

Definition 5.3 (Almost k-wise independence)**.** *Let n, k be integers such that $k \leq n$, and let $\delta > 0$. A random variable X over n bit strings is called (n,k,δ)-*independent*, if for any $S \subseteq [n]$, with $|S| \leq k$, the marginal distribution $X|_S$ is δ-close to uniform, in statistical distance.*

We use the following explicit construction of Alon et al. [AGHP92].

Theorem 5.2 ([AGHP92])**.** *For all $\delta > 0$ and integers n, k, there exists an explicit construction of an (n,k,δ)-independent sample space, with size $(k\log(n)/\varepsilon)^{2+o(1)}$.*

Proof of Lemma 5.2. For an integer k and $\delta > 0$ which will be determined later, let $Z \in \{0,1\}^{2^n \cdot m}$ be a sample from a $(2^n \cdot m, 2^k \cdot m, \delta)$-independent sample space. We index a bit of the sample Z by a pair composed of $x \in \{0,1\}^n$ and $i \in [m]$, and denote the respective random bit by $Z_{x,i}$. Define the (random) function $\mathsf{QuasiAffExt}\colon \{0,1\}^n \to \{0,1\}^m$ by $\mathsf{QuasiAffExt}(x) = (Z_{x,1}, \ldots, Z_{x,m})$.

Let $U \subseteq \{0,1\}^n$ be an affine subspace of dimension k, and let $f\colon \{0,1\}^m \to \{0,1\}$ be an arbitrary function, which we think of as a "test" function, or a distinguisher. Since $\mathsf{QuasiAffExt}$ restricted to U is a function of $2^k \cdot m$ bits of Z, it holds that $\{\mathsf{QuasiAffExt}(u)\}_{u \in U}$ are 2^k random variables over $\{0,1\}^m$ that are δ-close to uniform. Thus, the random variable (with randomness coming from Z)

$$\mathop{\mathbf{E}}_{u \sim U}\left[f(\mathsf{QuasiAffExt}(u))\right] = \mathop{\mathbf{E}}_{u \sim U}\left[f(Z_{u,1}, \ldots, Z_{u,m})\right]$$

is δ-close, in statistical distance, to the random variable $\mathbf{E}_{u \sim U}\left[f(R_{u,1}, \ldots, R_{u,m})\right]$, where $\{R_{u,i}\}_{u \in U, i \in [m]}$ are $2^k \cdot m$ uniformly distributed and independent random bits. Now, by the Chernoff bound,

$$\mathop{\mathbf{Pr}}_{R}\left[\left|\mathop{\mathbf{E}}_{u \sim U}\left[f(R_{u,1}, \ldots, R_{u,m})\right] - \mathop{\mathbf{E}}_{x \sim \{0,1\}^m}\left[f(x)\right]\right| > \varepsilon\right] \le 2^{-\Omega(\varepsilon^2 \cdot 2^k)}.$$

Thus,

$$\mathop{\mathbf{Pr}}_{Z}\left[\left|\mathop{\mathbf{E}}_{u \sim U}\left[f(\mathsf{QuasiAffExt}(u))\right] - \mathop{\mathbf{E}}_{x \sim \{0,1\}^m}\left[f(x)\right]\right| > \varepsilon\right] \le 2^{-\Omega(\varepsilon^2 \cdot 2^k)} + \delta\,.$$

By the union bound taken over all affine subspaces U of dimension k and functions $f\colon \{0,1\}^m \to \{0,1\}$, we get that as long as

$$\binom{2^n}{k+1} \cdot 2^{2^m} \cdot \left(2^{-\Omega(\varepsilon^2 \cdot 2^k)} + \delta\right) < 1,$$

there exists a point in the sample space for Z that induces an (n,k)-affine extractor $\mathsf{QuasiAffExt}$ with error ε. By taking $\delta = 2^{-\Omega(\varepsilon^2 \cdot 2^k)}$, one can verify that the equation above holds as long as

$$\begin{aligned} k &\ge \log(n/\varepsilon^2) + \log\log(n/\varepsilon^2) + O(1), \\ m &\le k - 2\log(1/\varepsilon) - O(1). \end{aligned}$$

We use the construction of a $(2^n \cdot m, 2^k \cdot m, \delta)$-independent sample space from Theorem 5.2. One can verify that the sample space size is $2^{O(n \cdot \log(n/\varepsilon))}$. One can then go over each point in the sample space and check whether the point induces an (n,k)-affine extractor with error ε. By the choice of parameters, such a point exists. Each point from the sample space should be compared against $\binom{2^n}{k+1} \cdot 2^{2^m}$ pairs of an affine subspace and a test function $f\colon \{0,1\}^m \to \{0,1\}$. Checking each fixed point in the sample space can be done in time $2^{2^m} \cdot 2^{O(n \cdot \log(n/\varepsilon))}$. Hence, the total running-time is $2^{2^m} \cdot 2^{O(n \cdot \log(n/\varepsilon))}$, as stated. □

Proof of Theorem 5.1. The construction of $\mathsf{QuasiBFExt}$ is very simple, and is defined by

$$\mathsf{QuasiBFExt}(x) = \mathsf{QuasiAffExt}(\mathsf{Cond}(x)),$$

for all $x \in \{0,1\}^n$, where Cond is Rao's linear lossless-condenser from Theorem 2.2. As for the analysis, let X be an (n,k)-bit-fixing source. By Theorem 2.2, $Y = \mathsf{Cond}(X)$ is a $(k \log n, k)$-affine source. Therefore, Lemma 5.2 implies that $\mathsf{QuasiAffExt}(Y)$ is ε-close to uniform. It is straightforward to verify that the running-time and number of output bits of QuasiBFExt is as claimed. □

6 Conclusion and Open Problems

The number of extractable bits in terms of the dependency of k in n

In this paper we study the intriguing behavior of the number of output bits one can extract from zero-fixing sources (and bit-fixing sources) in terms of the dependency of k in n. Theorem 1.2 and Theorem 1.3 imply that when $k > (1+o(1)) \cdot \log\log n$, one can extract essentially all the entropy of the source, whereas when $\mathsf{Tower}(k^{3/2}) < n$, one cannot extract more than a logarithmic amount of the entropy. The remaining entropy is inaccessible, information theoretically.

Is there a threshold phenomena behind this problem? Namely, is there some function $\tau \colon \mathbb{N} \to \mathbb{N}$, such that when $k > \tau(n)$, one can extract $\Omega(k)$ bits, whereas when $k < o(\tau(n))$, one can extract only $O(\log k)$ bits? Or perhaps the number of extractable bits in terms of the dependency of k in n is more gradual? Are there different behaviors for zero-fixing and bit-fixing sources? Theorem 1.3 shows that if there is such a threshold $\tau(n)$, then the function $\tau(n)$ is asymptotically not larger than $\log\log n$.

Explicit bit-fixing extractors for sub-logarithmic entropy

Theorem 1.3 gives a bit-fixing extractor QuasiBFExt that outputs essentially all the entropy of the source, even when the entropy is double-logarithmic in the input length. Although the running-time of evaluating QuasiBFExt is not polynomial in n, it is not very high, and we feel that constructing a polynomial-time bit-fixing extractor for sub-logarithmic, or even double-logarithmic entropy, should be attainable. We suspect that such a construction would require new ideas, as the ideas used in [GRS06, Rao09b] inherently require the entropy to be at least logarithmic in the input length. Furthermore, the split-in-the-median idea used in the proof of Theorem 1.2, is based on the "1 hits randomness" property that is unique to zero-fixing sources, and does not seem to be helpful for general bit-fixing sources.

Acknowledgement

We are thankful to Ran Raz and Avishay Tal for many fruitful discussions regarding this work.

References

[AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

[BBR85] C. H. Bennett, G. Brassard, and J. M. Robert. How to reduce your enemys information. In *Advances in Cryptology (CRYPTO)*, volume 218, pages 468–476. Springer, 1985.

[BIW06] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. *SIAM Journal on Computing*, 36(4):1095–1118, 2006.

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

[Bou07] J. Bourgain. On the construction of affine extractors. *GAFA Geometric And Functional Analysis*, 17(1):33–57, 2007.

[BSZ11] E. Ben-Sasson and N. Zewi. From affine to two-source extractors via approximate duality. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, pages 177–186. ACM, 2011.

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

[CGH$^+$85] B. Chor, O. Goldreich, J. Håstad, J. Freidmann, S. Rudich, and R. Smolensky. The bit extraction problem or t-resilient functions. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*, pages 396–407. IEEE, 1985.

[CKK$^+$13] R. Chen, V. Kabanets, A. Kolokolova, R. Shaltiel, and D. Zuckerman. Mining circuit lower bound proofs for meta-algorithms. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 20, page 57, 2013.

[DG10] M. DeVos and A. Gabizon. Simple affine extractors using dimension expansion. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity (CCC)*, pages 50–57. IEEE, 2010.

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

[EL75] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets*, 10:609–627, 1975.

[ER52] P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London Mathematical Society*, 3(2):417–439, 1952.

[Fri92] J. Friedman. On the bit extraction problem. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 314–319. IEEE, 1992.

[GR08] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. *Combinatorica*, 28(4):415–440, 2008.

[GRS06] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SIAM Journal on Computing*, 36(4):1072–1094, 2006.

[GUV09] V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Journal of the ACM*, 56(4):20, 2009.

[KRT13] I. Komargodski, R. Raz, and A. Tal. Improved average-case lower bounds for DeMorgan formula size. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 588–597. IEEE, 2013.

[KZ06] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2006.

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

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

[Li13] X. Li. Extractors for a constant number of independent sources with polylogarithmic min-entropy. In *Proceedings of the 54th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 100–109. IEEE, 2013.

[NN93] J. Naor and M. Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM journal on computing*, 22(4):838–856, 1993.

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

[Rao09a] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. *SIAM Journal on Computing*, 39(1):168–194, 2009.

[Rao09b] A. Rao. Extractors for low-weight affine sources. In *Proceedings of 24th Annual IEEE Conference on Computational Complexity, (CCC '09)*, pages 95–101. IEEE, 2009.

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

[RRV99] R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in Trevisan's extractors. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 149–158. ACM, 1999.

[Sha11] R. Shaltiel. An introduction to randomness extractors. In *Automata, languages and programming*, pages 21–41. Springer, 2011.

[Spe77] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20:69–76, 1977.

[SU01] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science, 2001*, pages 648–657. IEEE, 2001.

[Tre01] L. Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, 2001.

[TSU12] A. Ta-Shma and C. Umans. Better condensers and new extractors from Parvaresh-Vardy codes. In *Proceedings of the 27th Annual IEEE Conference on Computational Complexity (CCC)*, pages 309–315. IEEE, 2012.

[TSZS01] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 638–647. IEEE, 2001.

[Vaz85] V. U. Vazirani. Towards a strong communication complexity theory or generating quasi-random sequences from two communicating slightly-random sources. In *Proceedings of the seventeenth annual ACM symposium on Theory of Computing*, pages 366–378. ACM, 1985.

[Yeh11] A. Yehudayoff. Affine extractors over prime fields. *Combinatorica*, 31(2):245–256, 2011.